

Metasurface Design Using Electromagnetic Inversion

Trevor Brown

A Thesis submitted to the Faculty of Graduate Studies of the
University of Manitoba in partial fulfilment of the requirements for
the degree of

Doctor of Philosophy

Department of Electrical and Computer Engineering
University of Manitoba
Winnipeg, Manitoba, Canada

Copyright © 2020 by Trevor Brown

Abstract

This thesis presents the theory and development of a framework for the design of electromagnetic metasurfaces. These metasurfaces can be used to systematically transform an incident electromagnetic field into a different transmitted field, providing new levels of control not typically possible with conventional materials. Although a metasurface is made up of subwavelength scattering elements, it can be macroscopically represented as a homogenized model using effective surface susceptibilities. These tensorial surface susceptibilities define the relationship between the tangential electric and magnetic fields on either side of the metasurface through a set of generalized boundary conditions known as the generalized sheet transition conditions.

In this work we concern ourselves with *macroscopic* metasurface design, the goal of which is to find an appropriate set of surface susceptibilities to support a desired field transformation. Up until now, macroscopic design methods have been mostly limited to ideal cases in which analytical expressions of the input and output fields are fully known. While this limitation is acceptable for simpler applications such as refraction, reflection, or polarization manipulation of plane waves, more general transformations, such as producing a desired far-field radiation pattern, pose a challenge.

To address the above limitation, we propose framing macroscopic metasurface design as an electromagnetic inverse source problem. We show that the equivalent currents produced by solving an appropriately constructed inverse source problem are directly related to the tangential transmitted fields required to compute the surface susceptibility parameters that characterize the metasurface. The design method is developed for several different types of field specifications, namely complex (amplitude and phase) fields, phaseless (amplitude-only) power patterns, and far-field performance criteria (e.g., main beam direction, beamwidth, null locations, etc.). We then show that local power conservation can be enforced during the inversion process, allowing for the design of metasurfaces that only require passive, lossless, and reciprocal elements. Lastly, we extend the framework to the

design of cascaded metasurfaces. This introduction of a second metasurface removes the need to have equal input and output power distributions, thereby increasing the variety of supported field transformations.

Acknowledgements

First of all I would like to thank my academic advisor, Dr. Puyan Mojabi, for his invaluable guidance, support, and friendship over the past several years. Without his considerable time and effort this work would not have been possible. I would also like to sincerely thank my friends and colleagues Chai Narendra, Chen Niu, and Dr. Nozhan Bayat.

I also would like to extend my gratitude to my Ph.D. committee, Dr. Vladimir Okhmatovski, Dr. Masoud Asadzadeh, and Dr. Sean Hum for taking the time to evaluate and improve this work.

Lastly, I want to thank the Natural Sciences and Engineering Research Council of Canada, the University of Manitoba GETS program, and the Canada Research Chair program for financially supporting this work, as well as the Canadian Microelectronics Corporation (CMC) for the provision of ANSYS Electronics Desktop.

Table of Contents

Abstract	i
Acknowledgements	iii
List of Abbreviations	vi
List of Symbols	vii
1 Introduction	1
1.1 Motivation	4
1.2 Novelties of the Thesis	5
1.3 Thesis Outline	8
2 Background	9
2.1 Metasurface Fundamentals	9
2.2 Electromagnetic Inversion	15
3 Inversion Framework	22
3.1 Introduction	23
3.2 Problem Statement	28
3.3 Methodology	29
3.4 Motivation	30
3.5 Metasurface Fundamentals	33
3.6 Inverse Source Framework	36
3.7 Inversion Algorithm Implementation	40
3.8 FDFD-GSTC Solver	46
3.9 Illustrative Examples	47
3.10 Limitations	58
3.11 Conclusion	58

4	Enforcing Local Power Conservation	60
4.1	Introduction	61
4.2	Inverse Source Design Framework	63
4.3	Enforcing Local Power Conservation (LPC)	66
4.4	Methodology	69
4.5	Results	70
4.6	Conclusion	73
5	Cascaded Metasurface Design	74
5.1	Introduction	75
5.2	Problem Statement	78
5.3	Methodology	80
5.4	Numerical Implementation of the Inversion Framework	84
5.5	Full-Wave Simulated Examples	96
5.6	Limitations	104
5.7	Conclusion	104
6	Conclusions and Future Work	106
6.1	Conclusions	106
6.2	Future Work	109
	References	112
	Appendix A List of Publications	122
	Appendix B Cascaded Gradient Derivations	126
	B.1 Gradients with Respect to \mathbf{x}_2	126
	B.2 Gradients with Respect to \mathbf{x}_1	130
	Appendix C Total Variation Regularization	133

List of Abbreviations

Abbreviation	Description
1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
CG	Conjugate gradient
EFIE	Electric field integral equation
FF	Far-field
FDFD	Finite difference frequency domain
GSTCs	Generalized sheet transition conditions
HPBW	Half-power beamwidth
LPC	Local power conservation
MFIE	Magnetic field integral equation
NF	Near-field
PML	Perfectly matched layer
PCB	Printed circuit board
RWG	Rao-Wilton-Glisson
ROI	Region of interest
SLL	Side lobe levels
TPC	Total power conservation
TV	Total variation
TE	Transverse electric
TM	Transverse magnetic

List of Symbols

Herein we provide some remarks about the notation used in this document along with a list of commonly used symbols:

- Vector-valued functions are denoted using an arrow over uppercase or lowercase non-bold letters, such as $\vec{E}(\vec{r})$, $\vec{H}(\vec{r})$, $\vec{J}(\vec{r}')$, and $\vec{M}(\vec{r}')$.
- Discrete quantities are denoted using bold font. Matrices are represented by bold uppercase letters, such as $\mathbf{A}_{\xi 1}$, $\mathbf{A}_{\xi 2}$, and \mathbf{W} , while vectors are represented using bold lowercase letters, such as \mathbf{x} and \mathbf{f} .
- Cost functionals are represented using \mathcal{C} with appropriate subscripts.

Symbol	Description
$\hat{x}, \hat{y}, \hat{z}$	Unit vectors in the x , y , and z directions
\vec{r}	Position vector to a general observation point
\vec{r}'	Position vector to a general source point
j	Imaginary unit ($j^2 = -1$)
Σ^-	Internal metasurface boundary
Σ^+	External metasurface boundary
$\langle \cdot, \cdot \rangle$	Inner product operator
$\ \cdot \ _2$	L_2 -norm operator
\vec{E}	Electric field vector
\vec{H}	Magnetic field vector
\vec{J}	Electric current
\vec{M}	Magnetic current
k_0	Free space wavenumber
λ	Wavelength in free space
∇	Gradient operator
$\nabla \cdot$	Divergence operator
$\nabla \times$	Curl operator
∇^2	Laplacian operator
$(\cdot)^H$	Hermitian operator
$(\cdot)^*$	Complex conjugate operator

Chapter 1

Introduction

Ever since the concept of electromagnetic radiation was theorized and proved in the late 19th century, we have learned to control electromagnetic waves in various ways for our own benefit. The technologies resulting from these developments are ubiquitous in modern daily life, ranging from communication networks to a majority of consumer electronic devices. Furthering our pursuit of controlling electromagnetic waves, considerable time and effort has been put towards the research and development of electromagnetic metamaterials over the last two decades.

The definition of a ‘metamaterial’ has evolved significantly since the term was coined at the turn of the 21st century. In one of the first papers to use the term, the authors referred to a metamaterial as a “composite medium with simultaneous negative permeability and permittivity” [1]. In the years since, the consensus definition has expanded to include areas beyond electromagnetics, and the term ‘metamaterials’ now broadly encompasses “artificial materials for novel wave phenomena” [2]. These metamaterials generally feature subwavelength-scale features that produce some macroscopic behaviour unobtainable by a typical homogeneous material, such as the ‘left-handed’ medium in [1]. While numerous applications for metamaterials have been developed, the three-dimensional (3D) nature

results in several drawbacks including complex fabrication processes, increased losses, and a larger and heavier structure. To overcome these limitations, the research community began to focus on the development of electromagnetic *metasurfaces*, which can be thought of as the (quasi) two-dimensional (2D) counterpart to 3D volumetric metamaterials.

While the term ‘metasurface’ only appeared in literature recently, physicists and engineers have been using electrically thin devices to control electromagnetic fields for many decades. This includes reflectarrays [3] and frequency selective surfaces [4] developed in the latter half of the 20th century. The knowledge gained from the design and development of these precursory devices has led to modern metasurfaces, which are able to offer a more *systematic* level of control.¹ These metasurfaces are composed of an arrangement of scattering elements of subwavelength size, with a thickness much smaller than the wavelength of operation. The individual scattering elements can be designed to provide local control over the amplitude, phase, and polarization of the electromagnetic field passing through, producing a collective effect that would otherwise be unobtainable [6, 7, 8, 9, 10, 11]. The granular level of control allows metasurfaces to systematically transform a known incident electromagnetic field into a desired output (or transmitted) field, similar to an optical lens. This field transformation ability has led to numerous useful applications including polarization control [12, 13], plane wave refraction and reflection [14], thin absorbers [15], near-field (NF) focusing [16], surface wave control [17], spatial processing [18], and many others. Furthermore, there has recently been a significant focus on programmable or ‘intelligent’ metasurfaces, whose unit cells can be reconfigured dynamically (e.g., through the use of diodes) to modify the intended reflection and transmission [19, 20]. The concept of reconfigurable metasurfaces also has the capability to support smart radio environments, in which the environment (i.e., windows, walls, etc.) can be optimized or configured to enhance wireless communication, rather than relying solely on advancements in either the

¹For more discussion on the precursors to metasurfaces, see Section 1.2 of [5].

transmitter or receiver [21].

The design of metasurfaces is aided by the fact that they behave as typical *materials* in terms of their interaction with electromagnetic fields, that is to say their properties can be homogenized at the wavelength scale. Therefore a metasurface can be macroscopically characterized by effective surface properties, typically using one of the following three representations:

- Effective polarizabilities [14]
- Effective surface impedance [22]
- Effective surface susceptibilities [23]

In a practical sense, the three homogenized models are equivalent² with respect to a zero-thickness metasurface, and the susceptibility model is adopted in the work presented here.

In metasurface design, sometimes referred to as metasurface synthesis, there are two necessary yet distinct steps: (i) macroscopic design and (ii) microscopic design [22]. Macroscopic design is concerned with determining the effective surface parameters that characterize the homogenized metasurface model (e.g., surface susceptibilities) to support one or more specified field transformations. On the other hand, microscopic design refers to the physical design of the individual unit cell structures to achieve the desired effective surface parameters. In this work, we focus on challenges associated with macroscopic design; however, some practical limitations stemming from microscopic design are also considered. Therefore, the goal of this research is to determine the effective surface susceptibility representation of a metasurface that is able to produce some intended field transformation.

²The equivalence holds assuming the polarization in the direction normal to the metasurface is neglected. This simplification is commonly utilized in metasurface design and does not result in the loss of any generality as discussed in Section 2.1.

1.1 Motivation

Several macroscopic design methods have been developed to date [12, 22, 23, 24, 25]. Most of these design procedures allow for the determination of the metasurface elements required to support an arbitrary field transformation, with one major caveat: all fields (i.e., incident, reflected, *and* transmitted) must be known on the boundary imposed by the metasurface.³ This requirement is trivial for simple cases in which the transmitted field is known on the metasurface boundary. With plane wave refraction for example, analytical field expressions are easily formulated. However, the goal of many practical problems is to satisfy some (transmitted) field specifications at some distance away from the metasurface. For example, in antenna design the desired field is often a radiation pattern or pattern characteristics in the far-field (FF) region. In microwave imaging, one might desire a particular NF amplitude distribution, such as focused ‘hot-spot’ at some distance away to improve the achievable resolution [26]. In cases like these, the corresponding transmitted field on the metasurface boundary is not easily defined.

Additionally, the transmitted field specifications could be of several different forms. In the simplest scenario, the field specifications would contain amplitude *and* phase information of the desired electric and/or magnetic field. However, in many situations only a desired power pattern is known or required, and specifying appropriate phase information is challenging and unnecessarily restrictive. For this reason, the second scenario is when the specifications contain amplitude-only (phaseless) field information. The third scenario arises from considering the goals of common antenna design problems. Generally, the aim of antenna design is not to satisfy a particular well-defined radiation pattern but rather a set of (typically FF) performance criteria. These criteria could include main beam direction(s), half-power beamwidth(s) (HPBW), null location(s), polarization, and others. Thus, the

³The design method recently published in [25] does not suffer this limitation and takes a similar approach to the one presented in this thesis. However, the method differs in the sense that it uses equivalent electric currents only and employs a different optimization scheme.

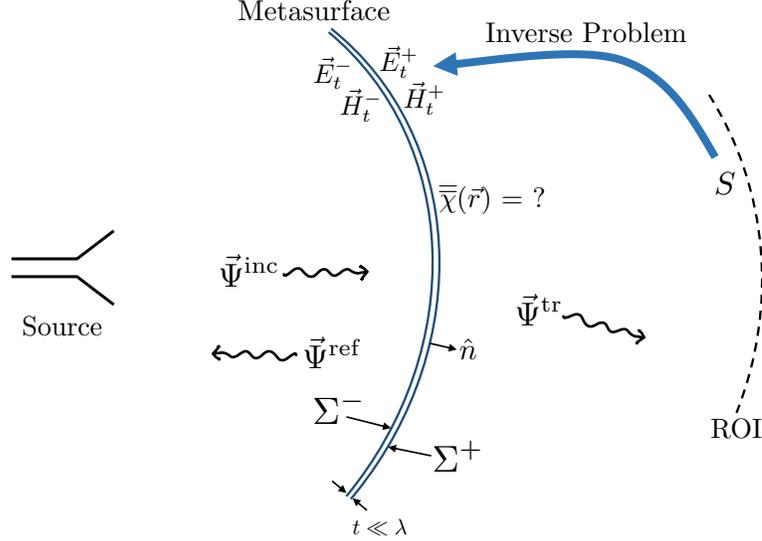


Fig. 1.1: Overview of the metasurface design problem. The subwavelength-thin metasurface has internal and external boundaries Σ^- and Σ^+ , respectively. The incident field $\vec{\Psi}^{\text{inc}}$ interacts with the metasurface, producing both a reflected field $\vec{\Psi}^{\text{ref}}$ and a transmitted field $\vec{\Psi}^{\text{tr}}$. The tangential components of the electric and magnetic fields on Σ^- are denoted as \vec{E}_t^- and \vec{H}_t^- , while the tangential fields on Σ^+ are denoted as \vec{E}_t^+ and \vec{H}_t^+ . The user-provided field specifications S are defined on some region of interest (ROI) external to the metasurface. An electromagnetic inverse source problem is solved to find \vec{E}_t^+ and \vec{H}_t^+ that produce the given field specifications S , from which the susceptibility distribution can be computed. © 2019 IEEE. Reprinted, with permission, from [27] with minor modifications.

third scenario is when the field specifications are given in terms of some combination of the aforementioned performance criteria. A design method that supports the three types of field specifications listed above has the potential to facilitate more practical applications of metasurfaces beyond the idealized examples commonly shown in the literature.

1.2 Novelties of the Thesis

With above limitations of the current macroscopic design methods in mind, the *single* metasurface design problem can be constructed as shown in Figure 1.1. To allow for more flexible design goals (such as the three scenarios mentioned above), we propose framing the meta-

surface design problem as an electromagnetic inverse source problem.⁴ In electromagnetic inversion, the goal is to determine a *cause* from some known electromagnetic effect. In electromagnetic inverse *source* problems, the cause is a set of electric and/or magnetic currents that produce the known (observed or desired) effect, often referred to as the ‘equivalent’ or ‘reconstructed’ currents. In this problem, the electromagnetic effect is the set of field specifications in some region of interest (ROI) external to the metasurface. For each of the three scenarios, a cost functional is constructed that quantifies the difference between the fields produced by the reconstructed currents and the field specifications. In each case, the (potentially nonlinear) cost functional is subsequently minimized to find an appropriate set of equivalent currents that produces the desired result. Furthermore, we show that under certain conditions, the reconstructed currents are directly related to the required tangential output electric and magnetic fields on the metasurface. The tangential electric and magnetic fields can then be used in conjunction with the known incident field to determine the effective surface susceptibility distributions required to support the transformation.

Formulating the problem using electromagnetic inversion results in several additional advantages beyond the added generality with regards to the type of field specifications. First, the field specifications can be arbitrarily located at any number of discrete points in the region external to the metasurface. This is in contrast to alternative techniques that rely on backpropagation which, among other limitations, require the complex fields to be defined on canonically-shaped surfaces. Secondly, the generality of the inverse source framework allows for arbitrarily-shaped metasurfaces, or even disconnected metasurfaces. For example, this framework could allow for the design of a pair of metasurfaces, each with its own incident field source, which together produce the desired output field. Thirdly, the inherent non-uniqueness of the inverse source problem is, perhaps surprisingly, beneficial in

⁴Two recently published methods also approach the design of metasurfaces using equivalent current formulations, although they are focused on the application areas of surface wave to leaky wave transformation [28] and metasurface reflectarrays [29].

this case. The non-uniqueness indicates there will be multiple solutions that produce the same output field, allowing us to choose the ‘best’ solution with regards to other criteria (such as ease of implementation).

Additionally, it is of practical importance to ensure that the resulting metasurfaces do not require loss and/or gain in order to perform the desired field transformation. To accomplish this, we ensure *local* power is conserved across the metasurface (i.e., the time-averaged real power density into and out of each unit cell are equal). In this work we express the local power conservation (LPC) condition [30] as a function of the electric and magnetic equivalent currents and introduce an additional cost functional that, when minimized, ensures that local power is conserved. Assuming this condition holds, the transformation can be implemented in a passive and lossless way by introducing magnetoelectric coupling [31] as discussed in more detail in Chapter 4.

Lastly, we extend the design framework to the scenario in which *two* metasurfaces are used in a cascaded fashion to produce a desired transmitted field, similar to the backpropagation-based technique in [32, 33]. A *single* passive and lossless metasurface is limited in the variety of field transformations that it can support due to the fact that local power must be conserved at each and every unit cell. The addition of a second metasurface allows for a redistribution of power in between the two metasurfaces, allowing for a greater variety of supported field transformations while still ensuring that each metasurface is passive and lossless. In development of this cascaded design method, we introduce additional regularization to the optimization procedure in the form of a total variation (TV) regularizer. This regularization term has a smoothing effect on the equivalent currents and provides several benefits discussed in more detail in Chapter 5. Furthermore, the optimization for the cascaded metasurface design is performed using the conjugate gradient (CG) technique, with analytical expressions provided for all required gradients.

1.3 Thesis Outline

This thesis is structured based on the grouped manuscript style (i.e., sandwich thesis) as defined by the University of Manitoba. Thus, Chapters 3 to 5 are the author’s published or submitted peer-reviewed journal papers. Due to this format, some background information and general concepts may be repeated throughout this thesis. A complete list of the author’s published and submitted papers can be found in Appendix A.

Chapter 2 of this thesis provides background information on metasurface theory and the electromagnetic inversion framework. Chapter 3 is a published peer-reviewed journal paper that is part of a special issue of the IEEE Transactions on Antennas and Propagation on “Recent Advances in Metamaterials and Metasurfaces”. This paper describes the details of the inversion framework used to produce the required tangential output fields for different types of field specifications, and how to subsequently compute the susceptibilities assuming a monoisotropic metasurface. Several 2D examples are presented that use a finite difference frequency domain (FDFD) solver to simulate the behaviour of the designed metasurfaces, along with a 3D example to demonstrate the arbitrary geometry capabilities. Chapter 4 addresses the main limitation of the previous chapter, extending the method to the design of passive, lossless, and reciprocal metasurfaces. The material in this chapter is a published peer-reviewed journal paper in IEEE Antennas and Wireless Propagation Letters. Chapter 5 is a journal paper that is under review for publication in IEEE Transactions on Antennas and Propagation, and extends the inversion framework to the design of cascaded metasurfaces to allow for a greater variety of supported field transformations. While the possible field transformations that a single passive and lossless metasurface can support are limited due to the LPC condition, adding a second metasurface avoids this restriction by allowing for a redistribution of power. This chapter also introduces the TV regularizer that is added to improve the smoothness of the solution. Lastly, Chapter 6 summarizes the contributions of this thesis and provides recommendations for future work.

Chapter 2

Background

In this chapter we present background information on (i) metasurfaces and (ii) electromagnetic inversion to aid in the understanding of the material in the following chapters.

2.1 Metasurface Fundamentals

A metasurface introduces an electromagnetic discontinuity that cannot be explicitly represented by conventional boundary conditions in their standard form [23]. However, alternative *rigorous* boundary conditions have been developed that extend the conventional boundary conditions to the case in which surface polarizations are present [34]. These boundary conditions are known as the generalized sheet transition conditions (GSTCs), and were first developed in [35]. Kuester *et. al* first applied these conditions to metasurfaces by considering a surface of electrically small scatterers that can be characterized by surface polarization densities [36].

Assuming a time-dependency of $e^{j\omega t}$ for the time-harmonic fields, where ω is the angular

frequency, the GSTCs can be expressed as [5]

$$\hat{n} \times \Delta \vec{H} = j\omega \vec{P}_t^e - \hat{n} \times \nabla_t (\hat{n} \cdot \vec{P}^m), \quad (2.1a)$$

$$\Delta \vec{E} \times \hat{n} = j\omega \mu_0 \vec{P}_t^m - \frac{1}{\epsilon_0} \nabla_t (\hat{n} \cdot \vec{P}^e) \times \hat{n}, \quad (2.1b)$$

$$\hat{n} \cdot \Delta \vec{D} = -\nabla \cdot \vec{P}_t^e, \quad (2.1c)$$

$$\hat{n} \cdot \Delta \vec{B} = -\mu_0 \nabla \cdot \vec{P}_t^m \quad (2.1d)$$

where \hat{n} is the unit outward normal vector to the metasurface as shown in Figure 1.1, the terms \vec{P}^e and \vec{P}^m represent the electric and magnetic surface polarization densities, and \vec{D} and \vec{B} represent the electric and magnetic flux densities, respectively. The terms ϵ_0 and μ_0 represent the permittivity and permeability of free space, respectively, and the subscript t denotes the vector components that are locally transverse relative to \hat{n} (i.e., locally tangential to the surface). Therefore, based on the local coordinates shown in Figure 1.1, \vec{P}_t can have two components: one along the unit vector \hat{u} and the other along the unit vector \hat{v} , where $\hat{u} \times \hat{v} = \hat{n}$ and $\hat{u} \perp \hat{v}$. The operator ∇_t denotes the gradient operator with respect to \hat{u} and \hat{v} , and the operator Δ indicates the difference between the fields on either side of the metasurface, defined in terms of the transmitted, incident, and reflected fields as

$$\Delta \vec{\Psi} \triangleq \vec{\Psi}^{\text{tr}}|_{\Sigma^+} - \left(\vec{\Psi}^{\text{inc}}|_{\Sigma^-} + \vec{\Psi}^{\text{ref}}|_{\Sigma^-} \right), \quad (2.2)$$

where Σ^- and Σ^+ denote the internal (i.e., incident field side) and external (i.e., output field side) boundaries of the metasurface as depicted in Figure 1.1. As can be seen, the GSTC formulation has been written for an arbitrarily-shaped metasurface, and is only valid for curved metasurfaces if the curvature is electrically large (i.e., the metasurface is electromagnetically quasi-planar) [37].

In the most general sense, the polarization densities of a linear bianisotropic medium

can be written in terms of the acting fields, \vec{E}_{act} and \vec{H}_{act} as [38]

$$\vec{P}^e = \epsilon_0 N \langle \bar{\alpha}_{ee} \rangle \vec{E}_{\text{act}} + N \langle \bar{\alpha}_{em} \rangle \sqrt{\mu_0 \epsilon_0} \vec{H}_{\text{act}}, \quad (2.3a)$$

$$\vec{P}^m = N \langle \bar{\alpha}_{mm} \rangle \vec{H}_{\text{act}} + N \langle \bar{\alpha}_{me} \rangle \sqrt{\frac{\epsilon_0}{\mu_0}} \vec{E}_{\text{act}}, \quad (2.3b)$$

where the N is the number of scattering elements per unit area and the $\langle \bar{\alpha} \rangle$ tensors represent the averaged electric/magnetic (first subscript) polarizabilities of a given scatterer resulting from an electric/magnetic excitation (second subscript). These acting fields can be thought of as the average of the fields on both sides of the metasurface while taking into account the scattering from all elements *except* the given scatterer. Kuester *et al.* have shown that the fields from the given scatterer can be modelled by a disk containing electric and magnetic dipoles which are related to the electric and magnetic polarization densities \vec{P}^e and \vec{P}^m [36]. These relations can be written in terms of the average fields across the metasurface, leading to the *macroscopic* representation

$$\vec{P}^e = \epsilon_0 \bar{\chi}_{ee} \vec{E}_{\text{av}} + \bar{\chi}_{em} \sqrt{\mu_0 \epsilon_0} \vec{H}_{\text{av}}, \quad (2.4a)$$

$$\vec{P}^m = \bar{\chi}_{mm} \vec{H}_{\text{av}} + \bar{\chi}_{me} \sqrt{\frac{\epsilon_0}{\mu_0}} \vec{E}_{\text{av}}, \quad (2.4b)$$

where $\bar{\chi}_{ee}$, $\bar{\chi}_{em}$, $\bar{\chi}_{mm}$, and $\bar{\chi}_{me}$ represent the electric/magnetic (first subscript) surface susceptibility tensors indicating the response to an electric/magnetic (second subscript) field excitation [39]. These susceptibilities are 3×3 dimensionless tensors and therefore represent a total of 36 susceptibility terms. The average fields are defined as

$$\vec{\Psi}_{\text{av}} \triangleq \frac{\vec{\Psi}^{\text{tr}}|_{\Sigma^+} + \left(\vec{\Psi}^{\text{inc}}|_{\Sigma^-} + \vec{\Psi}^{\text{ref}}|_{\Sigma^-} \right)}{2}. \quad (2.5)$$

The expression in (2.4) includes both the normal and tangential components of the polarization densities (and fields), and substituting this formulation into the GSTCs in (2.1) would

result in a system of differential equations that is not easily solved. However, the normal components of the susceptibility tensors (and hence the polarization densities) are typically ignored. This simplification results in a system of *linear* equations as the spatial derivatives in (2.1) are neglected, and can be justified in several ways [40]. One justification is that metasurfaces are commonly designed for an incident field that impinges with the metasurface at close to normal incidence, and the tangential components of the incident electric and magnetic fields would be much larger than any normal components. Another justification arises from the fact that any field can be equivalently represented in terms of its tangential components [23]. Thus, neglecting the normal components of the susceptibilities does not eliminate any potential solutions and does not result in any loss of generality. Therefore, we substitute the *tangential* components of (2.4) into (2.1), which results in [41]

$$\hat{n} \times \Delta \vec{H}_t = j\omega\epsilon_0 \bar{\bar{\chi}}_{ee} \vec{E}_{t,av} + j\omega\sqrt{\epsilon_0\mu_0} \bar{\bar{\chi}}_{em} \vec{H}_{t,av}, \quad (2.6a)$$

$$\Delta \vec{E}_t \times \hat{n} = j\omega\mu_0 \bar{\bar{\chi}}_{mm} \vec{H}_{t,av} + j\omega\sqrt{\epsilon_0\mu_0} \bar{\bar{\chi}}_{me} \vec{E}_{t,av}, \quad (2.6b)$$

where the terms $\bar{\bar{\chi}}_{ee}$, $\bar{\bar{\chi}}_{em}$, $\bar{\bar{\chi}}_{mm}$, and $\bar{\bar{\chi}}_{me}$ are now 2×2 tensors (with units of m) that act only on the tangential components of the electric and magnetic fields.

The expressions in (2.6) are the representations we utilize in the design framework described in the following chapters. If the tangential fields on either side of the metasurface boundary are known, the expressions in (2.6) can be solved to determine a set of susceptibility distributions that are able to support the transformation.¹

2.1.1 Requirements for Reciprocity, Losslessness, and Passivity

There are several practical aspects that are important to consider in the design of electromagnetic metasurfaces to ensure the final designs are physically practical, the first of

¹The system depicted in (2.6) has enough degrees of freedom to simultaneously support *multiple* field transformations [23]. If a single transformation is desired, not all of the degrees of freedom are needed and some components of the susceptibility tensors may be allowed to be zero.

which is that the metasurface should be reciprocal. A system is considered reciprocal if the location of the source and observer can be interchanged without any difference in the received signal. A nonreciprocal device can be achieved in several ways, such as biasing a ferrimagnetic medium with an external magnetic field, or incorporating some form of spatial or temporal modulation [42]. Breaking reciprocity using these techniques is often challenging to implement, thus, ensuring the metasurface is reciprocal results in a more practically feasible structure. The reciprocity conditions for a general bianisotropic medium in terms of the susceptibility tensors are [38]

$$\overline{\overline{\chi}}_{ee}^T = \overline{\overline{\chi}}_{ee}, \quad \overline{\overline{\chi}}_{mm}^T = \overline{\overline{\chi}}_{mm}, \quad -\overline{\overline{\chi}}_{em}^T = \overline{\overline{\chi}}_{me}, \quad (2.7)$$

where the transpose operator is denoted by the superscript T.

Next, we consider the presence of loss and/or gain in the metasurface. We begin by stating the time-averaged bianisotropic Poynting theorem [43], which for time-harmonic fields can be written as

$$\nabla \cdot \langle \vec{S} \rangle = -\langle \vec{I}_J \rangle - \langle \vec{I}_M \rangle - \langle \vec{I}_{Pe} \rangle - \langle \vec{I}_{Pm} \rangle \quad (2.8)$$

where \vec{S} is the Poynting vector and the $\langle \cdot \rangle$ operator represents the time-averaging operation over a single period. The terms \vec{I}_J , \vec{I}_M , \vec{I}_{Pe} , and \vec{I}_{Pm} represent the work done due to electric currents \vec{J} , magnetic currents \vec{M} , electric polarization, and magnetic polarization,

respectively. In the frequency domain these time-averaged quantities can be stated as

$$\langle \vec{I}_J \rangle = \frac{1}{2} \text{Re} \left(\vec{E}^* \cdot \vec{J} \right), \quad (2.9a)$$

$$\langle \vec{I}_M \rangle = \frac{1}{2} \text{Re} \left(\vec{H}^* \cdot \vec{M} \right), \quad (2.9b)$$

$$\langle \vec{I}_{Pe} \rangle = \frac{1}{4} \text{Re} \left[j\omega \left(\vec{E}^* \cdot \vec{P}^e - \vec{P}^{e*} \cdot \vec{E} \right) \right], \quad (2.9c)$$

$$\langle \vec{I}_{Pm} \rangle = \frac{1}{4} \text{Re} \left[j\omega\mu_0 \left(\vec{H}^* \cdot \vec{P}^m - \vec{P}^{m*} \cdot \vec{H} \right) \right], \quad (2.9d)$$

where the * superscript represents the complex conjugate operation. These expressions can be written in terms of the susceptibility tensors as [5]

$$\langle \vec{I}_J \rangle = \frac{1}{4} \text{Re} \left(j\omega\epsilon_0 \vec{E}^* \cdot (\bar{\chi}_{ee} - \bar{\chi}_{ee}^*) \cdot \vec{E} \right), \quad (2.10a)$$

$$\langle \vec{I}_M \rangle = \frac{1}{4} \text{Re} \left(j\omega\mu_0 \vec{H}^* \cdot (\bar{\chi}_{mm} - \bar{\chi}_{mm}^*) \cdot \vec{H} \right), \quad (2.10b)$$

$$\langle \vec{I}_{Pe} \rangle = \frac{1}{4} \text{Re} \left[j\omega\epsilon_0 \left(\vec{E}^* \cdot (\bar{\chi}_{ee} - \bar{\chi}_{ee}^H) \cdot \vec{E} + 2\eta_0 \vec{E}^* \cdot \bar{\chi}_{em} \cdot \vec{H} \right) \right], \quad (2.10c)$$

$$\langle \vec{I}_{Pm} \rangle = \frac{1}{4} \text{Re} \left[j\omega\mu_0 \left(\vec{H}^* \cdot (\bar{\chi}_{mm} - \bar{\chi}_{mm}^H) \cdot \vec{H} - 2\vec{E}^* \cdot \bar{\chi}_{me}^H \cdot \frac{\vec{H}}{\eta_0} \right) \right], \quad (2.10d)$$

where the H superscript denotes the Hermitian (conjugate transpose) operation and the fields \vec{E} and \vec{H} can be replaced by the average fields across the metasurface [5]. The term $\nabla \cdot \langle \vec{S} \rangle$ in (2.8) represents the divergence of the power flow at a single location on the metasurface. If $\nabla \cdot \langle \vec{S} \rangle < 0$, then the metasurface is lossy as there is a net power flow *into* the metasurface. Conversely, $\nabla \cdot \langle \vec{S} \rangle > 0$ represents a net power flow *out* of the metasurface, representing gain. If $\nabla \cdot \langle \vec{S} \rangle = 0$ then the metasurface is overall lossless and passive. This being said, there is an important caveat. Even if $\nabla \cdot \langle \vec{S} \rangle = 0$, we may actually have loss and gain in the individual terms of (2.10) that compensate for each other. Therefore, we require that all terms of (2.10) are zero in order to ensure that the corresponding metasurface is completely passive and lossless. Setting the four expressions in (2.10) to zero and assuming reciprocity holds, we get the following conditions for a passive, lossless, and reciprocal

metasurface [5]

$$\overline{\overline{\chi}}_{ee}^T = \overline{\overline{\chi}}_{ee}^*, \quad \overline{\overline{\chi}}_{mm}^T = \overline{\overline{\chi}}_{mm}^*, \quad \overline{\overline{\chi}}_{em}^T = \overline{\overline{\chi}}_{me}^*. \quad (2.11)$$

Since we have neglected the normal component of the polarization, the four susceptibility tensors will each be 2×2 tensors for a general bianisotropic metasurface. For the case in which $\overline{\overline{\chi}}_{ee}$, $\overline{\overline{\chi}}_{mm}$, $\overline{\overline{\chi}}_{em}$, and $\overline{\overline{\chi}}_{me}$ are all scalar, then the conditions for a reciprocal, lossless, metasurface simplify to

$$\chi_{ee}, \chi_{mm} \in \mathbb{R}, \quad \chi_{em}, \chi_{me} \in \mathbb{I}, \quad \chi_{em} = -\chi_{me}, \quad (2.12)$$

where \mathbb{R} and \mathbb{I} represent the spaces of real and imaginary numbers, respectively. The inclusion of these conditions into the design process will be discussed in more detail in Chapter 4.

2.2 Electromagnetic Inversion

An inverse problem is defined as a problem in which we would like to determine the *cause* of an observed *effect*. Inverse problems are common in several different fields of science and engineering, such as image deblurring [44], acoustic sensing [45], geophysical modelling [46], and others. In this work we are concerned with electromagnetic inversion, in which the observable effect consists of electric and/or magnetic fields. Electromagnetic inversion can be decomposed into two categories, each classified by the type of ‘cause’ that is being determined. The first category is inverse *scattering*, in which the aim is to find the material properties (typically relative permittivity and/or permeability) of some unknown electromagnetic scatterer. The unknown scatterer is illuminated with an incident field, and the observed ‘effect’ is the measured field scattered from the object. The second category of electromagnetic inversion is inverse *source*, in which the goal is to determine a set of electric and/or magnetic currents that produce some measured or desired electromagnetic fields.

