9-14-2018

Design and Optimization of a 3-D Plasmonic Huygens Metasurface for Highly-Efficient Flat Optics

Bryan M. Adomanis

Design and Optimization of Plasmonic 3-D Huygens Metasurface Building Blocks for Highly-Efficient Flat Optics

DISSERTATION

Bryan M. Adomanis, Major, USAF

AFIT/GAP/ENP/PHD-18

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GAP/ENP/PHD-18

DESIGN AND OPTIMIZATION OF PLASMONIC 3-D HUYGENS METASURFACE BUILDING BLOCKS FOR HIGHLY-EFFICIENT PLASMONIC FLAT OPTICS

DISSERTATION

Presented to the Faculty

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Doctorate of Philosophy in Applied Physics

Bryan M. Adomanis, B.S., M.S.

Major, USAF

14 September 2018

AFIT/GAP/ENP/PHD-18

DESIGN AND OPTIMIZATION OF PLASMONIC 3-D HUYGENS METASURFACE BUILDING BLOCKS FOR HIGHLY-EFFICIENT PLASMONIC FLAT OPTICS

DISSERTATION

Bryan M. Adomanis, B.S., M.S.
Major, USAF

Committee Membership:

Dr. Michael A. Marciniak
Dissertation Advisor

Maj. Samual Butler, PhD
Committee Member

Dr. Augustine Urbas
Committee Member

Dr. Doug Collins
Committee Member

Table of Contents

Page

Page

List of Figures

Figure Page

DESIGN AND OPTIMIZATION OF PLASMONIC 3-D HUYGENS METASURFACE BUILDING BLOCKS FOR HIGHLY-EFFICIENT PLASMONIC FLAT OPTICS

I. Introduction

1.1 Motivation for Research

As future DoD systems move to more compact weapon systems, the design space of efficiency vs. efficacy becomes ever more critical to navigate. We need systems with low size, weight and power (SWaP), yet we also demand little to no sacrifice in warfighting capacity or competency. In spite of this, we sometimes find that the inherent physical design constraints put an upper limit on these warfighting metrics. Optical systems design highlights an example of this. In order to image moving objects, optical systems for Intelligence, Surveillance and Reconnaissance (ISR) applications often require "fast" lenses which are large with respect to their focal length; it is then easy to see how a lens diameter may be capped to reduce SWaP or to fit a particular form factor, but then would, in turn, inhibit the lens speed. What is desired is a design space which can reduce physical constraints while allowing independent tuning of performance characteristics. An intriguing new prospect for enabling such a design space for optical systems has emerged in the last seven years. Through the application of a new generalized form of Snell's Law, a lens has been demonstrated to provide an optical focusing function through a single planar interface, requiring no bulk material nor curvature of the interface[1]. These flat lenses operate through precise control of the gradient of the phase shift ($\nabla\Phi$) the incident light experiences as it

crosses the interface. In the earliest flat lens works, control of $\nabla\Phi$ is enabled through the application of engineered surfaces, such as metasurfaces[2], also called "discontinuous phase surfaces," since the phase of the wave traveling across the interface is modulated within a wavelength of light and thus experiences an abrupt discontinuity in phase, dictated along the surface by $\nabla\Phi$. For these primal metasurface design architectures, periodic arrays of subwavelength antennas were utilized to dictate the broadside angle of the plane waves scattered from incident light, termed the so-called "anomalous" refraction mode[3, 4, 5, 6, 7].

Since the first seminal two-dimensional (2D) plasmonic works in infrared (IR) and telecom wavelengths [1, 4, 5, 8], metasurface research has expanded to nearly every possible corner of electromagnetic design. In terms of spectrum, efforts have spanned RF [9, 10, 11], THz [12, 13, 14], IR [1, 4, 15, 16] and near-IR/visible regimes[17, 18, 19]; in terms of application, aside from the multitude of lensing and beamsteering functions, metasurfaces have been implemented for other novel wavefront manipulations, such as compound lensing, optical vortex beam formation, and polarization conversion[2, 20]; in terms of constituent material base, plasmonics have generally ceded popularity to dielectrics, as interests have shifted towards optical applications in the visible regime[21]; and finally, in terms of architecture, plasmonic 2D "wallpaper" works dominated early, whereas 2D arrays of high-aspect ratio dielectrics[22, 23, 24] and stacked multi-layer pseudo-3D plasmonics now reign.

1.2 Problem Statement

Despite the wide range of materials, architectures and designs, a primary challenge still remains in attaining an efficient comparison to conventional bulk optics. Single-layer 2D plasmonics have fared the worst in this, due to their fundamental limitations on efficiency for linear polarization conversion, capped at 25% [25, 26].

On the other hand, much higher efficiencies have been reported for the multi-layer plasmonic structures, with the most successful of these formed as "Huygens-like" source elements which boast forward scattering with extreme suppression of backward scatter. Though, these are built up as layers of planar stacks, and can suffer from internal reflections and limited magnetic mode control. A plasmonic out-of-plane (OOP) scatterer addresses both of these issues, which, by its nature of lying along the plane of incidence, can attain the magnetic component necessary for Huygens source generation.

However, only a few recent efforts are known to attempt plasmonic Huygens-like metasurfaces which possess a true OOP orientation. Perhaps the most fundamental OOP geometry—the omega particle—is a composition of a dipole and loop to provide the needed amplitude and phase responses (which are bi-anisotropic) and has been shown to work efficiently as reflective metasurfaces[27], but their utility as a transmit-array is shown to be impossible without active elements[28]. Out of the known OOP transmissive solutions[29, 28, 30], only one is not limited in function or optical application: the impressive work in 2015 by B. O. Zhu and Y. Feng. They developed a formulation for arbitrary transmission amplitude and phase control using realizable materials at RF frequencies; though, at higher frequencies, the lumped circuit elements required are dimensionally challenging to fabricate and the precise control of resistive and reactive elements are functionally difficult to implement. No attempt to translate their work to these spectral regimes is known. In short, neither the academic nor commercial industries have produced a consensus on the optimal design architecture, and thus research on metasurface design is still quite fertile ground.

1.3 Research Objectives

Does an optimally-efficient and realizable plasmonic Huygens metasurface solution exist at IR frequencies? To address this question, in this dissertation proposes and investigates a novel design concept for plasmonic OOP Huygens metasurface elements: starting with a scatterer "blank slate," use a genetic algorithm (GA) routine to optimize the topology of an OOP binary grid of voxels, which will form the geometries necessary to generate the proper electric and magnetic modes for efficient Huygens-like forward scattering at a targeted phase retardation (Φ_t). The overarching research objectives are:

- Design and simulate realizable GA-optimized Huygens metasurface elements to serve as the basic building blocks for full-scale flat optics.
- Construct a full-scale metasurface optical device (beamsteerer, lens, etc.) from these building blocks and validate they are operating efficiently, as designed.

The goal is for the individual elements to reach a technical objective of at 50% transmission efficiency for all Φ_t, spanning the entirety of the $0 - 2\pi$ phase space. Demonstrating a high-efficiency with full phase control would make these MS building blocks an attractive plasmonic alternative to dielectrics for practical flat optics applications.

1.4 Anticipated Impact

Future miniaturized unmanned and autonomous systems will require miniaturized components, and sensors and optical components may be the most important of these, as they support the primary function for small ISR platforms. The vision of the "Small Unmanned Aircraft Systems (SUAS) Flight Plan: 2016-2036" states four attributes that SUAS must possess to sustain the nation's military advantage

and meet operational needs; each SUAS must be: an Exponential Force Multiplier, an Easily Integrated Asset, a Cost Savings Enabler, and a Partnership Builder[31]. A low-SWaP, highly-efficient metasurface optic is a disruptive technology which can contribute directly to all of these attributes. Metamaterials and metasurfaces have shown to be capable of a dynamic response through active control of the inclusion in various ways, and thus can be significant force multipliers (not examined in this work, see several recent extensive reviews[32, 33, 34]). A planar optic would be significantly easier to integrate than a bulk, curved optic, especially with the compacted real-estate available in a SUAS. Though fabrication costs can be high for metasurfaces, there are continual improvements to this factor thanks to large investments in both plasmonic and CMOS-compatible metamaterials fabrication processes over the past two decades; and so—when factoring in SWaP savings and the force multiplication—a metasurface optic may produce an overall cost savings. Finally, metasurfaces are born out of the massively interactive metamaterials community, which has already engendered strong ties between academia, large-system integrators, and government agencies, as well as fostering interagency relationships between sister services and the greater ISR communities. However, hovering above all these attributes is the requirement to perform; metasurface devices will likely not be a successful SUAS component technology if it cannot reach similar performance metrics of a conventional device. Therefore, it is expected that this proposal for developing more efficient building blocks for plasmonic metasurfaces will be a key enabler in meeting the USAF's future SUAS vision.

1.5 Organization

Herein outlines the background, research objectives and anticipated methods, procedures and techniques availed in the pursuit of this endeavor. Chapter 2 provides

the theoretical orientation: the background, the theory and a summary of recent efforts needed to understand the state of development of metasurfaces and the nature of challenges currently in play. It covers the various metasurface architectures developed to-date, and systematically compares them against the GA-based OOP concept. A general overview of the GA optimization structure and philosophy and how it will be incorporated into the conceptual design is also entailed.

Chapter 3 encompasses the bulk of the methodologies developed for the design of the OOP metasurface elements, to include the development of the GA optimization routine and the computational methods employed in assessment of the metasurface element models. The capability of the simulation software to successfully design a metasurface element is validated against realized V-antenna metasurface lenses, while the GA routine is verified against a simple model, demonstrating that a binary grid of voxels can be organized in such a way—through genetic evolution—to form an optimal Huygens source.

Knowing that the computational foundation is determined sound, Chapter 4 proceeds with the investigation of a more complex, realizable 3D unit cell into which the OOP grid can integrate. The design architecture is finalized based on the "membrane projection lithography" (MPL) fabrication phenomenology, and a variety of voxel shapes and tilings are examined, demonstrating the flexibility in design. The GA optimization routine is shown to produce MPL designs which can span the full phase space at high transmittances and negligible reflectances, demonstrating the flexibility in performance.

With a solid foundation of design in place, full-scale devices are simulated and realized in Chapter 5. This chapter presents all findings and offers a thorough examination on the computational and experimental results. Modifications to the design deemed necessary in the course of attempting real-world fabrication were implemented

and assessed for impact to the metasurface performance. Finally, Chapter 6 provides an executive summary of these findings, and offers a discourse on the impact and future direction for continuing the research.

II. Theoretical Orientation

2.1 Introduction

Conventional refractive IR lenses for imaging, spectroscopy and laser applications are generally made of a bulk semiconductor or inorganic material with sufficient thickness and curvature to induce the desired focal length[35]. The fundamental mechanism for beam shaping is the phase accumulation generated through propagation of a wave through the bulk material. The curvature effectively modifies this propagation length to shape the wavefront at the exit interface, and the superposition of these point wavefronts produce a curved wavefront that focuses to a theoretical singularity or diverges to infinity. Optical elements that rely on phase accumulation must be optically thick ($\gtrsim \lambda$) to cover the full $0 - 2\pi$ phase space required for continuity in the refractive function along the surface. Obviously, this bulk material can become a hindrance to low-SWaP optical devices, especially considering the complex compound lenses currently in use, driving the need for low-profile lenses.

Recently, a new theory based on discontinuous phase shifts was demonstrated which introduces great flexibility in the Principle of Stationary Action[1]. This flexibility can be translated to novel optical designs, enabling an extremely low-profile, low-density solution compared to refractive optics. The key technology enabler for these flat optics is the utilization of electromagnetically-tailored scatterers of light: metasurfaces. The initial demonstration of a metasurface flat lens in 2012 was based on a design using gold V-shaped dipole antennas ("V-antennas") to address the entire phase space necessary to shape an incident beam to a focal point[4]. However, due to the inherent limitations in this design, the functional efficiency was on the order of 1%. Attempts have been made to improve these lenses; however, primarily due to design and fabrication challenges, much work remains in order to replace today's

optical systems with a low-profile variant.

This chapter summarizes the initial theory and development of metasurface lenses and their subsequent advancements, preparing the reader for a thorough examination of the OOP metasurface design concept presented in Chapter 3. First, the theory behind metasurface phase control using the Generalized Snell's Law is covered, as well as how this is applied to creating a flat optic. All the known metasurface architectures are discussed, including efficiencies and limitations for each. Lastly, the novel 3D Huygens metasurface concept is introduced, with an overview of the binary genetic algorithm (GA) approach implemented to evolve an out-of-plane (OOP) scatterer with full 2π phase control, to be used as building blocks for efficient metasurface optics.

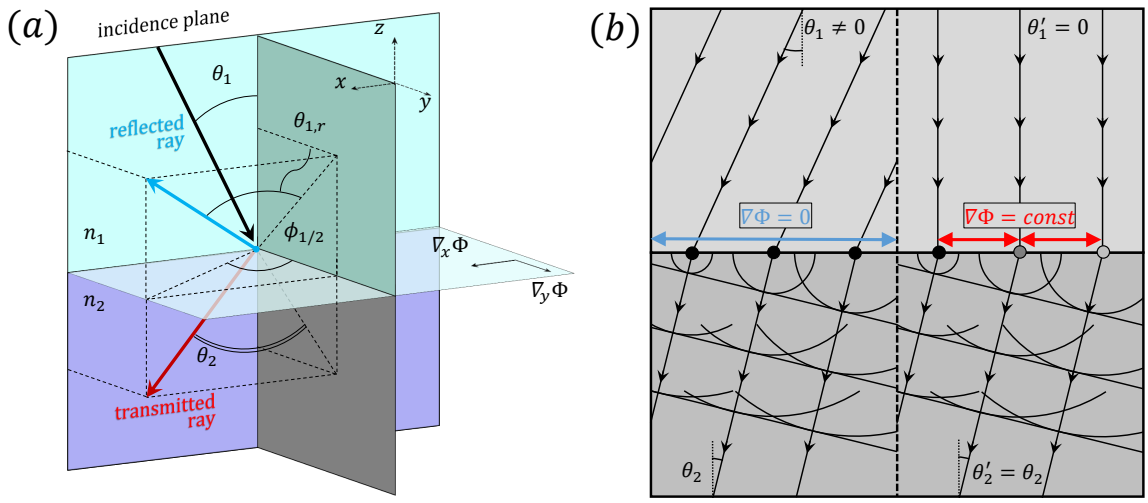

Figure 1. (a) Depiction of a generalized Snell's Law, including $\nabla_x\Phi$ and $\nabla_y\Phi$ terms representing a surface phase gradient. (b) Through application of the Principle of Stationary Action, an interface with $\nabla\Phi = 0$ (left, blue) behaves as expected via $n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$, with $\theta_1 \neq 0$; however, for $\nabla\Phi \neq 0$ (right, red)—in this case, some constant—this same θ_2 can be obtained at the incident angle $\theta_1' = 0$.

2.2 Metasurface Overview

2.2.1 Generalized Law of Reflection and Refraction.

The rapid development of metamaterials, photonics and plasmonics over the past two decades has forced many researchers to reassess their understanding of light propagation. New concepts have emerged from this body of work, such as negative or near-zero index of refraction and transformation optics. One new concept on which this research effort will focus is the generalization of Fermat's Principle, investigated in 2011 by several members of Federico Capasso's group at Harvard University. This seminal work offered a fresh perspective on Snell's Law: instead of engineering the wavefront of light propagating through a bulk, homogeneous medium of some different refractive index, the phase can be controlled through a phase gradient along a single interface[1, 2].

Refractive optics of a plane wave at the homogeneous interface of two differing bulk media is well understood: in short, referring to Figure 1(a), phase-matching conditions at a planar interface demand that an electromagnetic wave bend according to Snell's law, where for refractive indices n_1 and n_2 the well-known relation is $n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$, and the refracted angle θ_2 can be determined as a function of the incident angle θ_1[36]. However, for inhomogeneous interfaces, the phase delay (Φ) imposed upon a plane wave is no longer translationally-invariant across the surface; consequently, if the inhomogeneity is represented by a smooth function—as it is with curved bulk surfaces—a gradient in the phase ($\nabla\Phi$) is manifest.

Figure 1(b) demonstrates how a spatial phase gradient along a surface can modify the propagation characteristics of an incident light wave through the Principle of Stationary Action. Here, an oblique plane wave ($\theta_1 \neq 0$) incident on a planar surface ($\nabla\Phi = 0$) behaves as given by Snell's law, resulting in some refracted angle θ_2; whereas for a normal plane wave ($\theta_1' = 0$) upon an inhomogeneous surface there

exists some $\nabla\Phi = \text{constant}$ such that $\theta_2 = \theta_2'$. Ignoring any potential non-reciprocal behavior in the process, the converse is also true, where waves of identical incident angles will produce differing refracted angles. There is no extraordinary physical behavior going on; since the Principle of Stationary Action is the minimizing of a functional derivative with given constraints, this phase gradient is simply an additional constraint that modulates the minimization[5], and is a term needed to counter the lack of conservation of the incident wavevector across the interface. So the effect of a surface with $\nabla\Phi \neq 0$ is what is termed an "anomalous" refraction which can occur over arbitrary angles, and is explained mathematically by a generalization of Snell's Law which exhibits this gradient phase term:

$$\begin{aligned} n_2 \sin\theta_2 - n_1 \sin\theta_1 &= \frac{1}{k_0}\nabla_x\Phi \quad \text{(for parallel field components)} \\ \cos\theta_2 \sin\phi_2 &= \frac{1}{n_2 k_0}\nabla_y\Phi \quad \text{(for perpendicular field components)} \end{aligned} \tag{1}$$

where n is the refractive index of the stated domain, θ is the angle with respect to surface normal (z-axis), ϕ is the angle with respect to the y-axis, Φ is the phase and ∇_i is the gradient in the i^{th} spatial dimension along the interface.

The form of $\nabla\Phi$ is critical in how light behaves passing the interface, and can be related to antenna array theory. As well known, changing the separation and/or the relative scattered phase of an array of point scatterers will change the resulting radiation pattern[37]. As demonstrated on the right in Figure 1(b), two point scatterers—for example, the black and medium grey points—radiating at equal amplitudes and a relative phase shift to one another will generate a broadside plane wave. Adding a third scatterer (light grey) of the same relative phase generates a linear phase gradient ($\nabla^2\Phi = 0$) and will continue to support this plane wave. This behavior has a

direct correlation to diffraction from a blazed grating[38], where the slope of the phase retardation is analogous to the angle of the blazing. As such, a key takeaway from Equation 1 is that for a surface with $\nabla\Phi$ = constant a normally-incident plane wave ($\theta_1 = 0°$) will couple to a propagating plane wave at some oblique angle ($\theta_2 \neq 0°$) into one of the diffraction modes supported via the linear phase discontinuity.

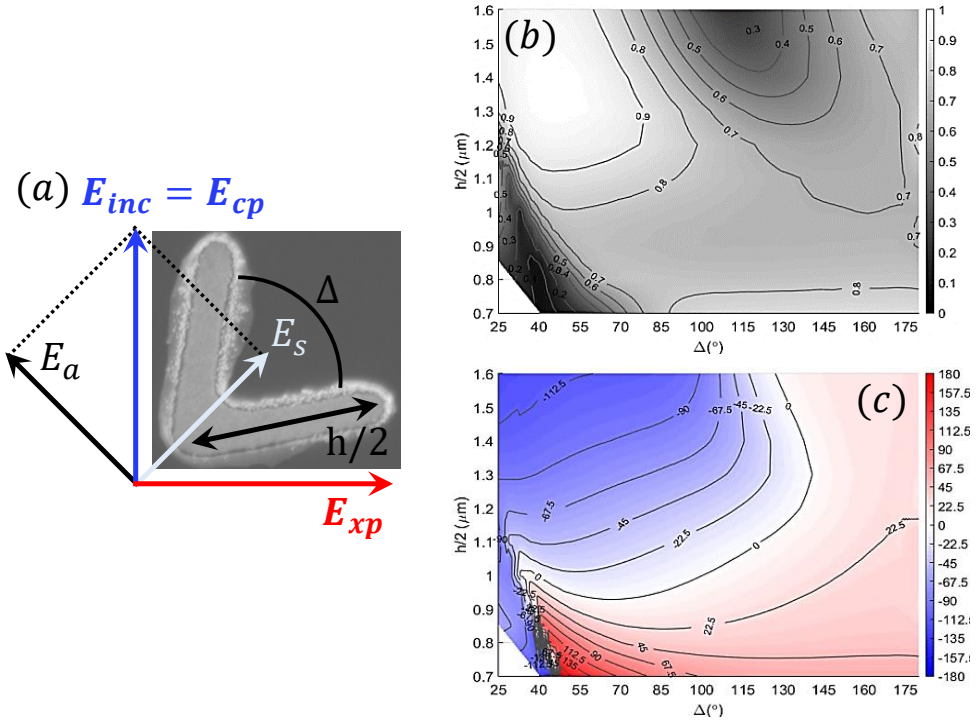

Figure 2. (a) SEM image of a representative V-antenna element, showing the fundamental geometric parameters of the vertex angle (Δ) and dipole arm length ($h/2$), the symmetric (E_s) and anti-symmetric (E_a) electrical modes between the two dipole arms, and how the incident field (E_{inc}), co-polarized (E_{cp}) and cross-polarized (E_{xp}) fields are related to these modes. (b,c) Computed profiles for the (b) amplitude and (c) phase of the V-antennas as a function of (Δ) and ($h/2$).

2.2.2 Discretized Phase Gradients Using Metasurfaces.

Up to this point, there has been no restriction placed on $\nabla\Phi$ yet. However, it becomes difficult to imagine a means to generate this control of $\nabla\Phi$ on a flat interface without discretizing the phase profile. This discretization is accomplished through modifications of the surface which scatter energy in a very particular way to mimic

the desired profile: the idea of representing $\nabla\Phi$ as a discrete array of electromagnetic scatterers is the basis of a metasurface. Metasurfaces are arrays of periodic subwavelength structures which can be engineered for a specific electromagnetic response, such as a particular resonance wavelength, bandwidth or scattering phase and amplitude. These factors become the impetus for designing a real-world flat lens, as an individual scattering element can be tailored to give a specific phase delay, and thus, correspondingly, an array of elements can be tailored to give a determinable $\nabla\Phi$ along an interface.

While the description of metasurfaces invokes to mind other structures similar in nature—and even some used in optical applications, such as frequency-selective surfaces or gradient-index metamaterials[39]—for the purposes of this document the first metasurface designed for use in anomalous refraction is from Yu *et al.* (2012). This metasurface was based on gold (Au) V-antennas on a silicon (Si) substrate backbone, as shown in Figure 2(a), and designed to operate at 8μm. The choice of V-antennas over a more simple dipole antenna was due to the lack of accessible phase space for the range of dipole lengths alone; comparatively, the V-antenna permits a wide phase space due to near-field coupling of symmetric (E_s) and anti-symmetric (E_a) electrical modes between the two dipole arms, which in this case are excited evenly by a vertical incident background field (E_{inc}. Being electrically small, other modes are negligible, as is coupling to neighboring elements due to the nature of the scattering cross-section, which is on the order of the largest antenna dimension ($\sim 1\mu$m). It is important to note this phase space exists only for the cross-polarized (cross-pol) field—as demanded by the V-antenna design—and thus requires the incident light to undergo linear polarization conversion. This factor will be discussed again when considering lens efficiencies.

As the arms vary in vertex angle (Δ) and dipole arm length ($h/2$), the current

vectors change magnitude and direction. Thus, the normalized scattering amplitude and phase delay can vary over $\{0, 1\}$ and $\{0, 2\pi\}$, respectively. Shown in Figures 2(b,c) are the amplitude and phase mappings as functions of (Δ) and $(h/2)$, from a variant of Yu's V-antenna recreated using the finite element analysis (FEA) software suite COMSOL Multiphysics®. It is highly desired for the amplitude contours to follow the gradient of the phase, so that a wide phase space can be accessed with uniform amplitude. Since each point in these profiles represents a unique scatterer geometry, a set of designs can be selected to construct a surface which possesses the desired $\nabla\Phi$ for a given optical function.

2.2.3 Metasurface-Based Lenses.

As of yet, the only behavior of anomalous refraction discussed is simple broadside bending of light; of course, what is of primary interest in this work is generating a lensing function from these planar surfaces. With this in mind, there are two means to array the metasurface elements to accomplish lensing: either the phase gradient or the element spacing must be non-linear. All known lenses have utilized the former, as the latter has low areal scatterer densities, reducing efficiency. The first metasurface lens, published by Aieta *et al.* (2012), used a version of Yu's 8-element V-antenna supercell scaled for 1.55μm short-wave IR (SWIR)[4].

So, what conditions are required for a metasurface to lens? The general concept for constructing a metasurface with $\nabla^2\Phi \neq 0$ is shown in Figure 3, where the black semicircle represents the contour spanned by a parabola, approximated as a circle with focal length f as the radius and centered at $(0, f)$. As long as all scatterers in the array are in phase along this contour, then they will converge constructively at $(0, f)$, since the waves will all travel the same distance to the center point. The requirement for the i^{th} scatterer in a plane to be in phase along this contour is given

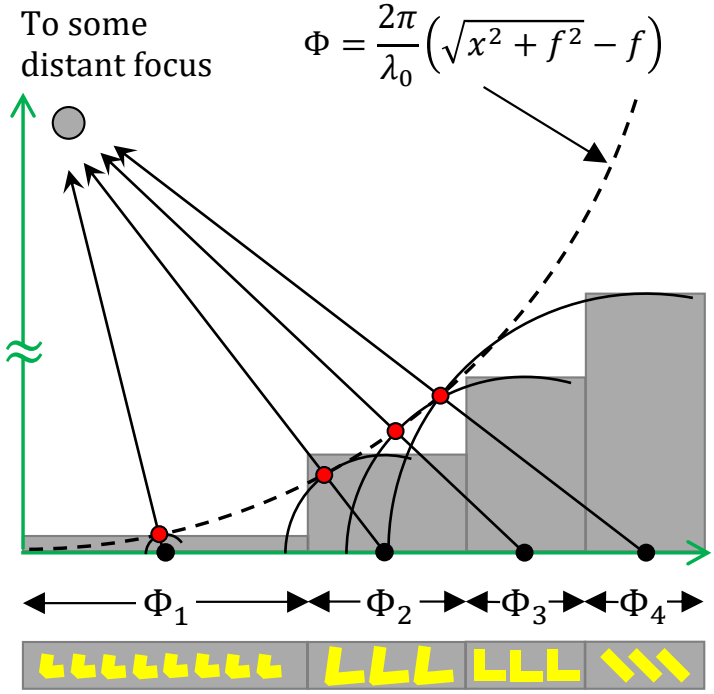

Figure 3. A parabolic phase profile (black, dashed) according to Eq. 2 can be discretized by phase zones (grey) of some determined phase Φ_i (black dot) which are all in phase along the contour (red dot), and thus all in phase at some distant focal point $(0, f)$.

by the discrete Fresnel lens formulation of the lensmaker's equation:

$$\Phi_i = \frac{2\pi}{\lambda_0}(\sqrt{x_i^2 + y_i^2 + f^2} - f) + \Phi_0 = \frac{2\pi}{\lambda_0}(\sqrt{r_i^2 + f^2} - f) + \Phi_0 \tag{2}$$

where λ_0 is the freespace wavelength, Φ_0 is the reference phase of the central $(i = 1)$ scatterer, the index $i \in \{1, N\}$, and N is the number of scatterers needed to meet the requirement:

$$\sum_{i=1}^{N} \Delta\Phi_i = \sum_{i=1}^{N} (\Phi_i - \Phi_{i-1}) = 2\pi \tag{3}$$

For a linear phase gradient, this is simply $\Delta\Phi = \frac{2\pi}{N}$. However, for a linear phase

gradient where the elements are equally spaced—i.e. a metasurface like from Figure 1(b), which only bends a plane wave—this condition cannot be met, as Φ_i follows a parabolic curve for small r.

In a real-world lens, the equiphase contour given by (2) cannot feasibly be populated by a continuum of unique scatterers, so a feasible approach is to implement discretized regions of constant phase (P_i, grey boxes) to approximate the phase matching conditions[6]. Each P_i is a non-uniform length and contains a uniform array of elements of a given phase delay. The higher the discretization (N), the better an approximation of the phase profile is achieved, and the more the point spread function (PSF) resolves as diffraction-limited. The Strehl ratio (S) is used to quantify the effect of N on lens performance, and can be shown that $N = 4$ gives $S = 0.8$, which is considered sufficient for diffraction-limited performance[40, 6]. Additionally, $N = 8$ gives $S = 0.96$, while $N = 16$ gives $S = 0.99$—a modest increase for the additional fabrication considerations, which might indicate why $N = 8$ is such a popular choice for demonstration[2, 25].

2.3 Performances of Various Metasurface Architectures

2.3.1 Two-Dimensional Planar Architectures.

Aieta's lens proved successful, though not efficient. The fabricated lens showed a clear focal region which matches well to theory, but the focusing efficiency was on the order of 1%. The primary reason was due to the inherently low maximum transmission efficiency available to a small 2D planar scatterer transmitting in the cross-pol mode. For light illuminating a surface with only constant, discretized electric current densities (i.e. no magnetic or intra-cell coupling), it can be shown using Fresnel transmission and reflection coefficients $\{t, r\}$ that the phase range of light is bounded

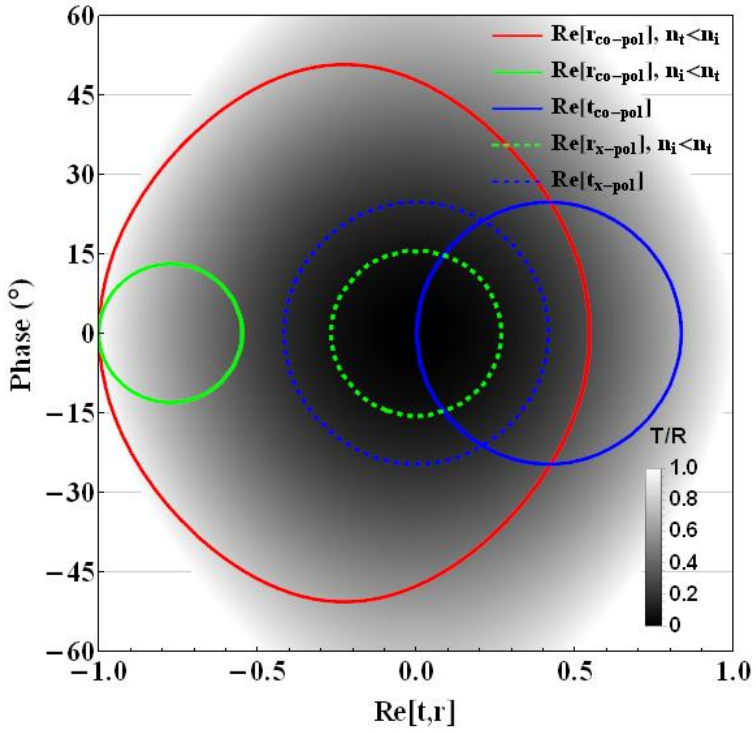

Figure 4. Complex Fresnel transmission (t) and reflection (r) coefficients for a sample 2D planar metasurface, with the horizontal axis the real part and the vertical axis the imaginary part, in terms of phase. The limits on the amplitude and phase of t (blue) and r (red, green) for co-polarized (line) and cross-pol (dashed) light are given as contours. The background density (greyscale) shows the resultant intensities $\{T, R\} = |\{t, r\}|^2$, and $\{n_i, n_t\}$ are the refractive indices for the incident and transmitting media, respectively. Limitation on Re$[t_{x-pol}]$ is shown to be $\sim 25\%$.

at $(-90°, 90°)$ for two lossless media of arbitrary refractive indices[25, 26]. Using the seminal Au/Si V-antenna planar structures ($n = 3.42$ at 8μm), the limits on the amplitude and phase of t (blue) and r (red, green) are shown in Figure 4 for co-polarized (co-pol; line) and cross-pol (dashed) light. The background density (greyscale) shows the resultant transmission and reflection intensities $\{T, R\} = |\{t, r\}|^2$, and $\{n_i, n_t\}$ are the refractive indices for the incident and transmitting media, respectively. This figure indicates that for co-pol light there is high potential T, though T suffers when spanning the phase space; conversely, the cross-polarized light spans phase at a constant, but possesses a low T.

Ultimately, Figure 4 indicates that passive, planar metasurfaces suffer from in-

herent physical limitations; for cross-polarized scattering, this is 25%. Nevertheless, the seminal work led to other developments of other metasurface a designs and applications, such as compound zone plates/lenses[41], lenses in the visible regime (VIS) based on Babinet structures[15], THz lenses[42], and vortex plates[43, 44]. Though, transmission efficiencies of these earlier structures were not very high, nor even near theory in most cases. Ni's lens was 10%, Wan's compound lens was 12%, Genevet's vortex plate was 30%, while Karimi's was 3%. In addition to the physical limitations, reasons cited in these texts are the lack of filling fraction of the inclusion (Aieta claims pushing the periodicity from 750nm to 220nm would engender a 10% efficiency), absorption losses, design of the resonator to couple only specified wave components (e.g., E-field only, limited polarizations, narrowband, etc.), and fabrication errors.

The most efficient planar metasurfaces availed the Pancharatnam-Berry (PB) geometric phase manipulation concept, involving an ensemble of identical aperture antennas which could obtain full phase coverage through rotation of the antenna about its central axis[45, 14, 10]. This functionality generally applies to any polarization state, but the conversion is near 100% for circular-to-circular polarization, with the only limiting factor being from the scattering limit. Ding *et al.*(2015) used an electrically-resonant LC lumped circuit aperture (common in metamaterial works for strong coupling into electrical modes[46]) to attain an efficiency near the theoretical limit, at 24.7%.

2.3.2 Psuedo Three-Dimensional Architectures.

Due to the aforementioned 25% cap on 2D planar architectures, researchers began expanding the metasurface concept to more complex architectures. The key ingredient missing is control over a magnetic response, which is required for effective wavefront manipulation from a planar surface[47, 37]. Without it, the impedance cannot be

tuned for efficient coupling into the interface, and the results are high insertion loss, primarily due to reflection, and no suppression of backward propagating modes. There are three primary design approaches that are employed to generate both the electric and magnetic modes required for full phase control, all requiring multiple layers of scatterers: the stacked-layer approach, the guided-wave approach and the lumped circuit element approach. All of them in one form or another utilize the concept of the surface equivalence principle, which states that the fields $\{E, H\}$ produced by all contributions to the field distribution on a closed surface $\{E_s, H_s\}$ can be fully reproduced by equivalent currents on that surface[48]:

$$\begin{Bmatrix} \overline{J}_s \\ \overline{M}_s \end{Bmatrix} = \begin{Bmatrix} \hat{n} \times \overline{H}_s \\ \overline{E}_s \times \hat{n} \end{Bmatrix}. \tag{4}$$

This means that material losses, magnetoelectric coupling, and modes from any number of scatterers can be represented by a single set of complex current terms, as long as these effects occur over a region which is small with respect to the wavelength.

The stacked-layer approach involves stacking of these 2D planar structures—using microwave engineering terms, these are called sheet impedances (or admittances). The electrical responses are driven by the capacitive elements as common in most engineered antennas, such as frequency-selective surfaces, metamaterials and the aforementioned metasurfaces, while the magnetic responses are generated by circulation between the capacitive sheets. The responses of each sheet can be independently tuned through scatter geometry (electric) and sheet spacing (magnetic), and thus each can be ultimately represented by separate parallel LC circuits. Knowing the input fields, the circuit transformation, and the desired amplitude and phase of the scattered field, closed form solutions are possible via an ABCD transfer matrix for determination of the required admittances as a function of the geometries and spacings.

Excellent early examples of these metasurfaces are from Paul *et al.*(2010) and Pfeiffer & Grbic (2013)[13, 49]. While theoretically this approach can yield 100% transmission efficiency, these structures produced still-impressive 85% and 62% efficiencies, respectively. Paul's design was simulated as a full 3-layer lens, with simulated efficiency of 38% and measured 20%[12], while for a 3-layer beamsteering variant of Pfeiffer's design it was 33%/20%[47], with a majority of the deficiency blamed on Ohmic loss and fabrication errors. Nevertheless, this was still a major improvement over the single-digit efficiencies of earlier fabricated works.

Metasurfaces based on the guided-wave approach are primarily used for bi-anisotropic control of polarization states. They use a similar approach, but tailor the incident impedance sheet to also accept a certain polarization state with low insertion loss, then couple to an intermediate functional sheet, and then to a sheet tuned for efficient out-coupling of desired polarization state into a propagating mode. Pfeiffer *et al.*(2014) again led the field on a thorough demonstration of this approach, attaining a measured 50% transmittance at 1.5μm from circular-to-circular (C2C) conversion.

The final pseudo-3D architecture uses lumped circuit elements (LCE) to generate the electric and magnetic responses, generally independently. Many of these designs consist of a capacitively-loaded dipole-like element aligned to the incident E-field and an inductively loaded loop-like element aligned out-of-plane for coupling of the H-field[7]. Once again, Pfeiffer & Grbic (2013) presented a thorough analysis of this architecture, specifically tuning the structures to behave as Huygens' sources with extremely low backscatter, attaining a peak transmission efficiency of 86% into the desired diffraction mode.

A Huygens source is mathematically formulated as follows[48]. Consider a domain is split into two regions separated by a surface S at $z = 0$, with a propagating free space plane wave sourced at $z = -\infty$ with amplitude E_0, wavenumber $k_0 = 2\pi/\lambda_0$

and in free space impedance Z_0, where the incident fields are

$$\begin{Bmatrix} \overline{E}^i \\ \overline{H}^i \end{Bmatrix} = \begin{Bmatrix} \hat{x} E_0 e^{-ik_0 z} \\ \hat{y} \frac{E_0}{Z_0} e^{-ik_0 z} \end{Bmatrix}. \tag{5}$$

Then the equivalent currents on S must be such that this wave is continuing to propagate forward, with no backward propagation. Carrying out the current calculations in Equation 4 gives

$$\begin{Bmatrix} \overline{J}_s \\ \overline{M}_s \end{Bmatrix} = \begin{Bmatrix} -\hat{x} \frac{E_0}{Z_0} \\ -\hat{y} E_0 \end{Bmatrix}, \tag{6}$$

and by solving for the well-known vector potential equation, the fields in each region are given by, for the electric current

$$\begin{Bmatrix} \overline{E}_e \\ \overline{H}_e \end{Bmatrix} = \begin{cases} \begin{Bmatrix} \hat{x} \frac{E_0}{2} e^{ik_0 z} \\ -\hat{y} \frac{E_0}{2Z_0} e^{ik_0 z} \end{Bmatrix} & z < 0 \\ \begin{Bmatrix} \hat{x} \frac{E_0}{2} e^{-ik_0 z} \\ \hat{y} \frac{E_0}{2Z_0} e^{-ik_0 z} \end{Bmatrix} & z > 0 \end{cases} \tag{7}$$

and for the magnetic current,

$$\begin{Bmatrix} \overline{E}_m \\ \overline{H}_m \end{Bmatrix} = \begin{cases} \begin{Bmatrix} -\hat{x} \frac{E_0}{2} e^{ik_0 z} \\ \hat{y} \frac{E_0}{2Z_0} e^{ik_0 z} \end{Bmatrix} & z < 0 \\ \begin{Bmatrix} \hat{x} \frac{E_0}{2} e^{-ik_0 z} \\ \hat{y} \frac{E_0}{2Z_0} e^{-ik_0 z} \end{Bmatrix} & z > 0. \end{cases} \tag{8}$$

Quite obviously, when the contributions are added from the symmetric fields due to the electric current source and the anti-symmetric fields due to the magnetic current source, the remaining fields are only in the $z > 0$ region. This gives additional insight on the physical limitation noted previously for 2D planar metasurfaces, whereas a surface generating only a symmetric field cannot suppress backward propagation, losing at least 50% of incident power in this direction.

The enlightenment that Equations (7) and (8) provide is that a Huygens source can be generated from any surface interfacing with an incident plane wave, as long as the surface currents can be controlled as described in Equation (4). All three of these pseudo-3D architectures offer this control of $\{\overline{J}_s, \overline{M}_s\}$ in different ways by tailoring the impedance sheets, accomplished in practice through patterning of the metasurface with elements which are engineered for optimal mode coupling (and in the case of the guided-wave approach, for a desired polarization state). However, they all also require a multilayer design—whether stacked vertically, or sandwiched horizontally on opposite sides of a dielectric spacer, as in the case of the LCE—and this introduces significant fabrication complexity with respect to the simple planar architectures. Many alternative designs exist for each approach, with some even a hybridization of multiple approaches, such as the work by Monticone *et al.*(2013) using optical LCEs in conjunction with stacked layers to implement a device which has impedance-matched external layers surrounding the central functional layer[25]. Their efficiency was upwards of 75%, but realization of the structure was challenging due to severe losses in the metallic adhesion interlayers.

Attempts at metasurfaces as transmitarrays with single-layer (or "single-wire"), out-of-plane (OOP) Huygens-like elements have been attempted, but few. Perhaps the most fundamental OOP geometry–the omega particle–is a composition of

a dipole and loop to provide the needed amplitude and phase responses (which are bi-anisotropic) and has been shown to work efficiently as reflective metasurfaces[27], but its utility in forming a transmitarray is shown to be impossible without active elements[28]. Similarly, chiral elements, such as a spiral element, cannot control arbitrary phase in transmission, and also cannot be used for wavefront manipulation. Out of the known OOP transmissive solutions which can modulate phase[28, 29, 30], none have yet demonstrated a tractable solution for optical frequencies. Hu & Wei (2017) used an array of plasmonic rectangular prisms with varying widths and separations to generate independently-tunable electric and magnetic modes with transmission efficiencies averaging 85% across all design parameters in the short-wave IR (SWIR). However, due to the heavy reliance on neighboring periodic elements, the maximum realizable phase range was limited to a $0 - 1.5\pi$ interval, and so this design cannot be used for prime optical applications such as lensing or beamforming. Two general formulations have been presented for implementation of Huygens' metasurface elements for arbitrary wavefront control, but both have only been demonstrated in the RF, using designs which are not as tractable at optical frequencies. The first, in the work by Zhu & Feng (2015), a formulation for arbitrary transmission amplitude and phase control using realizable materials was presented. Here, the LCEs required careful placement of resistors of known resistances at capacitive gaps to control the phase response, but at higher frequencies these type of components are dimensionally challenging to fabricate and the precise control of resistive and reactive elements are functionally difficult to implement. The second, works by Tretyakov's group (2015-2016) demonstrated an exceptional understanding of OOP element compositions for Huygens-like scattering, combining Omega and chiral elements of opposite handedness to mutually compensate the bi-anisotropic contributions from each other, permitting the unit cell to behave as a pair of orthogonal electric and magnetic dipoles[27, 28].

They achieved transmittances of 83% over a 7% bandwidth, centered at 4.32 GHz.

Though both of these formulations are spectrally independent, there are no known attempts to generate equivalent designs using these OOP geometries and architectures in the optical regime, and thus there is still a vast deficiency in efficient, plasmonic-based metasurfaces. On one hand, this apathy of metallic nanostructures is entirely justified, due to the runaway dissipative losses experienced in metals at optical frequencies. On the other hand, perhaps interest has been rather low in plasmonic architectures only due to the great successes in achieving high efficiencies with dielectric architectures. These structures will not be given great discourse in this dissertation, but the most successful designs generally operate in the same manner as the pseudo-3D plasmonic architectures, instead of controlling phase through formation of circulating currents along a curved wire or between capacitive sheets, rather they form these currents due to Mie resonances within the 1λ-thick, high-aspect dielectric structure in a manner quite analogous to Fabry-Perot resonances[22, 50, 51].[1] A major criticism of dielectric metasurfaces is the intrinsic narrowband resonant behaviors, as Mie resonances are highly dependent on geometry; nevertheless, they offer many benefits. Most importantly, they have shown to be extremely efficient, measured in the upward range of 80%, comparable to commercial bulk optics[52]. Additionally, realizations of test articles are exceptionally simplistic, considering current CMOS-compatible fabrication standards. As a result, developments of plasmonic architectures—which had not achieved over 40% in measured efficiency—were largely abandoned[21].

In summary, the 2D plasmonic reflectarrays and pseudo-3D stacked or LCE plasmonic transmitarrays have been reasonably successful in practice, demonstrating unique polarization control, broadband response, and wide angle operation. Despite this, all-dielectric metasurfaces have thoroughly exceeded the efficiencies of these plas-

[1] There are also analogous waveguide structures, allowing phase control by propagation delay, but these are not as popular—see references in [22].

monic architectures, and no fundamentally new architectures have been introduced that might overcome this obstacle. For high efficiency, it is quite apparent a Huygens-like behavior is imperative; and to accomplish this, it is quite apparent that some OOP component must exist to support generation of the anti-symmetric magnetic modes. All the while, some mechanism must be in place that modulates the phase of the total field. While pseudo-3D designs can theoretically achieve the exact impedances and admittances needed to accomplish perfect, phase-controlled Huygens' sources, limitations are introduced in the construction, such as reflections and absorption due to internal reflections between stacked layers. Therefore, a continued pursuit of metasurface design concepts is warranted. This dissertation covers the development of a novel 3D metasurface design architecture that offers an unprecedented flexibility in the functionality of the metallic scatterer.

