Deep Learning Techniques for Transmissive Metasurface Design

by

Chen Niu

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

in

Electrical and Computer Engineering

University of Manitoba

Winnipeg, Manitoba, Canada

Copyright © 2024 by Chen Niu

I dedicate this thesis to my loving grandpa.

献给我亲爱的姥爷洪春田。

Acknowledgements

I would like to express my deepest gratitude to my academic advisor, Dr. Puyan Mojabi, who has not only guided me through my Ph.D. but also served as my advisor during my master's studies. His unwavering support, mentorship, and friendship have been instrumental in my academic journey. His insight and encouragement have significantly shaped this research and my development as a scholar. I am also thankful to my colleagues, especially Mario Phaneuf, for their collaboration, and the many fruitful discussions.

A special thanks goes to my parents for their unconditional love and belief in me. I am equally grateful to my aunt and uncle in Oakville for their kindness and hospitality, and to my girlfriend for her patience, understanding, and constant encouragement.

I would also like to extend my sincere appreciation to my Ph.D. committee members, Dr. Ian Jeffrey (University of Manitoba), Dr. Christopher Henry (University of Manitoba), and my External Examiner Dr. Maokun Li (Tsinghua University) for taking the time to evaluate and improve this work.

Finally, I acknowledge the financial support from MITACS and the Natural Sciences and Engineering Research Council of Canada (NSERC). I also wish to thank the Canadian Microelectronics Corporation (CMC) for providing ANSYS Campus Solutions and express my gratitude to Mr. James Dietrich for his assistance with performing metasurface measurements.

Abstract

Electromagnetic (EM) metasurfaces are considered as two-dimensional artificial structures since they are thin with respect to the wavelength of operation. These thin devices are designed to tailor the incoming EM waves into the desired outgoing ones. This transformation is achieved through lattices of engineered subwavelength scattering particles, known as meta-atoms or unit cells. The properties of these unit cells provide the mechanism to control the amplitude, phase, and polarization of the incoming EM waves. It is through this control that the EM radiation pattern of the metasurface can be shaped in terms of different far-field (FF) criteria such as main beam directions and sidelobe levels. However, the current metasurface design process has been complex and iterative, often requiring extensive computational resources.

In light of this challenge, this thesis investigates and presents an end-to-end systematic approach for transmissive metasurface design, aims at shorten the design cycle. To achieve this, it develops deep learning (DL) frameworks and integrates them into the two key stages of metasurface design: macroscopic and microscopic. At the macroscopic design stage, the primary goal is to infer the necessary metasurface properties from the desired FF specifications while ensuring that the metasurface remains lossless, passive, and reciprocal. This thesis explores how these requirements can be met through DL techniques, either by addressing them in separate steps or simultaneously in a single step, with the benefits and drawbacks of each approach examined.

Subsequently, the microscopic design stage aims to yield physical designs that realize the properties obtained in the macroscopic stage. To achieve this, this thesis develops a DL framework to synthesize metallic traces on three-layered unit cells, which

can be fabricated using the printed circuit board (PCB) technology. Finally, this thesis integrates the macroscopic and microscopic DL frameworks to create an end-to-end design pipeline. The performance of this pipeline is verified through full-wave EM simulations and experimental evaluation.

Table of contents

List of figures **xiii**

List of tables **xxiii**

Nomenclature **xxv**

1 Introduction **1**

1.1 Motivation 13

1.2 Objectives and Contributions 19

1.3 Structure of the Thesis 22

2 Background **23**

2.1 Metasurface Fundamentals 23

2.1.1 Generalized Sheet Transition Conditions (GSTCs) 27

2.1.2 Constraints on Surface Susceptibilities 32

2.1.3 Two-port Network Model 35

2.1.4 Local Power Conservation 39

2.1.5 Microscopic Design 43

2.2 Deep Learning 44

2.2.1 Neural Networks and Backpropagation 45

2.2.2 Convolutional Neural Networks 46

2.2.3 Supervised Learning 47

2.2.4 Unsupervised Learning 48

2.2.5 Generative Modelling 48

3 Macroscopic Design from Far-Field Specifications 51

3.1 Introduction 52

3.2 Problem Statement and the Overall Methodology 55

3.3 Metasurface Background 56

3.4 Deep Learning Background and Overall Architecture 58

3.4.1 Encoder-only Transformers 59

3.4.2 Discrete Variational Autoencoders (dVAE) 61

3.5 Source Array Synthesis 64

3.5.1 Formatting Far-Field Specifications to Sequential Data 65

3.5.2 Adapting the dVAE to Source Array Synthesis 65

3.5.3 Final Network Architecture 66

3.5.4 Dataset Generation and Training 67

3.6 Auxiliary Surface Waves 69

3.7 Results 71

3.7.1 2D Design 72

3.7.2 3D Design 75

3.8 Conclusion 77

4 Macroscopic Design from 2D Far-Field Masks 79

4.1 Introduction 80

4.2 Methodology 82

4.3 Results 86

4.4 Conclusion 90

5 Macroscopic Design from 3D Far-Field Masks 91

5.1 Introduction 92

5.2 Methodology 93

5.3 Results 96

5.4 Conclusion 97

6 Microscopic Design from Far-Field Specifications 101

6.1 Introduction 102

6.2 Scope and Problem Statement 105

6.3 Background and Methodology 106

6.3.1 Diffusion Models (DM) 107

6.3.2 U-Net 108

6.3.3 Latent Diffusion Model (LDM) 109

6.4 Unit Cell Design 110

6.4.1 VAE Dataset Generation and Training 111

6.4.2 DM Dataset Generation and Training 112

6.5 Results 115

6.5.1 Plane wave refraction 116

6.5.2 Cylindrical wave to plane wave transformation 123

6.5.3 From far-field performance criteria to metasurface design 125

6.6 Conclusion 127

7 Microscopic Design from Far-Field Masks 129

7.1 Introduction 130

7.2 Methodology 131

7.3 Results 132

7.4 Measurements 133

7.4.1 Planar near-field range (PNFR) 135

7.4.2 Evaluation of the Reflection 136

7.5 Conclusions 140

8 Conclusions and Future Work 143

8.1 Main Conclusions 144

8.2 Summary and Limitations 145

8.2.1 Two-step DL based macroscopic metasurface design 145
8.2.2 Single-step DL based macroscopic metasurface design 146
8.2.3 DL based microscopic metasurface design 147
8.3 Future Works 148
8.3.1 Complete metasurface designs from 3D masks 149
8.3.2 Improve the flexibility of the microscopic design 149
8.3.3 Relax the locally periodicity assumption 149
8.4 Closing Remarks 150

Appendix A Cascaded Design in a Two-Step Scheme **151**

Appendix B Additional Data for Chapter 6 **155**

Appendix C Details on Integrated Loss Function **159**

Appendix D Additional DL Framework Information **163**

References **169**

List of figures

1.1 A sample metasurface placed within a wooden frame. 3

1.2 A close-up view of a sample metasurface showing its unit cell. 4

1.3 A sample fully periodic metasurface, originally designed by Z. Liu et al. 9

1.4 The metasurface design pipeline. The highlighted parts in red are the focus of this thesis. This figure is a modified version of Fig. 5.1. 10

1.5 A sample cascaded metasurface system designed by J. Gohel. 13

1.6 A classification of the macroscopic design approaches for metasurface design 16

1.7 A classification of the microscopic design approaches for unit cell design 18

2.1 The Omega meta-atom. The electric polarization density vector $\boldsymbol{P}$ at location u also depends on the magnetic field at $u + \Delta u$. 25

2.2 The metasurface is assumed to lie on the uv plane having a zero thickness. The normal unit vector to the metasurface is $\hat{n}$. The tangential fields on the top and bottom faces of the metasurface are denoted by ψ^+ and ψ^- respectively where ψ is either the electric or magnetic field. A conventional pillbox is used to derived the boundary conditions. 28

2.3 A sample metasurface with a size of $L_u \times L_v$ and a subwavelength thickness ($\delta \ll \lambda$ where λ denotes the wavelength). This metasurface consists of three-layered dogbone copper traces printed on a dielectric substrate. For clarity, this figure only shows one column of such dogbone copper traces. Assuming a zero-thickness model (surface model), the metasurface is located on the uv plane at $n = 0$. The metasurface will then have two sides: $n = 0^+$ (output) and $n = 0^-$ (input). The metasurface is designed to transform a specific incident wave $\boldsymbol{\psi}^{\mathrm{i}}(\boldsymbol{r})$ where $\boldsymbol{\psi}$ represents the electric $\boldsymbol{E}$ and magnetic $\boldsymbol{H}$ fields into desired reflected wave $\boldsymbol{\psi}^{\mathrm{r}}(\boldsymbol{r})$ and transmitted wave $\boldsymbol{\psi}^{\mathrm{t}}(\boldsymbol{r})$. The tangential fields on the input side are denoted as $\vec{E}_t^-$ and $\vec{H}_t^-$. The tangential fields on the output side are denoted as $\vec{E}_t^+$ and $\vec{H}_t^+$. Since the metasurface lies on the uv plane, the tangential components to the metasurface will have u and v components. 33

2.4 Transformation of a plane wave into two refracted plane waves using a single lossless and passive metasurface utilizing surface waves. The metasurface lies on the uv plane under the 2D assumption. 41

2.5 A cascaded metasurface system that utilizes the forward ψ^{fwd} and reflected ψ^{ref} waves in the region between the two metasurfaces for power redistribution. 42

2.6 (a) Three-layered dogbone unit cell. (b) Dimensions of the dogbone for a single layer. The parameters L, W, and T are the length, width and trace width of the dogbone. Variation of these dimensions can generate different values for surface susceptibilities. 44

3.1 The end-to-end metasurface macroscopic design pipeline utilized in this work. The lengths and directions of the arrows in the source array image depict the amplitude and phase of different infinitesimal dipole antennas respectively. 56

3.2 The neural network illustration of the encoder-only transformer. The output of the transformer encoder is normally referred to as the hidden state **H**, but in the case of the encoder-only transformer, **H** is the final output target sequence **T**. 59
3.3 The neural network illustration of variational autoencoder (VAE). . . . 62
3.4 The conceptual illustration of discrete variational autoencoders (dVAE) which shows encoding the University of Manitoba's logo to a discrete sequence and then decoding this sequence back in an attempt to arrive at the original logo. 63
3.5 The vectorized far-field performance criteria where the first row lists the theta (θ) angles of the far-field cut. The main beam directions, HPBWs of each beam, and null locations are represented in binary format, where 1 represents the presence and 0 represents the absence of the quantity, and are located at the second, third, and last rows respectively. The side lobe level (SLL) and maximum null levels (MNL) are arrays of constant value and are located at the fourth and fifth rows respectively. 64
3.6 The proposed neural network architecture of far-field specifications to source array synthesis for both (a) 2D and (b) 3D scenarios. It should be noted that in (a) the dSeqVAE encoder and the dSeqVAE decoder on the two sides of the transformer are of two separate models as indicated by the different colors. On the other hand, the dSeqVAE encoders in (a) and (b) are of the same model, and are therefore shown with the same color. 66
3.7 The illustration of the equivalent problem with respect to the input region for the scattered fields, which consist of the auxiliary surface waves used to meet the LPC constraint. Love's condition is assumed for the output region of this equivalent problem and is enforced by its equivalent additive constraint during the optimization process. 70
3.8 The two-port circuit model for a three-layer metasurface unit cell. . . . 72

3.9 Three scenarios for the 2D design cases. 73
3.10 Ansys HFSS simulation of the metasurface under the 2D transverse electric illumination, which plots E_z magnitude at a fixed phase. . . . 74
3.11 Ansys HFSS metasurface simulation set up for the 3D case with an illuminating horn antenna. 76
3.12 The 3D example where the desired specifications are given on two perpendicular far-field cuts. 77
3.13 The near-field simulation of the 3D metasurface in the E-plane cut (yz plane), which plots E magnitude at a fixed phase. 78

4.1 ©2024 IEEE. Forward path of unsupervised learning for the developed mask-based pattern synthesis with LPC (*MBPS-LPC*). 83
4.2 ©2024 IEEE. Desired power pattern (PP) in the shape of a trapezoid and identified by lower and upper masks (first example) along with the achieved far-field power patterns. 85
4.3 ©2024 IEEE. Plot of the electric field magnitude at a fixed phase in Ansys HFSS for the first example using the MBPS-LPC. The coordinate center is located at the center of the metasurface. 87
4.4 ©2024 IEEE. Desired square-shape power pattern (PP), along with the achieved far-field power patterns for the second example. 88
4.5 ©2024 IEEE. Plot of the electric field magnitude at a fixed phase in Ansys HFSS for the second example. 89
4.6 ©2024 IEEE. The reactance values [Ω] of the three-layered metasurface design for the trapezoid and square-shaped masks. 89

5.1 A transmissive metasurface design pipeline. The highlighted parts in orange are the focus of this chapter. 92
5.2 DL-based metasurface macroscopic design pipeline from 3D masks (based on the 2D work described in Chapter 4). 95

5.3 Projection of the hemisphere of interest with 27 circular cones onto a plane using the (a) uv coordinates and (b) Peirce quincuncial projection coordinates, which we refer to as $x|_{\text{Peirce}}$ and $y|_{\text{Peirce}}$. In both cases, the somewhat vertical contours indicate fixed φ angles, and the somewhat horizontal contours indicate the fixed θ angles. 96
5.4 Desired elliptical masks (solid contour) in Peirce quincuncial projected domain ($x|_{\text{Peirce}}$ and $y|_{\text{Peirce}}$ plane) and the achieved normalized power pattern in color for the (a) first and (b) second examples. 98
5.5 Desired upper and lower masks (green and blue) and the achieved PP (red) shown in Fig. 5.4 but plotted in 3D. (a) First and (b) second examples. 98
5.6 Local power conservation mismatch at each unit cell for the (a) first and (b) second examples. The zero value means perfect match and the unity indicates 100% mismatch. 99

6.1 An illustration of the diffusion process for diffusion models. The dashed and solid arrows represent the forward and reverse diffusion processes, respectively. 103
6.2 An illustration of the U-Net architecture. 106
6.3 An illustration of the latent diffusion model (LDM). The conditional information, in our case, is the desired two-port scattering parameters along with the frequency of operation. 107
6.4 The parameters L, W, and T are the length, width and trace width of the dogbone. As shown in the left figure, these three dogbone layers are supported by two dielectric substrates (Rogers RO3010). 109
6.5 The three-channel pixelated binary image of dogbones (not to scale). . 111

6.6 The fabricated metasurface (for the wide-angle refraction example). The output face of the metasurface is visible in this figure. The spatial modulation is along the y axis, with no variation along the x axis. Five periods have been used along the y axis, with eight unit cells in one period. 114

6.7 The absolute error of the achieved scattering parameters for each unit cell with respect to the desired ones for the plane wave refraction example. Each period consists of eight unit cells. The final unit cell design is based on $\mathcal{C}_1$. The values of $\mathcal{C}_2$ and $\mathcal{C}_3$ are also reported for the sake of completeness. 115

6.8 Plot of the electric field magnitude at a fixed phase in Ansys HFSS. In this simulation, one period (P) of the metasurface has been used and the metasurface is enclosed by the periodic boundary conditions. In addition, PEC baffles have been used for this simulation. The left side of the metasurface is its input side (illuminated by ψ_{inc} where ψ denotes $\{E, H\}$) and the right side of the metasurface is its output side yielding the refracted wave ψ_{trans}. 117

6.9 Plot of the electric field magnitude at a fixed phase for the truncated metasurface where five periods ($5P$) are used to form the truncated metasurface (comprising of 40 unit cells and no PEC baffles). From each side, the metasurface is extended by quarter-wavelength absorbers. . . . 118

6.10 The normalized far-field power pattern plots of the truncated metasurface ($5P$ aperture size) for three cases: (i) Ansys HFSS simulation of the metasurface consisting of dogbone-shaped unit cells, (ii) Ansys HFSS simulation of the metasurface consisting of ideal impedance sheet unit cells, and (iii) the measured far-field power pattern of the fabricated metasurface. (The "correction factor" of $\cos\theta$ has not been applied to any of these plots.) 120

6.11 The far-field measurements of the fabricated metasurface in a compact antenna test range (CATR). The output face of the metasurface, i.e., the face that is supposed to interact with a plane wave at 69.6°, is illuminated by the serrated reflector. 121

6.12 The metasurface mounting and the receiving horn antenna. The horn antenna is located on the input side of the metasurface (i.e., the side that is supposed to interact with a normally incident plane wave). . . . 121

6.13 The gain pattern obtained from the CATR measurements with and without the metasurface. The maximum gain of the horn antenna structure with and without the metasurface are 14.1 dB and 10.6 dB respectively. 122

6.14 Plot of the electric field magnitude at a fixed phase. The metasurface is illuminated by a line source and then transforms the incident cylindrical wave to a truncated plane wave. 123

6.15 The absolute error of the achieved scattering parameters for each unit cell with respect to the desired ones for the circular wave to plane wave transformation example. The metasurface consists of 31 unit cells. The final unit cell design is based on $\mathcal{C}_3$. The values of $\mathcal{C}_1$ and $\mathcal{C}_2$ are also reported for the sake of completeness. 125

6.16 The achieved far-field power pattern from the metasurface designed to meet the desired far-field performance criteria. 126

6.17 The absolute error of the achieved scattering parameters for each unit cell with respect to the desired ones for the performance criteria design example. The metasurface consists of 39 unit cells. The final unit cell design is based on $\mathcal{C}_3$. The values of $\mathcal{C}_1$ and $\mathcal{C}_2$ are also reported for the sake of completeness. 127