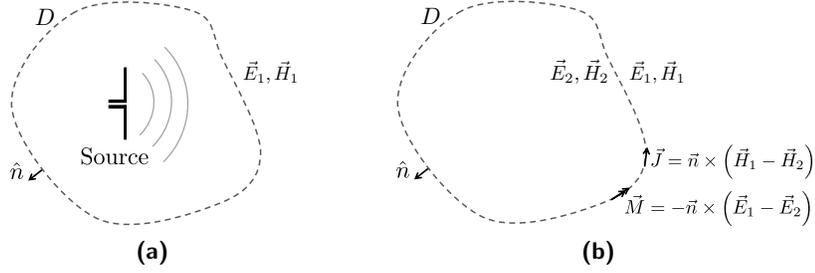


Fig. 2.1: The electromagnetic equivalence principle. The electromagnetic sources in (a) can be represented by the equivalent electric and magnetic currents \vec{J} and \vec{M} that produce the same fields \vec{E}_1 and \vec{H}_1 external to D .

In this work we frame the metasurface design problem as an inverse source problem, and thus the following section presents some additional details about the construction of this framework.

2.2.1 Inverse Source Problems

The inverse source problem is constructed by way of the electromagnetic surface equivalence principle [47]. Consider an electromagnetic source that produces fields \vec{E}_1 and \vec{H}_1 everywhere in space and is enclosed by some fictitious boundary D as shown in Figure 2.1(a). The equivalence principle states that we can remove the source and construct a scenario in which the fields *outside* D remain the same, but the fields inside D (denoted as \vec{E}_2 and \vec{H}_2) are different. The newly imposed discontinuity in the fields requires a set of electric and magnetic surface currents on D , denoted as \vec{J} and \vec{M} in Figure 2.1(b). These currents, which we often refer to as ‘equivalent’ or ‘reconstructed’ currents, are related to the difference between the tangential components of the interior and exterior fields according to the well-known boundary conditions

$$\vec{J} = \hat{n} \times (\vec{H}_1 - \vec{H}_2), \quad (2.13a)$$

$$\vec{M} = -\hat{n} \times (\vec{E}_1 - \vec{E}_2), \quad (2.13b)$$

where \hat{n} is the unit outward normal vector to D as shown in Figure 2.1. In other words, the equivalent currents in Figure 2.1(b) radiate in free space, producing the original fields \vec{E}_1 and \vec{H}_1 outside D and different fields \vec{E}_2 and \vec{H}_2 within D . In inverse source problems, the fields \vec{E}_1 and \vec{H}_1 are known in some capacity outside D (most often some distance away), and the goal is to find \vec{J} and \vec{M} that produce these fields.

2.2.2 Forward Problem

Before we can consider the inverse problem, we need to have a mathematical model defining the relationship between the equivalent electric and magnetic currents and the electric and magnetic fields that they produce. We will refer to this mapping as the *forward* problem, as we are computing the observable effects from the known cause. In free space, the mapping from currents to fields is computed using the electric and magnetic field integral equations (EFIE and MFIE)

$$\begin{aligned} \vec{E}(\vec{r}) = & -j\eta_0 k_0 \int_D \left[\vec{J}(\vec{r}') + \frac{1}{k_0^2} \nabla \nabla'_D \cdot \vec{J}(\vec{r}') \right] g(\vec{r}, \vec{r}') \vec{r}' \\ & + \int_D \vec{M}(\vec{r}') \times \nabla g(\vec{r}, \vec{r}') d\vec{r}', \end{aligned} \quad (2.14)$$

$$\begin{aligned} \vec{H}(\vec{r}) = & \frac{-jk_0}{\eta_0} \int_D \left[\vec{M}(\vec{r}') + \frac{1}{k_0^2} \nabla \nabla'_D \cdot \vec{M}(\vec{r}') \right] g(\vec{r}, \vec{r}') \vec{r}' \\ & - \int_D \vec{J}(\vec{r}') \times \nabla g(\vec{r}, \vec{r}') d\vec{r}', \end{aligned} \quad (2.15)$$

where $\vec{r}' \in D$ and the position vector \vec{r} belongs to the region of interest (ROI) in which \vec{E}_1 and \vec{H}_1 are known. The characteristic impedance of free space is denoted by η_0 , k_0 represents the free space wavenumber, and ' $\nabla'_D \cdot$ ' represents the surface divergence operator with respect to the primed coordinates (i.e., on D). The Green's function is denoted as

$g(\vec{r}, \vec{r}')$ and can be written in free space as

$$g(\vec{r}, \vec{r}') = \frac{e^{-jk_0|\vec{r}-\vec{r}'|}}{4\pi|\vec{r}-\vec{r}'|}. \quad (2.16)$$

It should be noted that we typically only require the use of (2.14) *or* (2.15), but not both. This is justified by the uniqueness theorem, which states: “A field in a lossy region is uniquely specified by the sources within the region plus the tangential components of \vec{E} over the boundary, or the tangential components of \vec{H} over the boundary, or the former over part of the boundary and the latter over the rest of the boundary” [48].²

Lastly, we often represent the integral equation associated with each problem (i.e., (2.14) or (2.15)) using operator notation. For example, if we are interested in the electric field, the forward problem would be represented as

$$\vec{E}(\vec{r}) = \mathcal{A}(\vec{J}(\vec{r}'), \vec{M}(\vec{r}')) \quad (2.17)$$

where \mathcal{A} is the linear operator representing (2.14). Therefore the forward problem can be summarized as: for a given \vec{J} and \vec{M} , find \vec{E} or \vec{H} in the ROI; however, in inverse source problems we know \vec{E} or \vec{H} and would like to find the currents.

2.2.3 Inverse Problem

With a defined model relating the equivalent currents and their corresponding fields, we can now consider the inverse problem. It is important to note that the inverse problem will be inherently ill-posed in the sense of Hadamard due to the fact that the underlying integral equations in (2.14) and (2.15) are Fredholm integral equations of the first kind [49]. More specifically, an ill-posed problem is one that fails to satisfy *all* of the following criteria [50]:

²Extending this theorem to the lossless case is done by taking the limit of the resulting fields as the loss goes to zero [48].

1. A solution exists.
2. The solution is unique.
3. The solution depends continuously on the data.

In order to solve the inverse problem numerically, the relation in (2.17) must be discretized, both in terms of the equivalent currents on D and the fields in the ROI. Because of this discretization, there is a loss of degrees of freedom in the equivalent currents, and a true solution may not exist. Furthermore, in design problems the desired fields are user-specified (not measured), and depending on the choice of D , a solution may not exist.³ Thus the first criterion may fail, causing the inverse source problem to be ill-posed. This being said, the second criterion also fails due to the presence of non-radiating currents [51]. Due to the inherent rank deficiency of the discretized inverse source problem (in part due to the finite number of measurement/observation points and the low pass filtering effect of the Green's function integration kernel), there exists 'non-radiating' currents that produce zero fields at the points of interest, but non-zero fields elsewhere.⁴ Since any combination of these non-radiating currents can be added to a valid solution without an observable effect, an infinite number of solutions are possible, and thus the uniqueness criterion is violated. Finally, the smoothness of the Green's function (the kernel of the integral equations) causes the operator to become highly ill-conditioned. Attempting to solve an inverse problem numerically with an ill-conditioned operator can be challenging and stability issues can arise.

Of the various ways to solve an ill-posed inverse problem, in this work we minimize a cost functional over the unknown equivalent currents that quantifies the difference between the fields produced by the equivalent currents and the desired or measured fields that we want to produce. Assuming we are attempting to produce a desired or measured electric

³For example, to achieve a certain far-field directivity a minimum aperture size is required.

⁴The full set of these non-radiating currents comprises the null space of the integral operator associated with the inverse problem.

field \vec{E}_{des} , this cost functional would be of the general form

$$\left\| \mathcal{A} \left(\vec{J}(\vec{r}'), \vec{M}(\vec{r}') \right) - \vec{E}_{\text{des}}(\vec{r}) \right\|_{\text{ROI}}^2 \quad (2.18)$$

where the L_2 -norm over the ROI is defined as

$$\left\| \vec{E} \right\|_{\text{ROI}} = \left\langle \vec{E}, \vec{E} \right\rangle_{\text{ROI}}^{\frac{1}{2}} \quad (2.19)$$

with the inner product defined as

$$\left\langle \vec{E}_1, \vec{E}_2 \right\rangle_{\text{ROI}} = \int_{\text{ROI}} \vec{E}_1(\vec{r}) \cdot \vec{E}_2^*(\vec{r}) d\vec{r}. \quad (2.20)$$

Successfully minimizing (2.18) over \vec{J} and \vec{M} would result in a set of equivalent currents that produce the desired effect; however, the ill-posed nature of the problem makes this challenging. First of all, solving the system will invert any noise introduced either in \vec{E}_{des} (e.g., measurement noise) or \mathcal{A} (e.g., discretization or modelling error), and can result in nonphysical solutions with extremely rapid spatial variations. While such a solution technically represents a minimum of (2.18), it is the result of over-fitting the data and exhibits undesirable characteristics which we would prefer to avoid. For this reason, we utilize regularization to solve the problem in a stable manner.

The term ‘regularization’ refers to the inclusion of additional constraints or information in order to bias the solution away from those with undesirable characteristics. There are several common ways in which regularization is applied to inverse problems, including truncated singular value decomposition [52], Tikhonov regularization [50], multiplicative regularization [53], and others. With the exception of the material presented in Chapter 4, in this work we adopt the use of the truncated CG method, which is a gradient-based iterative optimization technique that acts as a Krylov subspace regularizer [54]. The iterative proce-

ture is truncated once the solution begins to stagnate, thereby preventing over-fitting. In Chapter 5, we introduce another form of regularization based on the L_2 -norm total variation regularizer which has been used in microwave imaging [53], NF antenna measurements [55], and image deblurring [44]. This additive regularizer has the effect of favouring smoother solutions and is discussed in more detail in Section 5.4.3 and Appendix C.

2.2.4 Internal Field Constraint

Up to this point, we have not imposed any restrictions on the fields produced by the equivalent currents on the *inside* of D , which we refer to as \vec{E}_2 and \vec{H}_2 . Any combination of \vec{J} and \vec{M} that produces the desired external fields \vec{E}_1 and \vec{H}_1 is deemed acceptable, regardless of the internal fields produced. However, imposing a restriction on the internal fields can provide several benefits. In this work we enforce the internal fields to be zero, which is known as Love’s equivalence condition [56]. Imposing this condition causes the boundary conditions in (2.13) to simplify to

$$\vec{J} = \vec{n} \times \vec{H}_1 \tag{2.21a}$$

$$\vec{M} = -\vec{n} \times \vec{E}_1. \tag{2.21b}$$

Thus, enforcing the internal fields to be zero results in the equivalent currents being directly proportional to the tangential components of the external electric and magnetic fields on D . In inverse source problems focused on characterization, such as NF antenna measurements, this property allows for the detection of defects or anomalies [57]. In this work, we utilize the relationships in (2.21) to determine the transmitted fields on the output side of the metasurface by first finding a set of equivalent currents. Additionally, enforcing Love’s condition has the added benefit of improving the conditioning of the system, thereby reducing the ill-posedness of the associated inverse problem [58].

Chapter 3

Inversion Framework

Preface

In this chapter, the general electromagnetic inversion framework is presented. Following an introduction in Section 3.1, the metasurface design problem is explained in Section 3.2. The general methodology and motivation for the approach are presented in Sections 3.3 and 3.4, respectively. Section 3.5 is a brief review of some relevant metasurface concepts, followed by the main contribution of the work, the inverse source framework, presented in Section 3.6. Details about the finite difference frequency domain solver are given in Section 3.8, which is then used for the 2D illustrative examples in Section 3.9. The limitations of the method (at the time of acceptance of the publication) are stated in Section 3.10, with some concluding thoughts following in Section 3.11.

The material presented in this chapter is based on the paper published in IEEE Transactions on Antennas and Propagation in March 2020 [27].¹

¹© 2020 IEEE. Reprinted, with permission, from T. Brown, C. Narendra, Y. Vahabzadeh, C. Caloz, and P. Mojabi, “On the Use of Electromagnetic Inversion for Metasurface Design”, *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 3, pp. 1812-1824, 2020.

Abstract

We show that the use of the electromagnetic inverse source framework offers great flexibility in the design of metasurfaces. In particular, this approach is advantageous for antenna design applications where the goal is often to satisfy a set of performance criteria such as half-power beamwidths and null directions, rather than satisfying a fully-known complex field. In addition, the inverse source formulation allows the metasurface and the region over which the desired field specifications are provided to be of arbitrary shape. Some of the main challenges in solving this inverse source problem, such as formulating and optimizing a nonlinear cost functional, are addressed. Lastly, some two-dimensional (2D) and three-dimensional (3D) simulated examples are presented to demonstrate the method, followed by a discussion of the method's current limitations.

3.1 Introduction

Electromagnetic metasurfaces are structures of subwavelength thickness that are composed of a subwavelength lattice of scattering elements [6, 7, 8, 9, 10, 11]. In particular, transmitting metasurfaces, which are the focus of this paper, offer a systematic transformation of an incident field (input field) to a desired transmitted field (output field) by imposing appropriate surface boundary conditions. From a circuit view point, these surfaces may be thought of as collections of several two-port networks transforming the input field to the output field [59]. In other words, a metasurface may be considered as an electromagnetic wave transformer [60]: similar to electric transformers that enable the *systematic* change of a given voltage level to a desired one, electromagnetic metasurfaces have the potential to enable the *systematic* transformation of a given electromagnetic field to a desired one. Thus, they can play an important role in antenna engineering applications.

Recently, a few metasurface design methods have been presented, e.g., see [22, 23].

In particular, two main design steps are outlined in [22]: macroscopic and microscopic. The macroscopic step finds the surface electric/magnetic impedance/admittance profiles, or equivalently, the electric and magnetic susceptibility profiles, required for the metasurface to transform the input field into the output field. In contrast, the microscopic step is concerned with the physical implementation of this surface impedance profile with the design of appropriate subwavelength elements such as a triple-dogbone topology [61]. The macroscopic aspect is the focus of this work, and therefore the physical implementation of subwavelength elements are not considered in this paper.

To the best of our knowledge, all these synthesis methods rely on the knowledge of the tangential fields on the input and output surfaces of the metasurface. However, in some applications these tangential fields are not readily available. For example, consider the case where the designer would like to transform the radiation pattern of a horn antenna to a different radiation pattern with the following design specifications: main beam direction, half-power beamwidth (HPBW), and side lobe levels (SLL). The knowledge of these design specifications does not directly translate into the knowledge of a tangential field on the output surface of the metasurface. This case clearly presents a limitation for the existing metasurface synthesis methods. This is important since many antenna design applications rely on these types of specifications, and not on an explicit expression of a desired complex field.

3.1.1 Electromagnetic Inversion

As can be seen from the discussion above, there is a need for greater flexibility in the design of metasurfaces, and in particular the ability to design from a set of desired performance criteria (e.g., HPBW, null directions, etc.) is necessary. This paper proposes to cast the metasurface design problem as an electromagnetic inverse source problem. The main goal in electromagnetic inverse problems is to find an unknown *cause* from its known *effect* [62].

For example, the known effect could be a desired HPBW, and the cause to be found could be a set of electric and magnetic currents that generate a radiation pattern exhibiting such a HPBW. (Once these currents are found, the surface properties of the metasurface can be obtained.) The act of processing the effect to find its cause is often referred to as *inversion*, and the obtained cause is often referred to as the *reconstructed* result (e.g., the reconstructed electric current). Broadly speaking, electromagnetic inversion can be classified into two categories. The first category is inverse *scattering*, in which the goal is to reconstruct complex permittivity and/or permeability profiles that produce a known scattered field when illuminated with a known incident field. The second category, which is the one used in this work, is inverse *source*, in which the goal is to reconstruct a set of electric and/or magnetic currents that generate a known electromagnetic signature.

3.1.2 Electromagnetic Inversion for Characterization

There are various applications for electromagnetic inversion within a wide frequency spectrum (from a few hertz [63] to optical frequencies [64]), most of which relate to some form of *characterization*. In such applications, electromagnetic inverse scattering or inverse source algorithms are applied to *measured* data. For example, in antenna diagnostics [65, 66], the near-field data of an antenna under test are measured. From this measured data, the equivalent currents of the antenna can be reconstructed using an electromagnetic inverse source algorithm, and can then, for example, be used for finding faulty elements in an antenna array. As another example, in microwave breast imaging [67], the breast is illuminated from different directions, and the resulting scattered fields are measured. From these measured scattered data, the complex permittivity profile of the breast can be reconstructed using an electromagnetic inverse scattering algorithm, and can then be potentially used for detecting tumours. One of the main advantages of using the electromagnetic inversion framework is that it can work with various forms of data. For example, electromagnetic inversion

algorithms can invert phaseless (amplitude-only) near-field measurements [55]. In addition, they are not merely limited to canonical measurement surfaces (e.g., planar, cylindrical, spherical), and work with arbitrarily shaped measurement domains [68]. Moreover, electromagnetic inversion algorithms can systematically incorporate prior information into their formulation to enhance the resulting reconstruction accuracy [69].

3.1.3 Electromagnetic Inversion for Design

As noted above, the electromagnetic inversion framework has been mainly used for characterization applications. To adapt it to design applications, we mainly need to replace the *measured* data with *desired* data in the mathematical formulation, and then reconstruct a cause that can generate the desired effect. For example, in [70], it was suggested that we can use electromagnetic inverse scattering algorithms for the design of dielectric profile lenses. Inspired by this work, inverse scattering algorithms have been recently used for the design of cloaking devices and lenses [71, 72, 73]. Similarly, inverse source algorithms have also been suggested for design applications [74], and have been used, for example, in the design of shaped beam reflectarrays [75].

Inverse scattering and inverse source problems are inherently ill-posed. That is, the solution to the associated mathematical problem has at least one of these features: (i) non-uniqueness, (ii) non-existence, or (iii) instability [50]. The main reason for this non-uniqueness is non-radiating sources: current sources that may exist in the investigation domain but create no electromagnetic fields at the observation domain [51], i.e., creating a null-space for the mathematical problem. Although this non-uniqueness is typically a disadvantage in characterization, it is actually *advantageous* for design as it provides extra degrees of freedom. For example, in [75], the non-uniqueness due to non-radiating sources has been used to satisfy user-defined geometrical constraints in reflectarray design. On the other hand, non-existence of the solution does not occur in characterization applications,

while it may occur in design applications, e.g., trying to design a highly directive radiation pattern from an electrically small aperture. In such situations, the electromagnetic inversion may still find a compromising solution that partially meets the desired design specifications. For example, consider a bianisotropic lens that transforms a given incident field to a desired field. The inverse scattering algorithm allows us to limit the solution space to purely isotropic dielectric lenses. In this situation, the solution to the mathematical problem may not exist; however, the inverse scattering algorithm may still result in an acceptable solution [73]. Finally, the instability of the mathematical problem is related to the so-called smoothing effect [76] of the free space Green's function operator, which acts as a low-pass filter for high spatial frequency information. Mathematically, this instability occurs due to the very small singular values of the operator, and is treated by the use of appropriate regularization techniques which attempt to stabilize the problem by removing the effect of small singular values in the inversion process [77, 78].

We have recently proposed and used the electromagnetic inverse source framework for the design of metasurfaces [79, 80]. In addition to applying the inverse source framework to metasurface design, these works extend it to more practical scenarios: those where the desired specifications are expressed in terms of some performance criteria (e.g., HPBW and null directions) rather than a completely-known desired field pattern. In this paper, we explore this approach in detail. To this end, we start with a problem statement, then briefly describe our methodology, which is followed by a section motivating the pursuit of this approach. We then present some fundamentals of metasurfaces, followed by presenting the inverse source framework for design in the operator notation. The details of our inversion algorithms are finally presented followed by a brief description of the simulation algorithm. We then present some illustrative examples, explain our current limitations, and finally present our conclusions.

3.2 Problem Statement

Consider the geometry shown in Figure 3.1, where the surfaces Σ^- and Σ^+ represent the input and output surfaces of the metasurface, and V^- and V^+ represent the volumetric domains located on the input and output sides of the metasurface respectively.² Let us assume that the following items are known: (i) an incident field $\vec{\Psi}^{\text{inc}}$ where $\Psi \in \{E, H\}$ that impinges on the metasurface from the input side, (ii) a surface geometry Σ of arbitrary shape for the metasurface, and (iii) a set of desired field specifications, denoted by S^{des} in a region of interest (ROI) within the output side of the metasurface V^+ . For example, the ROI can be a far-field region defined within a solid angle, or can be a near-field region. Our objective is to find the required electric and magnetic surface susceptibility profiles for the metasurface to transform $\vec{\Psi}^{\text{inc}}$ to a transmitted field, say $\vec{\Psi}^{\text{tr}}$, that exhibits the desired field specifications S^{des} . It should be noted that in contrast to existing design methods for metasurfaces, the proposed method does not assume the knowledge of $\vec{\Psi}^{\text{tr}}$; it only assumes the more practical knowledge of S^{des} , which could be for example a set of performance criteria such as HPBW or null directions. In summary, as depicted in Figure 3.2, we aim to infer the unknown properties of a system from a known input and its corresponding known output.

²To precisely define V^- and V^+ , we need to assume that Σ^- and Σ^+ extend to infinity or are closed surfaces. However, the method presented herein relaxes this requirement.

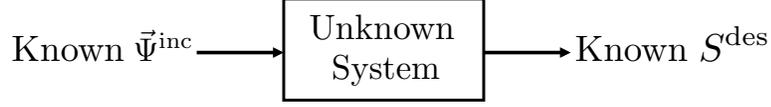


Fig. 3.2: The design problem from a system viewpoint. The unknown properties of the system (surface susceptibilities of a metasurface) are to be found from the knowledge of the input and output. The known input to the system is the incident field $\vec{\Psi}^{\text{inc}}$ and the known output is the desired specifications S^{des} . Note that S^{des} can be, for example, a set of performance criteria such as HPBW and null directions.

with a known antenna. In addition, the reflected field is assigned to be zero for reflectionless metasurfaces. Therefore, in this paper, we treat \vec{E}_t^- and \vec{H}_t^- as known quantities.

3.3.2 Finding Tangential Fields on the Output Surface

Let us denote the tangential electric and magnetic fields on Σ^+ by \vec{E}_t^+ and \vec{H}_t^+ respectively. The core of our methodology lies in inferring \vec{E}_t^+ and \vec{H}_t^+ from the knowledge of S^{des} . (Note that we do not assume the complete knowledge of $\vec{\Psi}^{\text{tr}}$; instead we merely assume the knowledge of S^{des} .) To this end, we use the electromagnetic inverse source framework to reconstruct required electric and magnetic currents (\vec{J} and \vec{M}) on Σ^+ from the knowledge of S^{des} . Once these currents are reconstructed, we can then find \vec{E}_t^+ and \vec{H}_t^+ . Having obtained the tangential electric and magnetic fields on Σ^- and Σ^+ , we can find the required surface electric and magnetic susceptibility profiles for the metasurface based on the generalized sheet transition conditions (GSTCs) [35, 36]. As will be discussed in Section 3.10, our current implementation does not enforce local power conservation or consider magnetoelectric coupling. Therefore, the resulting metasurfaces will include lossy and/or active elements.

3.4 Motivation

Let us now motivate the utilization of the electromagnetic inverse source framework for finding \vec{E}_t^+ and \vec{H}_t^+ . As will be seen, this has two main advantages: (i) the ability to work with arbitrarily-shaped geometries (in our case, various Σ and ROI geometries), and (ii) the

ability to handle various forms of S^{des} . To understand this, consider the following five cases.

3.4.1 Case I

If the desired output specification S^{des} is a refracted plane wave in V^+ , finding \vec{E}_t^+ and \vec{H}_t^+ is trivial as the expression for $\vec{\Psi}^{\text{tr}}$ is analytically known. (This is the case that has been mainly studied in the literature.) For this case, the electromagnetic inverse source framework is not needed.

3.4.2 Case II

Assume that the desired output specification S^{des} is given in terms of the desired complex (amplitude and phase) field on a canonical surface (ROI) within V^+ , and that the metasurface geometry (Σ) is of canonical shape. In such a case, \vec{E}_t^+ and \vec{H}_t^+ can be easily found by standard modal expansion algorithms. Such algorithms [81] are widely used in planar, cylindrical, and spherical near-field antenna measurements to back-propagate the measured data to the surface of the antenna under test for diagnostics. These algorithms can also be used to back-propagate the desired complex field data from the ROI to Σ^+ if both the ROI and Σ^+ are canonical shapes. Case II is in fact similar to Case I, and this can be justified based on the plane wave spectrum as follows. Assume that both Σ and ROI are parallel planar surfaces; then the modal expansion algorithm expands the desired complex field on the ROI in terms of a sum of plane waves (plane wave spectrum). We can then back-propagate all of these plane waves to the surface of the metasurface. Therefore, the use of electromagnetic inverse source algorithms is also not needed in Case II.³

³It should be noted that since in practical applications we are dealing with truncated surfaces (e.g., finite planes), the use of electromagnetic inverse source framework can still offer an advantage compared to the modal expansion algorithms (which assume infinite planes or closed surfaces). This will not be further discussed in this paper; see [82] for a discussion on this topic as applied to near-field antenna measurements.

3.4.3 Case III

Assume that the desired output specification S^{des} is given in terms of the complex field on an arbitrarily-shaped surface (ROI) within V^+ and/or the geometry of the metasurface (Σ) is of arbitrary shape [83]. This could, for example, occur when we aim to design conformal metasurfaces. Standard modal expansion algorithms cannot handle such geometrical scenarios. In this case, the use of electromagnetic inverse source algorithms is advantageous. (This advantage of the electromagnetic inverse source algorithms has been utilized in near-field antenna measurements, e.g., see [68].)

3.4.4 Case IV

Assume that the desired specification S^{des} is given in terms of amplitude-only (phaseless) fields. This could, for example, happen when the designer is interested in focusing the electromagnetic energy in a ‘hot spot’ for microwave hyperthermia applications, or to design a desired power pattern for antenna applications. For this case, even if Σ and the ROI are canonical surfaces, standard modal expansion algorithms are incapable of handling the phaseless nature of the data. One way to treat this is to assume a phase distribution; however, the phase assumption will automatically limit the number of achievable solutions. A better way is to avoid any phase assumptions and apply a phaseless electromagnetic inverse source algorithm to the data. We then typically obtain more solutions that satisfy the phaseless specifications, and single out the solution that results in the easiest physical implementation. (The use of phaseless electromagnetic inverse source algorithms has been previously considered in near-field antenna measurements to avoid the necessity of performing expensive phase measurements, e.g., see [55].)

3.4.5 Case V

Finally, case V aims at offering more flexibility to the designer. To this end, in Case V, the desired specifications S^{des} are some performance criteria rather than a field/power pattern. This is most practical for antenna engineering applications, e.g., setting the desired performance criteria to be the null directions or HPBW. This case cannot be handled by standard modal expansion algorithms, and we propose to use the electromagnetic inverse source framework to handle this case.

In summary, there are at least three cases that existing metasurface synthesis methods cannot handle, and where the use of the electromagnetic inverse source framework is therefore advantageous.

3.5 Metasurface Fundamentals

This section provides a brief overview of metasurface synthesis to show the importance of having the tangential field components on the input and output surfaces (Σ^+ and Σ^-) of the metasurface. The electromagnetic behaviour of a metasurface can be modelled by rigorous GSTCs, originally developed in [35]. The form that we adopt herein was developed in [36], and we follow the notation style of [23]⁴.

The electric and magnetic polarization densities can be expressed as

$$\vec{P}^e = \epsilon_0 \bar{\bar{\chi}}_{ee} \vec{E}_{\text{av}} + \bar{\chi}_{\text{em}} \sqrt{\mu_0 \epsilon_0} \vec{H}_{\text{av}}, \quad (3.1a)$$

$$\vec{P}^m = \bar{\chi}_{\text{mm}} \vec{H}_{\text{av}} + \bar{\chi}_{\text{me}} \sqrt{\frac{\epsilon_0}{\mu_0}} \vec{E}_{\text{av}}, \quad (3.1b)$$

where ϵ_0 and μ_0 are the permittivity and permeability of free space, respectively, and $\bar{\bar{\chi}}_{ee}$,

⁴For other homogenization models of metasurfaces (polarizability-based model and equivalent impedance matrix model) see Section 2.4 in [84].

$\bar{\bar{\chi}}_{\text{mm}}$, $\bar{\bar{\chi}}_{\text{em}}$, and $\bar{\bar{\chi}}_{\text{me}}$ represent the electric/magnetic (first subscript) surface susceptibility tensors describing the response to electric/magnetic (second subscript) field excitations [39].

The average fields on the surface Σ of the metasurface are defined as

$$\vec{\Psi}_{\text{av}} \triangleq \frac{\vec{\Psi}^{\text{tr}}|_{\Sigma^+} + (\vec{\Psi}^{\text{inc}}|_{\Sigma^-} + \vec{\Psi}^{\text{ref}}|_{\Sigma^-})}{2}. \quad (3.2)$$

We now assume that the normal components of the polarization densities are zero (i.e., $\hat{n} \cdot \vec{P}^e = \hat{n} \cdot \vec{P}^m = 0$, where \hat{n} is the outward unit vector normal to the metasurface as shown in Figure 3.1) to avoid the unnecessary complexity that would result from including spatial derivatives in the GSTCs [23]. This assumption does not result in any loss of generality, as the normal components of any field can be represented in terms of the corresponding tangential components (i.e., allowing non-zero normal components of the polarization densities would not introduce any independent solutions). Assuming a time-dependency of $e^{j\omega t}$, we can now construct a system of equations from the GSTCs in [23] and equations (3.1a) and (3.1b), relating the tangential electric and magnetic fields through the surface

susceptibilities as

$$\begin{aligned} \begin{pmatrix} -\Delta H_v \\ \Delta H_u \end{pmatrix} &= j\omega\epsilon_0 \begin{pmatrix} \chi_{ee}^{uu} & \chi_{ee}^{uv} \\ \chi_{ee}^{vu} & \chi_{ee}^{vv} \end{pmatrix} \begin{pmatrix} E_{u,av} \\ E_{v,av} \end{pmatrix} \\ &+ j\omega\sqrt{\epsilon_0\mu_0} \begin{pmatrix} \chi_{em}^{uu} & \chi_{em}^{uv} \\ \chi_{em}^{vu} & \chi_{em}^{vv} \end{pmatrix} \begin{pmatrix} H_{u,av} \\ H_{v,av} \end{pmatrix}, \end{aligned} \quad (3.3a)$$

$$\begin{aligned} \begin{pmatrix} -\Delta E_u \\ \Delta E_v \end{pmatrix} &= j\omega\mu_0 \begin{pmatrix} \chi_{mm}^{vv} & \chi_{mm}^{vu} \\ \chi_{mm}^{uv} & \chi_{mm}^{uu} \end{pmatrix} \begin{pmatrix} H_{v,av} \\ H_{u,av} \end{pmatrix} \\ &+ j\omega\sqrt{\epsilon_0\mu_0} \begin{pmatrix} \chi_{me}^{vv} & \chi_{me}^{vu} \\ \chi_{me}^{uv} & \chi_{me}^{uu} \end{pmatrix} \begin{pmatrix} E_{v,av} \\ E_{u,av} \end{pmatrix}, \end{aligned} \quad (3.3b)$$

where the subscripts and superscripts u and v refer to the tangential components of the local coordinate system defined by $\hat{u} \times \hat{v} = \hat{n}$ and $\hat{u} \perp \hat{v}$. The Δ operator denotes the difference between the fields on either side of the metasurface, defined in terms of the transmitted, incident, and reflected fields as

$$\Delta \vec{\Psi} \triangleq \vec{\Psi}^{\text{tr}} - (\vec{\Psi}^{\text{inc}} + \vec{\Psi}^{\text{ref}}). \quad (3.4)$$

For a single field transformation (i.e., one combination of incident, reflected, and transmitted fields), the system in (3.3) is clearly underdetermined. It should also be noted that solving (3.3) for a desired field transformation does not guarantee that the resulting susceptibilities are physically practical or realizable. Avoiding these impractical solutions requires enforcing additional constraints and/or modifying the generated solution. Finally, we note that all the field quantities in (3.3a) and (3.3b) can be found based on the knowledge of

\vec{E}_t^+ , \vec{H}_t^+ , \vec{E}_t^- , and \vec{H}_t^- . For example,

$$\Delta E_v = \hat{v} \cdot (\vec{E}_t^+ - \vec{E}_t^-), \quad (3.5)$$

$$H_{u,av} = \hat{u} \cdot \left(\frac{\vec{H}_t^+ + \vec{H}_t^-}{2} \right). \quad (3.6)$$

Therefore, as noted in Section 3.3, having the knowledge of \vec{E}_t^+ , \vec{H}_t^+ , \vec{E}_t^- , and \vec{H}_t^- enables us to find the required susceptibility profiles.

3.6 Inverse Source Framework

Herein, we extend Section 3.3 by detailing how the inverse source framework for metasurface design is formulated. This is explained in the following seven steps.

3.6.1 Preparing the Data to be Inverted

The first step in the development of the inverse source design formulation is to ensure that the desired specification S^{des} is in a proper form for use by inverse source algorithms. We consider the following three scenarios. (i) If S^{des} is in the form of complex fields, we directly use it in the inverse source formulation. (ii) If S^{des} is in the form of power patterns (phaseless), we also directly use it in the formulation. (iii) If S^{des} is in the form of some performance criteria (e.g., main beam direction, HPBW and null directions), we first form a weighted normalized power pattern (phaseless) that satisfies the required performance criteria. We will then use this pattern in the inverse source algorithm. (This step will be covered in more details in Section 3.7.3.) Therefore, in summary, the data to be used in the inversion algorithm can be represented in the form of $\mathcal{K}(S^{\text{des}})$, where \mathcal{K} is the operator that converts S^{des} to a data set that can be used by the inversion algorithm.

3.6.2 Creating the Data Operator

The so-called data operator, denoted by \mathcal{G} , takes a given set of electric and magnetic currents (\vec{J} and \vec{M}) on Σ^+ and outputs the corresponding complex fields on the ROI, i.e., $\vec{\Psi}^{\text{tr}}|_{\text{ROI}} = \mathcal{G}(\vec{J}, \vec{M})$. In our case, this operator is constructed based on the electric field integral equation (EFIE) [85]. In the three-dimensional (3D) case, the operator \mathcal{G} can be expressed as [86]

$$\begin{aligned} \mathcal{G}(\vec{J}, \vec{M}) = & -j\eta_0 k_0 \int_{\Sigma^+} \left[\vec{J}(\vec{r}') + \frac{1}{k_0^2} \nabla \nabla'_s \cdot \vec{J}(\vec{r}') \right] g(\vec{r}, \vec{r}') ds' \\ & - \nabla \times \int_{\Sigma^+} \vec{M}(\vec{r}') g(\vec{r}, \vec{r}') ds' \end{aligned} \quad (3.7)$$

where the position vectors \vec{r} and \vec{r}' belong to the ROI and Σ^+ respectively. In addition, $g(\cdot, \cdot)$ denotes the Green's function, and ' $\nabla'_s \cdot$ ' is the surface divergence operator with respect to the prime coordinate (i.e., position vector on Σ^+). Finally, η_0 and k_0 are the wave impedance and wavenumber in free space.⁵

3.6.3 Forming the Data Misfit Cost Functional

For various sets of \vec{J} and \vec{M} on Σ^+ , we need to compare the simulated effects with the desired effect to converge at an appropriate set of equivalent currents. In other words, we need to compare $\mathcal{G}(\vec{J}, \vec{M})$ with $\mathcal{K}(S^{\text{des}})$ to evaluate how appropriate a set of (\vec{J}, \vec{M}) is. Before doing so, we first need to make sure that $\mathcal{G}(\vec{J}, \vec{M})$ is in the form that is consistent with $\mathcal{K}(S^{\text{des}})$. For example, for the power pattern synthesis problem, we need to compare $|\mathcal{G}(\vec{J}, \vec{M})|^2$ with $\mathcal{K}(S^{\text{des}})$. To this end, let us assume that the operator \mathcal{L} takes $\mathcal{G}(\vec{J}, \vec{M})$ and outputs the form that is consistent with $\mathcal{K}(S^{\text{des}})$. We then form a data misfit cost functional

⁵It should be noted that the use of Green's function formulation enables us to perform the design in other media which are not free space, assuming that the Green's function in those media are known either analytically or numerically.

that is the L_2 -norm discrepancy between the desired effect $\mathcal{K}(S^{\text{des}})$ and the simulated effect $\mathcal{L}(\mathcal{G}(\vec{J}, \vec{M}))$ for a given electric and magnetic currents.⁶ The data misfit cost functional \mathcal{C} is then

$$\mathcal{C}(\vec{J}, \vec{M}) = \left\| \mathcal{L}(\mathcal{G}(\vec{J}, \vec{M})) - \mathcal{K}(S^{\text{des}}) \right\|_2^2, \quad (3.8)$$

where $\|\cdot\|_2$ denotes the L_2 -norm which is defined over the ROI at which the desired specifications are defined.⁷

3.6.4 Enforcing Love's Equivalence Condition

Based on the electromagnetic equivalence principle [88], when solving for \vec{J} and \vec{M} on Σ^+ from the knowledge of S^{des} we have the following two general options regarding the fields internal to Σ^+ : (i) no assumptions, or (ii) assuming a particular field such as a null field (Love's equivalence condition). It was shown in [86] for antenna diagnostics applications that the inverse source problem can result in different sets of \vec{J} and \vec{M} solutions based on different assumptions regarding the inner side of the reconstruction surface. However, they generate similar external fields on the ROI (hence the non-uniqueness of the inverse problem). Herein, we apply Love's equivalence condition when solving for \vec{J} and \vec{M} . To enforce this condition, we create a virtual surface that is inwardly offset with respect to Σ^+ , and then enforce zero tangential electric fields on this surface when solving for \vec{J} and \vec{M} . (For details, see [86].)