2.4 Binary Genetic Algorithms (GA) for Discrete Metasurface Optimization

2.4.1 Genetic Algorithm (GA) Overview.

In the search for the ideal scatterer topology, perhaps the best first step is to relax the constraint that compositions of wires and loops is are the most ideal geometries for phase-controlled Huygens' sources. This begs the question: with a blank canvas, how would nature design this unit cell? One answer can be found in an optimization technique called a "genetic algorithm" (GA), which is a subset of the broader span of evolutionary algorithms. The concept is simple: using a large random set of models, evolve the parameters which define a particular design space using a biologically-inspired "survival of the fittest" approach, similiar to the way Darwinism cites natural selection as the natural mechanism to impress small, positive variations of genetic traits to offspring in order to improve the chance to survive; conversely,

non-beneficial genes reduce the probability that the offspring will continue to propagate those genes. Analogously, the best performing model parameters are combined to form new designs, which should promote the propagation of designs with better results and suppress those with poorer results, statistically. Genetic algorithms have been used for optimization of antenna arrays since the 1980's, and experienced a resurgence in the advent of metamaterials, where the focus was now placed on optimizing the scatterer geometry and material, rather than the array layout; for example, the antenna factor of an array of uniform Hertzian dipoles[53]. With the rise of metasurface applications, the architectural complexity has been vastly increased over previous uniform 2D planar metamaterials—the non-uniform arrays of subwavelength inclusions often requires changes in phase response, orientation, geometry over a distance as small as one period. It makes sense, therefore, to apply a GA approach to the design of an OOP Huygens metasurface.

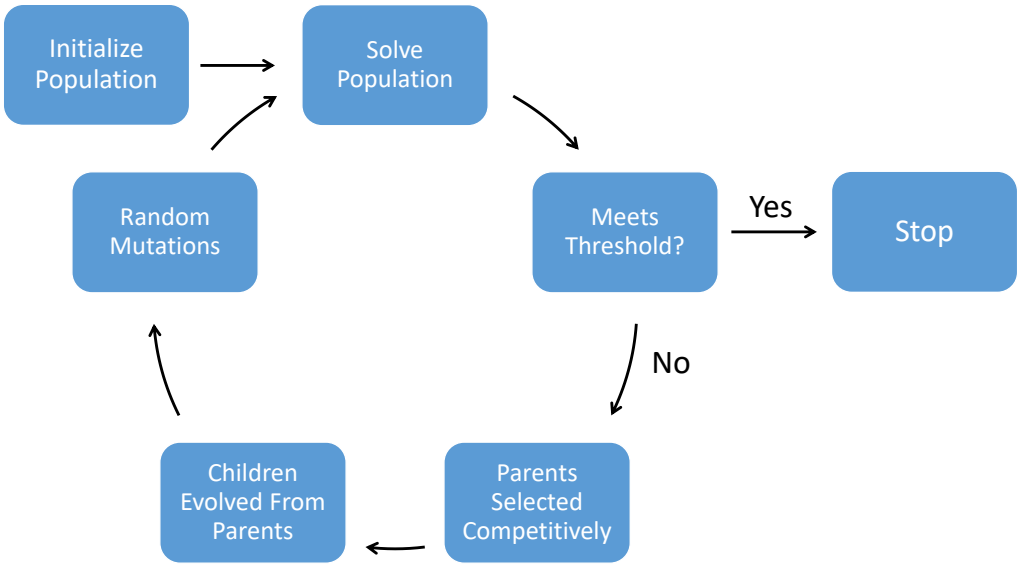

Figure 5. Generalized process flow for a genetic algorithm.

Details for the specific implementation of the GA routine will be covered in Chapter 3.3, as it involves integration into a computational electromagnetics (CEM) solver;

but for now, it will suffice to cover the basic GA definitions and a generalized process flow, shown in Figure 5. Each design, or "individual" in a parameter space is defined by a unique set of parameters, or "genes," which can be represented in a conventional encoding—say some length $L = 0.85\mu$m—or a binary coding, where that same length can be represented in a string of bits, such as $L = \{101101\}$. Any number of parameters can be strung together to form a "chromosome," the full set of genes used to uniquely identify an individual, and a full group of individuals uniquely identifies a "population" or "generation." The objective of the GA process to randomize an initial population and solve in a manner appropriate for the problem at hand, whether the genes be the inputs to an analytical or computational model. The chosen output results are fed into a "fitness" function (also termed "cost" or "objective" function), which is designed to give individuals with more desirable results a higher score. The fitness function is usually weighted for preference of one or more outputs, and normalized to between {0,1}. From here, a selection process unfolds to determine which individuals will pass on genes. Primarily, these are either deterministic, where the absolute most fit in each generation are chosen as parents, and the rest are eliminated; or, stochastic, where the most fit individuals are given increased opportunities of being chosen as parents over less fit individuals, though not guaranteed. The genes of these selected parents are evolved in one of two primary methods: "crossovers" and "mutations;" this is where statistical probabilities begin to significantly affect future populations, even for deterministic selection strategies.

Crossovers take part of each parent and combine their genes in a manner that makes a new "child" design with a set of genes unique from either parent. Why crossovers improve a solution search is not understood entirely[54], but the idea is that the merging of parameters from better-performing designs have a better chance of producing more fit results than the merging from poorly-performing ones. After

all children have been created, a low rate of mutations is applied to the population (typically 1-10%); for binary encodings, this simply means that a bit is flipped from $1 \rightarrow 0$ or $0 \rightarrow 1$. The primary purpose of a mutation is to kick a few individuals out of some local extrema by making a significant modification to the chromosome, as opposed to the crossover tendency near local extrema to engender similar children from similarly fit parents. The mutation operation forces a probing of new regions of parameter space, and so the results of some of the children may reside along previously-unsampled gradients in the solution space, which, over enough iterations, should progress to a more optimal extremum. Once all the children are determined, they constitute a new generation, and the general process of solve-check-evolve is repeated. After each solution set, the fitnesses can be checked against a fitness threshold as a go/no-go switch for termination of the routine. Additional heuristics are often implemented to help convergence of the GA routine, and the specific ones used in this GA routine will be covered in Chapter 3.3.3.

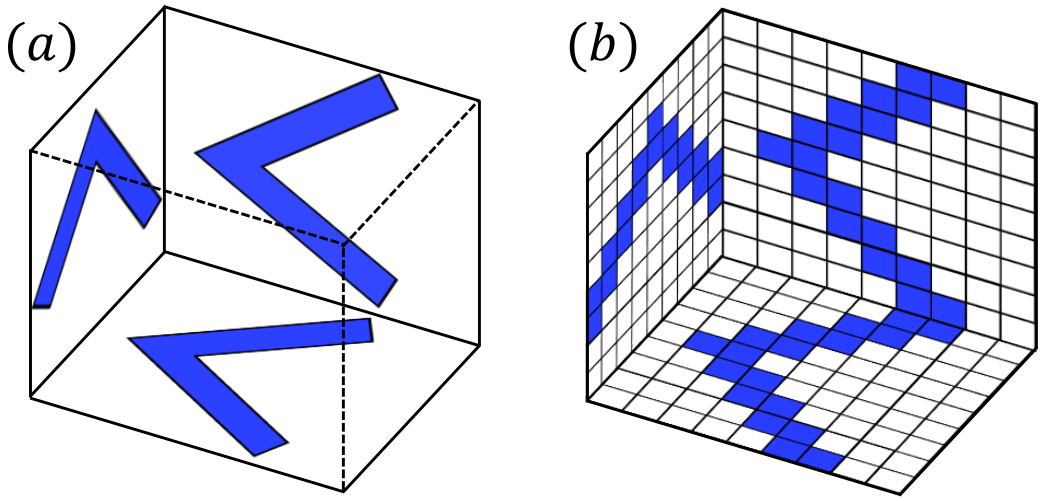

Figure 6. Concept 3D architecture for supporting out-of-plane (OOP) scatterer. (a) An analog scatterer can be supported on vertical walls and the floor of a cavity. (b) By deconstructing the walls into $M \times N$ grids of 3D pixels—i.e. voxels—then the grids can be optimized using a genetic algorithm routine by switching the voxels in a binary operation to either a "1" for metal or a "0" for air.

2.4.2 Voxel-Based Building Blocks.

How can this process assist in creating an ideal Huygens source? The concept begins with the choice of a 3D architecture that can support an OOP scatterer. Ideally, the inclusion would be suspended in air, not bonded to a substrate which can interfere with scattering behavior. This can be easily simulated, of course; but it was intended that this research would be designed with realizable fabrication conditions in mind. Therefore, a viable means of support would be as displayed in Figure 6(a): an open cavity with a set of orthogonal walls, upon which a scattering inclusion can be placed in alignment with, or perpendicular to, the incidence plane (for p- and s-polarization states, respectively). To implement the GA routine, imagine deconstructing a small volumetric region adjoining any of the interior walls into an $M \times N$ grid of voxels[2], with each voxel permitted to exist as either a metal or as air, as in Figure 6(b). The GA routine can treat each of those voxels as a parameter with a binary range of either "1" as metal or "0" as air, and the entire voxel grid of binary values constitutes a chromosome representing a unique scatterer. From here, the routine can evolve the grid into the optimal topology for a Huygens source emitting at a desired phase at a maximum transmittance, as per calculated by the CEM solver. The parameter space of such a design, considering all combinations of voxel grid layouts, is $2^{M \times N}$ according to the Rule of Product. The shear size of this space reinforces the preference of GA optimization method over any calculus-based method or many of the random search methods, because calculus-based methods would only span a small region of the space before converging to a local extrema; and because global convergence of many random search methods tends to be highly dependent on starting point (see Nelder-Mead method in [53]), so the larger parameter space, the lower the likelihood

[2]Some works use the term "pixels," even though there is some finite thickness to the geometry; here, "voxel"—the more accurate term for a 3D pixel—will be used whenever there is a non-negligible thickness.

of picking a viable start point. Moreover, some these methods are based on pure statistical sampling, such as a Monte Carlo method, require an intractable number of solutions must be made to have any significant probability of finding a suitable extremum in extremely large spaces.

This concept of discretizing a spatial region into an array of binary pixels/voxels and then applying an evolutionary optimization routine is not new. The first reported pixel-based GA optimization was by Manara *et al.*(1999) applied to a frequency-selective surface (FSS)—in many ways a precursor to metasurface[55]. In several examples from Prof. Werner's group at the Pennsylvania State University, they applied a similar technique to an FSS in order to control spectral position of stop-bands[56], bulk material properties[57], and create artificial magnetic conducting ground planes[58]. Werner *et al.*(2011) also created a switchable voxel out of a chalcogenide glass phase-change material, and was able to demonstrate three dual-function filters (reflection/absorption at 3.5μm, transmission/reflection at 3.5μm and transmission/absorption at 1.55μm) out of this single building block[59]. In order to calculate the desired results (generally transmission or reflection coefficients), usually a computational method is utilized, such as a periodic method-of-moments (PMoM), where either the pixel is treated independently with rooftop current approximations, or a finite element method (FEM), where all pixels are meshed as a unit cell and solved using field integral equations.

However, to-date no OOP optical scatterer has been attempted in this discrete genetically-evolved manner, and few have attempted complete phase control. A singular example of using a pixelated-GA routine for full-2π phase control was produced by Bossard *et al.*(2016) for high-powered microwave reflector antennas, though it is important to note this was for a 2D planar surface[60]. This lack of GA implementations for OOP designs could be due to the reason why many researchers pursue an

optimization routine: they intend to fabricate the structure. If so, an OOP scatterer in the IR regime has not historically been an easily realizable design, and may have precluded adaptation of the pixelated GA implementation.

2.5 Summary

Metasurfaces are a relatively new concept, and as a result, it was necessary to thoroughly identify and assess each type of metasurface architecture developed to-date in order to firmly establish the novelty of the proposed OOP concept. It was shown that the most promising plasmonic architectures were Huygens-like pseudo-3D stacked layers of planar 2D metasurfaces, and a few analog OOP designs; however, none of these have proven viable for realization in the IR regime. Instead of relying on an analytical approach to solve the problem, this research employs a computational natural selection to determine the optimal architecture. The overarching concept is to apply the binary genetic algorithm to a grid of voxels embedded inside a 3D unit cell which allows for OOP interaction with an incident wave, and in turn, the electric and magnetic modes required for Huygens-like scattering. The next chapter will detail the development of the GA routine in full—to include the computational design and implementation—and it will be verified that the GA can generate such a Huygens scatterer from an OOP grid.

III. Research Methodology

3.1 Introduction

This chapter covers the methodology behind the OOP Huygens metasurface. First, the structural architecture is described, leading to the development of the 3D unit cell with voxelated grids. Next, the initialization and verification of the COMSOL-based GA routine are discussed in detail, to include a breakdown of each GA process and included heuristics. Then the fabrication phenomenology is outlined, showing how the demonstration beamsteerer was produced. Finally, a description of the experimental measurement is included, highlighting the process in which the device was validated.

After the initial concept formulation—the idea that a highly-efficient transmissive metasurface architecture might be formed from an OOP voxel grid optimized by a GA routine—a logical, step-by-step progressive development process was enacted to ensure success in this endeavor. Before jumping directly into a complex 3D model, it was necessary to establish a viable "cradle-to-grave" methodology for taking a design from concept to optimization, then to fabrication, and finally to experimental validation. The first step was to identify a numerical technique to solve for the needed metrics—the far-field phase and amplitude from the populated voxel grid—and then verify that this choice could accurately reproduce a flat optical device from the simulated individual elements. This included laying the foundations for fabrication and measurement; the successful completion of this process would engender a high confidence in the ability to accomplish all tasks in the research effort. The second step involves the development of the GA routine and the initial test and evaluation phase, certifying its functionality using a simple model; in particular, by optimizing the voxel layout of a single grid suspended in air to behave as a Huygens-like scatterer.

The third step was the full implementation and refinement of the GA routine for the periodic 3D unit cell, consisting of the cavity and a populated grid on at least one of the interior faces. Finally, the last step involved the final design determination for a full-scale device and experimental validation. The first three steps are covered in this chapter, whereas the last step will be addressed in the proceeding two chapters.

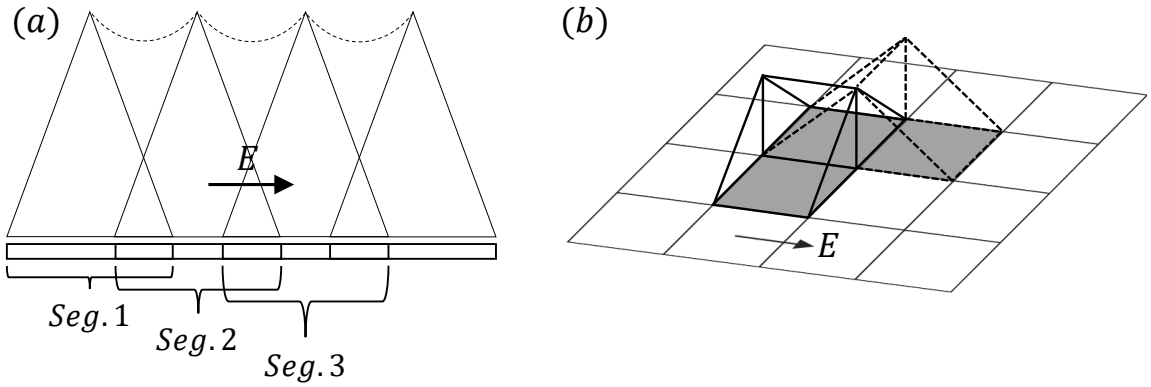

Figure 7. (a) Periodic method of moments (PMoM) requires a piecewise linear distribution for a wire, such as this commonly-used rooftop function of overlapping triangular distributions. (b) Extending this concept for pixels along a 2D surface, the rooftop function can simply be translated in the direction orthogonal to the dominant current distribution to form rectangular prisms.

3.2 Identification and Verification of Numerical Approach

3.2.1 Common Computational Methods for Pixelated Arrays.

As mentioned in the last chapter, most pixelated GA implementations utilize a periodic method of moments (PMoM) or finite element method (FEM) full-wave analysis numerical technique to solve for the desired sources and fields, since there are generally no analytical solutions to such a robust electromagnetic problem. The PMoM deals with expressing the scalar and vector potentials as the well-known integral equations, with the sources inside the integral and boundary conditions (BC's) determined by the specifics of the environment (Floquet BC's, symmetry BC's, PEC, applied currents/potentials, etc.). These sources are broken into basis functions representing

the current in some local spatial region of the interface, such as the commonly-used rooftop function shown in Figure 7(a) for a piecewise linear distribution for a wire[61]. Treating the pixels as an electrically-small antenna elements with their own BC's, in its most simple formulation the basis function is essentially a rectangular base triangular distribution, which can cover multiple pixels as shown in Figure 7(b), and with the function just translated in the direction orthogonal to the dominant current vector; more sophisticated formulations can be used to consider additional behaviors like, for example, edge effects[61, 62]. The distributions are then mapped to the entire structure, oftentimes overlapping, and the integrals are solved in a piecewise fashion. Once the surface currents are determined, the near- and far-field components can be computed from known propagation methods.

The benefit of the PMoM is that only the interface in which the sources are induced needs to be analyzed, and the fields in the remaining regions can be evaluated by periodic Green functions on the boundaries and free-space Green functions for all other domains external to the interface. However, it is the least accurate of the full-wave analysis techniques, and is limited to rather simple planar geometries (these pixelated designs being made of small, simple geometries to form complex geometries). For the proposed OOP Huygens metasurface, there is a great deal of complexity in the structure, possessing one to five voxelated faces which may be either in-plane or out-of-plane, and may be aligned with or perpendicular to the incident plane. Additionally, it has been shown that for any polarization state a scattering response of a single scatterer can be modified when placed in an opposing orientation and in close proximity, due to the increased coupling[63]. A high level of intra-cell coupling is certainly a possibility in this 3D unit cell, and thus the basis functions would poorly represent the true surface currents. Therefore, it was decided that the FEM would be the optimal tool for accomplishing the unit cell analysis.

The FEM approach is quite similar to the PMoM approach in computation (basis functions, integral equations), with two stark differences. First, the entire computational domain is meshed, usually by triangles (2D) and tetrahedrals (3D) which can be distributed non-uniformly and with varying resolution, permitting high accuracy in regions of importance. This also easily accounts for coupling, as information across space is transferred from volumetric element boundary to volumetric element boundary, instead of propagated between segregated surfaces via Green functions. Second, the integral equations used are based on the differential form of Maxwell's equations, and the basis functions represent field or potential distributions, as opposed to the current distribution basis functions of the PMoM. Being not limited to surface currents and charges, this approach is useful for handling OOP interactions and inhomogeneous media—such as can be expected from the proposed design—and is considered more accurate than any PMoM approach[64]. The major sacrifice for this accuracy ends up being larger matrices to solve (though very sparse due to the small "finiteness" of the element), and thus increased resource requirements.

Evidently, it is unavoidable to use the FEM analysis technique for this 3D design. Many proprietary and public-domain FEM simulation codes exist, and several commercial FEM solvers have emerged as leaders in industry and academia (Abaqus, Ansys, Autodesk). This effort will use the commercial COMSOL Multiphysics ("COMSOL") product exclusively to solve for the complex total fields emitted in the far-field region by each individual model in the GA routine. COMSOL uses several FEM techniques (it can handle time-dependent PDE's), such that is cannot simply be stated as a FEM solver, but more of a software suite of FEM-based numerical analysis and solver tools. As opposed to pure programming-based applications of FEM—say, through Mathematica or MATLAB—COMSOL has the advantage of a graphical user interface, which helps immensely with computational meshing of complicated topolo-

gies. All needed features for this effort are embedded in the software, to include periodic boundary conditions (PBC's), material loss, multi-port networks (for extracting co- and cross-polarizations), diffraction order calculations, far-field calculations, and results-based visualizations.

In addition to COMSOL's GUI-based control, a Java based Application Programming Interface (API) can be used to script all aspects of the modeling and simulation process in MATLAB, through an add-on module called "LiveLink for MATLAB." This permits a shared model to exist between the two programs, where changes to the model in one environment are dynamically updated in the other, and programmatic control can switch seamlessly, such as extraction of results from COMSOL, analyzing in MATLAB, then updating the model according to the analytical outputs and re-solving in COMSOL. The overall impact is that a user has the power of a fully-autonomous programmatic FEM algorithm, with the sophistication of software-based FEM feature implementation—an impact which will be highlighted in Chapter 3.3.3 when covering the GA development.

3.2.2 Verification of COMSOL Metasurface Design.

The first step in adapting COMSOL was to verify the software by attempting to reproduce a reported metasurface device, which would serve as a benchmark of what could be expected from the simulations in terms of precision, accuracy, limitations and computation time. The easiest design choice for this verification task was the V-antenna made famous by Capasso's group[1, 4]. The goal of these initial simulations was to obtain the scattered cross-polarized field (E_s^{xp}) of the V-antennas as a function of the vertex angle (Δ) and dipole length (h), where

$$\overline{E}_s^{xp} = \overline{E}_{total}^{xp} - \overline{E}_{background}^{xp} = E_{total}^{xp}. \tag{9}$$

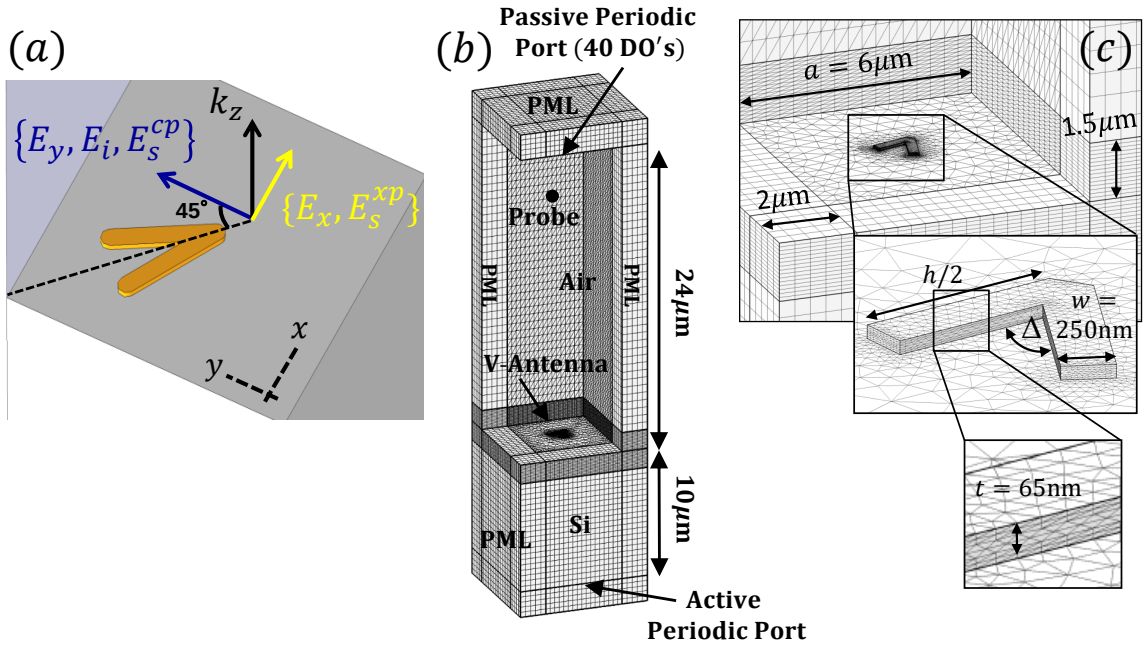

Figure 8. (a) The field components of interest for V-antenna model, where the incident field $\overline{E}_i$, co-polarized scattered field $\overline{E}_s^{cp}$ and near-field y-component $\overline{E}_y$ are all aligned, and the same for the cross-polarized scattered field $\overline{E}_s^{xp}$ and near-field x-component $\overline{E}_x$. (b) Full-scale model of the V-antenna unit cell, highlighting interface thicknesses, boundary conditions, active/passive ports and extraction probes. (c) Close-up of scatterer, showing all pertinent parameters and dimensions.

This, of course, was the intention in the seminal work: to isolate the cross-pol total field from any interference from the background field, which has zero amplitude in the cross-pol basis. From Equation (9), the phase ($\Phi = Arg[\overline{E}_s^{xp}]$) and amplitude ($|\overline{E}_s^{xp}|$) were extracted to map the scatterer response from a full range of parametrized V-antenna geometries.

The field components, full COMSOL model and metrics of interest are shown in Figure 8(a-c), respectively. The model consists of a 65nm-thick gold (Au) V-antenna scatterer with a trace width of $w = 250$nm on a 10μm-thick silicon (Si) substrate embedded in 24μm-thick air domain. Two solution steps were required to acquire $\overline{E}_s$ for a scatterer residing on a dielectric half-space: the first environment solves for the total field of the undecorated air/Si interface ($\overline{E}_t'$) using PBC's and periodic

ports; while the second adds in the scatterer and replaces the ports and PBC's with perfectly-matched layers (PML's) to solve for $\overline{E}_s$ using $\overline{E}'_t$ from the first environment as the background field of the second, such that $\overline{E}_s = \overline{E}_t - \overline{E}'_t$. Note that this second environment does not include any periodicity, which is acceptable for this work. The assumption is that each scatterer will be placed far enough from nearest neighbors so that mutual coupling can be ignored, and has been a staple approximation in previous metasurface work[1, 45, 15]. As a result, the unit cell width was set to $a = 6\mu$m to be large enough to not interact with the PML, but small enough to limit the diffraction orders to ± 1; see [65] for additional details on the simulation environment.

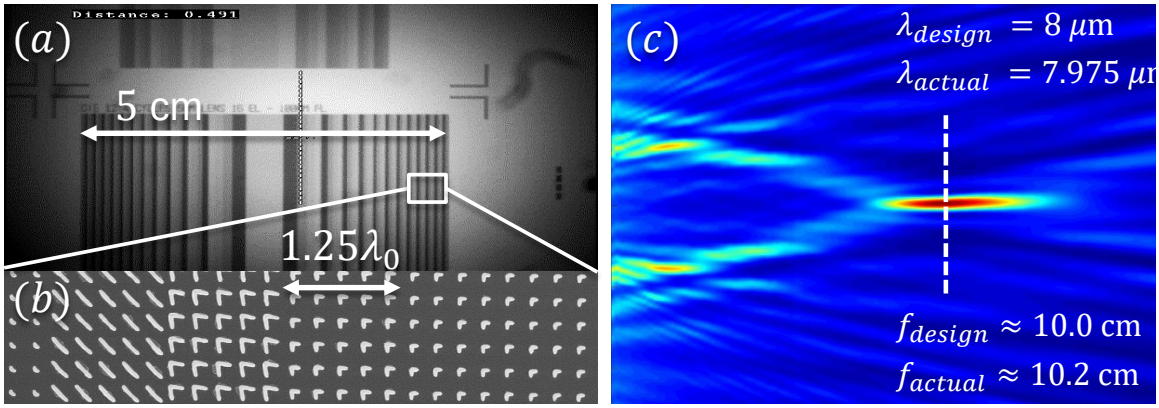

Figure 9. (a) SEM images of a 16-element, 5cm V-antenna metasurface lens, with (b) a magnified region near the extents of the lens, showing the sharp gradient of antenna elements. (c) The measured depth-of-field (DOF) of a lens with designed focal length of $f = 10$cm at 8μm, showing excellent agreement with COMSOL simulations.

Results of the simulated parametric profiles were already shown in Figure 2(b,c) for the (b) amplitude and (c) phase, and match extremely well with that from the seminal work[1], though, with a slight shift in both profiles due to the slightly larger trace width (up from $w = 220$nm). From here, in collaboration with Sandia National Laboratories, ten rectangular variants of the V-antenna lens and nine metasurfaces were fabricated, with slightly varying design parameters, such as number of discretizations (N), focal lengths (f), periodicities (a, which affects areal density) and number of phase cycles (which affects diameter and lens speed). A 16-element lens, a magni-

fied region near the extent of the lens, and the measured depth-of-field (DOF) of a lens with designed focal length of $f = 10$cm are shown in Figure 9(a-c), respectively. The experimental setup and additional characterization studies will be covered in Chapter V, but for now it suffices to say that that a FLIR camera was used to scan the DOF around the focal point, and Figure 9(c) shows the calculated average of the translated 2D slices. Clearly, the DOF exhibits that the metasurface development process was successful, coming within 0.2cm of the designed focal length and 25nm of the designed wavelength. This, then, provided a high confidence in the modeling and simulation accuracy of COMSOL, and served as the impetus for proceeding with the GA development.

3.3 Development and Verification of GA Foundation

3.3.1 Definition of Verification for a GA Routine.

Any good and reliable algorithm must undergo the same steps of verification and validation during the development process. Verification asks the question "was it built correctly?" and, once verified it was built correctly, the next question is of validation, asking "does it work as intended?" This section focuses on the former, and the latter will be reserved for the results presentation in Chapters IV and V. For simulation code, one means of carrying out a verification task is ensuring that the inputs to a simple, well-understood model produce the expected outputs. In the previous section, it was validated that COMSOL could accurately reproduce scattering amplitudes and phases from a scatterer, and so whichever outputs arise from a simple GA-based model should be trusted—what is more important in this task is that they behave as expected. For example, a scatterer of length L which is electrically small ($L << \lambda$) should behave like a dipole, with equal forward and backward scatter; whereas a scatterer with $L > \lambda/2$ should have strong side lobes

from multiple current peaks distributed across the surface[37].

With the GA routine, the measure of verifiability can be if it repeatably can converge to a near-global extremum. While the GA optimization is stated as a "global" optimization routine, this is only true in the most simple or ideal cases; in reality, there is no way to sensibly verify the GA has hit a global value, because an exhaustive search of the both the parameter and solution spaces is what the user is attempting to avoid in the first place by implementing the GA. If the entire space is not sampled, then there is a chance a better solution might exist! There are two mechanisms that force the GA to operate with this global "limiter," but both are related to the same root cause of user compromise. The first mechanism is a termination threshold. Thresholds are set in the GA to terminate the routine before it could possibly search the entire parameter space, whether it be because a minimum fitness value was reached, or because a reasonable number of generations have passed. As stated in Chapter 2, the size of the parameter space of, for example, a 4×4 voxel grid is $2^{4\times4} = 65536$, and every possible grid layout cannot feasibly be assessed. The second is population size. This determines the diversity inherent in the GA routine, such that a larger population will permit more chromosomes to be in play for the selection process; as will be explained later in this section, even low-scoring individuals have a probability to be chosen as a parent. Lower populations tend to prematurely converge to a local extremum from which it cannot escape, while higher populations take exponentially larger resources and will likely converge much slower[66]. Unfortunately, as many optimization problems are unique in nature, rules for determining the optimal population sizes tend to be founded on empirical evidence, based on the specifics of a given problem type[67, 68].

So the best an effective GA routine can do to be "verifiable" is to demonstrate the capacity to progressively attain larger and larger magnitudes of the extremum as

the routine advances forward. This entails confirmation that the GA is kicking itself out of local extrema and subsequently converging, which could be examined quantitatively by relatively sharp jumps in the fitness factor followed by regions of near-zero derivative. Repeatability is also a key verifiability factor. While each GA routine will be unique—different random starting population, different parental selection and mutation probabilities—statistically the routine should converge to a reasonably similar solution, given the terms in the fitness function. In the case of the OOP Huygens metasurface, this means attaining a high transmittance at the target phase, within tolerances, for each GA routine run. The voxel grid layout of the optimal design may be vastly different in its topology, but the end result is the same.

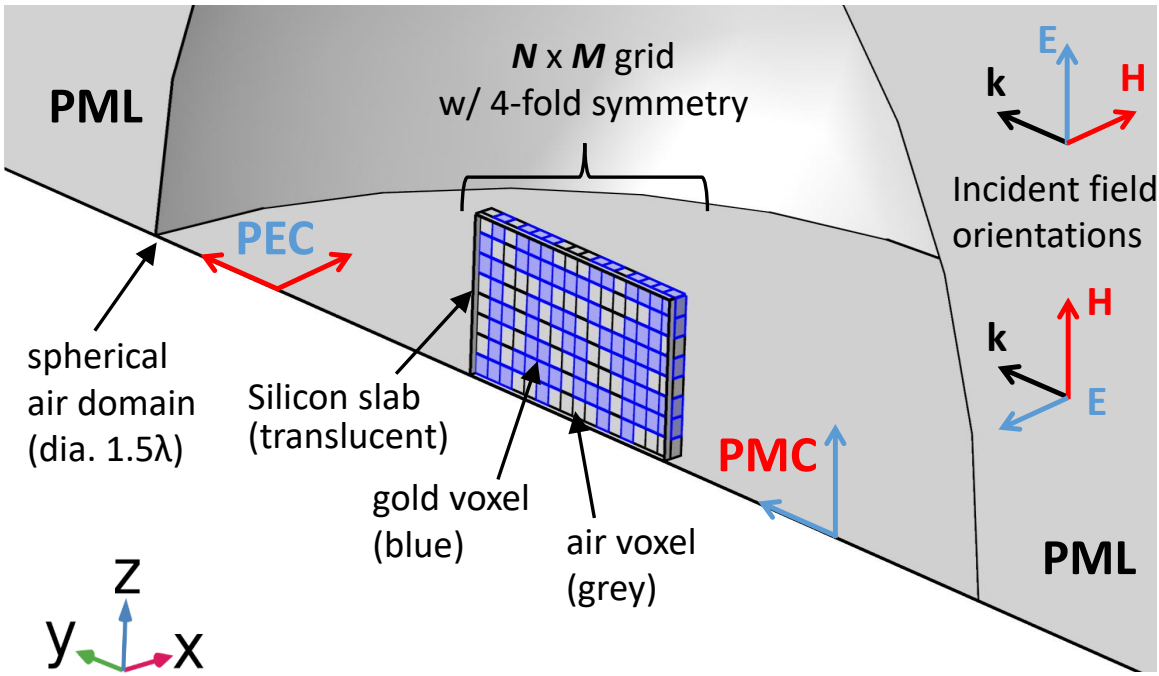

Figure 10. A general $M \times N$ COMSOL model for verifying GA routine, consisting of two grids with a Si spacer. This model was used to ensure the routine could converge the binary voxel layout to a Huygens-like scatterer. The model has 4-fold symmetry, so only the upper quadrant is simulated, and the BC's are either PEC in the $x/y-$plane and PMC in the $y/z-$plane for TM-polarized light (as shown) or vice-versa for TE-polarized light (not shown). The slab is translucent so as to detail the voxel distribution.

3.3.2 Verification Model.

The GA routine was verified using a simple model of two back-to-back grids suspended in air, as shown in Figure 10. The goal of the verification model was to optimize the grid layout so that the structure behaved like a Huygens source, seeking a maximum forward intensity of the scattered far-field intensity ($|\overline{E}^f_{far,s}|^2$), and maximum difference between the forward and backward intensities, given as

$$|\overline{E}^{\Delta}_{far,s}|^2 = |\overline{E}^f_{far,s}|^2 - |\overline{E}^b_{far,s}|^2, \tag{10}$$

where, for brevity, the (far, s) designation to the scattered far-field will be dropped. The corresponding multi-objective fitness function for this desired optimization is given by

$$F(\overline{E}^f, \overline{E}^{\Delta}) = w_f \frac{|\overline{E}^f|^2 - \min|\overline{E}^f|^2}{\max|\overline{E}^f|^2 - \min|\overline{E}^f|^2} + w_{\Delta} \frac{|\overline{E}^{\Delta}|^2 - \min|\overline{E}^{\Delta}|^2}{\max|\overline{E}^{\Delta}|^2 - \min|\overline{E}^{\Delta}|^2}, \tag{11}$$

where $w_{f,\Delta}$ are the weights for the forward and difference terms, respectively, and $w_{\Delta} = 1 - w_f$. The intensities are normalized by the maximum calculated from across the entire population. It is best to maximize the difference term $|\overline{E}^{\Delta}|^2$ rather than minimize the backward term $|\overline{E}^b|^2$, as this can lead to solutions that can have low $|\overline{E}^b|^2$, but at the cost of low $|\overline{E}^f|^2$. Of course, the weights could be adjusted, but there would still some compromise between these two directions. By maximizing $|\overline{E}^{\Delta}|^2$ instead, the two terms mutually support each other: an increase in $|\overline{E}^{\Delta}|^2$ leads to either an increase in $|\overline{E}^f|^2$ or a decrease in $|\overline{E}^b|^2$, or both; meanwhile, an increase in $|\overline{E}^f|^2$ may either increase $|\overline{E}^{\Delta}|^2$ or have no effect, if $|\overline{E}^b|^2$ also increased the same amount—no other trends are possible.

Each grid was a 16×16 square array of square voxels 177nm on edge, with a

320nm-thick Si layer sandwiched in between. The spherical air domain of diameter $1.5\lambda_0$ was surrounded by a PML, and perfect magnetic conductors (PMC's) and perfect electrical conductors (PEC's) were used on the boundaries to invoke a 4-fold symmetry, meaning only the upper 8×16 half of a single grid needed to be simulated. A background plane wave was launched as

$$\overline{E}_{bg} = (\cos(\phi)\hat{x} + \sin(\phi)\hat{z})E_0 e^{-i\frac{2\pi}{\lambda_0}y}, \tag{12}$$

with ϕ as the polarization angle, such that $\phi = 0°$ for the transverse electric (TE) polarization, also known as s-polarization, and $\phi = 90°$ for the transverse magnetic (TM) polarization, also known as p-polarization. The background field amplitude E_0 was set to 1V/m.

The desired fields were sampled from probes placed directly in front of and behind the structure, and the far-field was calculated at the air/PML boundaries from the near-field boundary values using a Stratton-Chu projection:

$$\overline{E}(\overline{r}) = \hat{r} \times i\frac{1}{2\lambda} \iint [\hat{n} \times \overline{E}' - Z\hat{r} \times (\hat{n} \times \overline{H}')]e^{i\frac{2\pi}{\lambda}\overline{r}' \cdot \hat{r}} d\overline{S}', \tag{13}$$

where the tangential fields $\{\overline{E}', \overline{H}'\}$ are taken from the homogeneous spherical boundary $\overline{S}'$ separating the air domain and the PML, $\hat{n}$ is the unit normal to $\overline{S}'$, $\overline{r}'$ is the vector from the origin to $\overline{S}'$, Z is the impedance $\sqrt{\mu/\epsilon}$, and $\hat{r}$ is the unit vector from the origin to the point $\overline{r}$. Specific implementation of this equation for simulation full-scale optical devices will be detailed in Chapter 5.3.

Aside from the GA verification, this initial model could also provide some level of validation that an OOP scatterer would be superior to a planar geometry. To starkly

demonstrate this, the field values were normalized against fields of an optimized in-plane V-antenna, using geometries from the COMSOL validation work. This leads to the final form of the results used in the GA, given as the differential cross-section (DCS), relative to the V-antenna (dBv):

$$\sigma = \lim_{r\to\infty} 20\log_{10}\left(4\pi r^2 \frac{|\overline{E}_{far}|}{|\overline{E}_i|}\frac{1}{|\overline{E}^V_{far}|}\right) \approx 20\log_{10}\left(4\pi\frac{|\overline{E}_{far}|}{|\overline{E}^V_{far}|}\right), \tag{14}$$

where the far field is taken at $r = \pm 1.0$m for either the forward or backward scattering, and the incident (background) field is 1V/m. Thus, $\sigma = 0$ dBv indicates the same value as that of the V-antenna. Being on a dB scale, the min() functions in Equation (11) were required to account for negative values, and the denominators therefore scale the normalization so that $0 \leq F \leq 1$. Results from the field probes are input into Equation (14), providing everything needed for the GA calculation of Equation (11) and the verification task.

3.3.3 Development of GA Routine.

While COMSOL offers an Optimization Module, it only contains gradient-based and random search optimizers, which are all designed to halt computation upon finding a localized solution. However, through the use of their Java-based "Application Programming Interface (API)," algorithms can be written to enable a variety of additional functionalities, such as a global optimizer. The API is a feature of the "LiveLink for MATLAB" module, and is an means to communicate between MATLAB between COMSOL. The overarching process flow for the GA routine shown in Figure 11 describes how this communication takes place, and which program controls the process. Processes entirely accomplished in MATLAB are shown in light blue, entirely run in COMSOL are in red, and those which involve some means of transferring information

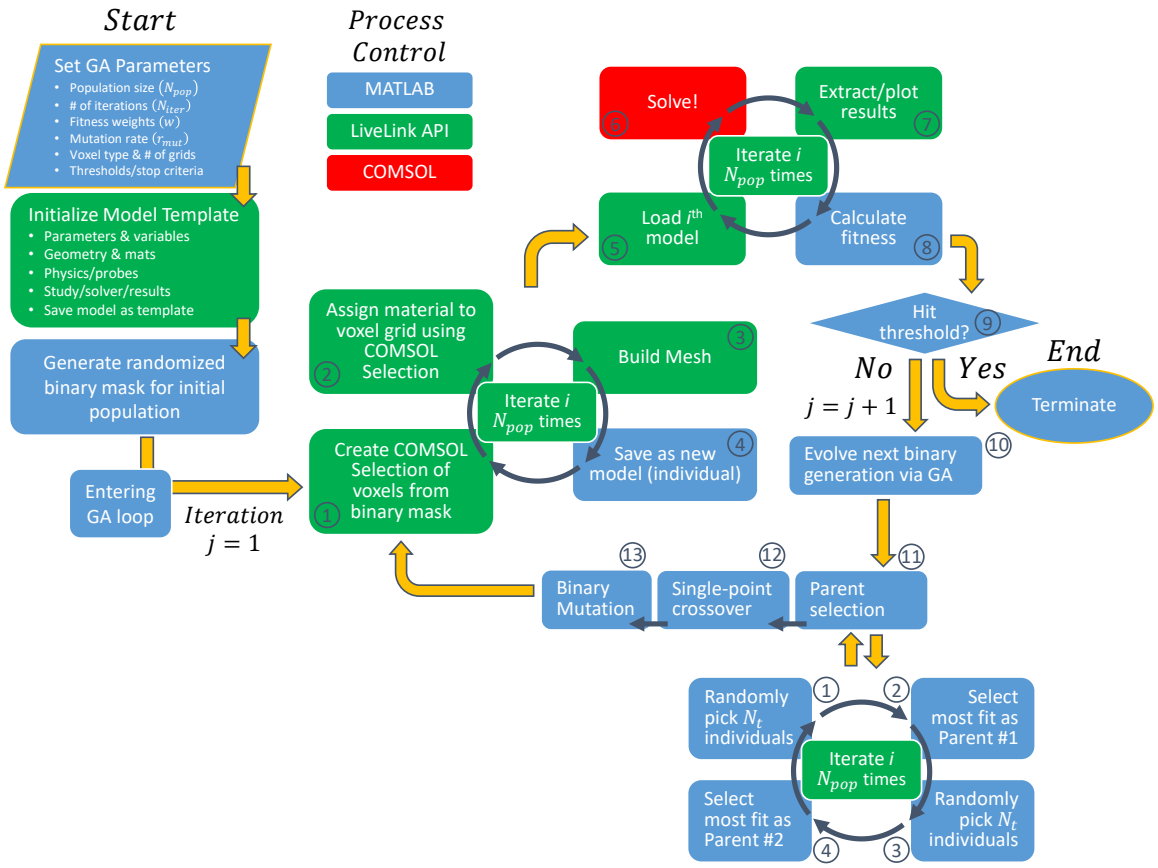

Figure 11. Overarching process flow for the GA routine based on communication between COMSOL Multiphysics and LiveLink for MATLAB. The algorithm is built using LiveLink's Application Programming Interface (API), and control is color-coded according to whether the processes is handled by MATLAB (blue), COMSOL (red), or some combination of the two through the LiveLink interplay (green). Numbered steps are described in the main text.

between the two are controlled by LiveLink, in green. Herein, the programming of the GA routine is detailed; as a consequence, this subsection will continually refer to this chart.

Building a COMSOL model follows a logical process that often must be accomplished in a serial order, with only a few exceptions. The boundary conditions in the cannot be applied, for example, without the physical boundaries created, obviously. In general, the build process is: (1) define the parameters and functions, (2) build the geometry, (3) set materials, (4) establish the physics environment, (5) create the

mesh, (6) develop the study and (7) initialize the outputs. From here, the model is ready to be solved.