6.18 Plot of the electric field magnitude at a fixed phase. The metasurface is illuminated by a line source and then transforms this incident field to an output aperture fields in an attempt to meet the desired far-field performance criteria. 128

7.1 ©2024 IEEE. DL-based end-to-end design pipeline. (The macroscopic and microscopic stages are based on work described in Chapter 4 and Chapter 6 respectively.) The first neural network is termed MBPS-LPC, which indicates a mask-based pattern synthesis (MBPS) approach with LPC. 131
7.2 ©2024 IEEE. Trapezoidal-shaped far-field (FF) mask and power patterns (PP) obtained from Ansys HFSS under three scenarios. 133
7.3 The fabricated metasurface in a wooden frame fastened by nylon screws. 134
7.4 The planar near-field range measurement setup for the fabricated metasurface. (Some absorbers are removed for a better view of the system.) 134
7.5 Trapezoidal-shaped far-field (FF) mask and power patterns (PP) obtained from Ansys HFSS simulation and PNFR measurements at 9.9 GHz and 10 GHz. 136
7.6 (a) The metasurface under test illuminated by a horn antenna; (b) calibration measurement where the face of the metasurface is covered by a copper tape. 137
7.7 The time-domain reflected signal captured by the VNA for two cases: metasurface under test, and the same metasurface covered by a copper tape. 138
7.8 The reflected signal captured by the VNA after applying time gating for two cases: the metasurface under test and the same metasurface covered by a copper tape. 139
7.9 The frequency-domain reflection response for the metasurface and the metasurface with a copper tape after applying time gating. 140

7.10 The frequency-domain reflection response for the metasurface after applying time gating and calibration. Note that due to calibration errors, some parts of the reflection are erroneously greater than 0 dB. . 141

A.1 Desired power pattern (PP) with a cosecant-squared shape, which is identified by lower and upper masks. The performances of the neural network are denoted by *PP Source* and *PP Truncated* whereas the performances of the metasurface (simulated by Ansys HFSS) employing dielectric substrates with different thickness values are denoted by *PP HFSS* and *PP HFSS (thinner substrate)*. 152

A.2 Plot of the electric field magnitude at a fixed phase in Ansys HFSS for Example II where a cascaded metasurface system employing a thinner substrate is used. The coordinate center is located at the center of Metasurface 1. 153

A.3 The forward path of unsupervised learning for the developed mask-based pattern synthesis (MBPS). 153

D.1 Training and validation loss curves for the 2D FF mask used in macroscopic metasurface design as described in Chapter 4. The orange line represents the training loss, while the blue line shows the validation loss over the training epochs. 164

D.2 Training and validation loss curves for the 3D FF mask used in macroscopic metasurface design as described in Chapter 5. The orange line represents the training loss, while the blue line shows the validation loss over the training epochs. 164

D.3 Training and validation loss curves for the Diffusion Model used in microscopic metasurface design as described in Chapter 6. The orange line represents the training loss, while the blue line shows the validation loss over the training epochs. 165

D.4 The three neural network architectures for the components of the developed DL framework for the far-field specifications to planar source array synthesis as described in Chapter 3, and more specifically in Fig. 3.6. . 166

D.5 The neural network architecture for the DL framework developed for mask-based pattern synthesis with LPC (MBPS-LPC) as describe in Chapter 4, and more specifically in Fig. 4.1. 167

List of tables

B.1 Unit cell dimensions for the first design example. All dimensions are in mm. 155
B.2 Unit cell dimensions for the second design example. All dimensions are in mm. 156
B.3 Unit cell dimensions for the third example. All dimensions are in mm. . 157

Nomenclature

Symbols

$\boldsymbol{B}$	Magnetic flux density
$\boldsymbol{D}$	Electric flux density
$\boldsymbol{E}$	Electric field
$\boldsymbol{H}$	Magnetic field
$\boldsymbol{M}$	Magnetic polarization density
$\boldsymbol{P}$	Electric polarization density
χ	Surface susceptibilities
η_0	Impedance of free space
μ_0	Permeability of free space
$\nabla\cdot$	Divergence Operator
$\nabla\times$	Curl Operator
∇	Gradient Operator
χ	Susceptibilities (volumetric)
ε_0	Permittivity of free space

c_0	Speed of light in free space
k_0	Wavenumber

Acronyms / Abbreviations

CATR	Compact antenna test range
CNN	Convolutional neural network
DL	Deep learning
EBG	Electromagnetic band gap
EM	Electromagnetic
FF	Far-field
FSS	Frequency-selective surface
GAN	Generative adversarial network
GPU	Graphics processing unit
GSTC	Generalized sheet transition condition
HPBW	Half-power beamwidth
LPC	Local power conservation
NF	Near-field
PCB	Printed circuit board
PNFR	Planar near-field antenna range
PP	Power pattern
SLL	side lobe levels

TE	Transverse electric
TM	Transverse magnetic
VAE	Variational autoencoders
VNA	Vector network analyzer

Other Symbols

ABCD parameters	Transmission parameters
S-parameters	Scattering parameters
Z-parameters	Impedance parameters

Chapter 1

Introduction

James Clark Maxwell predicted the existence of electromagnetic waves in 1864 by writing electromagnetic equations with an additional term, known as the displacement current density [1]. In 1884-1885, Oliver Heaviside rewrote the original 20 equations of Maxwell into a more compact form, which is currently being used [2]. In 1887, Heinrich Hertz was able to conduct his pioneering experiment to validate the existence of electromagnetic waves according to Maxwell's predication, which was followed by Guglielmo Marconi's wireless experiment and his first wireless patent application in 1896 [1]. In 1901, the first British ship employed Marconi apparatus [3]. Around 1902, Marconi experimentally demonstrated the possibility of wireless transmission, based on electromagnetic radiation, between Canada and UK [3]. Since then, scientists and engineers have made significant strides in creating different electromagnetic radiation patterns that have enabled modern society to communicate information at the speed of light. Today, wireless technology has become ubiquitous in our daily life. At any moment, we find ourselves surrounded by a multitude of wireless devices that send and receive information simultaneously.

At the heart of this technology lies the antenna: a component for converting guided waves (electric signals within a circuit) into electromagnetic waves propagating in free space, and vice versa. Despite the amazing advances in the field of antennas, the task of designing and optimizing modern antenna systems to meet competitive standards of

lightweight construction, conformal design, broadband capability, and cost efficiency has become increasingly challenging. In addition, the desired specifications can take many different forms in the far-field (FF) and near-field (NF) zones[1]. For example, in the FF zone, the desired specifications can be in the form of main beam directions, half-power beamwidths (HPBWs), and maximum side lobe levels (SLLs). Alternatively, the FF criteria can be described by lower and upper masks such as a radiation pattern that follows a trapezoidal shape.

When faced with such diverse performance criteria, we may encounter situations where there are no established equations or design paths for antenna development. This non-systematic nature makes the design and optimization process of new antenna systems an iterative process whose starting point heavily relies on the designer's intuition and experience.

Over the past decades, the field of metamaterials and metasurfaces have garnered significant research interest due to its potential ability to systematically manipulate electromagnetic waves. The term 'metamaterial' was originally coined around the turn of 21st century by Walser [5–7] to describe artificial materials that posses electromagnetic properties beyond those naturally occurring in conventional materials [8–10]. However, humans have been working with artificial materials long before the term was introduced. One notable example is the stained glass use on many religious buildings' windows. The incorporation of metallic inclusion (metallic salts) during the manufacturing process facilitated the creation of the various colors in the stained glass. The metallic salt embedded within the glass can selectively absorb certain wavelengths of light, contributing to the specific hues observed in the finished artwork [5]. Today, the general consensus on the definition of "metamaterial" as "an arrangement of artificial structural elements, designed to achieve advantageous and unusual electromagnetic properties" [11]. This definition draws the following analogy between metamaterials and natural materials. The constituent elements of natural materials are atoms whereas

[1]The near-field zone is a region immediately around the antenna which extends for at least a few wavelengths. After the near-field zone, the far-field zone begins. For a more detailed description, see [4].



Fig. 1.1 A sample metasurface placed within a wooden frame.

those of metamaterials are engineered sub-wavelength inclusions often referred to as meta-atoms. The proper design of these meta-atoms enables the metamaterials to effectively manipulate the incident electromagnetic wavefront. For instance, researchers in [12] successfully engineered a 'left-handed' or double-negative metamaterial (simultaneously having negative permeability and permittivity) using a periodic array of subwavelength split-ring resonators and wires. This work was based on the concept developed by Veselago in 1968 [13].

EM metasurfaces, e.g., see Fig. 1.1, are considered as the two-dimensional (surface) counterpart of three-dimensional (volumetric) metamaterials [14–17]. They are considered as "surfaces" since their thickness is small compared to the wavelength of operation. Similar to metamaterials, they consist of subwavelength scatterers, known as meta-atoms or unit cells, which are incorporated within their thin structures. For example, see a close-up view of a sample metasurface in Fig. 1.2 showing a few of its unit cells. The interaction of these meta-atoms with the incident electromagnetic wave

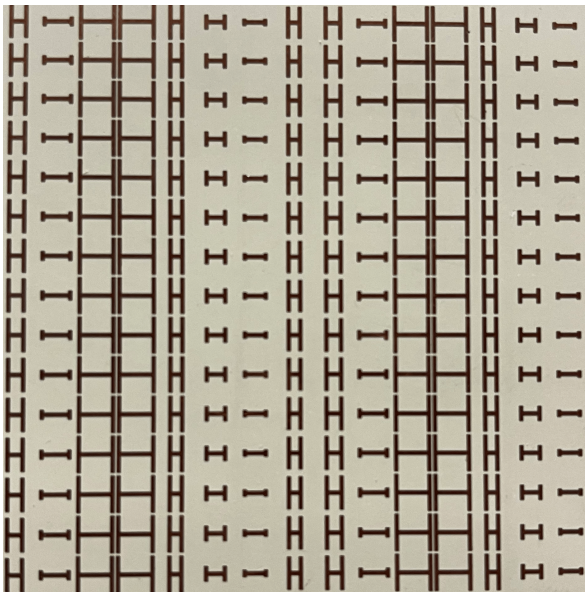


Fig. 1.2 A close-up view of a sample metasurface showing its unit cell.

can tailor the resulting electromagnetic wave in various ways including unconventional manners. An example for an unconventional performance is achieving perfect refraction (i.e., refraction without incurring any reflections) [18, 19]. In particular, each scattering element is engineered to exert control over the amplitude, phase, and/or polarization of its local electromagnetic field, leading to a collective desired impact [15–17].

Similar to metamaterials whose overall properties can be described by their homogenized permittivity and permeability values, the presence of subwavelength scatterers within the metasurface structure allows for the metasurface to be characterized using homogenized surface properties.[2] These surface properties can then be included in zero-thickness sheet transition conditions which describe the boundary conditions associated with the metasurface. These boundary conditions are known as the generalized sheet transition conditions [22, 21]. Compared to metamaterials, metasurfaces present practical advantages such as reduced fabrication complexity, decreased bulkiness and weight, lower losses, and improved scalability. As a result, they have attracted increased research attention over the last decade. For example, they are considered as

[2]Note that due to the subwavelength thickness of the metasurface, it is not appropriate to model it with permittivity and permeability values [20, 15]. Instead, surface susceptibilities [21] can be used.

candidates in next generation wireless communication systems to tailor communication channels [23].

The performance of metasurfaces can be justified based on the electromagnetic surface equivalence principle. This principle allows us to replace a closed region containing actual sources (e.g., radiating antennas) by a set of equivalent surface currents on its boundary. This principle, introduced by Schelkunoff in 1936 [24, Sec 7.8], is also referred to as Huygens equivalence principle in honour of Dutch physicist Christiaan Huygens, who introduced the idea of "secondary source concept" in 1690 [25, Sec 6.3.1]. The surface equivalence principle is the theoretical foundation for near-field antenna measurements that enables us to calculate far-field patterns from near-field measurements [26]. Similarly, it is used to find the radiation patterns of aperture antennas such as horn antennas. In the case of metasurfaces, the goal is to design metasurfaces that support the existence of necessary surface equivalent currents to achieve the desired transformation. That is why metasurfaces may be viewed as physical implementation of Huygens principle. Note that in antenna engineering, the term "Huygens source" refers to perpendicular electric and magnetic currents [27, Sec 4.4.3]. Therefore, metasurfaces that implements such currents are sometimes referred to as Huygens metasurfaces [28, 29].

Although the metasurfaces terminology has emerged only in recent years, the conceptual notion of manipulating the propagation of electromagnetic waves with thin artificial structures has been present for many years. As noted in [10], one of the earliest examples of metasurface-like structures can be traced back to electromagnetic wave transmission and reflection from a grid of metallic wires in a 1898 paper by H. Lamb [30]. In this paper, Lamb showed that by changing the geometrical features of the grid, the transmission and reflection of electromagnetic waves can be controlled [30]. This draws analogy to designing proper meta-atoms for metasurfaces to tailor electromagnetic radiation [10].

More recently, frequency-selective surfaces (FSSs), heavily developed in the latter half of the 20th century, consist of a periodic arrangement of resonant-length struc-

tures or elements, typically etched onto a thin metallic film or printed on a dielectric substrate [31, 14]. FSSs are engineered to selectively permit specific electromagnetic waves to pass through while reflecting or absorbing others based on their frequency and angle of incidence. They have been used in a spectrum of applications including frequency-selective filters, beam-steering devices, polarization converters, and electromagnetic shields [32]. Another example include the reflectarray and transmitarray antennas. Reflectarray antennas, or simply reflectarrays, consist of an array of scattering elements (often patches or slots) placed on a surface. These elements are designed to reflect incident electromagnetic waves with controlled phase shifts. By adjusting the size, shape, and spacing of these elements, reflectarrays can manipulate the reflected wavefront to achieve desired far-field responses [33]. Transmitarray antennas, or simply transmitarrays, operate similarly to reflectarray antennas but are designed to transmit electromagnetic waves rather than reflecting them [34]. This gives them the benefit or avoiding the feed blockage issue [35]. The key distinction between the conventional form of these devices and metasurfaces lies in the size of their unit cells, which are typically in the order of half a wavelength. Due to this, in contrast to metasurfaces, they cannot be characterized as homogenized materials [15, 14]. Similarly, other electrically thin structures such as electromagnetic band gaps (EBGs)[3] [36] with unit cell sizes comparable to the wavelength cannot be considered as metasurfaces [37, Sec 1.1]. Finally, it should be noted that we have focused on the conventional forms of these devices. On the other hand, if these devices employ subwavelength unit cells, they can be characterized as metasurfaces. For example, a reflectarray with subwavelength unit cells is a reflective metasurface.

Thin structures such as near-field plates, which comprise an array of closely positioned subwavelength scatterers, can be classified under metasurfaces. These near-field plates were originally proposed to create very focused beams in the near-field zone of the near-field plate [38]. In addition, holographic artificial impedance surfaces were proposed to create desired radiation patterns [39, 14]. For example, the unit cell size

[3]EBGs can be used to suppress unwanted radiation or to improve the performance of antennas by reducing sidelobes.

in [39] was about one-sixth of the wavelength, which means that the surface can be categorized as a metasurface. These types of structures [39, 40] (i.e., a printed metallic subwavelength pattern supported by a grounded thin dielectric substrate excited by sources such as a monopole) were later referred to as metasurface antennas [41, 14]. These devices provided important theoretical foundations to the current metasurface development.

An important development in the area of electromagnetic metasurfaces was published in 2011 by Capasso's group [42] in which they demonstrated the generalized laws of reflection and refraction by employing a two-dimensional single-layer array of subwavelength elements. The ability to obtain such anomalous reflection and refraction was of high interest in the Antennas and Electromagnetic community since we can steer beams in desired directions with passive artificial surfaces. In particular, anomalous refraction in transmissive metasurfaces received significant attention but encountered a problem: undesired reflections from the surface. To alleviate these reflections, Huygens' metasurfaces, also known as metamaterial Huygens' surfaces, were proposed in 2013 by Grbic's group [16]. Their proposed metasurface structure used two layers of copper traces to realize Huygens' sources (i.e., perpendicular electric and magnetic currents). As noted in [43], it was shown in 2014 by Epstein and Eleftheriades that small undesired reflections are still needed to be present for a fully lossless and passive metasurface design even with such Huygens' metasurfaces [44]. Later in 2016, it was shown that by further increasing the degrees of freedom in the lossless and passive metasurface design, it is possible to achieve anomalous refraction without undesired reflections [45–47, 43]. This was then referred to as perfect refraction. The extra required degree of freedom was using omega bianisotropy, thus, having omega bianisotropic[4] metasurfaces. In most designs, this extended the double-layer design of Huygens' metasurfaces [16] to three-layer designs [18]. More discussion on omega bianisotropy will be presented in Section 2.1.

[4]As noted by Kong in his 1972 paper [48], "*the constitutive relations of a bianisotropic medium relates* $\mathbf{D}$ *to both* $\mathbf{E}$ *and* $\mathbf{B}$*, and* $\mathbf{H}$ *to both* $\mathbf{E}$ *and* $\mathbf{B}$" where $\mathbf{D}$ is the electric flux density, $\mathbf{B}$ is the magnetic flux density, $\mathbf{E}$ is the electric field intensity, and $\mathbf{H}$ is the magnetic field intensity.