⁶While the L_2 -norm is a common choice for cost functional formulations in inverse problems, other options do exist [87]. For example, using the L_1 -norm would increase the relative weighting of smaller data-misfit values, which may be preferred in some applications. In our approach, this has been addressed by the use of the weighting matrix \mathbf{W} discussed in Section 3.7.3.

⁷As will be seen in Section 3.7.3, we may add an extra cost functional to the data misfit cost functional to enforce additional specifications.

3.6.5 Minimizing the Data Misfit Cost Functional

To reconstruct an appropriate set of equivalent electric and magnetic currents, we minimize the data misfit cost functional, i.e.,

$$\text{appropriate } (\vec{J}, \vec{M}) = \underset{\vec{J}, \vec{M}}{\operatorname{argmin}} \left\{ \mathcal{C}(\vec{J}, \vec{M}) \right\}. \quad (3.9)$$

To this end, we use the conjugate gradient (CG) algorithm [89]. In this algorithm, \vec{J} and \vec{M} are iteratively updated as

$$(\vec{J}_{k+1}, \vec{M}_{k+1}) = (\vec{J}_k, \vec{M}_k) + \beta_k \mathbf{v}_k \quad (3.10)$$

where $\beta_k \in \mathbb{R}$ is the step length and \mathbf{v}_k is the conjugate gradient direction. To find \mathbf{v}_k , we need to know the gradient of the cost functional with respect to the currents at the k^{th} iteration which is denoted by \mathbf{g}_k .

3.6.6 Regularization

When minimizing (3.9), we need to ensure that the instability of the ill-posed problem is properly handled. The methods used to treat this instability are often referred to as regularization methods [50, 90, 77], and the methods that control the regularization weight are often called regularization parameter choice methods [91]. One common method in treating this instability is to augment the data misfit cost functional by a penalty cost functional, e.g., by the L_2 -norm of \vec{J} and \vec{M} . This method is referred to as Tikhonov regularization, which has been widely used in various applications. Another regularization method is to use an iterative algorithm such as the conjugate gradient least squares method and limit (truncate) the number of iterations [92],[78]. Herein, we use this truncated CG approach,

although the truncation point is chosen in an *ad hoc* way in the current implementation.⁸

3.6.7 Finding the Susceptibility Profiles

Having found the equivalent currents, and noting that Love’s equivalence condition has been applied in solving the inverse source problem, we can easily obtain the required tangential fields on Σ^+ from the reconstructed currents as

$$\vec{H}_t^+ = -\hat{n} \times \vec{J} \quad \text{and} \quad \vec{E}_t^+ = \hat{n} \times \vec{M}. \quad (3.11)$$

As noted in Section 3.3.1, we assume that we already know \vec{E}_t^- and \vec{H}_t^- . Now that \vec{E}_t^+ , \vec{H}_t^+ , \vec{E}_t^- , and \vec{H}_t^- are known, we can use (3.3) to find the susceptibility profiles. This completes the description of our inverse source framework for design.

3.7 Inversion Algorithm Implementation

The previous section outlined the main steps for the inverse source framework for metasurface design. Herein, we explain how the inversion algorithm is implemented in the discrete domain. To simplify the discussion, consider the following specific three scenarios, where \mathbf{x} denotes the (unknown) concatenated vector of equivalent electric and magnetic currents at the discretized surface Σ^+ . As described in Section 3.6.7, once \mathbf{x} is reconstructed, the required susceptibilities can be found.

⁸The main reasons for selecting the conjugate gradient technique is that it can introduce additional regularization and its computational efficiency [50, 93]. For example, for linear problems, at the k^{th} iteration of CG, the solution space is constrained to a Krylov subspace of dimension k . Constraining the solution of ill-posed problems to such subspaces in early iterations improves the solution, as most noise would also require a higher-dimensional basis and is therefore excluded. However, in later iterations, the solution can begin to deteriorate as the dimension of the solution space increases and allows for noise to be inverted. This behaviour is known as *semi-convergence* [50, 94], and highlights the importance of selecting a proper truncation point for the iterations. In this work we employ an *ad-hoc* approach of truncating the iterations when the solution stagnates to a certain degree, but standard regularization parameter-choice methods could be used as well (e.g., the *L*-curve method [95]).

3.7.1 Scenario I

For this scenario, assume that the desired specifications S^{des} are a set of complex field values which can be either in the near-field or far-field zones. That is, $S^{\text{des}} = \mathbf{f}$ where \mathbf{f} is a vector of complex electric (or, magnetic field) values representing the desired field at the ROI. Then, the cost functional (in the discrete domain) is simply

$$\mathcal{C}(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{f}\|_2^2 \quad (3.12)$$

where \mathbf{A} is the discretized integral operator that when operating on a given \mathbf{x} produces the fields at the locations of interest on the ROI. (We note that this formulation is not restricted to canonical ROI and Σ^+ .) To enforce Love's condition (see Section 3.6.4), we augment \mathbf{f} by a number of zeros as $\mathbf{f} = [\mathbf{f}; \mathbf{0}]$, where the operator “;” denotes vector concatenation and $\mathbf{0}$ denotes a column vector of zeros whose length is the same as the number of points at which we enforce null fields at the virtual Love's surface. Once \mathbf{f} is augmented, the matrix \mathbf{A} needs to be augmented accordingly. To reconstruct \mathbf{x} , minimization of (3.12) over \mathbf{x} is performed using the CG method, which is identical to CG minimization commonly used in the source reconstruction method for near-field antenna measurements [85].

3.7.2 Scenario II

In this scenario, a specific far-field power pattern (phaseless quantity) is desired. That is, $S^{\text{des}} = |\mathbf{f}|^2$ where $|\mathbf{f}|^2$ is a vector of the desired squared field amplitude (i.e., power) at the ROI. Then, the cost functional becomes

$$\mathcal{C}(\mathbf{x}) = \left\| |\mathbf{A}\mathbf{x}|^2 - |\mathbf{f}|^2 \right\|_2^2 \quad (3.13)$$

where $|\mathbf{A}\mathbf{x}|^2$ denotes the magnitude-squared of the simulated field due to a predicted \mathbf{x} . In contrast to (3.12), the cost functional (3.13) is nonlinear due to the absence of phase

information.⁹ This cost functional is minimized using the CG algorithm where the gradient vector at the k^{th} iteration (evaluated at \mathbf{x}_k) can be derived as [96, Section 4.2.1]

$$\mathbf{g}_k = 2\mathbf{A}^H [\mathbf{r}_k \odot \mathbf{A}\mathbf{x}_k]. \quad (3.14)$$

The superscript H denotes the Hermitian (complex conjugate transpose) operator, $\mathbf{r}_k = |\mathbf{A}\mathbf{x}_k|^2 - |\mathbf{f}|^2$ is the residual vector and \odot represents a Hadamard (element-wise) product. Once the gradient vector \mathbf{g}_k is found, the conjugate gradient direction \mathbf{v}_k can be obtained. The step length β_k at each CG iteration can also be found by minimizing $\mathcal{F}(\beta_k) = \mathcal{C}(\mathbf{x}_k + \beta_k\mathbf{v}_k)$ with respect to β_k [96, Section 4.2.2]. This cost functional is a fourth-degree polynomial with respect to β_k as

$$\begin{aligned} \mathcal{F}(\beta_k) &= \beta_k^4 \left\| |\mathbf{A}\mathbf{v}_k|^2 \right\|_2^2 + 4\beta_k^3 \operatorname{Re} \langle |\mathbf{A}\mathbf{v}_k|^2, (\mathbf{A}\mathbf{x}_k) \odot (\mathbf{A}\mathbf{v}_k)^* \rangle \\ &\quad + 2\beta_k^2 \left[\operatorname{Re} \langle \mathbf{r}_k, |\mathbf{A}\mathbf{v}_k|^2 \rangle + \left\| \operatorname{Re}((\mathbf{A}\mathbf{x}_k) \odot (\mathbf{A}\mathbf{v}_k)^*) \right\|_2^2 \right] \\ &\quad + 4\beta_k \operatorname{Re} \langle \mathbf{r}_k, (\mathbf{A}\mathbf{x}_k) \odot (\mathbf{A}\mathbf{v}_k)^* \rangle + \|\mathbf{r}_k\|_2^2 \end{aligned} \quad (3.15)$$

where Re is the real-part operator, $\langle \cdot, \cdot \rangle$ denotes the inner product operator, and the superscript $*$ denotes the complex conjugate operator. To find the step length β_k , we differentiate (3.15) with respect to β_k , and then numerically solve for the roots of the resulting third-degree polynomial, which has a pair of complex conjugate roots and one real root. The step length will then be the real root.

⁹It is instructive to note that when inverse source algorithms are used in phaseless near-field antenna measurement applications, the phaseless data $|\mathbf{f}|^2$ often need to be measured in two different regions so that the phase information can be implicitly derived using the correlation of the two phaseless data sets. However, in design applications, this is not necessary since we do not need to worry about the true phase data. In fact, we are only interested in a solution that satisfies the amplitude-only requirement.

3.7.3 Scenario III

This scenario considers a more practical case in which the desired field is only known in terms of certain far-field performance criteria, such as main beam direction, HPBW, and null directions. (We will consider polarization later in this section.) Thus, for example,

$$S^{\text{des}} = \{\text{main beam directions, HPBWs, nulls}\}. \quad (3.16)$$

Creating a Normalized Power Pattern from Performance Criteria

The above desired performance criteria first need to be converted to a form that is usable by the inverse source algorithm, which is a normalized power pattern.¹⁰ To this end, we form a power vector denoted by $|\mathbf{f}|^2$. In the direction of each main beam, the associated element of $|\mathbf{f}|^2$ is set to a value of 1, representing the normalized maximum power of the field. Nulls are represented in $|\mathbf{f}|^2$ by specifying the desired power level of the null relative to the main beam (e.g., a value of 10^{-3} would indicate a desired null at 30 dB below the main beam). The desired HPBWs for each main beam are enforced by approximating the main beam(s) using a cosine distribution. Desired power values are then interpolated using this approximation in a symmetric fashion around each main beam, resulting in corresponding elements in $|\mathbf{f}|^2$ ranging from 0.5 (at the HPBW limit) to 1 at the main beam. As an example, consider a desired far-field power pattern with a main beam at $\theta = -18^\circ$, a HPBW of 36° , and a -30 dB null at 40° away from the main beam. For a 2D problem, this set of performance criteria would result in $|\mathbf{f}|^2$ taking the normalized power values shown in Figure 3.3.

Balancing the Normalized Power Pattern

As can be seen in Figure 3.3, the quantitative values of the nulls can be significantly smaller than those of the main beams. In such situations, the inversion algorithm may favour main

¹⁰Normalization of the power pattern is not absolutely necessary, but allows for easier balancing if multiple cost functional terms are involved.

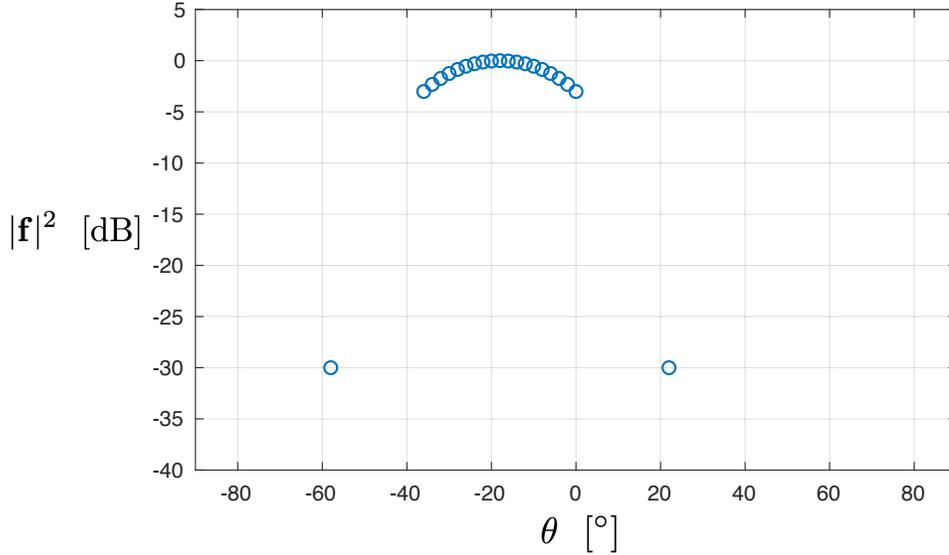


Fig. 3.3: The normalized power pattern elements of $|\mathbf{f}|^2$ for a desired 2D far-field pattern with a main beam at $\theta = -18^\circ$, a HPBW of 36° , and a -30 dB null at 40° away from the main beam. Note the two blue circles at -30 dB indicating the desired nulls.

beam directions and HPBWs as compared to null directions. To avoid this scenario, we balance $|\mathbf{f}|^2$ by the use of a diagonal ‘weighting’ matrix denoted by \mathbf{W} . Now instead of inverting $|\mathbf{f}|^2$, we invert $\mathbf{W}|\mathbf{f}|^2$. The elements of \mathbf{W} are selected to increase the relative contributions of the desired nulls to the functional, which can often be rendered insignificant compared to the main beam(s). More specifically, the i th diagonal element of \mathbf{W} , denoted by w_i , is computed as

$$w_i = \begin{cases} \frac{\zeta}{|f_i|^2}, & |f_i|^2 \leq \gamma \\ 1, & \text{otherwise} \end{cases} \quad (3.17)$$

where γ is a real positive threshold parameter used to determine which data ($|f_i|^2$) is classified as a ‘null’ and ζ is a real positive parameter that controls the amount of weighting applied, relative to the main beam.¹¹

¹¹This balancing idea has also been used in near-field antenna measurements [66] and microwave imaging applications [97].

Creating a Data Misfit Cost Functional

The data misfit cost functional is formed as

$$\mathcal{C}_1(\mathbf{x}) = \left\| \mathbf{W} \left(|\mathbf{A}_{\xi_1} \mathbf{x}|^2 + |\mathbf{A}_{\xi_2} \mathbf{x}|^2 \right) - \mathbf{Wf} \right\|_2^2. \quad (3.18)$$

Note that since these performance criteria are far-field quantities, $|\mathbf{A}\mathbf{x}|^2$ has been written as the summation of two terms: $|\mathbf{A}_{\xi_1} \mathbf{x}|^2$ and $|\mathbf{A}_{\xi_2} \mathbf{x}|^2$ where the subscripts ξ_1 and ξ_2 denote two *locally* orthogonal polarizations of the field.

Enforcing the Polarization Specification

Assume the desired specification S^{des} also includes a polarization constraint. This can be handled by adding the following cost functional to (3.18)

$$\mathcal{C}_2(\mathbf{x}) = \|\mathbf{A}_{\text{X-pol}} \mathbf{x}\|_2^2 \quad (3.19)$$

where $\mathbf{A}_{\text{X-pol}}$ is the discrete integral operator related to the undesired polarization (the cross-polarized component) of the field. Therefore, the total cost functional may now be represented as

$$\mathcal{C}(\mathbf{x}) = \mathcal{C}_1(\mathbf{x}) + \tau \mathcal{C}_2(\mathbf{x}) \quad (3.20)$$

where τ is a positive real parameter that controls the relative weight of the two cost functionals.

Minimization

To minimize (3.20) using the CG algorithm, the gradient vector at the k^{th} iteration can be obtained as

$$\begin{aligned} \mathbf{g}_k = & 2\mathbf{A}_{\xi 1}^H (\mathbf{W}^H \mathbf{r}_{1,k} \odot \mathbf{A}_{\xi 1} \mathbf{x}_k) + 2\mathbf{A}_{\xi 2}^H (\mathbf{W}^H \mathbf{r}_{1,k} \odot \mathbf{A}_{\xi 2} \mathbf{x}_k) \\ & + 2\tau \mathbf{A}_{\text{X-pol}}^H \mathbf{A}_{\text{X-pol}} \mathbf{x}_k \end{aligned} \quad (3.21)$$

where $\mathbf{r}_{1,k}$ is equal to the quantity inside the norm of (3.18) evaluated at \mathbf{x}_k . A closed-form expression for the step length is not available, so instead an appropriate step length is found numerically using a technique known as Armijo's condition [98]. It should be noted that there are several tuning parameters involved in Armijo's condition that affect the convergence of the minimization process. The stability of the iterative method can be very sensitive to these parameters, and finding appropriate values can often be a challenge. Another difference when solving (3.20) compared with the two previous scenarios is how Love's condition is applied. The functional in (3.20) is first minimized *without* Love's condition to generate an initial guess for the equivalent currents that is then used to solve (3.20) again *with* Love's condition. We have found this two step procedure to greatly improve the convergence and stability of the solution. Lastly, it should be noted that the functionals in (3.18) and (3.19) are for the general case with fields radiating in 3D. For 2D problems, $\mathbf{A}_{\xi 2}$ is simply set to zero and the same formulation can be used as written. In a similar fashion, $\mathbf{A}_{\text{X-pol}}$ is set to zero if a desired polarization is not required.

3.8 FDFD-GSTC Solver

Once we have reconstructed \mathbf{x} and have subsequently obtained the required surface susceptibilities, we need to verify that they can, in fact, transform the incident field to a transmitted field that satisfies the required specifications S^{des} . To do this for 2D examples, we use a

finite difference frequency domain (FDFD) GSTC solver, which is referred to as the FDFD-GSTC solver. This solver is an extension of the standard FDFD formulation with added boundary conditions to enforce the GSTC relationships. For details of the FDFD-GSTC solver, see [39].

It should be noted that the FDFD-GSTC solver is currently unable to simulate active metasurface elements (i.e., susceptibilities that correspond to gain). To overcome this limitation, we scale the E_t^+ and H_t^+ achieved in (3.11), which is justified by noting that any scaled forms of (3.11) still satisfy S^{des} . Therefore, in our implementation, we scale both \vec{H}_t^+ and \vec{E}_t^+ by the coefficient α (i.e., $\vec{H}_t^+ = -\alpha\hat{n} \times \vec{J}$ and $\vec{E}_t^+ = \alpha\hat{n} \times \vec{M}$) to ensure that no active elements are required (resulting in increased loss). Finally, we note that since our current implementation of the FDFD-GSTC solver is limited to 2D problems, for 3D problems, we use a 3D forward EFIE electromagnetic solver to verify our reconstruction results. However, as opposed to the FDFD-GSTC solver, our 3D EFIE solver does not incorporate the surface susceptibilities into its formulation, and merely calculates the effects of the reconstructed currents.

3.9 Illustrative Examples

This section presents several 2D and 3D synthetic examples demonstrating the flexibility of the inverse source method in dealing with different forms of desired specifications. Although not a requirement, in the examples presented here we consider the design of reflectionless metasurfaces and thus enforce $\vec{\Psi}^{\text{ref}} = 0$.

3.9.1 2D Examples

We first consider the design of several metasurfaces to transform an incident TE_z wave, with fields propagating in 2D in the xy plane at 10 GHz. The metasurface is 1D in this case, placed along the line $x = 0$ with elements (unit cells) that are $\lambda/6$ in length for each of

the examples. (λ denotes the wavelength.) Since all fields are TE_z , the system in (3.3) can be simplified and the metasurface can be represented by only four unknown susceptibility terms. Furthermore, we stipulate that there is no magnetoelectric coupling for simplicity, resulting in $\bar{\chi}_{\text{em}} = \bar{\chi}_{\text{me}} = 0$ [39]. With this choice, the unknown susceptibility tensors reduce to

$$\chi_{\text{ee}}^{yy} = \frac{2}{j\omega\mu_0} \frac{E_y^{\text{inc}} + E_y^{\text{ref}} - E_y^{\text{tr}}}{H_z^{\text{inc}} + H_z^{\text{ref}} + H_z^{\text{tr}}} \Big|_{\text{on the metasurface}}, \quad (3.22a)$$

$$\chi_{\text{mm}}^{zz} = \frac{2}{j\omega\epsilon_0} \frac{H_z^{\text{inc}} + H_z^{\text{ref}} - H_z^{\text{tr}}}{E_y^{\text{inc}} + E_y^{\text{ref}} + E_y^{\text{tr}}} \Big|_{\text{on the metasurface}}. \quad (3.22b)$$

Note that in the above equations, all of the field components are evaluated tangential to the surface of the metasurface (i.e., $x = 0$). Also, as explained earlier, E_y^{tr} and H_z^{tr} are not necessarily known on the surface of the metasurface, thus, justifying the use of the inverse source framework. In each example, the tangential components of the desired transmitted field (i.e., E_y^{tr} and H_z^{tr}) are found using the inverse source step of the design procedure, and then (3.22a) and (3.22b) are used to compute the required χ_{ee}^{yy} and χ_{mm}^{zz} to support the transformation from a specified incident field with no reflections. The 2D examples are then verified using the FDFD-GSTC solver, which simulates the interaction between the incident field and the susceptibilities in a solution domain of 20λ by 30λ in the xy plane. The solution domain uses a discretization size of $\lambda/50$ and is bounded by a perfectly matched layer (PML) of thickness λ on all sides.

Complex field

For the first example, we attempt to produce a desired near-field (NF) distribution using complete (amplitude and phase) field information over a given ROI. The near-field distribution shown in Figure 3.4 is generated from a pyramidal horn antenna in ANSYS HFSS at 10 GHz, and we use the complex field values (H_z) along the intersecting dashed white

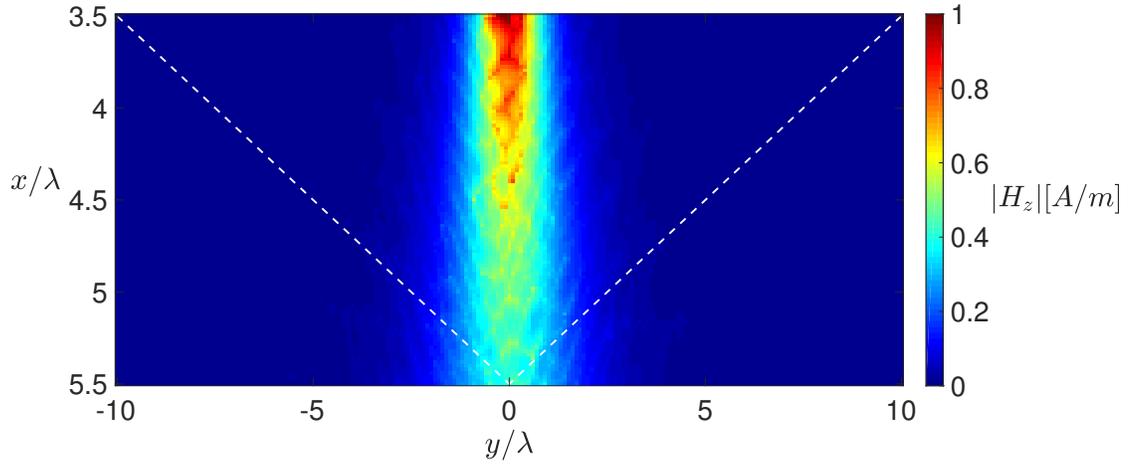


Fig. 3.4: The desired NF distribution which is generated from a pyramidal horn in ANSYS HFSS. The amplitude and phase of H_z are extracted along the dashed white lines (ROI) and used as input (S^{des}) for the metasurface design.

lines as input to the inverse source algorithm. That is, in this case, the intersecting dashed white lines are the ROI, and the complex near-field H_z data on these white lines are S^{des} . Moreover, intentionally, the two white lines have been chosen to represent an ROI which is not of canonical shape. The designed portion of the metasurface extends from $y = -5\lambda$ to $y = 5\lambda$ along the line $x = 0$, with absorbing susceptibilities placed along the remainder of the $x = 0$ line. The surface over which we reconstruct the equivalent currents coincides with the metasurface boundary Σ and is discretized into elements of length $\lambda/6$. Love's equivalence condition is enforced along the line $x = -\lambda/10$ with a resolution of $\lambda/9$. The proposed method is used to first find equivalent currents J_y and M_z on Σ that produce the desired near-field by minimizing (3.12), and then (3.11) is used to calculate the desired tangential transmitted fields on the metasurface. The incident field in each of the 2D examples is a normally incident TE_z plane wave, which is 'tapered' along y to avoid interactions with the PML that would introduce numerical error. (The H_z amplitude of the plane wave is 1 A/m for $|y| \leq 7\lambda$ and linearly decreases to zero for $7\lambda < |y| \leq 10\lambda$.) The necessary susceptibilities for the transformation (assuming no reflection) are then found using (3.22a)

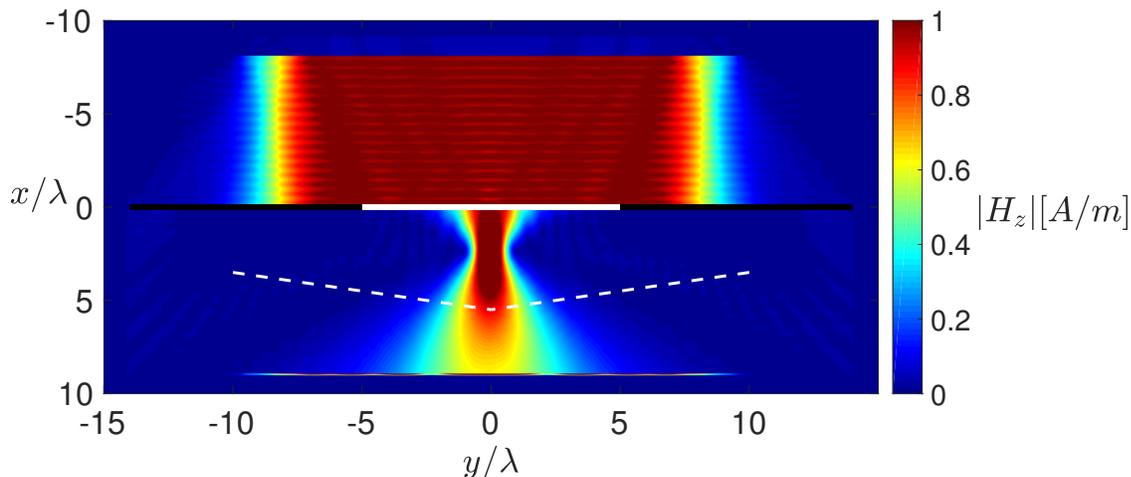


Fig. 3.5: The magnitude of H_z in the solution domain when the metasurface designed for the near-field distribution is illuminated by a normally incident tapered plane wave, simulated using the FDFD-GSTC solver. The designed metasurface extends from $y = -5\lambda$ to $y = 5\lambda$ along $x = 0$ as indicated by the solid white line, while absorbing susceptibilities have been added along $x = 0$ for $|y| > 5\lambda$ as shown by the solid black line. The dashed white lines (ROI) indicate the location where the desired near-field H_z is specified.

and (3.22b). The total field produced by simulating the designed metasurface and the incident field with the FDFD-GSTC solver is shown in Figure 3.5. The efficiency, calculated as the ratio of power transmitted through the metasurface to the power incident on the metasurface, is 18.6% for this example, thus, showing the presence of lossy regions within the metasurface.¹² Figure 3.6 compares the desired (red) and produced (blue) amplitude and phase of H_z along the dashed white lines, parameterized along y only. As can be seen from the plot, both the desired amplitude and phase of H_z have been produced with excellent accuracy.¹³

¹²The efficiency of the examples in this chapter can be compared to the examples in Chapters 4 and 5 in which we design passive and lossless metasurfaces with transmission efficiency $> 70\%$.

¹³The accuracy agreement has been evaluated based on visual inspection.

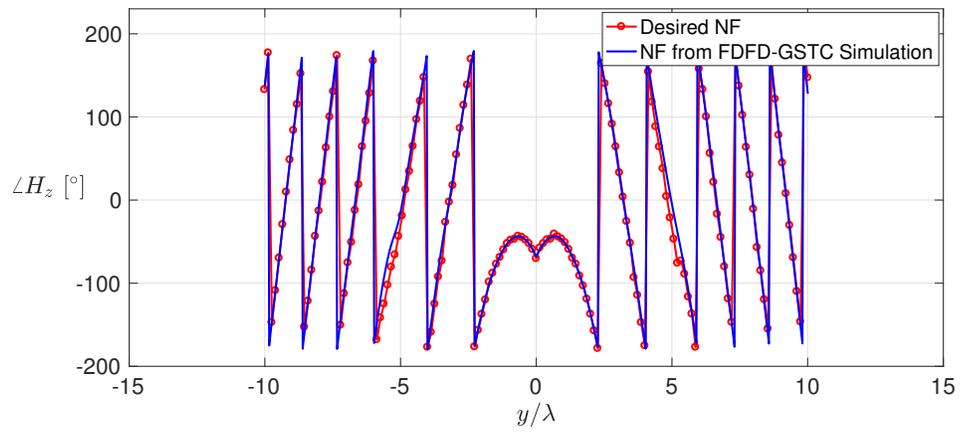
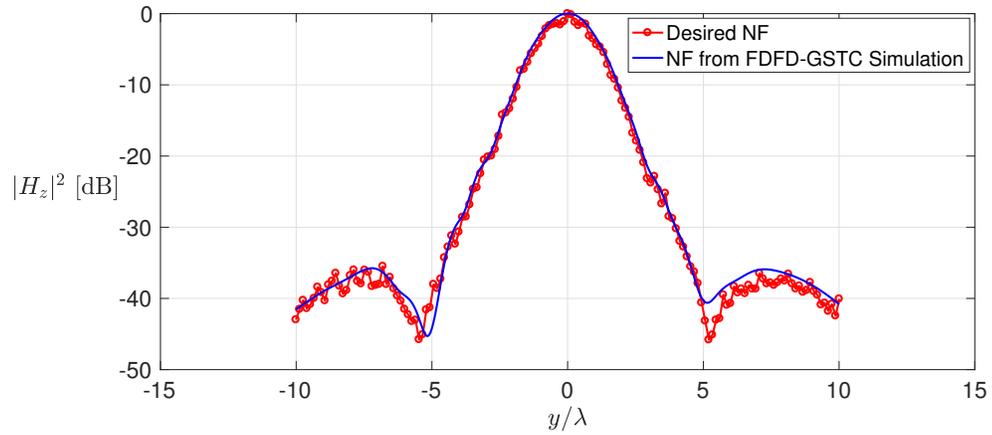


Fig. 3.6: A comparison of the desired normalized near-field (NF) distribution (solid red curve with circular markers) and the near-field distribution produced by the designed metasurface using both amplitude and phase information (solid blue curve).

Phaseless Power Pattern

The second example consists of attempting to produce a desired power pattern (phaseless) in the far-field (FF) region. In this example, the metasurface dimensions and incident field are the same as the previous example (relative to wavelength), with a frequency of operation of 1 GHz. The desired far-field power pattern for this example has been generated using a uniformly spaced array of 13 \hat{z} -directed elementary dipoles placed along the line $x = 0$ from $y = -3\lambda$ to $y = 3\lambda$. The squared amplitude (power) of H_z produced by this array is computed on a semicircular domain of radius 500λ for $-90^\circ \leq \varphi \leq 90^\circ$ as shown in Figure 3.7 (solid red curve with circular markers). The required field components E_y^{tr} and H_z^{tr} (tangential to the metasurface) are found by minimizing (3.13) and applying (3.11), and then the required susceptibilities are found using (3.22a) and (3.22b). To evaluate the inversion performance, the resulting susceptibilities are simulated using the FDFD-GSTC solver. The total field in the solution domain is shown in Figure 3.8 to demonstrate the nearly reflectionless performance of the metasurface. The far-field power pattern produced by the incident field passing through the metasurface is shown in Figure 3.7 (solid blue curve), with an efficiency of 17.0%. As can be seen from the plot, the produced far-field power pattern exhibits very good agreement with the desired power pattern.

Performance Criteria

The goal of the third example is to satisfy a set of desired far-field performance criteria. These specifications consist of two main beams with associated HPBWs and nulls as shown in Table 3.1. The metasurface and incident field in this case are the same as the previous example for the phaseless power pattern. The desired normalized power pattern created based on the desired performance criteria is shown in Figure 3.9 by red circular markers. In particular, note the three circular markers at -60 dB level indicating the desired null directions. The inversion algorithm finds the required equivalent currents by minimizing

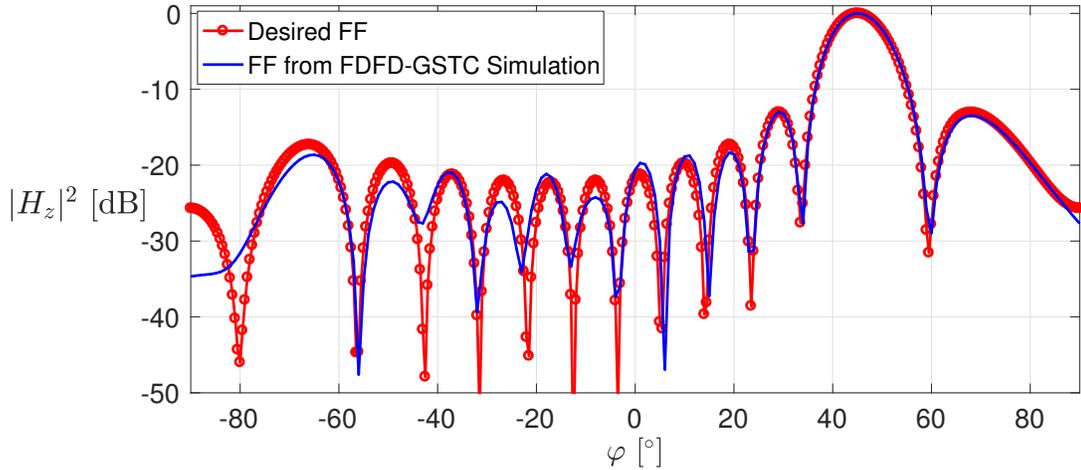


Fig. 3.7: A comparison of the normalized far-field (FF) power pattern (phaseless) produced by the FDFD-GSTC simulation of the designed metasurface (solid blue curve) and the desired power pattern (solid red curve with circular markers).