A rational approach to programming these processes is to create an individual sub-routine (i.e. a MATLAB function or script) to handle each one. This assists in debugging and reducing sweeping changes to the model to a single parameter switch. For example, altering the geometry of the voxel from, say, a square to a hexagonal building block requires a great many changes to the model across the entire process flow: boundary conditions change, the mesh changes, material allocations change, etc. However, if each build process is tied to a function call sub-routine, these changes can be fed into all the sub-routines as an input argument in the form of a string or a numerical value. For this voxel geometry example, simply designating a switch will redirect entire sub-routines:

```
function [] = geom_build(voxel_type)
  switch voxel_type
    case 'square'
        % geometry commands for square voxel
    case 'hex'
        % geometry commands for hex voxel
        end
```

This function "geom_build()" contains the Java-based commands specific to COMSOL for building the geometry. For building a simple square pixel, for example, these commands would be:

```
voxel=model.component('comp1').geometry('geom1').create('voxel', 'Block');
    voxel.set('size',{'wx', 'wy', 'wz'});
    voxel.set('pos',{'posx', 'posy',  'posz'});
model.component('comp1').geometry('geom1').run('r1');
```

where w is the pixel width in $x/y/z$, and $posx/posy/posz$ are parameters representing the center position of the first voxel in the array, which depends on w. To create a

square voxel array, this geometry is copied and shifted by period w to the desired size, and the entire array is extruded along the surface normal. In general, the sizing and position of each voxel in an array can be automated for a shape of any $n-$fold symmetry and any applicable tiling.

The rest of the model build goes similarly as described, using the API language to build the rest of the unit cell geometry, the physics, study, and outputs as indicated in the previous subsection. However, before the materials can be assigned and the mesh built, the voxel grid must be linked to the binary chromosome.

At this point, it is important to discuss how COMSOL handles indexing. It terms a 3D volumetric geometry as a "domain," a 2D surface as a "boundary," and a 1D line as an "edge," assigning each geometry an index, with independent sets of indices for each dimension. These indices are used to identify which geometries will possess certain features, such as particular boundary conditions (e.g. which are a PMC and which are a PEC), mesh control (where to refine) and material assignment. However, COMSOL assigns these indices by rastering the model, starting at $\{x, y, z\} = \{-\infty, -\infty, -\infty\}$ and shifting in the $+y-$direction while scanning in z until the $y/z-$plane is fully indexed, and then shifting vertically in the $+x-$direction. This means that an array of voxels are indexed in a different numerical pattern if they lie in one plane than if they lie in another, as shown in Figure 12(a) for a 3-face unit cell. The first grid to be indexed would be the floor, as it counts from $\{1 \rightarrow M \times N\}$ (here $M = 9, N = 7$), and then only the first row of the grid in the $x/z-$plane to the right from $\{M \times N + 1 \rightarrow M \times N + M\}$, then the first row of the grid in the $x/z-$plane on the left $\{M \times N + M + 1 \rightarrow M \times N + 2M\}$, then back to the second row on the right grid, and so on. The indexing for this 3-face unit cell will be different for a 3-face unit cell oriented on other faces, such as if they were aligned in the $\pm x/y-$plane instead of $\pm x/z$, or any other unique combination of grids and alignments. While

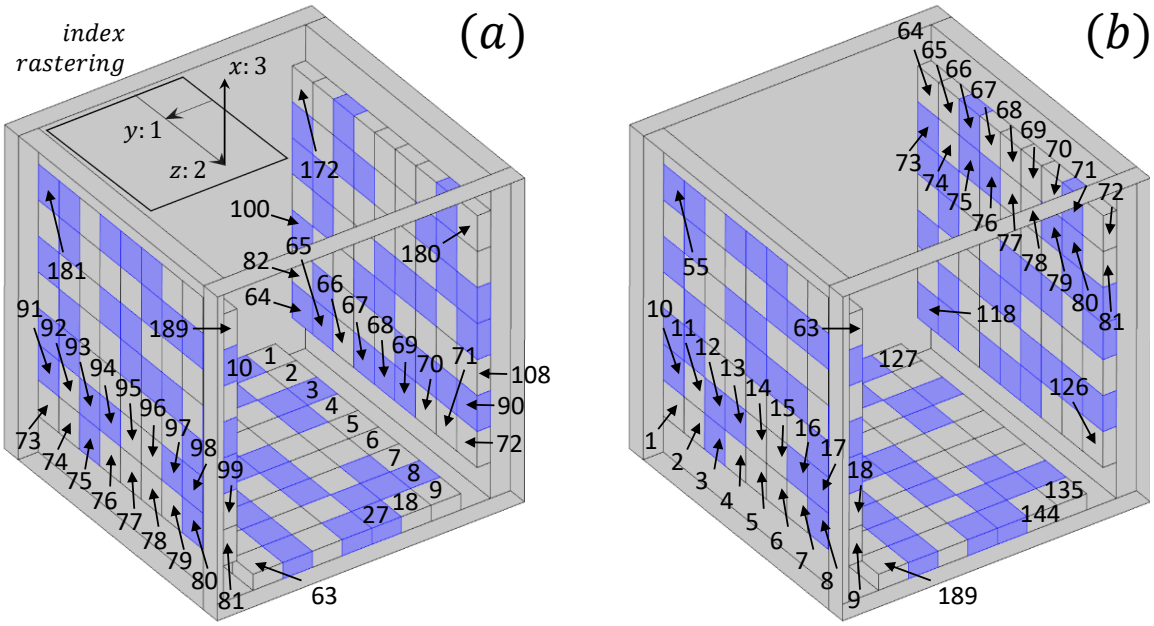

Figure 12. (a) An example of how COMSOL indexing of domains can present issues for multiple grids arrayed on faces which are aligned along planes with different basis vectors. COMSOL counts in a given fashion, $y \to z \to x$, generating an indexing scheme which is not simply ordered for certain grid orientations, such as shown in the upper right and left walls. (b) By changing how COMSOL indentifies these domains (see Fig. 13), they can be ordered consecutively for more efficient matrix operations in the GA routine.

this is not an insurmountable accounting task, it certainly complicates the tracking algorithm. Ideally, the indexing convention would be driven by the user; however, COMSOL does not offer this capability inherently. So a solution to this problem was created by utilizing the COMSOL "selections" and "variables" feature. A selection can group numbered indices together either explicitly by its given index, or spatially by targeting all geometries that interact with a virtual cubic, cylindrical or spherical region, and can be programmed using the API to quickly reference complex geometric sets. To accurately pinpoint which domain index belonged to each voxel in the grid, an infinitesimally thin cylindrical geometric selection was placed at the center position of each voxel, selecting only the particular domain of that voxel, and no other. That voxel was then assigned a new variable according to how many grids were formed, and on what faces. An illustration of this is shown in Figure 13, where a voxel which

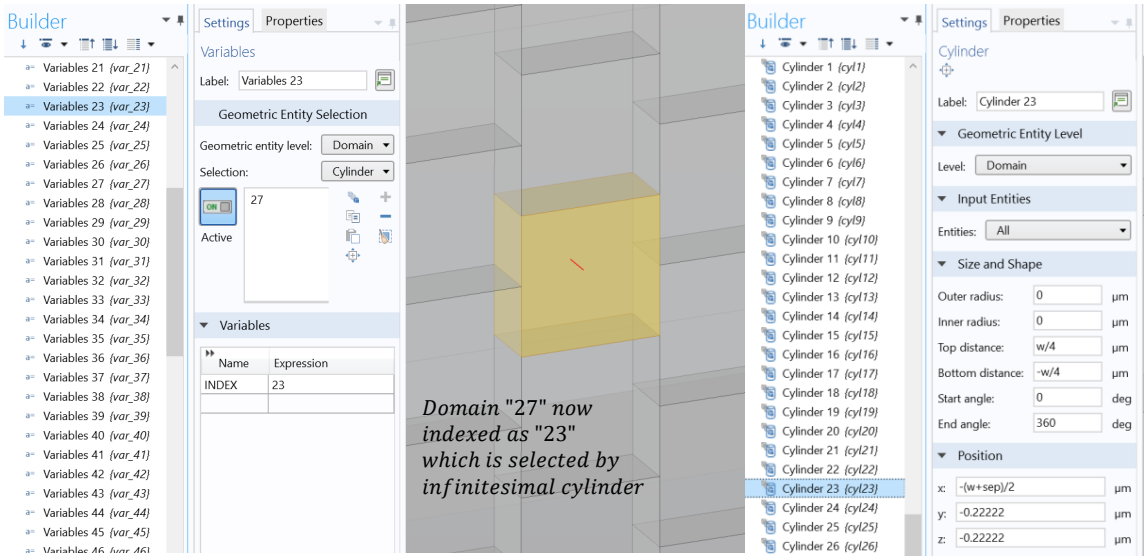

Figure 13. In this example of how to rename the COMSOL domain index, a variable is created in the program (left) which assigns domain 27 (center) to another numerical index (23, in this case), such that it is ordered consecutively with its neighbors in a given grid. Then an infinitesimally small cylindrical selection is created (right) which is labelled as this new index, and selects only this domain, such that all $M \times N$ voxels can be predictably accounted by the starting index for that particular grid.

COMSOL indexed as domain number 27 could now be referenced using a geometric selection named "cyl23". Continuing from the example above, this 3-face unit cell can now be indexed in a manner seen in Figure 12(b), such that each face can be assigned a number n, and the n^{th} grid counts from $\{(n-1)M \times N + 1 \rightarrow nM \times N\}$. Note how the right grid now counts downward; as will be explained later, opposing faces must be flipped to account for the 90° shift in metal deposition angle, and this new indexing allows a direct mapping from voxels $\{1 \rightarrow M \times N + 1, 2 \rightarrow M \times N + 2, ...\}$, even though they belong to a 2-fold rotation-reflection point group[1].

These group selections are a key enabler of the GA routine, serving as a straightforward means to link the voxel grid to the binary assignment of 1's and 0's for a given chromosome—representing the particular layout of metal or air for a given

[1]Using Schoenflies notation, this is an S_2 point group, with 2-fold rotation with reflection across a plane perpendicular to the rotation axis[69].

model—and then tracking the evolution of the chromosomes. How this works through the API for the initial population is that a random number is generated to determine how many m voxels will contain metal for a given individual, then this number is randomly dispersed as m 1's throughout an otherwise zero-valued array of size $M \times N$. The numbered indices of the 1's are linked to a new group selection of all the named cylindrical selections created previously, such that if the binary chromosome reads as $\{0, 1, 0, 0, 1, ...\}$, then a selection with {"cyl2", "cyl5", ...} is created. This "selection of selections" is assigned to a metallic material in COMSOL, and the transferral from a randomly generated binary array in MATLAB to a metalized voxel layout in COMSOL is complete. This accomplishes steps (1) and (2) in the GA process flow chart in Figure 11, and the following script is a simplified version of how this is programmed using the API:

```
arr_size=M*N;
arr_count=1:1:arr_size;
for i=1:Npop % Number of individuals in population
    m(i)=round((arr_size-1)*rand(1)+1); % Number of voxels with metal
    for j=1:m
    % Idenfity indices which are metal for initial population:
    A_init{i}(j)=arr_count(round((arr_size-1)*rand(1)+1));
    % Set j^th index of A_init equal to 1 (metal) for the i^th individual of ...
        initial population of entire GA routine (A_pop (:,:,1)):
    A_pop(i,A_initl{i}(j),1)=1;
    % Tie cylinder selection to index:
    vox_list{j}=sprintf('cyl\%03g',A_init{i}(j));
    end % Loops through all j metallic genes (1's) in i^th individual
% Create selection of all metal voxels:
vox_on=model.component('comp1').selection.create('voxel on', 'Union');
% Groups all metal voxels from vox_list
vox_on.set( 'input',vox_list);
% Link selection to created COMSOL material:
Au=model.component('comp1').material.create('Au', 'Common');
Au.selection.named('vox_on');
end % Loops through i^th individual
```

Though left out for brevity, this scripted loop must also contain the meshing

commands (3), and then save off each model as a new individual in the population (4). The code then moves into the next iterative loop, where it loads the first individual of the population (5) and solves using the command "model.sol('sol1').runAll" (6). The solver best for large sparse linear systems of equations often found in complex electromagnetics models is the parallel sparse direct solver (PARDISO); the power of which will be evident in the 3D unit cell where it was used to solve multiple GA routines at once using multiple COMSOL instances.

After each solution is determined, the results (forward CS, σ_f; backward CS, σ_b) are calculated from the fields (7) as given in Equation (14) and passed onto the main GA calculations. After the fitness in Equation (11) is calculated for the entire population (8), the max fitness can be checked against a threshold value for a decision to continue or terminate (9). In all GA routines in this work, the only termination threshold used was a maximum number of iterations—usually between 20-40, depending on well-understood the likelihood a particular design would converge. For the decision of a continuance of the GA routine, the next step is the evolution of the binary chromosome (10).

The first step in the evolutionary process is to select parents (11), and this step tends to be the most critical and modified of all other steps. A stochastic "tournament" selection strategy was chosen for the GA routine, which starts with taking a small random subset of N_t individuals from the current population. These N_t individuals are ranked according to their fitness, and the most fit of the subset is chosen as the first parent of the i^{th} child, P_i^1. This process repeats to choose a second parent, P_2^i; however, the algorithm was modified to prevent $P_2^i = P_1^i$, as this would inhibit diversity early-on in the GA, effectively relying only on the sparse mutations to evolve the child. For small populations, the tournament size is often chosen to be the minimum, $N_t = 2$, because using a large N_t will tend to consistently contain the

highest performers, reducing the diversification. However, the population size for the verification model must be on the order of the chromosome length of $8 \times 16 = 128$, so it was chosen as $N_{pop} = 120$. This is too large for $N_t = 2$, as it would too often grab individuals in the lower half of fitnesses, which would make the GA converge too slowly, or worse, stagnate; therefore it was determined $N_t = 5$ or $N_t = 6$ worked adequately to promote diversity while averting stagnation.

There are a great many extensions to the simple GA optimization routine which can be availed in refining the parental selection process. The first used in this work is the *elitist strategy*, which tries to retain the top performers from each generation. For stochastic selections, there is a chance that either the most fit individuals are not chosen to pass on their genetics, or that they are chosen but the child is evolved to a significantly poorer fitness. Thus, the most beneficial genes are lost. To prevent this loss, some small percentage of the top performers are retained as new children, and not mutated in any way. For the works in this research, only the singular best fit of a generation is kept. Another extension employed is a *steady state* replacement, which—as defined in this work—as a similar approach to the elitist strategy, but allowing mutation of the retained population. This ensures some level of retention of good genetic traits, while still permitting the underlying functionality of global optimization that the mutations empower. One last extension utilized was not found in references, but it involves the injection of a new random population if the average fitness of the population fails to meet a given threshold after a certain number of iterations. This threshold was generally set to a mid-range value, around $0.50 - 0.75$, and was set to replace the worst-performing $N_{inject} = 25\%$ after $10 - 15$ iterations.

After all parents are identified, the crossovers are performed (12). Again, there is some variation which can be applied to this step, such as a multi-point crossover, or a fully stochastic method picking a random gene from the two parents. By far,

the most popular choice is the single-point crossover, which is also utilized in this work, as it was found to provide enough genetic evolution to converge in a timely manner. No other crossover strategies were investigated. The crossover point was chosen randomly for each child as follows:

```
Nxo=Npop-Nss % # of crossover population is Npop - steady state population
p_xo=randi([1 arr_size],1,Nxo);  % Randomizes Nxo crossover points between {1, ...
    arr_size}
for i=1:Nxo
    A_xo(i,:)=[A_sort(parent_index(i,1),1:p_xo(i)), ...
        A_sort(parent_index(i,2),p_xo(i)+1:end)];
end
```

where the number of crossovers (N_{xo}) are determined by subtracting off from N_{pop} the highest performers by the steady-state strategy, A_{xo} is the matrix of chromosomes for each population after crossover, A_{sort} is the entire current generation sorted by highest fitness and $parent\_index(i, 1:2)$ is the index that picks the parent chromosome out of A_{sort}, with N_{pop} and arr_size as previously defined.

With all the crossovers performed, the last step is to modify A_{xo} by applying random mutations (13). For this work, mutation rates were kept to between $1\% \leq r_{mut} \leq 5\%$, as it was determined that higher rates would prevent timely convergence due to the extremely large design space. This a simple script, written as:

```
A_mut=A_xo;
Nmut=floor(arr_size*Npop*r_mut) % Mutation rate:  r_mut
for i=1:Nmut
    mut_ind(i,:)=[randi([2, Npop]); randi([1 arr_size])];
    if A_mut(mut_ind(i,1),mut_ind(i,2))==1
        A_mut(mut_ind(i,1),mut_ind(i,2))==0 % If 1, make 0
    else
        A_mut(mut_ind(i,1),mut_ind(i,2))==1 % If 0, make 1
    end
end
```

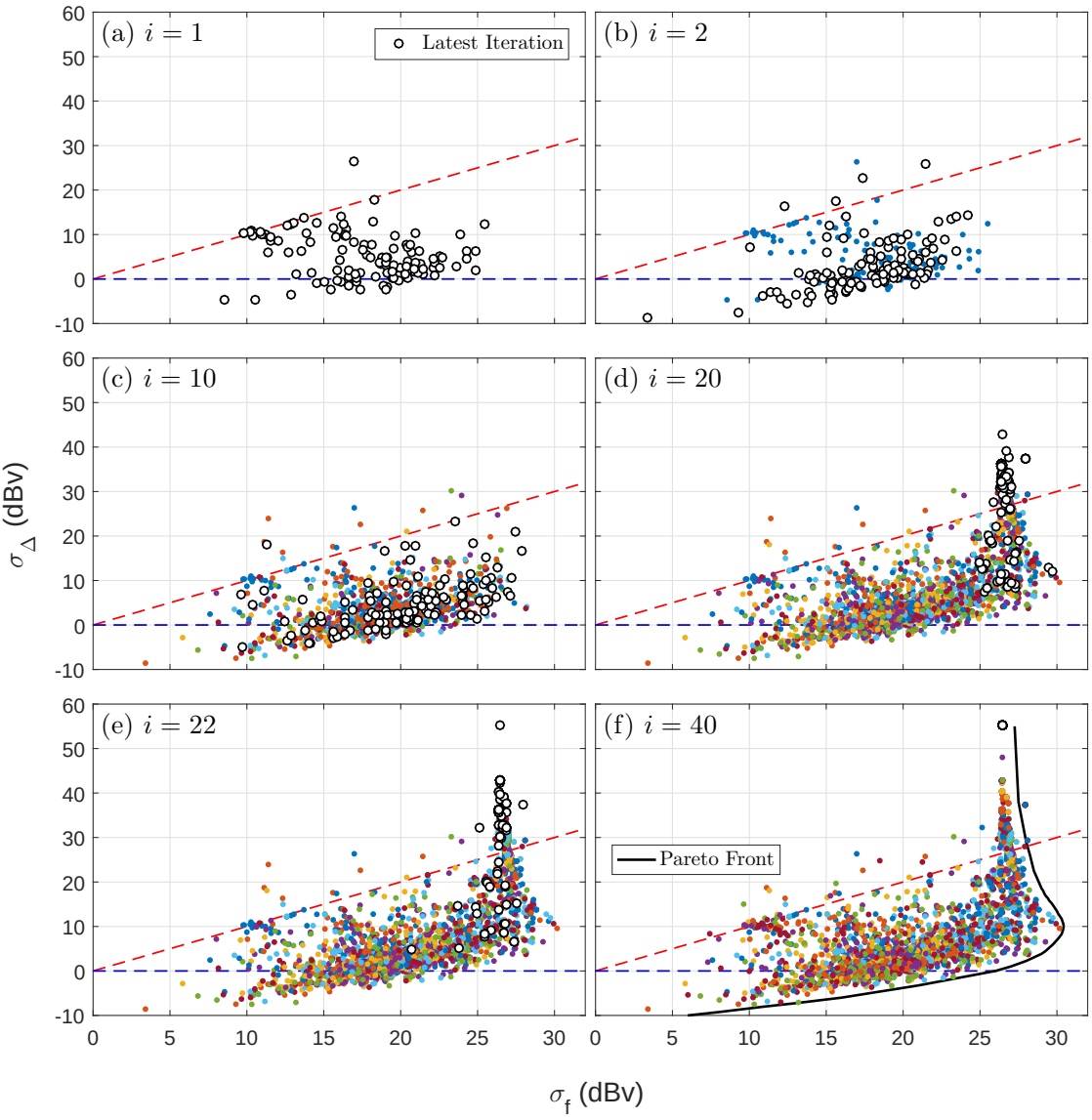

Figure 14. Snapshots of the 2D solution spaces extracted at various iterations of a $N_{pop} = 120, N_{iter} = 40$ GA routine for the verification model. The horizontal axis shows the forward scattering DCS σ_f, and the difference between forward and backward scattering DCS σ_Δ is on the vertical axis, both in dBv, referenced to a nominal V-antenna intensity. The red line indicates designs with a Huygens-like behavior, having a value of 0dBv in backscatter, while the blue line at $\sigma_\Delta = 0$ indicates a dipole-like behavior of equal parts forward and backward. This shows how the individual models—initially sparsely populated in early iterations (a,b) converged toward the optimal solution at $j = 22$ (e), but if the GA were continue to run the population would be comprised entirely of the optimal value by $j = 40$ (f).

where the key subprocess here is the stochastic $mut\_ind(i,:)$, which chooses a random chromosome (index 1) and a random gene in that chromosome (index 2) to mutate, leaving out the first chromosome from any mutation as per the elitist strategy. The A_{mut} matrix consists of all evolved children, and fully represents a new generation. After some steps of backwards accounting of A_{mut} to determine which "1" genes belong to which COMSOL domain indices (essentially ① for the new generation), the "voxel on" selection is updated with the new "vox" selections (essentially ②), where at this point the process has come full-circle and continues to iterate as previously described.

3.3.4 GA Input Parameters.

Many test runs were carried out for the verification model, examining the effects of the various GA inputs: fitness weights (w), population sizes (N_{pop}), max iterations (N_{iter}), mutation rates (r_{mut}), tournament selection size (N_t), steady-state percentage (N_{ss}/N_{pop}) and injection size (N_{inject}), in addition to some model variations, such as mesh strategy, freespace wavelength (λ_0), voxel thickness (t) and incident polarizations (E_x, E_z from Equation (12)).

It was found that some input parameters were ideal, and remained constant or within a small range. The selection size was set to $5 \leq N_t \leq 6$ to sample a large enough set of potential parents to find a statistically "good" candidate, and the steady-state percentage was kept to $20\% \leq N_{ss} \leq 25\%$ to ensure the best genetic traits were retained. This assisted heavily in convergence, with a draw-back of the solution space possessing less diversity—these types of trade-offs are common in evolutionary algorithms.

The population size remained between $80 \leq N_{pop} \leq 120$ to keep computational resources low. The smaller N_{pop}, the less confidence a converged solution will be the

global solution; this is an unavoidable problem in this non-analytical approach, as the large design spaces cannot be sufficiently probed by a small N_{pop}, yet a large N_{pop} is not feasible for reasonable convergence. However, as long as $N_{pop} > M \times N$, along with the 1-5% r_{mut} as mentioned previously, then there was confidence that a sufficiently strong *local* solution would be found within 100-150 hours and within $N_{iter} \leq 40$.

Any other inputs or strategies which warrant discussion or were changed from values given in Chapter 3.3.2 will be covered either in this section or in future sections on the 3D cavity structure.

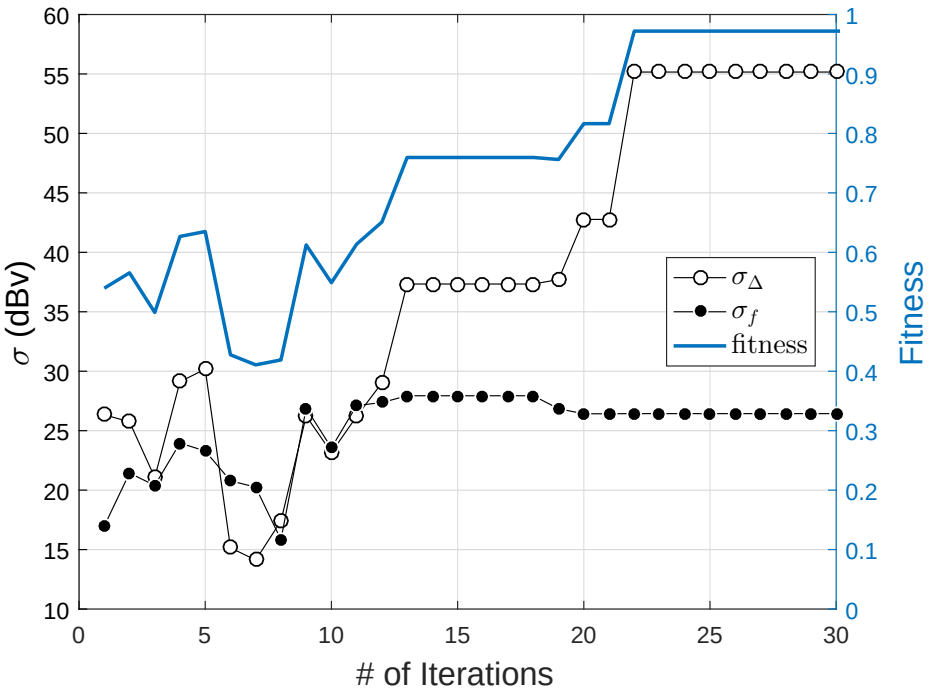

Figure 15. Convergence plot from the GA routine shown in Figs. 14(a-f), tracking the most fit solution for each iteration (line, right axis) and separated into the individual values of σ_Δ (white dot, left axis) and σ_f (black dot, left axis). While some higher σ_f values existed, the σ_Δ was lower; however, due to the weight on the latter, this was not the most fit.

3.3.5 Verification Results.

The 2D solution spaces of a full $N_{pop} = 120, N_{iter} = 40$ GA routine are displayed in Figures 14(a-f), using $w_f = 0.2, w_\Delta = 0.8, \lambda_0 = 8\mu$m, $t = 150$nm and a $p-$polarized incident wave (E_z). The horizontal axis shows the forward scattering DCS, σ_f, and the difference between forward and backward scattering DCS, σ_Δ, is on the vertical axis, both in dBv, referenced to the V-antenna intensity of 4.261×10^{14} (V/m)^2. Each series of dots represents a unique generation of solved models, with the most current generation in the GA showing as large white dots with black borders. The red line indicates designs with a Huygens-like behavior, having a value of 0dBv in backscatter, while the blue line at $\sigma_\Delta = 0$ indicates a dipole-like behavior of equal parts forward and backward. Ideally, the solutions should be as far above and right of the line as possible to give high forward and minimal backward scattering.

The graphs demonstrate how the solutions evolve from (a) the initial population $j = 1$ to (b-e) future generations at $j = 2, 10, 20, 30$ and (f) the final iteration at $j = 40$. A quick survey of these snapshots of the evolution indicate the GA routine is working as intended. At first, the randomized initial population is scattered all over the range, with most lying somewhere between dipole-like and Huygens-like behavior. At $j = 2$ there was an average decrease in the average fitness, as the high density of low-performing designs provided a viable opportunity to be selected. By $j = 10$, some creep in mass towards the upper right is seen, and by $j = 20$ the mass has formed a spike near $\sigma_f \sim 26$ dBv. By $j = 22$, the maximum fitness is attained at solution point $P_1 = \{\sigma_\Delta = 55.16, \sigma_f = 26.44\}$dBv, and by $j = 40$, all solutions have migrated to this maximum, which appears to be either a strong localized maximum or the global maximum.

The routine was run for 40 iterations as demonstration of global convergence; obviously, this routine would have benefited from a threshold in which it terminated

after no change in the fitness has been detected for several iterations, under the constraint that the fitness was acceptably high. Examining the convergence chart in Figure 15, the most fit solution for each iteration was tracked (line, left axis) and separated into the individual values of σ_Δ (white dot, right axis) and σ_f (black dot, right axis). The fitness most closely follows the path of σ_Δ because of the heavy weight on the difference, and converges at $j = 22$. The region between $j = 13$ and $j = 19$ is a good example of a local maximum, and how evolutionary operations kicked the stagnant solution out into a different region in the solution space, one with ultimately a higher fitness. Care must be taken in any threshold implemented, such that weaker local maximums do not trigger termination; for example in this case, if the threshold was set to require both $F_j > 0.9$ and $|F_j - F_{j-1}| < 0.1$, then this routine would have terminated properly around $j = 25$.

A multi-objective fitness will lead to multi-dimensional solution spaces, which makes difficult defining the concept of a "best" solution. Note in Figure 14(f) the presence of many solutions with higher σ_f, but lower σ_Δ, such as the one at $P_2 = \{\sigma_\Delta = 9.57, \sigma_f = 30.17\}$ dBv. The solutions between the two regions of max σ_f and max σ_Δ have formed a Pareto front, defining the boundary of possible solutions for the given GA and model parameters. Any of these solutions are the "best" in terms of some manner of trade-off between the two σ terms, and the design choice is subjective to the problem at hand. Changing some of these parameters can change the shape of the front—such as the fitness function weights, or the grid resolution (voxel size or shape)—as will be demonstrated in the final 3D design.

Examining the full grid layout of the P_1 solution in Figure 16(a) (mirrored from the symmetrical grid to show the full 16×16), there are some partially and fully segregated geometries that represent a mix of dipole-like straight sections and loop-like curvatures—two key ingredients to forming a Huygens-like source. The out-of-

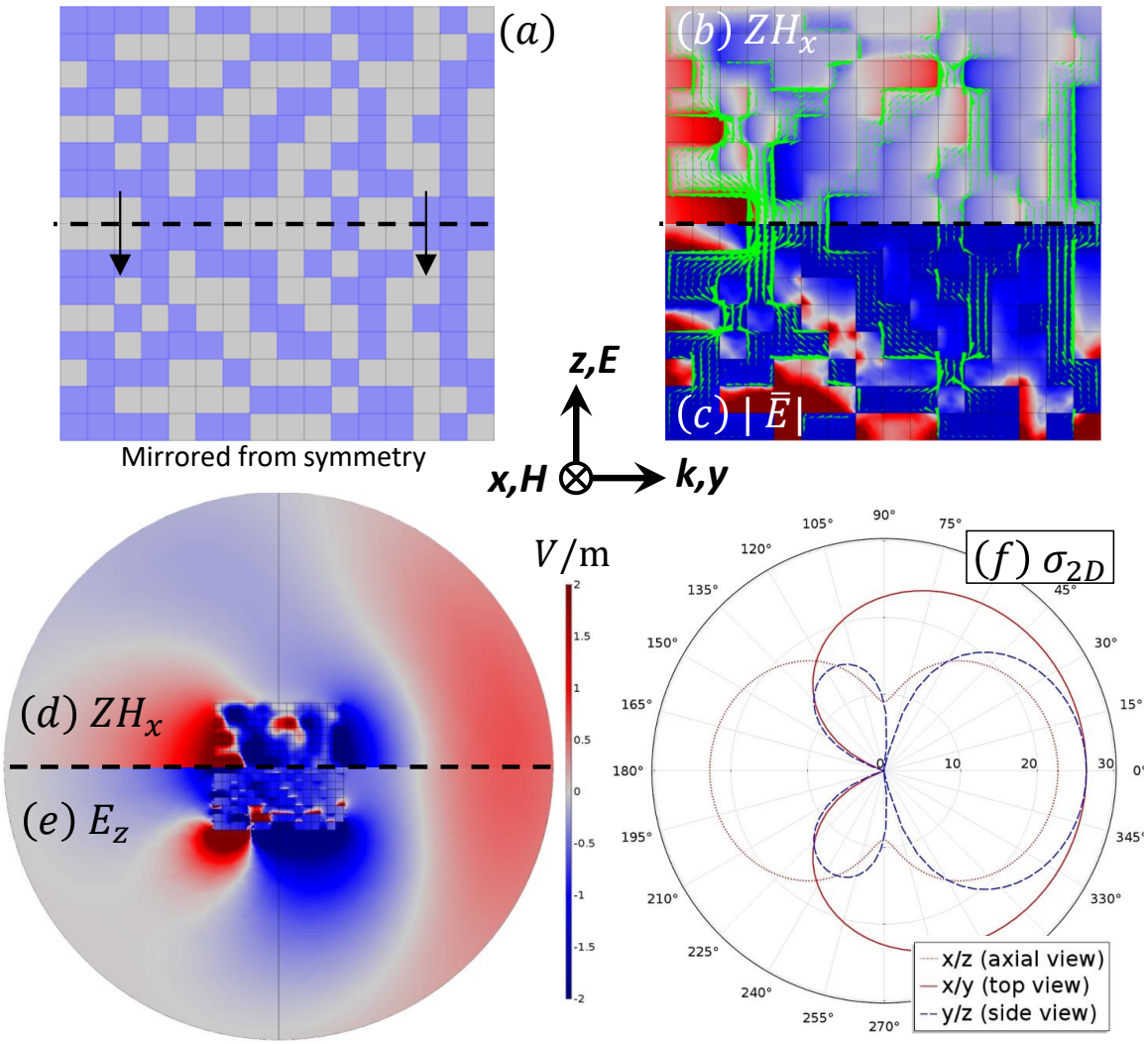

Figure 16. View of the (a) 16×16 grid layout, (b-e) near-fields and (f) projected 2D differential cross section (DCS, σ_{2D}) of the optimized verification model in each cardinal plane. The lower half has been mirrored, as per symmetry, indicated by the dashed line, and fields are split between upper and lower half for brevity. Many modes are seen induced in the H_z field (b) as a result of the current rotation and in (c) where there are abrupt terminations of conduits for current to flow. The result in the near field (d,e) is a strong preferrential forward scattering, and in the far-field (f) a clear Huygens-like behavior is seen.

plane scattered magnetic field $Z * H_x$ (in V/m) and the normalized scattered electric field E_{norm} are shown in Figure 16(b,c), respectively, with the current density in green arrows. Both field profiles show multiple spatially localized concentrations separated by the various geometries. For the H-field, opposing colors indicate opposing sign of the phase, and strong variations are occurring over the span of a single 177nm voxel.

Backing off to view the entire domain in Figure 16 for the (d) $Z * H_x$ and (e) the E_z fields, a clear forward scattering is observed, beginning at about 3μm from the center of the grid.

Projecting the fields out to $r = 1$m provide the polar plots shown in Figure 16(f), offering a view of the radiation pattern from the side (blue, dashed), top (red, solid) and down along the axis (red, dotted), in terms of the DCS in Equation (14), over a 2π observation angle in each plane. This is clearly Huygens-like in the scattering behavior, with both top and side views showing the significant reduction in backscatter ($\sigma_b = -28.73$dBv, not visible on the given $0 - 30$dBv scale) and a forward scattering intensity $10^{\sigma_f/10} = 440\times$ larger than the V-antenna!

The backward-angled side lobes are undesirable, but remember that the fitness function did not address this behavior explicitly—it only targeted the back lobe at 180° for ease of demonstration. A function added to Equation (11) that would be effective in reducing all side lobes is

$$F = w_f F_f + w_\Delta F_\Delta + w_\Omega \left(1 - \frac{\int_{\Omega_0} \sigma \; d\Omega}{\int \sigma \; d\Omega}\right), \tag{15}$$

where the integration is carried out over some large solid angle Ω_0 encompassing at least the rearward hemispherical shell, and normalized by the integration over the entire shell to match the previous scaling of the forward and difference terms F_f and F_Δ as defined in Equation (11). Even though adding another objective may require larger N_{pop} for more diversity or increase time to convergence, for real-world developments this approach may make sense, depending on the application.

Unfortunately, the P_1 layout exhibits some design flaws that must be addressed. Looking at Figure 16(a), several regions where the voxels form a closed loop, there are many voxels which are touching corner-to-corner (CtoC). Both of these are infeasible design aspects, but a discussion of the latter will be placed on hold until the next

section. For this verification task, it was more imperative to examine the closed loop issue. It was expected that the complexity of these designs would demand a lithographic fabrication technique to be realized, and when using a photoresist there are some constraints on the mask that must be considered. An important one is that the mask requires physical support; there cannot be a point where the apertures of the mask—in this case representing where the metal would exist on the structure—completely enclose a point on the mask. It is the equivalent of cutting a circle in paper with scissors.

There are at least two ways the existence of closed loops can be addressed. One way to prevent formation of closed loops is through a check sub-routine embedded in the GA routine itself, where after the chromosome evolution the grid is scanned for these formations. However, for large grids such as the 128-voxel verification model it would involve a great deal of nested control loops and additional computational delay over the entire $N_{pop} \times N_{iter}$ solutions, and thus the sub-routine was deemed too outside the scope of this effort to warrant development.

The other way is to simply perform a short, manual post-GA operation on the model, switching voxels as necessary to open up the topology. An example of how this was implemented is shown in Figure 17(a), placing yellow dots on the upper half to indicate which voxels were flipped in the symmetric model from $0 \rightarrow 1$ or $1 \rightarrow 0$, with the modified model displaying in Figure 17(b). This implementation of forced mutation brings up a concern on the stability of the solution space in general—would a vast majority of solutions shift wildly if this technique were applied to the entire set? Before examining the results of thus singular model, consider the form of the solution space itself. The Pareto front defines the boundaries of a given subset of solutions, usually dictated by the GA parameters and the general topological framework the evolution has established over many iterations. In short, the low rate of mutations

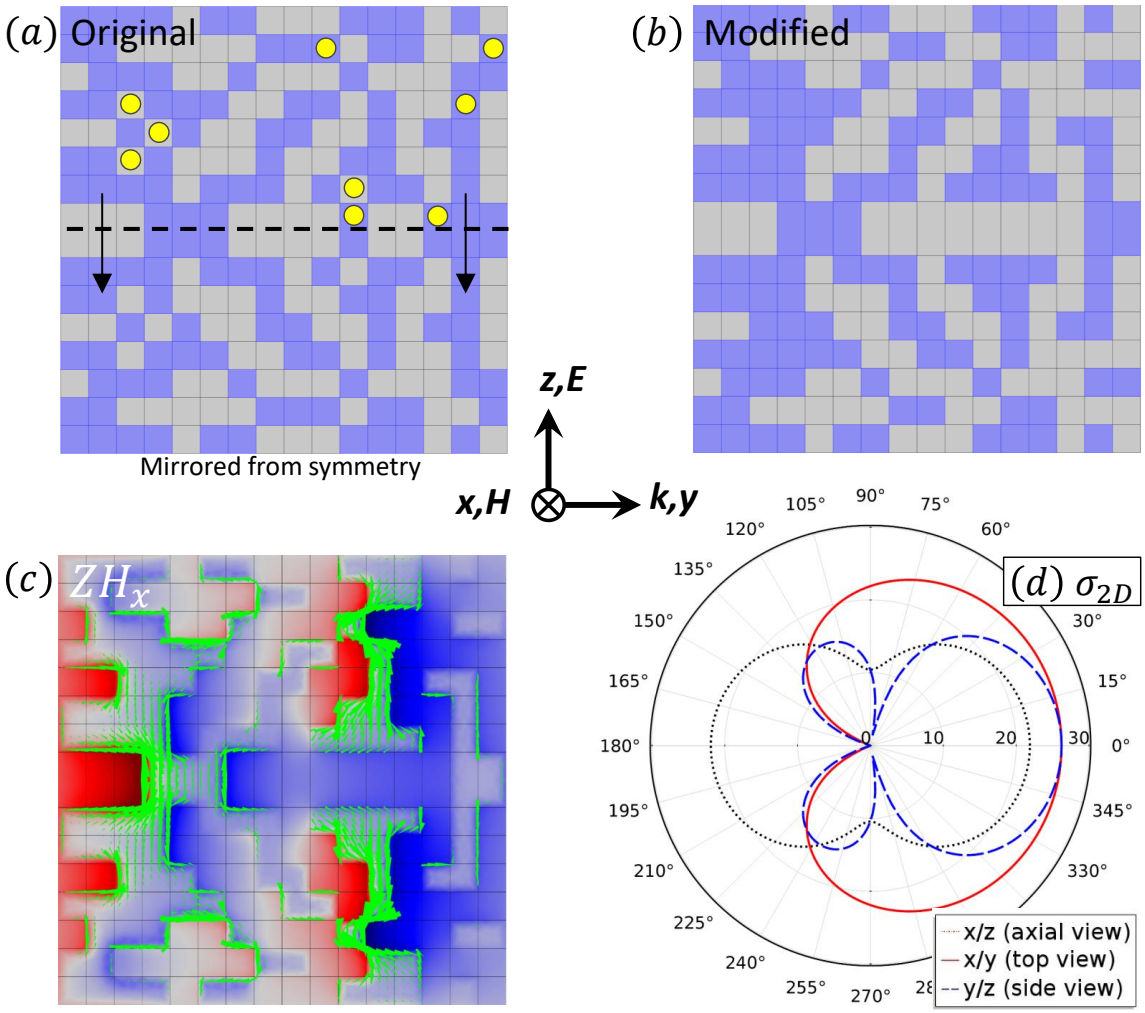

Figure 17. **(a) Modifications to the verification model layout in Fig. 16(a), placing yellow dots to show where bits were flipped to open the closed loops, with the modified model shown (b). (c) Near-fields (ZH_x) are stronger in some regions, yet, the scatterer still behaves Huygens-like—nearly identical to the original in Fig. 16(d)—according to the far-field plots (d).**

can ultimately go only so far in changing the topology after the parents become more and more homogeneous over many generations. Therefore, by changing a few bits of a far-evolved model it should not be expected to vastly alter the result. This idea does not hold true for all models, especially those with less resolution, as some will have critical geometries that are key to controlling the necessary current flow or field coupling for the intended function (think: phase or polarization control).

While keeping in mind this is only a single example, the results of the modified model are shown in Figure 17(c) for the $Z * H_x$ near-field component and in Figure 17(d) for the far-field. These show similarities in both the field and current profiles, so it is not surprising that this modified model produces a slightly better output of $P_3 = \{\sigma_\Delta = 61.00, \sigma_f = 26.24\}$dBv, nearly 5dBv down in backscatter from P_1 while only sacrificing 0.2dBv forward. Since this was not a full study, there cannot be any claim that this similarity of outcome should necessarily be expected for every modification, though a more robust assessment of topology sensitivity will be investigated in the 3D unit cell designs and fabricated devices.

Overall, the verification task affirmed two key results: first, that the GA routine can successfully and repeatedly target a globally optimal solution out of a massive design space using a commercial FEM solver—these were the defined criteria for "verifiability" from Chapter 3.3.1; second, this demonstration affirmed the theory that an OOP voxelated grid can *indeed* form a topology which will generate the electric and magnetic moments necessary for highly-efficient Huygens' sources. But this is not a metasurface, yet. The verification model was a simple structure of a sandwiched grid suspended in air; furthermore, there was no phase control demanded in this demonstration, which is the key functionality of a metasurface. The next section expands upon the simple verification model by applying the now-certified GA routine to a realizable 3D metasurface unit cell.

3.4 Three-Dimensional Unit Cell Design Methodology

3.4.1 Topology Considerations.

As described briefly in Chapter 2.4.2, the proposed OOP Huygens metasurface design concept consists of a cavity with a "blank slate" grid of voxels arrayed on any or all of the interior faces, and in general there is a great deal of freedom permitted in

this approach. A fundamental characteristic of the design is that there is no *a priori* assumption of what topology the optimal scatterer will take, other than the fact that curved geometries lying along the incident plane produce some type of magnetic response, and linear geometries aligned with the incident polarization produce some type of electric response. As a Huygens source, there is no requirement that these modal responses be generated from a singular, contiguous topology; different regions may contain separate electric or magnetic modes and still meet the requirement of spatially co-located, as they exist within a subwavelength dimension. Though, there are some constraints: it is imperative that both responses exist simultaneously (there cannot be solely an electric response), be of near equal strengths, and are in the proper phases such as to mimic the symmetric and anti-symmetric fields in Equations (7) and (8) and ensure cancellation of backward propagation.

Any tiling of voxel building block and grid array that forms a tessellation (no gaps or overlaps) could potentially be implemented in this design, and some examples are shown in Figures 18(a-h). The most straightforward layout would be a "regular" tiling, such as (a) the square lattice with square voxels that was used in the verification model, or (b) its rotated variant. However, as mentioned, these tilings induce potential CtoC contact at the corners of the voxels—shown in black—and creates significant challenges in stability of simulation and in accurate reproduction of these contacts in fabrication.