Tailoring electromagnetic fields by the metasurface can be done in the reflective mode, transmissive mode, and a combination thereof. In this thesis, we are concerned with the transmissive mode, which implies that we prefer to have zero reflection from the metasurface. That is why the metasurfaces considered in this thesis are referred to as transmissive metasurfaces. As of today, a myriad of metasurfaces have been developed for a wide variety of applications by different researchers. This includes beam shaping applications [19, 46, 47], polarization control [49–51], electrically thin absorbers [52, 53], near-field focusing [38], spatial processing [54], analog computation [55, 56], energy harvesting and imaging [57, 58], impedance matching [59], and many others. Furthermore, recent advancements have been made in reconfigurable metasurfaces, where unit cells can be reconfigured (e.g., utilizing diodes or varactors) to adjust their response [60, 61]. If this reconfigurability can be automated, these metasurfaces can be referred to as 'intelligent' metasurfaces. They can then be installed on windows or walls to create smart radio environments to enhance wireless communications, avoid blockage issues, and offering a solution beyond solely relying on advancements in either the transmitter or receiver sides [23].

In addition, temporal modulation can be used in unit cells to create new frequency components. This can have application such as camouflaging and security [62]. We also note that for beam shaping applications noted above, we need to utilize aperiodic metasurfaces, which are known as spatially modulated metasurfaces in our context. That is, the properties of neighboring unit cells are not identical. (A sample periodic metasurface is shown in Fig. 1.3 showing identical unit cells in each direction. This was originally designed in [53].) If spatial modulation and temporal modulation are combined, we arrive at spatiotemporally modulated metasurfaces, or space-time metasurfaces [63]. These metasurface can not only perform beam shaping but can also tailor the frequency contents [64].

Despite the complexity of metasurface structures, their performance can be described by appropriate sheet (zero-thickness) boundary conditions. This provides a significant advantage for the design process of metasurfaces. The use of boundary conditions to

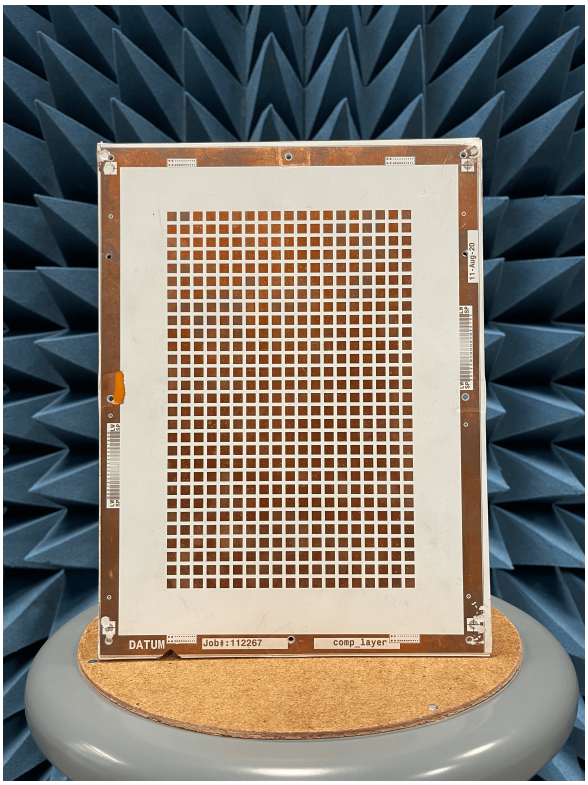


Fig. 1.3 A sample fully periodic metasurface, originally designed by Z. Liu et al.

describe the performance of thin structures is not new in electromagnetics. For example, as noted in [14], Kildal considered transversely and longitudinally corrugated surfaces and modeled them with boundary conditions associated with soft and hard surfaces, respectively in 1988 [65, 66]. These corrugated surfaces have different applications, for example in the design of horn antennas with rotationally symmetric beams with low cross polarization [67]. In addition, simple impedance boundary conditions were used in the design of artificial surfaces such as near-field plates [38]. Let us now consider the development of metasurface boundary conditions as described in [68]. An important contribution to the generalization of conventional electromagnetic boundary conditions, which was later adapted to metasurfaces, was the work of Idemen and Serbest in 1987 [69]. They expressed the discontinuity of electromagnetic fields across a thin interface based on impressed currents *and* surface polarization densities. We now refer to these boundary conditions as the generalized sheet transition conditions or simply GSTCs. In 2003, Kuester et al. used these GSTCs and then expressed the surface

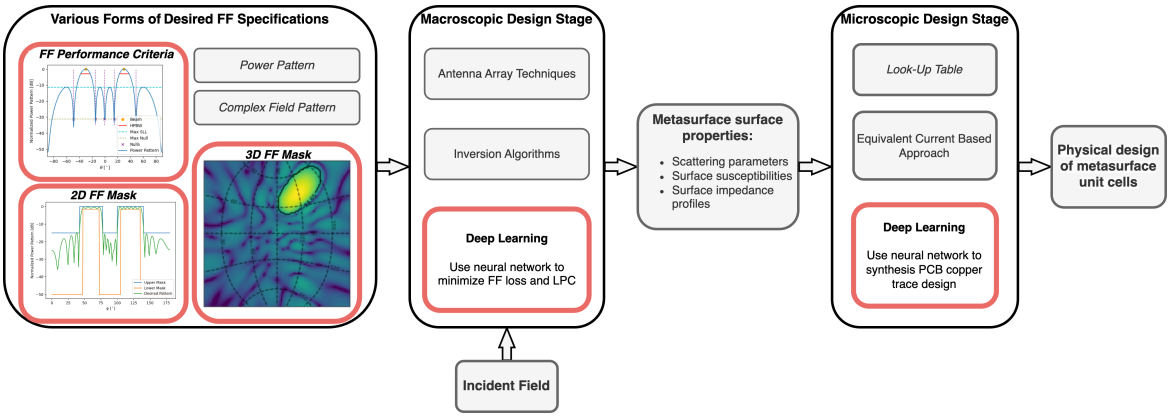


Fig. 1.4 The metasurface design pipeline. The highlighted parts in red are the focus of this thesis. This figure is a modified version of Fig. 5.1.

polarization densities within the GSTCs based on the surface polarizabilities of meta-atoms and average fields [22]. As noted in [68], Holloway et al. subsequently expressed the GSTCs in terms of surface susceptibilities and average fields in 2012 [15]. Although Holloway et al. acknowledged that their formulation can be extended to bianisotropic metasurfaces (i.e., with magnetoelectric coupling terms), their presented formulation was limited to monoanisotropic metasurfaces (i.e., without magnetoelectric coupling terms). More details on magnetoelectric coupling will be presented in Section 2.1. Then, in 2015, Achouri et al. presented the GSTC formulation for bianisotropic metasurfaces using surface susceptibilities and average fields [21]. Finally, it should be noted that metasurface boundary conditions can also be written in other forms including those that incorporate surface electric impedance, surface magnetic admittance and the magnetoelectric coupling coefficient in conjunction with average fields [29].

Now that we have covered the background on metasurfaces and their boundary conditions, let us briefly describe how metasurfaces are designed. At the current state of the art, the metasurface design process can be partitioned into two major steps: 1) macroscopic design, and 2) microscopic design [28] as illustrated in Fig. 1.4. In the macroscopic design step, the main goal is to calculate the necessary metasurface properties based on the desired specifications to be achieved. These specifications can take various forms, including (FF) performance criteria such as main beam directions,

half-power beamwidths, null locations, and sidelobe levels [70, 71]. Alternatively, the desired response may be expressed through FF lower and upper masks, within which the resulting power pattern (PP) must reside [72–74]. Additionally, complex field patterns may be employed to describe the desired response, which requires simultaneous specification of amplitude and phase patterns [75]. The necessary metasurface properties can be represented in various forms such as scattering parameters, surface susceptibilities, surface impedance profiles, and surface polarizabilities based on the input-output electromagnetic field requirements [16, 76, 47, 77]. Then, in the microscopic design, the designer needs to develop a physical implementation of the metasurface that realizes the requirements of the previous step [16, 47, 78]. This typically requires the choice of dielectric substrates, dielectric bondply, and copper trace design on these substrates. The copper trace elements serve as the subwavelength inclusions of the metasurface. These copper traces, in conjunction with the chosen dielectric substrate, aim to implement the required macroscopic properties. Typically, the microscopic design of metasurfaces is based on the local periodicity assumption. This assumption means that the properties of neighboring unit cells do not change significantly. For spatially modulated metasurfaces, this assumption can be violated, which then results in a non-ideal performance. More recently, an approaches that does not rely on local periodicity has been proposed. This method uses surface equivalent currents around the unit cell as opposed to widely-used periodic boundary conditions [79].

Over the past decade, the field of deep learning (DL) [80] in the form of deep neural networks [81], has shown incredible results in performance and has become the defacto method in many machine learning applications. As noted in [82, Sec 2.2], the development of neural networks can trace its roots back to the 1940s by the works of Warren McCulloch and Walter Pitts that modeled how neurons in the brain might work [83]. However, it wasn't until the 1980s that DL began to gain significant traction. A pivotal moment occurred in 1986 when Geoffrey Hinton co-authored the backpropagation algorithm, which revolutionized neural network training methodologies [84]. This breakthrough laid the foundation for many future

advancements in DL. In the early 2000s, DL faced a period of stagnation due to limitations in computing power and data availability. However in 2012, AlexNet [85], a deep convolutional neural network (CNN) developed by Alex Krizhevsk of Hinton's group, won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) that has reignited interest in the field. Another pivotal role of accelerating the progress of DL research is the advent of Graphics Processing Units (GPUs) around 2010 by providing the computational power necessary for training deep neural networks at scale. Since then, DL has continued to flourish, driving innovations across various domains, such as image classification [85], object detection [86], and natural language processing [87]. More interestingly is that DL models are not only capable of conventional discriminate tasks but have also shown success in various generative tasks such as text generation [88], text-to-image generation [89],and super resolution [90] just to name a few. Such generative models include generative adversarial networks (GANs) [91] and variational autoencoders (VAEs) [92]. Furthermore, DL techniques have also started to contribute to other disciplines as well, such as biology and drug discovery [93], materials science [94], and physics [95].

In addition, recent research results show that the DL based methods also hold great promise in addressing challenges in electromagnetic related applications [96–98]. For example, DL has enabled advances in electromagnetic modeling and computational electromagnetics where DL models can efficiently approximate solutions to complex electromagnetic problems and solve Maxwell's equations faster with reasonable accuracy, significantly speeding up the modeling process by reducing reliance on time-consuming full-wave electromagnetic simulations [99–101]. Another example is antenna and array designs, the DL-based approaches offer significantly faster performance while maintaining high accuracy compared to traditional methods and satisfying complex constraints [102, 103]. Moreover, DL-based approaches have been used in electromagnetic imaging and sensing applications such as reconstruction of permittivity images [104] and other properties [105]. Additionally, DL-based approaches have been used for data calibration in electromagnetics. For example, a DL-based approach

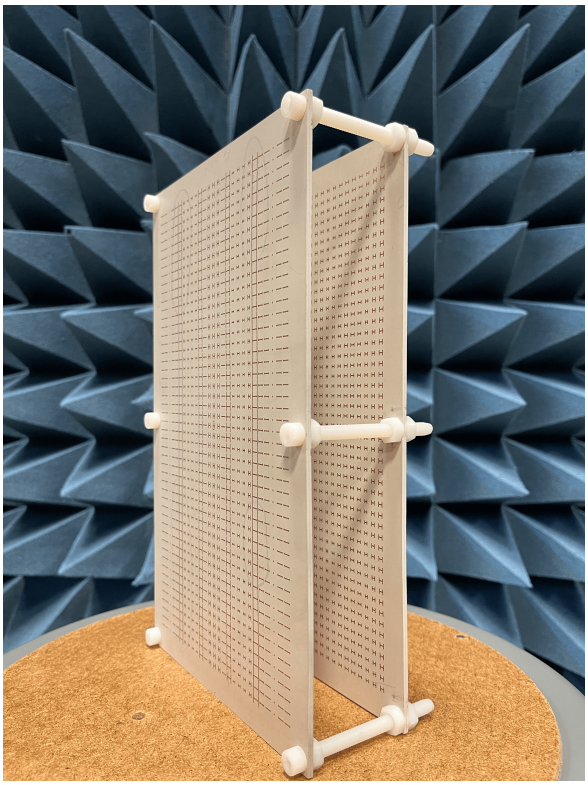


Fig. 1.5 A sample cascaded metasurface system designed by J. Gohel.

has been used to calibrate the data collected in a three-dimensional environment to be used with tomographic (two-dimensional) imaging algorithms [106]. Finally, DL approaches have also been explored and utilized in metasurface designs. In particular, different DL models have been developed for microscopic design of different types of metasurfaces [78, 107–109]. As noted earlier, in this thesis, we are focused on transmissive metasurfaces. In this context, our goal is to investigate an end-to-end design pipeline for transmissive metasurface design that is able to handle various forms of FF specifications using DL techniques. The motivation behind this is explained in the following section.

1.1 Motivation

In the macroscopic design of metasurfaces, several methods have been developed to determine the necessary surface properties when user-defined near-field or far-

field specifications are given. This step becomes straightforward for scenarios where analytical expressions for the tangential fields on the faces of the metasurface are available. Examples of these scenarios are plane wave refraction where the tangential fields on the faces of the metasurface can be found based on plane wave expressions. In some other special cases, antenna array techniques can be used. For example, we can use the so-called Taylor line-source (one parameter) to achieve a radiation pattern with decaying minor lobes where we use the single parameter to control the maximum sidelobe level [4, Sec 7.7]. This has been, for example, used in the design of cascaded metasurfaces [110]. (A sample cascaded metasurface is shown in Fig. 1.5, which is originally reported in [111].)

However, such approaches are limited to the special cases supported by the antenna array techniques. To generalize the approach (i.e., inferring necessary tangential fields and surface properties) to various forms of user-defined specifications, researchers have proposed formulating the problem as an inverse problem. One of the early works in proposing inverse formulation for design applications can be found in [112]. This work essentially formulates the antenna design problem as an electromagnetic inverse source problem. In the area of metasurfaces, this idea was used to infer the necessary equivalent currents of the metasurface to generate a desired set of near-field or far-field specifications. Once these equivalent currents are determined, they can be used to find the required tangential fields on the metasurface, and finally the required metasurface properties can be determined [113, 75]. This approach is also called inverse design, inversion for design, or source reconstruction method for design [71, 114]. The implementation of these approaches are mostly done through gradient based optimization techniques. Alternatively, DL techniques can be used for inferring the tangential fields and metasurface properties [70]. In particular, DL techniques provide their solution almost in real time, and can also adapt easily to other design scenarios such as reconfigurable metasurface design [115].

The above approaches are presented without considering additional design constraints associated with metasurface properties. In particular, for our transmissive

metasurface design, we are interested in lossless and passive structures. To this end, we need to ensure that the normal real power density entering each metasurface unit cell is equal to the normal real power density leaving the same unit cell. This condition is often referred to as local power conservation (LPC) constraint [47]. Therefore, all of the above techniques may need to be modified when constraints such as LPC are considered. For example, imagine a metasurface that is to transform a line source (cylindrical waves) to a plane wave. Although both cylindrical and plane wave expressions are available analytically, we cannot design a lossless and passive metasurface for this transformation in a straightforward manner. This will then require more advanced treatment so as to enforce the LPC constraint [116]. Similarly, this is also the case for transforming a cylindrical wave to a Taylor (one-parameter) aperture distribution using a lossless and passive metasurface system. For example, the authors of [110] have used cascaded metasurface systems to perform this transformation.

In this thesis, the LPC constraint for lossless and passive metasurfaces is considered. Thus, it is worthwhile to classify how the LPC constraint can be integrated into the design process. For the design of lossless and passive transmissive metasurfaces, the macroscopic design process that accounts for the LPC constraint, can be broadly classified into (i) two-step and (ii) single-step schemes.[5]

In the two-step approach, the desired far-field specifications and the LPC constraint are considered separately. To this end, an appropriate tangential aperture field distribution is first generated that satisfy the desired far-field specifications. The fields then need to be further optimized to meet the LPC constraint using methods such as the incorporation of auxiliary surface waves [116, 118, 119] and utilization of cascaded metasurface pairs [110, 120]. The first stage has been mostly done using various non-DL based methods such as analytical expressions, antenna array techniques [110] and inversion (inverse source) algorithms [26]. As noted earlier, techniques based on analytical expressions are limited to special cases, and even in those cases, they typically do not take into account practical features such as the finite aperture size

[5]This part is based on the Introduction section of the author's publication at the *IEEE Antennas and Wireless Propagation Letters*, 2024 [117].