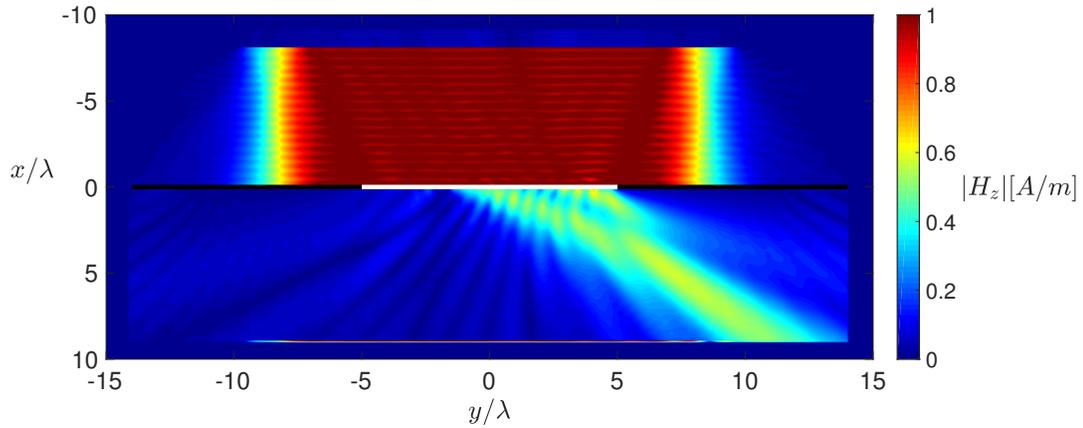


Fig. 3.8: The magnitude of H_z in the solution domain when the metasurface designed for the phaseless power pattern is illuminated by a normally incident tapered plane wave, simulated using the FDFD-GSTC solver.

(3.18), and then the required tangential fields and subsequently the required susceptibilities are found. The resulting susceptibilities are then given to the FDFD-GSTC method to evaluate the performance of the metasurface. The far-field normalized power pattern produced by the metasurface designed for these specifications is shown in Figure 3.9 (blue curve)

Table 3.1: Desired Far-field (FF) Specifications – 2D Example ($\theta = 90^\circ$)

Specifications	Main Beam 1	Main Beam 2
Direction	$\varphi = -26^\circ$	$\varphi = 34^\circ$
HPBW	38°	12°
Nulls (relative to the main beam)	28° at -60 dB	32° at -60 dB

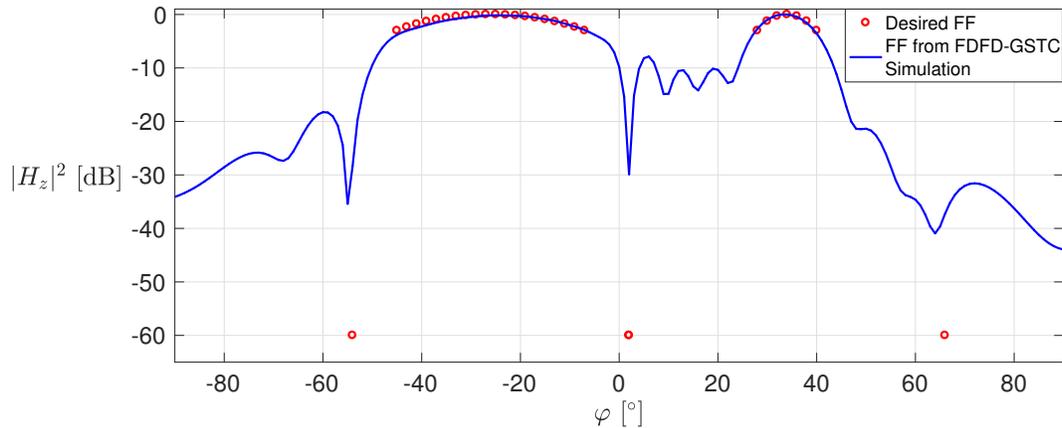


Fig. 3.9: The red circular markers show the desired normalized far-field (FF) power pattern obtained from a set of performance criteria S^{des} outlined in Table 3.1. Note the three red circles at -60 dB level indicating the desired null directions. The solid blue curve shows the normalized FF power pattern produced by the FDFD-GSTC simulation of the designed metasurface.

and represents an efficiency of 13.0%. This achieved FF power pattern has main beams in $\varphi = -25^\circ$ and $\varphi = 34^\circ$, with HPBWs of 36.1° and 11.8° respectively. The desired nulls, although not at the specified level of -60 dB, are clearly present in the FF power pattern as well. Comparing the metrics of the produced FF power pattern to those specified in Table 3.1 confirms that the resulting metasurface satisfies the design constraints with only a few minor differences.

3.9.2 3D Example

To display the generality of the proposed method, we now present an example with desired far-field performance criteria in 3D. The metasurface in this case is a spherical cap that

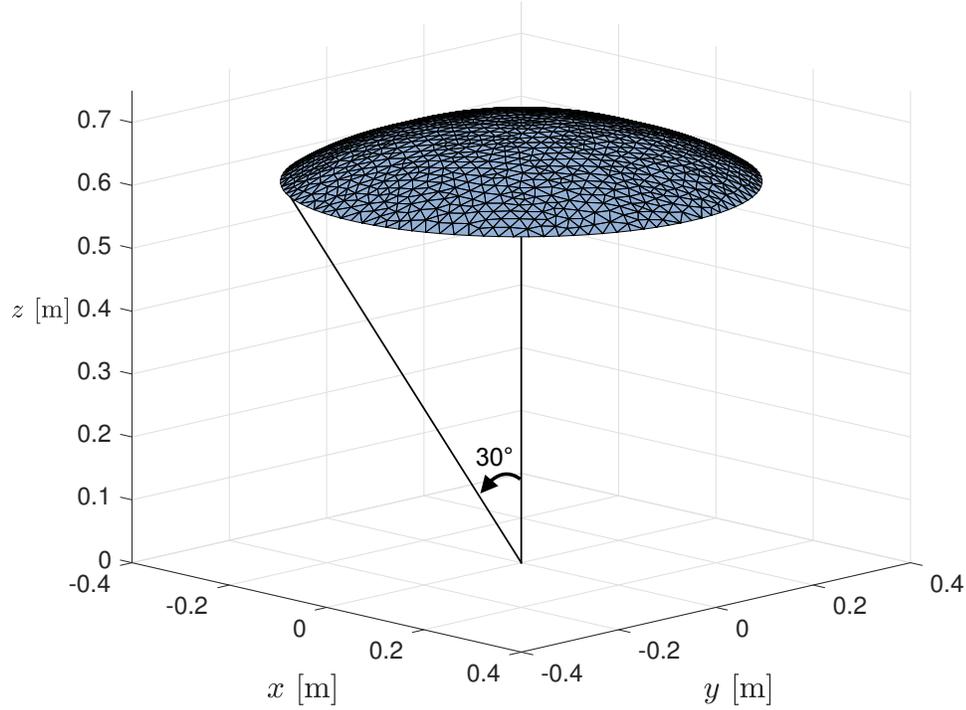
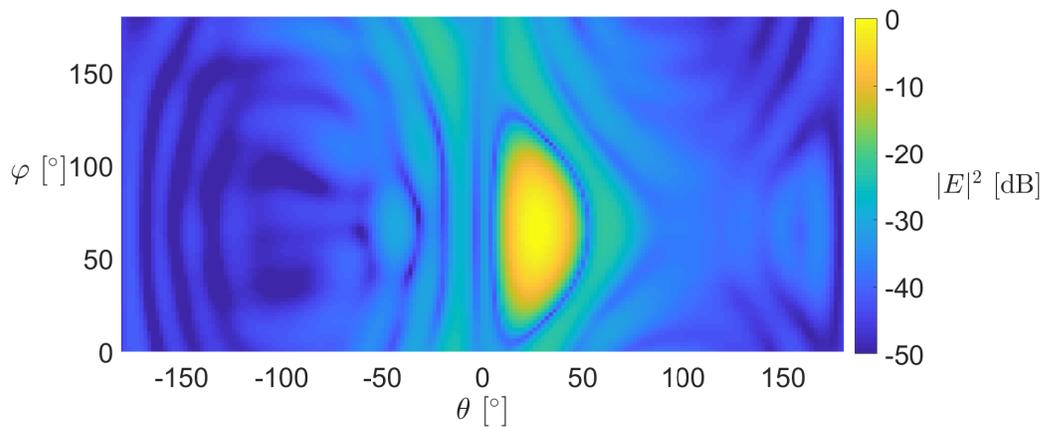


Fig. 3.10: The spherical cap metasurface boundary Σ used for the 3D example, which extends from $\theta = 0^\circ$ to $\theta = 30^\circ$ at a radius of 0.70 m. The equivalent currents are reconstructed on this surface.

extends from $\theta = 0$ to $\theta = 30^\circ$ at a radius of 0.70 m, as shown in Figure 3.10. Love's condition is enforced on an inward-offset surface of the same shape but at a radius of 0.68 m, which results in a separation of approximately $\lambda/10$ between the two surfaces at the operating frequency of 1.7 GHz. The reconstruction surface Σ in this example (i.e., the spherical cap) is discretized into triangular mesh elements and the equivalent currents are represented using Rao-Wilton-Glisson (RWG) basis functions [99]. The desired FF performance criteria are shown in Table 3.2. These specifications are enforced over a (partial) spherical domain of radius 500λ and an angular resolution of 1° . The inverse source algorithm is applied to (3.20) to find the equivalent currents from the desired specifications. Once these currents are reconstructed, their power pattern is simulated using a 3D EFIE solver, which is shown in Figures 3.11 and 3.12. As can be seen, the main beam produced

Table 3.2: Desired Far-field (FF) Specifications – 3D Example

Specifications	Main Beam
Direction	$\theta = 28^\circ, \varphi = 66^\circ$
HPBW	20°
Nulls (relative to the main beam)	24° at -60 dB
Polarization	Main beam linearly polarized in $\hat{\theta}$ direction

**Fig. 3.11:** The normalized far-field power pattern produced by the equivalent currents obtained by the inverse source algorithm for the desired specifications in Table 3.2.

by the equivalent currents is along $(\theta = 28^\circ, \varphi = 66^\circ)$ as desired, and nearly satisfies the beamwidth requirement with a HPBW of 17.8° . Also, the first null is 24° away from the main beam in a symmetric fashion. Moreover, within the main beam, the $\hat{\theta}$ component of the electric field is a *minimum* of 31.5 dB higher than the undesired $\hat{\varphi}$ polarization. In summary, all of the specifications have been met with only a minor deviation in the HPBW.

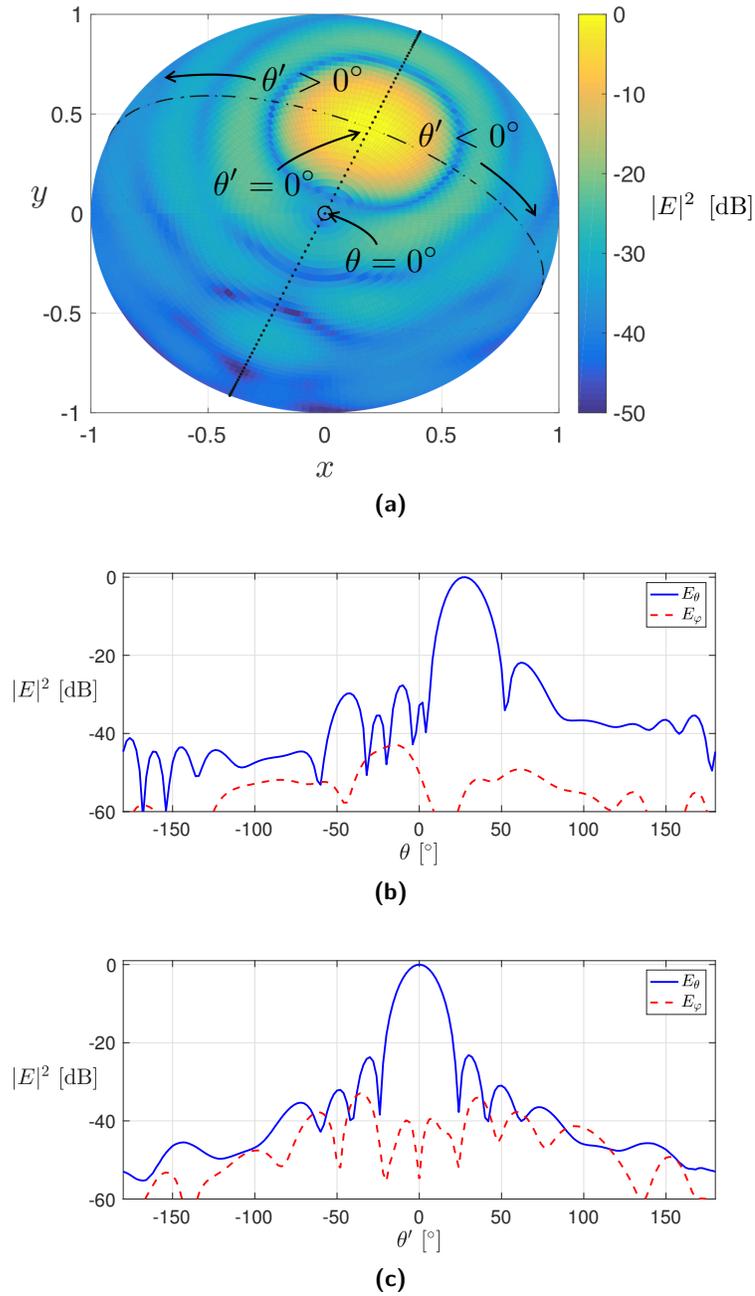


Fig. 3.12: The E-plane (b) and H-plane (c) cuts of the far-field pattern obtained by the inverse source algorithm for the desired specifications in Table 3.2. For clarification, the FF power pattern is shown in (a) over a normalized sphere as viewed from the positive z axis ($\theta = 0^\circ$), with the E-plane cut represented by the dotted line and the H-plane cut represented by the dash-dotted line. To accommodate plotting of the H-plane cut, a new angle denoted by θ' has been introduced such that $\theta' = 0^\circ$ coincides with $(\theta = 28^\circ, \varphi = 66^\circ)$ as shown in (a) and used as the horizontal axis in (c).

3.10 Limitations

While the presented method has the potential to allow for more practical, non-idealistic metasurface designs, it is important to note the limitations of the current implementation. The main limitation arises from the fact that local power conservation is not enforced at any point during the design procedure, necessarily resulting in lossy/active elements. Future work will be focused on enforcing local power conservation during the inversion process, made possible by prior knowledge of the incident field. Furthermore, supporting such a transformation without loss or gain requires extra degrees of freedom, either by allowing for nonreciprocity or bianisotropy [100, 101] (e.g., by introducing magnetoelectric coupling). The reduction of the degrees of freedom from enforcing local power conservation will limit the solution space, and will most likely result in a worse fitting of the design specifications.

Another limitation arises from the nature of the inverse problem, which is that an ‘appropriate’ solution may not exist for a given set of design specifications and metasurface size and geometry. We currently do not have a method for determining if a ‘feasible’ solution exists prior to the optimization process and therefore tuning of the metasurface parameters can be required in some cases.

3.11 Conclusion

A macroscopic metasurface design procedure has been presented that allows for a field transformation between incident, reflected, and transmitted fields. The desired transmitted field can be specified in a variety of ways, including complex field quantities, amplitude-only (phaseless) field data, or in terms of far-field performance criteria such as main beam directions, HPBW, null locations, or polarization. Additionally, there are no restrictions on the locations at which the desired field is specified and the metasurface geometry, allowing for increased design flexibility. Several 2D examples were shown to demonstrate the capabilities

of the proposed method for each of the different types of desired fields, and the designed metasurfaces were simulated using a FDFD-GSTC solver. The 3D example, although not verified through simulation of surface susceptibilities, was presented to demonstrate the most general form of the proposed method.

Many challenges remain, including ensuring the resulting metasurface susceptibilities are physically realizable. This may include incorporating restrictions into the inversion to ensure that the resulting elements are passive and lossless (as mentioned in Section 3.10), while also reducing high spatial variations. These restrictions can, for example, be implemented by adding extra terms to the data misfit cost functional. Similarly, we have not investigated how we can take advantage of the non-uniqueness of the inverse source problem to yield solutions that are more desirable. Lastly, it is possible to specify a desired field that cannot be produced by the metasurface. With this in mind, a procedure that can determine whether or not the desired field can be supported by the metasurface geometry would be a useful future addition.

Chapter 4

Enforcing Local Power Conservation

Preface

The main limitation of the framework presented in Chapter 3 is that the resulting metasurface is (in general) lossy and/or active. In order to facilitate the design of practical unit cells, the design procedure must be altered to ensure the generated susceptibilities are completely lossless and passive. This requires introducing additional degrees of freedom through new susceptibility terms as detailed in the following sections. As will be shown, the introduction of additional susceptibility terms results in a bianisotropic metasurface which can potentially be nonreciprocal. Thus, care is also taken to ensure that the now bianisotropic metasurface also satisfies the conditions for reciprocity.

The material presented in this chapter is based on a paper published in *IEEE Antennas and Wireless Propagation Letters* in August 2020 [102].¹

¹© 2020 IEEE. Reprinted, with permission, from T. Brown, Y. Vahabzadeh, C. Caloz, and P. Mojabi, “Electromagnetic Inversion with Local Power Conservation for Metasurface Design”, *IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 8, pp. 1291-1295, 2020.

Abstract

A method based on electromagnetic inversion is extended to facilitate the design of passive, lossless, and reciprocal metasurfaces. More specifically, the inversion step is modified to ensure that the field transformation satisfies local power conservation, using available knowledge of the incident field. This paper formulates a novel cost functional to apply this additional constraint, and describes the optimization procedure used to find a solution that satisfies both the user-defined field specifications and local power conservation. Lastly, the method is demonstrated with a two-dimensional (2D) example.

4.1 Introduction

Metasurfaces have emerged as useful devices for systematically controlling electromagnetic fields [6, 7, 8, 9, 10, 11]. These subwavelengthly thin structures can perform arbitrary field transformations by imposing appropriate generalized boundary conditions, providing a level of control over some *desired* field produced by a known *incident* field. This fundamental ability has led to a variety of applications, including generalized refraction and reflection [14], polarization manipulation [12, 13], spatial processing [18], impedance matching [103], radiation pattern tailoring [32], and others.

In order to design a metasurface to perform a given field transformation, the tangential electric and magnetic fields must be known on either side of the boundary imposed by the metasurface. Most existing design procedures are limited to problems in which the output field is known analytically on the output side of the metasurface. However, this is satisfactory only for well-defined problems such as plane wave refraction [23]. In order to generalize this to incompletely defined problems, we recently developed a design method which allows for more flexible output field specifications [27]. Using this method, referred to as electromagnetic inversion for metasurface design, the field can be specified at arbitrary locations

external to the metasurface, either with or without phase (amplitude-only) information. Furthermore, the desired field can also be specified as a *set of performance* criteria, such as main beam direction(s), null location(s), beamwidth, or polarization. While this method allows for more general field specifications, it does not take advantage of prior knowledge of the incident field and typically requires loss and/or gain to perform the specified field transformation.

In this work, we extend the electromagnetic inversion algorithm of [27] to allow for the design of lossless, passive, and reciprocal metasurfaces. This method uses electromagnetic inversion to solve for a set of tangential output (transmitted) fields that produce some user-specified field, but modifies the inversion process by incorporating an additional step that penalizes solutions that do not satisfy local power conservation (LPC). Note that LPC dictates that for reflectionless metasurfaces the power entering a given unit cell is equal to the power exiting that unit cell in the direction normal to the metasurface. Once an appropriate solution is found that satisfies both the field specifications and LPC, surface susceptibilities (or, other types of parameters such as surface impedances) can be computed to perform the required transformation.

Enforcing LPC necessarily limits the ability to tailor the amplitude of the wavefronts. This can be understood by noting that LPC is concerned with power at the unit cell level, thus, immediately placing a constraint on the amplitude of the wave at the output port of each unit cell. This can be an issue for complex design problems such as antenna pattern synthesis as they often need full control over both amplitude and phase. Recently, different methods have been suggested to enable more amplitude control by considering total power conservation (TPC); e.g., see [104] for different examples. One of these methods [32],[33] uses two metasurfaces instead of one, each of which individually satisfies LPC. However, collectively, the two-metasurface system only satisfies TPC. Herein, we restrict our attention to single metasurfaces and do not consider cascaded structures.

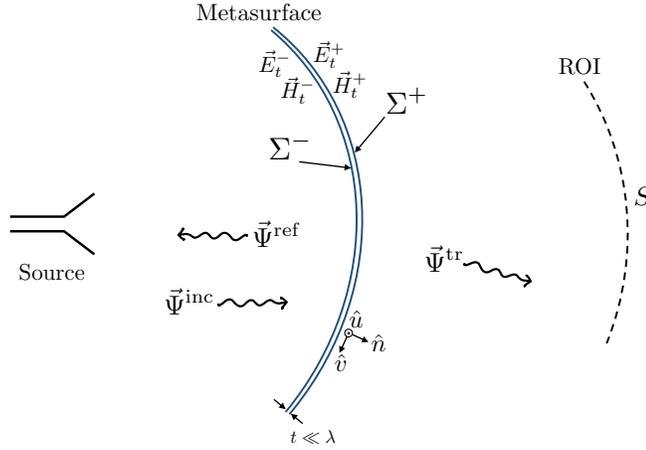


Fig. 4.1: Overview of the metasurface design problem. The input and output surface boundaries of the metasurface are denoted by Σ^- and Σ^+ , respectively. Some source generates an incident field $\vec{\Psi}^{\text{inc}}$ which interacts with the metasurface, producing both a reflected field $\vec{\Psi}^{\text{ref}}$ and a transmitted field $\vec{\Psi}^{\text{tr}}$. The tangential components of the electric and magnetic fields on Σ^- are denoted as \vec{E}_t^- and \vec{H}_t^- , while the tangential fields on Σ^+ are denoted as \vec{E}_t^+ and \vec{H}_t^+ . The user-defined field specifications S are defined on some of interest (ROI) external to the metasurface. Since the metasurface may be of arbitrary shape, we define the local coordinate system $(\hat{u}, \hat{v}, \hat{n})$ on Σ^+ , where \hat{n} is the unit outward normal to Σ^+ . © 2019 IEEE. Reprinted, with permission, from [27] with minor modifications.

We begin by briefly reviewing the electromagnetic inversion design procedure without LPC. We then discuss and derive the constraint used to enforce LPC, and then describe how the inversion process is modified to account for this new constraint. An example is then presented to demonstrate this idea. Finally, some conclusions and a discussion of possible extensions to this work are presented.

4.2 Inverse Source Design Framework

Herein, we present a brief review of the design method presented in [27], where the main goal is to find tangential fields on the *output* side of the metasurface that satisfy some set of user-defined field specifications S in some external region of interest (ROI). An overview of the problem is depicted in Figure 4.1. We denote the input and output surface boundaries of the metasurface as Σ^- and Σ^+ , respectively. The tangential fields (denoted as such by

the subscript t) that we require to design the metasurface consist of the total fields on Σ^- , \vec{E}_t^- and \vec{H}_t^- (consisting of the incident and reflected fields), and the transmitted fields on Σ^+ , \vec{E}_t^+ and \vec{H}_t^+ . The user-defined specifications S in the ROI fall into three general categories, ordered from most to least specific (i.e., most to least information):

1. Complex (amplitude and phase) field distributions (either in the near-field or far-field regions),
2. Phaseless field distributions (i.e., amplitude-only, power pattern),
3. Far-field performance criteria (i.e., main beam directions(s), null locations, beamwidth, etc.).

The specifications S in the ROI will be inputted to an electromagnetic inverse source algorithm, which we simply refer to as an inversion algorithm.² This inversion algorithm solves for a set of equivalent electric (\vec{J}) and magnetic (\vec{M}) currents that produce the field specifications in the ROI. It should be noted that a *unique* solution for the equivalent currents does not exist due to the inherent ill-posedness of the inverse source problem; however, this non-uniqueness is actually *advantageous* for design problems as it increases the available degrees of freedom [105, 75]. The domain upon which the equivalent currents are determined, commonly referred to as the ‘reconstruction surface’, is chosen to coincide with the physical boundary imposed by the metasurface. These currents are found by minimizing a data misfit cost functional, which we denote herein as $\mathcal{C}_1(\vec{J}, \vec{M})$, using the conjugate gradient method. (This cost functional is a mapping from complex equivalent currents to a real value.) This functional essentially quantifies the difference between the fields generated by the equivalent currents and the field specifications, with the exact form depending on the category of field specifications listed above. (For more details on this data misfit cost functional, see (12), (13), and (20) in [27].)

²In addition to electromagnetic inverse source algorithms, electromagnetic inverse scattering algorithms have also been used for design applications; e.g., see [72].

If Love's equivalence condition is enforced (i.e., null fields on the input side of the metasurface), then the resulting equivalent currents are related to the desired transmitted fields as

$$\vec{H}_t^+ = -\alpha \hat{n} \times \vec{J} \quad \text{and} \quad \vec{E}_t^+ = \alpha \hat{n} \times \vec{M}, \quad (4.1)$$

where α is a real-valued scaling parameter which does not affect the characteristics of the *normalized* radiated field, but allows for some flexibility that will be utilized in the next section. Once the desired tangential transmitted fields are known, the generalized sheet transition conditions (GSTCs) [35] can be utilized to determine a set of surface susceptibilities modelling the discontinuity from the (known) incident/reflected field and (desired) transmitted field [11]. Assuming a time-dependency of $e^{j\omega t}$ and free space on either side of the metasurface, the relationship can be written as

$$\begin{aligned} \begin{pmatrix} -\Delta H_v \\ \Delta H_u \end{pmatrix} &= j\omega\epsilon_0 \begin{pmatrix} \chi_{ee}^{uu} & \chi_{ee}^{uv} \\ \chi_{ee}^{vu} & \chi_{ee}^{vv} \end{pmatrix} \begin{pmatrix} E_{u,av} \\ E_{v,av} \end{pmatrix} \\ &+ j\omega\sqrt{\epsilon_0\mu_0} \begin{pmatrix} \chi_{em}^{uu} & \chi_{em}^{uv} \\ \chi_{em}^{vu} & \chi_{em}^{vv} \end{pmatrix} \begin{pmatrix} H_{u,av} \\ H_{v,av} \end{pmatrix} \end{aligned} \quad (4.2a)$$

$$\begin{aligned} \begin{pmatrix} -\Delta E_u \\ \Delta E_v \end{pmatrix} &= j\omega\mu_0 \begin{pmatrix} \chi_{mm}^{vv} & \chi_{mm}^{vu} \\ \chi_{mm}^{uv} & \chi_{mm}^{uu} \end{pmatrix} \begin{pmatrix} H_{v,av} \\ H_{u,av} \end{pmatrix} \\ &+ j\omega\sqrt{\epsilon_0\mu_0} \begin{pmatrix} \chi_{me}^{vv} & \chi_{me}^{vu} \\ \chi_{me}^{uv} & \chi_{me}^{uu} \end{pmatrix} \begin{pmatrix} E_{v,av} \\ E_{u,av} \end{pmatrix}, \end{aligned} \quad (4.2b)$$

where ω is the angular frequency of the time harmonic fields, and ϵ_0 and μ_0 are the permittivity and permeability of free space.³ The subscripts and superscripts u and v denote the

³The formulation shown here assumes that the normal components of the polarization densities are zero for mathematical simplicity.

tangential components of the local coordinate system of each unit cell defined by $\hat{u} \times \hat{v} = \hat{n}$ and $\hat{u} \perp \hat{v}$. The χ terms represent the electric/magnetic (first subscript) surface susceptibility components in the presence of an electric/magnetic (second subscript) field excitation [83]. The spatial dependencies of the E , H , and χ terms, which span the metasurface geometrical surface, have been dropped for brevity. The difference and average fields are defined for an arbitrary field $\vec{\Psi}$ as

$$\Delta\vec{\Psi} \triangleq \vec{\Psi}^{\text{tr}} - (\vec{\Psi}^{\text{inc}} + \vec{\Psi}^{\text{ref}}) \quad (4.3)$$

$$\vec{\Psi}_{\text{av}} \triangleq \frac{\vec{\Psi}^{\text{tr}}|_{\Sigma^+} + (\vec{\Psi}^{\text{inc}}|_{\Sigma^-} + \vec{\Psi}^{\text{ref}}|_{\Sigma^-})}{2}. \quad (4.4)$$

The final step in the design procedure is solving (4.2) for the non-zero susceptibility terms (depending on the problem, some χ terms may be assumed to be zero). Once the susceptibilities are determined, they can be implemented in different ways, e.g., as three-layered (dogbone) impedance sheets [61, 101], as will be considered later. Note that in the three-layer impedance sheet approach, it is often assumed that the impedance sheets are purely reactive, and that they are printed on (nearly) lossless dielectric substrates. Therefore, LPC needs to hold when considering this unit cell design approach.

4.3 Enforcing Local Power Conservation (LPC)

The main limitation of the electromagnetic inversion design procedure presented in [27] is that the synthesized susceptibilities may require (undesirable) loss and/or gain. To overcome this issue, we first require that the input and output fields must satisfy LPC [30, 101]. That is, the real power incident on each unit cell must be equal to the real power transmitted from each unit cell, as enforced by the following equation that must hold along the

metasurface, i.e., at each unit cell:

$$\frac{1}{2} \text{Re}(\vec{E}_t^- \times \vec{H}_t^{-*}) = \frac{1}{2} \text{Re}(\vec{E}_t^+ \times \vec{H}_t^{+*}) \quad (4.5)$$

where ‘Re’ denotes the real-part operator and the superscript ‘*’ denotes the complex conjugate operator. From this point onwards, we will assume 2D TE_z polarized fields and a 1D metasurface along the line $x = 0$ (i.e., $\hat{u} = \hat{y}$, $\hat{v} = \hat{z}$, and $\hat{n} = \hat{x}$) for simplicity, although the formulation would still hold for arbitrarily-shaped metasurfaces and 3D fields. We denote the left hand side of (4.5) evaluated at the i^{th} unit cell as

$$p_i = \frac{1}{2} \text{Re}(E_y^- \times H_z^{-*}) \Big|_{\text{unit cell } i} . \quad (4.6)$$

Note that p_i for all i values is known since it corresponds to the input power density at each unit cell of the metasurface. (Herein, we assume a reflectionless metasurface; thus, E_y^- and H_z^- correspond to the incident electromagnetic field.) Using (4.6) and the relationship between the equivalent currents and the tangential transmitted fields in (4.1), we can write the LPC constraint in (4.5) as the following vector equality

$$\begin{bmatrix} p_1 \\ \vdots \\ p_N \end{bmatrix} = \frac{\alpha^2}{2} \text{Re} \left(\begin{bmatrix} M_z J_y^* |_{\text{unit cell 1}} \\ \vdots \\ M_z J_y^* |_{\text{unit cell N}} \end{bmatrix} \right) \quad (4.7)$$

where N is the total number of unit cells. The above vector equality can then be compactly written as

$$\mathbf{p} = \frac{\alpha^2}{2} \text{Re}(\mathbf{M} \odot \mathbf{J}^*) \quad (4.8)$$

where \mathbf{J} and \mathbf{M} are discretized complex vectors of the equivalent currents J_y and M_z at each unit cell; i.e., $\mathbf{J} \in \mathbb{C}^N$ and $\mathbf{M} \in \mathbb{C}^N$. In addition, ‘ \odot ’ represents the element-wise Hadamard

product. The vector $\mathbf{p} \in \mathbb{R}^N$ is a discrete vector of the real incident power calculated at each unit cell, with the i^{th} element of \mathbf{p} equal to p_i .

Separating the equivalent currents into their real and imaginary parts, denoted by the subscripts ‘R’ and ‘I’, (4.8) becomes

$$\begin{aligned}\mathbf{p} &= \frac{\alpha^2}{2} \text{Re} \{ (\mathbf{M}_R + j\mathbf{M}_I) \odot (\mathbf{J}_R - j\mathbf{J}_I) \} \\ &= \frac{\alpha^2}{2} (\mathbf{J}_R \odot \mathbf{M}_R + \mathbf{J}_I \odot \mathbf{M}_I).\end{aligned}\tag{4.9}$$

At this point, the restriction imposed by the LPC constraint becomes obvious. In [27], the equivalent currents had *four* degrees of freedom (i.e. \mathbf{J}_R , \mathbf{J}_I , \mathbf{M}_R , and \mathbf{M}_I) to satisfy the field specifications, but as shown in (4.9), the LPC constraint reduces the degrees of freedom to *three*. In other words, enforcing LPC results in a reduction in the dimension of the solution space, and may exclude some solutions that would otherwise satisfy the field constraints in an optimal manner.

We can now formulate a cost functional, say \mathcal{C}_2 , to quantify the LPC constraint and include it in the design procedure. This term is formulated from (4.9) as

$$\mathcal{C}_2(\mathbf{J}, \mathbf{M}) = \frac{\|\mathbf{J}_R \odot \mathbf{M}_R + \mathbf{J}_I \odot \mathbf{M}_I - \frac{2}{\alpha^2} \cdot \mathbf{p}\|_2^2}{\|\frac{2}{\alpha^2} \cdot \mathbf{p}\|_2^2}\tag{4.10}$$

where $\|\cdot\|_2$ represents an L_2 norm. Since we have complete freedom in selecting the scaling parameter α , it should be chosen in a way that minimizes (4.10) for a given set of currents. Therefore, when (4.10) is evaluated, the parameter α that results in the minimum of (4.10) is used.

4.4 Methodology

Our unknowns consist of the separated real and imaginary parts of the electric and magnetic equivalent currents, which we collectively write for convenience as

$$\mathbf{x} = \left[\mathbf{J}_R ; \mathbf{J}_I ; \mathbf{M}_R ; \mathbf{M}_I \right] \in \mathbb{R}^{4N} \quad (4.11)$$

where ‘;’ denotes the column-wise vector concatenation. First, as described in [27], $\mathcal{C}_1(\mathbf{x})$, which includes the desired field specifications, is minimized *without* the LPC constraint using the conjugate gradient method. This provides an estimate for \mathbf{x} that satisfies the field constraints and Love’s condition prior to applying the LPC constraint. This estimate of \mathbf{x} is denoted as \mathbf{x}_0 . Next, we minimize

$$\mathcal{C}(\mathbf{x}) = \underbrace{\mathcal{C}_1(\mathbf{x})}_{\text{specifications}} + \kappa \underbrace{\mathcal{C}_2(\mathbf{x})}_{\text{LPC}}, \quad (4.12)$$

where κ is a real-valued scalar weighting parameter used to balance the contribution of \mathcal{C}_2 (LPC) with respect to \mathcal{C}_1 (field specifications). Particle swarm optimization (a *global* optimization technique) is used to minimize (4.12), rather than a gradient-based technique, due to the increased nonlinearity introduced by $\mathcal{C}_2(\mathbf{x})$.⁴ The initial particle states are set to \mathbf{x}_0 and the absolute search space bounds are set to be 10% above the maximum absolute value of \mathbf{x}_0 .⁵ Once convergence is reached, the required tangential fields are obtained using (4.1).⁶

⁴Particle swarm was selected because the convergence behaviour was observed to be more consistent compared to other global optimization methods (e.g., genetic algorithm and simulated annealing.)

⁵Limiting the search space to within 10% of the initial guess was sufficient due to the quality of the initial guess from the previous optimization stage.

⁶In this framework, the number of discrete points used to enforce the field specifications does not typically increase with problem size. Therefore, for larger problems, the number of ‘observation’ points (related to the rows of the matrix operator involved in computing \mathcal{C}_0) will become dominated by the ‘virtual’ points at which Love’s equivalence condition (null field) is enforced. With this assumption, the computational complexity of *evaluating* \mathcal{C}_1 is $O(N^2)$ and \mathcal{C}_2 is $O(N)$, where N is the number of unit cells.