An infinitesimal contact point is a geometry which only exists in a simulation, and cannot be trusted as a connected topology; yet, COMSOL has no problem solving a model with such a layout. A demonstration of the instability this can cause is shown in Figures 19(a-c), presenting a 7×9 grid layout which has a vital CtoC connection (a, white circle). It can be seen that a large density of current flows through these regions, indicating the program recognizes this contact. This design exhibits a transmittance

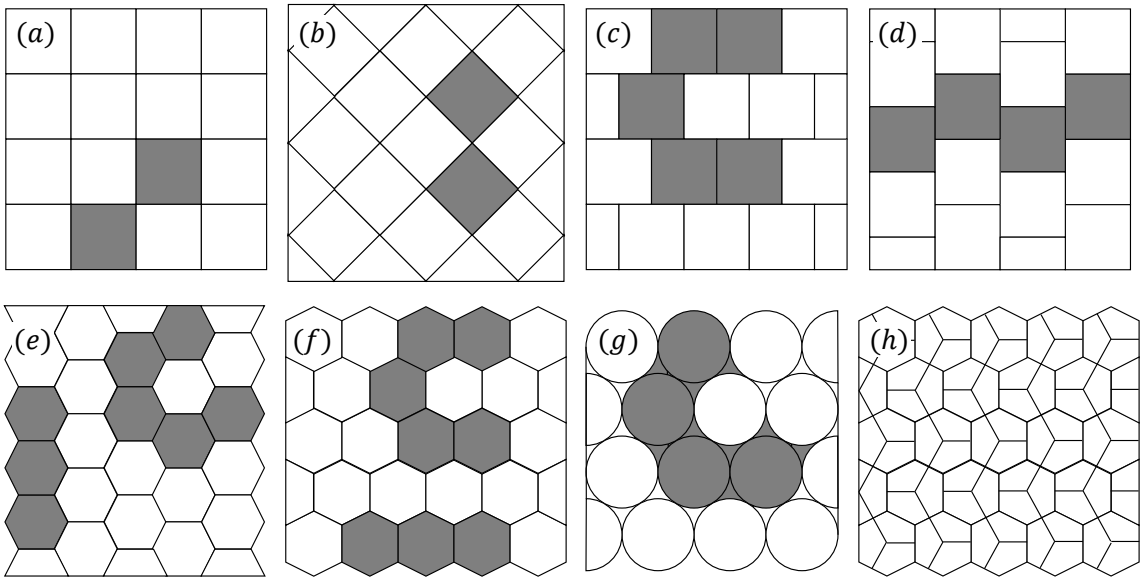

Figure 18. The proposed OOP Huygens metasurface design concept supports nearly any tiling of voxel building block and grid array that forms a tessellation (no gaps or overlaps), such as the examples shown. The square voxels in a square lattice (a,b) proved infeasible, while the "brick" (c,d) and "hex" (e,f) voxels are to become the staple design architectures used throughout this work. Other designs (g,h) are considered too challenging in either GA implementation and/or fabrication.

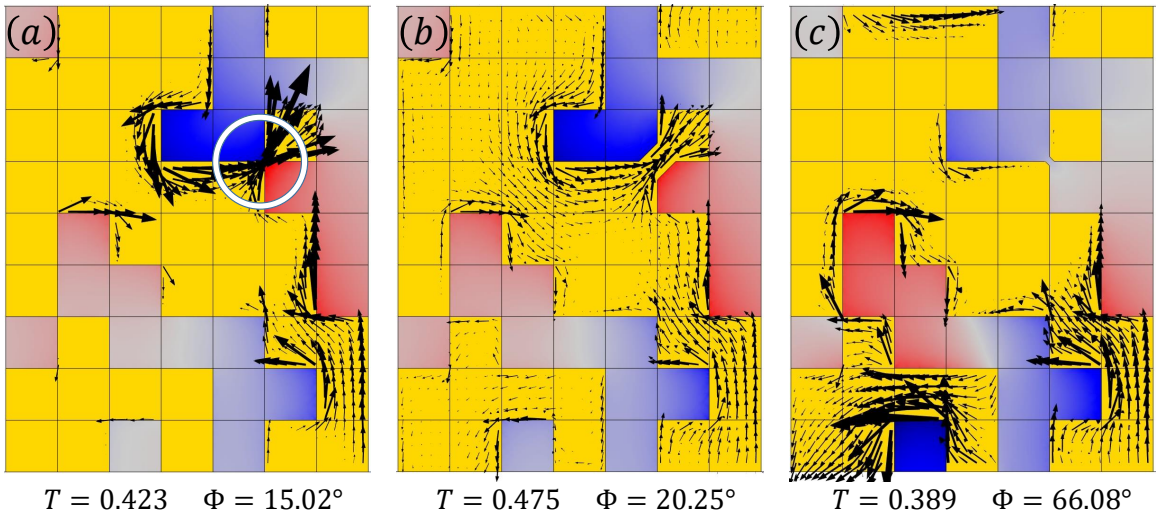

Figure 19. (a) A 7×9 square grid layout which has a vital CtoC connection (circle), resulting in high current density (arrows) across an infinitesimal width. (b) By enlarging the width of this connection, the design becomes more stable and improves in transmittance (T), while only suffering a manageable $5°$ phase shift (Φ). (c) However, even the slightest break in this connection completely alters the performance metrics, leading to a low confidence in realized structure.

of $T = 0.423$ with a phase shift in the transmitted total field of $\Phi = 15.02°$. If this CtoC connection is bridged by inserting some additional material, as in Figure 19(b), then there is a slight change in density, as the current has a wider path to flow, but this produces a manageable change in the results, giving $\{0.475, 20.25°\}$. A 5° shift in phase is generally not significant in metasurfaces, where the discretizations are often in steps of 22.5° or 45°, or more. Conversely, if this CtoC contact were severed, introducing a gap between voxels as in Figure 19(c), then current ceases to flow to this region, and subsequently the localized magnetic modes become weaker and less spatially distinguished in phase; meanwhile, other regions begin forming strong modes. The end results were $\{0.389, 66.08°\}$—a 50° phase shift! Since the location of these CtoC contacts are so unpredictable, there is no reason to believe there would be a uniform shift in all the different metasurface unit cells, and so this instability becomes impossible for an optical engineer to manage.

The remaining tilings in Figures 16(c-h) are examples of how to avoid this issue. A simple solution is to invoke a "brick" pattern: a shift in alternating rows of the square/square array, seen in (c,d) for horizontal and vertical shifts, respectively. This is analogous to a triangular lattice of square voxels. Unfortunately, these simple solutions form geometries with the coarsest possible curvatures (c), and cannot form pure linear geometries across the shifted bricks (d). By changing to another set of regular tilings, a triangular lattice with hexagonal voxels (honeycomb structure), seen in (e,f), a reasonable trade-off between curvature and linearity is achieved, though at the cost of containing slight ridges along all boundaries. These saliencies can introduce additional impedances, and the effect they have on current densities and, ultimately, amplitudes and phase delays will be examined in Chapter 5.6.

More exotic combinations are possible, such as two-shape "dual tilings," using a circle-packing example (g) with circular and 3-point star voxels in a hexagonal

lattice, or (h) a dissected tiling of a honeycomb made from uniform pentagonal voxels; though, both of these are problematic, as well. The dual-tiling can improve both curvature and linearity by filling these gaps with the small secondary shapes, but it also increases the resolution requirement and can introduce challenging geometries, such as with the finely-pointed tips shown in black in (g). Dissection, on the other hand, not only increases resolution, but adds little to no improvement in forming the necessary geometries. Due to the aforementioned limitations in tilings shown in (a,b,g,h), voxels were limited to the tilings in (c,d,e,f): square or hexagonal shapes in hexagonal lattices, referred herein as "brick" or "hex" layouts. Some square/square layouts are shown as used in initial testing of the GA routine, but were not considered for any design down-selection.

While evident that the higher the resolution of the grid, the easier it is to form the necessary discrete approximations to a curved topology, there exists some lower resolution that can no longer resolve the curvature and linearity sufficiently. The 4×4 brick grid in Figures 16(c,d) cannot form both a full c-shaped curve (c) and a linear structure (d) without some contact occurring; whereas the 5×5 hex grid in Figures 16(e,f) can separate these features, but raise questions on whether there is enough real estate to form the diversity of topologies needed for efficient and complete phase control. An exhaustive study could be performed to accomplish this, but a rough estimate could easily be found through initial testing. Before this estimate could be determined, a more thorough understanding of the metasurface architecture and associated fabrication constraints was needed.

3.4.2 A 3D Unit Cell Based on the Membrane Projection Lithography (MPL) Technique.

Up to now, the specifics of the metasurface architecture have not been described, nor linked to a realizable structure. A promising architecture that could support the OOP voxel grid was based on a fabrication phenomenology called "membrane projection technology" (MPL), developed by Sandia National Laboratories in 2010[70, 71] and filed under U.S. Patent No. 8,197,887 [72]. Originally developed using an SU-8 polymer, the MPL structure has recently been adapted for CMOS compatibility using a silicon (Si) basis[73]. Silicon is excellent for scattering at LWIR due to its low absorption loss; though, it suffers from high index—see the complex refractive index in Figure 20 for the real part, n (blue) and the imaginary part, k (red), using data from SNL measurements on the constitutive parameters.

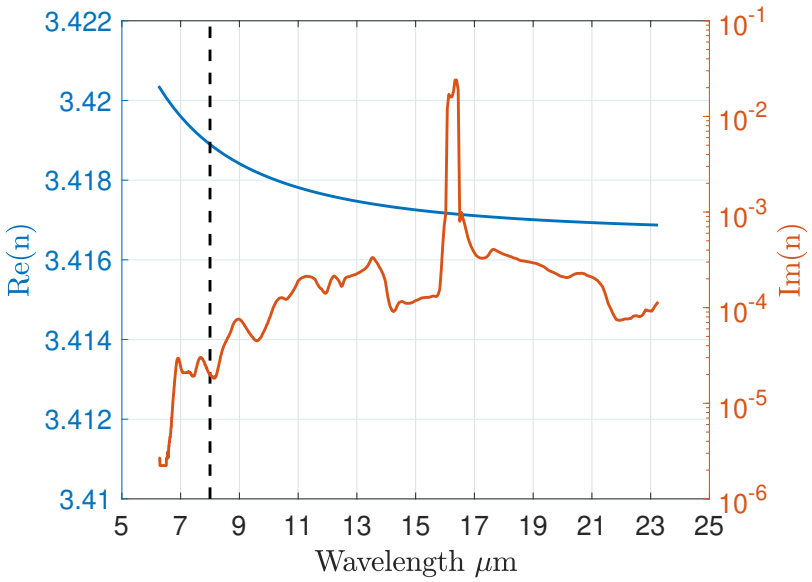

Figure 20. Real (blue) and imaginary (red) parts of the refractive index of silicon (Si), indicating only mild dispersion and very low loss at $\lambda_0 = 8\mu$m (dashed).

A cartoon of the MPL fabrication process and examples of MPL structures are shown in Figure 21(a,b), respectively. First, (1,2) a cavity array is formed from a

substrate material using holographic lithography, which is then (3) backfilled with a compatible organic or inorganic material and (4) planarized to prepare for (5) deposition of the membrane. After (6) e-beam patterning of the membrane, this permits the aperture to be used for (7) evacuation of the backfill. At this point, the cavity is empty, and (8-9) metal can be evaporated through the pattern on the membrane at any desired angle. The membrane (now coated in metal) is (10) lifted off and the structure is complete: an open cavity with OOP scatterers on any of the internal faces.

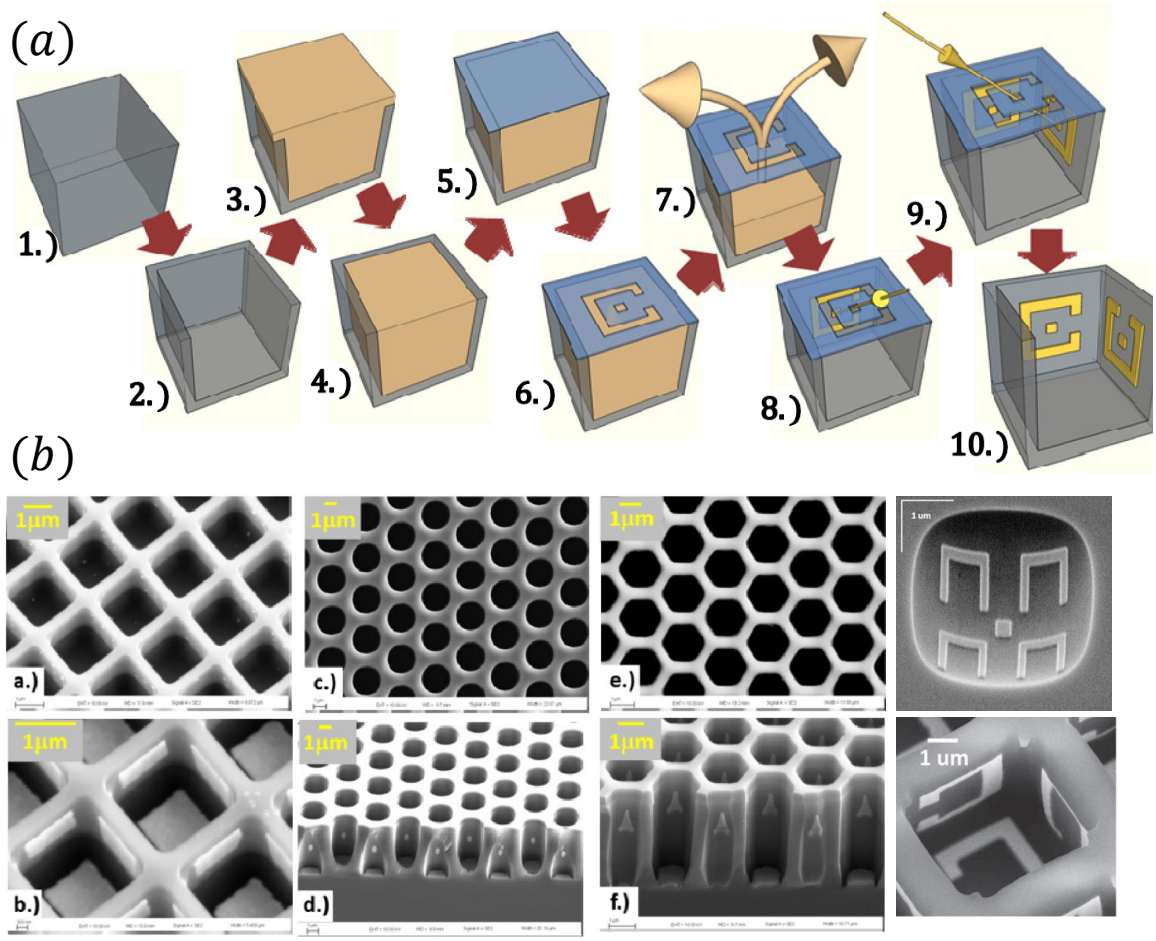

Figure 21. (a) The 10-step membrane projection lithography (MPL) process, showing how a 3D OOP scatterer can be deposited on any of the interior faces of a cavity. (b) Fabricated examples MPL structures show the great diversity in cavity shape (top row) and inclusion (bottom). Images taken from Refs. [71, 73, 74] with permission.

Evidently, the combination of cavity shape, inclusion shape and evaporation angle makes the MPL structure the most versatile 3D architecture known; to date, cavities of inverted quadrilateral pyramids, hemispheres[75], and cylindrical, hexagonal and rectangular prisms have been demonstrated in square and triangular lattices[73]. In conjunction with proper programming through COMSOL's API, the GA routine can exploit this versatility, modulating any of these aspects by adding genes that can represent a given structural combination. However, for simplicity, in this work the cavity is restricted to either a cubic or a rectangular prism, permitting perfectly flat vertical walls that isolate the voxel grids such that they are aligned exclusively along either the incidence plane or the planes orthogonal to it.

Since the membrane is orthogonal to these walls as well, a 45° angular projection through a pattern on the membrane can deposit an array of perfectly square voxels in the cavity—as shown in Figure 22(a)—by elongating the pattern accordingly, depending on the membrane thickness s. This concept is expanded to a full $M \times N$ grid to translate the verification model to a fully realizable OOP plasmonic Huygens metasurface, as shown in Figure 22(b) for a 9×9 square array of square voxels. Graphically shown in Figure 22(c), it can be seen that for a single voxel of width w the membrane must be biased to $w + s$ in the plane of the angular deposition, in the direction normal to the wall being deposited.

3.4.3 MPL Design Specifications.

Every effort was made to incorporate features into the design that could be realized with established MPL fabrication processes; this includes the many constraints that exist, as well. Sandia has developed a large set of unprocessed cubic and rectangular cavity arrays for various internal efforts—these are backfilled, but unmasked and unmetalized, ready for a rapid transition to the imprinting and metal deposi-

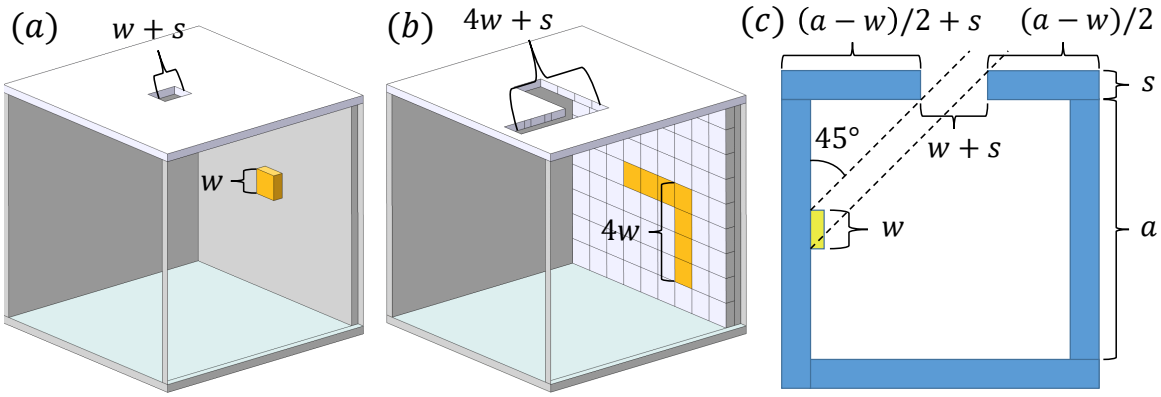

Figure 22. (a) A square voxel of length w can be deposited on a vertical wall by projecting the metal evaporation at $45°$ through a rectangular aperture of length $w+s$. (b) Expanding the single voxel concept to a $M \times N$ grid, translating the verification model in Fig. 10 to a fully realizable OOP plasmonic Huygens metasurface. (c) Diagram showing how the membrane occludes the deposition, and thus the extent of each pattern must be biased by the thickness s in order to project onto the vertical wall.

tion steps. They have permitted liberal use of some elementary MPL structures to develop laboratory-scale devices for this research effort. As such, there are several explicit physical and design characteristics which will remain fixed, based on SNL's specifications of these precursors. None of these adaptations affected the fundamental GA routine as described in Chapter 3.3.3, other than a new formulation of F (given in Chapter 3.4.4).

The cavities are arrayed in a square lattice, with $a = 2.3\mu$m periodicities in both dimensions, wall thickness $t_w = 300$nm, 2μm lateral cavity dimensions, and either 2μm in depth for cubic cavities, or 2.6μm for rectangular boxes. By keeping the incident freespace wavelength to $\lambda_0 = 8\mu$m, this periodicity remains just below the diffraction edge of $2.3\overline{3}\mu$m. Sandia's fabrication tolerance on a single voxel was determined by the limits of the imprinting method and the membrane thickness, and was set to 222nm for square voxels and 385nm for hex voxels.

Note in Figure 22(b) that the additional membrane biasing of length s would prevent the lowest row of voxels to be formed; moreso, the upper row would not be resolved for depositions through the same mask onto the opposing wall at $-45°$

for 2-face metasurface elements. Even further, due to tolerances in the holographic lithography there exists curvature in the cavity where the wall meets the floor, which can warp the projected metal deposition. Therefore, the upper and lower rows of the grid were eliminated entirely to mitigate fabrication risk. This limits the grid sizes in the final fabricated designs to 7×9 for the brick layout, containing $v = 68$ voxels (counting half-voxels) and 5×8 ($v = 49$) for the hex layout. However, for demonstration of the upper limits of transmission efficiency in the phase control of the metasurface design, higher resolution hex grids of 7×10 ($v = 78$) will be presented.

With the voxel grid now integrated into the MPL cavity, the COMSOL model required a significant overhaul from the verification model. One stark difference in the MPL model compared to the verification results is in the field quantities. In Chapters 3.3.2 and 3.2.2, it was described how only the scattered far field was sampled, with the background field being subtracted from the total field as given in Equation (9). However, the intent of this work was to demonstrate that a metasurface can strongly transmit with full phase control, and in order to do that, the total field must be sampled—not just the scattered field. The idea is that the MPL structure would perturb the background field just slightly enough that the total field emerges from this discontinuous interface also only slightly perturbed.

Now, the simulation is approached is a typical unit cell analysis of a periodic structure, using Floquet periodic boundary conditions (PBC's) and an S-parameter extraction of a multi-port network of the form

$$\begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \begin{bmatrix} E_i^{cp} \\ \tilde{E}_i^{cp} \\ E_i^{xp} \end{bmatrix} = \begin{bmatrix} r^{cp} \\ t^{cp} \\ t^{xp} \end{bmatrix} \begin{bmatrix} E_i^{cp} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} E_r^{cp} \\ E_t^{cp} \\ E_t^{xp} \end{bmatrix}, \tag{16}$$

where the Si-side co-polarized input field (E_i^{cp}) is the only non-zero component, nul-

lifying the cross-polarized input field (E_i^{xp}) and the air-side co-polarized input field ($\tilde{E}_i^{cp}$). The air-side cross-polarized input and response terms are left out of the matrix (which would have produced a 4-port network), because it was known the COMSOL port would not produce this field, and none of the outputs were of interest in this work. This leaves only the co-polarized complex reflection and transmission coefficients ($S_{11} = r^{cp}, S_{21} = t^{cp}$) and the cross-polarized complex transmission coefficient ($S_{31} = t^{cp}$). This last term can result from potential polarization rotation—otherwise known as chirality or optical activity—through interaction with the metasurface element, and is important to understanding the overall behavior of the design.

Similar to the manner in which the V-antenna model was set up in Figure 8(b), the excitation port is aligned with the $x/z-$plane and placed in the Si backbone $1.5\lambda_0$ away, with the field propagating in air along $+\hat{y}$ to a receiving port the same distance away. The complex refractive index of the lossy and dispersive gold (Au) and silicon (Si) material was taken from measurements performed by SNL. Mesh resolution in the air and substrate were set to $\lambda_0/6n$, while resolution of the voxel grid was set to $1/4^{th}$ the smallest voxel dimension. Meshing metallic domains near the PBC requires an extremely fine mesh, due to the strong field concentrations; therefore, the wall is shifted to the center of the unit cell in $\hat{x}$, reducing computational cost. Only either a single grid, or a back-to-back set of grids were investigated in this effort, and thus no shift was performed in $\hat{z}$ since it was known no grids would be placed on these walls.

3.4.4 Fitness Function for Efficient Phase Control.

The MPL model utilized all of the same subroutines as the verification model, so the GA process flow remained intact from the prior description, with only minor variations to the LiveLink scripts to incorporate the brick/hex tilings and the aforementioned changes to the BC's and physics environments. The most significant

change was in the fitness function. Since the goal was to design a metasurface element which could maintain a high transmittance (T) at some targeted phase retardation (Φ_t), a useful fitness would take the form:

$$F(\Phi, T) = w_\Phi \frac{\sigma^2}{|\Phi - \Phi_T|^2 + \sigma^2} + w_T \frac{T - T_{\text{low}}}{T_0 - T_{\text{low}}}, \tag{17}$$

where w is the weight applied to Φ or T, σ is a standard deviation defining a Lorentzian distribution centered at Φ_T and T_{low} is a minimum acceptable value for T, chosen to be 0.3. The terms Φ_0/T_0 are the phase/transmittance of the bare (unmetalized) Si cavity array (though Φ_0 is not used in Equation (17)). The value of Φ_0 and T_0 is a function of a and w_t, and its parametrization will be presented in Chapter 4. The Lorentzians of linewidths $10 \leq \sigma \leq 30$ were deemed appropriate functions to drive all solutions towards the line $\Phi = \Phi_t$, and the transmittance term is a simple normalization that force solutions outward from T_{low}. Generally, higher weight was placed on the phase $0.8 \leq w_\Phi \leq 0.9$, as it is much more imperative that the solutions drift toward the proper Φ_t than high T. This is because the topologies between a low T solution and a high T solution near Φ_t tend to share a much more similar framework than for high T solutions across some phase difference, and thus the GA has a much easier time tweaking a few voxels to attain a high T at the targeted phase than to change a large section of the topological framework to get to Φ_t.

A density plot of Equation (17) using the values $\Phi_t = 120°$ (red, dashed), $\Phi_0 = -62°$ (black, dashed) is shown in Figure 23(a), along with contours of equal F that vary from convex to concave as $F \to 1$. Additional constraints were applied to the fitness calculation, setting $F = 0$ for any solutions $T_{\text{low}} \leq T \leq T_0$ and $|\Phi - \Phi_T| < 0.9|\Phi_0 - \Phi_T|$, to mitigate the possibility of solutions with high T/low Φ and low T/high Φ from dominating.

However, this fitness was deemed too slow to reach convergence. This may be ex-

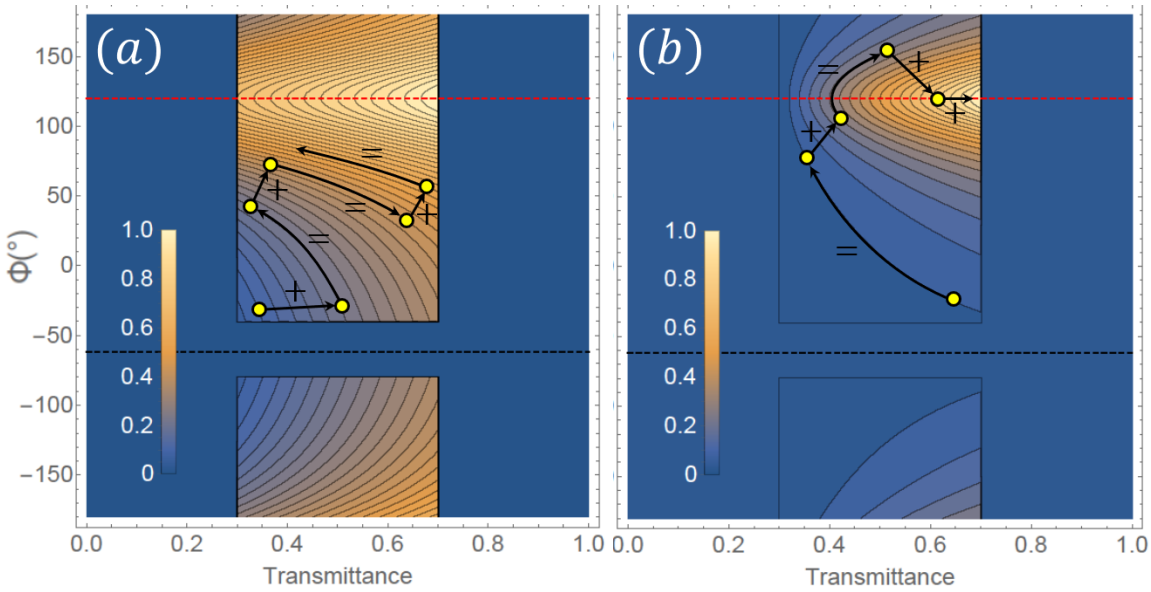

Figure 23. Cost functions (F) as defined by (a) Eq. 17 and (b) Eq. 18, with contours of equal F and additional constraints as described in the main text. Sample target phase is $\Phi_t = 120°$ (dashed, red) and reference phase is $\Phi_0 = -62°$ (dashed, black). In (a) the contours transition from convex to concave, which can engender slow convergence—an example is given, following the arrows. Thus, an additional term was added to form (b), allowing for complete concavity and better convergence.

plained by the convex contours; as can be seen in Figure 23(a), only within about 45° do the contours start to transition to concave, which engenders two mechanisms for slow convergence. The first is that the convexity can push solutions nearly laterally—that is, along a constant Φ—as seen between points $1 \to 2$. The second is that near the point of inflection the solutions can stagnate, moving along contours with little change in Φ or very shallow gradients, as seen between points $4 \to 5$ and $5 \to 6$. What is desired is a fitness which continually forces solutions towards the right side of the Φ_t red line, which can be accomplished by factoring a gradient in T against the Φ term, and a gradient in Φ against the T term, giving the fitness function

$$F(\Phi, T) = \left(w_\Phi \frac{\sigma^2}{|\Phi - \Phi_T|^2 + \sigma^2} + w_T \frac{|\Phi - \Phi_0|}{|\Phi_T - \Phi_0|} \right) \frac{T - T_{\text{low}}}{T_0 - T_{\text{low}}}, \tag{18}$$

where the only new term is Φ_0, used to drive solutions away from this line. Displayed in Figure 23(b), Equation (18) is convex everywhere, with varying convexity everywhere thanks to the two new biasing gradients. Following an example path, the solutions are pushed quite rapidly until they reach a region near Φ_t, and then follow a strong concavity towards $F = 1$. This provided faster results on the order of about 25%, and allowed better ejection from local minima.

3.5 Summary

In this chapter, the methodology behind the development of a functioning GA routine was laid out. After considering the FEM-based computational electromagnetics platform COMSOL, it was validated that the program could accurately design a large-scale metasurface device. Then the validated software was used to verify the build of a two-objective GA routine which could repeatedly converge to an global solution, optimizing the topology of a set of sandwiched, symmetrical OOP grids in air to mimic the scattering profile of a Huygens source. Finally, a realizable 3D structure was introduced, based on Sandia's MPL fabrication phenomenology, and it was shown how the design and fitness function changed accordingly. In the next chapter, the GA routine will process a myriad of MPL designs in an attempt to identify and analyze the best topologies for highly-efficient and complete phase control in transmission.

IV. Simulations of a Optimized Huygens Metasurface Elements

4.1 Introduction

With the both the COMSOL-linked GA routine and the binary voxelated OOP MPL structure firmly established, the question still remains whether the two can combine to develop a series of efficient elements with precise phase retardations that can serve as the building blocks of metasurface applications. In this chapter, various tilings and configurations are examined to provide a basis to answer this question, as it will be shown that indeed the GA optimization can repeatedly generate high-T elements that can cover the entire 2π phase range. Herein, simulation results are presented for the baseline MPL structure and the hex and brick elements which will be used to construct full-scale flat optical devices.

4.2 Baselines of Unmetalized MPL Cavities

Before proceeding with the GA routine, first the reference phase retardation Φ_0 and baseline transmittance T_0 must be sampled from a bare MPL model that contains no metal. This will be used to determine the phase shift $\Delta\Phi = \Phi_t - \Phi_0$ and the normalized transmittance $T_n = T/T_0$, the latter of which defines how much insertion loss is incurred at the Si/air interface of periodic cavities upon exiting the structure. The GA routine assesses only the absolute transmittance T, but frequently T_n will be cited, as it is a truer comparison of metasurface efficiency against published works; theoretically, if the empty MPL structure could be made lossless, the metasurface would transmit at T_n.

Parametrized profiles of T_0 and Φ_0 are shown in Figures 24(a,b), respectively, for the square unmetalized cavity and Figures 24(c,d) for the rectangular cavity, as

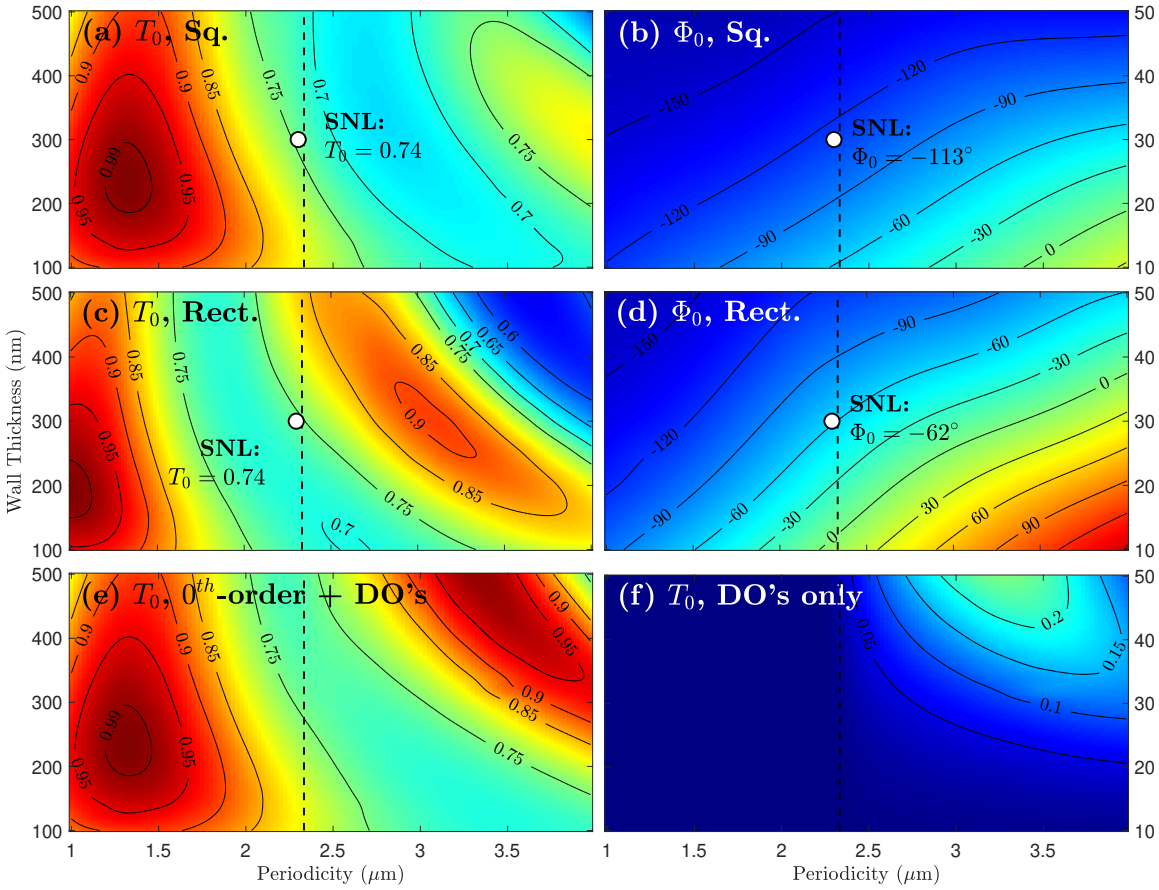

Figure 24. Parametrized profiles of (a,c,e,f) the reference transmittance (T_0) and (b,d) the reference phase (Φ_0) for (a,b) the square unmetalized cavity and (c,d) the rectangular cavity, as a function of periodicity (a) and wall thickness (t_w). The values for the $a = 2.3\mu$m MPL periodicity is indicated (white dot). In (a,c) only the 0-order transmittance was required, as the a falls below the diffraction edge (dashed), but the T_0 in (a) is also shown in (e) with all diffraction orders (DO's) and in (f) with only higher DO's.

a function of periodicity (a) and wall thickness (t_w). The transmittance in (a,c) are of the 0-order diffraction mode, but for completeness, the transmittances for 0-order and 16 higher-order modes (e) and the contribution from only these 16 higher-orders (f) are also given; as expected, there is no contribution below $2.3\overline{3}\mu$m from anything other than the 0-order transmittance–to include the cross-polarized (S_{31}) contribution. At the design parameters fixed by the available SNL patterned Si wafers $\{a = 2.3\mu\text{m}, t_w = 300\text{nm}\}$, the values are given as $\{T_0 = 0.746, \Phi_0 = -113°\}$ for the 2μm cubic cavity and $\{0.746, -62°\}$ for the 2.6μm rectangular cavity (white dot).

Extrema were found at design parameters {1.325μm, 233nm} (cubic) and {1.00μm, 190nm} (rectangular), giving near-lossless transmission; however, these designs are much harder to implement as a metasurface, for two reasons. For one, these walls are very thin, and may not have enough mechanical support in the pre-backfill or post-evacuation steps of the MPL fabrication process. The second reason is that the available surface areas for patterning would be $< 1\mu$m, which at SNL's fabrication tolerances likely will not provide enough of a grid resolution to form a topology diverse enough to access the entire phase space. So while an option for a near-lossless cavity array is possible, it is not feasible for realization with current fabrication technologies, and the $\sim 25\%$ insertion loss from the {2.3μm, 300nm} design is an acceptable trade-off for enabling an experimental validation of the OOP metasurface. It is important to note this loss does not include any insertion loss from air into the Si substrate, which, assuming contribution only from reflection, is between $30 - 45\%$ at 8μm, depending on the polishing[76].

4.3 Validation of MPL-Based 3D Metasurface Design

4.3.1 Demonstration of 3D Huygens-like OOP Grid.

Perhaps the best demonstration of efficacy of the binary GA optimization is to show a design which boasts a full 180° phase shift; for a cubic cavity, this seeks a target phase of $\Phi_t = 65°$ (rounding the reference phase down slightly to $\Phi_0 = -115°$). By showing the GA can identify a scatter to retard the phase by 180°, it is implied that it can hit *any* phase value in the range $-180°$ to $+180°$.

The basis for this example layout is a 7×10 hex grid in a hex tiling, and the 2D solution space from the GA routine with an initialization of $N_{pop} = 80$ and $N_{iter} = 39$ is shown in Figure 25. The dashed lines show T_0 (red), Φ_0 (blue) and Φ_t (black), the black-bordered white dots represent the solutions of the final generation, grey dots

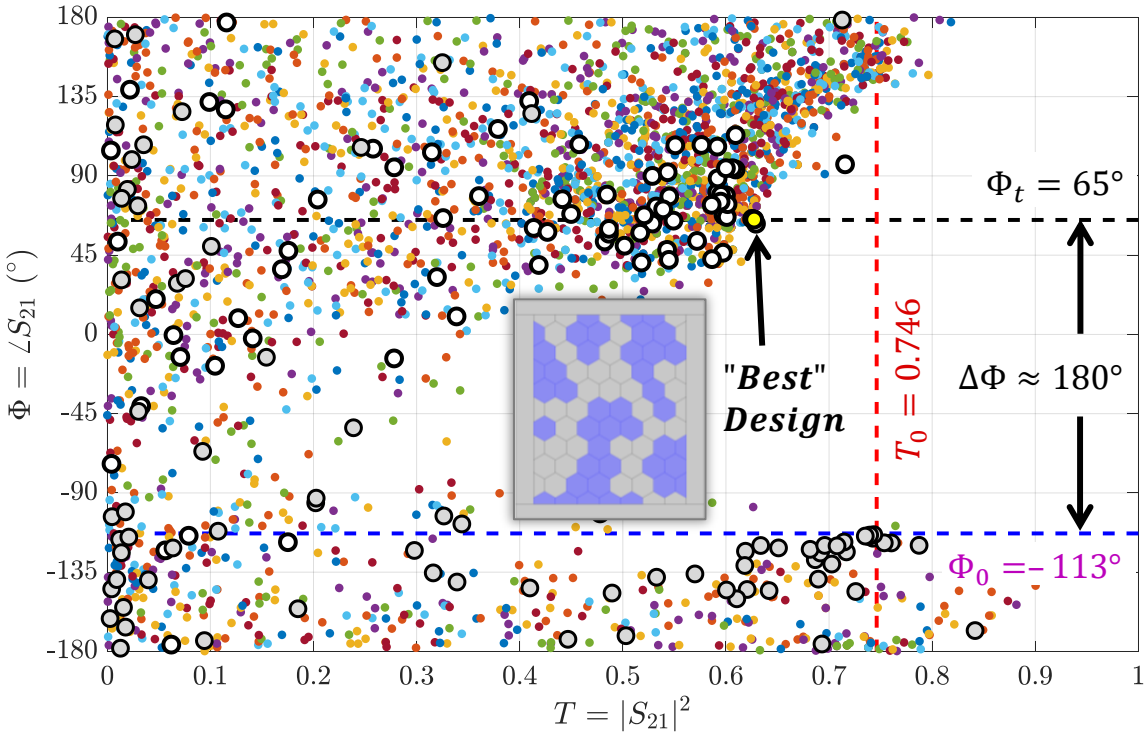

Figure 25. Solution space from a GA optimzation routine as a function of transmittance ($T = |S_{21}|^2$) and phase ($\Phi = \angle S_{21}$) for an example 7×10 hex layout, with initialization metrics $N_{pop} = 80$ and $N_{iter} = 39$. The target phase ($\Phi_t = 65°$) (black) is $\sim 180°$ phase shift from Φ_0 (blue). The final generation is shown (white dots) along with solutions from all previous generations (colored dots), and the "best" solution was found at $T = 0.63, \Phi = 65.2°$ (yellow dot), with the associated binary grid layout inset.

represent the initial population, and each other colored set of smaller dots represent a previous generation.

As similar to the solution space of the verification model Figure 14, there is a high density of solutions around the target phase value, indicating the GA routine has successfully evolved the grid into the "best" solution—however, which is "best" can be subjective in a 2D solution space. Also as similar to the results from the verification model, a loose Pareto front has formed, defining where the boundary of "best" solutions lies, based on the desired tolerance of either T or Φ. For this example, a low tolerance of the phase was chosen to show how well the GA could accurately converge to a given Φ_t, selecting a design (inset) as close as possible to

$\Phi_t = 65°$, indicated by the large yellow dot in Figure 25 at $T = 0.63, \Phi = 65.2°$. Normalized to the unmetalized structure this is $T_n = 0.84$, and with a reflectance of $R = |S_{11}|^2 = 0.02$, the metalized grid is behaving as a strong Huygens-like scatterer the total field.

Early generations (grey dots) tend to swarm around the T_0 and Φ_0 baselines, as sparsely populated grids will tend to poorly exhibit the resonant coupling needed for phase retardation. However, as the voxels begin to evolve into capacitive and inductive geometries, coupling improves and the continuum of T, Φ values becomes more accessible. This can be seen by tracking the evolution of the individual with the highest F over several iterations, shown in Figure 26. The grid layout (gold) of the randomized initial ($j = 1$) population is the least ordered, and provides very few geometries for interaction with the incident light. However, very quickly the first strong modal responses can be seen at $j = 2$, already producing significant phase shifts at high-T and setting the framework for some of the fundamental topological forms used in the proceeding iterations: by $j = 20$, these three or four fundamental forms repeat themselves in every generation. Though not apparent in this particular example design, sometimes this early framework may lead to Pareto fronts that would never reach thresholds—i.e., locally optimized—which is where mutations assist heavily in evolving from one framework to another.

For this example solution, the convergence to $j = 39$ required solving 3120 individuals, and the computation time for each full-wave solution was 2-3 minutes on a 28-core/256GB workstation; therefore, a single GA optimization routine can take on the order of 100-150 hours. Though this seems daunting, there are several aspects to consider. First, as can be seen in Figure 25, many other high-T phase values can be attained from a single routine during the design evolution towards Φ_t, precluding the need to run N routines for N discretizations. Second, looking at the convergence plot

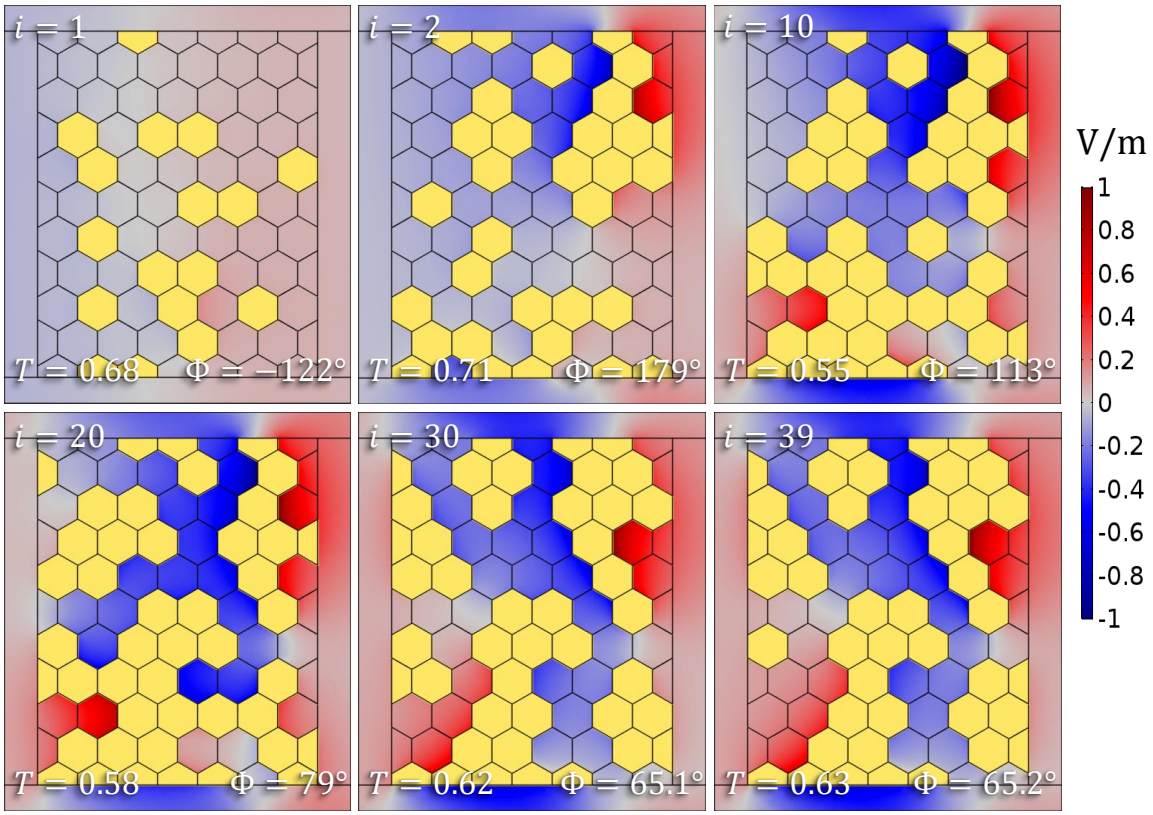

Figure 26. Evolution of the individual with the highest F in a generation over a range of iterations. The grid layout (gold) of the randomized initial ($j = 1$) population is the least ordered, and provides very few geometries for interaction with the incident light. However, strong modal responses can be seen as early as $j = 2$, and this layout begins setting the framework for some of the fundamental topological forms used in the proceeding iterations ($j = 10, 20, 30, 39$).

in Figure 27, giving the T, Φ values from the maximum F for each iteration, it can be seen that a solution which meets the $T \geq 0.6$ threshold (red, dashed) is attained a full nine iterations (24-36 hrs) prior at $j = 30$, giving $T = 0.618, \Phi = 65.1°$. This example continued on ti $j = 39$ only for demonstration of the GA routine to attain a global maximum F, which lead to a marginal increase in T; in real implementation, a threshold would likely have been triggered. Third, the parameter space is vast: in this case, the hex layout consists of 82 unique voxels, and each voxel is a parameter with a possible value of 1 or 0. To sweep the entire space of possible arrangements would, by the Rule of Product, require an unfathomable $2^n \approx 5 \times 10^{24}$ solutions.