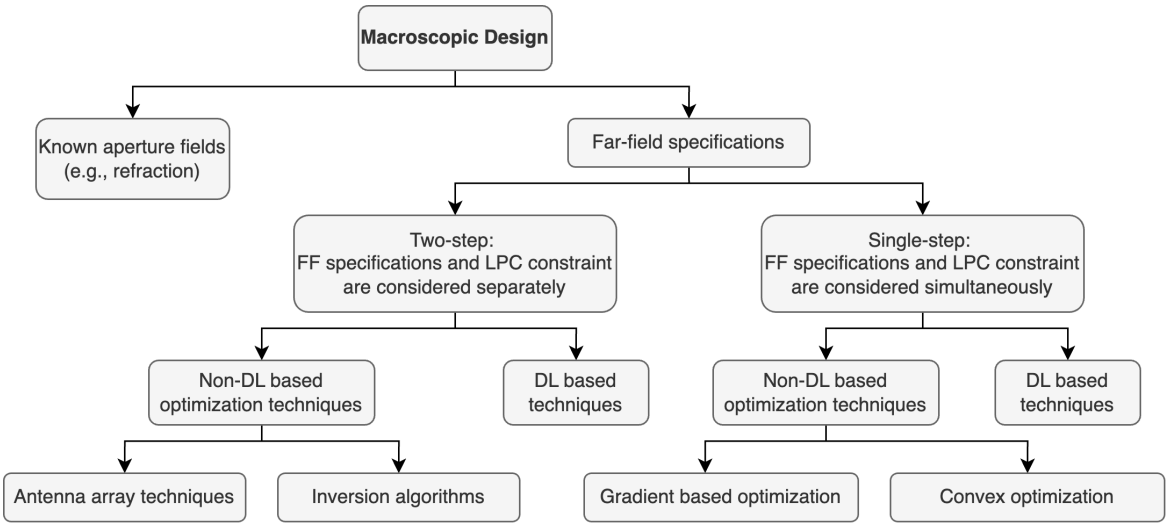


Fig. 1.6 A classification of the macroscopic design approaches for metasurface design

of the metasurface. For the other techniques, handling non-conventional far-field and near-field criteria such as maximum null levels and non-conventional beam shapes can still be challenging. As an example, as noted in [71, Sec III.D.1], their proposed inverse solver needs to make an assumption about both the phase and magnitude of the electric field to favour a certain beam level at a desired angle. This is not ideal since in practice the user only specifies the desired beam level in dB (a phaseless quantity). As another example, the inverse solver presented in [75] requires the user to defined all the sidelobes. In many cases, this is too restrictive as the user may only care about the maximum sidelobe level, rather than their exact shape. In such scenarios, the DL-based techniques can be advantageous. In addition, DL-based techniques are almost real-time in their prediction mode, which is also advantageous if the metasurface performance needs to adjust dynamically (e.g., reconfigurable metasurfaces for smart radio environments).

Let us now focus on the main disadvantage of the two-step schemes. We note that the first stage of this two-step approach constitutes an electromagnetic inverse source problem, which can have non-unique solutions. This non-uniqueness could result in an aperture field whose power profile significantly differs from that of the incident field. As a result, the optimization algorithm utilized in the second stage to enforce the LPC

may face challenges in aligning these two power profiles through power redistribution. If this situation occurs, this can be conceptually viewed as the first stage providing a non-ideal *initial guess* for the second step.

To alleviate the above challenge, the single-step schemes have been proposed. This approach directly integrates the LPC constraint into the formulation when generating the metasurfaces' aperture fields. This requires the algorithm to find an aperture field that would simultaneously satisfies both the far-field specifications and the LPC constraint. Several non-DL optimization based method have been developed such as non-linear conjugate gradient approaches [121, 122] and modified form of convex optimization [71]. For example, in [121, 122], a regularization term is added to the cost functional that relates the real and imaginary parts of equivalent currents to the incident power at each unit cell so as to meet the LPC constraint. This can then be viewed as incorporating auxiliary surface waves into the solution in an implicit manner [122]. Due to the incorporation of the LPC constraint, we are not aware of any analytical expressions and array techniques for such the single-step scheme. Moreover, similar to the case for the two-step scheme, it can be challenging for inversion methods to handle non-conventional far-field or near-field specifications. In summary, the taxonomy of the methods for metasurface macroscopic design is illustrated in Fig. 1.6.

Having discussed the macroscopic design step, let's now turn our attention to the microscopic stage. The microscopic design stage of metasurface involves designing the physical construction of unit cells to achieve the desired surface properties obtained in the macroscopic step. The method for the microscopic step is still not systematic since it relies on an iterative optimization process using a large lookup table and adjusting various design parameters (such as the dimensions of the sub-wavelength inclusions) through full-wave simulations [28, 123]. For transmissive metasurfaces, the unit cell design often involves employing multi-layered unit cell structures, such as three-layer [18, 123] or four-layer [124] designs. However the lookup tables are typically generated from single-layer unit cell simulations, and thus disregarding the near-field

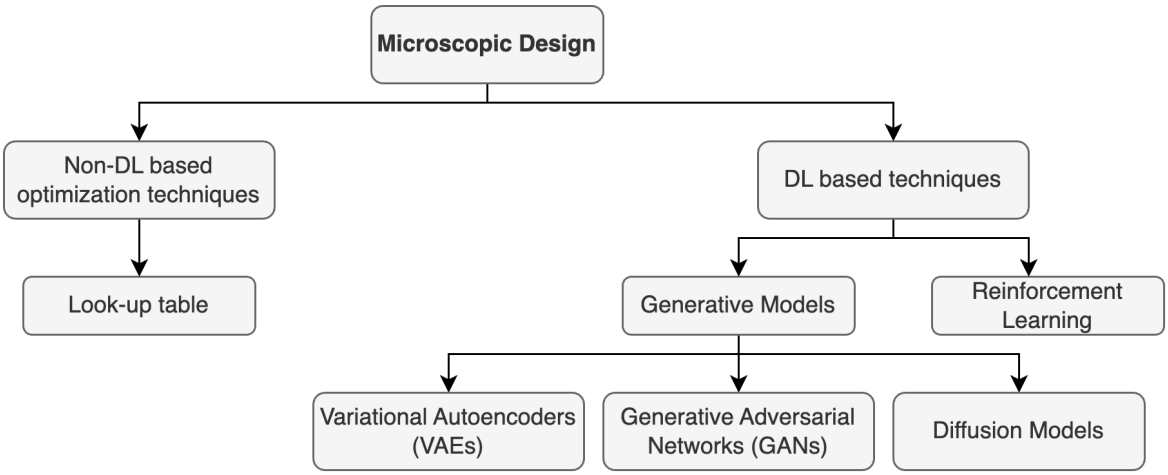


Fig. 1.7 A classification of the microscopic design approaches for unit cell design

coupling effects in scenarios involving multi-layered metasurface designs. To address this concern, approaches like those described in [125] have been developed to integrate the near-field coupling effect into the unit cell design process, which requires further optimization and fine tuning in a full-wave forward solver.

The current state-of-the-art approach utilizes generative DL models to provided an appropriate design (or, initial design) for a multi-layer unit cell configuration. This approach utilizes a range of generative DL techniques for metasurface design, including Generative Adversarial Networks (GANs) [108, 126] and Variational Autoencoders (VAEs) [78, 107]. However, there are challenges associated with these two methods as well. One disadvantage of GANs is their inherent instability during training, often resulting in mode collapse or oscillations. Additionally, GANs are known to be sensitive to hyperparameters and initialization choices, which can make training difficult and require extensive tuning [127]. On the other hand, VAEs for generative tasks tends to produce blurry or low-quality designs. This blurriness is often attributed to the loss function used in VAEs, which penalizes the deviation of generated samples from the mean of the learned distribution [128]. As a result, VAEs struggle to capture fine-grained details and produce sharp images and in this case metasurface unit cell designs. As will be seen later in the thesis, we have used diffusion models for the microscopic design to address the shortcomings of VAEs and GANs [129]. Finally,

for metasurfaces employing reconfigurable unit cells, reinforcement learning has been used to determine the necessary states of the reconfigurable elements, e.g., determining the on or off states of PIN diodes [130]. In summary, the taxonomy of methods for metasurface microscopic design is illustrated in Fig. 1.7.[6]

1.2 Objectives and Contributions

Given the aforementioned limitations, the main objective of this undertaken research is to investigate and address the challenges associated with the current metasurface design process by creating a systematic end-to-end DL based metasurface design framework. This proposed design framework aims to guide antenna designers to go from user-defined far-field specifications (e.g., beam directions, null locations, side lobe levels, and half-power beam widths) or far-field masks to the appropriate physical design of metasurfaces (i.e., copper traces on PCB).

There are various forms of metasurfaces such as reflective (usually with a metallic ground plane) and transmissive metasurfaces. To limit the scope of this thesis, the research in particular is concerned with a specific type of transmissive metasurfaces known as omega-type bianisotropic metasurfaces [47, 18]. Consequently, the objective is to transform a known incident field such as a known radiating antenna into a desired power pattern that meets the user-defined far-field specifications. To this end, the user input for the design framework are: (i) the frequency of operation (within the X-band in this thesis), (ii) desired far-field specifications, (iii) electrical size of the metasurface (planar metasurfaces with the typical size of about 5 to 15 wavelengths in each direction), (iv) the dielectric substrate to be used in the design of the metasurface (e.g., Rogers RO3010 with the relative permittivity of about 10.2, the loss tangent of about 0.0022, and thickness of 1.28 mm), and (v) the incident field with which the metasurface is illuminated. Since we are designing transmissive metasurfaces,

[6]It should also be mentioned that all the techniques presented for the microscopic design stage are based on the local periodicity assumption. More recently, this assumption has been relaxed through the use of equivalent surface currents instead of periodic boundary conditions [79]. This technique can be used in conjunction with all the techniques mentioned above.

the reflections from the metasurface need to be minimized so as to maximize the transmission efficiency. Furthermore, it should be noted that metasurfaces can be employed to change the polarization of electromagnetic waves (e.g., linear polarization to circular polarization). In this thesis, we focus on linearly polarized scenarios where no change of polarization is desired. To meet the overall objective, we consider the following sub-objectives.

Sub-Objective 1 Herein, we aim to build a DL framework that synthesizes metasurface aperture fields based on user-defined far-field specifications (such as desired beam directions, null angles and half-power beamwidths). This DL framework does not consider the LPC constraint. Thus, the LPC constraint is enforced in a subsequent stage using auxiliary surface waves in conjunction with a gradient-based optimization approach. Therefore, this approach belongs to two-step schemes noted in Section 1.1. The performance of the metasurface is evaluated using full-wave simulations with unit cells formed by impedance sheets.

- This topic will be discussed in Chapter 3.
- This chapter is taken from the following journal paper: **C. Niu**, M. Phaneuf, T. Qiu and P. Mojabi, "A Deep Learning-Based Approach to Design Metasurfaces From Desired Far-Field Specifications," *IEEE Open Journal of Antennas and Propagation*, vol. 4, pp. 641-653, 2023.
- The main contributions of the author in this chapter (paper) are the development of the DL framework, full-wave simulations, and writing the paper.

Sub-Objective 2 Herein, we aim to build a DL framework for macroscopic design that considers desired far-field specifications in the form of lower and upper far-field masks and LPC constraint simultaneously. Therefore, this approach belongs to single step schemes defined in Section 1.1. The metasurface is then evaluated using full-wave simulations with unit cells constructed with impedance sheets.

- This topic will be discussed in Chapters 4 and 5.
- Chapter 4 is taken from the following journal paper: **C. Niu**, M. Phaneuf and P. Mojabi, "Transmissive Metasurface Synthesis from Far-Field Masks Using Unsupervised Learning," *IEEE Antennas and Wireless Propagation Letters*, vol. 23, no. 8, pp. 2371-2375, Aug. 2024.
- Chapter 5 is taken from the following conference paper: **C. Niu** and Puyan Mojabi, "Synthesis of Aperture Fields for Transmissive Metasurface Design from 3D Masks Using Deep Learning," URSI Atlantic Radio Science Meeting (AT-RASC), pp. 1-4, Gran Canaria, Spain, May 19-24, 2024 (Recipient of a Young Scientist Award).
- The main contributions of the author in these chapters (papers) are the development of the DL framework, full-wave simulations, and writing the papers.

Sub-Objective 3 Herein, we aim to build a DL framework that can generate the microscopic design (i.e., three-layer metallic trace design) of metasurfaces based on the required macroscopic surface properties. In addition to performing full-wave simulations to verify the metallic trace design, we experimentally verify the performance of two metasurfaces designed using this approach at the University of Manitoba's Compact Antenna Test Range (CATR) and planar near-field antenna range (PNFR).

- This topic will be discussed in Chapters 6 and 7.
- Chapter 6 is taken from the following journal paper: **C. Niu**, M. Phaneuf and P. Mojabi, "A Diffusion Model for Multi-Layered Metasurface Unit Cell Synthesis," *IEEE Open Journal of Antennas and Propagation*, vol. 4, pp. 654-666, 2023.
- Chapter 7 is partially based on the following conference paper: **C. Niu**, M. Phaneuf, and P. Mojabi, "End-to-End Transmissive Metasurface Design from

Far-Field Masks Using Deep Learning," *IEEE International Symposium on Antennas and Propagation and ITNC-USNC-URSI Radio Science Meeting*, pp. 1-2, Florence, Italy, 2024.

- The experimental part of Chapter 7 is yet to be submitted for publication.
- The main contributions of the author in this chapters (papers) are the development of the DL framework, full-wave simulations, experimental work, and writing the papers.

In summary, it can be seen that Sub-Objectives 1 and 2 are concerned with macroscopic design, while Sub-Objective 3 deals with microscopic design followed by experimental evaluation. Therefore, this thesis investigates and develops a DL based approach for end-to-end metasurface design. That is, the contribution of this thesis is the development of DL-based macroscopic and microscopic procedures for reciprocal, lossless and passive transmissive metasurface design along with its full-wave and experimental verifications.

1.3 Structure of the Thesis

As indicated above, this thesis is presented in the grouped manuscript style (sandwich thesis). To this end, we first provide background information about metasurfaces and DL in Chapter 2. The next three chapters, Chapters 3, 4, and 5, focus on the macroscopic design stage of metasurfaces. Then, Chapters 6 and 7 investigate (i) the microscopic design stage, (ii) its integration with the macroscopic design stage, and (iii) experimental evaluation. Finally, Chapter 8 concludes this thesis and present some future works.

Chapter 2

Background

In this chapter, we present the background information on (i) metasurfaces and (ii) deep learning to facilitate the understanding of the materials in the following chapters.

2.1 Metasurface Fundamentals

In this work, the metasurfaces discussed belong to the category of linear bi-anisotropic media. We mainly follow [131] for presenting the background required for metasurfaces. To derive the constitutive relations for linear bianisotropic metasurfaces, let us consider a medium exposed to an electric field $\boldsymbol{E}$ (V/m) and magnetic field $\boldsymbol{H}$ (A/m).[1] Consequently, there will be an induced electric polarization density $\boldsymbol{P}$ (C/m^2) and magnetic polarization density $\boldsymbol{M}$ (A/m). The overall induced electric and magnetic flux densities $\boldsymbol{D}$ (C/m^2) and $\boldsymbol{B}$ (Wb/m^2) are then expressed as

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}, \tag{2.1a}$$

$$\boldsymbol{B} = \mu_0 (\boldsymbol{H} + \boldsymbol{M}), \tag{2.1b}$$

where ε_0 (F/m) and μ_0 (H/m) are the permittivity and permeability in vacuum respectively. For linear isotropic media, we could also express these two equations

[1] The units of these quantities are expressed within the parentheses.

using susceptibilities by replacing $\boldsymbol{P} = \varepsilon_0 \chi_{\text{ee}} \boldsymbol{E}$ and $\boldsymbol{M} = \chi_{\text{mm}} \boldsymbol{H}$. We then get

$$\boldsymbol{D} = \varepsilon_0 (1 + \chi_{\text{ee}}) \boldsymbol{E}, \tag{2.2a}$$

$$\boldsymbol{B} = \mu_0 (1 + \chi_{\text{mm}}) \boldsymbol{H}. \tag{2.2b}$$

The subscripts in χ_{ee} and χ_{mm} indicate the electric-to-electric and magnetic-to-magnetic responses. Therefore, they may be referred to as electric-to-electric and magnetic-to-magnetic susceptibilities. The unitless relative permittivity and permeability are then defined as $\varepsilon_r = 1 + \chi_{\text{ee}}$ and $\mu_r = 1 + \chi_{\text{mm}}$ respectively. It should be noted that at this point, the susceptibilities are scalars, indicating that the material dielectric and magnetic properties are isotropic. The linear nature of the medium is indicated by the fact that the polarization densities are directly proportional to the field quantities.

For anisotropic media where the material properties are tensors, we can express susceptibilities $\bar{\bar{\chi}}$ as 3×3 complex tensors and write the polarization densities as

$$\boldsymbol{P} = \varepsilon_0 \bar{\bar{\chi}}_{\text{ee}} \cdot \boldsymbol{E}, \tag{2.3}$$

$$\boldsymbol{M} = \bar{\bar{\chi}}_{\text{mm}} \cdot \boldsymbol{H}, \tag{2.4}$$

where '$\cdot$' is the dot product, which in this case is the dot product between a dyad and a vector. As can be seen, for anisotropic media, the polarization densities and the field quantities are no longer parallel [132, pg. 146].