Next, the susceptibility components required to support the desired transformation must be computed. As noted in [101], if we want to support a transformation of this nature without using loss and/or gain, we require more degrees of freedom than afforded by only $\bar{\chi}_{ee}$ and $\bar{\chi}_{mm}$ (tensors collectively representing the various χ_{ee} and χ_{mm} terms in (4.2), respectively). One way to overcome this limitation is by allowing the bianisotropic terms $\bar{\chi}_{em}$ and $\bar{\chi}_{me}$ to be non-zero, introducing magnetoelectric coupling to the metasurface.

Assuming 2D TE_z fields with a 1D metasurface along $x = 0$, (4.2) simplifies to

$$-\Delta H_z = (j\omega\epsilon_0 E_{y,av}) \chi_{ee}^{yy} + (j\omega\sqrt{\mu_0\epsilon_0} H_{z,av}) \chi_{em}^{yz} \quad (4.13a)$$

$$-\Delta E_y = (j\omega\mu_0 H_{z,av}) \chi_{mm}^{zz} + (j\omega\sqrt{\mu_0\epsilon_0} E_{y,av}) \chi_{me}^{zy}. \quad (4.13b)$$

In order to avoid loss and gain, we first stipulate that χ_{ee}^{yy} and χ_{mm}^{zz} must be purely real [101]. Next, we note that since $\chi_{ee}^{yy} \in \mathbb{R}$ and $\chi_{mm}^{zz} \in \mathbb{R}$ satisfy the first two conditions for losslessness and passivity, the remaining (third) condition for losslessness and passivity, i.e., $(\chi_{em}^{yz})^* = \chi_{me}^{zy}$, must also hold since the field transformation satisfies LPC. If we enforce χ_{em}^{yz} and χ_{me}^{zy} to be purely imaginary, then any lossless and passive solution will also satisfy the condition of reciprocity, $\chi_{em}^{yz} = -\chi_{me}^{zy}$ [23]. This results in four real unknowns (per unit cell) that must satisfy the two complex equations in (4.13), which can be directly computed assuming the tangential fields on both sides of the metasurface are known.

4.5 Results

To illustrate the proposed method, we design a reflectionless 1D metasurface specified to transform an incident TE_z plane wave into a desired power pattern (phaseless field information only) specified in the far-field region. The fields are assumed to propagate in 2D in the xy plane, with the metasurface placed at $x = 0$. The frequency is 10.5 GHz and the metasurface unit cells are $\lambda/6$ in length, where λ represents the free space wavelength. The

designed region of the metasurface extends from $y = -2.5\lambda$ to $y = 2.5\lambda$, with absorbing elements placed along the rest of the $x = 0$ line. The incident field is a normally incident uniform plane wave. The desired phaseless power pattern is produced by simulating an array of 9 uniformly spaced elementary dipoles along the y -axis between $y = -2\lambda$ and $y = 2\lambda$. The specified desired far-field power pattern is computed for $-90^\circ \leq \varphi \leq 90^\circ$ and shown (in red) in Figure 4.2.

The surface on which the equivalent currents are reconstructed is chosen to coincide with the metasurface, with the same $\lambda/6$ discretization. First, a solution is found by minimizing the cost functional *without* the LPC constraint. This solution is then used to initialize a particle swarm optimization algorithm that minimizes (4.12) using a swarm size of 200 and a scaling factor of 0.1 for κ . The far-field pattern corresponding to the equivalent currents from the particle swarm optimization stage is shown in Figure 4.2 to demonstrate the decrease in pattern accuracy that must be made to enforce the LPC constraint. The resulting solution is then used to compute the four susceptibility terms using (4.1) and (4.13), which are necessarily passive, lossless, and reciprocal.

We then simulate the designed metasurface in ANSYS HFSS using the method described in [106]. This technique implements each unit cell using a three-layer admittance sheet topology [107], using the conversion between susceptibilities and admittances explicitly stated in [106]. The unit cells are then modelled in HFSS using three impedance boundary conditions, between which exist two substrate layers. The substrate used in this example is Rogers RO3010 ($\epsilon_r = 10.2$, $\tan \delta = 0.0022$) with each layer having a thickness of 50 mil. Additionally, metallic baffles have been placed between the unit cells, as in [61]. The total electric field resulting from this simulation is shown in Figure 4.3, and the far-field pattern associated with this simulation is shown in Figure 4.2.⁷ These results show that the main

⁷It should be noted that the cost functional consists of two weighted terms: a data misfit term and a local power conservation term. The degradation in FF performance compared to the results in Chapter 3 is largely due to the introduction of the local power conservation constraint, which requires some compromise in FF performance for the metasurface to be reciprocal, passive, and lossless.

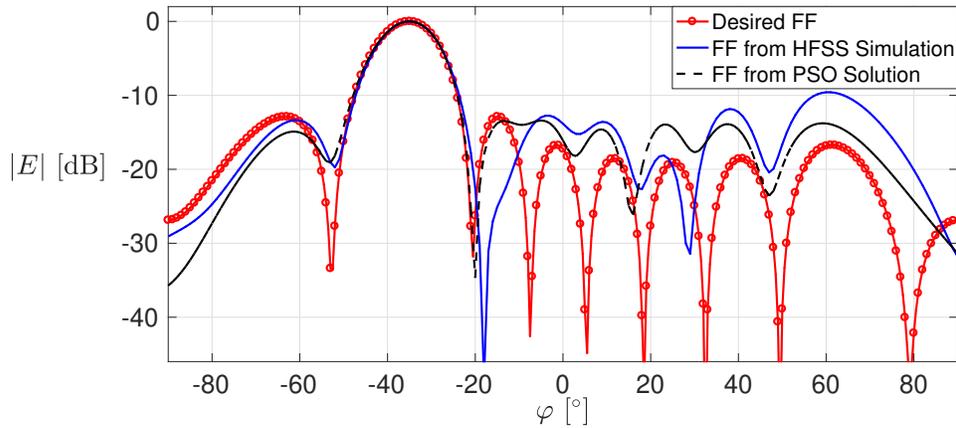


Fig. 4.2: Far-field power pattern produced by the equivalent currents generated after enforcing LPC using PSO (dashed black curve), the power pattern produced by the HFSS simulation of the designed metasurface (solid blue curve), and specified power pattern (solid red curve with circular markers).

features of the desired power pattern have been generated with only minor reflections, despite some deviation in the sidelobes. As noted earlier, this might be attributed to the fact that LPC does not allow for full control over the amplitude of the transmitted field. The transmission efficiency, defined as the ratio of the real power transmitted through the metasurface to the real power incident on the metasurface, is 78.7%.⁸

⁸The transmission efficiency observed in this example is significantly higher than the metasurfaces designed in Chapter 3, which all exhibited an efficiency under 20%. This difference is mainly due to the metasurfaces in Chapter 3 having loss, while the metasurface designed in this chapter is passive and lossless. The transmission efficiency can be improved further by incorporating a ‘smoothing’ regularizer into the optimization process, as shown in Appendix C.

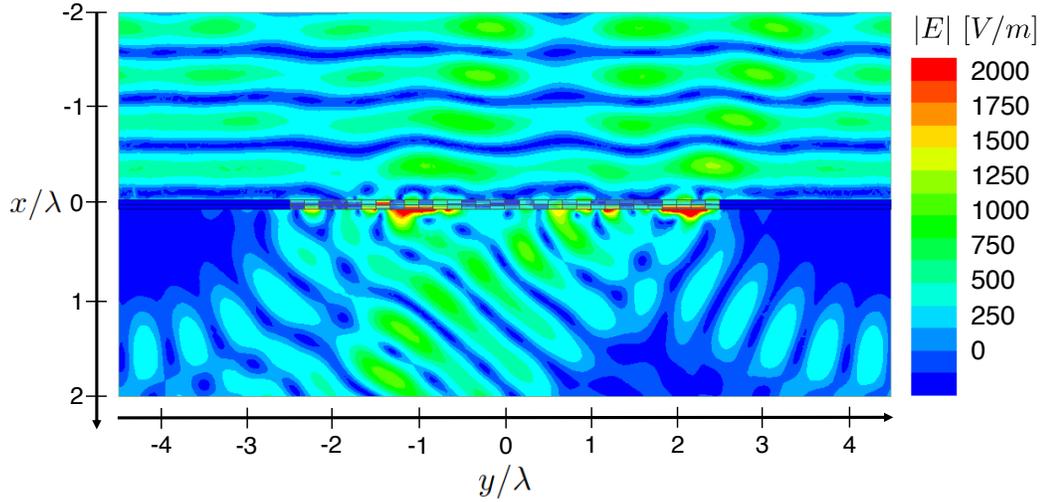


Fig. 4.3: Electric field amplitude when the designed metasurface is implemented using the three-layer admittance sheet topology and illuminated by a normally incident uniform plane wave, simulated using ANSYS HFSS.

4.6 Conclusion

A general metasurface design method was extended to ensure that the resulting field transformation satisfies local power conservation, allowing for the design of passive, lossless, and reciprocal metasurfaces. A constraint on the equivalent currents was derived from the local power conservation relationship, and incorporated into the design procedure using a secondary optimization step. Non-zero magnetoelectric coupling terms are introduced to compensate for the loss of degrees of freedom resulting from excluding loss and gain. A preliminary 2D example was shown for the design of a passive, lossless, and reciprocal metasurface attempting to produce a specified (phaseless) power pattern, with relatively good agreement.

Chapter 5

Cascaded Metasurface Design

Preface

The introduction of magnetoelectric coupling in Chapter 4 allows for the support of field transformations using lossless, passive, and reciprocal metasurfaces as long as local power conservation holds. However, enforcing local power conservation restricts the possible field transformation due to the requirement that the distribution of power across the metasurface must not change from incident to transmitted field (assuming no reflections). As will be shown, introducing a second metasurface allows for a redistribution of power to occur between the two metasurfaces, thereby increasing the variety of field transformations that can be supported. The work presented in this chapter also improves the optimization procedure by using gradient-based methods for all minimization, while also adding a total-variation regularizer to increase the ‘smoothness’ of the solution.

The material presented in this chapter is based on a paper under review for publication in *IEEE Transactions on Antennas and Propagation* titled “Cascaded Metasurface Design Using Electromagnetic Inversion with Gradient-Based Optimization” [108].¹

¹Reprinted, with permission, from T. Brown and P. Mojabi, “Cascaded Metasurface Design Using Electromagnetic Inversion with Gradient-Based Optimization”, *IEEE Trans. Ant. Prop.*, submitted July 2020.

Abstract

This paper presents an electromagnetic inversion algorithm for the design of cascaded metasurfaces that enables the design process to begin from more practical output field specifications such as a desired power pattern or far-field performance criteria. Thus, this method combines the greater field transformation support of multiple metasurfaces with the flexibility of the electromagnetic inverse source framework. To this end, two optimization problems are formed: one associated with the interior space between two metasurfaces, and the other for the exterior space. The cost functionals corresponding to each of these two optimization problems are minimized using the nonlinear conjugate gradient algorithm with analytic expressions for the gradient operators. A total variation regularizer is incorporated into the optimization procedure to favour smooth field variations from one unit cell to the next. The numerical implementation of the developed design procedure is presented in detail along with several two-dimensional (2D) simulated examples to demonstrate the capabilities of the method.

5.1 Introduction

Metasurfaces offer a level of systematic control over electromagnetic fields not typically possible with conventional materials [6, 8, 7, 10, 9, 11]. These thin metamaterials of sub-wavelength thickness can be designed to support arbitrary field discontinuities, leading to applications such as radiation pattern control [32], polarization control [12, 13], impedance matching [103], cloaking [83], and others. Fabrication of these metasurfaces is often much simpler than bulk metamaterials, typically using established printed circuit board (PCB) techniques and materials.

Metasurface design can be decomposed into two distinct steps: macroscopic and microscopic [22]. The first step, macroscopic design, is to determine a homogenized representation

of the metasurface that is able to effectively model the desired field transformation. Several representations are common in the literature, including effective surface susceptibilities [23], surface impedances/admittances [8], or surface polarizabilities [9]; herein, we adopt the use of surface susceptibilities. The second step, microscopic design, is to determine physical unit cell structures that exhibit the behaviour of the metasurface model. For example, in [61] and [101], three metallic dogbone layers have been used on two dielectric substrates connected by a bondply layer to implement the desired metasurface model. Microscopic design, although important, is not considered in the work presented here.

Several macroscopic metasurface design methods can be found in recent literature [22, 23]; however, most methods require knowledge of both the incident and the desired transmitted field *on the metasurface* itself. While the incident field is either known analytically or easily measured, an analytical representation of the transmitted field is typically only possible in ideal cases involving plane waves or other simple fields. We recently presented a macroscopic design method that allows for more practical design criteria, such as far-field (FF) performance criteria (e.g., main beam direction, half-power beamwidth (HPBW), null locations, etc.) [79, 27]. Since this method is based on solving an electromagnetic inverse source problem, we have referred to this approach as an electromagnetic inversion algorithm for metasurface design. Subsequently, in [102], we modified this inversion method by augmenting its associated cost functional to enforce local power conservation (LPC) [31, 30] which ensures the resulting metasurfaces can be implemented using passive, lossless, and reciprocal elements.

While practically necessary, the drawback of enforcing LPC is that the supported field transformations are restricted. One possible solution is to utilize *multiple* metasurfaces in succession, as described in [32, 33, 109, 110, 111]. In this scenario, the field transformations at each metasurface still satisfy LPC, but the transmitted field from the last metasurface no longer has to have the same power distribution as the field incident on the first meta-

surface. This extra freedom can support a greater variety of output fields by allowing for a redistribution of the incident field power onto the final (output) metasurface. Considering the cascaded metasurfaces as a single structure, *total* power is still conserved from incident to output fields but the power is redistributed locally.

The design methods presented in [32, 33, 109, 110, 111] utilize metasurface pairs to perform field transformations that would not be possible with a single metasurface.² Similar to the macroscopic design methods mentioned above, these methods require explicit knowledge of the transmitted field *on* the output metasurface. Therefore, the aim of this paper is to adapt the electromagnetic inversion metasurface design method to the design of cascaded metasurfaces (in our case, two metasurfaces) to allow for more general output field specifications. For example, the design objective can be to meet a specific power pattern (phaseless) or to meet performance criteria such as the HPBW and null locations.

To this end, we cast the design problem as two optimization problems. The first optimization problem aims to infer the required tangential fields on the output surface of the second metasurface so as to meet the field specifications. As will be seen, these inferred tangential fields need to satisfy total power conservation (TPC) with respect to the incident power impinging on the first metasurface. The second optimization problem is concerned with finding the tangential fields in the interior space between the two metasurfaces. As will be discussed, these fields need to satisfy LPC for each individual metasurface to ensure that they can be fabricated using passive and lossless elements. It is worth noting that while in [102] we used a stochastic method (particle swarm) to enforce LPC in the inversion process, in this work we use gradient-based optimization for improved convergence and computational efficiency. Furthermore, we introduce regularization into the optimization process based on the L^2 -norm total variation (TV) regularizer commonly used for the inverse problem associated with microwave imaging [113, 114]. This regularizer has a smoothing

²The concept of using multiple patterned surfaces to systematically control the amplitude and phase of a wavefront is reminiscent of early work on dual reflectarrays [112].

effect on the achievable solution which translates into less field variations from one unit cell of the metasurface to the neighbouring unit cell. This results in several benefits which will be discussed in the paper.

We begin with a description of the problem statement followed by a high-level overview of the proposed methodology. We then explain the numerical implementation of each step of the method in detail. (The derivation of the required gradient operators are provided in the appendices.) This is followed by a series of full-wave simulated examples in both 2D transverse magnetic and transverse electric cases. Lastly, we identify the existing limitations of the proposed method and present our conclusions.

5.2 Problem Statement

We consider the design of a pair of metasurfaces, denoted respectively as Σ_1 and Σ_2 as shown in Figure 5.1. In this paper we restrict our discussion to that of planar, parallel metasurfaces separated by a distance d , but the theory presented is consistent with more complicated geometries. A known electromagnetic source produces an incident field $\vec{\Psi}^{\text{inc}}$ where $\vec{\Psi} \in \{\vec{E}, \vec{H}\}$ that impinges on the *input* metasurface, Σ_1 . The interaction of this incident field with the pair of metasurfaces will create a reflected field $\vec{\Psi}^{\text{ref}}$ (which may be zero) emanating from Σ_1 and a transmitted field $\vec{\Psi}^{\text{tr}}$ emanating from the *output* metasurface, Σ_2 .

It is important to note that this design method does not require prior knowledge of $\vec{\Psi}^{\text{tr}}$ as is common in alternative methods, but rather a set of user-defined field specifications, denoted as \mathbf{f} on some region of interest S . As noted in [27], the field specifications could be provided in any of the following forms:

- complex fields (amplitude and phase information) in either the near-field (NF) or far-field (FF) region,

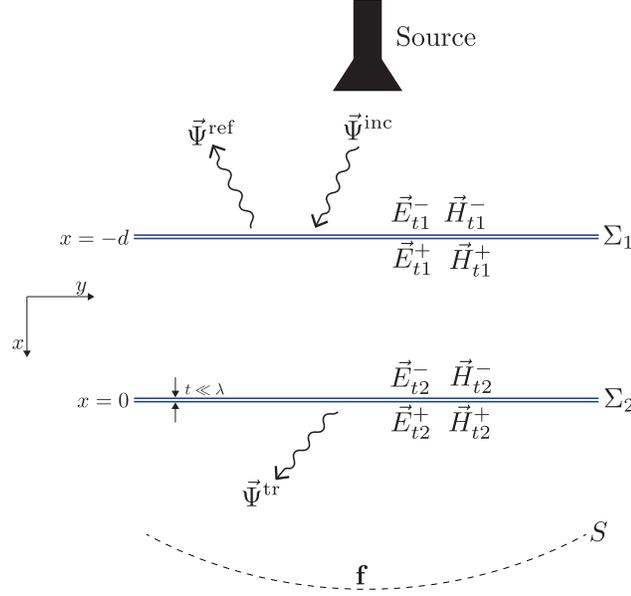


Fig. 5.1: Overview of the cascaded metasurface design problem. The two metasurfaces are denoted as Σ_1 and Σ_2 and separated by a distance of d . The inputs to the design method are the incident field $\vec{\Psi}^{\text{inc}}$ created by the source which impinges on the first metasurface, and the field specifications \mathbf{f} defined on the region of interest S . The fields \vec{E}_{t1}^- and \vec{H}_{t1}^- denote the tangential components of the electric and magnetic field on the *input* side of Σ_1 ('-' superscript), while \vec{E}_{t1}^+ and \vec{H}_{t1}^+ represent the tangential fields on the *output* side of Σ_1 ('+' superscript). The tangential fields on Σ_2 are defined analogously but with a '2' subscript instead of a '1'. The transmitted and reflected fields produced by the interaction are depicted as $\vec{\Psi}^{\text{tr}}$ and $\vec{\Psi}^{\text{ref}}$, respectively.

- phaseless fields (amplitude-only) in either the NF or FF region³, or
- FF performance criteria such as main beam direction(s), half-power beamwidth(s) (HPBW), and null locations.⁴

Each metasurface is characterized by a set of surface susceptibility distributions that relate the induced electric and magnetic polarization densities to the average electric and magnetic fields on the metasurface [35, 36]. The goal of the design procedure is then to

³If the desired phaseless fields are specified in the FF zone, they will be equivalent to a desired power pattern. To emphasize that the desired phase data have *not* been provided to the algorithm, $|\mathbf{f}|^2$ is used in the remaining of this paper to represent the amplitude-only specification for such cases. ($|\cdot|$ denotes the amplitude operator.)

⁴These performance criteria are first converted to a desired power pattern and are then incorporated in the inversion algorithm.

find susceptibility profiles for each metasurface such that, when illuminated with the given incident field, produce a transmitted field $\vec{\Psi}^{\text{tr}}$ that closely satisfies the field specifications. Attention must also be paid to ensure that the resulting susceptibility profiles can be implemented in a passive, lossless, and reciprocal way.

5.3 Methodology

This section provides a high-level overview of the procedure employed to design the cascaded metasurface pair. (More details on the implementation of the procedure will be presented in Section 5.4.) The goal of this procedure is to define the tangential electric and magnetic fields on either side of *both* metasurfaces, from which the susceptibility profiles can be calculated [23]. That is, we need to determine

- tangential output fields, i.e., \vec{E}_{t2}^+ and \vec{H}_{t2}^+ on $x = 0^+$,
- tangential input fields, i.e., \vec{E}_{t1}^- and \vec{H}_{t1}^- on $x = -d^-$, and
- tangential interior fields, i.e., \vec{E}_{t1}^+ and \vec{H}_{t1}^+ on $x = -d^+$, and \vec{E}_{t2}^- and \vec{H}_{t2}^- on $x = 0^-$.

Noting Figure 5.1, in what follows, we refer to $x = 0^+$ and $x = 0^-$ as Σ_2^+ and Σ_2^- respectively, and we refer to $x = -d^+$ and $x = -d^-$ as Σ_1^+ and Σ_1^- respectively.

5.3.1 Finding the Tangential Output Fields (on Σ_2^+)

The first step involves determining tangential electric and magnetic fields on the output side of the second (output) metasurface, denoted as \vec{E}_{t2}^+ and \vec{H}_{t2}^+ , that produce corresponding fields on S that match the user-defined field specifications \mathbf{f} . This step is formulated as an inverse source problem as shown in Figure 5.2, in which the unknowns are equivalent electric and magnetic currents, \vec{J}_2 and \vec{M}_2 . A cost functional is minimized to find \vec{J}_2 and \vec{M}_2 such that these equivalent currents:

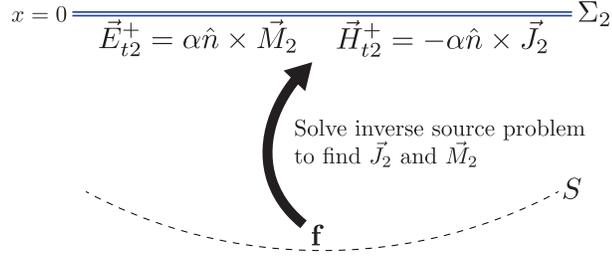


Fig. 5.2: First step of the cascaded metasurface design procedure. An inverse source problem is solved to find equivalent electric and magnetic currents \vec{J}_2 and \vec{M}_2 on Σ_2^+ that (i) produce fields on S satisfying the field specifications \mathbf{f} , and (ii) produce null fields (Love's condition) on inward offset of Σ_2 . Since Love's equivalence condition is enforced, the tangential output fields \vec{E}_{t2}^+ and \vec{H}_{t2}^+ can be directly computed from \vec{J}_2 and \vec{M}_2 .

- produce electric and magnetic fields that satisfy the field specifications \mathbf{f} on the region of interest S , and
- satisfy Love's equivalence condition (i.e., produce null fields on the input side of Σ_2).

Noting that Love's condition is enforced, the resulting equivalent currents will be related to the required tangential electric and magnetic fields as

$$\vec{H}_{t2}^+ = -\alpha \hat{n} \times \vec{J}_2 \quad \text{and} \quad \vec{E}_{t2}^+ = \alpha \hat{n} \times \vec{M}_2, \quad (5.1)$$

where \hat{n} is the unit outward normal to Σ_2 (i.e., towards the output side) and α is a real-valued scaling factor. ($\hat{n} = \hat{x}$ in Figure 5.1.) The choice of α does not affect the *normalized* output field. As will be seen, we select α to ensure that *total* power is conserved across the cascaded metasurface structure.

5.3.2 Tangential Input Fields (on Σ_1^-)

Since the cascaded metasurface system is assumed to be reflectionless, the tangential fields on the input side of the first metasurface (i.e., at $x = -d^-$) are assumed to be the same as the known incident field (i.e., the tangential components of $\vec{\Psi}^{\text{inc}}$). Thus, \vec{E}_{t1}^- and \vec{H}_{t1}^- are

$$\begin{aligned}
x = -d & \frac{\vec{E}_{t1}^-}{\vec{E}_{t1}^+ = \hat{n} \times \vec{M}_1} \quad \frac{\vec{H}_{t1}^-}{\vec{H}_{t1}^+ = -\hat{n} \times \vec{J}_1} \Sigma_1 \\
x = 0 & \frac{\vec{E}_{t2}^- = B_1(\vec{J}_1, \vec{M}_1)}{\vec{E}_{t2}^+} \quad \frac{\vec{H}_{t2}^- = B_2(\vec{J}_1, \vec{M}_1)}{\vec{H}_{t2}^+} \Sigma_2
\end{aligned}$$

Fig. 5.3: Finding the interior fields in the cascaded metasurface design procedure. An inverse source problem is solved to find equivalent electric and magnetic currents \vec{J}_1 and \vec{M}_1 on $x = -d^+$ such that (i) they result in fields satisfying local power conservation on both metasurfaces, and (ii) they produce null fields on the input side of Σ_1 (Love's condition). The fields \vec{E}_{t1}^+ and \vec{H}_{t1}^+ can then be directly computed from \vec{J}_1 and \vec{M}_1 since Love's condition is enforced, and \vec{E}_{t2}^- and \vec{H}_{t2}^- can be computed as a function of \vec{J}_1 and \vec{M}_1 through forward propagation.

known.

5.3.3 Finding the Tangential Interior Fields (on Σ_1^+ and Σ_2^-)

The next step is to determine the tangential electric and magnetic fields in the interior region between the two metasurfaces. More specifically, we want to determine \vec{E}_{t1}^+ , \vec{H}_{t1}^+ , \vec{E}_{t2}^- , and \vec{H}_{t2}^- . This is accomplished by formulating and solving a second inverse source problem, as shown in Figure 5.3. In this step, the goal is to find equivalent currents \vec{J}_1 and \vec{M}_1 on Σ_1^+ such that

- their corresponding fields satisfy LPC on both metasurfaces, and
- they satisfy Love's condition (i.e., produce null fields on the input side of Σ_1).

On Σ_1 , satisfying LPC means that the time-averaged real power density in the \hat{x} direction of \vec{E}_{t1}^+ and \vec{H}_{t1}^+ must be *locally* (i.e., at each unit cell) equal to the real power density of \vec{E}_{t1}^- and \vec{H}_{t1}^- in the \hat{x} direction. As in Section 5.3.1, since Love's condition is enforced the necessary fields are related to the equivalent currents as

$$\vec{H}_{t1}^+ = -\hat{n} \times \vec{J}_1 \quad \text{and} \quad \vec{E}_{t1}^+ = \hat{n} \times \vec{M}_1. \quad (5.2)$$

On Σ_2 , the LPC condition is between the previously defined output fields, $\vec{E}_{t_2}^+$ and $\vec{H}_{t_2}^+$, (Section 5.3.1) and the fields $\vec{E}_{t_2}^-$ and $\vec{H}_{t_2}^-$ that are computed from the equivalent currents using

$$\vec{E}_{t_2}^- = B_1(\vec{J}_1, \vec{M}_1) \quad \text{and} \quad \vec{H}_{t_2}^- = B_2(\vec{J}_1, \vec{M}_1) \quad (5.3)$$

where B_1 and B_2 are electric and magnetic field integral equation operators, respectively.

5.3.4 Computing the Susceptibility Profiles

Once the tangential fields on either side of both metasurfaces have been found, we can apply the generalized sheet transition conditions (GSTCs) [35, 36] to compute the surface susceptibility profiles required to support the two intended field discontinuities. Assuming a planar metasurface on the zy plane for convenience, adopting a time-dependency of $e^{j\omega t}$, and neglecting the normal component of the polarization densities for mathematical simplicity, the GSTCs (at a given location on the metasurface) can be simplified to

$$\begin{aligned} \begin{pmatrix} -\Delta H_z \\ \Delta H_y \end{pmatrix} &= j\omega\epsilon_0 \begin{pmatrix} \chi_{ee}^{yy} & \chi_{ee}^{yz} \\ \chi_{ee}^{zy} & \chi_{ee}^{zz} \end{pmatrix} \begin{pmatrix} E_{y,av} \\ E_{z,av} \end{pmatrix} \\ &+ j\omega\sqrt{\epsilon_0\mu_0} \begin{pmatrix} \chi_{em}^{yy} & \chi_{em}^{yz} \\ \chi_{em}^{zy} & \chi_{em}^{zz} \end{pmatrix} \begin{pmatrix} H_{y,av} \\ H_{z,av} \end{pmatrix} \end{aligned} \quad (5.4a)$$

$$\begin{aligned} \begin{pmatrix} -\Delta E_y \\ \Delta E_z \end{pmatrix} &= j\omega\mu_0 \begin{pmatrix} \chi_{mm}^{zz} & \chi_{mm}^{zy} \\ \chi_{mm}^{yz} & \chi_{mm}^{yy} \end{pmatrix} \begin{pmatrix} H_{z,av} \\ H_{y,av} \end{pmatrix} \\ &+ j\omega\sqrt{\epsilon_0\mu_0} \begin{pmatrix} \chi_{me}^{zz} & \chi_{me}^{zy} \\ \chi_{me}^{yz} & \chi_{me}^{yy} \end{pmatrix} \begin{pmatrix} E_{z,av} \\ E_{y,av} \end{pmatrix} \end{aligned} \quad (5.4b)$$

where ϵ_0 and μ_0 are the permittivity and permeability of free space and ω is the angular frequency of the time-harmonic fields. The surface susceptibility terms χ represent the

electric/magnetic response (denoted by the first subscript) to an applied electric/magnetic field (denoted by the second subscript). For a metasurface Σ_i (where $i = 1, 2$), the average and difference fields in (5.4) are defined as

$$\vec{\Psi}_{\text{av}} \triangleq \frac{\vec{\Psi}^+ + \vec{\Psi}^-}{2}, \quad (5.5)$$

$$\Delta\vec{\Psi} \triangleq \vec{\Psi}^+ - \vec{\Psi}^-. \quad (5.6)$$

Under certain assumptions that are discussed in Section 5.4.8, (5.4) can be reduced to a well-defined system of equations that allows for (unique) analytic solutions for the remaining susceptibility terms for a given field transformation.

5.3.5 Conversion to Three-Layer Admittance Model

Once the surface susceptibility profiles are determined, they can be converted to three-layer admittance-sheet models given a known dielectric substrate (e.g., Rogers RO3010). This can then be simulated (in our case, in ANSYS HFSS) to verify the performance of the cascaded metasurface system with respect to the desired specifications.

5.4 Numerical Implementation of the Inversion Framework

The following numerical implementation assumes one-dimensional (1D) metasurfaces along y with TE_z fields propagating in the xy plane, although the framework can be applied to other geometries and field polarizations. (A TM_z example will be considered in Section 5.5.3.)

5.4.1 Forming the Vectors of Unknowns

Each of the two metasurfaces is discretized into N unit cells (although the number of unit cells on each metasurface do not necessarily need to be equal). The discrete vector of

unknown electric and magnetic equivalent currents on Σ_1^+ (i.e., on $x = -d^+$) are denoted as \mathbf{J}_1 and \mathbf{M}_1 , which are discretized using pulse basis functions. These currents can be written in terms of their real and imaginary components as

$$\mathbf{J}_1 = \mathbf{J}_{1,R} + j\mathbf{J}_{1,I}, \quad (5.7a)$$

$$\mathbf{M}_1 = \mathbf{M}_{1,R} + j\mathbf{M}_{1,I}. \quad (5.7b)$$

We then concatenate the real and imaginary parts of the equivalent currents into a single vector as

$$\mathbf{x}_1 = \begin{bmatrix} \mathbf{J}_{1,R} \\ \mathbf{J}_{1,I} \\ \mathbf{M}_{1,R} \\ \mathbf{M}_{1,I} \end{bmatrix}. \quad (5.8)$$

With this organization, we will perform the optimization over the real-valued vector $\mathbf{x}_1 \in \mathbb{R}^{4N}$ as opposed to the complex equivalent currents $\mathbf{J}_1 \in \mathbb{C}^N$ and $\mathbf{M}_1 \in \mathbb{C}^N$.

Similarly, we perform the same re-structuring of the equivalent currents \mathbf{J}_2 and \mathbf{M}_2 on Σ_2^+ (i.e., on $x = 0^+$) to construct the second vector of unknowns denoted by \mathbf{x}_2 . Assuming the same number of unit cells for the second metasurface, then $\mathbf{x}_2 \in \mathbb{R}^{4N}$. Thus, we now have two (real-valued) vectors of unknowns: \mathbf{x}_1 and \mathbf{x}_2 . As described in Section 5.3, we first need to determine \mathbf{x}_2 , and then we will find \mathbf{x}_1 .

5.4.2 Forming a Data Misfit Cost Functional with Respect to \mathbf{x}_2

The first step involves forming a cost functional with respect to \mathbf{x}_2 representing the discrepancy between the field specifications \mathbf{f} on S and those generated by a predicted \mathbf{x}_2 , as described in Section 5.3.1. Let us assume that the desired specifications are *amplitude-only* fields on S . That is, let the field specifications be $|\mathbf{f}|^2$, a vector of squared field amplitude at

the discrete points of S .⁵ The data misfit cost functional associated with these specifications is a mapping from \mathbb{R}^{4N} to \mathbb{R} , and is expressed as

$$\mathcal{C}_F(\mathbf{x}_2) = \frac{\left\| \left| \mathbf{A}_F \mathbf{x}_2 \right|^2 - |\mathbf{f}|^2 \right\|_S^2}{\left\| |\mathbf{f}|^2 \right\|_S^2} \quad (5.9)$$

where \mathbf{A}_F is the discretized integral operator that produces H_z (due to TE_z assumption) at the locations of the field specifications in $|\mathbf{f}|^2$ from the equivalent currents \mathbf{x}_2 .⁶ In addition, $\|\cdot\|_S$ denotes the L^2 -norm over S . Later on in Section 5.5.2, we will also consider a case where the field specifications are instead FF performance criteria such as main beam directions, HPBW, and null locations. In such cases, these performance criteria need to be first converted to a desired power pattern $|\mathbf{f}|^2$. The procedure to perform this conversion can be found in [27].

In addition, we note that Love's condition is enforced by adding 'virtual' data to $|\mathbf{f}|^2$. This is done by augmenting $|\mathbf{f}|^2$ with a vector of zeros associated with the locations at which the null field is desired [27].⁷ In this case, these nulls are enforced on a line which is an inward offset of Σ_2^+ . Similarly, the matrix \mathbf{A}_F needs to be augmented such that when it operates on \mathbf{x}_2 , the operation results in the fields produced by \mathbf{x}_2 in the region of interest S and in the domain at which null fields are desired.

5.4.3 Forming a Total Variation Regularizer for \mathbf{x}_2

We now introduce an additive regularization term based on the L^2 -norm total variation (TV) regularizer.⁸ This functional, denoted as $\mathcal{C}_{\text{TV}}(\mathbf{x}_2)$, penalizes solutions with rapid

⁵If S is in the FF zone, $|\mathbf{f}|^2$ is equivalent to a desired power pattern.

⁶For TM_z polarization the operator \mathbf{A}_F would produce E_z at the locations of interest.

⁷This method to enforce Love's condition is based on enforcing Love's condition for the antenna diagnostic problem [58], and has also been applied to the inverse source problem associated with noninvasive specific absorption rate characterization [115].