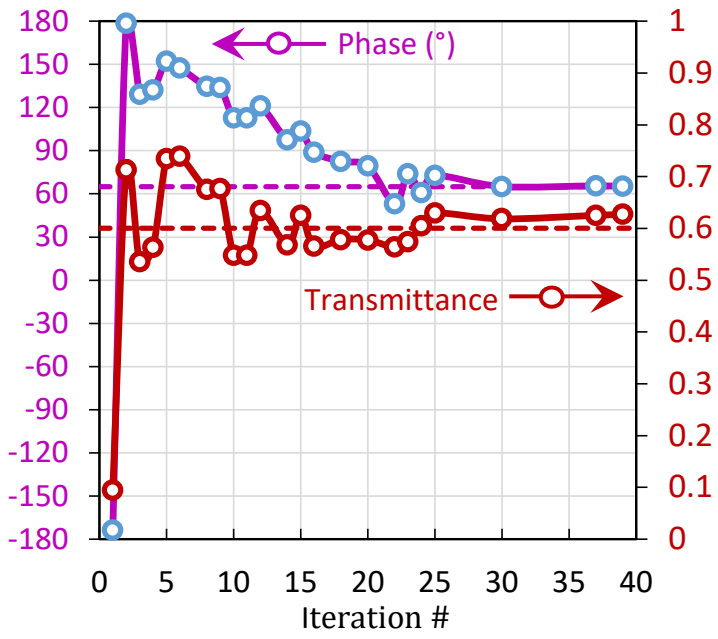

Figure 27. Convergence plot from the 2D solution space in Fig. 25, giving the T, Φ values from the maximum F for each iteration. It can be seen that a solution which meets the $T \geq 0.6$ threshold (red, dashed) is attained a full nine iterations (24-36 hrs) prior at $j = 30$, giving $T = 0.618, \Phi = 65.1°$. This example continued on to $j = 39$ only for demonstration of the GA routine to attain a global maximum F, which lead to a marginal increase in T.

Whenever there is a strong $\Delta\Phi$ there is an inextricable presence of strong absorption, as shown in Figure 28, where the absorptance is taken from $A = 1 - R - T$, which is a reasonable approximation since both the cross-polarized terms and the higher-order diffraction terms are ~ 0. The optimal design shown in Figure 25 is highlighted (yellow dot), and due to the extremely low reflectance (R), is near the $R = 0$ boundary (dashed, red). The link between phase accumulation and absorption can be explained through examination of Figures 29(a,b) of the (a) H_x and the (b) $|E|$ near field components around the metasurface interface, along with current densities (yellow arrows). Wherever there is strong rotation in the induced surface currents, there exists strong magnetic modes which can exhibit extreme spatial variation in phase; likewise at many narrow gaps in the scatterer geometry there are strong electric modes. In these regions, Ohmic loss will be prevalent in the metallic voxels due

to a high current density generated by the electric and magnetic modes, and thus, a large absorptance.

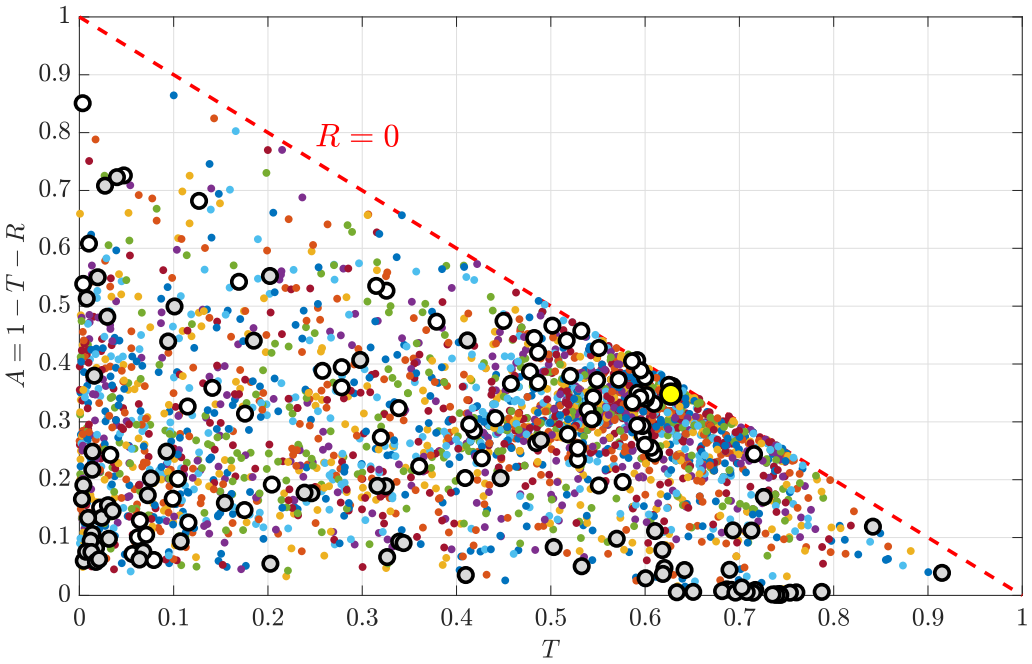

Figure 28. Absorptance $(A = 1 - R - T)$ as a function of transmittance (T), with the boundary $R = 0$ marked (dashed, red). The optimal design (yellow dot) is extremely close to this line, due to the Huygens-like nature of the scatterer.

By zooming out from Figure 29, the entire near-field in the model can be examined in Figures 30(a,b) for (a) H_x and (b) E_z. By applying the canonical equation for impedance (Z) for a propagating plane wave in a non-magnetic, homogeneous medium, the H-field can be scaled by Z in terms of V/m. Taking the ratio of the transverse field components,

$$Z = \frac{n}{\epsilon_{rel}} \frac{\overline{E}}{\overline{H} \times \hat{k}} = \frac{1}{\sqrt{\epsilon_{rel}}} \frac{E_z \hat{z}}{H_x \hat{x} \times \hat{y}} = \frac{1}{\sqrt{\epsilon_{rel}}} \frac{E_z}{H_x}, \tag{19}$$

where n is the index and ϵ_{rel} is the permittivity, the scaling is given by $H_x = E_z / Z\sqrt{\epsilon_{rel}}$. In this form, it can be seen the maximum strengths of the multiple modes of the two

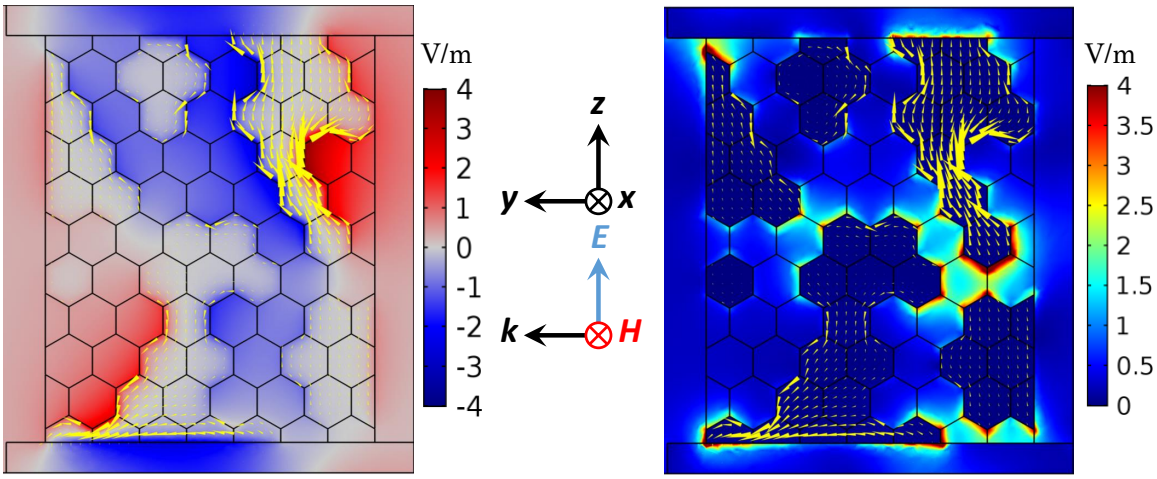

Figure 29. The (a) H_x and the (b) $|E|$ near-field components around the metasurface interface of the chosen "best" design, along with current densities (yellow arrows). Wherever there is strong rotation in the induced surface currents, there exists strong magnetic modes which can exhibit extreme spatial variation in phase; likewise at many narrow gaps in the scatterer geometry there are strong electric modes. Each contributes to the large absorption seen in Fig. 28.

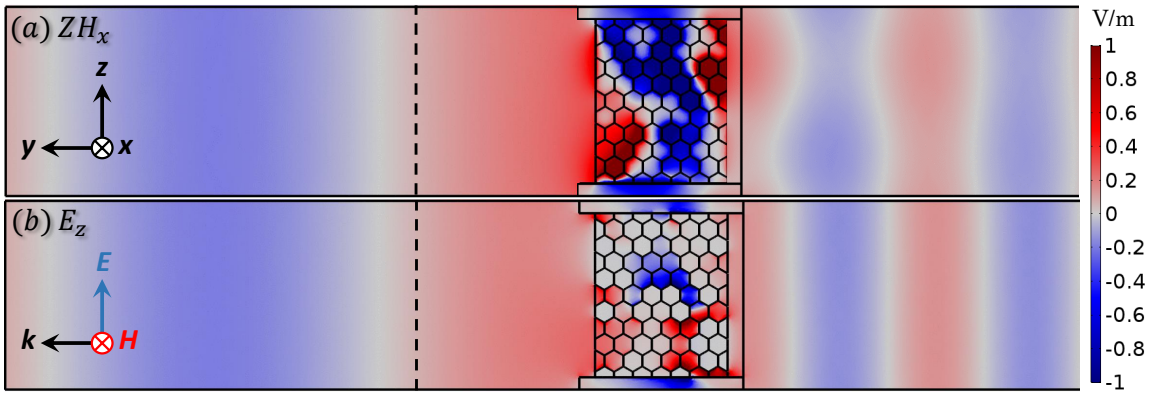

Figure 30. Near-fields (a) ZH_x and (b) E_z from the optimized design, showing how the maximum strengths of the multiple modes of the two components are of the same order. This, along with the close proximity of the orthogonal modes, the Huygens criteria are achieved, and the resulting transmitted total field (left of the unit cell) is also of near-equal strength and phase.

components are of the same order. With the close proximity of the electric and magnetic modes, and their near-equal strengths, the Huygens criteria are achieved, and the resulting transmitted total field (left of the unit cell) is also of near-equal strength and phase. An intermediate near-field from an antenna is considered to

occur, by IEEE convention, at distance

$$r_{FF} \geq 2L^2/\lambda \tag{20}$$

which, for the largest scatterer dimension in the unit cell $L = 2\sqrt{2}$ this gives $r_F F >= 2\mu$m. At this distance, a plane wave has clearly begun to form, marked as a dotted line in Figures 30(a,b), which should allow for proper phase construction very near to the metasurface.

As mentioned in Chapter 3.3.5 for the verification model, closed loops can occur occasionally, and no additional algorithms were written to manage these formations. With smaller grid resolutions, these appear less frequently; and when they do, they tend to take the form of only one or two voids in the center of a large body, as shown in Figure 31(a). Because surface current density (black, arrow) predominately is high only on exterior regions of larger bodies, these voids are generally isolated from the localized fields which form from these currents. As can be clearly seen, the fields interior to these voids are weak; therefore, their presence has little impact on the overall behavior, and so there is little impact whether the space is an air void or filled by metal. Of course, the obvious choice is to fill them in to eliminate the closed loop. Metrics for the demonstration hex model in Figure 31(a) at a target phase of $\Phi_t = 65°$ are: $T = 0.671$, $\Phi = 67.0°$ and $R = 0.01$ By filling in these voids as in Figure 31(b) the transmittance improves slightly to $T = 0.691$, though at the cost of increased phase separation of $\Phi = 71.4°$. Oftentimes when a design is only a few off the optimal target this can be corrected manually, as was done in Figure 17; applying the same technique to this design yields the final layout in Figure 31(c) at $T = 0.699$, $\Phi = 66.0°$ and $R = 0.01$, which is within an acceptable $1°$ of Φ_t.

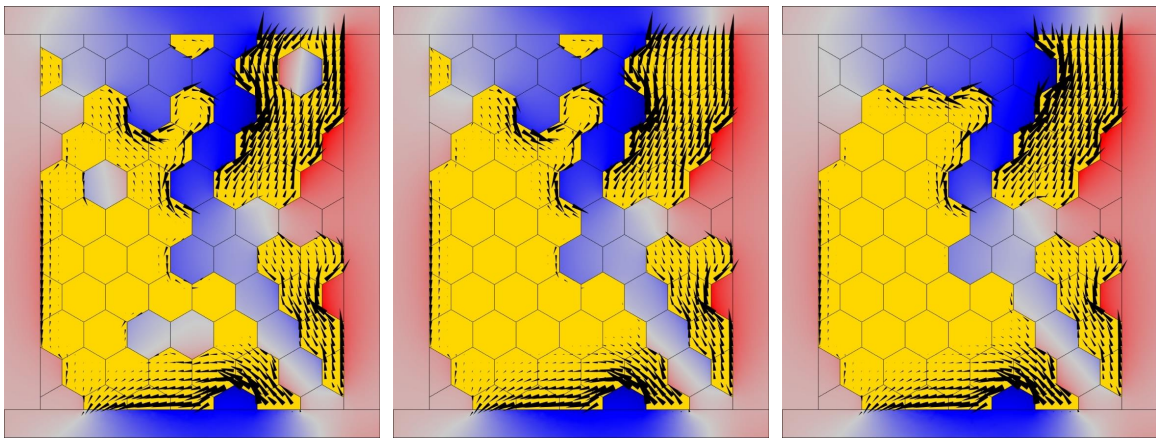

Figure 31. (a) Voids can appear during the GA evolution, and can persist the entire routine. (b) These voids are shielded by the surrounding metal, so by filling in these voids, the only detriment is a 6% shift in phase; however, this shift can be recovered (generally) by additional adjustment of a few voxels (c).

4.3.2 Design Repeatability.

As will be explained in Chapter 5.2.1, it was eventually discovered that the 5×8 hex resolution is the maximum allowable by MPL fabrication processes; additionally, SNL only provided rectangular cavities, so there will be no further use in examining the 7×10 hex structure beyond the initial demonstration. Nonetheless, it is not enough to show a single design performing a single phase shift at high transmittance; the metasurface architecture must demonstrate enough modularity in the grid layouts such that at least a complete 2π phase space can be traversed at nearly the same high intensity. Since a full 180° phase shift was demonstrated in the example 7×10 design, and the absorption should be strongest for the largest $\Delta\Phi$, it is not an unreasonable assumption that there should exist many grid layouts for which lesser phase shifts could be obtained at a similar or better efficiency. However, a performance characteristic of such great importance should not be left to an assumption; the OOP binary grid metasurface architecture is such a nascent approach that it demands a thorough

investigation on the flexibility of the MPL design phase control and repeatability of the GA-based optimization.

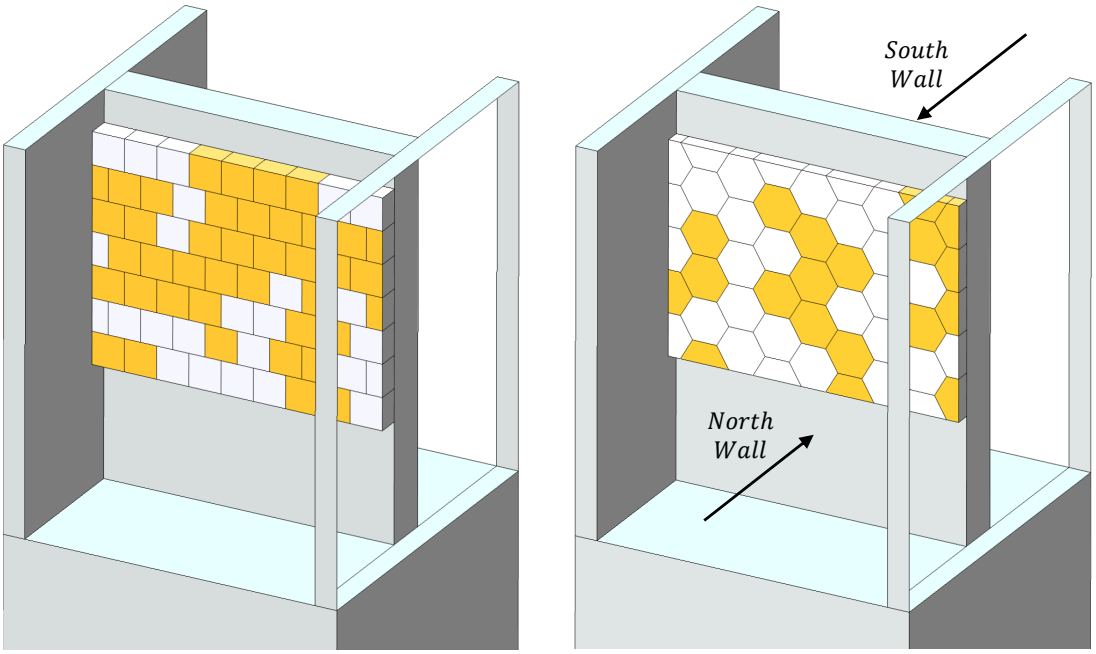

Figure 32. (left) A 7×9 brick and (right) 5×8 hex unit cell at the known resolution limit for fabrication at Sandia National Laboratories (detailed in Chapter 5.2.1). These architectures are used to demonstrate flexibility of the MPL design phase control and repeatability of the GA-based optimization.

Therefore, for complete robustness, it is most imperative to show that the hex and brick grids can not only (1) span the phase space, but (2) do so at the fabrication-limited resolution of 5×8 (for hex) and 7×9 (for brick), and (3) at the rectangular cavity depth. As opposed to displaying the capability to attain each of these points individually, in a rather redundant fashion, it can be done in a single, sweeping design modification using corresponding unit cell models in Figure 32. It is extremely important to note that some of these elements were chosen as "best" in part because of their ability to support the fabricated membrane, requiring at least two points of contact for long air voids, preventing any high-torque lever arms from collapsing during the MPL processes.

Results of the brick layout/rectangular cavity for a full 8-element discretiza-

tion of Φ_t are arrayed in Figures 33(a-f), showing phase shifts of $\{\Phi_t, \Delta\Phi\}$ = (a) $\{-150°, -90°\}$, (b) $\{-135°, -45°\}$, (c) $\{-60°, 0°\}$ and $\{-15°, 45°\}$, (d)$\{30°, 90°\}$, $\{75°, 135°\}$ and $\{120°, 180°\}$, and (e) $\{165°, -135°\}$ Element numbers are linked to their grid layout below (blue=metal), with elements $\{2, 3\}$ and $\{5, 6, 7\}$ taken from the same GA run, as shown in (b) and (d). Yellow dots indicate the performance metrics of the elements chosen for experimental validation, presented in a linear plot in (f).

All of the GA runs have very similar characteristics: the same clustering of initial populations (grey, dot) around Φ_0 and T_0 and final populations (white, dot) around Φ_t; large groupings from which a near-continuum of Φ values can be extracted (improving computation efficiency), and formation of a loose Pareto front—keep in mind some fronts are less well-defined than others due to hitting a termination threshold before it could solidify. Generations will evolve towards $+180°$ or $-180°$, depending on which direction is the smallest $\Delta\Phi$; as can be seen in (e), this sometimes means passing over the negative (positive) phase boundary to reach a positive (negative) phase value.

Amongst all elements, the standard deviation from the $\Delta\Phi = 45°$ interval was $1.12°$, with an average deviation from Φ_t of $0.50°$. The mean transmittance (red, dashed) in Figure 33(f) is $\langle T \rangle = 0.651$, or $\langle T_n \rangle = 0.873$, and, as assumed earlier, each GA routine was able to attain transmittances greater than the $\Delta\Phi = 180°$ element (#7), with the exception of the $\Delta\Phi = 90°$ element (#5). This may be because they came from the same GA run (that was targeted for maximum efficiency at $\Phi_t = 120°$, not $30°$). Element 5 could likely be improved by incorporating a GA run dedicated to maximizing T at that phase target.

Also in Figure 33(f) note the presence of modest reflectance (black) for elements $2 - 4$. These phase shifts are the closest to Φ_0, and consequently required less metal

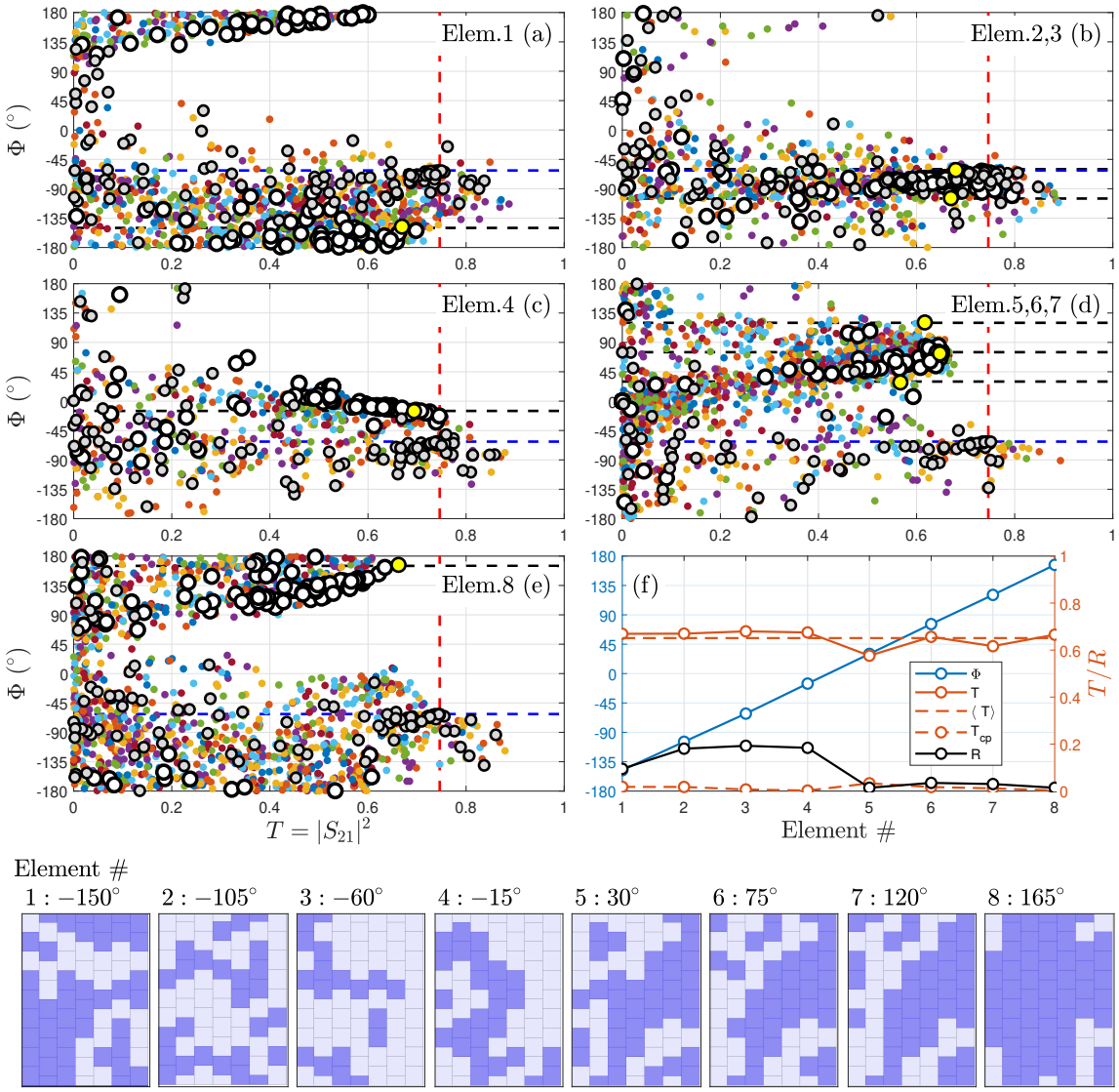

Figure 33. Results of the brick layout/rectangular cavity for a full 8-element discretization of Φ_t are arrayed in Figs. 33(a-f), showing phase shifts of $\{\Phi_t, \Delta\Phi\}$ = (a) $\{-150°, -90°\}$, (b) $\{-135°, -45°\}$, (c) $\{-60°, 0°\}$ and $\{-15°, 45°\}$, (d)$\{30°, 90°\}$, $\{75°, 135°\}$ and $\{120°, 180°\}$, and (e) $\{165°, -135°\}$ Element numbers are linked to their grid layout below (blue=metal), with elements $\{2, 3\}$ and $\{5, 6, 7\}$ taken from the same GA run, as shown in (b) and (d). Yellow dots indicate the performance metrics of the elements chosen for experimental validation, presented in a linear plot in (f).

to achieve this relatively small $\Delta\Phi$, seen in the element layouts at the bottom of Figure 33, resulting in less absorption. By invoking an additional reflectance term to the fitness function—similar to how it was done in Equation (15) for reducing

sidelobes—R could potentially be minimized while T maximized. Lastly, the cross-polarized transmittance T_{cp} is included to demonstrate its relative insignificance.

Repeating these GA optimization simulations for the smaller-resolution 5×8 hex grid, the following results are obtained in Figures 34(a-f) in an identical format as for the brick elements. Generally, the hex elements performed slightly better than the brick elements. The standard deviation from the $\Delta\Phi = 45°$ interval was 1.26°, with an average deviation from Φ_t of 0.23°, while the mean transmittance (red, dashed) is $\langle T \rangle = 0.658$, or $\langle T_n \rangle = 0.882$. This time *all* GA routines were able to attain transmittances equal or greater than at the largest phase shift (#7). The most stark difference is the reduction in reflectance, with no element generating larger than $R = 0.075$, with an average of $\langle R \rangle = 0.03$; in truth, this was likely due to more diligence on the element selection than any advantage in the hex topology. Nevertheless, these low reflectances—along with near-zero cross-polarization—means that all elements can be approximated as excellent polarized Huygens' sources.

4.4 Model Modifications for Real-World Device Fabrication

4.4.1 Sensitivities to t and λ.

The next chapter deals with the attempts at fabricating a real-world metasurface device using these OOP MPL elements, to a level of complexity and resolution never attempted previously by SNL—or anywhere. Over the course of this research effort, both metasurface design and metasurface fabrication were advancing rapidly in tandem, the former at AFIT and the latter at SNL, to the point where sometimes the models were tweaked to adjust for new discoveries in the MPL manufacturing. One of these adjustments was reducing the voxel thickness t to mitigate effects of a phenomenon called "aperture clogging" which occurs during the deposition process[73]. Metal builds up over time on the top of the membrane, and the additional height

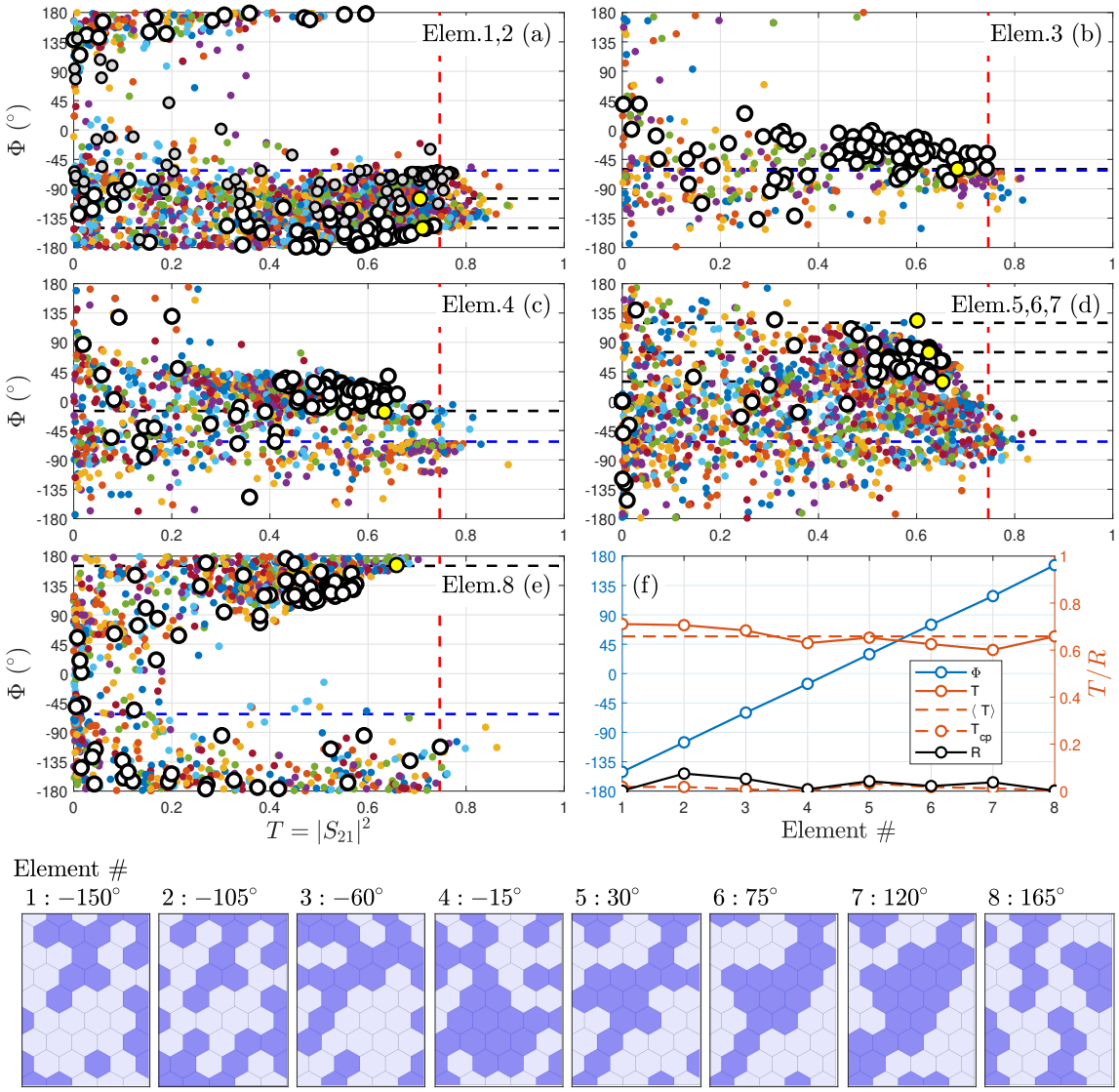

Figure 34. Repeat of Fig. 33 for the smaller-resolution 5×8 hex grid, in an identical format as for the brick elements. Generally, the hex elements performed slightly better than the brick elements. The most stark difference is the reduction in reflectance, with no element generating larger than $R = 0.075$, with an average of $\langle R \rangle = 0.03$.

reduces the projected area of the aperture due to the 45° deposition angle. This creates a taper on the extents of the voxels, as the aperture slowly closes.

Design sensitivity is expected to be high in the MPL design due to the extremely localized fields seen in Figures 29-31. So while slight variations in t or integrity of the voxel geometry likely would not have an effect in regions with weak localization,

conversely in regions of strong localization, variations in geometry can alter the fields to a significant degree. To curtail this effect, t was downsized from 150nm to 100nm. Gold has a skin depth calculated by[77, 78]

$$\delta = \frac{\lambda}{2\pi\Im(n)}, \tag{21}$$

which at $\lambda = 8\mu$m gives the imaginary index value of $\Im(n) = 54.663$ and Equation (21) comes out to $\delta = 23$nm, and at 100nm the fields produced on the Si and air sides of the voxel need only travel about two skin depths before interacting with each other. Therefore, some coupling between the two faces should be expected to exist, which may affect the overall response of the metasurface element.

Unfortunately, this change in t was identified after the exhaustive series of GA routines in Figure 33 and Figure 34 were produced, meaning that either these simulations needed to be re-accomplished, or the extent to which the reduced thickness would affect the phase and amplitude must be understood. If reducing t generates any non-uniform shifts in relative phase between neighbors, then the previously optimized designs may become impractical for real-world optical application; however, if the phases all shift uniformly, then $\Delta\Phi$ remains unchanged.

The variances in T and Φ as a function of t are shown in Figures 35(a,b), respectively, for the brick designs displayed in Figure 33, with $t = 150$nm. To maintain $T \geq 50\%$ for all Φ_t, the metal thickness must remain between $\sim 110\text{nm}-215\text{nm}$, or an approximately 70% span around t, which is highlighted in tan. In this region, each of the $k = 1, 2, ...8$ target phase values (Φ_t^k) shifts with varying—but approximately linear—gradients; nevertheless, a general separation is maintained, sufficient for many optical manipulations, such as lensing.

To quantify this claim, Figure 35(c) shows the relative difference between each

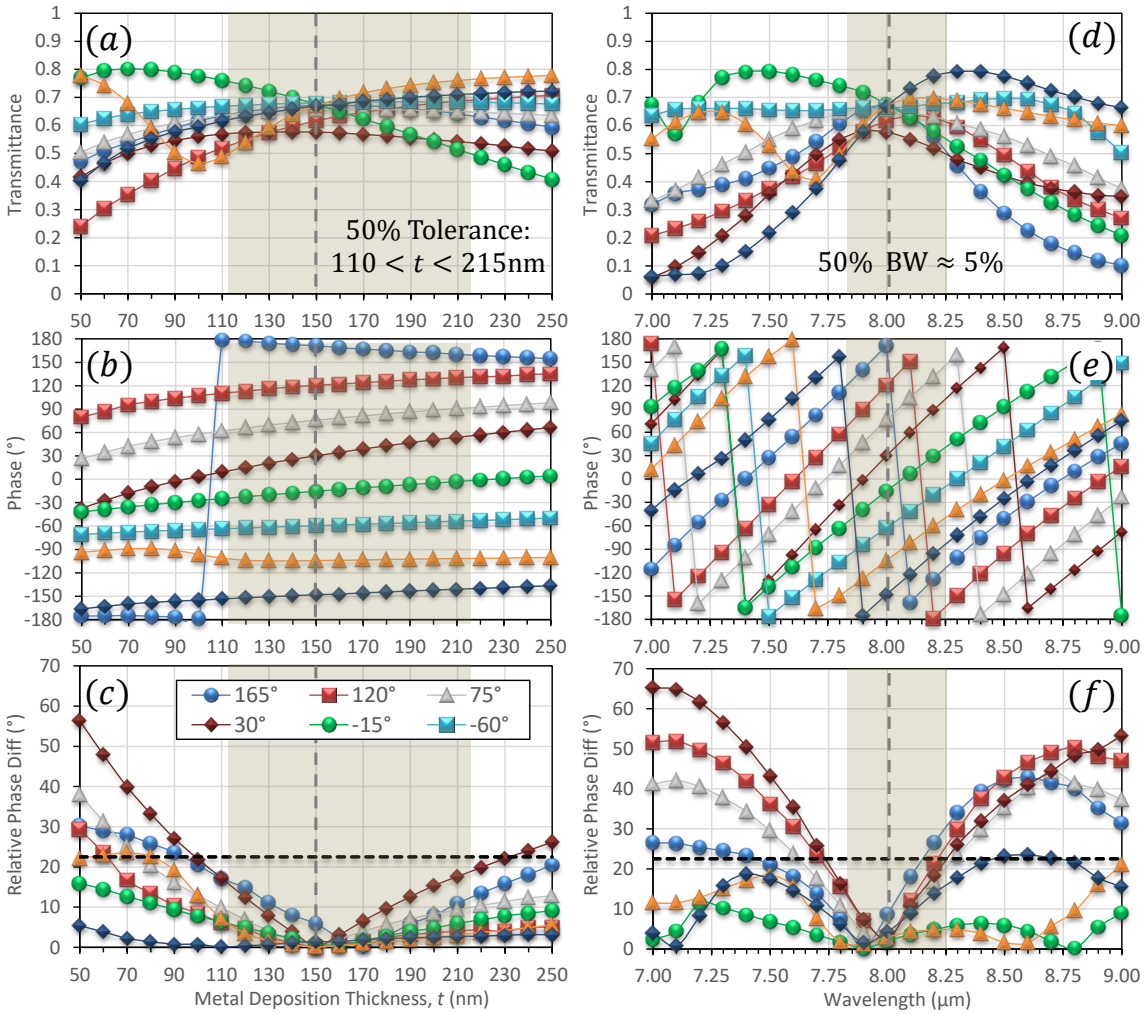

Figure 35. Variances in (a) T and (b) Φ as a function of t for the brick designs displayed in Fig. 33, with $t = 150$nm. To maintain $T \geq 50\%$ for all Φ_t, the metal thickness must remain between ~ 110nm-215nm. A sufficient phase separation is maintained in this region, quantified by examining (c) the relative difference between each phase value and the reference phase for each t, as explained in the main text. (d,e,f) Spectral dispersion for these same metrics. The 3dB bandwidth is limited to just under 5%.

phase value and the reference phase for each t, where this relative difference is defined as $\Delta\Phi_r^k = |\Phi_t^k(t) - \Phi_0(t)| - 360^\circ\ k/N$. For this $N = 8$ element discretization, as long as $\Delta\Phi_r^k \leq\ 360^\circ/(2N) = 22.5^\circ$, then the maximum separation between any neighboring elements is $\leq\ 90^\circ$ in the most extreme case where neighbors are offset from Φ_t by -22.5° and 22.5°, respectively. Therefore, the optical element would thus be no less discretized than $N = 4$, and the criteria for maintaining a diffraction-limited optic—where the Strehl ratio is above 0.8—is met[6, 79]. As can be seen, within the $T \geq 50\%$ range (tan), each $\Delta\Phi_r^k$ falls below 22.5° threshold (black, dashed).

In addition to aperture clogging, there is dispersion of the metal through the mask aperture, as well as metal relaxation, which means that metal topology in the MPL fabrication process cannot be controlled to extremely high precision. Since each design is based on resonant couplings, the MPL metasurface exhibits spectral dispersion, as expected; however, this effect may be erratic with these slight geometric variations. Unlike the seminal V-antennas, which have two coupled resonance modes[1], the increased freedom from geometrical constraints means the modal response from each MPL design is unique and varied. Figures 35(d,e) demonstrate how widely dispersive these designs can be, with a nearly flat spectral response near Φ_0 (60°), single-mode antenna-like responses at others ($120^\circ,\ 165^\circ$), and multi-mode responses at yet others ($-105^\circ, -15^\circ$). This limits the $T \geq 50\%$ (3dB) bandwidth to about 5% around λ_0. It can be seen that not all lines remain sufficiently separable, as in Figure 35(f), which reduces this bandwidth further. Nevertheless, additional scrutiny in design could eliminate the worst offenders to improve bandwidth.

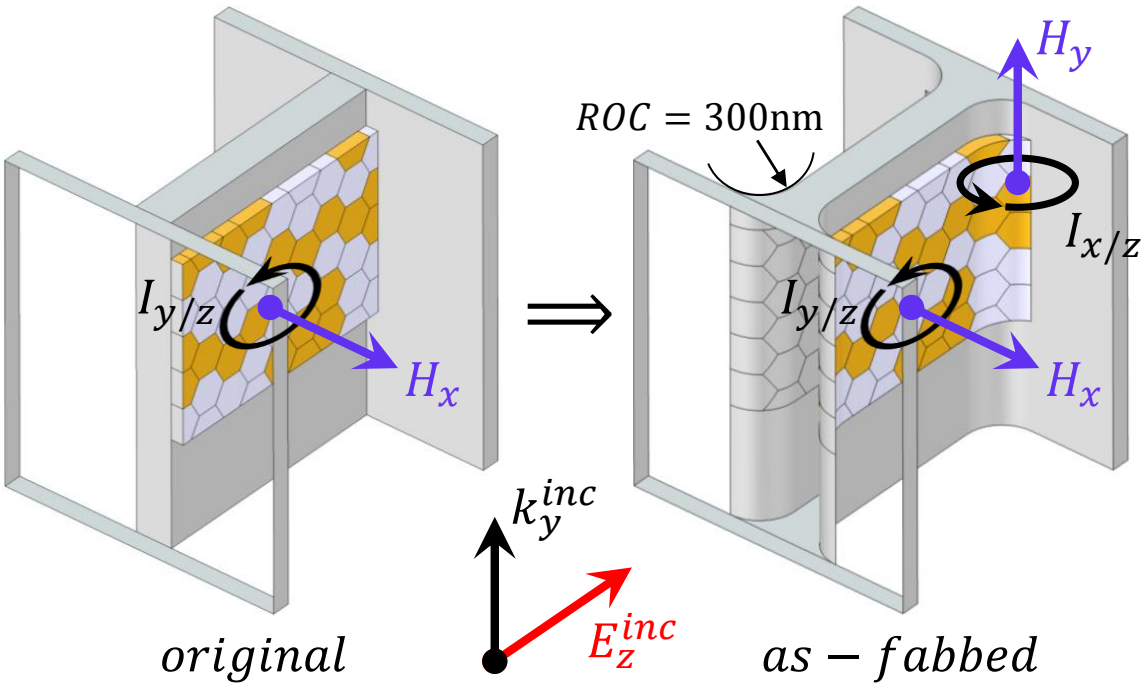

Figure 36. Original "ideal" model and the modified "as-fabbed" model, which exhibits an out-of-plane curvature to mimic what is seen in SEM images of the fabricated MPL structures—see Fig. 21(b) and Chapter 5.2. This curvature generates new induced components: a $\hat{x}$-directed current (I_x) and a $\hat{y}$-directed magnetic field (H_y).

4.4.2 Modeling "As-Fabbed" Curvature.

Viewing some of the previous examples of fabricated MPL structures from Figure 32, a major difference between the ideal COMSOL model and the "as-fabbed" structure is in the curvature in the vertical joints around the interior of the cavity. At first, this might not seem like a concern; however, the ideal model assumes the grid is illuminated by a uniform phasefront after exiting the cavity floor/air interface, and this curvature introduces an inhomogeneity across the interface. In addition to the phase non-linearity, a secondary issue arises, where now portions of the scatterer are curved out of the incidence plane, as can be seen in an example model in Figure 36. Previously, the current induced from the electric response of the OOP scatterer was

generally confined to the incidence plane ($y/z-$plane), and as a result, the magnetic response was generally confined to the orthogonal plane (H_x). Now, the potential exists for an additional component to be generated for each response: a $\hat{x}$-directed current (I_x) and a $\hat{y}$-directed magnetic field (H_y).