Using the above equations, we also have

$$\begin{bmatrix} D_u \\ D_v \\ D_n \end{bmatrix} = \varepsilon_0 \left(\bar{\bar{I}} + \begin{bmatrix} \chi_{\text{ee}}^{uu} & \chi_{\text{ee}}^{uv} & \chi_{\text{ee}}^{un} \\ \chi_{\text{ee}}^{vu} & \chi_{\text{ee}}^{vv} & \chi_{\text{ee}}^{vn} \\ \chi_{\text{ee}}^{nu} & \chi_{\text{ee}}^{nv} & \chi_{\text{ee}}^{nn} \end{bmatrix} \right) \begin{bmatrix} E_u \\ E_v \\ E_n \end{bmatrix}, \tag{2.5a}$$

$$\begin{bmatrix} B_u \\ B_v \\ B_n \end{bmatrix} = \mu_0 \left(\bar{\bar{I}} + \begin{bmatrix} \chi_{\text{mm}}^{uu} & \chi_{\text{mm}}^{uv} & \chi_{\text{mm}}^{un} \\ \chi_{\text{mm}}^{vu} & \chi_{\text{mm}}^{vv} & \chi_{\text{mm}}^{vn} \\ \chi_{\text{mm}}^{nu} & \chi_{\text{mm}}^{nv} & \chi_{\text{mm}}^{nn} \end{bmatrix} \right) \begin{bmatrix} H_u \\ H_v \\ H_n \end{bmatrix}, \tag{2.5b}$$

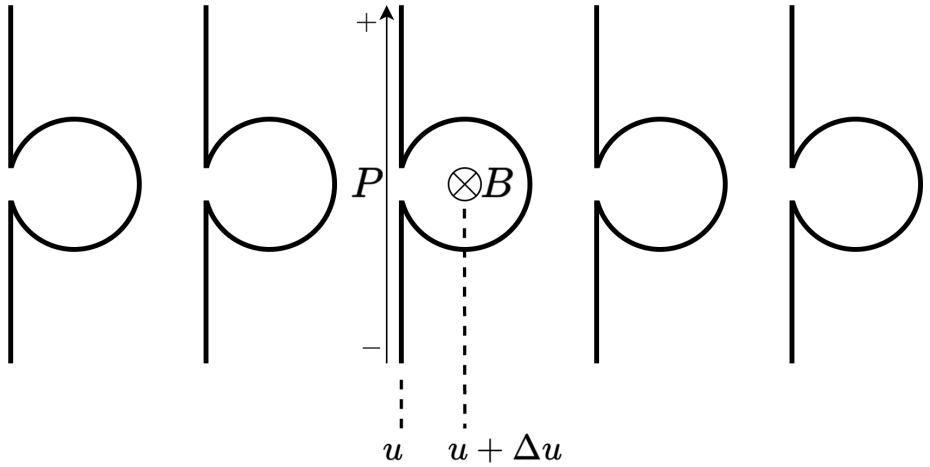


Fig. 2.1 The Omega meta-atom. The electric polarization density vector $\boldsymbol{P}$ at location u also depends on the magnetic field at $u + \Delta u$.

where $\overline{\overline{I}}$ is the unity dyadic tensor (i.e., identity matrix), and uvn represents a generic Cartesian coordinate system where $\hat{u} \times \hat{v} = \hat{n}$. (For example, we may have $\hat{u} = \hat{x}$, $\hat{v} = \hat{y}$, and $\hat{n} = \hat{z}$.)

The above discussion considers the medium to be spatially local. That is, the polarization density at a given location will only depend on the field quantity at the same location. With the advent of metamaterials, researchers have created media which are spatially dispersive [133]. That is, the polarization density at a given location can also depend on the neighboring field quantities. To understand this, let us consider a well-known artificial material made out of the so-called omega-shaped conducting meta-atoms [134] as shown in Fig. 2.1. In addition to an electric field at the location u, a (time-varying) magnetic field at the loop, say at location $u + \Delta u$, can induce currents in the loop according to Faraday's law, leading to an electric polarization density at the location u. Conversely, the electric field at u can create a current circulation in the loop, leading to magnetic polarization density at $u + \Delta u$. Such media can be represented as *bi-anisotropic* media where the polarization densities depend on both $\boldsymbol{E}$ and $\boldsymbol{H}$.

For bi-anisotropic media, the polarization densities can be expressed as

$$\boldsymbol{P} = \varepsilon_0 \overline{\overline{\chi}}_{\mathrm{ee}} \cdot \boldsymbol{E} + \frac{1}{c_0} \overline{\overline{\chi}}_{\mathrm{em}} \cdot \boldsymbol{H}, \tag{2.6}$$

$$\boldsymbol{M} = \overline{\overline{\chi}}_{\mathrm{mm}} \cdot \boldsymbol{H} + \frac{1}{\eta_0} \overline{\overline{\chi}}_{\mathrm{me}} \cdot \boldsymbol{E}, \tag{2.7}$$

where $\overline{\overline{\chi}}_{\mathrm{em}}$ and $\overline{\overline{\chi}}_{\mathrm{me}}$ denote magnetic-to-electric and electric-to-magnetic susceptibilities (unitless). In addition, c_0 (m/s) is the velocity of light in the vacuum, and η_0 (Ω) is the characteristic impedance of the vacuum. Thus, the electric and magnetic flux densities will become

$$\boldsymbol{D} = \varepsilon_0 (\overline{\overline{I}} + \overline{\overline{\chi}}_{\mathrm{ee}}) \cdot \boldsymbol{E} + \frac{1}{c_0} \overline{\overline{\chi}}_{\mathrm{em}} \cdot \boldsymbol{H}, \tag{2.8}$$

$$\boldsymbol{B} = \mu_0 (\overline{\overline{I}} + \overline{\overline{\chi}}_{\mathrm{mm}}) \cdot \boldsymbol{H} + \frac{1}{c_0} \overline{\overline{\chi}}_{\mathrm{me}} \cdot \boldsymbol{E}. \tag{2.9}$$

Up to this point, we have primarily modeled material properties within a volumetric framework. However, to accurately represent metasurfaces, it is important to have a suitable model for them. The metasurface is electrically thin[2], implying that it can be modelled as a zero-thickness sheet. To this end, the volumetric (3D) susceptibilities discussed above may be converted to surface (superscript 's') susceptibilities by employing the Dirac delta function denoted by '$\delta(\cdot)$'. That is, if we assume that the thickness of the metasurface is along the n direction, we may represent the 3D susceptibilities on the metasurface as [131, 135, 136]

$$\overline{\overline{\chi}}_{\mathrm{ee}} = \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{ee}} \delta(n), \tag{2.10a}$$

$$\overline{\overline{\chi}}_{\mathrm{mm}} = \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{mm}} \delta(n), \tag{2.10b}$$

$$\overline{\overline{\chi}}_{\mathrm{em}} = \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{em}} \delta(n), \tag{2.10c}$$

$$\overline{\overline{\chi}}_{\mathrm{me}} = \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{me}} \delta(n). \tag{2.10d}$$

[2]That is, $(d/\lambda_0 \to 0)$, where d is the thickness of the metasurface and λ_0 is the free space wavelength at the operating frequency.

2.1.1 Generalized Sheet Transition Conditions (GSTCs)

In this section, we present a derivations for the boundary conditions of the metasurface to be used in this thesis, which are known as generalized sheet transition conditions (GSTCs). This derivation follows the work presented in [131, 135, 136].

In order to design the surface susceptibilities of the metasurface based on the aperture fields on both faces of the metasurface, we must establish the boundary conditions that connect the aperture fields on both faces of the metasurface with its surface susceptibilities. For metasurfaces, these boundary conditions are commonly known as GSTCs. Let us again start from the frequency-domain Faraday's Law of Maxwell equations (assuming $\exp(j\omega t)$ time dependency[3]),

$$\nabla \times \boldsymbol{E} = -j\omega \boldsymbol{B}, \tag{2.11a}$$

$$\nabla \times \boldsymbol{E} = -j\omega\mu_0(\boldsymbol{H} + \boldsymbol{M}), \tag{2.11b}$$

where '$\nabla\times$' is the curl operator.

Herein, we assume that the metasurface lies on the uv plane, with the normal component to the metasurface being $+\hat{n}$, as depicted in Fig. 2.2. The tangential fields on the top face of the metasurface are denoted by ψ^+ where $\psi = \{E, H\}$. Similarly, the tangential fields on the bottom face of the metasurface are denoted by ψ^-. Since these ψ's deal with tangential fields, they can be either ψ_u or ψ_v since the metasurface lies on the uv plane. On other hand, the normal component of the fields is denoted by ψ_n^{left} and ψ_n^{right}. We further assume that we do not have impressed sources on the metasurface.

Considering the works presented in [131, 135, 136], we employ the conventional pillbox and apply the Stoke's theorem to (2.11b) over its front face. As can be seen in

[3]The angular frequency is ω; t denotes the time, and $j^2 = -1$.

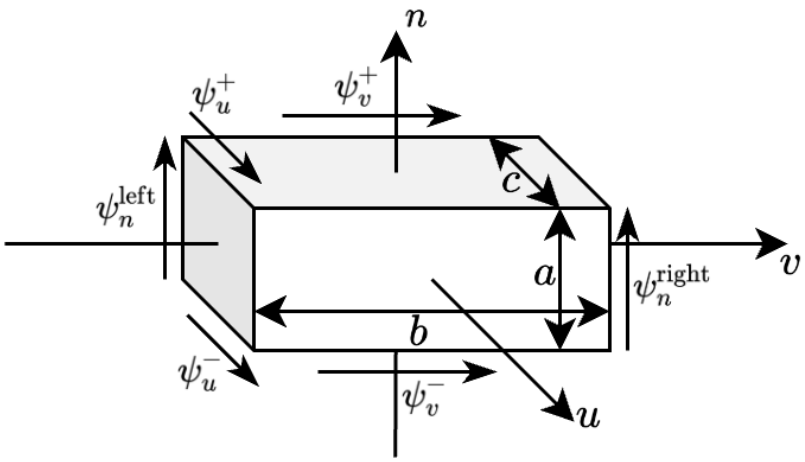


Fig. 2.2 The metasurface is assumed to lie on the uv plane having a zero thickness. The normal unit vector to the metasurface is $\hat{n}$. The tangential fields on the top and bottom faces of the metasurface are denoted by ψ^+ and ψ^- respectively where ψ is either the electric or magnetic field. A conventional pillbox is used to derived the boundary conditions.

Fig. 2.2, the front face is a rectangle parallel to the nv plane. This results in[4]

$$\int (\nabla \times \boldsymbol{E}) \cdot d\mathbf{s} = -j\omega\mu_0 \int (\boldsymbol{H} + \boldsymbol{M}) \cdot d\mathbf{s}, \tag{2.12a}$$

$$\oint \boldsymbol{E} \cdot d\boldsymbol{\ell} = -j\omega\mu_0 \int \boldsymbol{H} \cdot d\mathbf{s} - j\omega\mu_0 \int \boldsymbol{M} \cdot d\mathbf{s}, \tag{2.12b}$$

where $d\mathbf{s}$ is the element of surface vector and $d\boldsymbol{\ell}$ is the element of length vector along the contour of the integration area. (The so-called right hand rule is used to ensure that the directions of these two differential vectors are consistent.)

[4]Note that the following derivation works based on the approximation that the normal components of the electric and magnetic polarization densities supported by the metasurface are negligible. If they are not negligible, we need to modify this derivation by considering these normal components. For that, please see [136, Eq. 40 and Eq. 48]. As described in [137], this is justified by noting that this thesis considers thin metasurfaces illuminated by a fixed transmitter, typically with near normal incidence.

By letting $a \to 0$, we can further simplify (2.12) as

$$-E_v^+ b - \cancelto{0}{E_n^{\text{left}} a} + E_v^- b + \cancelto{0}{E_n^{\text{right}} a} = -j\omega\mu_o \int \boldsymbol{H} \cdot \hat{u} dv dn - j\omega\mu_o \int \boldsymbol{M} \cdot \hat{u} dv dn, \quad (2.13a)$$

$$-E_v^+ b + E_v^- b = -j\omega\mu_0 b \int_{-a/2}^{a/2} H_u dn - j\omega\mu_0 b \int_{-a/2}^{a/2} M_u dn. \quad (2.13b)$$

As a tends to zero, we can operate under the assumption of locally quasi-static fields across the metasurface. As noted in [138], this assumption entails that the field ψ within the metasurface is assumed to satisfy $\partial^2\psi/\partial n^2 = 0$ and thus ψ at a given (u, v) within the metasurface depends on n as $\psi(n) = \tau n + \rho$ where τ and ρ are constant coefficients. They can be found by noting the tangential fields on both faces of the metasurface. Consequently, it can be shown that integrating H_u and M_u over dn yields the average (subscript 'av') of these quantities on both sides of the limit multiplied by a, that is $\int_{-a/2}^{a/2} \psi dn = a(\psi^+ + \psi^-)/2 = a\psi_{\text{av}}$. Thus, (2.13) becomes

$$-E_v^+ b + E_v^- b = -j\omega\mu_0 ab \frac{H_u^+ + H_u^-}{2} - j\omega\mu_0 ab \frac{M_u^+ + M_u^-}{2}, \quad (2.14a)$$

$$-E_v^+ + E_v^- = -j\omega\mu_0 a H_{u,\text{av}} - j\omega\mu_0 a M_{u,\text{av}}. \quad (2.14b)$$

Let us now substitute (2.10) into (2.7) to obtain:

$$\boldsymbol{M}_{\text{av}} = \overline{\overline{\chi}}_{\text{mm}}^{\text{s}} \delta(n) \cdot \boldsymbol{H}_{\text{av}} + \frac{1}{\eta_0} \overline{\overline{\chi}}_{\text{me}}^{\text{s}} \delta(n) \cdot \boldsymbol{E}_{\text{av}}. \quad (2.15)$$

Given that we assume that the normal polarization densities supported by the metasurface are negligible, we consequently ignore all the surface susceptibilities that include the normal component [131, Sec 4.1.1]. That is,

$$\overline{\overline{\chi}}_{\text{mm}}^{\text{s}} = \begin{bmatrix} \chi_{\text{mm}}^{uu} & \chi_{\text{mm}}^{uv} & 0 \\ \chi_{\text{mm}}^{vu} & \chi_{\text{mm}}^{vv} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \overline{\overline{\chi}}_{\text{me}}^{\text{s}} = \begin{bmatrix} \chi_{\text{me}}^{uu} & \chi_{\text{me}}^{uv} & 0 \\ \chi_{\text{me}}^{vu} & \chi_{\text{me}}^{vv} & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (2.16)$$

Therefore, from the two previous equations, we can write

$$M_{u,\mathrm{av}} = \hat{u} \cdot (\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{mm}} \delta(n) \cdot \boldsymbol{H}_{\mathrm{av}} + \frac{1}{\eta_0} \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{me}} \delta(n) \cdot \boldsymbol{E}_{\mathrm{av}}) \\ = \chi^{uu}_{\mathrm{mm}} H_{u,\mathrm{av}} \delta(n) + \chi^{uv}_{\mathrm{mm}} H_{v,\mathrm{av}} \delta(n) + \frac{1}{\eta_0} \chi^{uu}_{\mathrm{me}} E_{u,\mathrm{av}} \delta(n) + \frac{1}{\eta_0} \chi^{uv}_{\mathrm{me}} E_{v,\mathrm{av}} \delta(n). \quad (2.17)$$

Substituting $M_{u,\mathrm{av}}$ into (2.14b) yields,

$$-E_v^+ + E_v^- = -j\omega\mu_0 a H_{u,\mathrm{av}} - j\omega\mu_0 a \hat{u} \cdot (\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{mm}} \delta(n) \cdot \boldsymbol{H}_{\mathrm{av}} + \frac{1}{\eta_0} \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{me}} \delta(n) \cdot \boldsymbol{E}_{\mathrm{av}}). \quad (2.18)$$

Since we model the metasurface as a sheet with zero thickness, we let 'a' approach zero. Given that the metasurface is situated at $n = 0$, the Dirac delta function can be expressed as $\delta(n)|_{n\to 0} = 1/a|_{a\to 0}$. Consequently, as '$a$' tends to zero, $\lim_{a\to 0} a\delta(n)|_{n\to 0} = a/a = 1$ [135]. Substituting this into the preceding equation, we obtain:

$$-E_v^+ + E_v^- = -j\omega\mu_0 \hat{u} \cdot (\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{mm}} \cdot \boldsymbol{H}_{\mathrm{av}} + \frac{1}{\eta_0} \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{me}} \cdot \boldsymbol{E}_{\mathrm{av}}). \quad (2.19)$$

The above boundary equation describes the discontinuity of the $\hat{v}$ component of $\boldsymbol{E}$ across the metasurface.

To obtain a similar boundary equation for the other tangential $\boldsymbol{E}$ component, i.e., $\hat{u}$, we need to integrate over the left face of the pillbox shown in Fig. 2.2, in a similar manner, yielding:

$$E_u^+ c - E_n^{\mathrm{right}} a - E_u^- c + E_n^{\mathrm{left}} a = -j\omega\mu_0 a c H_{v,\mathrm{av}} \\ - j\omega\mu_0 a c \hat{v} \cdot (\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{mm}} \delta(n) \cdot \boldsymbol{H}_{\mathrm{av}} + \frac{1}{\eta_0} \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{me}} \delta(n) \cdot \boldsymbol{E}_{\mathrm{av}}). \quad (2.20)$$

Letting a approach zero and dividing by c, the above equation becomes:

$$E_u^+ - E_u^- = -j\omega\mu_0 \hat{v} \cdot (\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{mm}} \cdot \boldsymbol{H}_{\mathrm{av}} + \frac{1}{\eta_0} \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{me}} \cdot \boldsymbol{E}_{\mathrm{av}}). \quad (2.21)$$

We now use the Δ notation to indicate the field differences between the two faces of the metasurface. That is,

$$\Delta\psi = \psi^{+} - \psi^{-}. \tag{2.22}$$

Then, combining (2.19) and (2.21), we have

$$\begin{bmatrix} \Delta E_v \\ -\Delta E_u \end{bmatrix} = j\omega\mu_0 \begin{bmatrix} \chi_{\mathrm{mm}}^{uu} & \chi_{\mathrm{mm}}^{uv} \\ \chi_{\mathrm{mm}}^{vu} & \chi_{\mathrm{mm}}^{vv} \end{bmatrix} \begin{bmatrix} H_{u,\mathrm{av}} \\ H_{v,\mathrm{av}} \end{bmatrix} + j\omega\frac{\mu_0}{\eta_0} \begin{bmatrix} \chi_{\mathrm{me}}^{uu} & \chi_{\mathrm{me}}^{uv} \\ \chi_{\mathrm{me}}^{vu} & \chi_{\mathrm{me}}^{vv} \end{bmatrix} \begin{bmatrix} E_{u,\mathrm{av}} \\ E_{v,\mathrm{av}} \end{bmatrix}. \tag{2.23}$$

Up to this point, we have established the first half of the GSTCs for the tangential fields. Now, employing a comparable approach, we can derive the second half of the GSTCs for the tangential magnetic fields starting from the Maxwell-Ampère's equation within a source-free medium. Thus, we begin with[5]

$$\nabla \times \boldsymbol{H} = j\omega\boldsymbol{D}, \tag{2.24a}$$

$$\nabla \times \boldsymbol{H} = j\omega(\varepsilon_0\boldsymbol{E} + \boldsymbol{P}). \tag{2.24b}$$

Similar to the previous cases, integrating around the front face and the left face of the pillbox and applying the Stoke's theorem, we obtain:

$$-H_v^{+} + H_v^{-} = j\omega\hat{u} \cdot (\varepsilon_0\overline{\overline{\chi}}_{\mathrm{ee}}^{\mathrm{s}} \cdot \boldsymbol{E}_{\mathrm{av}} + \frac{1}{c_0}\overline{\overline{\chi}}_{\mathrm{em}}^{\mathrm{s}} \cdot \boldsymbol{H}_{\mathrm{av}}), \tag{2.25a}$$

$$H_u^{+} - H_u^{-} = j\omega\hat{v} \cdot (\varepsilon_0\overline{\overline{\chi}}_{\mathrm{ee}}^{\mathrm{s}} \cdot \boldsymbol{E}_{\mathrm{av}} + \frac{1}{c_0}\overline{\overline{\chi}}_{\mathrm{em}}^{\mathrm{s}} \cdot \boldsymbol{H}_{\mathrm{av}}). \tag{2.25b}$$

[5]Similar to the previous case noted earlier, we assume that the normal polarization densities supported by the metasurface is negligible.