⁸The TV regularizer has been utilized in microwave imaging [113, 114], NF antenna measurements and characterization [55] as well as image deblurring [44]. One of the reasons behind the use of this regularizer

variations from cell to cell and is given by

$$\mathcal{C}_{\text{TV}}(\mathbf{x}_2) = w_{\text{TV}} \left\| D_y(\mathbf{x}_2) \right\|_{\Sigma_2}^2 \quad (5.10)$$

where w_{TV} is a positive, real-valued weighting parameter. The operator $D_y(\mathbf{x}_2)$ applies a derivative with respect to y to the individual current components of \mathbf{x}_2 *separately*, explicitly stated as

$$D_y(\mathbf{x}_2) \triangleq \begin{bmatrix} \frac{\partial}{\partial y} \mathbf{J}_{2,\text{R}} \\ \frac{\partial}{\partial y} \mathbf{J}_{2,\text{I}} \\ \frac{\partial}{\partial y} \mathbf{M}_{2,\text{R}} \\ \frac{\partial}{\partial y} \mathbf{M}_{2,\text{I}} \end{bmatrix}. \quad (5.11)$$

This regularization term biases the solution towards a higher degree of continuity in the equivalent currents (and therefore fields) from unit cell to unit cell. This is especially important because in Section 5.4.6 we attempt to enforce LPC at each unit cell on Σ_2 , and rapidly varying output fields could make this more difficult or even practically impossible given the finite size of the metasurfaces. A secondary benefit of this regularization is apparent when considering *microscopic* metasurface design, i.e., physical unit cell design. The typical procedure is to design and simulate each unit cell independently while assuming infinite periodicity. When each unit cell is then placed in the final metasurface structure, the periodicity assumption no longer holds and the behaviour of each unit cell will deviate from what was expected. The inclusion of the TV regularizer reduces the variation from cell to cell, thereby reducing the error in the periodicity assumption.

has been its smoothing effect on spurious oscillations in the reconstructed solution.

5.4.4 Iterative Minimization to Determine \mathbf{x}_2

Combining (5.9) and (5.10) results in the total cost functional (or, regularized cost functional) that is minimized during the first optimization step, explicitly written as

$$\mathcal{C}_2(\mathbf{x}_2) = \underbrace{\mathcal{C}_F(\mathbf{x}_2)}_{\text{data misfit}} + \underbrace{\mathcal{C}_{\text{TV}}(\mathbf{x}_2)}_{\text{smoothness}}. \quad (5.12)$$

Minimization is performed in an iterative fashion using the nonlinear conjugate gradient (CG) method. At the k^{th} iteration, the solution update is

$$\mathbf{x}_{2,k+1} = \mathbf{x}_{2,k} - \beta_k \mathbf{v}_k \quad (5.13)$$

where β_k is the real-valued step length. The vector \mathbf{v}_k is the CG direction at the k^{th} CG iteration computed using Polak-Ribière formula [116]

$$\mathbf{v}_k = \mathbf{g}_{2,k} + \frac{\left(\mathbf{g}_{2,k} - \mathbf{g}_{2,k-1}\right)^{\text{H}} \mathbf{g}_{2,k}}{\mathbf{g}_{2,k-1}^{\text{H}} \mathbf{g}_{2,k-1}} \mathbf{v}_{k-1} \quad (5.14)$$

where the superscript “H” denotes the Hermitian operation (complex conjugate transpose) and $\mathbf{g}_{2,k}$ and $\mathbf{g}_{2,k-1}$ are the gradients of (5.12) with respect to \mathbf{x}_2 at iterations k and $k-1$, respectively. As derived in Appendix B.1, the gradient of (5.12) consists of a sum of the gradients of (5.9) and (5.10), the first of which is

$$\mathbf{g}_F(\mathbf{x}_2) = \frac{4 \operatorname{Re} \left\{ \mathbf{A}_F^{\text{H}} (\mathbf{r}_F \odot \mathbf{A}_F \mathbf{x}_2) \right\}}{\left\| |\mathbf{f}|^2 \right\|_S^2} \quad (5.15)$$

where “ \odot ” indicates a Hadamard (element-wise) product of two vectors, ‘Re’ denotes the real-part operator⁹, and \mathbf{r}_F is the residual vector of (5.9) defined as

$$\mathbf{r}_F = |\mathbf{A}_F \mathbf{x}_2|^2 - |\mathbf{f}|^2. \quad (5.16)$$

The gradient of (5.10) is

$$\mathbf{g}_{TV}(\mathbf{x}_2) = -2w_{TV} D_y^2(\mathbf{x}_2) \quad (5.17)$$

where the operator $D_y^2(\mathbf{x}_2)$, similarly to (5.11), computes the second derivative (Laplacian operator) with respect to y of the individual current components of \mathbf{x}_2 as

$$D_y^2(\mathbf{x}_2) = \begin{bmatrix} \frac{\partial^2}{\partial y^2} \mathbf{J}_{2,R} \\ \frac{\partial^2}{\partial y^2} \mathbf{J}_{2,I} \\ \frac{\partial^2}{\partial y^2} \mathbf{M}_{2,R} \\ \frac{\partial^2}{\partial y^2} \mathbf{M}_{2,I} \end{bmatrix}. \quad (5.18)$$

Both (5.11) and (5.18) are evaluated numerically using a central difference approximation. The gradient required to update the search direction \mathbf{v}_k in (5.14) can then be computed as

$$\mathbf{g}_{2,k} = \mathbf{g}_{F,k} + \mathbf{g}_{TV,k}. \quad (5.19)$$

The iterative procedure is terminated when the change in \mathbf{x}_2 from one CG iteration to the next drops below a preset tolerance, typically $10^{-6}\%$.

5.4.5 Enforcing Total Power Conservation to Find \vec{E}_{t2}^+ and \vec{H}_{t2}^+

Upon convergence, the above iterative minimization will yield an appropriate \mathbf{x}_2 . Subsequently, we need to compute the corresponding output tangential fields on Σ_2^+ using (5.1).

⁹Note that since \mathbf{x}_2 is a real-valued vector, the gradient vector is naturally a real-valued vector as well.

However, we first need to compute the scaling factor α needed to ensure that *total* power is conserved across the cascaded metasurface structure. (Note that this is *not* related to LPC.) To do so, we define $\mathbf{p}_{\text{in}} \in \mathbb{R}^N$ as the vector containing the time-average real power density of \vec{E}_{t1}^- and \vec{H}_{t1}^- normally incident on Σ_1^- , with the i^{th} element of \mathbf{p}_{in} defined as

$$p_{\text{in},i} = \frac{1}{2} \text{Re} \left(\vec{E}_{t1}^- \times \vec{H}_{t1}^{-*} \right) \Big|_{\text{unit cell } i}. \quad (5.20)$$

In this work, as noted in Section 5.3.2, we are concerned with the design of reflectionless metasurfaces, and thus \mathbf{p}_{in} can be calculated with knowledge of the incident field only. Thus, the *total* incident power will be the summation of the elements of the vector \mathbf{p}_{in} , denoted by $\text{sum}(\mathbf{p}_{\text{in}})$.

The time-average real power density of the output field can be written in terms of the equivalent currents as [102]

$$\mathbf{p}_{\text{out}} = \frac{\alpha^2}{2} (\mathbf{J}_{2,\text{R}} \odot \mathbf{M}_{2,\text{R}} + \mathbf{J}_{2,\text{I}} \odot \mathbf{M}_{2,\text{I}}). \quad (5.21)$$

Note that since we have already determined \mathbf{x}_2 in the previous step, we also know $\mathbf{J}_{2,\text{R}}$, $\mathbf{J}_{2,\text{I}}$, $\mathbf{M}_{2,\text{R}}$, and $\mathbf{M}_{2,\text{I}}$ based on the concatenation convention used in Section 5.4.1. If we now select α to be

$$\alpha = \left(\frac{\text{sum}(\mathbf{p}_{\text{in}})}{\frac{1}{2} \text{sum}(\mathbf{J}_{2,\text{R}} \odot \mathbf{M}_{2,\text{R}} + \mathbf{J}_{2,\text{I}} \odot \mathbf{M}_{2,\text{I}})} \right)^{\frac{1}{2}} \quad (5.22)$$

the total output power will be equal to the total input power, i.e., $\text{sum}(\mathbf{p}_{\text{out}}) = \text{sum}(\mathbf{p}_{\text{in}})$, thus satisfying TPC. Finally, having found α , $\mathbf{J}_{2,\text{R}}$, $\mathbf{J}_{2,\text{I}}$, $\mathbf{M}_{2,\text{R}}$, and $\mathbf{M}_{2,\text{I}}$, we can now use (5.1) to determine the required output tangential fields \vec{E}_{t2}^+ and \vec{H}_{t2}^+ on Σ_2^+ .

5.4.6 Forming the Cost Functional with Respect to \mathbf{x}_1

The goal of this step is to determine the fields between the two metasurfaces that result in the intended field transformation while satisfying LPC on both metasurfaces. This involves finding a set of equivalent electric and magnetic currents \mathbf{x}_1 on Σ_1^+ , as defined in (5.8). The cost functional that is minimized in this step is a mapping from \mathbb{R}^{4N} to \mathbb{R} , and consists of the following four terms

$$\mathcal{C}_1(\mathbf{x}_1) = \underbrace{\mathcal{C}_L(\mathbf{x}_1)}_{\text{Love's cond.}} + \underbrace{\mathcal{C}_{P,1}(\mathbf{x}_1)}_{\text{LPC on } \Sigma_1} + \underbrace{\mathcal{C}_{P,2}(\mathbf{x}_1)}_{\text{LPC on } \Sigma_2} + \underbrace{\mathcal{C}_{TV}(\mathbf{x}_1)}_{\text{smoothness}}. \quad (5.23)$$

The first term is defined as

$$\mathcal{C}_L(\mathbf{x}_1) = \|\mathbf{L}\mathbf{x}_1\|_{\Sigma_1}^2 \quad (5.24)$$

where \mathbf{L} is a discretized integral operator that produces H_z (or E_z for TM_z polarization) at the locations where Love's condition is enforced for Σ_1 . Minimizing $\mathcal{C}_L(\mathbf{x}_1)$ ensures that \mathbf{x}_1 produces null fields on the $x = -d^-$ side of Σ_1 and allows us to eventually calculate the tangential fields using (5.2).

The second term in (5.23) is used to enforce LPC on Σ_1 as originally derived in [102], defined as

$$\mathcal{C}_{P,1}(\mathbf{x}_1) = \frac{w_1 \left\| \frac{1}{2} (\mathbf{J}_{1,R} \odot \mathbf{M}_{1,R} + \mathbf{J}_{1,I} \odot \mathbf{M}_{1,I}) - \mathbf{p}_{\text{in}} \right\|_{\Sigma_1}^2}{\|\mathbf{p}_{\text{in}}\|_{\Sigma_1}^2} \quad (5.25)$$

which quantifies the difference between the real power density of \vec{E}_{t1}^+ and \vec{H}_{t1}^+ and the known incident power \mathbf{p}_{in} in the x direction at each unit cell on Σ_1 . The real-valued weighting parameter w_1 is included to balance the contribution of $\mathcal{C}_{P,1}(\mathbf{x}_1)$ to the overall functional. As noted earlier, enforcing LPC is necessary to ensure that the resulting metasurface can be implemented in a passive and lossless manner [30].

The third term in (5.23) is used to enforce LPC on Σ_2 . In terms of the tangential fields,

this condition can be explicitly written as

$$\frac{1}{2} \operatorname{Re} \left(\vec{E}_{t2}^- \times \vec{H}_{t2}^{-*} \right) \Big|_{\text{unit cell } i} = \underbrace{\frac{1}{2} \operatorname{Re} \left(\vec{E}_{t2}^+ \times \vec{H}_{t2}^{+*} \right)}_{\text{known from Section 5.4.5}} \Big|_{\text{unit cell } i} \quad (5.26)$$

which must hold at each unit cell on Σ_2 . Analogously to (5.20), we evaluate the right-hand side of (5.26) for *each* unit cell to generate the vector \mathbf{p}_{out} containing the local real power densities of the fields \vec{E}_{t2}^+ and \vec{H}_{t2}^+ . We also define discretized integral operators \mathbf{B}_1 and \mathbf{B}_2 corresponding to (5.3), which respectively produce the tangential electric and magnetic fields at each unit cell on Σ_2^- from the equivalent currents \mathbf{x}_1 . We can then write $\mathcal{C}_{P,2}(\mathbf{x}_1)$ as

$$\mathcal{C}_{P,2}(\mathbf{x}_1) = \frac{w_2 \left\| \frac{1}{2} \operatorname{Re} \left(\mathbf{B}_1 \mathbf{x}_1 \odot \mathbf{B}_2^* \mathbf{x}_1 \right) - \mathbf{p}_{\text{out}} \right\|_{\Sigma_2}^2}{\left\| \mathbf{p}_{\text{out}} \right\|_{\Sigma_2}^2} \quad (5.27)$$

where w_2 is a real-valued weighting parameter. The final term of (5.23), $\mathcal{C}_{\text{TV}}(\mathbf{x}_1)$, performs the same operation previously defined in (5.10) (but operating on \mathbf{x}_1 instead of \mathbf{x}_2).

5.4.7 Iterative Minimization to Determine \mathbf{x}_1

As in Section 5.4.4, the cost functional in (5.23) is minimized using the nonlinear CG method with update equations similar to (5.13) and (5.14). As derived in Appendix B.2, the gradient of $\mathcal{C}_L(\mathbf{x}_1)$ is

$$\mathbf{g}_L(\mathbf{x}_1) = 2 \operatorname{Re} \left(\mathbf{L}^H \mathbf{L} \mathbf{x}_1 \right) \quad (5.28)$$

and the gradient of $\mathcal{C}_{P,1}(\mathbf{x}_1)$ is

$$\mathbf{g}_{P,1}(\mathbf{x}_1) = \frac{w_1}{\left\| \mathbf{p}_{\text{in}} \right\|_{\Sigma_1}^2} \begin{bmatrix} \mathbf{r}_1 \odot \mathbf{M}_{1,R} \\ \mathbf{r}_1 \odot \mathbf{M}_{1,I} \\ \mathbf{r}_1 \odot \mathbf{J}_{1,R} \\ \mathbf{r}_1 \odot \mathbf{J}_{1,I} \end{bmatrix} \quad (5.29)$$

where the residual vector \mathbf{r}_1 is defined as

$$\mathbf{r}_1(\mathbf{x}_1) = \frac{1}{2} (\mathbf{J}_{1,R} \odot \mathbf{M}_{1,R} + \mathbf{J}_{1,I} \odot \mathbf{M}_{1,I}) - \mathbf{p}_{\text{in}}. \quad (5.30)$$

In addition, the gradient of $\mathcal{C}_{P,2}(\mathbf{x}_1)$ is

$$\mathbf{g}_{P,2}(\mathbf{x}_1) = \frac{w_2 \operatorname{Re} \left[\operatorname{diag}(\mathbf{B}_1 \mathbf{x}_1) \mathbf{B}_2^* + \operatorname{diag}(\mathbf{B}_2^* \mathbf{x}_1) \mathbf{B}_1 \right]^T \mathbf{r}_2}{\|\mathbf{p}_{\text{out}}\|_{\Sigma_2}^2} \quad (5.31)$$

where the ‘diag(\cdot)’ operator generates a diagonal matrix from a vector and the superscript ‘T’ denotes the transpose operator. The residual vector \mathbf{r}_2 is defined as

$$\mathbf{r}_2(\mathbf{x}_1) = \frac{1}{2} \operatorname{Re} (\mathbf{B}_1 \mathbf{x}_1 \odot \mathbf{B}_2^* \mathbf{x}_1) - \mathbf{p}_{\text{out}}. \quad (5.32)$$

The gradient of $\mathcal{C}_{TV}(\mathbf{x}_1)$ is computed in the same way as (5.17) but now operates on \mathbf{x}_1 . In addition, the iterative procedure is terminated in the same way as in Section 5.4.4, when the solution stagnates as determined by a preset tolerance.

Finally, once we have a solution for \mathbf{x}_1 , we use (5.2) to compute the tangential fields \vec{E}_{t1}^+ and \vec{H}_{t1}^+ . We then use (5.3) to compute the tangential fields \vec{E}_{t2}^- and \vec{H}_{t2}^- . Let us also remind ourselves that we have already determined \vec{E}_{t2}^+ and \vec{H}_{t2}^+ in Section 5.4.5, and we already know \vec{E}_{t1}^- and \vec{H}_{t1}^- from Section 5.3.2. Therefore, we now know the tangential fields on both metasurfaces, and can therefore determine the required surface susceptibility profiles for both of these metasurfaces. Note that since we have enforced LPC for each individual metasurface, we can implement the desired transformation using lossless and passive metasurfaces.

5.4.8 Susceptibility Profile Calculation

Once the tangential fields have been determined on both sides of Σ_1 and Σ_2 we can apply the GSTCs to determine the required susceptibility profiles to support the two field discontinuities. For 1D metasurfaces along y with TE_z fields, (5.4) at a given unit cell simplifies to¹⁰

$$-\Delta H_z = (j\omega\epsilon_0 E_{y,\text{av}}) \chi_{\text{ee}}^{yy} + (j\omega\sqrt{\mu_0\epsilon_0} H_{z,\text{av}}) \chi_{\text{em}}^{yz} \quad (5.33\text{a})$$

$$-\Delta E_y = (j\omega\mu_0 H_{z,\text{av}}) \chi_{\text{mm}}^{zz} + (j\omega\sqrt{\mu_0\epsilon_0} E_{y,\text{av}}) \chi_{\text{me}}^{zy}. \quad (5.33\text{b})$$

In order to ensure that the metasurfaces are passive, lossless, and reciprocal, we follow the procedure in [102] which stipulates that χ_{ee}^{yy} and χ_{mm}^{zz} are purely real (\mathbb{R}) while χ_{em}^{yz} and χ_{me}^{zy} are purely imaginary (\mathbb{I}). This restriction allows (5.33) to be solved analytically on each metasurface, with a unique solution for each unit cell.¹¹ (Any field transformation that satisfies LPC can be supported in this manner due to the inclusion of magnetoelectric coupling [31].) Note that since (i) we have already satisfied LPC in Section 5.4.6, and (ii) $\chi_{\text{ee}}^{yy} \in \mathbb{R}$ and $\chi_{\text{mm}}^{zz} \in \mathbb{R}$ cannot create any loss, this will inherently result in the losslessness condition $\chi_{\text{me}}^{zy} = (\chi_{\text{em}}^{yz})^*$ [117, Appendix B] being satisfied. Finally, the stipulation $\chi_{\text{me}}^{zy} \in \mathbb{I}$ and $\chi_{\text{em}}^{yz} \in \mathbb{I}$ causes the relation $\chi_{\text{me}}^{zy} = (\chi_{\text{em}}^{yz})^*$ to be equivalent to $\chi_{\text{me}}^{zy} = -\chi_{\text{em}}^{yz}$, which is the condition for reciprocity [117, Sec. 2.2], and therefore the resulting metasurface is also reciprocal. Thus, in the next section, we have substituted χ_{me}^{zy} with $-\chi_{\text{em}}^{yz}$.

5.4.9 Conversion to the Three-Layer Model

In order to simulate the metasurfaces using commercial software, we employ a three-layer admittance sheet topology [107, 100, 118, 110] for each unit cell as shown in Figure 5.4.

¹⁰Later on, in Section 5.5.3, we consider a TM_z case as well.

¹¹Note that (5.33) represents two complex equations which are equivalent to four real equations. On the other hand, since $\chi_{\text{ee}}^{yy} \in \mathbb{R}$, $\chi_{\text{mm}}^{zz} \in \mathbb{R}$, $\chi_{\text{em}}^{yz} \in \mathbb{I}$ and $\chi_{\text{me}}^{zy} \in \mathbb{I}$ we essentially have four real unknowns as well; thus, this system of equations has a unique solution.

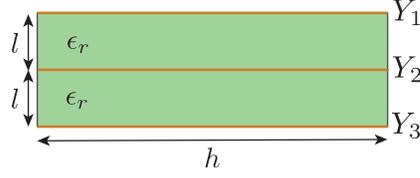


Fig. 5.4: The three-layer admittance sheet topology used for each unit cell [107]. Two substrate layers with relative permittivity ϵ_r and thickness l are placed between three admittance sheets Y_1 , Y_2 , and Y_3 . Also, h represents the size of the unit cell. (The bondply layer has not been included in this model.) From a circuit viewpoint, this can be seen as five cascaded two-port networks: three shunt admittances (Y_1 , Y_2 , and Y_3) and two transmission line sections of length l [61].

The conversion from the susceptibility profiles to the admittance sheet model is performed using the procedure in [106]. At a given unit cell, we start by rearranging (5.33) into the ABCD representation of a two-port network as

$$\begin{bmatrix} \vec{E}_{ti}^- \\ \vec{H}_{ti}^- \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \vec{E}_{ti}^+ \\ \vec{H}_{ti}^+ \end{bmatrix}, \quad (5.34)$$

where $i = 1, 2$ for the first and second metasurfaces respectively. This results in the following ABCD parameters in terms of the surface susceptibility values

$$A = G^{-1} \left[\frac{k_0^2}{4} \chi_{ee}^{yy} \chi_{mm}^{zz} - \left(1 - \frac{jk_0}{2} \chi_{em}^{yz} \right)^2 \right] \quad (5.35a)$$

$$B = G^{-1} [-j\omega\mu_0\chi_{mm}^{zz}] \quad (5.35b)$$

$$C = G^{-1} [-j\omega\epsilon_0\chi_{ee}^{yy}] \quad (5.35c)$$

$$D = G^{-1} \left[\frac{k_0^2}{4} \chi_{ee}^{yy} \chi_{mm}^{zz} - \left(1 + \frac{jk_0}{2} \chi_{em}^{yz} \right)^2 \right] \quad (5.35d)$$

where G is

$$G = - \left(\frac{k_0}{2} \chi_{em}^{yz} \right)^2 - \frac{k_0^2}{4} \chi_{ee}^{yy} \chi_{mm}^{zz} - 1. \quad (5.36)$$

It is worthwhile to note that $AD - BC = 1$ (the criterion for a reciprocal network). In addition, note that since $\chi_{ee}^{yy} \in \mathbb{R}$, $\chi_{mm}^{zz} \in \mathbb{R}$ and $\chi_{em}^{yz} \in \mathbb{I}$, then we will have $A \in \mathbb{R}$, $D \in \mathbb{R}$, $B \in \mathbb{I}$, and $C \in \mathbb{I}$, which further indicates a lossless two-port network.

The $ABCD$ parameters of the three-layer admittance sheet topology can be computed using transmission line theory [107] and equated with (5.35), allowing the three admittance sheet values Y_1 , Y_2 , and Y_3 to be calculated as¹²

$$Y_1 = \frac{D - jZ_0 \sin(\beta l) \cos(\beta l) Y_2 - \cos^2(\beta l) + \sin^2(\beta l)}{2jZ_0 \sin(\beta l) \cos(\beta l) - Z_0^2 \sin^2(\beta l) Y_2} \quad (5.37a)$$

$$Y_2 = \frac{B - 2jZ_0 \sin(\beta l) \cos(\beta l)}{-Z_0^2 \sin^2(\beta l)} \quad (5.37b)$$

$$Y_3 = \frac{A - jZ_0 \sin(\beta l) \cos(\beta l) Y_2 - \cos^2(\beta l) + \sin^2(\beta l)}{2jZ_0 \sin(\beta l) \cos(\beta l) - Z_0^2 \sin^2(\beta l) Y_2} \quad (5.37c)$$

where Z_0 , β , and l are the characteristic impedance, propagation constant, and thickness of the substrate layers, respectively. Since A and D are purely real and B and C are purely imaginary and noting that Z_0 and β are purely real (assuming a lossless substrate), we will have purely imaginary Y_1 , Y_2 and Y_3 . Therefore, from a circuit point of view, Y_1 , Y_2 and Y_3 represent either capacitive or inductive elements. In the next section, we calculate Y_1 , Y_2 and Y_3 for each unit cell on each metasurface, and subsequently simulate our cascaded metasurface system in ANSYS HFSS.

5.5 Full-Wave Simulated Examples

The examples presented here use ANSYS HFSS to simulate the cascaded metasurfaces designed using the proposed procedure along with the three-layer admittance sheet topology for each unit cell.

¹²Similar relations can also be found in the Supplemental Material of [110].

5.5.1 Phaseless Far-Field Power Pattern

In the first example we design a pair of metasurfaces to produce a desired FF power pattern from a normally incident TE_z plane wave (i.e., only E_x , E_y , and H_z field components) at 10.5 GHz. We assume 2D field propagation in the xy plane with 1D metasurfaces placed along $x = 0$ and $x = -1.5\lambda$, where λ denotes the free space wavelength. The metasurfaces are both 5λ in length and each metasurface is discretized into 30 unit cells of width $\lambda/6$. While the normally incident plane wave requires periodic boundaries (implemented through the use of master/slave boundaries in HFSS), the designed metasurface will actually be aperiodic. To accommodate the simulation of an aperiodic structure within periodic boundary conditions, we add absorbing elements in-line with both metasurfaces that extend out a distance of two wavelengths from either end. These absorbers ensure that the fields *close* to the metasurfaces are not affected by the adjacent “images” of the structure due to the periodic boundaries.

The field specifications consist of the phaseless FF power pattern shown in Figure 5.5, specified for the angular range $-90^\circ \leq \varphi \leq 90^\circ$. In the first optimization step, we use $w_{\text{TV}} = 10^{-12}$ and a step length of 100 and minimize (5.12) to determine the equivalent currents \mathbf{x}_2 on the *output* metasurface located at $x = 0$. The tangential output fields are then computed and scaled to match the total incident field power using (5.1) and (5.22). The second optimization step is then performed to minimize (5.23). The weighting parameters used in this step are $w_1 = 0.5$, $w_2 = 0.7$, and $w_{\text{TV}} = 8 \times 10^{-10}$, and a step length of 1 is used.¹³ Once \mathbf{x}_1 is determined, the remaining tangential fields are found using (5.2) and (5.3) and the susceptibility profiles can be computed using (5.33).

In this design, we use the Rogers RO3010 substrate ($\epsilon_r = 10.2$, $\tan \delta = 0.0022$)¹⁴ for each

¹³These weighting parameters—and those of following examples—are determined in an *ad-hoc* fashion to ensure the solution is not over-smoothed at the expense of satisfying local power conservation or the enforcement of Love’s condition. Additionally, the step length is selected to be as large as possible while avoiding any divergence behaviour.

¹⁴Since this loss tangent is small, it has not been taken into account in the calculation of the required admittances for each unit cell of each metasurface (Section 5.4.9). However, the presence of this loss tangent

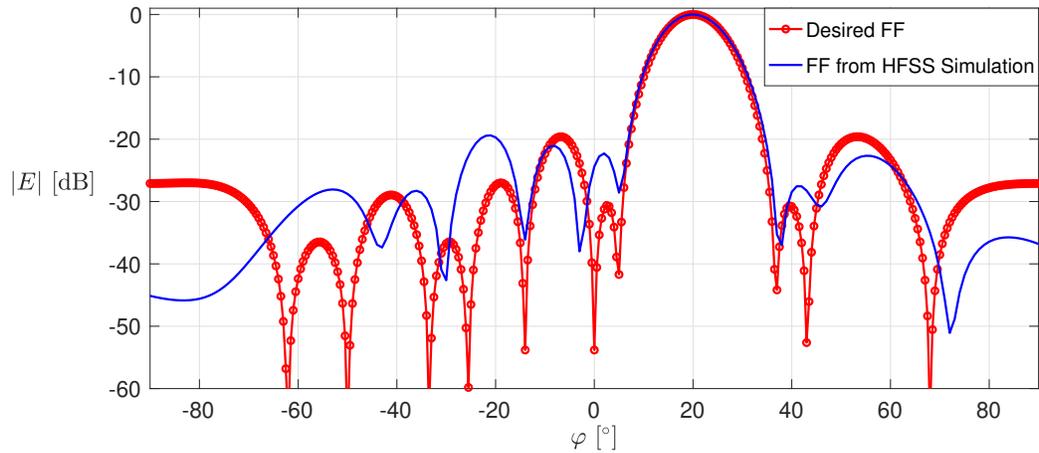


Fig. 5.5: Comparison of the specified desired far-field power pattern (solid red curve with circular markers) and the normalized far-field power pattern produced by the ANSYS HFSS simulation of the designed cascaded metasurfaces (solid blue curve).

of the two layers, each with a thickness of 50 mil. We now calculate the admittance sheet values Y_1 , Y_2 , and Y_3 for each unit cell using (5.35) through (5.37). The admittance sheets are then implemented in the HFSS simulation using impedance boundary conditions, with metallic baffles placed between each the unit cells as in [61]. The real part (absolute value) of the total electric field in the simulation domain is shown in Figure 5.6. The FF pattern generated from this simulation (using a NF to FF transformation) is shown in Figure 5.5, in which the simulated FF pattern exhibits excellent agreement with the desired FF pattern within the main beam. Although the generated sidelobes are not exactly matched with the desired sidelobes, the generated sidelobe level (SLL) remains almost below -20 dB. The overall transmission efficiency, which we define as the ratio between the total real power of the output field to the total real power of the incident field, is 77.76% in this example.

has been taken into account in ANSYS HFSS simulations.

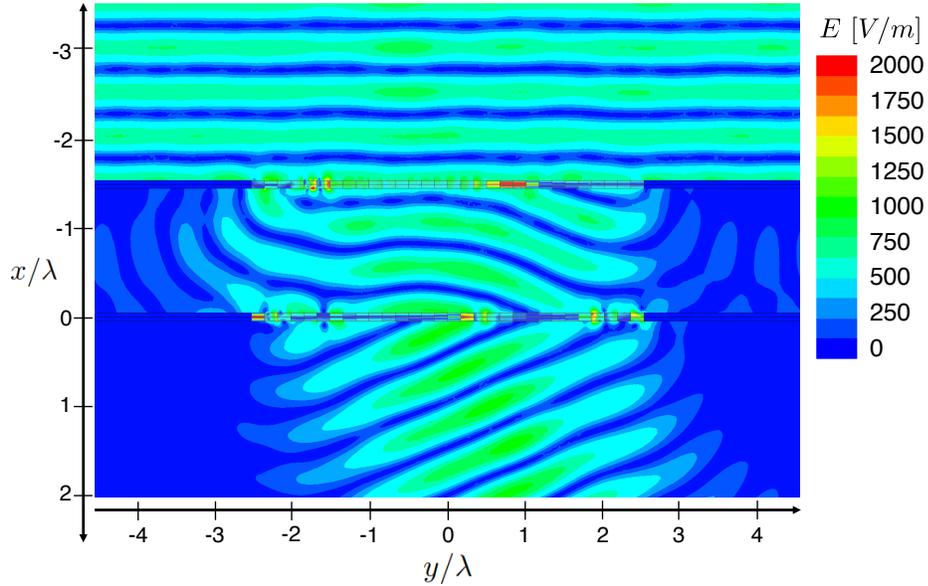


Fig. 5.6: Absolute value of the real part of the total electric field when the cascaded metasurface structure designed for the far-field power pattern in Figure 5.5 is illuminated by a normally incident plane wave.

5.5.2 Far-Field Performance Criteria

In the second example, the goal is to design a cascaded metasurface structure that produces a FF pattern that satisfies a collection of performance criteria. The only change required to support this type of field specifications is to replace the functional in (5.9) with

$$\mathcal{C}_{\text{FP}}(\mathbf{x}_2) = \frac{\left\| \mathbf{W} |\mathbf{A}_{\text{F}} \mathbf{x}_2|^2 - \mathbf{W} |\mathbf{f}|^2 \right\|_S^2}{\left\| \mathbf{W} |\mathbf{f}|^2 \right\|_S^2} \quad (5.38)$$

where \mathbf{W} is a diagonal prescaling matrix used to balance the relative contributions of the various performance criteria, calculated as in [27]. The details on converting the FF performance criteria into $|\mathbf{f}|^2$ can also be found in [27]. For this example the target FF performance criteria are shown in Table 5.1, which consist of two identical main beams (in different directions), each with the same HPBW and associated nulls. A visual representation of these performance criteria can be seen in Figure 5.7, in which the elements of \mathbf{f} are plotted using

Table 5.1: Desired Far-field (FF) Performance Criteria

Specifications	Main Beam 1	Main Beam 2
Direction	$\varphi = -34^\circ$	$\varphi = 34^\circ$
HPBW	16°	16°
Nulls (relative to main beam)	22° and 38°	22° and 38°

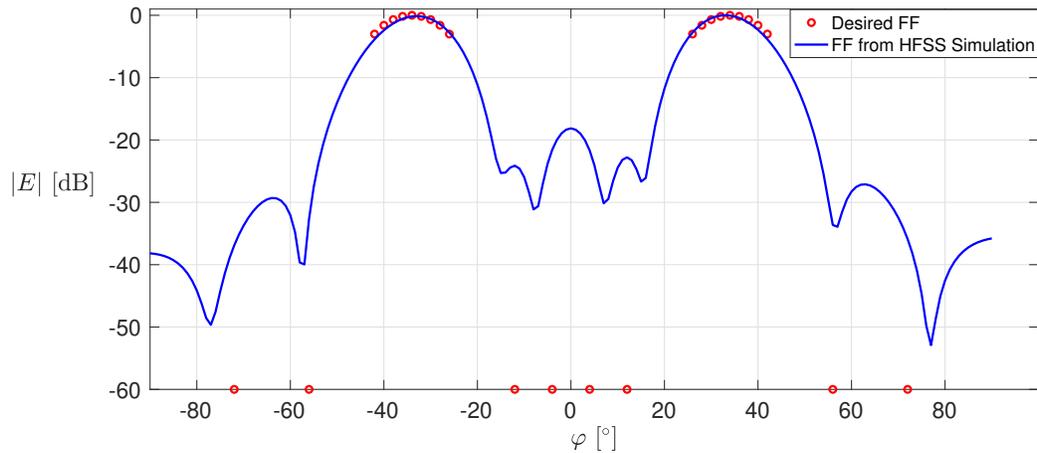


Fig. 5.7: Comparison of the specified far-field performance criteria (red circular markers) and the normalized far-field power pattern produced by the ANSYS HFSS simulation of the designed cascaded metasurfaces (solid blue curve). The red circular markers on the -60 dB line specify the desired null locations.

red circular markers. In particular, note the eight red circular markers on the -60 dB level that indicates the desired eight null locations.

We adopt the same geometry and incident field as the previous example in Section 5.5.1. In the first optimization step, we use parameters of $w_{TV} = 10^{-9}$ and a step length of 10, and in the second optimization step we use weighting parameters $w_1 = 0.5$, $w_2 = 0.7$, $w_{TV} = 8 \times 10^{-10}$, and a step length of 0.1. The real part (absolute value) of the total electric field resulting from simulating the designed metasurfaces in HFSS is shown in Figure 5.8, while the resulting FF is shown in Figure 5.7. The produced FF pattern has two main beams at $\varphi = \pm 33^\circ$ with HPBWs of 15.5° and 15.2° , exhibiting excellent agreement with the field

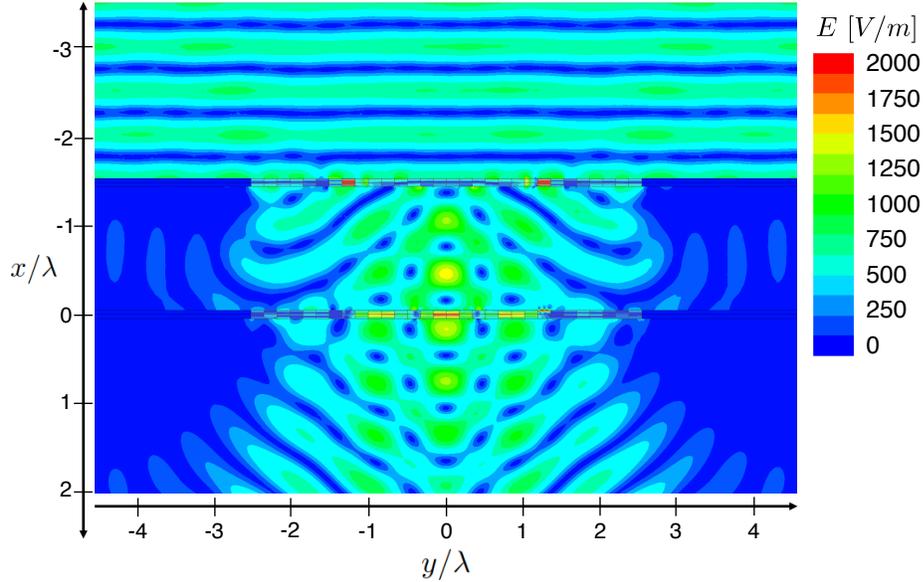


Fig. 5.8: Absolute value of the real part of the total electric field when the cascaded metasurface structure designed for the far-field performance criteria of Table 5.1 is illuminated by a normally incident plane wave.

specifications. The eight desired pattern nulls are also visible, although they deviate from the intended directions by up to 5° . In addition, although we did not achieve very deep nulls (set to -60 dB level in the specifications), all the generated nulls are still below -20 dB with the best null being below -50 dB. Finally, the transmission efficiency observed in this case is 83.31%.