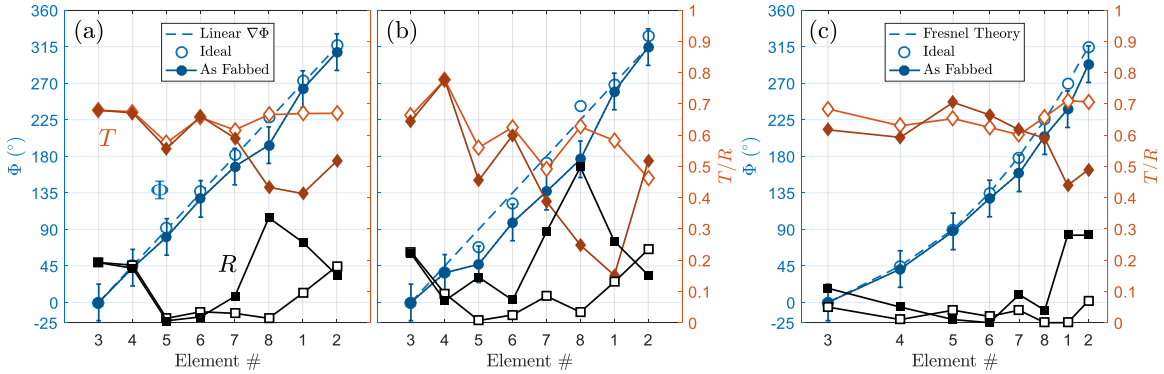

Figure 37. Comparisons between the Φ (blue, circle), T (red, diamond) and R (black, square) for the ideal (open) and as-fabbed (solid) models, spanning all eight elements of the (a) $t = 150$nm and (b) $t = 100$nm brick designs and (c) the hex design. Error bars on the phase points indicate the $22.5°$ upper tolerance for an $N = 8$ discretization. Several elements in (b) performed quite poorly in T and R because the parameters have been shifted too far from the optimized design—this is important because (b) best represents the metasurface beamsteerer as it was fabricated in Chapter 43.

As before, there are two options: repeat the optimization routines, or verify this change will not significantly affect the outcome. Simulations on the as-fabbed brick and hex models were run on the eight elements identified in Figure 33 and Figure 34, respectively, with the results in Figures 37(a-c) showing Φ (blue, circle), T (red, diamond) and R (black, square) for the original ideal (open marker) and modified as-fabbed (filled marker) models, with element #3 at $\Phi_t = -60°$ representing the Φ_0 reference phase. The brick GA routines were run before the reduction in t was decided, so both (a) $t = 150$nm and (b) $t = 100$nm are assessed for the same brick elements, with the phase points compared against a linear gradient (since the these will be

used for a beamsteering application), while the hex elements were only optimized at $t = 100nm$ and compared against a lensing profile. Therefore, (a) and (c) can be used to compare the effects of just modifying the cavity curvature, while (b) can be used to individually compare the effects of both the curvature and reducing the thickness to beyond the recommended tolerances as shown in Figure 35.

There are several things to note in Figure 37. First, an overarching commentary is that each modification seems to affect each element differently, which can be attributed to the widely varying topology of each design: some geometric configurations are more sensitive to perturbations than others in the regions which drive high current densities. Though, this is somewhat true even amongst elements which share the same general topology (coming from the same GA run). For example, brick elements $\{2, 3\}$ differ strongly in how the modifications affected T, though Φ was not considerably off-target; at the same time, both brick and hex elements $\{5, 6, 7\}$ remained largely uniform in response—in fact, these three hex elements showed overall improvement in T and R!

A second comment is that the thickness reduction is much more severe a modification than the cavity curvature. The 100nm as-fabbed results for the brick elements (b, open) are quite poor, and unrepresentative of the potential of the GA optimization technique. Though most elements fall within the $360°/N$ relative phase range (error bars) to support sufficient interference, transmittance drops to below 0.25 for elements $\{1, 8\}$; and with reflection above 0.25 for several elements, these behaviors are no longer Huygens-like in response.

The good news is that all but elements $\{1, 2\}$ for the hex design show resilience in efficiency to the as-fabbed modification, and even those two elements are well within phase tolerances. Since these designs are already at the $t = 100$nm fabrication limitation, and possess the curvature seen in the realized structure, this indicates that

the GA optimization can produce a superior metasurface building block if the proper model parameters are set at the beginning.

Balancing the time constraints of getting into the SNL Silicon Fab queue against running additional GA set for optimal performance, it was decided to push forward with the less-than-optimal brick designs, without any further changes. The relative error in phase shifts all fall within tolerances for what engenders a sufficient phase discretization, as described in the previous subsection; but unfortunately, it was expected that the transmission efficiency for some elements would likely drop considerably from what is given by the GA optimization, and there may be some slight degradation to the integrity of the beam pattern, resulting in less suppression of higher-order diffraction modes. After this determination, the brick and hex designs were submitted to SNL for fabrication of full-scale devices. Further simulation work was accomplished on the full-scale devices after the designs were finalized, and will be presented in the next chapter along with the experimental results.

4.5 Summary

Early in this chapter, baselines of the MPL architecture were identified, and it was validated that the GA could optimize a singular high-resolution hex grid in a square MPL cavity. Yet, the ultimate goals in the chapter were to validate that the COMSOL-based GA routine could reliably produce high-efficiency Huygens-like MPL elements at any target phase value across a 2π span, and to assess those designs for robustness as the development moves towards the next step of realization of metasurface devices. Clearly from Figure 33 and Figure 34, the first goal was successfully attained, with the GA reaching all phase targets at transmittances of the upwards of 60%—or 80% when normalized to the baseline T_0—along with near-zero reflectance. The GA proved to be an extremely reliable tool for converging a large and complex

parameter space into an easily-accessible multidimensional solution space, from which an optimized 3D metasurface element can be selected by a designer.

For the second goal, it was discovered that the response of the MPL designs have quite a bit of tolerance against changes in voxel thickness (t) and to the cavity curvature seen in fabrication (Figure 37). However, it was determined that t must be lowered to a value outside this tolerance, to 100nm, and at this severe modification the designs are no longer performing optimally. Unfortunately, this was discovered after the first set of brick element designs were sent to SNL, and so a full-scale simulation of a device using these elements will be performed in Chapter 5.3 to indicate how this will affect performance. Fortunately, the hex designs were modified in time, and show that the MPL binary grid architecture could still be optimized by the GA at this increased $t = 100$nm constraint. Therefore, the hex designs should produce the best representations of the power of the GA implementation.

V. Simulations and Experimental Validation of Full-Scale Optical Devices Based on Optimized Huygens Elements

5.1 Introduction

With all of the MPL elements identified to span the phase space, all that is left to validate the GA optimization of an OOP Huygens metasurface is to build and confirm efficient function of a representative optical device. Any metasurface application can be built from tailoring the surface to adhere to a specific phase profile, and the most simple demonstrations are a beamsteerer—which has a linear phase gradient profile—and a lens—which has a parabolic phase profile according to Equation (2). This chapter details the entire validation process for the design, build and simulation of these devices, followed by the experimental setup and measurement. The end result will prove the concept that plasmonic OOP scattere rs can compete with dielectric and pseudo-3D plasmonic stacks as viable metasurface architectures.

5.2 Fabrication of MPL-Based Metasurface Devices

5.2.1 Membrane Patterning.

Each MPL element was transferred from a COMSOL model to a fabricable membrane design by associating each voxel to a set of four (for brick) or six (for hex) coordinate pairs, as shown in Figure 38(a) for a sample 6×7 hex design. Whenever discussing a deposited pattern, there are three orientations which must be noted. First, the unit cell view is rotated 90° clockwise from Figures 39, 40, 43-45, and 47-48, where the open face of the cavity was adjacent to the left of the grid. Now, the top of the grid is adjacent to the open face of the cavity, as oriented in Figures 32 and 36. Knowing this, the second item to note is that the individual voxel orientation is also rotated 90°, generating the tiling pattern as in Figure 18(e); this, in fact, was the first

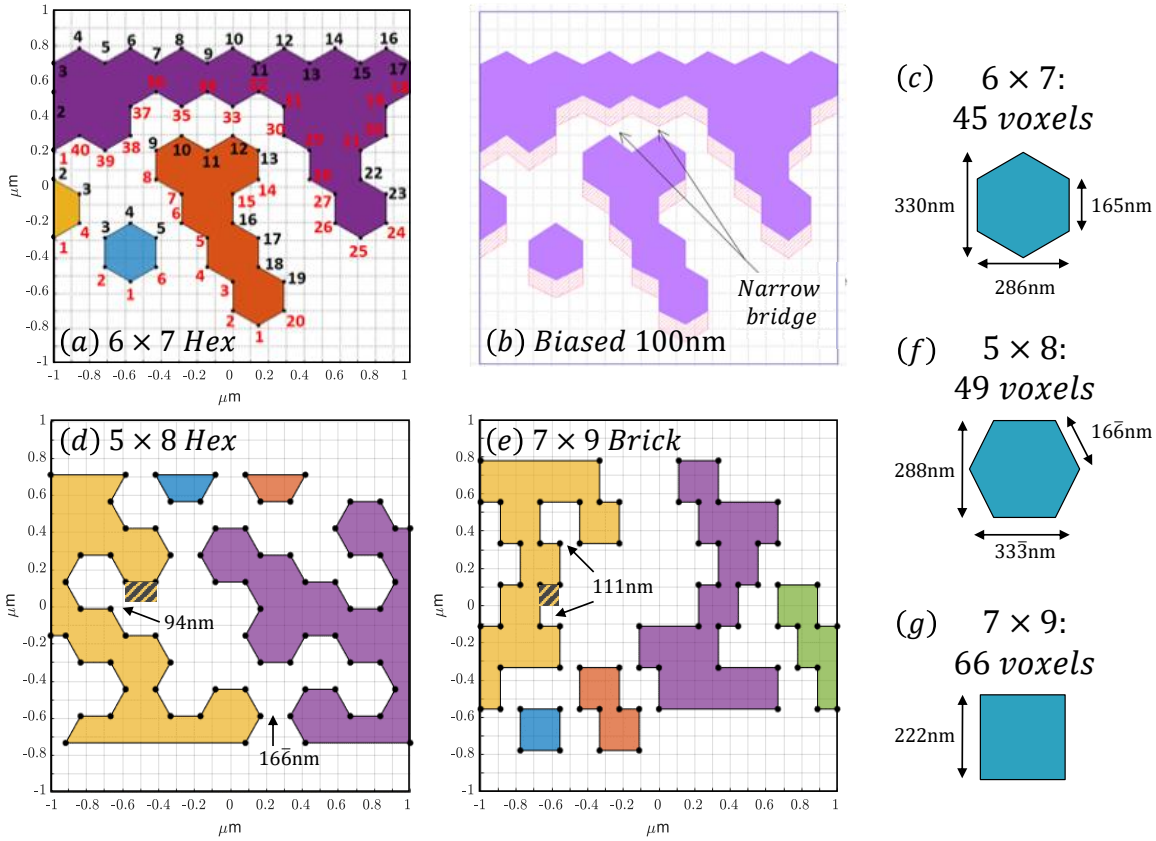

Figure 38. (a) Example 6×7 hex layout showing all non-redundant exterior points, representing a coordinate pair used to draw the contours for lithographic writing of the membrane. However, the original orientations (b,c for hex example) left extremely small gaps when biased 100nm, as required. These orientations were rotated by $90°$ (f,g) to permit more practical minimum gap widths of 94nm for hex (d) and 111nm for brick (e).

hex orientation attempted in this work. Third, it is important to note the deposition orientation; this is for a membrane projection onto the north face of the cavity, so the scatterer would be formed as-seen, as opposed to deposition onto the south face, which would flip the deposited pattern upside down.

For each element, a MATLAB-based program records all non-redundant exterior points, as numbered in Figure 38(a), and arranges them according to the desired phase profile; for a beamsteerer, this is a simple repetition of an $N-$element supercell, and for a lens, the program determines how many repetitions are needed to form each

phase zone for the given the focal length and diameter. The program takes the set of points for the full device and generates the Graphic Data System (GDS) library files, which are the standard inputs for the computerized systems that write lithographic masks. The GDS files were then viewed and edited using an interfacing software such as KLayout [80].

One of these needed edits was in the membrane biasing. As described in Chapter 3.4.2, the finite membrane thickness requires biasing a length of the membrane thickness s in order to account for proper geometric projection at the 45° deposition angle. With deposition on the north wall through a membrane of $s = 100$nm, each membrane pattern required a biasing to the south, meaning all the lower points of the pattern—the red numbers in Figure 38(a)—were extended a length s downward, with the outcome of this shown in Figure 38(b) via the red thatching. Because of this particular voxel orientation (Figure 38(c)), which also controls the geometry, the biasing ends up being along both the longest and the shortest dimensions of the hexagon. Consequently, this leaves only a 65nm piece of membrane wherever the pattern is only one pixel wide, such as the narrow bridges that isolate two metallic geometries (Figure 38(b), arrows). Though not shown here, the same issue occurred with the brick pattern displayed in Figure 18(d), where the bridge widths were reduced to 11nm—a completely impractical condition.

There was an extremely high concern from SNL that these short widths would be too thin to support the membranes, and so a possible solution was found by rotating the orientation of the voxels by 90°, as shown in Figures 38(d,f) for hex and Figures 38(e,g) for brick. This permits the biasing (orange thatch) to be drawn over more manageable distances, where now the smallest widths are indicated by the arrows in (d), measuring a 94nm minimum width diagonally and (e), measuring a 122nm minimum width vertically and 111nm horizontally.

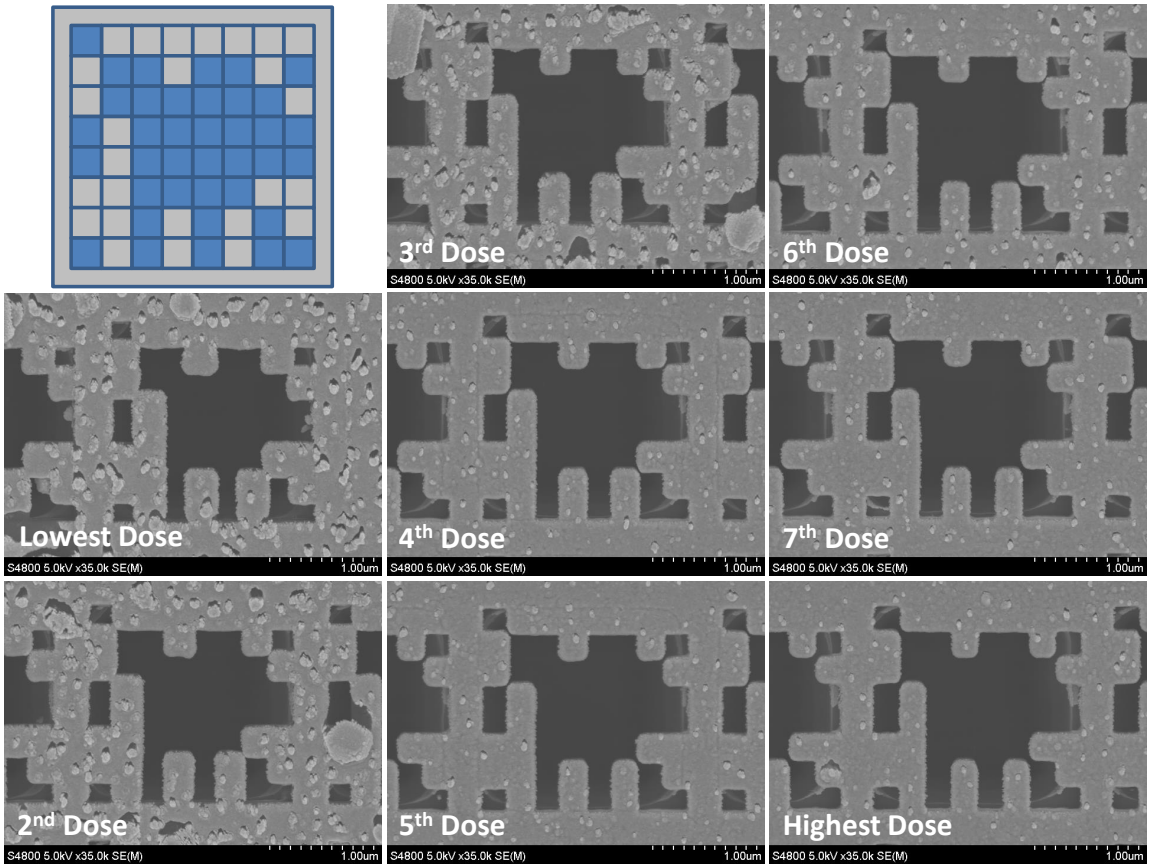

Figure 39. Resulting membranes of calibration test writes of a uniform array of an arbitrary 8×8 square grid of square voxels (upper left). Increasing doses widens the dimensions of each voxel, which can help account for fabrication effects such as blooming (an effect of dispersion through the aperture), metal relaxation and tapering due to aperture clogging.

5.2.2 Initial Fabrication Runs on Square Layout.

While the voxel and biased membrane dimensions were considered reasonable for e-beam lithographic (EBL) techniques (10nm or 20nm resolutions are common), the device fabrication was still considered quite risky because of the uncertainty in how accurately the metal deposition could reproduce the fine geometries of the complex pattern. As a result, a calibration test was performed on a uniform array of an arbitrary 8×8 square grid of square voxels, shown in Figure 39, in order to obtain a sense of what e-beam dosages would achieve the desired sharpness in voxel edges.

The dosage helps adjust the true voxel dimension needed for a replica square on the vertical face, accounting for fabrication effects such as blooming (an effect of dispersion through the aperture), metal relaxation and tapering due to aperture clogging. As dosage increased the individual voxel aperture became enlarged, and some gaps began to appear at the corner-to-corner (CtoC) contacts due to the finite beam size causing rounding of sharp corners.

Results of the aperture size on a single metal deposition on the north wall are shown in Figure 40 for each dose. It was known that the base cavity array used in this test was undersized, as the cavities had contracted a bit in storage; in fact, looking through the membrane in Figure 39, the top of the Si walls can be seen at the extents of the pattern. Accordingly, some clipping of the voxel layout was expected, as was some metal deposits on the top of the structure. Nevertheless, the point of the test was to examine the voxel formation, so this error was acceptable, and a base array with much more accurate dimensions was planned to be used in the final device fabrication.

All doses permitted remarkably sharp corners and good recreation of the main body of the design. CtoC contact breaks down at the 5^{th} dose and lower, which blocks current flow. At the higher doses, some semblance of contact appears, and while the COMSOL simulations in Chapter 3.3.5 demonstrated the response should remain steady for bridges of modest thickness, there was some significant risk this would not hold true experimentation—for example, the high current density may burn out what is essentially a nanometer-scale wire bond. This level of uncertainty resulted in the choice to move to the brick and hex voxels, though from this test it was determined the 6^{th} dose or higher are good starting points for all future attempts.

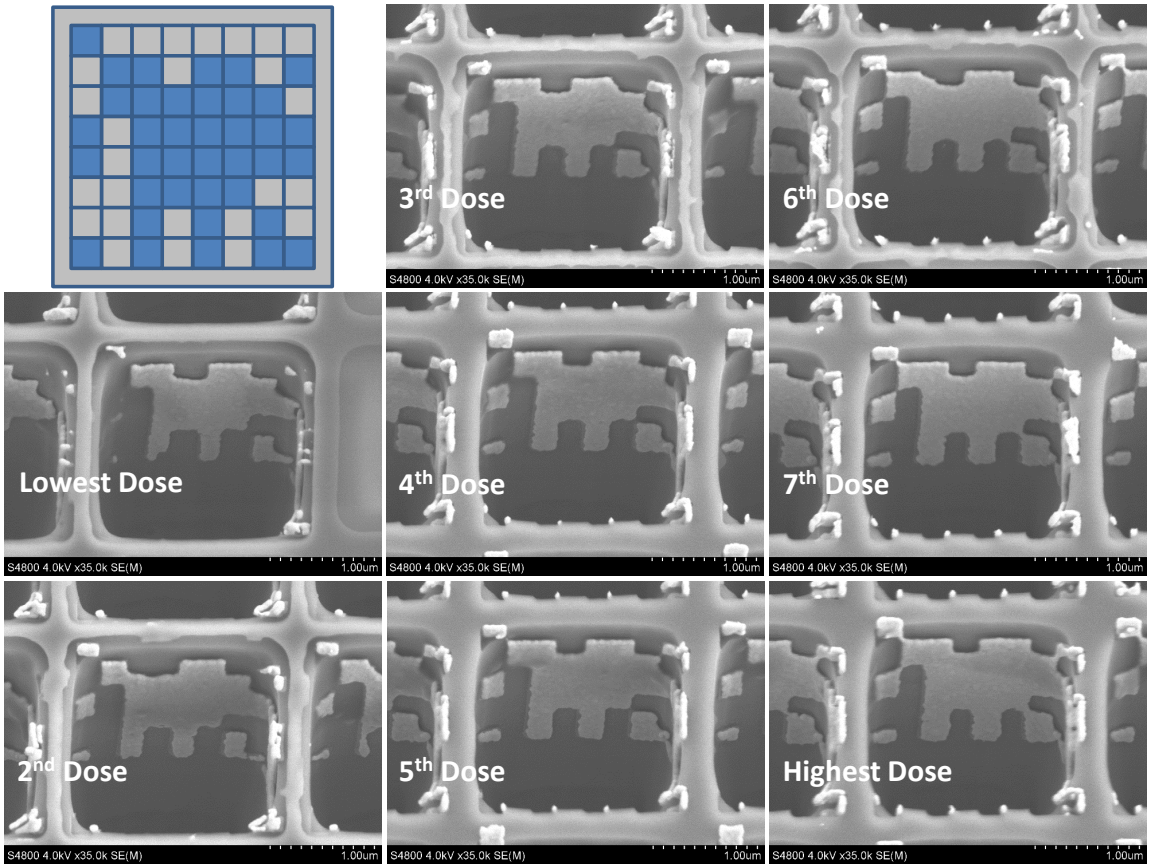

Figure 40. Resulting metal depositions through membranes in Fig. 39. It was known that the base cavity array used in this test was undersized, so some clipping of the voxel layout was expected, as was some metal deposits on the top of the structure. Corner-to-corner contact breaks down at the 5^{th} dose and lower, and current flow between these voxels is blocked.

5.2.3 MPL Fabrication of Elements and Full-Scale Device.

With the previous calibration test carried out, it was a simple transition to fabrication demonstrations of the individual MPL elements. SEM images from each of the major steps are shown in Figure 41(a-e) for the $\Delta\Phi = 180°$ brick design, proceeding through (a) the GDS layout, (b) the post-EBL backfill evacuation, (c) the metal deposition and (d) the post-liftoff final product. There is still some clipping of the pattern at the corners, which was attributed to both the aforementioned cavity relaxation and a slight downward misregistration of the membrane with respect to the

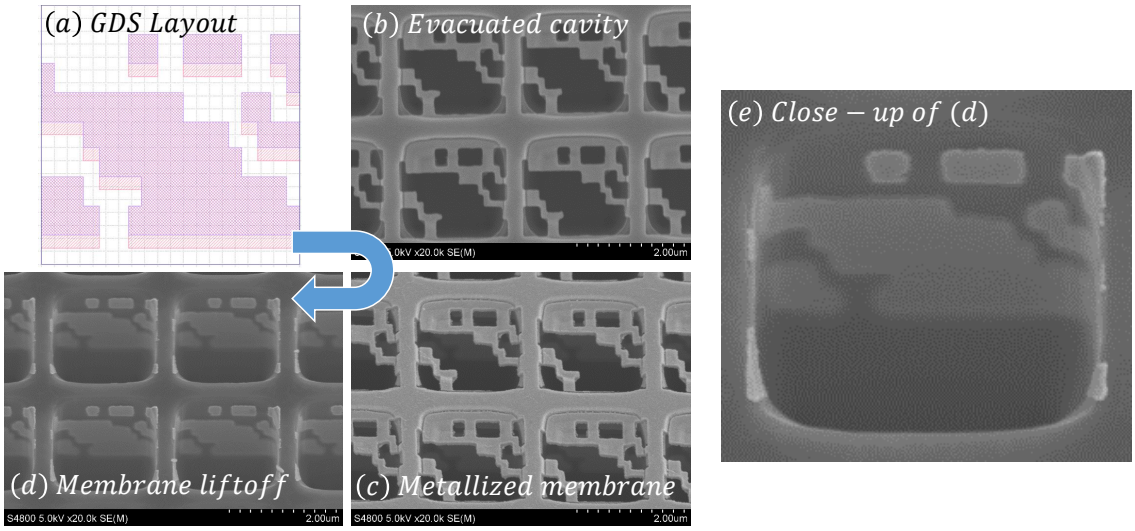

Figure 41. Fabrication of GA optimized brick element targeted for a $\Delta\Phi = 180°$ phase shift, proceeding through (a) the GDS layout, (b) the post-EBL backfill evacuation, (c) the metal deposition and (d) the post-liftoff final product. Some clipping is still apparent, due to misregistration and slight cavity relaxation, but not as severe as in the test articles (Fig. 40). SEM images for the remaining elements can be found in Appendix A.

open face of the cavity; however, the relaxation is not as severe as in the test sample, and the alignment would be improved in the final device fabrication. The overall result is much less metal on the top of the structure, and a near-complete recreation of the fine features in the design, even along the curved regions of the wall to the left and right (see (e) for a close-up view). While the curvature has the tendency to pull the electric and magnetic responses into an additional dimension and modestly shift phase, it was shown in Chapter 4.4.2 that this should still be expected to behave within experimental tolerances for proof-of-concept. Additional SEM images for the remaining fabrication runs on the individual elements can be seen in Appendix A.

From these initial assessments of the realized MPL elements, a proof of concept metasurface optic was designed. The most simple demonstration of a metasurface function is beamsteering, where the phase profile of the surface is that of a linear gradient—the electromagnetic equivalent of a blazed grating[38]. To compensate from having a large bank of available discretized elements to populate the increasing phase

accumulation, the elements are arranged in a supercell such that the phase cycles every 2π. Most published metasurface beamsteering devices have a discretization of $N = 8$ to attain a Strehl ratio of 0.96, which provides a nearly diffraction-limited PSF. However, in order to better control the fabrication uncertainties it was deemed best to simplify the structure to $N = 4$ elements—this gives a Strehl ratio of 0.80, which is considered the minimum for diffraction-limited behavior in any diffractive optic, including metasurfaces[40, 6].

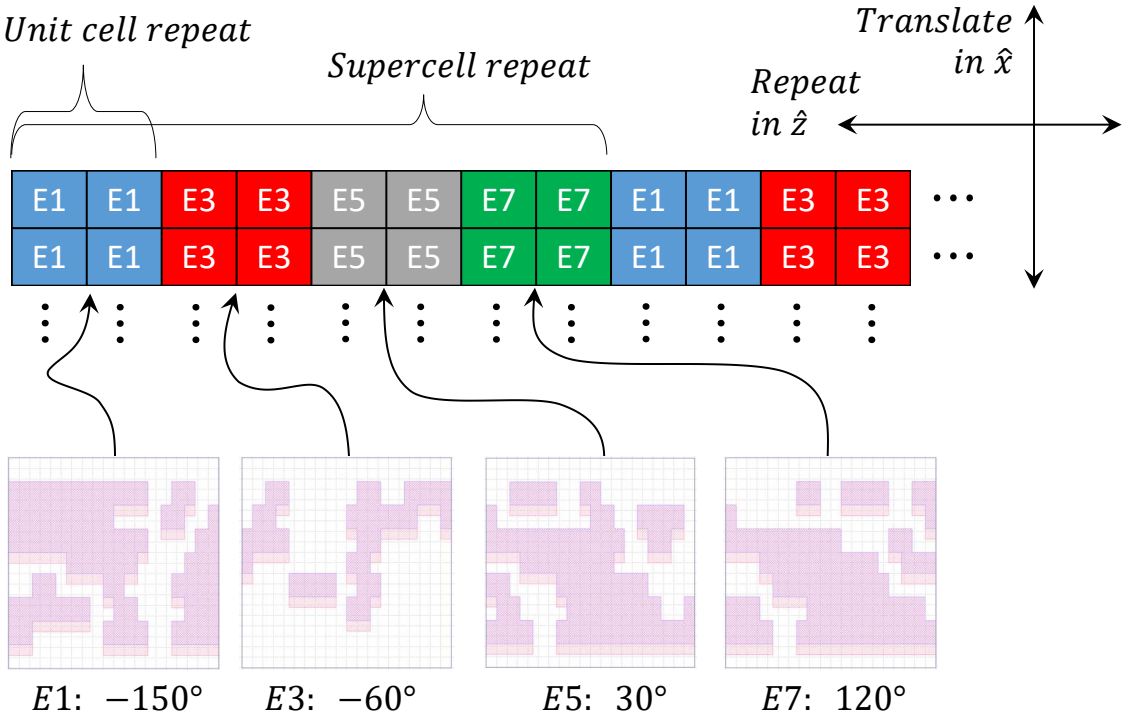

Figure 42. Design of an $8-$element metasurface beamsteering device based on the brick designs, where each of the four designs are repeated once over, using elements $\{1, 3, 5, 7\}$ that gives absolute phase values of $\Phi_t = \{-150, -60, 30, 120\}^\circ$, or referenced phase values of $\Delta\Phi_0 = \{-90, -0, 90, 180\}^\circ$. This 2-repeat array stretched the broadside steering angle to 23.50° from the optical axis.

The issue with a $4-$element beamsteerer is that the gradient is steep, which means the steering occurs at a sharp angle from the optical axis; for a periodicity of $a = 2.3\mu$m, this is observed at 41°. At this angle, there was a major concern that a significant portion of the response would be cut-off by the aperture in the sample

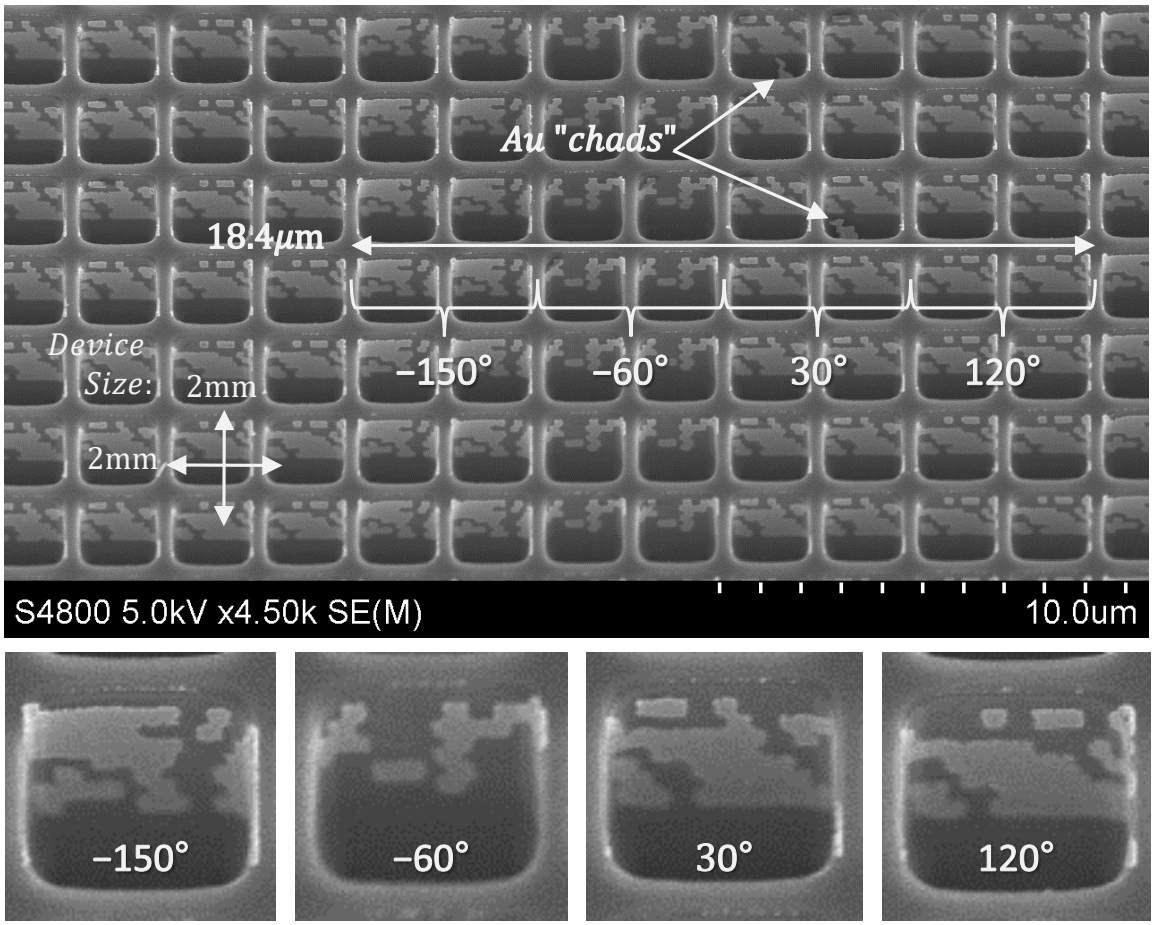

Figure 43. An SEM of the final, full-scale experimental beamsteering device, along with enlarged images of each element. Comparing to the GDS layouts in Fig. 42, the elements came out in excellent form, faithfully recreated with the least amount of clipping and improved registration over previous fab tests. The smallest features are the 111nm are clearly discernible, with good element-to-element reproducibility.

holder. Because of the extreme level of detail demanded by EBL writing of the MPL elements, they were not going to be as large as the 5mm V-antenna lenses—rather, the maximum size which could be fabricated was only 2mm×2mm, leading to a 43% reduction in the functional metasurface area for a 1mm-thick sample holder. In order to improve this reduction, an 8−element design was proposed based on the brick designs, where each of the four designs are repeated once over, as shown in Figure 42, using elements $\{1, 3, 5, 7\}$ that gives absolute phase values of $\Phi_t = \{-150, -60, 30, 120\}^\circ$, or referenced phase values of $\Delta\Phi_0 = \{-90, -0, 90, 180\}^\circ$. This 2-repeat array stretched

the broadside steering angle to $\theta = 90° - \arctan(8 \times 2.3\mu\text{m}/8\mu\text{m}) = 23.50°$ from the optical axis, which cuts the loss in half.

An SEM of the final, full-scale experimental device is shown in Figure 43, along with enlarged images of each element. Comparing to the GDS layouts in Figure 42, the elements came out in excellent form, faithfully recreated with the least amount of clipping and improved registration over previous fab tests. The smallest features are the 111nm are clearly discernible, with good element-to-element reproducibility. The major perturbations seen were what appear to be free-floating chads of gold that were formed on the membrane during deposition and had fallen off into the cavity, as indicated by the arrows. Given the small sample size, it is difficult to discern how prevalent this defect was, but viewing the larger image (Figure 43 is cropped) it only appeared in the 30° elements, and at a rate of 11%.

5.3 Simulation of Full-Scale Devices

5.3.1 Extracting Device-Level Far-Field Behavior.

Using the COMSOL models that were modified and solved for the effects of fabrication, full-scale devices can be simulated for comparison to experimental measurement. However, the realized 2mm beamsteerer requires 108 repetitions of the 18.4μm supercell, for a total of 864 elements, and so a model of this size could never be meshed in any numerical solver. However, many near-to-far field projections exist which can assist in recreating the full-scale device, such as the Stratton-Chu formulation given by Equation (13).

The near-field surface integrals given in this equation require tangential field components that must be extracted from the simulation. This was accomplished by discretizing a virtual boundary in the $x/z-$plane a distance $\lambda_0/4$ away from the open face of the cavity into a set of $i \times j$ probes $P^k_{i,j}$ that record the tangential fields

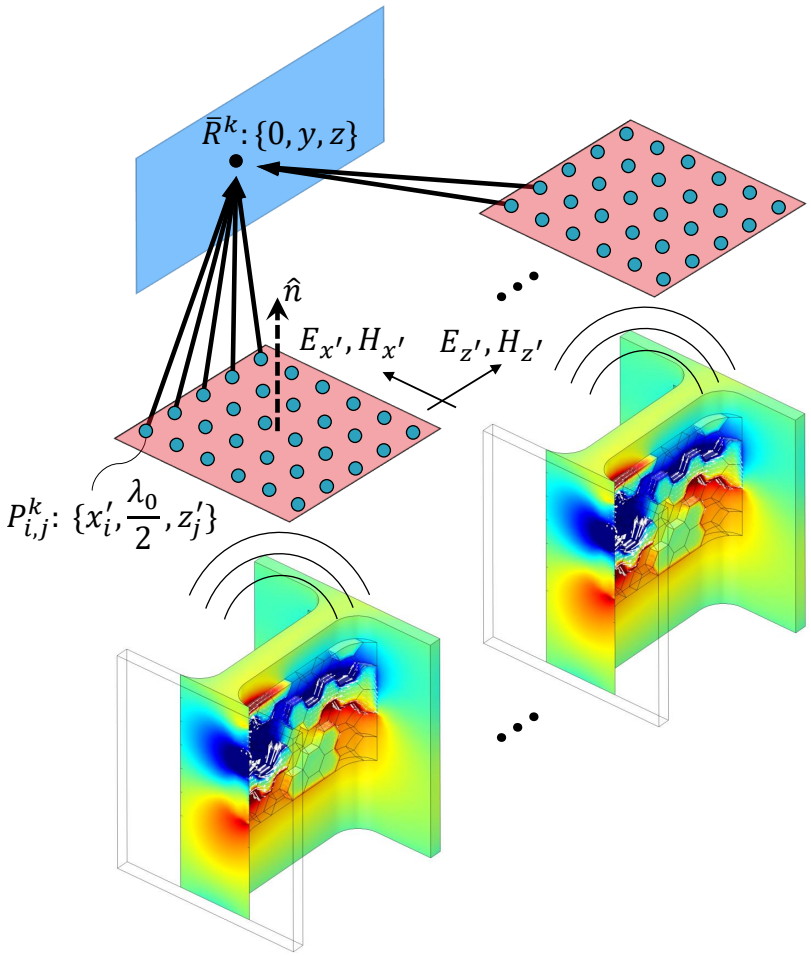

Figure 44. A depiction for calculating a discrete surface integral for the Stratton-Chu far-field (Eq. 13) using the COMSOL results, by discretizing a virtual boundary (pink) into a set of evenly distributed $i \times j$ probes $P^k_{i,j}$. These probes record the tangential fields $\{E^{i,j}_{x'}, E^{i,j}_{z'}, H^{i,j}_{x'}, H^{i,j}_{z'}\}$ at the surface coordinates $\{x'_i, \lambda_0/2, z'_j\}$ and are input into Eq. 22.

$\{E^{i,j}_{x'}, E^{i,j}_{z'}, H^{i,j}_{x'}, H^{i,j}_{z'}\}$ at the discrete surface coordinates $\{x'_i, \lambda_0/2, z'_j\}$, as shown in Figure 44. From here, there is a great amount of freedom in implementation, where any number of k elements can be arrayed in any manner, using the field from any element of a desired phase response to discretely build any phase profile necessary for a given optical function.

For a 1D array (along $\hat{z}$) of k elements at position z_k with uniform probe spacings $\Delta x'$ and $\Delta z'$ along the near-field boundary, Equation (13) is presented in a discrete form for the far field at any point in the plane $\{x = 0, y, z\}$ as

$$\overline{E}^k(x,y,z) = -i\frac{e^{-i\frac{2\pi}{\lambda_0}\sqrt{y^2+(z-z_k)^2}}}{2\lambda_0(y^2+(z-z_k)^2)}\Delta x'\Delta z' \times \sum_{i,j=1}^{5}\left\{\begin{array}{c}\left(E^{i,j}_{x'}y + \frac{Z}{\sqrt{y^2+(z-z_k)^2}}H^{i,j}_{z'}\left(y^2+(z-z_k)^2\right)\right)\hat{x} \\ \left(E^{i,j}_{z'}(z-z_k) + \frac{Z}{\sqrt{y^2+(z-z_k)^2}}H^{i,j}_{x'}y(z-z_k)\right)\hat{y} \\ \left(-E^{i,j}_{z'}y - \frac{Z}{\sqrt{y^2+(z-z_k)^2}}H^{i,j}_{x'}y^2\right)\hat{z}\end{array}\right\}. \quad (22)$$

The distance z_k accounts for the additional phase accumulated along the physical spacing, independent of the phase of the metasurface element. This calculation can easily accommodate 2D surfaces and 3D field profiles, but for the purposes of this comparison a beamsteerer or a rectangular lens can be sufficiently represented by sampling the 2D profile of a 1D array.

5.3.2 Metasurface Beamsteerer using Brick Elements.

To illustrate the effects that the model modifications have on the device functionality, the near fields from both the GA-optimized ideal model designed at $t = 150$nm and the $t = 100$nm as-fabbed model were extracted and input into Equation (22) for computation of the full-scale devices. The computed beamsteerer in Figure 45(a-f) is displayed for a full 108-supercell device of (a) the ideal and (b) the as-fabbed model, along with a 10-supercell close-up of the E_z field components for (c) the ideal and (d) the as-fabbed, as well as the E_y components (e,f, respectively). The diffraction order designations follow from the metasurface image in Figure 43: from the perspective of the propagation direction outward, the gradient in phase is increasing to the left, which means the beamsteering should be to the right and is designated as the +1 diffraction order.

Both the (a) ideal and (b) as-fabbed models achieve an identical steering angle of 25.8° in the +1-order, off by $\sim 10\%$ from the theoretical angle of 23.5°. Since this

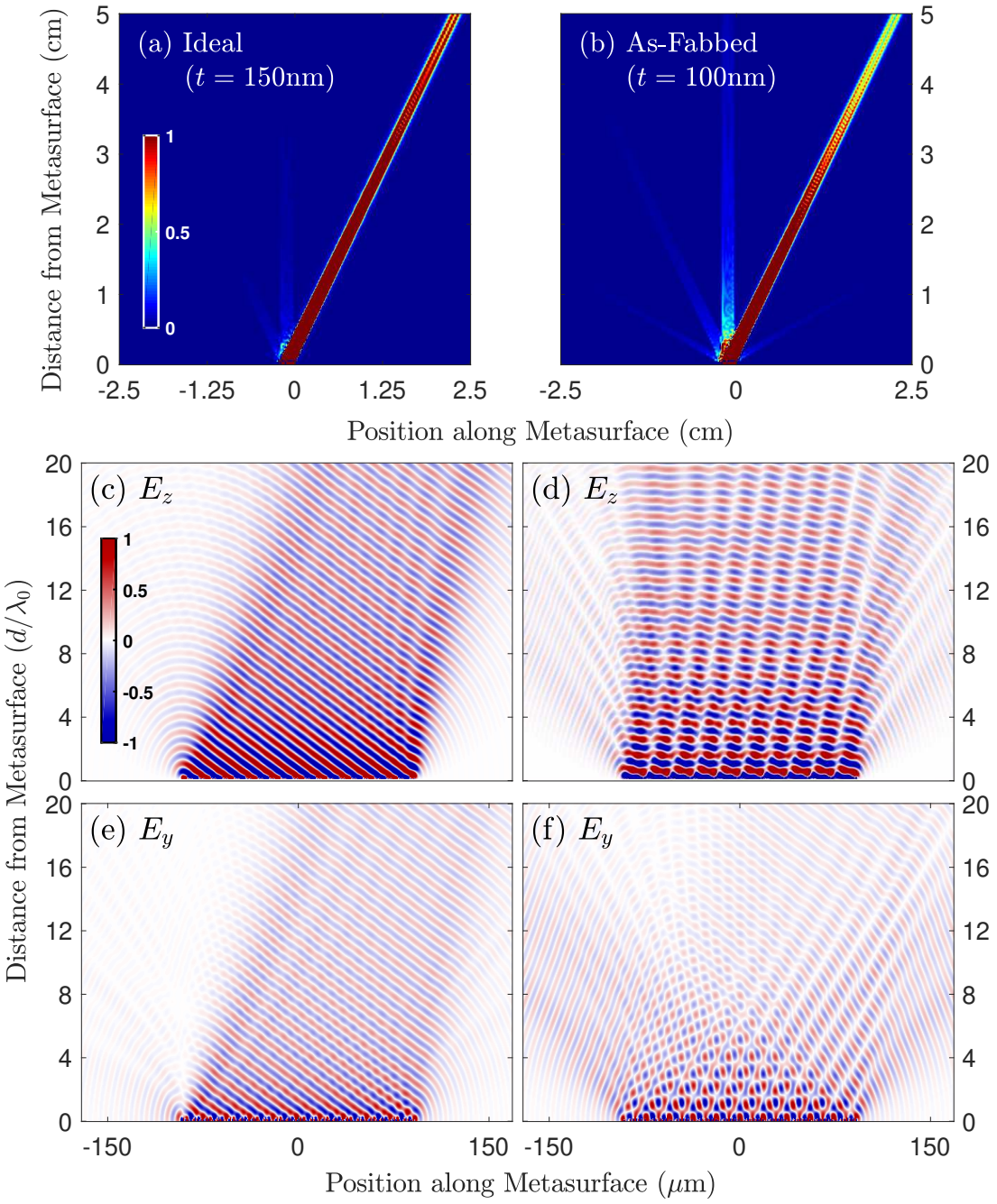

Figure 45. Computed results from far-field projection of a metasurface beamsteerer, as per Eq. 22. The intensity profiles of a full 108-supercell device using (a) the ideal, GA optimized, t = 150nm design and (b) the as-fabbed, t = 100nm modified design both show an identical steering angle of 25.8° in the +1-order ($\sim$ 10% from theory of 23.5°), and contributions into the 0-, −1-, and ±4-orders of $\leq$ 5%. These contributions are higher in (b), due to the modifications, which can seen to affect the near field of a 10-supercell device, comparing the ideal (c) E_z and (e) E_y to the as-fabbed counterparts (d,f). The cross-polarized E_x component does not contribute significantly.

beamsteering error is irrespective of the design, it is most likely due to truncation error or the inherent phasefront error from the coarse discretization. Error in metasurface steering angle is not commonly reported in literature; all examined works show comparisons between simulation and experimental measurements, but the experimental data points are only shown as existing within a few degrees of proximity of numerical or theoretical results—without exact values labeled—and simply state the device as "in good agreement" [18, 81, 7, 3]. One dielectric metasurface example stated a 4% error[82], but no articles explain sources of error outside fabrication, which is obviously not the case, here.