By combining the above equations along with the first portion of the GSTCs, and subsequently substituting $\mu_0/\eta_0 = 1/c_0 = \sqrt{\varepsilon_0\mu_0}$, the GSTCs will be

$$\begin{bmatrix} \Delta E_v \\ -\Delta E_u \end{bmatrix} = j\omega\mu_0 \begin{bmatrix} \chi_{\mathrm{mm}}^{uu} & \chi_{\mathrm{mm}}^{uv} \\ \chi_{\mathrm{mm}}^{vu} & \chi_{\mathrm{mm}}^{vv} \end{bmatrix} \begin{bmatrix} H_{u,\mathrm{av}} \\ H_{v,\mathrm{av}} \end{bmatrix} + j\omega\sqrt{\varepsilon_0\mu_0} \begin{bmatrix} \chi_{\mathrm{me}}^{uu} & \chi_{\mathrm{me}}^{uv} \\ \chi_{\mathrm{me}}^{vu} & \chi_{\mathrm{me}}^{vv} \end{bmatrix} \begin{bmatrix} E_{u,\mathrm{av}} \\ E_{v,\mathrm{av}} \end{bmatrix}, \tag{2.26a}$$

$$\begin{bmatrix} -\Delta H_v \\ \Delta H_u \end{bmatrix} = j\omega\varepsilon_0 \begin{bmatrix} \chi_{\mathrm{ee}}^{uu} & \chi_{\mathrm{ee}}^{uv} \\ \chi_{\mathrm{ee}}^{vu} & \chi_{\mathrm{ee}}^{vv} \end{bmatrix} \begin{bmatrix} E_{u,\mathrm{av}} \\ E_{v,\mathrm{av}} \end{bmatrix} + j\omega\sqrt{\varepsilon_0\mu_0} \begin{bmatrix} \chi_{\mathrm{em}}^{uu} & \chi_{\mathrm{em}}^{uv} \\ \chi_{\mathrm{em}}^{vu} & \chi_{\mathrm{em}}^{vv} \end{bmatrix} \begin{bmatrix} H_{u,\mathrm{av}} \\ H_{v,\mathrm{av}} \end{bmatrix}. \tag{2.26b}$$

In summary, (2.26) represents the GSTCs boundary conditions that describe how the metasurface supports the discontinuity of the the tangential electric and magnetic fields across its interface through its surface susceptibilities. In other words, noting Fig 2.3, The GSTCs in (2.26) shows how a metasurface, when modeled as a zero-thickness sheet, can transform the fields on its input face to the fields on its output face through its surface susceptibilites χ.

2.1.2 Constraints on Surface Susceptibilities

As can be seen, we have represented the metasurface with 16 surface susceptibility components, which can be complex numbers. Since (2.26) represents four equations, we can think of it as four equations and sixteen susceptibility unknowns, and we can therefore support four independent transformations with one metasurface [21]. However, in practice, the surface susceptibilities might not offer these many degrees of freedom. Herein, we discuss this mainly based on [131, 21].

In this thesis, we require the metasurface to be reciprocal. The metasurface exhibits reciprocity when the source and measurement device can be interchanged without affecting the received signal. Nonreciprocal devices can be realized through various methods, including biasing a ferromagnetic medium with an external static magnetic field or introducing temporal modulation [139]. However, implementing these

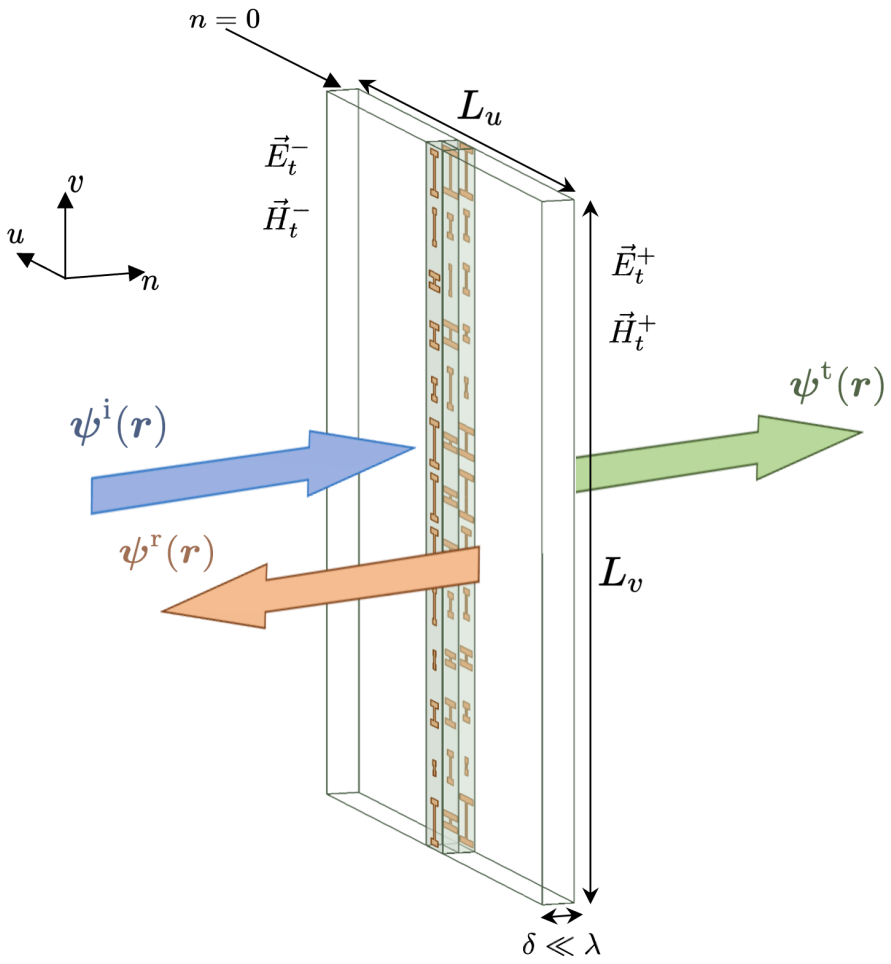


Fig. 2.3 A sample metasurface with a size of $L_u \times L_v$ and a subwavelength thickness ($\delta \ll \lambda$ where λ denotes the wavelength). This metasurface consists of three-layered dogbone copper traces printed on a dielectric substrate. For clarity, this figure only shows one column of such dogbone copper traces. Assuming a zero-thickness model (surface model), the metasurface is located on the uv plane at $n = 0$. The metasurface will then have two sides: $n = 0^+$ (output) and $n = 0^-$ (input). The metasurface is designed to transform a specific incident wave $\boldsymbol{\psi}^{\mathrm{i}}(\boldsymbol{r})$ where $\boldsymbol{\psi}$ represents the electric $\boldsymbol{E}$ and magnetic $\boldsymbol{H}$ fields into desired reflected wave $\boldsymbol{\psi}^{\mathrm{r}}(\boldsymbol{r})$ and transmitted wave $\boldsymbol{\psi}^{\mathrm{t}}(\boldsymbol{r})$. The tangential fields on the input side are denoted as $\vec{E}_t^-$ and $\vec{H}_t^-$. The tangential fields on the output side are denoted as $\vec{E}_t^+$ and $\vec{H}_t^+$. Since the metasurface lies on the uv plane, the tangential components to the metasurface will have u and v components.

techniques to break reciprocity can be challenging. In summary, ensuring reciprocity in our metasurface design leads to a more practically achievable structure, which needs

to meet the following constraints [140, 21, 139]:

$$\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{ee}} = (\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{ee}})^{\mathrm{T}} \quad , \quad \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{mm}} = (\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{mm}})^{\mathrm{T}} \quad , \quad \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{em}} = -(\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{me}})^{\mathrm{T}}, \tag{2.27}$$

where the transpose operator is indicated by the superscript 'T'. Consequently the degrees of freedom in the design process is reduced, with the following constraints:

$$\begin{gathered} \chi^{uv}_{\mathrm{ee}} = \chi^{vu}_{\mathrm{ee}} \quad , \quad \chi^{uv}_{\mathrm{mm}} = \chi^{vu}_{\mathrm{mm}}, \\ \chi^{uu}_{\mathrm{me}} = -\chi^{uu}_{\mathrm{em}}, \quad \chi^{vv}_{\mathrm{me}} = -\chi^{vv}_{\mathrm{em}}, \quad \chi^{uv}_{\mathrm{me}} = -\chi^{vu}_{\mathrm{em}}, \quad \chi^{uv}_{\mathrm{me}} = -\chi^{vu}_{\mathrm{em}}. \end{gathered} \tag{2.28}$$

Furthermore, it is generally desirable for the metasurface to be both lossless and passive. To establish the required condition, we can begin with the time-averaged Poynting theorem for a bianisotropic medium [37]. Using this approach, the time-average Poynting vector is written in terms of bianisotropic susceptibilities. Then, for a medium to be lossless and passive, the divergence of the time-average Poynting vector needs to be zero. This results in the following constraints for surface susceptibilities [131, Sec 2.5]

$$\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{ee}} = (\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{ee}})^{\mathrm{H}} \quad , \quad \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{mm}} = (\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{mm}})^{\mathrm{H}} \quad , \quad \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{em}} = (\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{me}})^{\mathrm{H}}. \tag{2.29}$$

where the superscript 'H' denotes the Hermitian (complex conjugate transpose) operator.

Combining (2.29) with the reciprocity conditions outlined in (2.27) yields the following relationships for a metasurface that is lossless, passive, and reciprocal:

$$\begin{gathered} \mathrm{Im}(\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{ee}}) = 0, \quad \mathrm{Im}(\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{mm}}) = 0, \quad \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{ee}} = (\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{ee}})^{\mathrm{T}}, \quad \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{mm}} = (\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{mm}})^{\mathrm{T}} \\ \mathrm{Re}(\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{em}}) = 0, \quad \mathrm{Re}(\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{me}}) = 0, \quad \overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{em}} = -(\overline{\overline{\chi}}^{\mathrm{s}}_{\mathrm{me}})^{\mathrm{T}}. \end{gathered} \tag{2.30}$$

where 'Re' and 'Im' denotes the real-part and imaginary-part operators.

We further limit ourselves to non-gyrotropic metasurfaces to avoid polarization alteration [18]. This results in [18]

$$\chi_{\mathrm{ee}}^{uv} = \chi_{\mathrm{ee}}^{vu} = 0, \quad \chi_{\mathrm{mm}}^{uv} = \chi_{\mathrm{mm}}^{vu} = 0, \quad \chi_{\mathrm{em}}^{uu} = \chi_{\mathrm{em}}^{vv} = \chi_{\mathrm{me}}^{uu} = \chi_{\mathrm{me}}^{vv} = 0. \tag{2.31}$$

Finally, as will be seen in the thesis, we consider transverse electric (TE) and transverse magnetic (TM) scenarios. For example, for the TE case with respect to the vn plane, we will have three field components, namely E_u, H_v, and H_n. Consequently, for this example, based on the constraints noted earlier and assuming the metasurface being on the uv plane, we deal with $\chi_{\mathrm{mm}}^{vv} \in \mathbb{R}$, $\chi_{\mathrm{me}}^{vu} \in \mathbb{I}$, $\chi_{\mathrm{ee}}^{uu} \in \mathbb{R}$. (Herein, $\mathbb{R}$ and $\mathbb{I}$ denote purely real and purely imaginary numbers respectively.)

2.1.3 Two-port Network Model

Based on the preceding discussion, it is evident that metasurfaces behave like "wave transformers" capable of converting one set of tangential fields into another. Consequently, we may model the metasurface's unit cells using two-port networks. This modeling approach simplifies its complex electromagnetic behavior to circuit theory principles, enabling efficient design, evaluation and optimization of the metasurfaces. Different two-port network parameters can be used to model metasurfaces such as scattering parameters (S-parameters) [141], impedance parameters (Z-parameters) [29], transmission parameters (ABCD parameters) [45]. Herein, as an example, we use the scattering parameters to model the unit cells and derive the expressions for scattering matrix [**S**] from surface susceptibilities $\overline{\overline{\chi}}$. We start from the expression of scattering parameters:

$$\begin{bmatrix} E_1^{\mathrm{out}} \\ E_2^{\mathrm{out}} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} E_1^{\mathrm{in}} \\ E_2^{\mathrm{in}} \end{bmatrix} \tag{2.32}$$

where the subscripts 1 and 2 indicate the port numbers (which in our case represent the two faces of the metasurface per unit cell) and the superscripts indicate the incoming and outgoing tangential electric fields at its corresponding port respectively. Let us,

for example, assign Port 1 to the side of the metasurface on which the incident source field impinges and Port 2 as the output aperture of the metasurface.

Let us also assume that the metasurface lies on the uv plane, and consider the TE assumption with respect to the un propagation plane with $\partial/\partial v = 0$. Thus, the electric field will have $\hat{v}$ polarization. In addition, for this example, Port 1 and Port 2 correspond to $n = 0^-$ and $n = 0^+$ faces, respectively. Assuming the incident electric field (plane wave) takes the form of $\boldsymbol{E}^i = e^{-jk_0 n}\hat{v}$, we can then describe the fields at the metasurface interface using the following relations:[6]

$$\boldsymbol{E}^i(n = 0^-) = \hat{v}, \tag{2.33a}$$

$$\boldsymbol{E}^r(n = 0^-) = S_{11}\hat{v}, \tag{2.33b}$$

$$\boldsymbol{E}^t(n = 0^+) = S_{21}\hat{v}, \tag{2.33c}$$

where the superscripts i, r, and t denotes the incident, reflected, and transmitted fields respectively. Thus, the electric field average $\boldsymbol{E}_{\text{av}}$ and difference $\Delta\boldsymbol{E}$ can be expressed as,

$$\boldsymbol{E}_{\text{av}} = \frac{(\boldsymbol{E}^i + \boldsymbol{E}^r) + \boldsymbol{E}^t}{2} = \frac{1 + S_{11} + S_{21}}{2}\hat{v}, \tag{2.34a}$$

$$\Delta\boldsymbol{E} = \boldsymbol{E}^t - (\boldsymbol{E}^i + \boldsymbol{E}^r) = (S_{21} - S_{11} - 1)\hat{v}. \tag{2.34b}$$

Assuming the metasurface is operating in free space, and assuming plane wave propagation in the $\hat{n}$ direction, we can similarly derive the associated incident, reflected,

[6]Herein, for the derivation of the scattering parameters with respect to surface susceptibilities, we follow [131]; however, we limit ourselves to non-gyrotropic metasurfaces. Consequently, for each scattering parameter, say S_{11}, we only consider one polarization component. For the general case, the cross-polarization effect also needs to be considered. In other words, our S_{11} is in face S_{11}^{vv}. However, in the general case, it can also have S_{11}^{uv} component.

and transmitted magnetic fields as

$$\boldsymbol{H}^i(n=0^-) = \hat{n} \times \frac{\boldsymbol{E^i}}{\eta_0} = \hat{n} \times \frac{\hat{v}}{\eta_0} = -\frac{1}{\eta_0}\hat{u}, \tag{2.35a}$$

$$\boldsymbol{H}^r(n=0^-) = -\hat{n} \times \frac{\boldsymbol{E^r}}{\eta_0} = -\hat{n} \times \frac{S_{11}\hat{v}}{\eta_0} = \frac{S_{11}}{\eta_0}\hat{u}, \tag{2.35b}$$

$$\boldsymbol{H}^t(n=0^+) = \hat{n} \times \frac{\boldsymbol{E^t}}{\eta_0} = \hat{n} \times \frac{S_{21}\hat{v}}{\eta_0} = -\frac{S_{21}}{\eta_0}\hat{u}. \tag{2.35c}$$

Subsequently, the magnetic field average $\boldsymbol{H}_{\text{av}}$ and difference $\Delta\boldsymbol{H}$ can be expressed as,

$$\boldsymbol{H}_{\text{av}} = \frac{(\boldsymbol{H}^i + \boldsymbol{H}^r) + \boldsymbol{H}^t}{2} = \frac{1}{2\eta_0}(-1 + S_{11} - S_{21})\hat{u}, \tag{2.36a}$$

$$\Delta\boldsymbol{H} = \boldsymbol{H}^t - (\boldsymbol{H}^i + \boldsymbol{H}^r) = \frac{1}{\eta_0}(-S_{21} + 1 - S_{11})\hat{u}. \tag{2.36b}$$