5.5.3 Comparison to a Single Metasurface with TM_z Fields

In this final example, we change the incident field to that of an electric current line source (infinite in the z direction), placed $\lambda/3$ away from the input metasurface. The fields produced by this source are TM_z (i.e., E_z , H_x , and H_y only). The frequency remains at 10.5 GHz, but each metasurface now consists of 50 unit cells of $\lambda/10$ width and the thickness of the Rogers RO3010 substrate has been decreased to 10 mil. Absorbing elements and periodic boundary conditions are not needed with this incident field, and so the simulation

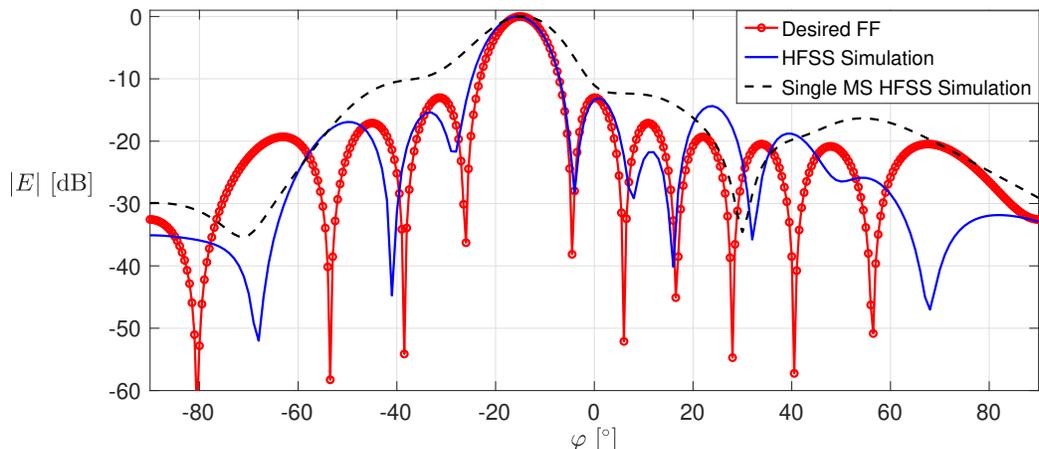


Fig. 5.9: Comparison of the specified far-field power pattern (solid red curve with circular markers) and the normalized far-field power patterns produced by the ANSYS HFSS simulations of the designed cascaded metasurface structure (solid blue curve) and the single metasurface (dashed black curve).

domain is bounded by an appropriate perfectly-matched layer (PML).¹⁵

The field specifications consist of a FF power pattern with a main beam at $\varphi = -15^\circ$ as shown in Figure 5.9. We use the proposed procedure to design a cascaded metasurface structure, with $w_{\text{TV}} = 10^{-5}$ and a step length of 0.005 in the first optimization step. In the second optimization step we use $w_1 = 25$, $w_2 = 25$, and $w_{\text{TV}} = 5 \times 10^{-11}$, and a step length of 1. Additionally, we attempt to design a *single* metasurface using the procedure in [102] to produce a satisfactory FF pattern from the same incident field. The real part (absolute value) of the total magnetic field in the simulation domains for the single metasurface and the cascaded metasurface system are shown in Figures 5.10(a) and (b), respectively. In addition, the corresponding FF patterns for the single metasurface (dashed black curve) and cascaded metasurface system (solid blue curve) are shown in Figure 5.9. While the single metasurface has a higher transmission efficiency (88.00% compared to 72.36%), the FF pattern produced by the cascaded metasurfaces is significantly closer to the target pattern.

¹⁵Periodic master/slave boundaries are still present on the faces perpendicular to the z axis due to the 2D nature of the simulation.

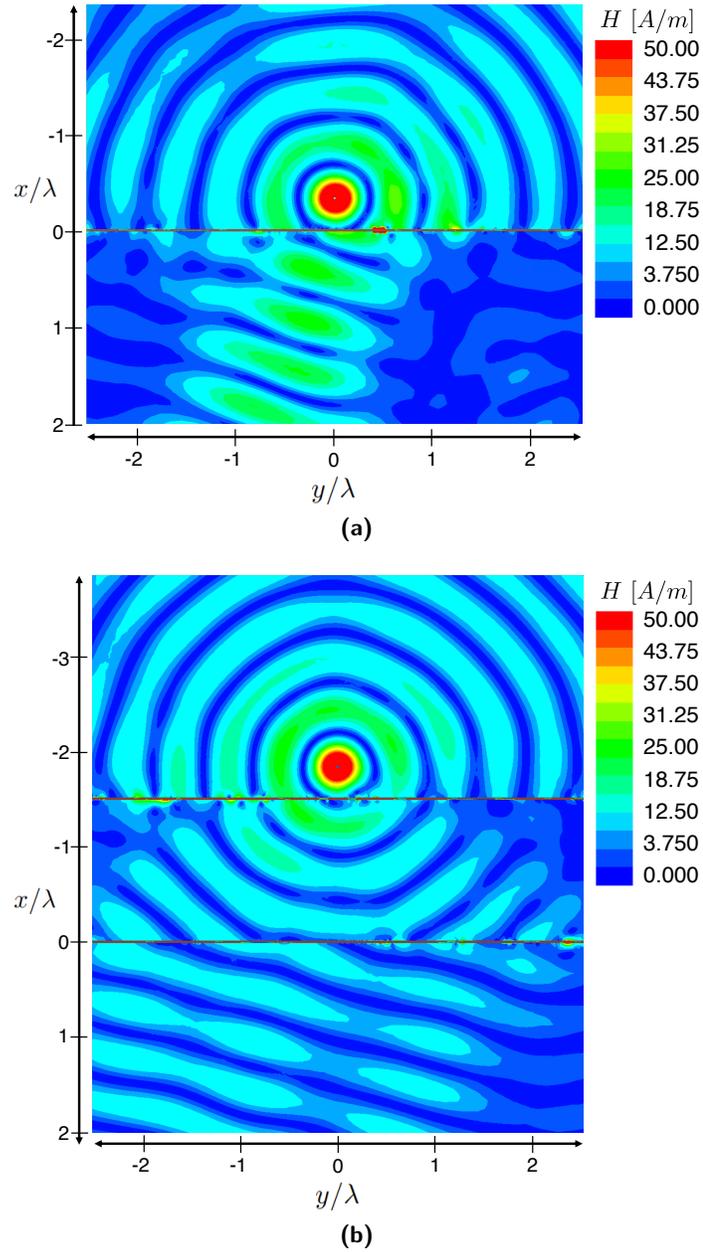


Fig. 5.10: Absolute value of the real part of the total magnetic field of the HFSS simulations using (a) a single metasurface and (b) two metasurfaces to produce the desired far-field power pattern shown in Figure 5.9 with an incident field produced by an electric current line source.

The narrow beamwidth requires a rather uniform field amplitude distributed over a large aperture, which a single metasurface is unable to produce because it is constrained by the focused input power distribution generated by the line source. In the cascaded scenario, the first metasurface is able to ‘redirect’ the power such that the second metasurface can more easily create the intended output field.

5.6 Limitations

The main limitation of the proposed design method is that we have not yet taken reflections between the two metasurfaces into account, and currently require that the separation be large enough such that reflections can be neglected. However, incorporating reflections into the model can allow for smaller separation between the metasurfaces and improve the transmission efficiency, as shown in [33] and [109]. A secondary limitation is that the CG step length and functional weighting parameters are determined in an *ad-hoc* fashion, and require some tuning to ensure convergence.

5.7 Conclusion

A framework for the design of passive, lossless, and reciprocal cascaded metasurfaces structures was presented. The framework utilizes electromagnetic inversion to allow for the flexibility to specify the desired output field in a variety of ways, including complex fields (amplitude and phase), phaseless fields (amplitude-only), or even in terms of far-field performance criteria such as main beam directions, half-power beamwidths, and null locations. This framework allows for full generality with respect to metasurface geometry and the locations at which the field specifications are defined. Local power conservation is enforced at each metasurface during the two-step optimization procedure, and we have introduced a total variation regularizer to improve the continuity of the resulting fields from unit cell to

unit cell. Several designed metasurfaces were simulated in ANSYS HFSS using a three-layer admittance sheet topology to demonstrate the validity and flexibility of the method.

The next logical step is to *physically* implement the designed metasurfaces using metallic ‘dogbone’ structures [101, 61] or some other unit cell design. The main challenge concerning the physical implementation is compensating for the mutual coupling between the layers of each unit cell. To this end, an iterative method for tuning unit cells to account for mutual coupling was recently presented in [119]. However, a method of this nature may not be practical for the large number of unique unit cells required for the aperiodic designs presented here.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

The main objective of this thesis was to construct a framework for the macroscopic design of electromagnetic metasurfaces that allows for more flexibility in terms of the output field specification. To this end, we formulated the metasurface design problem as an electromagnetic inverse source problem. This formulation allows us to find a set of equivalent electric and magnetic currents that produce the desired output fields, which can be directly related to the required tangential fields through the use of Love's equivalence condition. With the tangential input and output fields known, the macroscopic representation of the metasurface (i.e., susceptibility tensors) can be computed. The development of this framework consists of several novel contributions as listed below.

6.1.1 Summary of Contributions

- **Inverse source framework for metasurface design:** A design framework for approaching metasurface design as an inverse source problem is introduced in Chapter 3. We approach the problem by attempting to find equivalent electric and magnetic cur-

rents, which we have shown to be directly related to the required tangential electric and magnetic fields under certain conditions. An appropriate set of equivalent currents is found through the iterative minimization of a cost functional that quantifies the discrepancy between the fields produced by said currents and the desired fields. The desired fields can be specified in three different forms, including (i) complex fields (amplitude and phase) in either the NF or FF region, (ii) phaseless fields (amplitude-only, often referred to as a ‘power pattern’ if in the FF region) in either the NF or FF region, and (iii) FF performance criteria such as main beam direction(s), HPBW(s), null location(s), and polarization. The flexibility with regards to the field specifications improves the practicality of the method and allows the design procedure to be used for a wider variety of applications. Furthermore, the generality of the inverse source framework allows for the design of non-planar metasurfaces as shown in Chapter 3.

- **Enforcing local power conservation within the inversion algorithm:** Some alternative methods avoid loss and gain by simply removing the lossy or active parts of the resulting susceptibilities or impedances, but this has a detrimental and unavoidable effect on the field transformation. In this work, we modify the inversion algorithm such that the resulting field transformation can *always* be implemented in a passive and lossless way. This is accomplished by ensuring local power is conserved across the metasurface. In Chapter 4, we introduce an additive term to the cost functional which quantifies the difference between the input time-averaged real power density distribution (computed from the known incident field), and the output power distribution (computed from the equivalent currents). Minimizing this additional term along with the original functional results in a field transformation that *can* be implemented without requiring loss and/or gain. Furthermore, we impose restrictions on the susceptibilities to guarantee that the resulting macroscopic representation is in

fact passive and lossless.

This addition to the design framework was tested and validated by simulating the designed metasurfaces using ANSYS HFSS. This was facilitated by converting the homogenized susceptibility model to a three-layer admittance sheet topology for each unit cell.

- **Extension to the design of cascaded metasurfaces:** The inverse source design framework was extended to allow for the design of two cascaded metasurfaces. A single passive and lossless metasurface is limited in terms of the possible field transformations that it can support due to the local power conservation requirement. Adding an additional metasurface allows for a redistribution of power to take place between the input and output metasurfaces, thereby allowing for output fields whose power distributions differ from that of the input field. The concept of utilizing a second metasurface has recently been proposed by other groups [32, 110]. In this work, the use of the inverse source framework to perform the design of both metasurfaces is a novel contribution. Approaching the design of both metasurfaces using this framework incorporates the benefits listed above, and most notably the improved flexibility with regards to the output field specifications. The development of the cascaded design framework also led to the use of gradient-based optimization, which offers improved convergence and is less sensitive to initial parameters compared to global optimization techniques, such as the particle swarm algorithm employed in Chapter 4.

Another contribution that has a significant effect is the addition of the total-variation regularizer. The inclusion of this term in the cost functionals reduces the rapid variations in the equivalent currents (and therefore fields). A smooth output field is especially important in the case of cascaded metasurface design, because the input metasurface must redistribute the input power to satisfy local power conservation on the output metasurface. If the output field contains rapid spatial variations, satisfying

local power conservation becomes difficult or even physically impossible due to the finite size of the input metasurface.

6.2 Future Work

There are several improvements and extensions that can be made upon the work presented in this thesis. These developments can be categorized into (i) physical implementation, (ii) reflection control, and (iii) inversion algorithm improvements.

6.2.1 Physical Implementation

The last remaining step in the design process is to create the individual unit cell structures that exhibit the properties defined by the macroscopic model (i.e., susceptibility distributions). The three-layer admittance sheet topology adopted in Chapters 4 and 5 can naturally be implemented using a triple-dogbone structure [61, 101], although the mutual coupling between layers has been shown to be a major issue. There has been some recent work on ‘tuning’ the dogbone dimensions to account for the mutual coupling [119], but the iterative method may not be feasible for the large number of unit cells in the aperiodic designs presented here.

Once a physical implementation that properly accounts for the mutual coupling is developed, it should be fabricated and its output compared with the initial field specifications. The completion of this step will result in a full end-to-end metasurface design method from initial output field specifications to a physically implemented structure.

6.2.2 Reflection Control

In this work, we concerned ourselves with the design of reflectionless metasurfaces. However, incorporating non-zero reflections into the design may provide additional flexibility, especially in the cascaded metasurface scenario. In the work presented in Chapter 5, we

needed to ensure a large enough separation between the metasurfaces such that multiple reflections in the inner region could be neglected. If non-zero reflections are incorporated into the model, we may be able to reduce the separation distance significantly and improve the efficiency as shown in [33].

6.2.3 Inversion Algorithm Improvements

Lastly, there are several improvements that can be made to the inversion algorithm. The first is that the selection of the step length associated with cost functionals that incorporate LPC is currently performed in an *ad hoc* manner. Furthermore, the step length is fixed at the beginning of the inversion process and remains the same for all iterations of CG. A simple improvement would be to use a *dynamic* step length that takes the current magnitude of the gradient into account. This would allow the step length to increase as convergence is approached, counteracting the effect of ‘vanishing’ gradients near an optimum. A further improvement would be to compute the step length analytically, which would require a line-search minimization along the search direction at each iteration. If possible, this method would provide the optimal choice of step length at each iteration and result in the quickest convergence. In addition, multi-objective optimization can also be investigated for this inverse problem which can enable a more systematic treatment of the problem against multiple design constraints.

Another possible extension would be to add support for inhomogeneous media on the output side of the metasurface. This would involve changing the Green’s function evaluated in the forward operator to that of the new media. With this functionality, the design framework could be used to design metasurfaces that act as a matching layer between two different media while also creating some desired field distribution in the complicated medium on the output side. This could be useful in applications such as microwave imaging, in which it would be useful to ensure maximum transmission efficiency into the object of

interest (e.g., a human limb) from the transmitting antennas located in the surrounding medium. The metasurface could also be designed to provide a NF distribution inside the imaging domain that is advantageous for imaging, such as a highly-focused beam.

References

- [1] D. R. Smith, W. J. Padilla, D. Vier, S. C. Nemat-Nasser, and S. Schultz, “Composite medium with simultaneously negative permeability and permittivity,” *Physical review letters*, vol. 84, no. 18, p. 4184, 2000.
- [2] (2019, October) Metamorphose VI. [Online]. Available: <http://congress2019.metamorphose-vi.org>
- [3] D. Berry, R. Malech, and W. Kennedy, “The reflectarray antenna,” *IEEE Transactions on Antennas and Propagation*, vol. 11, no. 6, pp. 645–651, 1963.
- [4] J. Arnaud, “Resonant-grid quasi-optical diplexer,” *Electronics Letters*, vol. 9, no. 25, pp. 589–590, 1973.
- [5] K. Achouri, “Synthesis and applications of electromagnetic metasurfaces,” Ph.D. dissertation, École Polytechnique de Montréal, 2017.
- [6] C. L. Holloway, E. F. Kuester, J. A. Gordon, J. O’Hara, J. Booth, and D. R. Smith, “An overview of the theory and applications of metasurfaces: The two-dimensional equivalents of metamaterials,” *IEEE Antennas and Propagation Magazine*, vol. 54, no. 2, pp. 10–35, 2012.
- [7] C. Pfeiffer and A. Grbic, “Metamaterial Huygens’ surfaces: Tailoring wave fronts with reflectionless sheets,” *Phys. Rev. Lett.*, vol. 110, p. 197401, May 2013.
- [8] M. Selvanayagam and G. Eleftheriades, “Discontinuous electromagnetic fields using orthogonal electric and magnetic currents for wavefront manipulation,” *Optics Express*, pp. 14 409–14 429, 2013.
- [9] S. Tretyakov, “Metasurfaces for general transformations of electromagnetic fields,” *Philos. Trans. Royal Soc. A: Mathematical, Physical and Engineering Sciences*, vol. 373, no. 2049, p. 20140362, 2015.
- [10] G. Minatti, M. Faenzi, E. Martini, F. Caminita, P. De Vita, D. González-Ovejero, M. Sabbadini, and S. Maci, “Modulated metasurface antennas for space: Synthesis, analysis and realizations,” *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 4, pp. 1288–1300, April 2015.

-
- [11] K. Achouri and C. Caloz, "Design, concepts and applications of electromagnetic metasurfaces," *Nanophotonics*, vol. 7, no. 6, pp. 1095–1116, 2018.
- [12] C. Pfeiffer and A. Grbic, "Bianisotropic metasurfaces for optimal polarization control: Analysis and synthesis," *Physical Review Applied*, vol. 2, no. 4, p. 044011, 2014.
- [13] M. Selvanayagam and G. V. Eleftheriades, "Polarization control using tensor Huygens surfaces," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 12, pp. 6155–6168, 2014.
- [14] V. S. Asadchy, M. Albooyeh, S. N. Tsvetkova, A. Díaz-Rubio, Y. Ra'di, and S. Tretyakov, "Perfect control of reflection and refraction using spatially dispersive metasurfaces," *Physical Review B*, vol. 94, no. 7, p. 075142, 2016.
- [15] Y. Ra'Di, C. Simovski, and S. Tretyakov, "Thin perfect absorbers for electromagnetic waves: theory, design, and realizations," *Physical Review Applied*, vol. 3, no. 3, p. 037001, 2015.
- [16] A. Grbic, R. Merlin, E. M. Thomas, and M. F. Imani, "Near-field plates: Metamaterial surfaces/arrays for subwavelength focusing and probing," *Proceedings of the IEEE*, vol. 99, no. 10, pp. 1806–1815, 2011.
- [17] E. Martini, M. Mencagli, and S. Maci, "Metasurface transformation for surface wave control," *Phil. Trans. R. Soc. A*, vol. 373, no. 2049, p. 20140355, 2015.
- [18] K. Achouri, G. Lavigne, M. A. Salem, and C. Caloz, "Metasurface spatial processor for electromagnetic remote control," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 5, pp. 1759–1767, 2016.
- [19] H. Yang, X. Cao, F. Yang, J. Gao, S. Xu, M. Li, X. Chen, Y. Zhao, Y. Zheng, and S. Li, "A programmable metasurface with dynamic polarization, scattering and focusing control," *Scientific reports*, vol. 6, p. 35692, 2016.
- [20] Q. Ma, G. D. Bai, H. B. Jing, C. Yang, L. Li, and T. J. Cui, "Smart metasurface with self-adaptively reprogrammable functions," *Light: Science & Applications*, vol. 8, no. 1, pp. 1–12, 2019.
- [21] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and road ahead," *arXiv preprint arXiv:2004.09352*, 2020.
- [22] A. Epstein and G. V. Eleftheriades, "Huygens' metasurfaces via the equivalence principle: design and applications," *Journal of the Optical Society of America B*, vol. 33, no. 2, pp. A31–A50, 2016.
-

-
- [23] K. Achouri, M. A. Salem, and C. Caloz, "General metasurface synthesis based on susceptibility tensors," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 7, pp. 2977–2991, 2015.
- [24] T. Niemi, A. O. Karilainen, and S. A. Tretyakov, "Synthesis of polarization transformers," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 6, pp. 3102–3111, 2013.
- [25] S. Pearson and S. V. Hum, "Using augmented Lagrangian methods to design electromagnetic surfaces with far field constraints," in *2020 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, 2020, pp. 969–970.
- [26] N. Bayat and P. Mojabi, "A mathematical framework to analyze the achievable resolution from microwave tomography," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 4, pp. 1484–1489, 2016.
- [27] T. Brown, C. Narendra, Y. Vahabzadeh, C. Caloz, and P. Mojabi, "On the use of electromagnetic inversion for metasurface design," *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 3, pp. 1812–1824, March 2020.
- [28] M. Bodehou, C. Craeye, E. Martini, and I. Huynen, "A quasi-direct method for the surface impedance design of modulated metasurface antennas," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 1, pp. 24–36, 2019.
- [29] J. Budhu and A. Grbic, "Perfectly reflecting metasurface reflectarrays: Mutual coupling modelling between unique elements through homogenization," *IEEE Transactions on Antennas and Propagation*, 2020.
- [30] A. Epstein and G. V. Eleftheriades, "Passive lossless Huygens metasurfaces for conversion of arbitrary source field to directive radiation," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 11, pp. 5680–5695, 2014.
- [31] —, "Arbitrary power-conserving field transformations with passive lossless omega-type bianisotropic metasurfaces," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 9, pp. 3880–3895, 2016.
- [32] A. H. Dorrah and G. V. Eleftheriades, "Bianisotropic Huygens' metasurface pairs for nonlocal power-conserving wave transformations," *IEEE Antennas and Wireless Propagation Letters*, vol. 17, no. 10, pp. 1788–1792, Oct 2018.
- [33] V. G. Ataloglou, A. H. Dorrah, and G. V. Eleftheriades, "Design of compact Huygens' metasurface pairs with multiple reflections for arbitrary wave transformations," *IEEE Transactions on Antennas and Propagation*, 2020.
-

- [34] M. Dehmollaian, G. Lavigne, and C. Caloz, “Comparison of tensor boundary conditions (tbcs) with generalized sheet transition conditions (gstcs),” *arXiv preprint arXiv:1901.10414*, 2019.
- [35] M. M. Idemen, *Discontinuities in the electromagnetic field*. John Wiley & Sons (IEEE Press series on electromagnetic wave theory; 40), 2011.
- [36] E. F. Kuester, M. A. Mohamed, M. Piket-May, and C. L. Holloway, “Averaged transition conditions for electromagnetic fields at a metafilm,” *IEEE Transactions on Antennas and Propagation*, vol. 51, no. 10, pp. 2641–2651, 2003.
- [37] X. Jia, Y. Vahabzadeh, C. Caloz, and F. Yang, “Synthesis of spherical metasurfaces based on susceptibility tensor GSTCs,” *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 4, pp. 2542–2554, April 2019.
- [38] A. Sihvola, A. Viitanen, I. Lindell, and S. Tretyakov, “Electromagnetic waves in chiral and bi-isotropic media,” *Norwood, MA, USA: Artech House*, 1994.
- [39] Y. Vahabzadeh, K. Achouri, and C. Caloz, “Simulation of metasurfaces in finite difference techniques,” *IEEE Trans. Antennas Propag*, vol. 64, no. 11, pp. 4753–4759, 2016.
- [40] K. Achouri and O. J. Martin, “Angular scattering properties of metasurfaces,” *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 1, pp. 432–442, 2019.
- [41] K. Achouri and C. Caloz, “Design, concepts, and applications of electromagnetic metasurfaces,” *Nanophotonics*, vol. 7, no. 6, pp. 1095–1116, 2018.
- [42] V. Asadchy, M. S. Mirmoosa, A. Díaz-Rubio, S. Fan, and S. A. Tretyakov, “Tutorial on electromagnetic nonreciprocity and its origins,” *arXiv preprint arXiv:2001.04848*, 2020.
- [43] J. Kong, *Electromagnetic Wave Theory*. John Wiley & Sons, 1986.
- [44] A. Abubakar, P. M. van den Berg, T. M. Habashy, and H. Braunisch, “A multiplicative regularization approach for deblurring problems,” *Image Processing, IEEE Transactions on*, vol. 13, no. 11, pp. 1524–1532, 2004.
- [45] P. A. Nelson and S.-H. Yoon, “Estimation of acoustic source strength by inverse methods: Part I, conditioning of the inverse problem,” *Journal of sound and vibration*, vol. 233, no. 4, pp. 639–664, 2000.
- [46] R. Ellis and D. Oldenburg, “Applied geophysical inversion,” *Geophysical Journal International*, vol. 116, no. 1, pp. 5–11, 1994.
- [47] S. Schelkunoff, “Some equivalence theorems of electromagnetics and their application to radiation problems,” *Bell System Technical Journal*, vol. 15, no. 1, pp. 92–112, 1936.

-
- [48] R. F. Harrington, *Time-harmonic electromagnetic fields*. McGraw-Hill, 1961.
- [49] C. W. Groetsch, *The theory of Tikhonov regularization for Fredholm equations of the first kind*. Pitman Advanced Publishing Program, 1984, vol. 105.
- [50] P. C. Hansen, *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. SIAM, 1998, vol. 4.
- [51] A. Devaney and G. Sherman, “Nonuniqueness in inverse source and scattering problems,” *IEEE Transactions on Antennas and Propagation*, vol. 30, no. 5, pp. 1034–1037, Sep. 1982.
- [52] P. C. Hansen, “The truncated SVD as a method for regularization,” *BIT Numerical Mathematics*, vol. 27, no. 4, pp. 534–553, 1987.
- [53] A. Abubakar and P. M. van den Berg, “Total variation as a multiplicative constraint for solving inverse problems,” *Image Processing, IEEE Transactions on*, vol. 10, no. 9, pp. 1384–1392, 2001.
- [54] J. R. Shewchuk, “An introduction to the conjugate gradient method without the agonizing pain,” 1994.
- [55] T. Brown, I. Jeffrey, and P. Mojabi, “Multiplicatively regularized source reconstruction method for phaseless planar near-field antenna measurements,” *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 4, pp. 2020–2031, 2017.
- [56] A. Love, “The integration of the equations of propagation of electric waves.” *Proceedings of the Royal Society of London*, vol. 68, no. 442-450, pp. 19–21, 1901.
- [57] L. Foged, L. Scialacqua, F. Saccardi, J. Quijano, G. Vecchi, and M. Sabbadini, “Practical application of the equivalent source method as an antenna diagnostics tool [AMTA corner],” *Antennas and Propagation Magazine, IEEE*, vol. 54, no. 5, pp. 243–249, 2012.
- [58] J. L. A. Quijano and G. Vecchi, “Field and source equivalence in source reconstruction on 3D surfaces,” *Progress In Electromagnetics Research*, vol. 103, pp. 67–100, 2010.
- [59] M. Selvanayagam and G. V. Eleftheriades, “Circuit modeling of Huygens surfaces,” *IEEE Antennas and Wireless Propagation Letters*, vol. 12, pp. 1642–1645, 2013.
- [60] C. Caloz, K. Achouri, G. Lavigne, Y. Vahabzadeh, L. Chen, S. Taravati, and N. Chamanara, “A guided tour in metasurface land: Discontinuity conditions, design and applications,” in *2017 IEEE International Conference on Computational Electromagnetics (ICCEM)*, March 2017, pp. 310–311.
- [61] G. Xu, S. V. Hum, and G. V. Eleftheriades, “Augmented Huygens’ metasurfaces employing baffles for precise control of wave transformations,” *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 11, pp. 6935–6946, June 2019.
-

- [62] N. K. Nikolova, *Introduction to Microwave Imaging*. Cambridge University Press, 2017.
- [63] A. Abubakar, T. M. Habashy, V. L. Druskin, L. Knizhnerman, and D. Alumbaugh, “2.5D forward and inverse modeling for interpreting low-frequency electromagnetic measurements,” *Geophysics*, vol. 73, no. 4, pp. F165–F177, July–Aug 2008.
- [64] P. van den Berg and A. Abubakar, “Optical microscopy imaging using the contrast source inversion method,” *Journal of Modern Optics*, vol. 57, pp. 756–764, 05 2010.
- [65] Y. A. Lopez, F. Las-Heras Andres, M. R. Pino, and T. K. Sarkar, “An improved super-resolution source reconstruction method,” *Instrumentation and Measurement, IEEE Transactions on*, vol. 58, no. 11, pp. 3855–3866, 2009.
- [66] T. Brown, C. Narendra, C. Niu, and P. Mojabi, “On the use of electromagnetic inversion for near-field antenna measurements: A review,” in *2018 IEEE Conference on Antenna Measurements & Applications (CAMA)*. IEEE, 2018, pp. 1–4.
- [67] T. M. Grzegorzcyk, P. M. Meaney, P. A. Kaufman, R. M. diFlorio-Alexander, and K. D. Paulsen, “Fast 3-D tomographic microwave imaging for breast cancer detection,” *IEEE Transactions on Medical Imaging*, vol. 31, no. 8, pp. 1584–1592, Aug 2012.
- [68] C. Narendra, T. Brown, N. Bayat, and P. Mojabi, “Multi-plane magnetic near-field data inversion using the source reconstruction method,” in *2018 18th International Symposium on Antenna Technology and Applied Electromagnetics (ANTEM)*, Aug 2018, pp. 1–2.
- [69] A. H. Golnabi, P. M. Meaney, S. D. Geimer, and K. D. Paulsen, “3-D microwave tomography using the soft prior regularization technique: Evaluation in anatomically realistic MRI-derived numerical breast phantoms,” *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 9, pp. 2566–2575, Sep. 2019.
- [70] O. M. Bucci, I. Catapano, L. Crocco, and T. Isernia, “Synthesis of new variable dielectric profile antennas via inverse scattering techniques: a feasibility study,” *IEEE transactions on antennas and propagation*, vol. 53, no. 4, pp. 1287–1297, 2005.
- [71] R. Palmeri and T. Isernia, “Volumetric invisibility cloaks design through spectral coverage optimization,” *IEEE Access*, vol. 7, pp. 30 860–30 867, 2019.
- [72] L. Di Donato, T. Isernia, G. Labate, and L. Matekovits, “Towards printable natural dielectric cloaks via inverse scattering techniques,” *Scientific Reports*, vol. 7, no. 1, p. 3680, 2017.
- [73] R. Palmeri, M. T. Bevacqua, A. F. Morabito, and T. Isernia, “Design of artificial-material-based antennas using inverse scattering techniques,” *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 12, pp. 7076–7090, Dec 2018.

- [74] O. M. Bucci, G. D’Elia, G. Mazzarella, and G. Panariello, “Antenna pattern synthesis: a new general approach,” *Proceedings of the IEEE*, vol. 82, no. 3, pp. 358–371, March 1994.
- [75] M. Salucci, A. Gelmini, G. Oliveri, N. Anselmi, and A. Massa, “Synthesis of shaped beam reflectarrays with constrained geometry by exploiting nonradiating surface currents,” *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 11, pp. 5805–5817, Nov 2018.
- [76] P. C. Hansen, “Numerical tools for analysis and solution of Fredholm integral equations of the first kind,” *Inverse Problems*, vol. 8, no. 6, pp. 849–872, Dec 1992.
- [77] —, *Discrete inverse problems: insights and algorithms*. SIAM, 2010.
- [78] P. Mojabi and J. LoVetri, “Overview and classification of some regularization techniques for the Gauss-Newton inversion method applied to inverse scattering problems,” *Antennas and Propagation, IEEE Transactions on*, vol. 57, no. 9, pp. 2658–2665, 2009.
- [79] T. Brown, C. Narendra, and P. Mojabi, “On the use of the source reconstruction method for metasurface design,” *12th European Conference on Antennas and Propagation (EuCAP)*, pp. 302–306, April 2018.
- [80] T. Brown, C. Narendra, Y. Vahabzadeh, C. Caloz, and P. Mojabi, “Metasurface design using electromagnetic inversion,” in *2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, 2019, pp. 1817–1818.
- [81] C. Parini, S. Gregson, J. McCormick, and D. Janse van Rensburg, *Theory and Practice of Modern Antenna Range Measurements*. The Institution of Engineering and Technology, 2014.
- [82] P. Petre and T. K. Sarkar, “Differences between modal expansion and integral equation methods for planar near-field to far-field transformation,” *Progress In Electromagnetics Research*, vol. 12, pp. 37–56, 1996.
- [83] M. Dehmollaian, N. Chamanara, and C. Caloz, “Wave scattering by a cylindrical metasurface cavity of arbitrary cross section: Theory and applications,” *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 6, pp. 4059–4072, 2019.
- [84] F. Yang and Y. Rahmat-Samii, Eds., *Surface Electromagnetics: With Applications in Antenna, Microwave, and Optical Engineering*. Cambridge University Press, 2019.
- [85] Y. Álvarez, F. Las-Heras, and M. R. Pino, “Reconstruction of equivalent currents distribution over arbitrary three-dimensional surfaces based on integral equation algorithms,” *Antennas and Propagation, IEEE Transactions on*, vol. 55, no. 12, pp. 3460–3468, 2007.

- [86] J. L. A. Quijano and G. Vecchi, "Improved-accuracy source reconstruction on arbitrary 3-D surfaces," *Antennas and Wireless Propagation Letters, IEEE*, vol. 8, pp. 1046–1049, 2009.
- [87] C. Estatico, M. Pastorino, and A. Randazzo, "A novel microwave imaging approach based on regularization in L^p banach spaces," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 7, pp. 3373–3381, 2012.
- [88] S. R. Rengarajan and Y. Rahmat-Samii, "The field equivalence principle: Illustration of the establishment of the non-intuitive null fields," *Antennas and Propagation Magazine, IEEE*, vol. 42, no. 4, pp. 122–128, 2000.
- [89] R. Fletcher and C. Reeves, "Function minimization by conjugate gradients," *The Computer Journal*, vol. 7, no. 2, pp. 149–154, Jan 1964.
- [90] P. C. Hansen, "Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems," *Numerical algorithms*, vol. 6, no. 1, pp. 1–35, 1994.
- [91] —, "Analysis of discrete ill-posed problems by means of the L-curve," *SIAM Review*, vol. 34, no. 4, pp. 561–580, Dec 1992.
- [92] T. K. Jensen and P. C. Hansen, "Iterative regularization with minimum-residual methods," *BIT Num. Math.*, vol. 47, pp. 103–120, Jan 2007.
- [93] T. K. Jensen, *Stabilization algorithms for large-scale problems*. IMM, Informatik og Matematisk Modellering, Danmarks Tekniske Universitet, 2006.
- [94] J. Chung, J. G. Nagy, and D. P. O'Leary, "A weighted GCV method for Lanczos hybrid regularization," *Electronic Transactions on Numerical Analysis*, vol. 28, no. Electronic Transactions on Numerical Analysis, 2008.
- [95] M. E. Kilmer and D. P. O'Leary, "Choosing regularization parameters in iterative methods for ill-posed problems," *SIAM Journal on matrix analysis and applications*, vol. 22, no. 4, pp. 1204–1221, 2001.
- [96] T. Brown, "Antenna characterization using phaseless near-field measurements," Master's thesis, University of Manitoba, 2016. [Online]. Available: <http://hdl.handle.net/1993/31693>
- [97] P. Mojabi and J. LoVetri, "A prescaled multiplicative regularized Gauss-Newton inversion," *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 8, pp. 2954–2963, Aug 2011.
- [98] Y.-H. Dai, "Conjugate gradient methods with Armijo-type line searches," *Acta Mathematicae Applicatae Sinica*, vol. 18, no. 1, pp. 123–130, 2002.