Intensities (a,b) were normalized to the max of either model, and viewed at saturated scale to highlight the weak higher-order diffraction modes. The beam of the as-fabbed device was about 37% reduced in intensity compared to the ideal device, which is on the order of the transmittance reduction in the as-fabbed brick element #1. The diffraction strengths for the ideal (as-fabbed) device are about 2.5% (5%) into the 0-order, 0.5% (1%) into the -1-order, and 2% (1%) into the ± 4-orders, with the latter having a simulated value of $\pm 60.4°$, very close to the theoretical value of $\pm 60.1°$. The ideal near-field E_z component of the smaller simulated device (c) shows a near pristine phasefront, with a small contribution into the $0-$ and -1-orders due to edge effects, while the as-fabbed field is significantly more pronounced in effect due to the perturbations to the linearity of the phase gradient and the amplitude reduction in some elements. Some of the additional energy channeled into the higher orders may be attributed to the E_y component (e,f), which is directed along the optical axis. It has a reasonable strength compared to E_z (average $|E_y| = 0.37|E_z|$), and can be seen clearly to contribute to several orders; conversely, the cross-polarized E_x component is quite weak (average $|E_x| = 0.01|E_z|$—not shown), and would not be a prevailing factor for these modes.

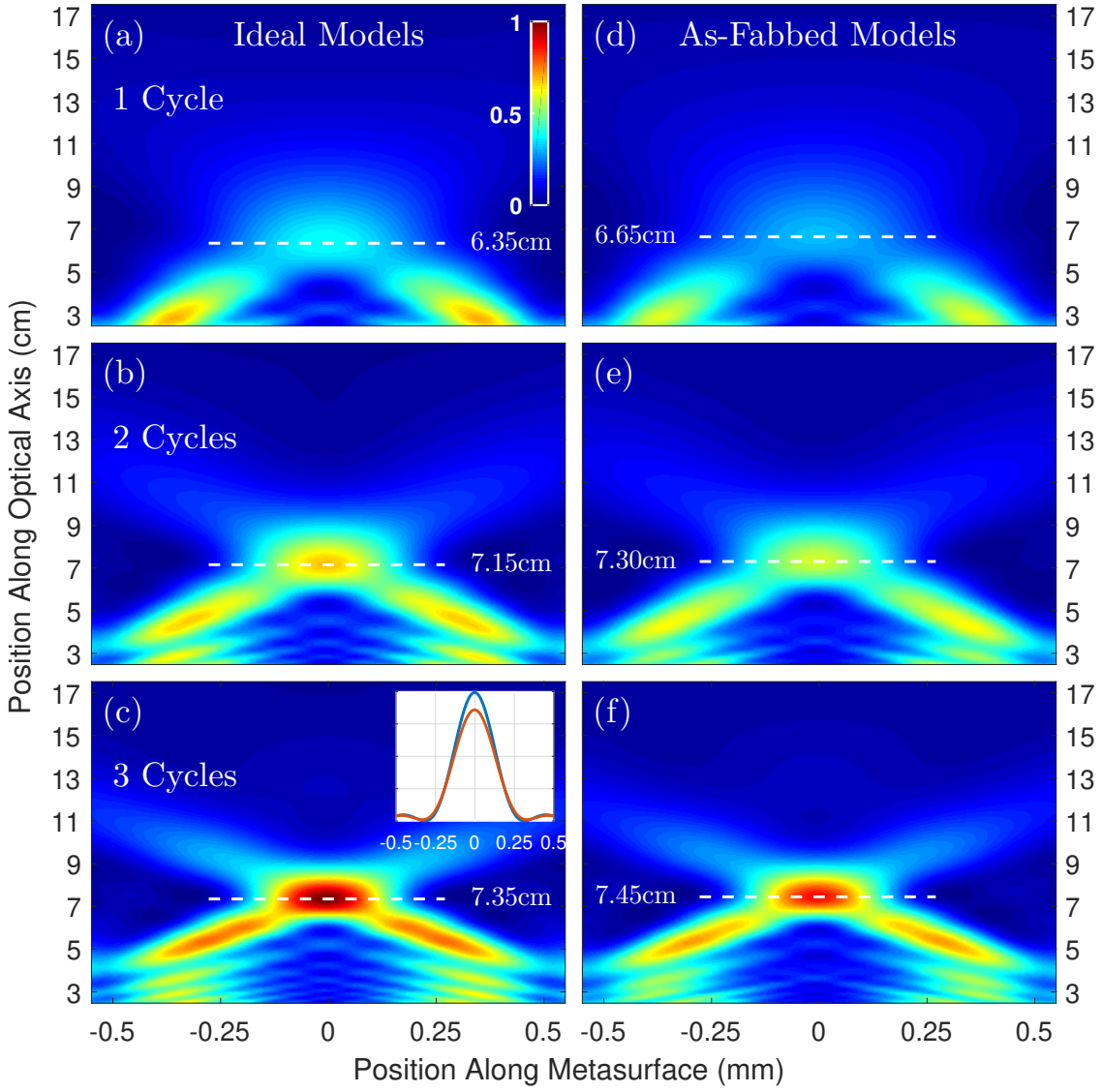

Figure 46. Computed results from far-field projection of a $f_0 = 7.5$cm lens metasurface lens at 8μm, as per Eq. 22, over increasing phase cycles for the ideal (a-c) and the as-fabbed (d-f) models. Accuracy in the focus improves as the cycles increase, to within $1-2\%$ of f_0 for the 3-cycle lenses (c,f). As expected due to the cavity modification, intensities dropped slightly for the as-fabbed lenses, by about $15-20\%$ compared to the ideal lenses.

5.3.3 Metasurface Lens using Hex Elements.

Since the hex designs were only optimized at $t = 100$nm, the difference between the ideal and as-fabbed hex models was only the curvature. How this affects the depth-

of-field for a $f_0 = 7.5$cm lens at 8μm is shown in Figures 46(a-f) over increasing phase cycles for the ideal (a-c) and the as-fabbed (d-f) models. Each image was normalized to the max of all images, which was the 3-cycle as-fabbed model (f), and several distinguishing characteristics can be identified between the two models.

The focal points begin significantly far from f_0, by 10-15% for the 1-cycle lens (a,d), down to $1-2\%$ for the 3-cycle lens (c,f), which follows the trends seen in the measurements of the V-antenna lenses developed for the COMSOL validation task, where a decrease in the number of full phase cycles brings about a decrease in f [83]. The 3-cycle lenses were the only set that maintained a diffraction-limited performance, presented in the inset of (e), defined as the diameter of the first set of nulls defined as

$$D_{PSF} = \frac{\lambda f}{D_{ap}/2}, \tag{23}$$

which leads to a theoretical $D_{PSF} = 632.4\mu$m with the aperture diameter being the 3-cycle cut-off at $D_{ap} = 3.795$mm. This matches well to the calculated $D_{PSF} = 631.7\mu$m, given the resolution of the calculation.

In addition, the diffraction pattern of the as-fabbed model between the two side lobes is not as pristine as in the ideal model, which was expected because of the slight deviation the phase profile in Figure 37. Nevertheless, this is much less pronounced than the effect of phase perturbations in the simulated beamsteerer, since the hex elements did not have large parameter modifications as did the brick elements.

5.4 Experimental Validation of Fabricated Devices

5.4.1 The Bi-Directional Transmittance Distribution Function (BTDF).

In this section, validation of a full-scale metasurface device is presented through experimental measurement of the optical function. While the V-antenna lenses were validated for proper focusing function by directly imaging the depth-of-field using a microbolometer array, a more informational measurement can be made for a beam-steerer through application of the bi-directional transmittance distribution function (BTDF), in conjunction with a device developed by Schmitt Measurement Systems called the "Complete Angle Scatter Instrument" (CASI®). Detailed expositions of the BTDF and the CASI are not carried out here; for this, the reader is referred to some original published works on the subjects[84, 85, 86]. However, a short overview in how the BTDF and the CASI relate to the validation measurements will be described henceforth.

The BTDF (τ) relates the differential transmitted radiance, expressed as the radiant flux density (power/unit area) per solid angle as a function of observation angle (in $W/\mathrm{m}^2 \cdot sr$), to the differential incident irradiance, expressed as the received radiant flux density as a function of a surface element orientation (in W/m^2); and thus τ is a function of the incident and outgoing polar and azimuthal angles $\Omega = \{\theta, \phi\}$, given here in the spectrally-independent form (in $1/sr$)

$$\tau(\Omega_i, \Omega_t) = \frac{dL_t^e(\Omega_t)}{dE_i^e(\Omega_i)} = \frac{dL_t^e}{d\left(\int_{\Omega_i} L_i^e \cos\theta_i d\Omega_i\right)} = \frac{dL_t^e/d\Omega_i}{L_i^e \cos\theta_i} = \frac{dP_t^e/d\Omega_t}{P_i^e \cos\theta_t}, \qquad (24)$$

where $L_{i,t}^e$ is the incident or transmitted radiance and the definition for the differential radiant flux (power) $dP = L\cos\theta dA d\Omega$ is used. The differential area dA falls out of the ratio, as the area receiving the power and the area radiating the power are the

same. Angles are measured from the normal of the surface upon where they interact: the incident face for $\{\theta_i, \phi_i\}$ and the transmitting face for $\{\theta_t, \phi_t\}$

The final form of τ in Equation (24) allows for straightforward application of a measurement technique wherein τ can be determined by sampling P_i and P_t through a sample as the observation solid angle Ω_t is swept over some defined space. For assessment of the metasurface beamsteerer performance, all that is needed is to illuminate the device at normal incidence ($\theta_i = 0, \phi_i = 0$) to obtain a constant incident power ($P_i(0,0)$), and record the power variation over only the polar angles $P_t(\theta_t, 0)$, making sure to have the device aligned such that θ_t scans in the plane that comprises the surface phase gradient and the optical axis.

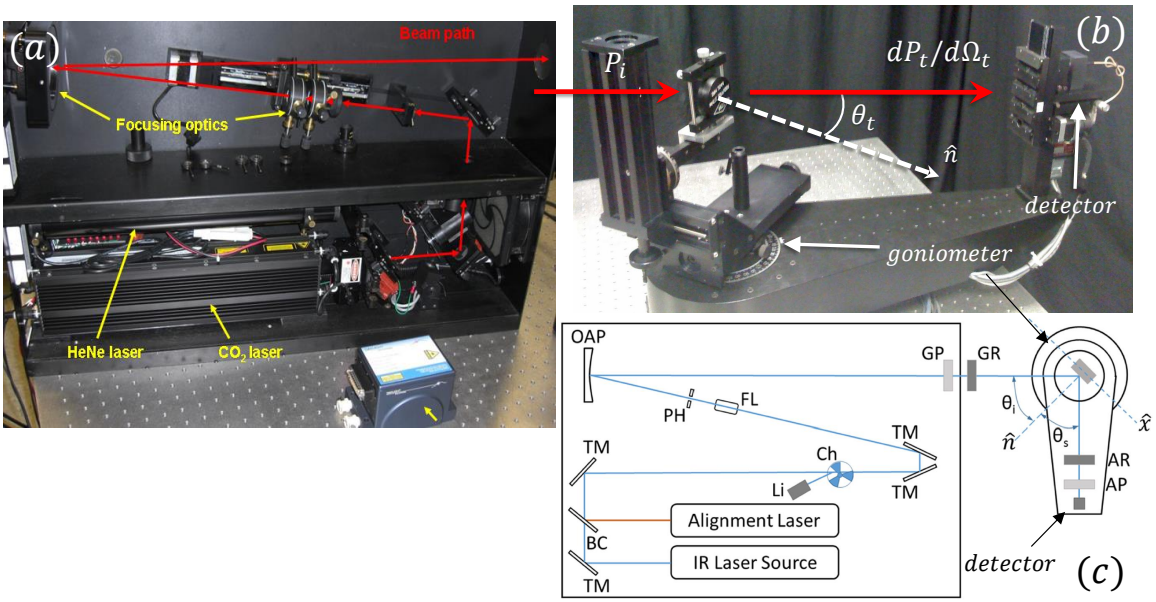

Figure 47. (a,b) Photos and (c) a system diagram of the Complete Angle Scatterer Instrument (CASI). (a) Light from a collimated quantum cascade laser (QCL) operating in the range of $7.7 - 8.2\mu$m passes through a series of components for optical alignment and beam manipulation, which re-collimates the beam into the sample. (b) The sample sits at the center of rotation of a goniometer (b), at the end of which is the detector that can revolve about the sample mount. Polarimetric data can be generated and analyzed through the polarizer-retarder generator pair (GP/GR) and the retarder-polarizer analyzer pair (AR/AP) (shown in (c)). Images (a,c) taken from Ref. [86] and image (b) taken from Ref. [85], by permission.

5.4.2 CASI Description.

The CASI located at AFIT was designed to carry out exactly such a measurement. Referencing the system diagram (c) and photos (a,b) in Figure 47, light from a collimated quantum cascade laser (QCL) operating in the range of $7.7-8.2\mu$m passes through a series of components for optical alignment—including a beam combiner (BC), turning mirrors (TM), a chopper (Ch) with lock-in amplifier (Li)—then a focusing lens (FL) passes the beam through a pinhole (PH), diverging the light onto an off-axis parabolic mirror (OAP), which re-collimates the beam into the sample. The sample sits at the center of rotation of a goniometer, which serves as the origin for referencing all angles, and at the end of the goniometer arm is the detector, which has a freedom of revolution about the sample mount. Polarimetric data can be generated and analyzed through proper implementation of the polarizer-retarder generator pair (GP/GR) and the retarder-polarizer analyzer pair (AR/AP), though the particular combinations depend on the polarization state of the incident and sampled signals—as will be inferred in the following paragraphs, the AR was not utilized, in this work. Though the close-up image of the goniometer in (b) was set-up for a reflectance measurement (BRDF), fixing the stage at $\theta_i = 0$ such that a BTDF measurement can be made is a trivial adjustment.

Oftentimes, limited lithographic fabrication resources demand multiple samples be imprinted on a single wafer, which was indeed the case both for the V-antenna and the MPL samples. In thermal infrared, this situation might cause unintended radiation if any samples on the shared wafer also respond in this spectral regime, where heating of the wafer from the incident light can excite the sample. To prevent this for the V-antenna samples, a mask with a single exit aperture designed for the device under test was used. However, due to the small MPL sample size, an aperture would cut down a great deal of observable area, thus reducing both radiance as a function of

angle and the range of measurable observation angles. Therefore, a custom sample holder was designed, seen in Figure 48(a,b) which consists of a slab with an entrance aperture that isolates the irradiation of the sample, and a recessed region in which the sample sits. The recess is cut such that the 2mm metasurface exactly aligns with the entrance aperture. Instead of another slab with an exit aperture on the transmission side, a bracket was custom fit to slide over the Si wafer and brace it down, allowing unobstructed view of the entire 180° observation range of θ_t. A second aperture was cut to illuminate the baseline, unmetalized MPL cavity array, for normalization to the BTDF measurement.

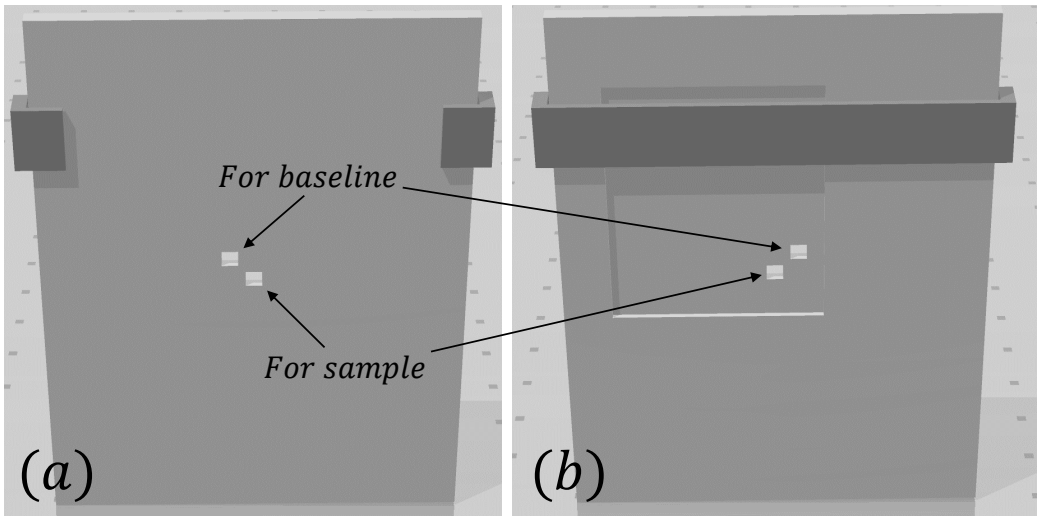

Figure 48. The (a) incidence side and (b) transmittance side of a custom sample holder designed to isolate the 2mm×2mm beamsteerer. It consists of a slab with two entrance apertures: one that isolates the irradiation of the sample, and one that isolates the baseline, unmetalized cavity array. A recessed region in which the sample sits is cut such that the metasurface is exactly aligned with the center aperture. A bracket was custom fit to slide over the Si wafer and brace it down, allowing unobstructed view of the entire 180° observation range of θ_t

The BTDF was measured for both 2mm×2mm MPL metasurface created from four unique OOP brick elements and the unmetalized MPL cavity array, which will serve as a normalization baseline to highlight the effects of the metasurface function. Each was sampled at the four linear polarization states: H/H, for horizontal on

sample, horizontal on sensor; H/V, for horizontal on sample, vertical on sensor; along with V/V and V/H. The setup required modification for each of these states. For H/H, only the GP and AP were needed to clean up the incident beam and emitted signal, correcting for any slight misalignments. For H/V, the AP was simply rotated for vertical polarization. For V/V, the GR was placed before the GP to rotate the incident beam 90°. By setting the GP to the vertical orientation, the proper retardation for vertical incident light can be obtained by rotating the GR until a maximum throughput was obtained. Finally, for V/H, the setup was identical to V/V, but with the AP rotated to horizontal polarization.

The grid layouts of the sample device were aligned to the H polarization and illuminated at the incidence angle $\theta_i = 0°$. There was no cross-polarized transmission observed in either the individual element (via S_{31}) or device (via E_x) simulations, nor did the models respond to a cross-polarized incident field, so it was expected that only H/H would provide any significant response. Furthermore, the unmetalized cavities are symmetric about the optical axis, so there should be little difference in the H/H vs. V/V and H/V vs. V/H responses, and it should be nearly all in the 0-order mode, as the periodicity of $a = 2.3\mu$m is below the ± 1-order diffraction cutoff of 2.33μm.

5.5 Experimental Results for the Metasurface Beamsteerer

5.5.1 CASI Measurement.

Shown in Figures 49(a-d) are the BTDF results for (a) the raw, polarized measurement of the unmetalized MPL cavity array (serving as a baseline), (b) the raw, polarized measurements of the MPL metasurface, (c) the normalized H/H and (d) the normalized V/V. The baseline (a) shows a strong 0-order across all polarization states, with H/H (blue) and V/V (green) having the same amplitudes, as do H/V (red) and V/H (purple) at an intensity an order 10^{-3} that of the H/H. All of these

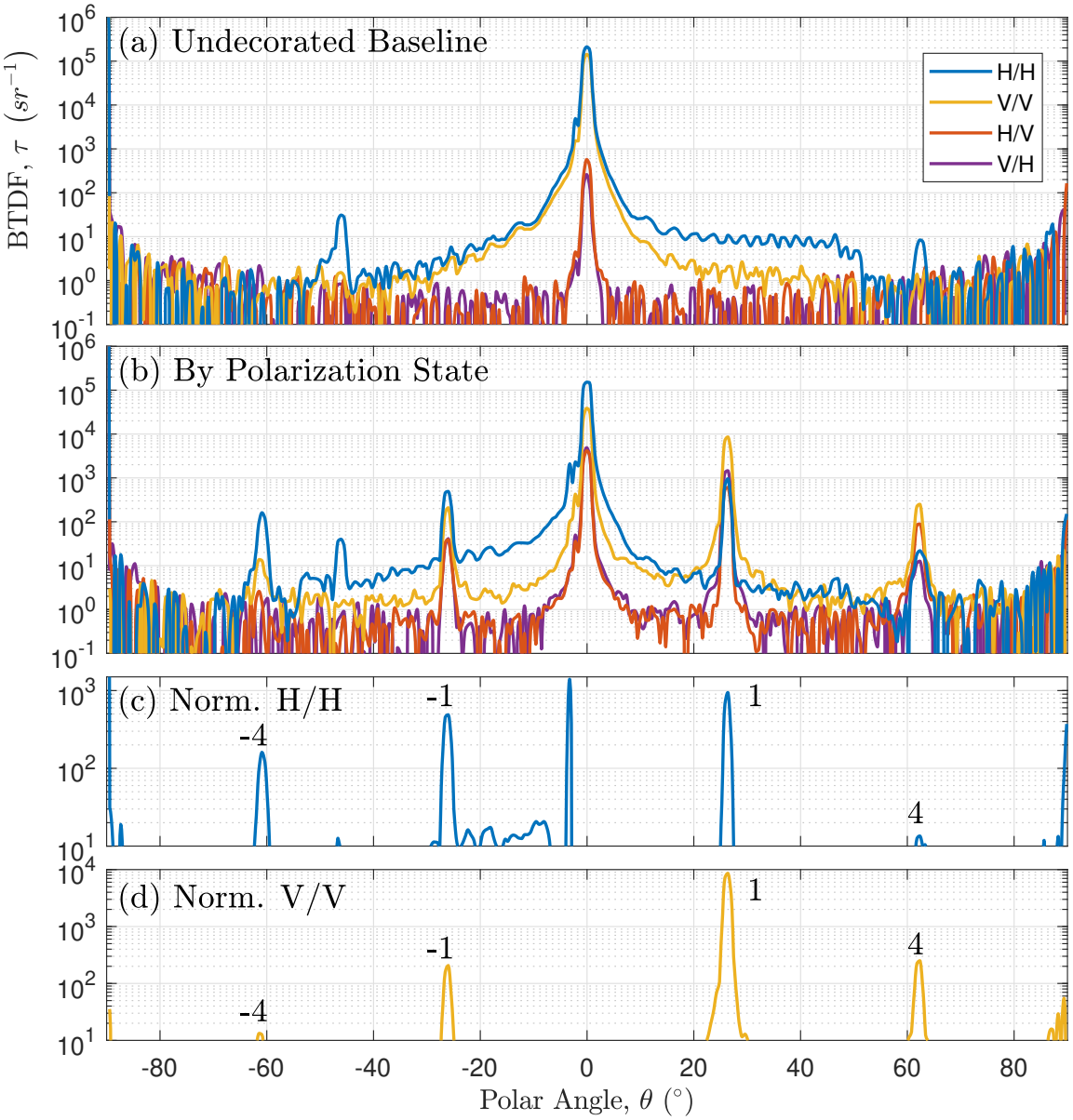

Figure 49. BTDF results for (a) the raw, polarized measurements of the undecorated MPL cavity array, (b) the raw, polarized measurements of the MPL metasurface, and (c) the normalized H/H and (d) normalized V/V polarization states. The metasurface beamsteering function is clearly present at the −1-order (correctly reversed from simulations), and larger than any mode other than the central peaks. The increase in intensity at the 0-order may be explained by internal reflections, which were not considered in the simulations.

results were expected, since the MPL periodicity was designed for operation below the diffraction edge, and there was no cross-polarization detected in any simulations.

However, looking at the device measurements (b), the 0-order remains strong, instead of the expected result of a dominant +1-order. Nevertheless, each polarization state has salient higher-order diffraction features at the $\mp$1-orders at $\theta_t = -26, 26.4°$, and $\mp$4-orders at $\theta_t = -60.9, 62.3°$, which are all within 2.1% of both the theoretical and as-fabbed simulated angles.

Since the responses of the simulated devices also contain the contribution from the cavity, a comparative magnitude of the 0-order beam should also appear in the computed intensity profile (Figure 45(b)). Indeed, it does appear, though it is only at 0.05 relative intensity to the steered beam—a far cry from the observed difference. So although the metasurface is functioning as designed at the proper orders, there is a weak coupling on the order of 10^{-2} which indicates there is a discrepancy between the device and the COMSOL models.

Two other anomalies exist in the CASI measurement. The first is a strong spike along the peak near -3°, which is unaccounted for in the simulations and is likely due to slight misalignment of the sample. The second is the presence of considerable H/V and V/H responses, which, according to Figures 33(f) and 34(f), the brick and hex elements should exhibit no cross-polarization. This indicates some optical activity not captured by the simulated periodic arrays.

5.5.2 Potential Contributions to Experimental Error.

Several aspects may be contributing to this large 0-order device behavior. The first consideration was that the aperture in the sample holder was misaligned with the device, allowing some portion of the surrounding unmetalized MPL structure to be illuminated. Inferring from what is seen in the baseline, this could provide a scan as

what is seen in the device measurement, with some energy coupling into the device, but some into the baseline.

A second consideration is the existence Fabry-Perot modes generated from backscatter into the substrate. The simulated device could not account for backscatter or the resultant multiple internal reflections from the $\sim 660\mu$m wafer, and the sample did not include an anti-reflection coating; with less-than-optimized as-fabbed designs, these effects could be quite significant to the overall function of the beamsteerer. After traversing the wafer and back, the F-P modes would be at a difference reference phase at the MPL interface; subsequently, these reflected, uniform plane waves will interact with the linear phase gradient. If most of the energy is channeled into the +1-order, as intended, these re-reflected waves would be channeled similarly into +1 upon interacting with the metasurface interface, but interfering destructively. Now functioning at a diminished efficiency, this may explain why the $\pm$1-orders are not easily distinguishable, as seen in the normalized H/H in Figure 49(c).

A third possible explanation for the overall beamsteering performance is the device design and realization itself. A four element beamsteerer left little room for error—a poor-performing element can make a large contribution to error, whereas a high discretization can compensate somewhat for these uncertainties. As mentioned, the brick MPL elements submitted for fabrication were significantly off the optimized phases and amplitudes, due to reduced voxel thickness and cavity curvature, and so this may be a considerable contribution to error. With the lack of a dominant higher-order mode, the device is behaving less like a blazed grating and more like a simple grating—this would be much more uniformly in phase than Figure 37(b) predicted. The as-fabbed brick element #5 ($\Delta\Phi = 90°$) has a phase retardation more like element #4 ($\Delta\Phi = 45°$), and element #7 ($\Delta\Phi = 180°$) more like element #6 ($\Delta\Phi = 135°$). Additionally, element #1 was so much less efficient than the

others, to the point where it may not even contribute effectively to the beamsteering in the realized device. Taking this element out of the simulated device, and using the ideal, optimized models to construct a 3-element device of a non-uniform phase gradient $\Delta\Phi = \{0, 45, 135\}^\circ$, then the diffraction efficiencies shift towards the 0-order and non-beamsteering -1-order modes, shown in Appendix B. Despite some energy redistributed away from the $+1$-order, it remains the strongest mode; nevertheless, this behavior is trending towards the observed result, indicating improper phasing from some elements might be contributing to the beamsteering function.

A fourth possibility is that the COMSOL models are inaccurate for the selected beamsteering design. They were simulated as having infinite periodicity with the beamsteerer only being considered infinite in the translated direction and with the direction along the phase gradient being a supercell of four unique, paired elements. This should affect the phase of any individual element due to mutual coupling of nearby neighbors of different phase and amplitude, and this coupling is not captured in the simulated device construction.

One last explanation for the unpredictable response might be due to a strong spectral sensitivity. While the spectral response of each element of the $t = 150$nm brick design was previously examined with respect to phase and amplitude (Figure 35), this was not revisited after the reduction to $t = 100$nm. Initial characterization work on the V-antenna lenses discovered significant fluctuations in the performance over extremely fine $\Delta\lambda$, with a drop in intensity of nearly 75% over just a 25nm spectral shift from λ_0[83].

While the MPL design permits up to five populated grids, for purposes of demonstration only a singular OOP grid was patterned on a wall aligned with the plane of incidence. It happens to be that in this particular configuration—an only this configuration—any incident electric field perpendicular to the grid cannot couple

strongly to an electric mode. Nor is the orientation of the magnetic field—which is parallel to the grid—sufficient for generation of a magnetic mode[1]. This explains why a low τ is expected for the V/V and V/H states; despite this claim, there is a significant V/V response: in fact, the normalized V/V in Figure 49(d) represents an ideal diffraction profile, with a strong +1-order and all others suppressed.

5.6 Progress Towards Hex Fabrication

To this point, the hex designs have been largely unused, save for the simulation of a full-scale lens in Figure 46; yet, the hex designs are the more promising building blocks for high-efficiency metasurface optics, for two reasons. First, from a design standpoint, they have more flexibility in forming magnetic modes due to the natural connectivity of the regular hexagonal tiling. Second, from a performance standpoint, the hex optimizations were run *after* SNL reported the $t = 100$nm fabrication constraint on the voxel thickness, and so there was no need to modify the optimized designs, as was done on the brick designs. This modification forced some of the brick elements to perform well below the optimized metrics (at $t = 150$nm), as shown in Figure 37, with significantly reduced transmittance, increased reflectance and non-linearly skewed phase. Ultimately, the effects on device functionality were demonstrated in the last section, with the prevalence of higher-order modes and reduced intensity in the co-polarized primary beam. Without this severe perturbation to t, the hex designs only needed to account for the curvature in the cavity, which was also seen in Figure 37 to maintain good performance overall.

The primary reason for this lack of attention is due to the difficulties encountered in the fabrication process. Writing a membrane mask using EBL may take additional

[1]The other configurations may engender very complex coupled electromagnetic modal behaviors, but these are not studied here. See [63] for an indication of expected modes from back-to-back sandwiched grids

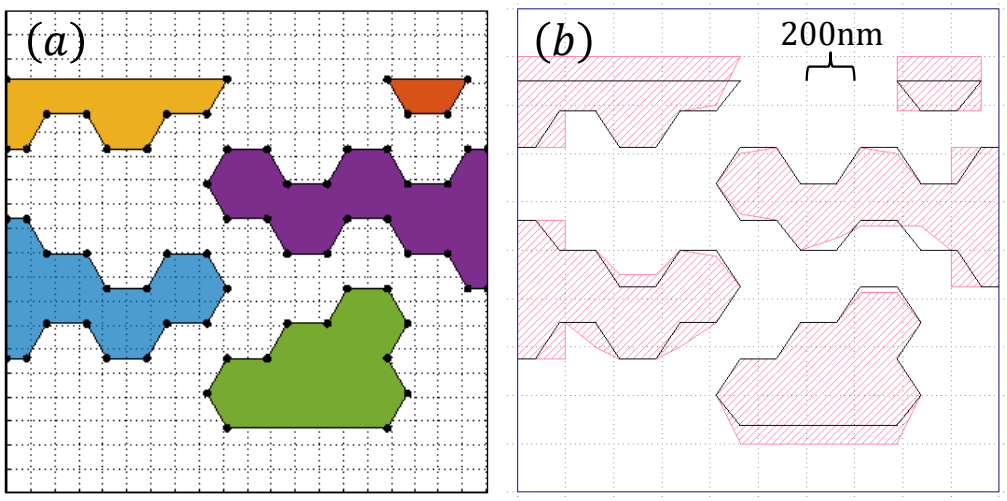

Figure 50. (a) Original hex design for element #8 ($\Phi_t = 165°$) and (b) its morphed topology, indicated by the thatched region. Due to use of a stepper for fabrication, rather than e-beam writing of the membrane, non-uniform biasing was required. Biasing was performed in both upward and downward—between $75 - 100$nm, as opposed to only biasing 100nm downward as required for the EBL membranes of the brick elements (Figs. 38-43).

time than the brick designs, due to the additional points in the hex voxel compared to the brick voxel. So, to eliminate the risk of rejection by SNL's Silicon Fab personnel for requesting excessive e-beam time, it was decided that the initial runs on the hex elements were to be accomplished using deep-UV (DUV) lithography on a step-and-scan camera, or "stepper" [87]. This is a relatively fast lithographic process, but reduces the maximum resolution to about 185nm; this introduces a major concern in fabricating the hex voxels, which have a minimum geometric feature of $166\overline{6}$nm. To better control the error, it was decided to manually morph the topology as needed. This was not to be a uniform alteration, as was the membrane biasing outlined in Chapters 3.4.2 and 5.2.1 for the brick elements; rather, the geometry of each sub-element was addressed individually for necessary deformation. An example of this shown in Figure 50 for (a) the original topology for hex element #8 ($\Phi_t = 165°$) and (b) the morphed topology, indicated by the thatched region. Notice that biasing was performed in both upward and downward—between $75 - 100$nm, as opposed to only

biasing 100nm downward as required for the EBL membrane.

To assess the impact of altering the hex topology, several models were adapted and simulated, with the solved H_x near-fields and associated performance metrics shown in Figures 51(a-d) for (a) the original model, (b) a model where most of the exterior curves (wherever the geometry forms convex edges) have been smoothed, (c) a model with each hex voxel replaced by a circle with the same diameter as the major hex axis, and (d) the same as (c), but for a square voxel. Comparing (b-d) to the optimized results of $\{\Phi = 166.2°, T = 0.658, R = 0.002\}$, both the exterior and circle models are extremely immune to the changes, boasting fractional differences of $\{0.02, 0.00, 1.5\}$ and $\{0.05, 0.01, 9.0\}$, respectively; however, the square model did not fare well, suffering a fractional difference of $\{0.254, 0.272, 118.0\}$. Absolute values of these metrics are included as insets.

The current densities (green) give a hint as to why (b, c) are so robust. Consider a dipole if a finite width bent into a curve (assuming a malleable metal), illuminated by a uniform field polarized in the plane of the dipole. The induced current density would be larger along the interior curvature due to the principle of least action—in particular, current flows such that the rate of heat loss is a minimum, which means since current following the longer exterior curve incur more Ohmic loss, current is distributed more towards the shorter interior curve[88]. The topology changes in (b,c) leave the interior curves largely intact, and so the current densities are generally maintained; conversely, the changes to the exterior curves are often severe—especially in (b)—which has little impact. On the other hand, the changes to the interior curves in the square geometry generates currents that are of a stronger amplitude in the lower half of the geometry, but of a lesser amplitude in the upper half, ultimately resulting in a vastly altered phase profile.

As of the completion of this dissertation, the fabrication of the hex elements had

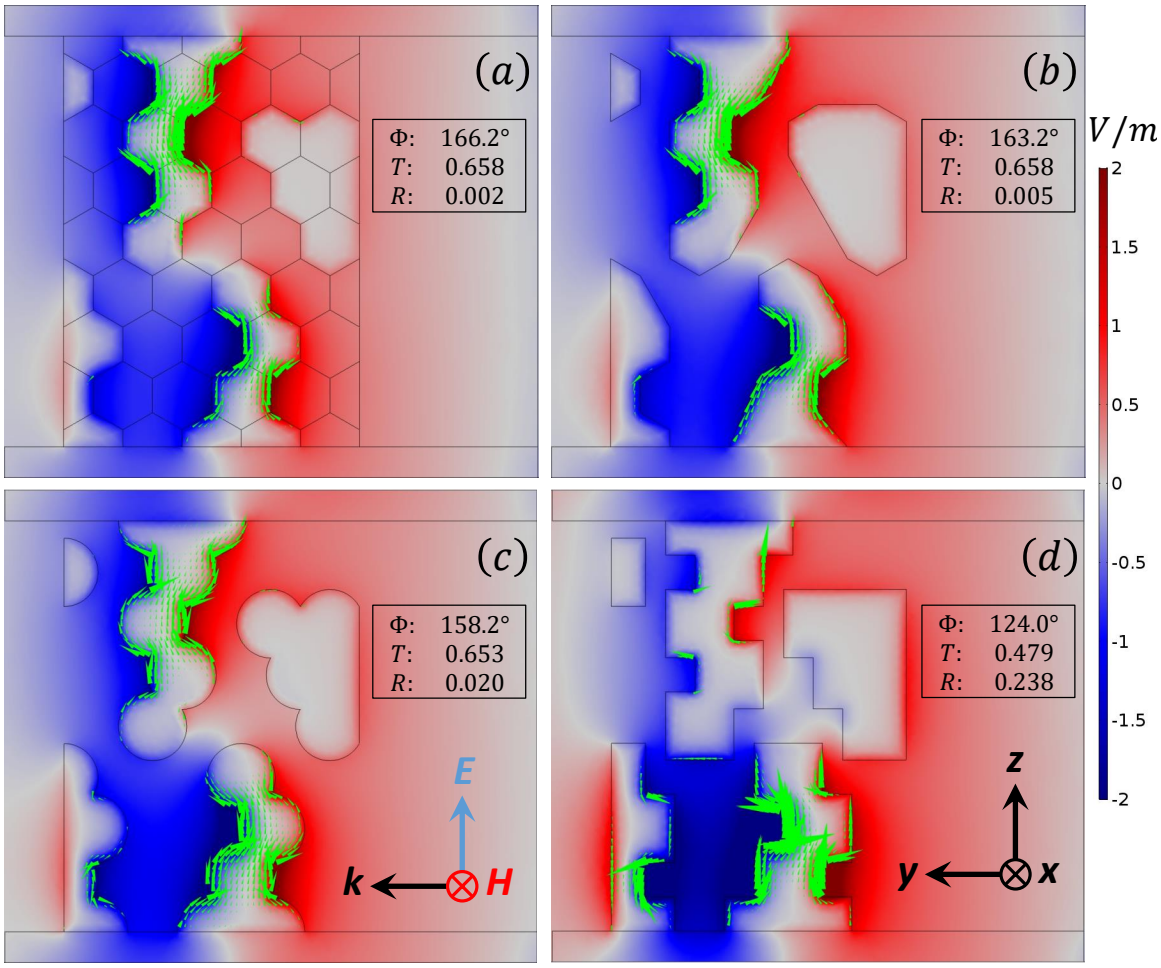

Figure 51. To assess the impact of altering the hex topology, several models were adapted and simulated, with the solved H_x near-fields and associated performance metrics shown for (a) the original model, (b) a model where most of the exterior curves (wherever the geometry forms convex edges) have been smoothed , (c) a model with each hex voxel replaced by a circle with the same diameter as the major hex axis, and (d) the same as (c), but for a square voxel. Designs in (b,c) are insensitive to these changes since they did not significantly alter the geometry on the interior curves, preserving current density (green); conversely, the square geometry (d) suffered due to large changes in the interior curvature.

not yet begun, but were in queue at SNL's Silicon Fab foundry. Plans are to examine each element as was done for the brick elements, then build an 8-element lens over one or two cycles. If the stepper fabrication approach produces too rough a resolution of the hex elements—to the point where interior curves become significantly morphed—then a switch to the EBL membrane write will be accomplished.

5.7 Summary

The focus of this chapter was to build and measure full-scale metasurface devices using the MPL designs, in order to validate that the GA optimization routine could produce highly-efficient Huygens metasurface building blocks. The fabricated designs were not exact replicas of the optimized designs, due to some changes in the voxel thickness and the lack of true "squareness" of the cavity. Nevertheless, these "as-fabbed" models were simulated in Chapter 4.4.2 and used to compute the far-field intensities of full-scale devices. These devices were shown to produce the desired optical effect—beamsteering, using the brick elements; and lensing, using the hex elements—and that the optical function did not vary significantly from the ideal, optimized models.

After some initial fabrication runs were performed on the MPL metasurface elements, a full-scale 4-element beamsteerer was built at SNL, and measured at AFIT. While it showed some anomalous behavior near the 0-order transmission, the beamsteering angle was in good agreement with simulation, and the intensity ratio was an order of magnitude higher in the beamsteering order (+1, from the perspective of the sample) than in any other higher order. However, most energy remained in the 0-order, meaning the incident field was not coupling strongly into the metasurface scatterers.

The hex elements were to be fabricated using a lower-resolution stepper process, and to account for this change the topology was biased non-uniformly. The effects on morphing the topology were assessed, and it was found that as long as the integrity of original interior curvature was maintained, there was little effect on the overall response. A lens design based on the hex elements was submitted to SNL, but was not completed in time for this work. Based on these results, the implications of the overall research effort, and recommendations for potential future work is discussed in

the next, and final, chapter.

VI. Discussion of Results, Impacts and Recommendations

6.1 Introduction

With all tasks accomplished in the research project, this chapter will conclude the dissertation. It begins with a summary of all notable aspects encountered in the development of these plasmonic 3D Huygens metasurfaces for highly-efficient flat optics, to include the problem statement, methodology, and results. Following this is a discussion on the research impact, highlighting the contributions this work has provided to the greater academic community. As this work was treading very fertile ground, there is still a great deal of work to accomplish; therefore, a final proposal is offered on potential work to be completed by future researchers and suggestions on improvements which might be employed.

6.2 Executive Summary

6.2.1 Problem Overview and Methodology.

While many metasurface architectures have been proposed over the past six years, few plasmonic solutions have demonstrated any feasibility to replace conventional optical materials. One of the most promising solutions have been architectures which possess some out-of-plane (OOP) feature, such that electric and magnetic modes can be supported; these offer high efficiency, due to a Huygens-like forward scattering behavior. But phase control and fabrication have remained a challenge beyond the RF regime—as a result, none have been demonstrated in the infrared (IR) or higher frequencies. This research was focused on developing an efficient plasmonic metasurface architecture in the IR by applying an evolutionary approach to the optimization of an OOP scatterer embedded in a 3D unit cell.

Rather than begin this investigation with fundamental forms which are known to produce the necessary electric and magnetic modes—such as dipoles or loops—it was a key imperative to start with a blank slate and let the nature of statistics formulate an answer to the question "which scatterer topology is best?" By breaking down a region of space into a grid of binary voxels which is in-plane with the incident wave, and out of plane to a planar interface, then an OOP scatter can be assembled by switching these voxels to either metal (bit "1") or dielectric (bit "0"). (Figure 6) But how to know which bits to flip? A genetic algorithm was ideally suited for such a task; however, these are typically used for analytical problems with massive design (i.e. parameter) spaces. These scatterers may possess any number of electric and magnetic modes for efficient Huygens-like behavior with full phase control—the starting geometries did not matter, as long as it produced the desired result. Therefore, a computational finite-element method was employed, using COMSOL Multiphysics, but there was no means to implement the binary GA routine in this software. Therefore, a great bulk of this work was in creating a method for enabling COMSOL to interact with a MATLAB-based GA in order to carry out the optimization of a non-analytical model (Figure 11).

After the COMSOL-based GA routine was verified against a simple, free-standing set of back-to-back grids (Figure 10), it was time to incorporate the grid in a realizable architecture from which large-scale metasurface samples could be fabricated. The ideal host for the grids was found from a Sandia National Laboratories (SNL) development called "membrane projection lithography" (MPL), which consists of an empty cavity that can have scatterers of any size and shape deposited onto its interior faces. Using this structure as a basis, a complex model was generated that represented the ideal MPL architecture. In this, any number of faces of a rectangular cavity could be populated with grids; however, for a simple demonstration of the concept, only a

single OOP face was investigated (Figure 32).

6.2.2 Results and Discussion.

There were many tasks that needed to be accomplished before attempting to build a real metasurface device based on these MPL elements. The first was to demonstrate COMSOL was capable of accurately targeting phase and amplitude to a sensitivity required for design of a metasurface lens, and that AFIT possessed the resources for building and measuring the device. Using the V-antenna planar architecture as a basis, 19 lenses and 8 metasurfaces were designed and fabricated at SNL according to COMSOL simulated profiles. These were successfully validated at AFIT by sampling the depth-of-field around the designed focal point, and found to be in excellent agreement with theory (Figure 9).

Next was to show that an OOP was even capable of forming the proper geometries needed for efficient forward scattering—and that the GA routine could converge a large parameter space to an optimal layout forming these geometries. Figures 14-17 give clear evidence that both of these tasks were successful, optimizing a symmetrical space of size 4×10^9 into a layout which produced forward scattering 55dB higher than the backward scatter, and when compared to a nominal V-antenna gave a forward intensity 440 times larger (see Chapter 3.3.5).