For simplicity of the notation, the following definitions (with *hat*) may be used

$$\hat{\chi}_{\text{ee}} \triangleq j\omega\varepsilon_0\chi_{\text{ee}}^{vv}, \tag{2.37a}$$

$$\hat{\chi}_{\text{em}} \triangleq j\omega\sqrt{\varepsilon_0\mu_0}\chi_{\text{em}}^{vu}, \tag{2.37b}$$

$$\hat{\chi}_{\text{mm}} \triangleq j\omega\mu_0\chi_{\text{mm}}^{uu}, \tag{2.37c}$$

$$\hat{\chi}_{\text{me}} \triangleq j\omega\sqrt{\varepsilon_0\mu_0}\chi_{\text{me}}^{uv}. \tag{2.37d}$$

From (2.26), if we only considers H_u and E_v we will have,

$$\Delta H_u = \hat{\chi}_{\text{ee}}E_{v,\text{av}} + \hat{\chi}_{\text{em}}H_{u,\text{av}}, \tag{2.38a}$$

$$\Delta E_v = \hat{\chi}_{\text{mm}}H_{u,\text{av}} + \hat{\chi}_{\text{me}}E_{v,\text{av}}. \tag{2.38b}$$

Substituting the (2.34) and (2.36) into the above equation we will have,

$$\frac{1}{\eta_0}(-S_{21} + 1 - S_{11}) = \hat{\chi}_{\text{ee}}\frac{1 + S_{11} + S_{21}}{2} + \hat{\chi}_{\text{em}}\frac{1}{2\eta_0}(-1 + S_{11} - S_{21}), \tag{2.39a}$$

$$(S_{21} - S_{11} - 1) = \hat{\chi}_{\text{mm}}\frac{1}{2\eta_0}(-1 + S_{11} - S_{21}) + \hat{\chi}_{\text{me}}\frac{1 + S_{11} + S_{21}}{2}. \tag{2.39b}$$

Then, re-arranging the above equation, we will have,

$$\left(\frac{\hat{\chi}_{\mathrm{ee}}}{2}+\frac{1}{\eta_0}+\frac{\hat{\chi}_{\mathrm{em}}}{2\eta_0}\right)S_{11}+\left(\frac{\hat{\chi}_{\mathrm{ee}}}{2}+\frac{1}{\eta_0}-\frac{\hat{\chi}_{\mathrm{em}}}{2\eta_0}\right)S_{21}=-\frac{\hat{\chi}_{\mathrm{ee}}}{2}+\frac{1}{\eta_0}+\frac{\hat{\chi}_{\mathrm{em}}}{2\eta_0}, \tag{2.40a}$$

$$\left(\frac{\hat{\chi}_{\mathrm{mm}}}{2\eta_0}+\frac{\hat{\chi}_{\mathrm{me}}}{2}+1\right)S_{11}+\left(-\frac{\hat{\chi}_{\mathrm{mm}}}{2\eta_0}+\frac{\hat{\chi}_{\mathrm{me}}}{2}-1\right)S_{21}=\frac{\hat{\chi}_{\mathrm{mm}}}{2\eta_0}-\frac{\hat{\chi}_{\mathrm{me}}}{2}-1. \tag{2.40b}$$

Applying Cramer's rule to the above equation, we have,

$$S_{11}=\frac{(-\eta_0\hat{\chi}_{\mathrm{ee}}+\hat{\chi}_{\mathrm{em}}+2)(-\hat{\chi}_{\mathrm{mm}}+\eta_0\hat{\chi}_{\mathrm{me}}-2\eta_0)-(\hat{\chi}_{\mathrm{mm}}-\eta_0\hat{\chi}_{\mathrm{me}}-2\eta_0)(\eta_0\hat{\chi}_{\mathrm{ee}}-\hat{\chi}_{\mathrm{em}}+2)}{(\eta_0\hat{\chi}_{\mathrm{ee}}+\hat{\chi}_{\mathrm{em}}+2)(-\hat{\chi}_{\mathrm{mm}}+\eta_0\hat{\chi}_{\mathrm{me}}-2\eta_0)-(\hat{\chi}_{\mathrm{mm}}+\eta_0\hat{\chi}_{\mathrm{me}}+2\eta_0)(\eta_0\hat{\chi}_{\mathrm{ee}}-\hat{\chi}_{\mathrm{em}}+2)}, \tag{2.41a}$$

$$S_{21}=\frac{(\eta_0\hat{\chi}_{\mathrm{ee}}+\hat{\chi}_{\mathrm{em}}+2)(\hat{\chi}_{\mathrm{mm}}-\eta_0\hat{\chi}_{\mathrm{me}}-2\eta_0)-(\hat{\chi}_{\mathrm{mm}}+\eta_0\hat{\chi}_{\mathrm{me}}+2\eta_0)(-\eta_0\hat{\chi}_{\mathrm{ee}}+\hat{\chi}_{\mathrm{em}}+2)}{(\eta_0\hat{\chi}_{\mathrm{ee}}+\hat{\chi}_{\mathrm{em}}+2)(-\hat{\chi}_{\mathrm{mm}}+\eta_0\hat{\chi}_{\mathrm{me}}-2\eta_0)-(\hat{\chi}_{\mathrm{mm}}+\eta_0\hat{\chi}_{\mathrm{me}}+2\eta_0)(\eta_0\hat{\chi}_{\mathrm{ee}}-\hat{\chi}_{\mathrm{em}}+2)}. \tag{2.41b}$$

The above equations enable us to find two scattering parameters from the knowledge of surface susceptibilities.

To also find the expressions for S_{12} and S_{22}, we similarly assume an incident electric field (normal plane wave) takes the form of $\boldsymbol{E}^i=e^{+jk_0n}\hat{v}$ that illuminating the metasurface from $n>0$ (i.e. Port 2). The corresponding incident, reflected, and transmitted fields can be expressed as:

$$\boldsymbol{E}^i(n=0^+)=\hat{v}, \tag{2.42a}$$

$$\boldsymbol{E}^r(n=0^+)=S_{22}\hat{v}, \tag{2.42b}$$

$$\boldsymbol{E}^t(n=0^-)=S_{12}\hat{v}, \tag{2.42c}$$

$$\boldsymbol{H}^i(n=0^+)=-\hat{n}\times\frac{\boldsymbol{E^i}}{\eta_0}=-\hat{n}\times\frac{\hat{v}}{\eta_0}=\frac{1}{\eta_0}\hat{u}, \tag{2.42d}$$

$$\boldsymbol{H}^r(n=0^+)=\hat{n}\times\frac{\boldsymbol{E^r}}{\eta_0}=\hat{n}\times\frac{S_{22}\hat{v}}{\eta_0}=-\frac{S_{22}}{\eta_0}\hat{u}, \tag{2.42e}$$

$$\boldsymbol{H}^t(n=0^-)=-\hat{n}\times\frac{\boldsymbol{E^t}}{\eta_0}=-\hat{n}\times\frac{S_{12}\hat{v}}{\eta_0}=\frac{S_{12}}{\eta_0}\hat{u}. \tag{2.42f}$$

Thus, the field average and difference can be expressed as

$$\boldsymbol{E}_{\text{av}} = \frac{(\boldsymbol{E}^i + \boldsymbol{E}^r) + \boldsymbol{E}^t}{2} = \frac{1 + S_{22} + S_{12}}{2}\hat{v}, \tag{2.43a}$$

$$\boldsymbol{H}_{\text{av}} = \frac{(\boldsymbol{H}^i + \boldsymbol{H}^r) + \boldsymbol{H}^t}{2} = \frac{1}{2\eta_0}(1 - S_{22} + S_{12})\hat{u}, \tag{2.43b}$$

$$\Delta\boldsymbol{E} = (\boldsymbol{E}^i + \boldsymbol{E}^r) - \boldsymbol{E}^t = (1 + S_{22} - S_{12})\hat{v}, \tag{2.43c}$$

$$\Delta\boldsymbol{H} = (\boldsymbol{H}^i + \boldsymbol{H}^r) - \boldsymbol{H}^t = \frac{1}{\eta_0}(1 - S_{22} - S_{12})\hat{u}. \tag{2.43d}$$

Substitute the above equations into (2.38) we have,

$$\left(-\frac{\hat{\chi}_{\text{mm}}}{2\eta_0} + \frac{\hat{\chi}_{\text{me}}}{2} - 1\right) S_{22} + \left(\frac{\hat{\chi}_{\text{mm}}}{2\eta_0} + \frac{\hat{\chi}_{\text{me}}}{2} + 1\right) S_{12} = \left(-\frac{\hat{\chi}_{\text{mm}}}{2\eta_0} - \frac{\hat{\chi}_{\text{me}}}{2} + 1\right), \tag{2.44a}$$

$$\left(\frac{\hat{\chi}_{\text{ee}}}{2} - \frac{\hat{\chi}_{\text{em}}}{2\eta_0} + \frac{1}{\eta_0}\right) S_{22} + \left(\frac{\hat{\chi}_{\text{ee}}}{2} + \frac{\hat{\chi}_{\text{em}}}{2\eta_0} + \frac{1}{\eta_0}\right) S_{12} = \left(-\frac{\hat{\chi}_{\text{ee}}}{2} - \frac{\hat{\chi}_{\text{em}}}{2\eta_0} + \frac{1}{\eta_0}\right). \tag{2.44b}$$

Applying Cramer's rule to the above equations we have,

$$S_{22} = \frac{(-\hat{\chi}_{\text{mm}} - \eta_0\hat{\chi}_{\text{me}} + 2\eta_0)(\eta_0\hat{\chi}_{\text{ee}} + \hat{\chi}_{\text{em}} + 2) - (\hat{\chi}_{\text{mm}} + \eta_0\hat{\chi}_{\text{me}} + 2\eta_0)(-\eta_0\hat{\chi}_{\text{ee}} - \hat{\chi}_{\text{em}} + 2)}{(-\hat{\chi}_{\text{mm}} + \eta_0\hat{\chi}_{\text{me}} - 2\eta_0)(\eta_0\hat{\chi}_{\text{ee}} + \hat{\chi}_{\text{em}} + 2) - (\hat{\chi}_{\text{mm}} + \eta_0\hat{\chi}_{\text{me}} + 2\eta_0)(\eta_0\hat{\chi}_{\text{ee}} - \hat{\chi}_{\text{em}} + 2)}, \tag{2.45a}$$

$$S_{12} = \frac{(-\hat{\chi}_{\text{mm}} + \eta_0\hat{\chi}_{\text{me}} - 2\eta_0)(-\eta_0\hat{\chi}_{\text{ee}} - \hat{\chi}_{\text{em}} + 2) - (-\hat{\chi}_{\text{mm}} - \eta_0\hat{\chi}_{\text{me}} + 2\eta_0)(\eta_0\hat{\chi}_{\text{ee}} - \hat{\chi}_{\text{em}} + 2)}{(-\hat{\chi}_{\text{mm}} + \eta_0\hat{\chi}_{\text{me}} - 2\eta_0)(\eta_0\hat{\chi}_{\text{ee}} + \hat{\chi}_{\text{em}} + 2) - (\hat{\chi}_{\text{mm}} + \eta_0\hat{\chi}_{\text{me}} + 2\eta_0)(\eta_0\hat{\chi}_{\text{ee}} - \hat{\chi}_{\text{em}} + 2)}. \tag{2.45b}$$

This concludes the derivation of the scattering parameters from given surface susceptibilities. Note that for a reciprocal metasurface, we will arrive at the well-known constraint $S_{12} = S_{21}$. Furthermore, the lossless and passive condition translates to $[\mathbf{S}]^H[\mathbf{S}] = [\mathbf{I}]$ where $[\mathbf{I}]$ is the identity matrix, and the superscript 'H' is the Hermitian (complex conjugate transpose) operator [142].

2.1.4 Local Power Conservation

To ensure that a metasurface can be lossless and passive, the local power conservation (LPC) constraint needs to hold [47, 121]. This essentially means that the power

entering each unit cell needs be equal to the power leaving the same unit cell. The LPC constraint will then limit the variety of the transformations that can be supported by lossless and passive metasurfaces. To handle this limitation, two techniques have been used in this thesis, which are briefly described in the following two subsections.

Use of Surface Waves

According to [143, Chapter 11], surface waves are intimately bound to a surface, such as the interface between air and an impedance surface. Thus, their fields are exponentially decaying in the direction normal to the surface while propagating along the surface. For example, let us assume a capacitive impedance surface located on the uv plane. Furthermore, let us assume a 2D problem where $\partial/\partial u = 0$. As shown in [144, Sect. 2.3], the following fields satisfy Maxwell's equations and the boundary condition in the $n \geq 0$ region:

$$H_v = e^{-jk_0 v} e^{-\alpha n}, \tag{2.46}$$

$$E_u = \frac{j\omega\mu_0}{\alpha} e^{-jk_0 v} e^{-\alpha n}, \tag{2.47}$$

where $\alpha \in \mathbb{R}^+$ is a real-valued positive attenuation coefficient. Due to the exponential decay of E_u and H_v, these are evanescent in the $+\hat{n}$ direction and represent surface waves which are tightly bound to uv plane.

Let us now consider the example shown in [116]. In this example, a normal plane wave with the following components

$$E_u = e^{-jk_0 n} \qquad \text{for } n \leq 0^-, \tag{2.48}$$

$$H_v = \frac{1}{\eta_0} e^{-jk_0 n} \qquad \text{for } n \leq 0^-, \tag{2.49}$$

illuminates the metasurface from its bottom face. The purpose is to refract this plane wave into two plane waves of equal amplitudes, say A, as shown in Fig. 2.4 without

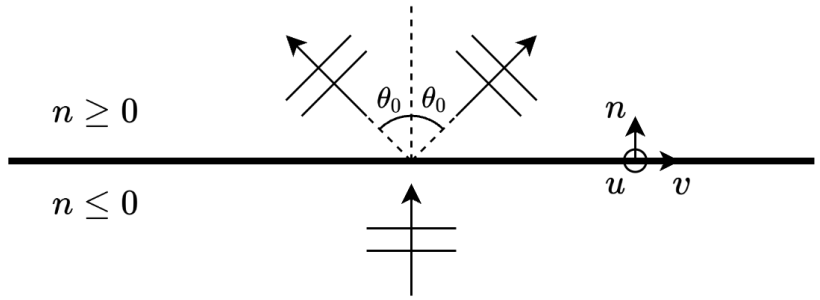


Fig. 2.4 Transformation of a plane wave into two refracted plane waves using a single lossless and passive metasurface utilizing surface waves. The metasurface lies on the uv plane under the 2D assumption.

incurring any reflections. Therefore, the desired fields on $n > 0$ will be

$$E_u = Ae^{-jk_0 n\cos\theta_0}(e^{-jk_0 u\sin\theta_0} + e^{+jk_0 u\sin\theta_0}) \qquad \text{for } n \geq 0^+, \tag{2.50}$$

$$H_v = \frac{A\cos\theta_0}{\eta_0}e^{-jk_0 n\cos\theta_0}(e^{-jk_0 u\sin\theta_0} + e^{+jk_0 u\sin\theta_0}) \qquad \text{for } n \geq 0^+, \tag{2.51}$$

$$H_n = \frac{-A\sin\theta_0}{\eta_0}e^{-jk_0 n\cos\theta_0}(e^{-jk_0 u\sin\theta_0} + e^{+jk_0 u\sin\theta_0}) \qquad \text{for } n \geq 0^+, \tag{2.52}$$

where the angle θ_0 is shown in Fig. 2.4.

Noting (2.48)-(2.51), it is clear that the LPC constraint at the metasurface cannot be satisfied. That is,

$$\frac{1}{2}\mathrm{Re}\{E_u(n=0^-)\times H_v^*(n=0^-)\} \neq \frac{1}{2}\mathrm{Re}\{E_u(n=0^+)\times H_v^*(n=0^+)\}. \tag{2.53}$$

where the superscript '$*$' denotes the complex conjugate operator.

As shown in [116], we can add surface waves to the $n = 0^-$ face of the metasurface to arrive at the following fields on $n \leq 0^-$ side:

$$E_u = e^{-jk_0 n} + E_{\mathrm{sw}}e^{+\alpha n}(e^{-j2k_0 u\sin\theta_0} + e^{+j2k_0 u\sin\theta_0}) \qquad \text{for } n \leq 0^-, \tag{2.54}$$

$$H_v = \frac{1}{\eta_0}e^{-jk_0 n} + j\frac{\alpha E_{\mathrm{sw}}}{k_0\eta_0}e^{+\alpha n}(e^{-j2k_0 u\sin\theta_0} + e^{+j2k_0 u\sin\theta_0}) \qquad \text{for } n \leq 0^-, \tag{2.55}$$

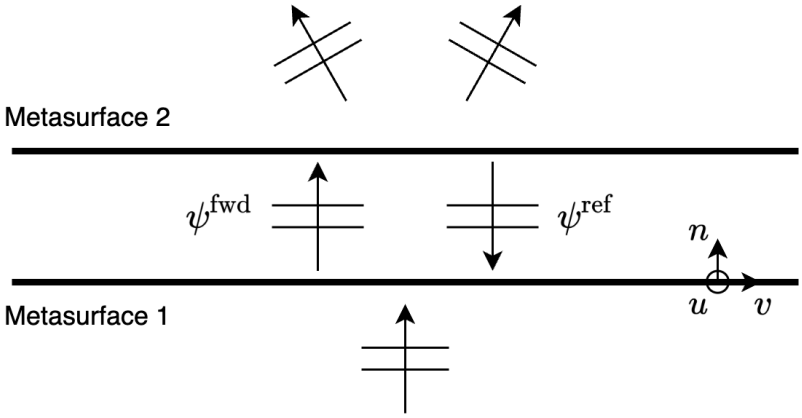


Fig. 2.5 A cascaded metasurface system that utilizes the forward ψ^{fwd} and reflected ψ^{ref} waves in the region between the two metasurfaces for power redistribution.

where $\alpha \in \mathbb{R}^+$. Note that $e^{\alpha n}$ in the two above equations represents exponential decay as n is negative in $n \leq 0^-$ region. Therefore, the new terms in (2.54)-(2.55) are evanescent in the normal direction (i.e., $-\hat{n}$) and the metasurface remains reflectionless. Subsequently, it has been shown in [116] that with the proper choice of A and E_{sw}, the LPC constraint can be met and a lossless and passive metasurface can be designed.