- [99] S. M. Rao, D. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *Antennas and Propagation, IEEE Transactions on*, vol. 30, no. 3, pp. 409–418, 1982.
- [100] M. Chen, E. Abdo-Sánchez, A. Epstein, and G. V. Eleftheriades, "Theory, design, and experimental verification of a reflectionless bianisotropic Huygens' metasurface for wide-angle refraction," *Physical Review B*, vol. 97, no. 12, p. 125433, 2018.
- [101] G. Lavigne, K. Achouri, V. S. Asadchy, S. A. Tretyakov, and C. Caloz, "Susceptibility derivation and experimental demonstration of refracting metasurfaces without spurious diffraction," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 3, pp. 1321–1330, 2018.
- [102] T. Brown, Y. Vahabzadeh, C. Caloz, and P. Mojabi, "Electromagnetic inversion with local power conservation for metasurface design," *IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 8, pp. 1291–1295, 2020.
- [103] A. H. Dorrah, M. Chen, and G. V. Eleftheriades, "Bianisotropic Huygens' metasurface for wideband impedance matching between two dielectric media," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 9, pp. 4729–4742, Sep. 2018.
- [104] A. H. Dorrah, G. Egorov, and G. V. Eleftheriades, "Non-local power wave transformations using omega bianisotropic Huygens' metasurface pairs," in *2019 URSI International Symposium on Electromagnetic Theory (EMTS)*, May 2019, pp. 1–4.
- [105] G. Oliveri, E. T. Bekele, M. Salucci, and A. Massa, "Array miniaturization through QCTO-SI metamaterial radomes," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 8, pp. 3465–3476, 2015.
- [106] T. Brown, Z. Liu, and P. Mojabi, "Full-wave verification of an electromagnetic inversion metasurface design method," *TechRxiv preprint*, Jan 2020.
- [107] J. P. Wong, A. Epstein, and G. V. Eleftheriades, "Reflectionless wide-angle refracting metasurfaces," *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 1293–1296, 2015.
- [108] T. Brown and P. Mojabi, "Cascaded Metasurface Design Using Electromagnetic Inversion with Gradient-Based Optimization," *TechRxiv preprint*, August 2020.
- [109] V. G. Ataloglou, A. H. Dorrah, and G. V. Eleftheriades, "Realizing antenna arrays with Huygens' metasurface pairs based on a moment-method-like design," in *2020 14th European Conference on Antennas and Propagation (EuCAP)*. IEEE, 2020, pp. 1–5.
- [110] B. O. Raeker and A. Grbic, "Compound metaoptics for amplitude and phase control of wave fronts," *Phys. Rev. Lett.*, vol. 122, p. 113901, Mar 2019. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.122.113901>

- [111] B. O. Raeker, A. Grbic, Y. Zhou, and J. Valentine, "All-dielectric compound metaoptics," in *2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, 2019, pp. 431–432.
- [112] R. Leberer and W. Menzel, "A dual planar reflectarray with synthesized phase and amplitude distribution," *IEEE transactions on antennas and propagation*, vol. 53, no. 11, pp. 3534–3539, 2005.
- [113] A. Abubakar, P. M. van den Berg, and J. J. Mallorqui, "Imaging of biomedical data using a multiplicative regularized contrast source inversion method," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 7, pp. 1761–1771, 2002.
- [114] P. Mojabi and J. LoVetri, "Microwave biomedical imaging using the multiplicative regularized Gauss–Newton inversion," *IEEE Antennas and Wireless Propagation Letters*, vol. 8, pp. 645–648, 2009.
- [115] M. Phaneuf and P. Mojabi, "Electromagnetic inversion for noninvasive specific absorption rate characterization," *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, 2020.
- [116] E. Polak and G. Ribiere, "Note sur la convergence de méthodes de directions conjuguées," *Revue française d'informatique et de recherche opérationnelle, série rouge*, vol. 3, no. 1, pp. 35–43, 1969.
- [117] K. Achouri, "Synthesis and applications of electromagnetic metasurfaces," Ph.D. dissertation, École Polytechnique de Montréal, Montreal, Canada, 2017.
- [118] F. Monticone, N. M. Estakhri, and A. Alù, "Full control of nanoscale optical transmission with a composite metascreen," *Phys. Rev. Lett.*, vol. 110, p. 203903, May 2013. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.110.203903>
- [119] A. E. Olk and D. A. Powell, "Accurate metasurface synthesis incorporating near-field coupling effects," *Physical Review Applied*, vol. 11, no. 6, p. 064007, 2019.
- [120] P. Mojabi, "Investigation and development of algorithms and techniques for microwave tomography," Ph.D. dissertation, University of Manitoba, Winnipeg, Manitoba, Canada (Available Online), 2010.
- [121] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. John Wiley & Sons, 1999.
- [122] D. Brandwood, "A complex gradient operator and its application in adaptive array theory," in *IEE Proceedings F: Communications Radar and Signal Processing*, vol. 130, 1983, pp. 11–16.
- [123] A. Van Den Bos, "Complex gradient and Hessian," *IEE Proceedings-Vision, Image and Signal Processing*, vol. 141, no. 6, pp. 380–383, 1994.

Appendix A

List of Publications

Herein is a list of the author's publications produced during his Ph.D. studies.

Journal Papers:

1. **Trevor Brown** and Puyan Mojabi, "Cascaded Metasurface Design Using Electromagnetic Inversion with Gradient-Based Optimization", *IEEE Transactions on Antennas and Propagation*, submitted July 2020, under review.
 - Trevor Brown: manuscript preparation, algorithm development and evaluation, simulations
 - Puyan Mojabi: advisory role
2. **Trevor Brown**, Yousef Vahabzadeh, Christophe Caloz, and Puyan Mojabi, "Electromagnetic Inversion with Local Power Conservation for Metasurface Design", *IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 8, pp. 1291-1295, 2020.
 - Trevor Brown: manuscript preparation, algorithm development and evaluation, simulations
 - Yousef Vahabzadeh, Christophe Caloz, and Puyan Mojabi: advisory role

3. **Trevor Brown**, Chaitanya Narendra, Yousef Vahabzadeh, Christophe Caloz, and Puyan Mojabi, “On the Use of Electromagnetic Inversion for Metasurface Design”, *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 3, pp. 1812-1824, 2020.
 - Trevor Brown: manuscript preparation, inversion framework development and evaluation, simulations
 - Chaitanya Narendra: manuscript preparation, conjugate gradient algorithm
 - Yousef Vahabzadeh: finite-difference frequency domain simulation code
 - Christophe Caloz and Puyan Mojabi: advisory role

Conference Papers:

1. **Trevor Brown**, Ziqi Liu, and Puyan Mojabi, “Full-Wave Verification of an Electromagnetic Inversion Metasurface Design Method”, *2020 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, 2020.
 - Trevor Brown: manuscript preparation, algorithm development and evaluation, simulations
 - Ziqi Liu: ABCD-parameter conversion
 - Puyan Mojabi: advisory role
2. **Trevor Brown**, Chaitanya Narendra, Yousef Vahabzadeh, Christophe Caloz, and Puyan Mojabi, “Metasurface Design Using Electromagnetic Inversion”, *2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, pp. 1817-1818, 2019.
 - Trevor Brown: manuscript preparation, algorithm development and evaluation, simulations
 - Chaitanya Narendra: conjugate gradient algorithm

- Yousef Vahabzadeh: finite-difference frequency domain simulation code
 - Christophe Caloz and Puyan Mojabi: advisory role
3. **Trevor Brown**, Chaitanya Narendra, Chen Niu, and Puyan Mojabi, “On the Use of Electromagnetic Inversion for Near-Field Antenna Measurements: A Review”, *2018 Conference on Antenna Measurements and Applications (CAMA)*, pp. 1-4, 2018.
- Trevor Brown: manuscript preparation, phaseless measurements and data-prescaling algorithms
 - Chaitanya Narendra: irregular measurement domain algorithms
 - Chen Niu: development and construction of a pendulum-based planar near-field antenna measurement system
 - Puyan Mojabi: advisory role
4. Chaitanya Narendra, **Trevor Brown**, Nozhan Bayat, and Puyan Mojabi, “Multi-Plane Magnetic Near-Field Data Inversion Using the Source Reconstruction Method”, *2018 18th International Symposium on Antenna Technology and Applied Electromagnetics (ANTEM)*, pp. 1-2, 2018.
- Chaitanya Narendra: manuscript preparation, algorithm development and evaluation
 - Trevor Brown: inversion framework
 - Nozhan Bayat: HFSS simulations
 - Puyan Mojabi: advisory role
5. **Trevor Brown**, Chaitanya Narendra, and Puyan Mojabi, “On the Use of the Source Reconstruction Method for Metasurface Design”, *2018 18th European Conference on Antennas and Propagation (EuCAP)*, pp. 302-306, 2018.

- Trevor Brown: manuscript preparation, algorithm development and evaluation, simulations
- Chaitanya Narendra: conjugate gradient algorithm
- Puyan Mojabi: advisory role

Appendix B

Cascaded Metasurface Design Gradient Derivations

This appendix is associated with the material presented in Chapter 5.¹ The gradients required for each of the two optimization steps are derived below.

B.1 Derivation of Required Gradients with Respect to \mathbf{x}_2

The CG minimization to update \mathbf{x}_2 requires finding the gradient of two cost functionals, namely, $\mathcal{C}_F(\mathbf{x}_2)$ and $\mathcal{C}_{TV}(\mathbf{x}_2)$, with respect to \mathbf{x}_2 . Keep in mind that although \mathbf{x}_2 vector represents the complex-valued equivalent electric and magnetic currents on Σ_2^+ , it has been constructed such that it is a purely-real vector; see Section 5.4.1. Therefore, all the gradient vectors will also be purely real vectors.

¹Reprinted, with permission, from T. Brown and P. Mojabi, “Cascaded Metasurface Design Using Electromagnetic Inversion with Gradient-Based Optimization”, *IEEE Trans. Ant. Prop.*, submitted July 2020.

B.1.1 Gradient of $\mathcal{C}_F(\mathbf{x}_2)$

We begin deriving the gradient of $\mathcal{C}_F(\mathbf{x}_2)$ using standard vector differentiation rules as

$$\begin{aligned} \mathbf{g}_F(\mathbf{x}_2) &= \frac{\partial}{\partial \mathbf{x}_2} \frac{\left\| |\mathbf{A}_F \mathbf{x}_2|^2 - |\mathbf{f}|^2 \right\|_S^2}{\left\| |\mathbf{f}|^2 \right\|_S^2} \\ &= \frac{2}{\left\| |\mathbf{f}|^2 \right\|_S^2} \frac{\partial}{\partial \mathbf{x}_2} \left[(\mathbf{A}_F \mathbf{x}_2) \odot (\mathbf{A}_F \mathbf{x}_2)^* - |\mathbf{f}|^2 \right]^T \mathbf{r}_F \end{aligned} \quad (\text{B.1})$$

where \mathbf{r}_F is given in (5.16) and \odot represents the Hadamard (elementwise) product. To evaluate the derivative in (B.1), we use the identity

$$\frac{\partial}{\partial \mathbf{x}} (\mathbf{A} \mathbf{x} \odot \mathbf{B} \mathbf{x}) = \text{diag}(\mathbf{A} \mathbf{x}) \mathbf{B} + \text{diag}(\mathbf{B} \mathbf{x}) \mathbf{A}. \quad (\text{B.2})$$

Noting that \mathbf{x}_2 is a purely real vector ($\mathbf{x}_2 \in \mathbb{R}^{4N}$), the use of (B.2) in conjunction with (B.1) results in

$$\begin{aligned} \mathbf{g}_F(\mathbf{x}_2) &= \frac{2}{\left\| |\mathbf{f}|^2 \right\|_S^2} \left[\text{diag}(\mathbf{A}_F \mathbf{x}_2) \mathbf{A}_F^* + \text{diag}(\mathbf{A}_F^* \mathbf{x}_2) \mathbf{A}_F \right]^T \mathbf{r}_F \\ &= \frac{2}{\left\| |\mathbf{f}|^2 \right\|_S^2} \left[2 \text{Re} \left\{ \text{diag}(\mathbf{A}_F \mathbf{x}_2) \mathbf{A}_F^* \right\} \right]^T \mathbf{r}_F. \end{aligned} \quad (\text{B.3})$$

Noting that the residual \mathbf{r}_F is purely real and applying the transpose operation we can simplify this expression to

$$\begin{aligned} \mathbf{g}_F(\mathbf{x}_2) &= \frac{4 \text{Re} \left\{ \mathbf{A}_F^H \text{diag}(\mathbf{A}_F \mathbf{x}_2) \mathbf{r}_F \right\}}{\left\| |\mathbf{f}|^2 \right\|_S^2} \\ &= \frac{4 \text{Re} \left\{ \mathbf{A}_F^H (\mathbf{r}_F \odot \mathbf{A}_F \mathbf{x}_2) \right\}}{\left\| |\mathbf{f}|^2 \right\|_S^2}. \end{aligned} \quad (\text{B.4})$$

B.1.2 Gradient of $\mathcal{C}_{\text{TV}}(\mathbf{x}_2)$

Next, consider the gradient of the total variation regularizer given in (5.10). As an analogous derivation, we instead consider the functional

$$\tilde{\mathcal{C}}_{\text{TV}}(u(y)) = \|\nabla_y u(y)\|_{\Sigma_2}^2 \quad (\text{B.5})$$

in which $u(y)$ is a *continuously* defined real-valued function of y on Σ_2 and the norm is defined as

$$\|\vec{f}(y)\|_{\Sigma_2}^2 = \langle \vec{f}(y), \vec{f}(y) \rangle_{\Sigma_2} = \int_{\Sigma_2} \vec{f}(y) \cdot \vec{f}(y) dy. \quad (\text{B.6})$$

Note that since we are dealing with real-valued functions (i.e., the continuous form of the real-valued \mathbf{x}_2 vector), the presence of the complex conjugate operator has been dropped in the above norm definition.² For the sake of notational simplicity, the y dependency of u will be implied from now on. We start by finding the first variation of $\tilde{\mathcal{C}}_{\text{TV}}(u)$, i.e., the derivative of $\tilde{\mathcal{C}}_{\text{TV}}(u)$ with respect to u when u is slightly varied by some function ψ (which is defined on the same domain as u)

$$\begin{aligned} \partial \tilde{\mathcal{C}}_{\text{TV}} &= \lim_{\varepsilon \rightarrow 0} \frac{\tilde{\mathcal{C}}_{\text{TV}}(u + \varepsilon \psi) - \tilde{\mathcal{C}}_{\text{TV}}(u)}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{\|\nabla_y (u + \varepsilon \psi)\|_{\Sigma_2}^2 - \|\nabla_y u\|_{\Sigma_2}^2}{\varepsilon}. \end{aligned} \quad (\text{B.7})$$

Using the definition of the norm in (B.6), we can expand (B.7) as

$$\begin{aligned} \partial \tilde{\mathcal{C}}_{\text{TV}} &= \lim_{\varepsilon \rightarrow 0} \frac{\langle \nabla_y u + \varepsilon \nabla_y \psi, \nabla_y u + \varepsilon \nabla_y \psi \rangle - \|\nabla_y u\|_{\Sigma_2}^2}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon^2 \|\nabla_y \psi\|_{\Sigma_2}^2 + 2 \langle \nabla_y u, \varepsilon \nabla_y \psi \rangle}{\varepsilon} \end{aligned} \quad (\text{B.8})$$

²For a similar derivation, but in the complex domain, see [120, Appendix D.3].

and evaluating the limit results in

$$\begin{aligned}\partial\tilde{\mathcal{C}}_{\text{TV}} &= 2 \langle \nabla_y u, \nabla_y \psi \rangle \\ &= 2 \int_{\Sigma_2} (\nabla_y u) \cdot (\nabla_y \psi) dy.\end{aligned}\tag{B.9}$$

If we note that

$$\nabla_y \cdot \left((\nabla_y u) \psi \right) = \psi \nabla_y \cdot \nabla_y u + \nabla_y u \cdot \nabla_y \psi\tag{B.10}$$

then (B.9) becomes

$$\partial\tilde{\mathcal{C}}_{\text{TV}} = 2 \int_{\Sigma_2} \left[\nabla_y \cdot \left((\nabla_y u) \psi \right) - \psi \nabla_y \cdot \nabla_y u \right] dy.\tag{B.11}$$

Applying the divergence theorem [121] results in

$$\partial\tilde{\mathcal{C}}_{\text{TV}} = 2 \int_{\partial\Sigma_2} (\nabla_y u) \psi \cdot \hat{x} dy - 2 \int_{\Sigma_2} \psi \nabla_y \cdot (\nabla_y u) dy\tag{B.12}$$

where $\partial\Sigma_2$ refers to the edges of the metasurface. We assume that the function u vanishes on the boundary, i.e., $u(y \in \partial\Sigma_2) = 0$, which is equivalent to assuming that the equivalent currents are assumed to be zero on the edges of the metasurface in our case. Since ψ exists in the same function space as u , then $\psi(y \in \partial\Sigma_2) = 0$. Subsequently, (B.12) simplifies to

$$\begin{aligned}\partial\tilde{\mathcal{C}}_{\text{TV}} &= -2 \int_{\Sigma_2} \psi \nabla_y \cdot (\nabla_y u) dy \\ &= \left\langle -2 \nabla_y \cdot (\nabla_y u), \psi \right\rangle_{\Sigma_2} \\ &= \left\langle -2 \nabla_y^2 u, \psi \right\rangle_{\Sigma_2}.\end{aligned}\tag{B.13}$$

If we were to discretize (B.5) such that the functional operates on the equivalent currents in \mathbf{x}_2 as in (5.10), and considering the regularization weight w_{TV} , the corresponding *discrete* gradient operator [122, 123] can then be written as

$$\mathbf{g}_{\text{TV}}(\mathbf{x}_2) = -2w_{\text{TV}} \begin{bmatrix} \frac{\partial^2}{\partial y^2} \mathbf{J}_{2,\text{R}} \\ \frac{\partial^2}{\partial y^2} \mathbf{J}_{2,\text{I}} \\ \frac{\partial^2}{\partial y^2} \mathbf{M}_{2,\text{R}} \\ \frac{\partial^2}{\partial y^2} \mathbf{M}_{2,\text{I}} \end{bmatrix}. \quad (\text{B.14})$$

B.2 Derivation of Required Gradients with Respect to \mathbf{x}_1

This requires calculation of the gradient of the four cost functionals, namely \mathcal{C}_{L} , $\mathcal{C}_{\text{P},1}$, $\mathcal{C}_{\text{P},2}$, and \mathcal{C}_{TV} , with respect to \mathbf{x}_1 . Keep in mind that although \mathbf{x}_1 represents the complex-valued equivalent electric and magnetic currents on Σ_1^+ , it has been constructed to be a purely real vector; see Section 5.4.1.

B.2.1 Gradient of $\mathcal{C}_{\text{L}}(\mathbf{x}_1)$

The gradient of $\mathcal{C}_{\text{L}}(\mathbf{x}_1)$ can be derived using vector differentiation rules as

$$\begin{aligned} \mathbf{g}_{\text{L}} &= \frac{\partial}{\partial \mathbf{x}_1} \|\mathbf{L}\mathbf{x}_1\|_{\Sigma_1}^2 \\ &= \frac{\partial}{\partial \mathbf{x}_1} \left[(\mathbf{L}\mathbf{x}_1)^{\text{H}} (\mathbf{L}\mathbf{x}_1) \right] \\ &= \frac{\partial}{\partial \mathbf{x}_1} \left[\mathbf{x}_1^{\text{T}} \mathbf{L}^{\text{H}} \mathbf{L} \mathbf{x}_1 \right] \\ &= \left[\mathbf{L}^{\text{H}} \mathbf{L} + (\mathbf{L}^{\text{H}} \mathbf{L})^{\text{T}} \right] \mathbf{x}_1 \\ &= \left(\mathbf{L}^{\text{H}} \mathbf{L} + \mathbf{L}^{\text{T}} \mathbf{L}^* \right) \mathbf{x}_1 \\ &= 2 \text{Re} \left(\mathbf{L}^{\text{H}} \mathbf{L} \mathbf{x}_1 \right) \end{aligned} \quad (\text{B.15})$$

B.2.2 Gradient of $\mathcal{C}_{P,1}(\mathbf{x}_1)$

To derive the required gradient of $\mathcal{C}_{P,1}(\mathbf{x}_1)$, we first find the gradient of $\mathcal{C}_{P,1}(\mathbf{x}_1)$ with respect to $\mathbf{J}_{1,R}$ only. Using vector differentiation rules this can be computed as

$$\begin{aligned}
\frac{\partial \mathcal{C}_{P,1}}{\partial \mathbf{J}_{1,R}} &= \frac{\partial}{\partial \mathbf{J}_{1,R}} \frac{w_1 \left\| \frac{1}{2} (\mathbf{J}_{1,R} \odot \mathbf{M}_{1,R} + \mathbf{J}_{1,I} \odot \mathbf{M}_{1,I}) - \mathbf{p}_{\text{in}} \right\|_{\Sigma_1}^2}{\|\mathbf{p}_{\text{in}}\|_{\Sigma_1}^2} \\
&= \frac{2w_1 \frac{\partial}{\partial \mathbf{J}_{1,R}} \left[\frac{1}{2} (\mathbf{J}_{1,R} \odot \mathbf{M}_{1,R} + \mathbf{J}_{1,I} \odot \mathbf{M}_{1,I}) - \mathbf{p}_{\text{in}} \right]^T \mathbf{r}_1}{\|\mathbf{p}_{\text{in}}\|_{\Sigma_1}^2} \\
&= \frac{w_1}{\|\mathbf{p}_{\text{in}}\|_{\Sigma_1}^2} \text{diag}(\mathbf{M}_{1,R}) \mathbf{r}_1 \\
&= \frac{w_1}{\|\mathbf{p}_{\text{in}}\|_{\Sigma_1}^2} \mathbf{r}_1 \odot \mathbf{M}_{1,R}
\end{aligned} \tag{B.16}$$

where \mathbf{r}_1 is given in (5.30). Performing the a similar derivation for the gradients with respect to $\mathbf{J}_{1,I}$, $\mathbf{M}_{1,R}$, and $\mathbf{M}_{1,I}$ and concatenating the results according to the order of (5.8) produces the final gradient with respect to \mathbf{x}_1 , i.e., $\mathbf{g}_{P,1}$, which is given in (5.29).

B.2.3 Gradient of $\mathcal{C}_{P,2}(\mathbf{x}_1)$

We now derive the gradient of $\mathcal{C}_{P,2}(\mathbf{x}_1)$ as

$$\begin{aligned}
\mathbf{g}_{P,2} &= \frac{\partial}{\partial \mathbf{x}_1} \frac{w_2 \left\| \frac{1}{2} \text{Re}(\mathbf{B}_1 \mathbf{x}_1 \odot \mathbf{B}_2^* \mathbf{x}_1) - \mathbf{p}_{\text{out}} \right\|_{\Sigma_2}^2}{\|\mathbf{p}_{\text{out}}\|_{\Sigma_2}^2} \\
&= \frac{2w_2 \frac{\partial}{\partial \mathbf{x}_1} \left[\frac{1}{2} \text{Re}(\mathbf{B}_1 \mathbf{x}_1 \odot \mathbf{B}_2^* \mathbf{x}_1) - \mathbf{p}_{\text{out}} \right]^T \mathbf{r}_2}{\|\mathbf{p}_{\text{out}}\|_{\Sigma_2}^2}.
\end{aligned} \tag{B.17}$$

Using the identity in (B.2) this simplifies to

$$\mathbf{g}_{P,2} = \frac{w_2 \left[\text{Re} \left(\text{diag}(\mathbf{B}_1 \mathbf{x}_1) \mathbf{B}_2^* + \text{diag}(\mathbf{B}_2^* \mathbf{x}_1) \mathbf{B}_1 \right) \right]^T \mathbf{r}_2}{\|\mathbf{p}_{\text{out}}\|_{\Sigma_2}^2}. \tag{B.18}$$

B.2.4 Gradient of $\mathcal{C}_{\text{TV}}(\mathbf{x}_1)$

This is the same as Section B.1.2. Thus,

$$\mathbf{g}_{\text{TV}}(\mathbf{x}_1) = -2w_{\text{TV}} \begin{bmatrix} \frac{\partial^2}{\partial y^2} \mathbf{J}_{1,\text{R}} \\ \frac{\partial^2}{\partial y^2} \mathbf{J}_{1,\text{I}} \\ \frac{\partial^2}{\partial y^2} \mathbf{M}_{1,\text{R}} \\ \frac{\partial^2}{\partial y^2} \mathbf{M}_{1,\text{I}} \end{bmatrix}. \quad (\text{B.19})$$

Appendix C

Total Variation Regularization

In this appendix, we present a comparative example demonstrating the effect of the total variation (TV) regularization detailed in Section 5.4.3. This example uses the same desired field specifications (i.e., the far-field power pattern shown in Figure 4.2) and metasurface geometry as the example shown in Section 4.5. We design two *single* metasurfaces using the proposed method to perform the same field transformation from a normally incident TE_z plane wave. The first design is produced *without* using the total variation regularizer, and the second *with* the total variation regularizer. In both cases we use the conjugate gradient method as the optimization algorithm with the same number of iterations and weighting parameters (as opposed to the example in Section 4.5 in which particle swarm optimization was used). The amplitude and phase of the tangential output fields (or equivalently, the currents) resulting from each case are shown in Figure C.1. The electric field in the simulation domain for both cases is shown in Figure C.2, while the far-field patterns produced are compared in Figure C.3. The smoothing effect TV regularization has on the currents (and therefore the output fields) is readily apparent from Figure C.1. In the case of multiple metasurfaces, the reduction of the rapid spatial variations is more impactful and reduces the difficulty of satisfying local power conservation at the output metasurface. The *normalized*

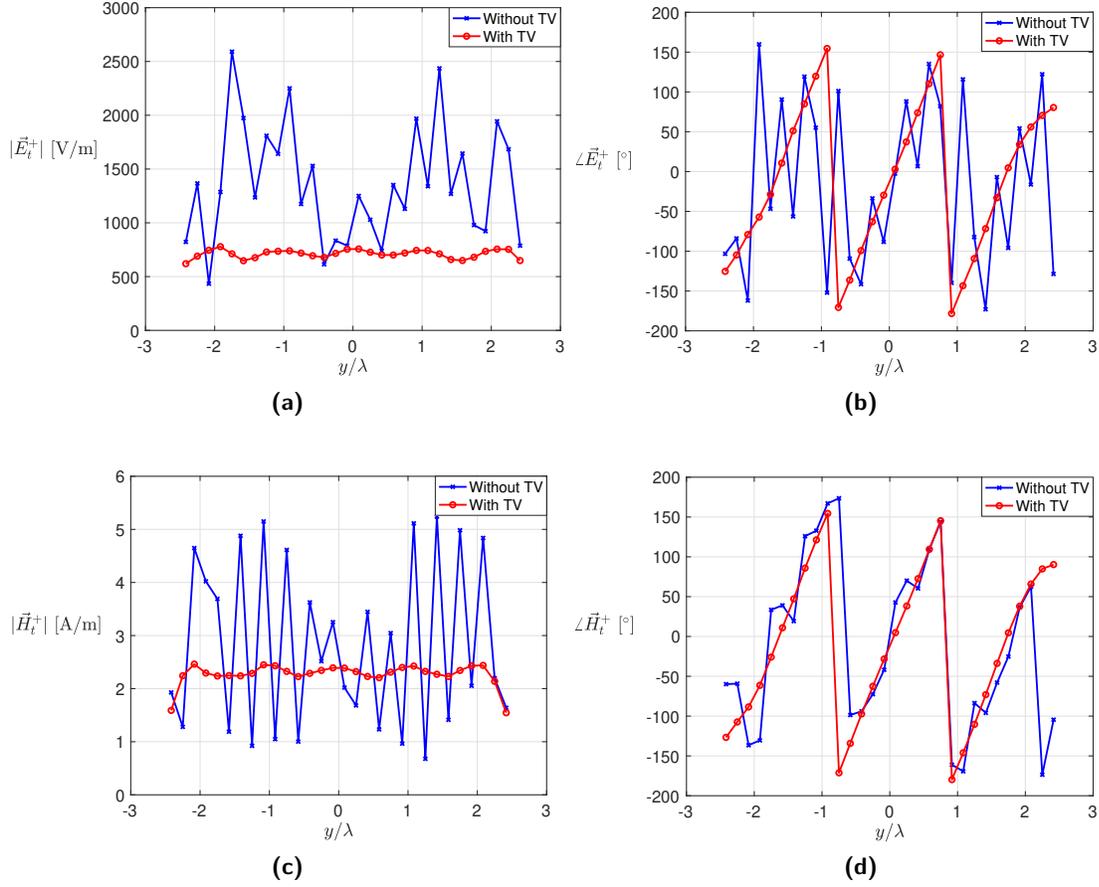


Fig. C.1: Comparison of the tangential fields on the output side of the metasurface (i.e., \vec{E}_t^+ and \vec{H}_t^+) resulting from the design procedure with (red line with circular markers) and without (blue 'x' markers) total variation regularization. The amplitude and phase of the fields are plotted for each of the 30 unit cells along the metasurface.

FF radiation patterns produced by both designs agree well with the desired FF pattern, with the design without TV regularization fitting the main beam and the first few sidelobes slightly better (as expected due to over-fitting). However, the transmission efficiency of the metasurface without TV regularization is significantly worse at 80.25% compared to an efficiency of 97.44% for the design using TV regularization.

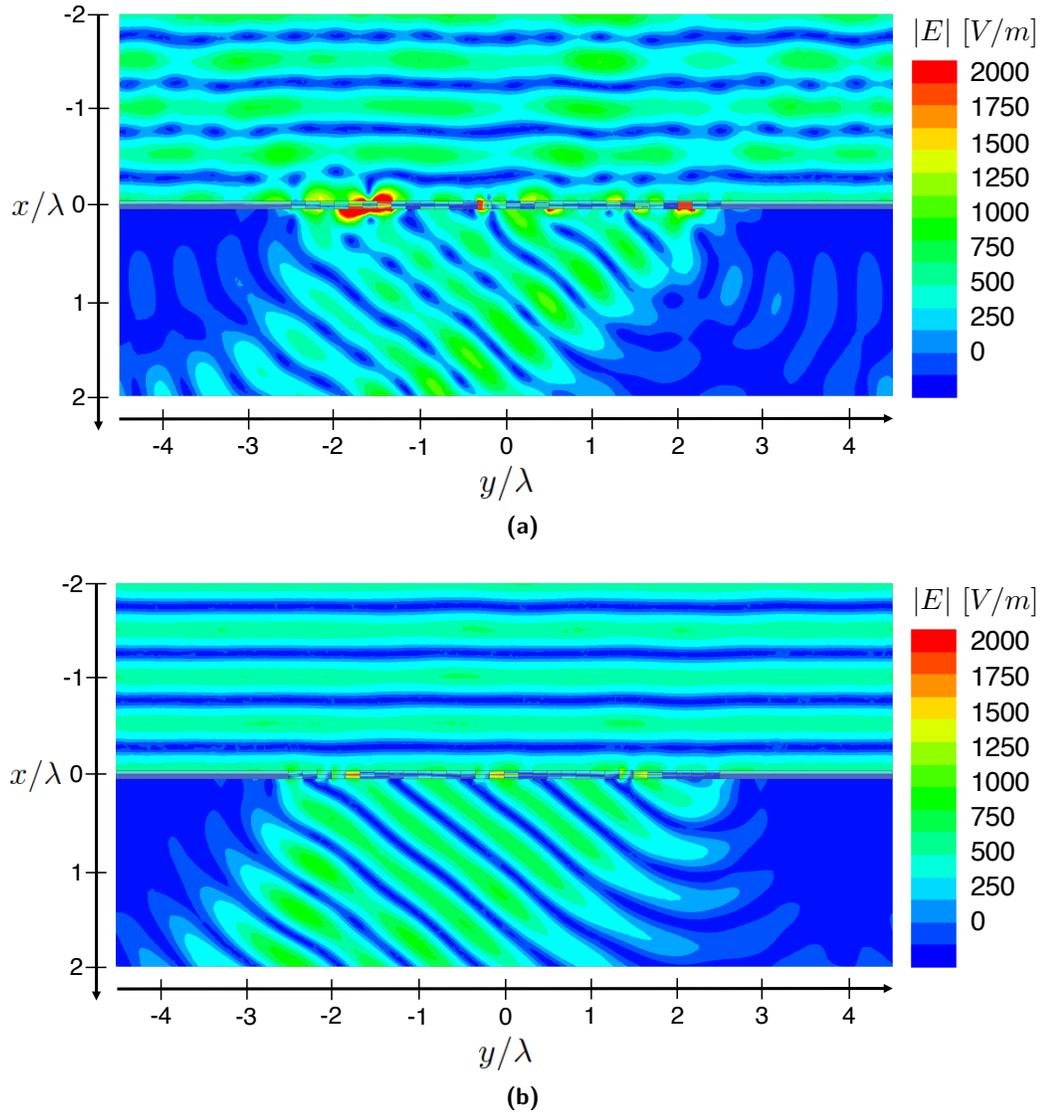


Fig. C.2: Absolute value of the real part of the total electric field when the metasurfaces designed (a) without and (b) with total variation regularization are illuminated by a normally incident plane wave.

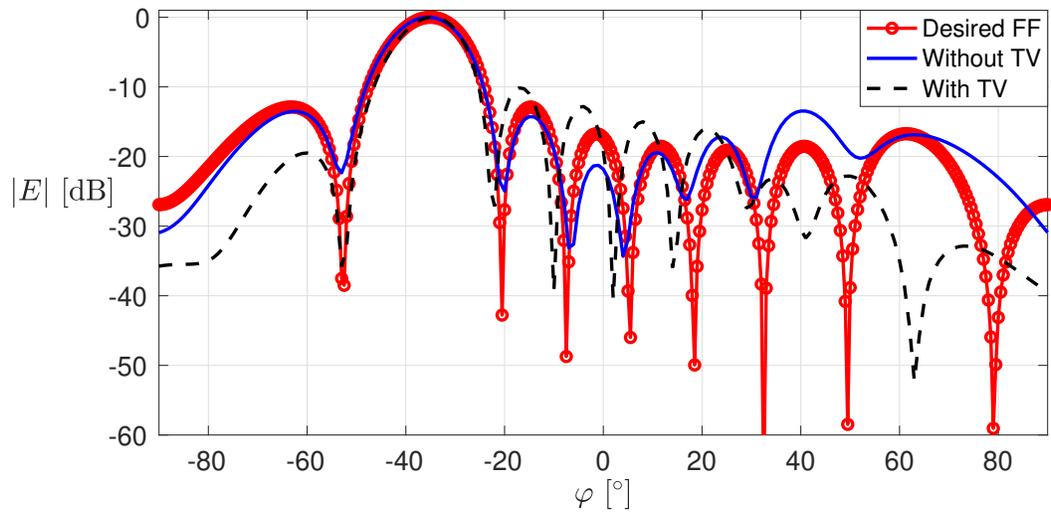


Fig. C.3: Comparison of the specified desired far-field power pattern (solid red curve with circular markers) and the normalized far-field power patterns produced by ANSYS HFSS simulations of the designed metasurfaces with (black dashed curve) and without (solid blue curve) total variation regularization.