However, this was not a metasurface, as no phase control had been demonstrated. With the knowledge that an OOP grid could be optimized for efficient scattering, the last step before fabrication was to validate that the grid could be modulated by the GA routine to retard the phase sufficiently for optical function—all while maintaining the Huygens-like response. The initial test proved the largest shift of $\Delta\Phi = 180^\circ$ was possible while achieving a transmittance of $T = 0.63$ and reflectance of $R = 0.02$ in the total field (Figure 25). To give an indication of how efficient this would be in

a more ideal, lossless structure, the normalization to the baseline transmittance of the unmetalized MPL cavities shows a transmittance of $T_n = 0.84$—this is within 2% of the highest-efficiency Huygens metasurface yet reported, by Pfeiffer and Grbic (2013), but this was in the RF regime with a near-perfect conductor. To justify *why* this embedded grid architecture was able to perform so well, the near-fields were examined to show that many localized electric and magnetic orthogonal modes were being generated with near-equal strengths, providing the impetus for Huygens-like behavior (Figures 29-30).

To demonstrate the flexibility of the MPL design architecture and repeatability of the GA routine to accurately target *any* point in the 2π phase space, series of both brick and hex elements were optimized inside a deeper, rectangular MPL cavity. These boasted mean T/R metrics of $\{0.65/0.10\}$ and $\{0.66/0.03\}$, repectively, while being within 2° of accuracy of all phase targets Φ_t(Figures 33-34). This result clearly establishes that the developed binary GA routine can produce plasmonic OOP scatterers capable of high-efficiency phase control, using an architectural foundation that can be realized in the infrared regime.

While the fabrication test runs showed very fine resolution of the voxel geometries (Figures 39-41), there was a major concern that this degradation would not be a great representation of the massive potential of the GA-based OOP metasurface concept—or worse yet, not even manifest the optical function. Nevertheless, the brick elements were still used to form a full-scale (4mm^2), but coarse, 4-element beamsteering metasurface (Figure 43). A far-field projection allowed an examination of how this as-fabbed device (at voxel thickness $t = 100\text{nm}$) would perform compared to the ideal one (at $t = 150\text{nm}$), and both successfully transmitted a beam with over 90% of the intensity into the +1-order (Figure 45), with only small fractions in the undesired $\{-4, 0, -1, 4\}$ orders and a negligible cross-polarized response.

Despite the design inaccuracies and associated losses, the diffraction orders predicted by the simulated device are visible, with the higher orders displaying the proper relative intensities (Figures 45(c-d)), as per bi-directional transmittance distribution function (BTDF) measurements from AFIT's Complete Angle Scatterer Instrument (CASI). However, there is an unexpectedly strong 0-order peak, and several reasons for this behavior were postulated, to include sample misalignment, extemporaneous Fabry-Perot modes from internal reflections, fabrication error, poorly represented brick MPL designs, poorly represented COMSOL models, and spectral sensitivity.

6.3 Research Impact

From the introduction, the main research objectives were stated to (1) design and simulate realizable GA-optimized Huygens metasurface elements to serve as the basic building blocks for full-scale flat optics, and then (2) build and experimentally validate that these devices are operating efficiently, as designed. What defined success in (1) was the ability to hit 50% transmission efficiency for all Φ_t, spanning the entirety of 2π phase space, and what defined "validated" in (2) was having a measurement in good agreement with the simulated devices constructed from the designs generated in (1). Each of these objectives, and their contributions, is described in detail.

6.3.1 Robust 3D Metasurface Design and Optimization Tool.

Addressing the first objective, an autonomous metasurface design and optimization tool was developed that was proven to accurately generate a Huygens-like scatterer transmitting at *any* chosen phase delay from an entirely generic 3D unit cell. The routine repeatedly hit 60% transmittance across the phase space—exceeding the objective—and it is important to note that 25% of this loss is attributed to the baseline cavity, indicating this efficiency can be much higher if the supporting structure

possesses low insertion loss. This was shown possible with the MPL at certain periodicities and wall thicknesses; though the grid resolution would be quite low and impractical.

The routine takes a large set of GA parameters (voxel shape, number of grids, population size, number of iterations, cavity dimension, incident wavelength, function weights, etc.) input by the user at the initialization, and outputs a solution space based on the dimensionality of the cost function. One of the most important features of the tool is the massive diversity this offers to an optical engineer: diversity in design, diversity in desired result, and diversity in application. This idea is expounded in the proceeding paragraphs.

While only a few simple voxel shapes and cavity geometries were employed to address a 3D electromagnetics scattering problem, the design diversity of the optimization tool is unprecedented; this statement is further elucidated by the choice of the example MPL architecture as a basis. Though a single OOP grid was presented in detail, any number of available faces could have been activated to attain a more complex response, where several orthogonal grids could be made to interact with high near-field coupling. Actually, a great deal of previous metasurface works could be replicated using this tool: a planar metasurface could be investigated by only setting a grid on the floor and removing the upper walls; while a pseudo-3D stacked metasurface could be investigated by adding a second planar grid on the top of the cavity. Even more, the cavity and voxel shapes have no hard constraints, and could have easily be mated to the more exotic MPL architectures (Figures 18 and 21).

With regards to diversity in desired result, it is vital to note that although a simple two-objective cost function was proposed here, any number of results which are extractable from the computational solver can be optimized. One obvious addition for this work was proposed in Eq. 15, looking for reduced sidelobes; another would be to

add a term for minimizing R (it was accomplished manually, for all presented results). While this diverse feature is an intrinsic property of an evolutionary optimizer in general, it was made especially effective in this effort by incorporating a means to address models in which the user may have no prior intuition on how the model will behave. This was demonstrated in the OOP binary grid, where it was unknown (1) whether the grid could even produce Huygens-like scattering given the preset voxel resolution, or (2) whether the grid layout could even fulfill phase control requirements at these high transmittances.

Considering diversity in application: this tool is not limited to electromagnetics problems. The highly symbiotic optimization routine is especially suited to handle non-analytical solutions since it is mated to both a commercial computational language and a commercial FEM solver able to handle multiphysics environments—which makes this tool even more accessible to the greater computational electromagnetics (CEM) community. More to the point, this product fills a major deficiency to 10000+ COMSOL users worldwide, as they only provide local optimizers (gradient-based, random search); there is no global optimizer available in the software. With this program, unfathomably large parameter spaces can be accessed ($\sim 10^{24}$ demonstrated) and compressed into a small subset of near-global and non-analytical solutions after only a meager fraction of attempts ($\sim 10^{-21}$ demonstrated). This allows a designer to either solve a complex model more rapidly, or add complexity to a model that was simplified for lack of resources. In fact, the design tool was considered such a useful contribution that the author was invited by COMSOL and Laser Focus World to give an instructional webinar on the implementation of the GA optimization routine on 7 June 2018. A generalized code was provided to their users via their web-based Application Exchange so they may tailor to their particular CEM problem [89].

6.3.2 Fabrication and Measurement of 3D Metasurface.

Addressing the second objective, a 4mm^2 beamsteering device was fabricated based on the output from the 3D metasurface design and optimization tool, which, to the knowledge of SNL personnel, was the amongst the most complex and highest-fidelity metasurface structure ever produced in this regime. One of the biggest questions of this research effort was whether fabrication of these elements would be possible—could the grid layouts could be faithfully reproduced using the MPL process, such that there could be high confidence in the simulated results? The device SEM image (Figure 43) makes a strong statement in favor of that sentiment.

However, as mentioned, two reactionary measures were enacted to get to that point, where t was reduced to 100nm and the discretization was set at a minimal $N = 4$ elements. Consequently, it should be understood that the beamsteer does not embody an ideal product of the design and optimization tool. Furthermore, there were experimental uncertainties which did not give high confidence in the BTDF. Indeed, it was determined in simulation that several of the brick elements performed quite poorly under this modification, dropping transmittance to as low as 15% and increasing reflectance as high as 50%. Therefore, to claim this metasurface design was fully validated would be inaccurate; though, it is justifiable to say a partial validation was accomplished. What is unmistakable from comparing the total (Figure 49(a)) and baseline (Figure 49(b)) BTDF is that the presence of the inclusions solely contributes to these higher-order diffraction lobes, and that they are of the proper $\{-4, -1, +1, +4\}$ order to within $\sim 2\%$ of the simulated angle; to this degree, it can be said the metasurface is performing an optical beamsteering function, albeit at an extremely poor efficiency.

Lack of full validation notwithstanding, this voxelated OOP grid design was considered so revolutionary that it is currently used by SNL as a basis for their DARPA

Extreme Optics and Imaging (EXTREME) program, which has a goal to "develop new optical components, devices, systems, architectures and design tools using Engineered Optical Materials (EnMats)" with an aim "to understand the trade-offs, and harness the possibilities, afforded by EnMats"[90]. A considerable deal of work has been performed in conjunction with SNL on the design/fab trade-offs and how it affects metasurface functionality: the impact of design variation on the unit cell performance was investigated in several ways, looking at changes to voxel thickness and cavity shape (Figures 35, 37), corner-to-corner contacts (Figure 19), corrections to voxel voids (Figure 31) and alterations to topology (Figure 51). Given the stated objectives, a meaningful contribution has been made to advancing DARPA's goals.

6.4 Recommendations for Further Investigations

Ultimately, the objective of this research can be stated succinctly as this: build a better plasmonic metasurface. The means to do this was carried out through the development of an optimized, efficient Huygens-like scatterer. While greatly successful in simulation, one of the biggest deficiencies in this research effort is that a true representative high-efficiency Huygens metasurface was not produced, and thus could not fully validate the GA-based design and optimization tool.

Therefore, the primary recommendation for further development is to generate more archetypal designs and devices for validation. This might start with optimizing the "as-fabbed" model, to include cavity curvature and reduced voxel thickness (for brick elements). The hex designs appear a much more promising basis, permitting better geometric curvature in the scatterer, but the use of the stepper in fabrication is expected to significantly weaken the performance potential; perhaps reverting to the EBL would produce a better specimen. One change which can make all of this easier is to move to a slightly longer wavelength; current capabilities may prove incapable

resolving the $1/4\lambda_0$ cavity and $1/24\lambda_0$ voxels to an accuracy sufficient to drive proper in- or out-coupling. Adding an anti-reflection coating (approximately 585nm of silicon oxide (SiO)) may also address some of the witnessed anomalies.

A secondary recommendation is to continue investigating the sources of error outlined in Chapter 5.5.2. Several experimental and design actions can be implemented to address these considerations: to assess whether the sample is misaligned, the current sample holder (Figure 48) can be removed and the sample can be stopped-down by aluminum tape to be sure the device is framed properly, precluding any illumination of the unmetalized portion; an anti-reflection coating of silicon oxide (SiO) can be sputtered on the double-polished backing of the Si wafer to assess contributions of the internal reflections; and finally, a straight-through transmission measurement can be accomplished using an FTIR to scan the spectrum, where any drop in the 0-order transmission might indicate the wavelength that energy is maximally distributed into the beamsteering angle. Knowing a potential wavelength of interest, the BTDF measurement can be accomplished in a small range around this value, and over the entire angular space. Additional to the potential experimental investigations, the COMSOL MPL model can be reassessed against the specific beamsteering device characteristics. The models can be simulated in only a partially-periodic environment, examining the effects of a lack of periodicity in the elements along the phase gradient. Nearest-neighbor coupling can be investigated by modeling pairs, or even a larger subset of the supercell—though, this may require access to high-performance, clustered computational resources.

A tertiary recommendation deals with the massive underbelly of 3D design not addressed in this work. Here, the final MPL framework was a single OOP grid in a rectangular cavity; yet, this example only skims the surface of a deep well of potential 3D architectures which can be explored using the GA-based design and

optimization tool in conjunction with the MPL fabrication phenomenology. Many exotic designs could be investigated without the constraint of reducing the model to some analytical form, and might provide insight to novel forms of electromagnetic control. For example, can we better understand complex intercell coupling in a 3D unit cell between multiple orthogonal scatterers, and its relationship to phase control and far-field intensity? Such a consideration has only been touched upon lightly by the author, previously[91, 63], but is made much more exploitable with this tool. As the GA routine was built for minimal example, it is expected additional programming will be required to enable these expanded investigations.

One last recommendation is to pursue an alternative means of validation for the GA routine through another CEM method, such as a finite-difference time domain (FDTD) solver. If this method were to attain T/Φ results for each element comparative to the COMSOL results, it might provide a secondary means of validating the design and optimization tool that the measurements could not provide, eliminating the any consideration that the model itself is the source of error in the realized device. Additionally, while having less robustness in model meshing, FDTD has the benefit of attaining broad-spectrum results more efficiently than an FEM solver, which must solve for each wavelength individually. Some commercial FDTD solvers, such as Lumerical, offer a similar symbiotic interfacing with MATLAB as was implemented with COMSOL, and can be used to mimic the design and optimization tool developed in this work[92].

6.5 Closure

Now is an exciting time to be on the forefront of evolutionary optical design. Due to heavy global investment in electromagnetically-engineered optical materials and processes, both computational tools and fabrication methods have advanced rapidly

and in tandem over the past two decades, in an attempt to keep pace with the emerging theories of subwavelength light-matter interactions. Never before has such great firepower been at our fingertips to design, analyze and realize these structures that bridge the gap between the atomic and the macroscopic. This research effort attempted to harness that power to push the technology edge a little further, progressing the ideas that fully-3D unit cell designs can be building blocks for practical metasurface architectures, and that lossy, plasmonic scatters have a utility in high-efficiency optics. While challenges remain in the full validation of these concepts, the design, simulation and optimization contributions of this work established that plasmonic 3D Huygens metasurfaces offer significant potential for flat optics applications.

Appendices

DESIGN AND OPTIMIZATION OF PLASMONIC 3-D HUYGENS METASURFACE BUILDING BLOCKS FOR HIGHLY-EFFICIENT PLASMONIC FLAT OPTICS

A. Appendix

1.1 SEM Images of Fabricated MPL Brick Elements

After several test fabrication runs were accomplished on a square layout (see Figs. 39 and 40), the GA-optimized MPL brick elements were fabricated to examine how well the fine 111nm features resolved. In Chapter 5.2, a single example of these elements was presented: the design which was targeted for a phase $\Phi_t = 120°$ which shifts in phase $\Delta\Phi = 180°$ (see Fig. 41). In this appendix, the SEM images of all eight finished brick elements are presented. The $\Phi_t = 150°$ element was accidentally metalized in the wrong direction, so the design was patterned upside down. However, this was corrected in the metasurface beamsteerer, as can be seen in Fig. 43.

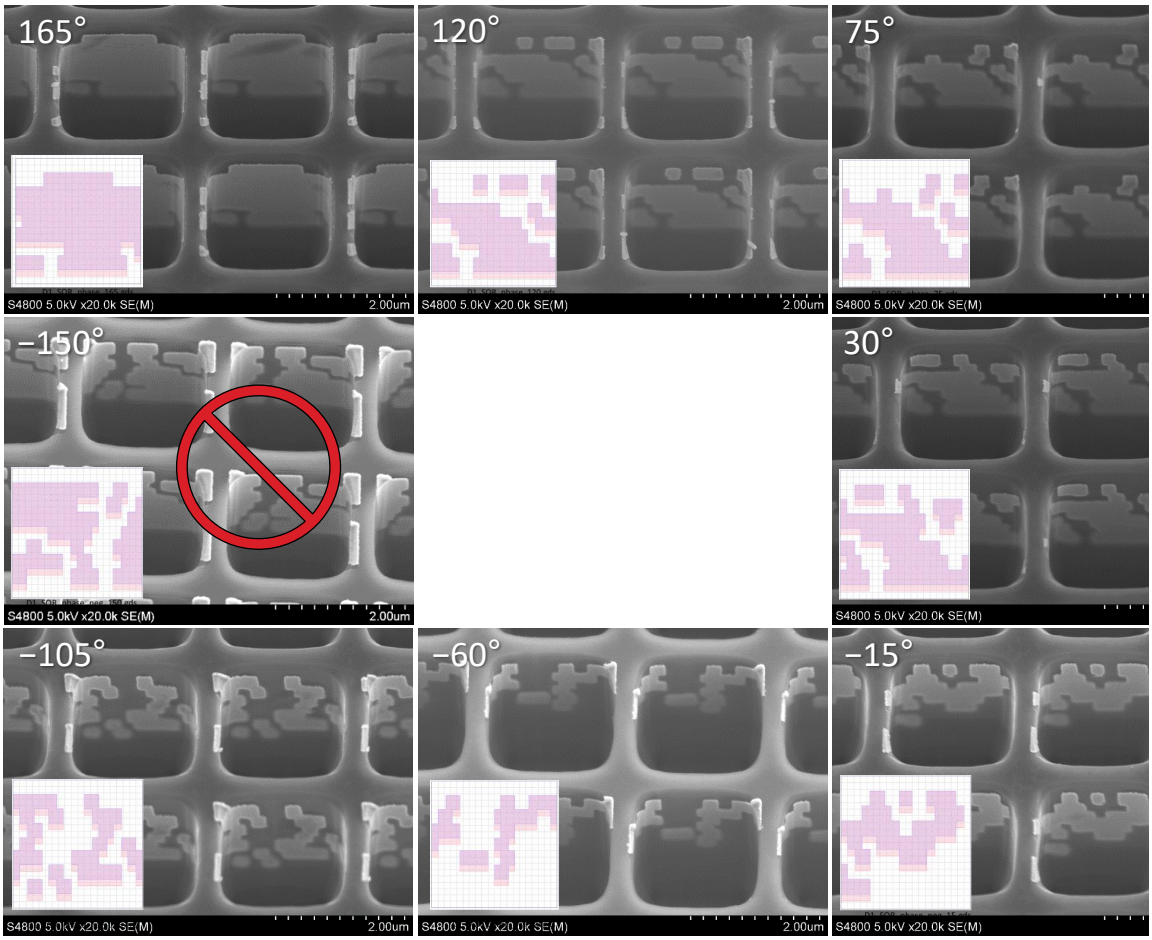

Figure 52. SEM images of all eight fabricated MPL brick elements. The $\Phi_t = 150°$ element was accidentally patterned upside down, but this was corrected in the full-scale device fabrication.

B. Appendix

2.1 Beamsteerers with Non-Uniform Phase Profile

To attempt to explain the undesired diffraction orders seen in the measured beamsteerer (Figure 49), a device with perturbations were calculated which exhibits a non-uniform phase profile, and is presented in Figure 53. This device (a) used only $N = 3$ elements, and they were of the ideal, optimized ($t = 150$nm) model in order to accurately craft the relative phase shifts of $\Delta\Phi = \{0, 45, 135\}^\circ$. For comparision, the as-fabbed ($t = 100$nm) device from Figure 45(b) is shown, and both images are scaled to the same range as in this reference. Despite a considerable amount of energy redistributed from the $+1$-order into the -1- and 0-orders, the steered beam remains the strongest, unlike what is seen in measurements; nevertheless, this result indicates improper phasing from some elements might be contributing to the beamsteering function as measured.

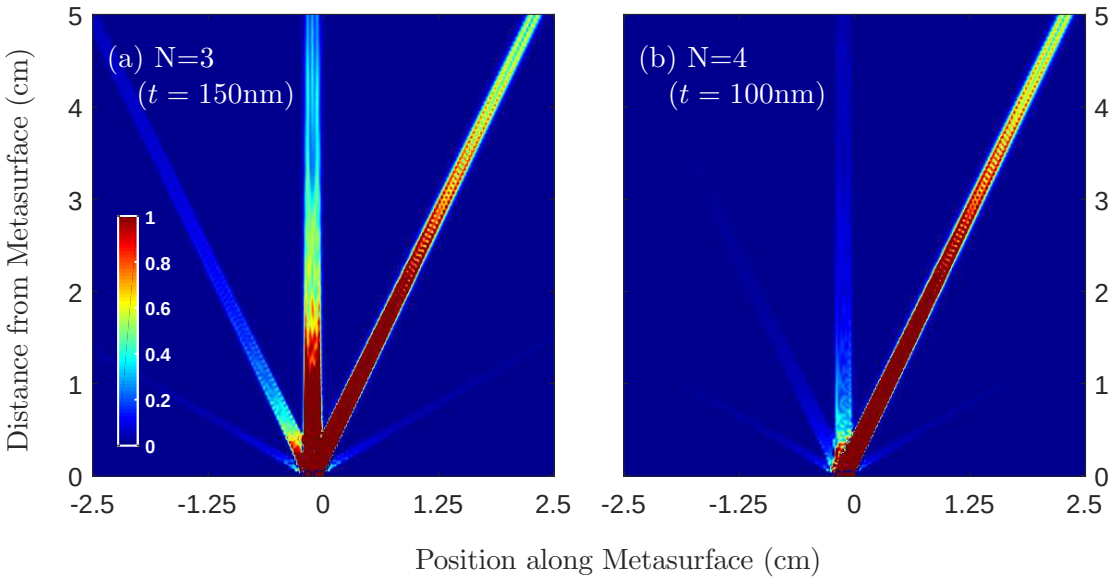

Figure 53. Calculated far-field intensity for (a) a 3-element metasurface beamsteerer made from ideal element, but with a non-linear phase gradient $\Delta\Phi = \{0, 45, 135\}^\circ$, and (b) the previous as-fabbed $N = 4$ device from Fig. 45(b), in order to consider what potential effects fabrication and design inaccuracies might have on experimental device. More energy is funneled into the 0- and $+1$-orders, yet the -1-order still remains the strongest.

Bibliography

1. Nanfang Yu, Patrice Genevet, Mikhail A. Kats, Francesco Aieta, Jean-Philippe Tetienne, Federico Capasso, and Zeno Gaburro. Light propagation with phase discontinuities: Generalized laws of reflection and refraction. *Science*, 334(6054):333–337, 2011.

2. Nanfang Yu and Federico Capasso. Flat optics with designer metasurfaces. *Nature Materials*, 13(2):139 – 150, 2014.

3. N. Yu, P. Genevet, F. Aieta, M. A. Kats, R. Blanchard, G. Aoust, J. P. Tetienne, Z. Gaburro, and F. Capasso. Flat optics: Controlling wavefronts with optical antenna metasurfaces. *IEEE Journal of Selected Topics in Quantum Electronics*, 19(3):4700423–4700423, May 2013.

4. Francesco Aieta, Patrice Genevet, Mikhail A. Kats, Nanfang Yu, Romain Blanchard, Zeno Gaburro, and Federico Capasso. Aberration-free ultrathin flat lenses and axicons at telecom wavelengths based on plasmonic metasurfaces. *Nano Letters*, 12(9):4932–4936, 2012. PMID: 22894542.

5. Francesco Aieta, Ali Kabiri, Patrice Genevet, Nanfang Yu, Mikhail A. Kats, Zeno Gaburro, and Federico Capasso. Reflection and refraction of light from metasurfaces with phase discontinuities. *Journal of Nanophotonics*, 6(1):063532–063532, 2012.

6. Francesco Aieta, Patrice Genevet, Mikhail Kats, and Federico Capasso. Aberrations of flat lenses and aplanatic metasurfaces. *Opt. Express*, 21(25):31530–31539, Dec 2013.

7. Carl Pfeiffer and Anthony Grbic. Metamaterial huygens' surfaces: Tailoring wave fronts with reflectionless sheets. *Phys. Rev. Lett.*, 110:197401, May 2013.

8. Romain Blanchard, Guillaume Aoust, Patrice Genevet, Nanfang Yu, Mikhail A. Kats, Zeno Gaburro, and Federico Capasso. Modeling nanoscale v-shaped antennas for the design of optical phased arrays. *Phys. Rev. B*, 85:155457, Apr 2012.

9. Wujiong Sun, Qiong He, Jiaming Hao, and Lei Zhou. A transparent metamaterial to manipulate electromagnetic wave polarizations. *Opt. Lett.*, 36(6):927–929, Mar 2011.

10. Xumin Ding, Francesco Monticone, Kuang Zhang, Lei Zhang, Dongliang Gao, Shah Nawaz Burokur, Andre de Lustrac, Qun Wu, Cheng-Wei Qiu, and Andrea Alù. Ultrathin pancharatnam–berry metasurface with maximal cross-polarization efficiency. *Advanced Materials*, 27(7):1195–1200, 2015.

11. Shulin Sun, Qiong He, Shiyi Xiao, Qin Xu, Xin Li, and Lei Zhou. Gradient-index meta-surfaces as a bridge linking propagating waves and surface waves. *Nature Materials*, 11(5):426 – 431, 2012.

12. J. Neu, B. Krolla, O. Paul, B. Reinhard, R. Beigang, and M. Rahm. Metamaterial-based gradient index lens with strong focusing in the thz frequency range. *Opt. Express*, 18(26):27748–27757, Dec 2010.

13. Oliver Paul, Benjamin Reinhard, Bernd Krolla, Ren Beigang, and Marco Rahm. Gradient index metamaterial based on slot elements. *Applied Physics Letters*, 96(24):241110, 2010.

14. Xueqian Zhang, Zhen Tian, Weisheng Yue, Jianqiang Gu, Shuang Zhang, Jiaguang Han, and Weili Zhang. Broadband terahertz wave deflection based on c-shape complex metamaterials with phase discontinuities. *Advanced Materials*, 25(33):4567–4572, 2013.

15. Xingjie Ni, Satoshi Ishii, Alexander V Kildishev, and Vladimir M Shalaev. Ultra-thin, planar, babinet-inverted plasmonic metalenses. *Light: Science & Applications*, 2(4):e72, 2013.

16. Anders Pors, Michael G. Nielsen, Ren Lynge Eriksen, and Sergey I. Bozhevolnyi. Broadband focusing flat mirrors based on plasmonic gradient metasurfaces. *Nano Letters*, 13(2):829–834, 2013. PMID: 23343380.

17. Xianzhong Chen, Lingling Huang, Holger Mühlenbernd, Guixin Li, Benfeng Bai, Qiaofeng Tan, Guofan Jin, Cheng-Wei Qiu, Shuang Zhang, and Thomas Zentgraf. Dual-polarity plasmonic metalens for visible light. *Nature communications*, 3:1198, 2012.

18. Lingling Huang, Xianzhong Chen, Holger Muhlenbernd, Guixin Li, Benfeng Bai, Qiaofeng Tan, Guofan Jin, Thomas Zentgraf, and Shuang Zhang. Dispersionless phase discontinuities for controlling light propagation. *Nano letters*, 12(11):5750–5755, 2012.

19. Haisheng Hou, Guangming Wang, Haipeng Li, Wenlong Guo, and Tangjing Li. Highly efficient multifunctional metasurface for high-gain lens antenna application. *Applied Physics A*, 123(7):460, Jun 2017.

20. Lei Zhang, Shengtao Mei, Kun Huang, and Cheng-Wei Qiu. Advances in full control of electromagnetic waves with metasurfaces. *Advanced Optical Materials*, 4(6):818–833, 2016.

21. Mohammadreza Khorasaninejad and Federico Capasso. Metalenses: Versatile multifunctional photonic components. *Science*, 358(6367):eaam8100, 2017.

22. Philippe Lalanne and Pierre Chavel. Metalenses at visible wavelengths: past, present, perspectives. *Laser & Photonics Reviews*, 11(3), 2017.

23. Liyi Hsu, Matthieu Dupré, Abdoulaye Ndao, Julius Yellowhair, and Boubacar Kanté. Local phase method for designing and optimizing metasurface devices. *Optics Express*, 25(21):24974–24982, 2017.

24. Aytekin Ozdemir, Zeki Hayran, Yuzuru Takashima, and Hamza Kurt. Polarization independent high transmission large numerical aperture laser beam focusing and deflection by dielectric huygens' metasurfaces. *OPTICS COMMUNICATIONS*, 401:46–53, 2017.

25. Francesco Monticone, Nasim Mohammadi Estakhri, and Andrea Alù. Full control of nanoscale optical transmission with a composite metascreen. *Phys. Rev. Lett.*, 110:203903, May 2013.

26. Amir Arbabi and Andrei Faraon. Fundamental limits of ultrathin metasurfaces. *arXiv preprint arXiv:1411.2537*, 2014.

27. VS Asadchy, Y Ra'Di, J Vehmas, and SA Tretyakov. Functional metamirrors using bianisotropic elements. *Physical review letters*, 114(9):095503, 2015.

28. Amr A Elsakka, Viktar S Asadchy, Ihar A Faniayeu, Svetlana N Tcvetkova, and Sergei A Tretyakov. Multifunctional cascaded metamaterials: Integrated transmitarrays. *IEEE Transactions on Antennas and Propagation*, 64(10):4266–4276, 2016.

29. Bo O Zhu and Yijun Feng. Passive metasurface for reflectionless and arbitary control of electromagnetic wave transmission. *IEEE Transactions on Antennas and Propagation*, 63(12):5500–5511, 2015.

30. Xiaobin Hu and Xin Wei. High efficiency broadband- 90 to 90 arbitrary optical rotation realized with meta reflectarray. *Optics express*, 25(5):5641–5650, 2017.

31. Small Unmanned Aircraft Systems (SUAS) Flight Plan: 2016-2036. World Wide Web Page. *https://www.af.mil/Portals/1/documents/isr/Small_UAS_Flight_Plan_2016_to_2036.pdf*.

32. Xiaoyu Miao, Brandon Passmore, Aaron Gin, William Langston, Shivashankar Vangala, William Goodhue, Eric Shaner, and Igal Brener. Doping tunable resonance: Toward electrically tunable mid-infrared metamaterials. *Applied Physics Letters*, 96(10):101111, 2010.

33. Jeremiah P Turpin, Jeremy A Bossard, Kenneth L Morgan, Douglas H Werner, and Pingjuan L Werner. Reconfigurable and tunable metamaterials: a review of the theory and applications. *International Journal of Antennas and Propagation*, 2014, 2014.

34. Kebin Fan and Willie J Padilla. Dynamic electromagnetic metamaterials. *Materials Today*, 18(1):39–50, 2015.

35. Photonics handbook. World Wide Web Page. Available at *https://www.photonics.com/EDU/Handbook.aspx*.

36. Max Born and Emil Wolf. {Chapter} {III} - {Foundations} {of} {Geometrical} {Optics}. In Max Born and Emil Wolf, editors, *Principles of Optics (Sixth (Corrected) Edition)*, pages 109 – 132. Pergamon, sixth (corrected) edition edition, 1980.

37. Constantine A. Balanis. *Antenna Theory: Analysis and Design, 3rd Edition*. Wiley-Interscience, 3 edition, April 2005.

38. Stéphane Larouche and David R. Smith. Reconciliation of generalized refraction with diffraction theory. *Opt. Lett.*, 37(12):2391–2393, Jun 2012.

39. D. A. Roberts, N. Kundtz, and D. R. Smith. Optical lens compression via transformation optics. *Opt. Express*, 17(19):16535–16542, Sep 2009.

40. Donald C O'Shea, Thomas J Suleski, Alan D Kathman, and Dennis W Prather. *Diffractive optics: design, fabrication, and test*, volume 62. Spie Press Bellingham, WA, 2004.

41. Weiwei Wan, Changbao Ma, and Zhaowei Liu. Control the dispersive properties of compound plasmonic lenses. *Optics Communications*, 291:390 – 394, 2013.

42. Dan Hu, Xinke Wang, Shengfei Feng, Jiasheng Ye, Wenfeng Sun, Qiang Kan, Peter J Klar, and Yan Zhang. Ultrathin terahertz planar elements. *Advanced Optical Materials*, 1(2):186–191, 2013.

43. Patrice Genevet, Nanfang Yu, Francesco Aieta, Jiao Lin, Mikhail A. Kats, Romain Blanchard, Marlan O. Scully, Zeno Gaburro, and Federico Capasso. Ultrathin plasmonic optical vortex plate based on phase discontinuities. *Applied Physics Letters*, 100(1):013101, 2012.

44. Ebrahim Karimi, Sebastian A Schulz, Israel De Leon, Hammam Qassim, Jeremy Upham, and Robert W Boyd. Generating optical orbital angular momentum at visible wavelengths using a plasmonic metasurface. *Light: Science and Applications*, 3(5):e167, 2014.

45. Ming Kang, Tianhua Feng, Hui-Tian Wang, and Jensen Li. Wave front engineering from an array of thin aperture antennas. *Optics express*, 20(14):15882–15890, 2012.

46. D. Schurig, J. J. Mock, and D. R. Smith. Electric-field-coupled resonators for negative permittivity metamaterials. *Applied Physics Letters*, 88(4):041109, 2006.

47. Carl Pfeiffer, Naresh K Emani, Amr M Shaltout, Alexandra Boltasseva, Vladimir M Shalaev, and Anthony Grbic. Efficient light bending with isotropic metamaterial Huygens' surfaces. *Nano letters*, 14(5):2491–2497, 2014.

48. Sembiam R Rengarajan and Yahya Rahmat-Samii. The field equivalence principle: Illustration of the establishment of the non-intuitive null fields. *IEEE Antennas and Propagation Magazine*, 42(4):122–128, 2000.

49. Carl Pfeiffer and Anthony Grbic. Cascaded metasurfaces for complete phase and polarization control. *Applied Physics Letters*, 102(23):231116, 2013.

50. Qian Zhao, Ji Zhou, Fuli Zhang, and Didier Lippens. Mie resonance-based dielectric metamaterials. *Materials Today*, 12(12):60–69, 2009.

51. Decker Manuel, Staude Isabelle, Falkner Matthias, Dominguez Jason, Neshev Dragomir N., Brener Igal, Pertsch Thomas, and Kivshar Yuri S. High-efficiency dielectric huygens surfaces. *Advanced Optical Materials*, 3(6):813–820, 2015.

52. Mohammadreza Khorasaninejad, Wei Ting Chen, Robert C Devlin, Jaewon Oh, Alexander Y Zhu, and Federico Capasso. Metalenses at visible wavelengths: Diffraction-limited focusing and subwavelength resolution imaging. *Science*, 352(6290):1190–1194, 2016.

53. Randy L Haupt and Douglas H Werner. *Genetic algorithms in electromagnetics*. John Wiley & Sons, 2007.

54. Yahya Rahmat-Samii and Eric Michielssen. Electromagnetic optimization by genetic algorithms. *Microwave Journal*, 42(11):232–232, 1999.

55. G. Manara, A. Monorchio, and R. Mittra. Frequency selective surface design based on genetic algorithm. *Electronics Letters*, 35(17):1400–1401, Aug 1999.

56. Jeremy A Bossard, Douglas H Werner, Theresa S Mayer, Jacob A Smith, Yan U Tang, Robert P Drupp, and Ling Li. The design and fabrication of planar multiband metallodielectric frequency selective surfaces for infrared applications. *IEEE Transactions on Antennas and Propagation*, 54(4):1265–1276, 2006.

57. Jeremy A Bossard, Do-Hoon Kwon, Yan Tang, Douglas H Werner, and Theresa S Mayer. Low loss planar negative index metamaterials for the mid-infrared based on frequency selective surfaces. In *Antennas and Propagation Society International Symposium, 2007 IEEE*, pages 2873–2876. IEEE, 2007.

58. Douglas J Kern, Douglas H Werner, Agostino Monorchio, Luigi Lanuzza, and Michael J Wilhelm. The design synthesis of multiband artificial magnetic conductors using high impedance frequency selective surfaces. *IEEE Transactions on Antennas and Propagation*, 53(1):8–17, 2005.

59. Douglas H Werner, Theresa S Mayer, Clara Rivero-Baleine, Nikolas Podraza, Kathleen Richardson, Jeremy Turpin, Alexej Pogrebnyakov, J David Musgraves, Jeremy A Bossard, Hee Jung Shin, et al. Adaptive phase change metamaterials

for infrared aperture control. In *Unconventional Imaging, Wavefront Sensing, and Adaptive Coded Aperture Imaging and Non-Imaging Sensor Systems*, volume 8165, page 81651H. International Society for Optics and Photonics, 2011.

60. Jeremy A Bossard, Clinton P Scarborough, Qi Wu, Sawyer D Campbell, Douglas H Werner, Pingjuan L Werner, Scott Griffiths, and Matthew Ketner. Mitigating field enhancement in metasurfaces and metamaterials for high-power microwave applications. *IEEE Transactions on Antennas and Propagation*, 64(12):5309–5319, 2016.

61. Raj Mittra, Chi H Chan, and Tom Cwik. Techniques for analyzing frequency selective surfaces-a review. *Proceedings of the IEEE*, 76(12):1593–1615, 1988.

62. Ling Li, Douglas H Werner, Jeremy A Bossard, and Theresa S Mayer. A model-based parameter estimation technique for wide-band interpolation of periodic moment method impedance matrices with application to genetic algorithm optimization of frequency selective surfaces. *IEEE transactions on antennas and propagation*, 54(3):908–924, 2006.

63. D Bruce Burckel, Bryan M Adomanis, Michael B Sinclair, and Salvatore Campione. Three-dimensional cut wire pair behavior and controllable bianisotropic response in vertically oriented meta-atoms. *Optics express*, 25(25):32198–32205, 2017.

64. Jian-Ming Jin. *The finite element method in electromagnetics*. John Wiley & Sons, 2015.

65. Bryan M Adomanis, D Bruce Burckel, and Michael A Marciniak. Comsol multiphysics® software as a metasurfaces design tool for plasmonic-based flat lenses. Presented at COMSOL Conference 2016 (Boston), available at `https://www.comsol.com/paper/download/363641/adomanis_paper.pdf`.

66. George Harik, Erick Cantú-Paz, David E Goldberg, and Brad L Miller. The gambler's ruin problem, genetic algorithms, and the sizing of populations. *Evolutionary Computation*, 7(3):231–253, 1999.

67. STANLEY GOTSHALL BART Rylander. Optimal population size and the genetic algorithm. *Population*, 100(400):900, 2002.

68. Pedro A. Diaz-gomez and Dean F. Hougen. Empirical study: Initial population diversity and genetic algorithm performance.

69. N.W. Ashcroft and N.D. Mermin. *Solid State Physics*. HRW international editions. Holt, Rinehart and Winston, 1976.

70. D Bruce Burckel, Joel R Wendt, Gregory A Ten Eyck, A Robert Ellis, Igal Brener, and Michael B Sinclair. Fabrication of 3d metamaterial resonators using

self-aligned membrane projection lithography. *Advanced Materials*, 22(29):3171–3175, 2010.

71. D Bruce Burckel, Joel R Wendt, Gregory A Ten Eyck, James C Ginn, A Robert Ellis, Igal Brener, and Michael B Sinclair. Micrometer-scale cubic unit cell 3d metamaterial layers. *Advanced Materials*, 22(44):5053–5057, 2010.

72. David Bruce Burckel. Three-dimensional metamaterials. 2012. U.S. Patent Number 8,197,887.

73. D Bruce Burckel, Paul J Resnick, Patrick S Finnegan, Michael B Sinclair, and Paul S Davids. Micrometer-scale fabrication of complex three dimensional lattice+ basis structures in silicon. *Optical Materials Express*, 5(10):2231–2239, 2015.

74. JR Wendt, DB Burckel, GA Ten Eyck, AR Ellis, I Brener, and MB Sinclair. Fabrication techniques for three-dimensional metamaterials in the midinfrared. *Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena*, 28(6):C6O30–C6O33, 2010.

75. D Bruce Burckel, Joel R Wendt, Igal Brener, and Michael B Sinclair. Dynamic membrane projection lithography [invited]. *Optical Materials Express*, 1(5):962–969, 2011.

76. D Gupta and SP Varma. Specular reflectance of silicon and germanium at variable angles of incidence in the infrared region. *Infrared physics*, 34(1):55–60, 1993.

77. E. Hecht. *Optics*. Addison-Wesley, 2002.

78. MA Ordal, LL Long, RJ Bell, SE Bell, RR Bell, RW Alexander, and CA Ward. Optical properties of the metals al, co, cu, au, fe, pb, ni, pd, pt, ag, ti, and w in the infrared and far infrared. *Applied optics*, 22(7):1099–1119, 1983.

79. Donald C O'Shea, Thomas J Suleski, Alan D Kathman, and Dennis W Prather. *Diffractive optics: design, fabrication, and test*, volume 62. Spie Press Bellingham, WA, 2004.

80. KLayout high-performance layout viewer and editor. Available at `https://www.klayout.de/`.

81. Shulin Sun, Kuang-Yu Yang, Chih-Ming Wang, Ta-Ko Juan, Wei Ting Chen, Chun Yen Liao, Qiong He, Shiyi Xiao, Wen-Ting Kung, Guang-Yu Guo, Lei Zhou, and Din Ping Tsai. High-efficiency broadband anomalous reflection by gradient meta-surfaces. *Nano Letters*, 12(12):6223–6229, 2012. PMID: 23189928.

82. Mikhail I. Shalaev, Jingbo Sun, Alexander Tsukernik, Apra Pandey, Kirill Nikolskiy, and Natalia M. Litchinitser. High-efficiency all-dielectric metasurfaces for ultracompact beam manipulation in transmission mode. *Nano Letters*, 15(9):6261–6266, 2015. PMID: 26280735.

83. Bryan M Adomanis, Matthew R Miller, Stephen E Nauyoks, and Michael A Marciniak. Sensitivities of large-aperture plasmonic metasurface flat lenses in the long-wave infrared. In *High Contrast Metastructures VII*, volume 10542, page 1054210. International Society for Optics and Photonics, 2018.

84. FO Bartell, EL Dereniak, and WL Wolfe. The theory and measurement of bidirectional reflectance distribution function (brdf) and bidirectional transmittance distribution function (btdf). In *Radiation scattering in optical systems*, volume 257, pages 154–161. International Society for Optics and Photonics, 1981.

85. Bradley Balling. A comparative study of the bidirectional reflectance distribution function of several surfaces as a mid-wave infrared diffuse reflectance standard. Technical report, AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH GRADUATE SCHOOL OF ENGINEERING AND MANAGEMENT, 2009.

86. Jason C Vap. Design and characterization of optical metamaterials using tunable polarimetric scatterometry. Technical report, AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH GRADUATE SCHOOL OF ENGINEERING AND MANAGEMENT, 2012.

87. Michael Quirk and Julian Serda. *Semiconductor manufacturing technology*, volume 1. Prentice Hall Upper Saddle River, NJ, 2001.

88. Richard P Feynman, Robert B Leighton, and Matthew Sands. The feynman lectures on physics: Mainly electromagnetism and matter, vol. 2, 1977.

89. COMSOL Multiphysics® application exchange. World Wide Web Page. https://www.comsol.com/community/exchange/612/.

90. DARPA extreme optics and imaging (EXTREME) program. World Wide Web Page. https://www.darpa.mil/program/extreme-optics-and-imaging.

91. Bryan M Adomanis, Paul J Resnick, and D Bruce Burckel. Reconciling measured scattering response of 3d metamaterials with simulation. *EPJ Applied Metamaterials*, 2:9, 2015.

92. Lumerical FDTD®. World Wide Web Page. https://www.lumerical.com/tcad-products/fdtd/.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
14-09-2018	PhD Dissertation	14-09-2018 - 14-09-2018

4. TITLE AND SUBTITLE
Design and Optimization of Plasmonic 3D Huygens Metasurfaces for Highly-Efficient Flat Optics

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Adomanis, Bryan M, Major

5d. PROJECT NUMBER
18ENP239

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management
(AFIT/EN)
2950 Hobson Way, Building 640
WPAFB OH 45433-8865

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT-ENP-DS-18-S-018

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Intentionally Left Blank

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution Statement A. Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
For miniaturization of future USAF unmanned aerial and space systems to become feasible, accompanying sensor components of these systems must also be reduced in size, weight and power (SWaP). Metasurfaces can act as planar equivalents to bulk optics, and thus possess a high potential to meet these low-SWaP requirements. However, functional efficiencies of plasmonic metasurface architectures have been too low for practical application in the infrared (IR) regime. Huygens-like forward-scattering inclusions may provide a solution to this deficiency, but there is no academic consensus on an optimal plasmonic architecture for obtaining efficient phase control at high frequencies. This dissertation asks the question: what are the ideal topologies for generating Huygens-like metasurface building blocks across a full 2π phase space? Instead of employing any *a priori* assumption of fundamental scattering topologies, a genetic algorithm (GA) routine was developed to optimize a "blank slate" grid of binary voxels inside a 3D cavity, evolving the voxel bits until a near-globally optimal transmittance (T) was attained at a targeted phase. All resulting designs produced a normalized $T \geq 80\%$ across the entire 2π range, which is the highest metasurface efficiency reported to-date for a plasmonic solution in the IR regime.

15. SUBJECT TERMS
Metasurfaces, Flat Optics, Plasmonic Metamaterials, Infrared, Genetic Algorithm

16. SECURITY CLASSIFICATION OF:	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
			Dr. Michael A. Marciniak, AFIT/ENP

a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	U	176	(937) 255-3636 x4529 michael.marciniak@afit.edu

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18