The above was an example wherein the surface waves can be added *analytically* to the solution to help with the power redistribution. However, in general, this requires an optimization approach to incorporate surface waves into our methodology. These optimization approaches typically minimize an appropriate cost functional using gradient based techniques, e.g., see [118, 119, 145].

Use of Cascaded Metasurfaces

Another approach to redistribute the power is to use a metasurface pair, also known as cascaded metasurfaces, as shown in Fig. 2.5 [110, 120, 146, 147] and Fig. 1.5 [111]. In this configuration, the multiple reflections that can occur between the two metasurfaces can redistribute the power such that the LPC constraint is met at each unit cell of both metasurfaces [110]. To this end, we need to infer the appropriate electromagnetic waves in the space between the two metasurfaces, i.e., $(E^{\text{fwd}}, H^{\text{fwd}})$ and $(E^{\text{ref}}, H^{\text{ref}})$, such that

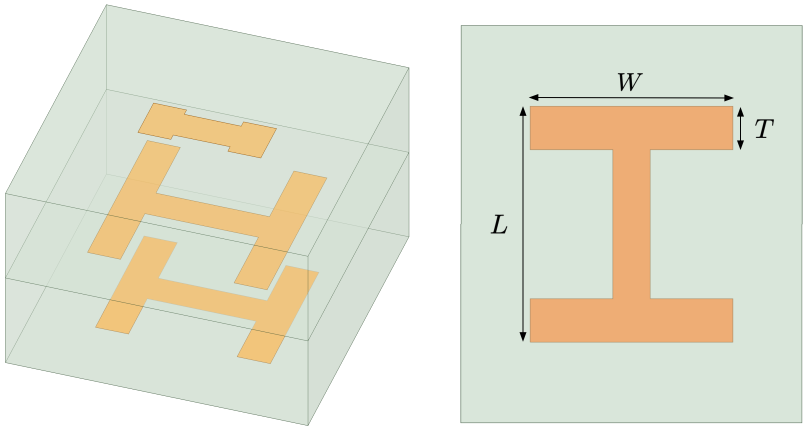
the LPC is met at each unit cell. Given that we already know the incident wave and the desired transmitted wave, if we can infer the appropriate internal waves, we can obtain the required tangential fields on both metasurfaces and complete the macroscopic design process. Inferring the internal fields requires an optimization algorithm that minimizes a cost functional that evaluates the LPC fitness at each unit cell. This cost functional can then be minimized using different techniques including gradient based approaches. Herein, we have used the algorithm presented in [120, 145]. (An example of a cascaded design is shown in Appendix A.)

2.1.5 Microscopic Design

Now that we have covered the macroscopic properties, let us clarify with an example the microscopic implementation of a metasurface. In Figure 2.3, we can see a three-layered structure (with a small thickness of δ) that is used to implement a given set of microscopic properties (surface susceptibilities). As can be seen, the microscopic design is performed by designing copper traces of various dimensions (all of which in the shape of "dogbones" in this particular example) printed on dielectric substrates. In this particular example, two dielectric substrates have been used and therefore the thickness of the metasurface δ is equal to the thickness of two dielectric substrates plus a very thin bondply layer to bond the two substrates. In other words, a given unit cell for this example consists of

unit cell = copper trace+substrate+copper trace+bondply+substrate+copper trace.

It has been shown that with the change of geometrical parameters of the so-called three-layer dog bone inclusions, one can achieve various surface susceptibilities [18, 123]. Thus, in this particular example, implementing required surface susceptibilities are done by controlling the width and length of dogbone copper traces as shown in Figure 2.6. This will be the mechanism with which we perform the microscopic design aspect in this thesis.



(a) (b)

Fig. 2.6 (a) Three-layered dogbone unit cell. (b) Dimensions of the dogbone for a single layer. The parameters L, W, and T are the length, width and trace width of the dogbone. Variation of these dimensions can generate different values for surface susceptibilities.

2.2 Deep Learning

Having established the background related to the metasurface theory, we now turn our attention to the other aspect of this thesis: deep learning. Machine learning, a branch of artificial intelligence, empowers computers to learn from data without explicit programming. This field includes a range of techniques, such as supervised, unsupervised, reinforcement learning, and generative methods [148].

DL [80], which is a subset of machine learning, employs neural networks inspired by the human brain's structure. Since neural networks, and in particular convolutional neural networks, play an important role in DL techniques and this thesis, we will first explore these foundational concepts in the next two subsections. Following this, we will look into three core categories of problems within machine learning, as well as DL in the context of this thesis: unsupervised learning, supervised learning, and generative modeling.

2.2.1 Neural Networks and Backpropagation

Neural networks are the building blocks of DL algorithms [80]. These networks consist of interconnected nodes, called neurons, which are organized in different layers. Each neuron receives input, processes it, and generates an output that contributes to the network's overall computation. These networks are designed to gradually transform its input data into more abstract and useful representations through each layer. The ultimate goal of a neural network is to learn a mapping from input data to the desired output. This mapping is achieved in the training process which captures and differentiates subtle details about the input, while remaining robust to variations of irrelevant information.

An example of a neural network is the multilayer perceptron (MLP) [149], which consists of an input layer, one or more hidden layers, and an output layer. The name "hidden" layer refers to the layers which are neither in the input layer nor the output layer, consequently being "hidden" from the outside. Each layer is fully connected to the next, meaning every neuron in one layer is connected to every neuron in the next layer. As the data passes through the network, each neuron computes a weighted sum of its inputs from the previous layer and then applies a non-linear activation function to produce its output. This non-linear transformation is crucial because it allows the network to capture and model complex, non-linear relationships within the data.

During training, the network adjusts its internal parameters (such as weights and biases) iteratively through a process called backpropagation [150]. In this process, the performance of the network is evaluated by a loss[7] function to update its parameters. To this end, the gradients of the loss function are computed with respect to the network's parameters. To achieve this efficiently, the chain rule is utilized, allowing the calculation of how each parameter affects the loss by breaking down the complex dependencies between layers into simpler and local derivatives. The backpropagation

[7]In this thesis, the "loss" may refer to the loss function in the context of machine learning as discussed in this section. In addition, it may refer to the physical loss due to heat generated in copper traces and dielectric substrates of metasurfaces. This "physical" loss will then affect the transmission efficiency of metasurfaces.

process begins at the output layer, where the network's prediction can be for example compared to the true target value, leading to the computation of the loss. The loss gradient is then propagated backward through the network, layer by layer. At each layer, the gradient of the loss function with respect to the parameters is calculated, allowing for the adjustment of these parameters in a direction that minimizes the loss. This backpropagation of gradients continues until it reaches the input layer.

By repeatedly applying this procedure, the neural network gradually learns to make more accurate predictions. The ability to compute and propagate gradients efficiently through all layers is what enables DL models to capture and learn complex functions from large datasets.

2.2.2 Convolutional Neural Networks

Beside the previously mentioned MLP, another important building block for the DL frameworks developed in this thesis are the convolutional neural networks (CNNs). CNNs are a specialized class of neural networks designed to process data with a grid-like topology, such as images, which can be represented as multi-dimensional arrays [148, Ch. 9]. For example, a color image consists of three 2D arrays corresponding to the red, green, and blue (RGB) color channels. In the case of metasurface microscopic design, the copper trace of a three-layer unit cell can be considered as a three-channel binary image, where the values of one and zero in each pixel represent the presence or absence of copper, and each channel represents a different layer of the unit cell [129]. The architecture of a typical CNN is composed of several stages, each with convolutional and pooling layers.

In a convolutional layer, neurons are connected to local regions of the input through a set of weights known as a filter or a kernel. Herein, for the sake of clarity, let us assume that the input is an image. This local connectivity then exploits the spatial structure of the data, based on the assumption that nearby pixels are more likely to be related. The same filter, with its weights shared across different spatial locations, is applied across the entire input image. The term 'convolutional' is derived from this

mathematical operation. This shared weight structure enables the detection of features such as edges, textures, or other patterns, regardless of their position in the image, making CNNs particularly effective at recognizing patterns across varying contexts. The convolution operation involves sliding the filter over the input array and computing the dot product between the filter and the corresponding local patches of the input. The result of this operation is a feature map, where each point indicates the presence of a specific feature at a given location in the input.

The pooling layer, often placed after a convolutional layer, reduces the spatial dimensions of the feature maps. This is usually done through operations like 'max' pooling, which selects the maximum value from a local patch in the feature map. Pooling helps to make the representation invariant to small translations and distortions in the input data, which is crucial for recognizing patterns in varying contexts. By stacking multiple convolutional and pooling layers, CNNs build a hierarchy of features. Lower layers capture simple patterns like edges or textures, while higher layers capture more abstract concepts such as parts of objects or even entire objects.

2.2.3 Supervised Learning

Having covered key aspects of DL relevant to this thesis, we now shift our focus to the three main types of machine learning problems addressed here. We begin with supervised learning.

Supervised learning is one of the most common paradigms in machine learning (DL included) where algorithms learn from labeled data [151, Sec 1.2]. In this setting, each data point in the training set is associated with a corresponding label or outcome. The labels of the dataset are typically generated by human or simulation software which can be time consuming and costly. The fundamental objective of supervised learning is then to infer a mapping or relationship between input features and output labels based on the provided examples. This type of learning is commonly applied in classification and regression tasks. For classification tasks, the goal is to assign input data to predefined categories such as image classification. A more relevant

application in the context of this thesis would be to predict the state of the unit cell for reconfigureable metasurfaces [115] based on desired far-field requirements. For regression tasks, the aim is to predict a continuous value. For example, in applied electromagnetics, one typical application would be the surrogate modeling [152], where one wants to bypass the time consuming full-wave simulation and predict the field responses based on the input model.

2.2.4 Unsupervised Learning

Unsupervised learning is a type of machine learning where the algorithm learns patterns from input data without explicit labeled responses. Instead, the algorithm is tasked with extracting meaningful patterns, structures, representations, or transformations from the input data without explicit guidance [151, Sec 1.3]. Unsupervised models universally share this characteristic, though their objectives vary widely. They are used in solving tasks like clustering, where similar data points are grouped together based on some similarity measure [153], or dimensionality reduction, which involves reducing the number of input features while retaining essential information [154]. An example for unsupervised learning applied to the macroscopic design of metasurfaces is presented in Chapter 4. In this approach, a forward solver (near-field to far-field transformation) is integrated into the loss function. Consequently, the generated results can be evaluated without explicit labels, enabling the training of the neural network using unsupervised learning.

2.2.5 Generative Modelling

In generative modelling, the objective is to create new data samples that resemble those present in the training dataset. A particular subset of generative modelling is conditional generative modelling which involves a generation process influenced by additional information or context provided during both training and inference

stages [155, Sec 20.3]. This is ideal for design applications such as metasurface designs in both microscopic and macroscopic stages.

It is also important to distinguish between generative modeling and prediction tasks, such as classification and regression. Prediction tasks involve estimating or inferring an output value given an input, which can be formulated as a function approximation problem. These models are usually trained to minimize a loss function that measures the disparity between the predicted and true values. Conversely, generative modeling focuses on generating new data samples that bear resemblance to those present in the training data. Most generative models learn the underlying probability distribution of the data and generate new samples by sampling from it [155, Sec 20].

It should also be noted that a generative model can be categorized as either supervised or unsupervised depending on its training methodology and the intended task. In unsupervised generative modeling, such as the basic VAEs and GANs, the model is trained solely on unlabeled data. In such approaches, the objective is to learn the inherent distribution of the data and subsequently generate new samples without any supervision. On the other hand, in supervised generative modeling, for example conditional generative modeling, the model is trained using labeled data where both input data and corresponding labels (conditions in conditional generative modeling) are provided during the training process. The goal in these approaches is to understand the underlying distribution of the data, alongside the conditional probability distribution of the labels given the data. In conditional generative modeling, the aim is to generate a new sample appropriate based on the given condition, utilizing models such as conditional VAEs and conditional GANs.

Based on the above, it is possible to frame metasurface design tasks as conditional generative tasks. In microscopic metasurface design, the goal is to generate physical unit cell designs based on their macroscopic properties. In particular, in Chapter 6, a conditional diffusion network was developed and trained on a labeled dataset for this purpose with the conditions being the desired scattering parameters. On the other hand, in macroscopic metasurface design, the objective is to generate a set of

metasurface macroscopic properties based on the condition of far-field specifications. As can be seen, in Chapter 3, a transformer model was developed and trained on a labeled dataset to achieve this.

Chapter 3

Macroscopic Design from Far-Field Specifications

Preface

In this chapter, we begin by presenting a deep learning framework for macroscopic metasurface design based on user-defined FF specifications. The FF specifications can, for example, be desired beam directions, null angles, and half-power beamwidths. The contents of this chapter are from a journal paper published in *IEEE Open Journal of Antennas and Propagation* in 2023 [70].[1]

Abstract

A deep learning neural network model in conjunction with a method to incorporate auxiliary surface waves is developed for the macroscopic design of transmitting metasurfaces. The main input to the neural network model is the user-defined desired far-field specifications. This network is used to calculate the required tangential electromagnetic fields on the metasurface. These fields will then be augmented by incorporating

[1] **C. Niu**, M. Phaneuf, T. Qiu and P. Mojabi, "A Deep Learning-Based Approach to Design Metasurfaces From Desired Far-Field Specifications," *IEEE Open Journal of Antennas and Propagation*, vol. 4, pp. 641-653, 2023.

auxiliary surface waves along the metasurface for power redistribution to satisfy the requirement for having lossless and passive metasurfaces. The designs will then be evaluated using full-wave simulations of metasurfaces with three-layer unit cell topology in both 2D and 3D scenarios.

3.1 Introduction

Electromagnetic metasurfaces enable the systematic transformation of incoming electromagnetic (EM) waves to desired outgoing ones [15, 14]. These thin devices can tailor the absorption [53], reflection [156], transmission [157] of EM waves, or a combination thereof. In addition to tailoring the spatial properties of EM waves (e.g., their beamwidths and sidelobe levels), metasurfaces can tailor the spectral properties (frequency contents) of EM waves, e.g., through temporal modulation of their surface properties [62]. Moreover, the simultaneous use of temporal and spatial modulations of metasurface properties enables us to control both spatial and spectral attributes of EM waves [64]. These transformations are governed by the generalized sheet transition conditions (GSTCs) [21] which relate the EM fields on both sides of the metasurface to its surface properties. These properties can be represented in different forms, e.g., through surface susceptibilities [21], polarizabilities [158], or impedances [47]. To design a metasurface, we first need to determine these surface properties. This step, referred to as macroscopic design [159], is then followed by the microscopic design step which is focused on the physical implementation, e.g., via patterned metallic claddings supported by dielectric substrates.

This chapter focuses on the macroscopic aspect of the metasurface design. In particular, it considers the macroscopic design of transmitting omega bianisotriopic metasurfaces under linear polarization. To achieve omega bianisotropy, the unit cells of these metasurfaces are designed to have magnetoelectric coupling such that an acting electric (magnetic) field induces a perpendicular magnetic (electric) current in the unit cell [29]. This can be for achieved, for example, by having non-symmetrical

three-layered unit cells [47]. The macroscopic design of metasurfaces requires the knowledge of the complex (magnitude and phase) tangential EM fields on both faces of the metasurface. Once these are determined, the input and output tangential fields can be linked via two-port networks [19], and subsequently, the required surface properties can be determined. However, the tangential EM fields are typically not known in practical situations [75]. For example, when the designer is only interested in achieving a desired power pattern (phaseless quantity) in the far-field zone of the metasurface, inferring the required complex tangential fields is non-trivial and requires solving an EM inverse problem [75]. Furthermore, the desired performance specification might not even be a power pattern, and can instead consist of a set of far-field performance criteria such as beam directions, half-power beamwidths (HPBW), side lobe levels (SLL), and nulls [71].

In the above situation, which is the primary focus of this chapter, these performance criteria are to be used to infer the required tangential EM fields. To this end, at least three approaches can be pursued. Firstly, the desired far-field performance criteria can be converted to a power pattern using *ad hoc* [75] or deep learning [160] techniques; then, the resulting power pattern can be processed by inversion algorithms to yield the tangential fields on the metasurface [75, 121]. Secondly, the performance criteria can be directly incorporated into an inversion algorithm to infer the required metasurface properties [71]. These two approaches typically require empirical tuning of the algorithms' parameters. Thirdly, standard antenna array techniques such as Chebyshev and Taylor distributions may be used to find the required tangential fields [110]. Such array techniques are efficient and provide great insight into the design process; however, they are often applied to planar and linear uniform arrays for conventional far-field performance criteria. Herein, we use a DL-based approach to obtain the required tangential fields from the knowledge of desired far-field performance criteria. DL techniques, which have been previously used for antenna array synthesis, e.g., design based on far-field masks [161, 102], offer a fast and automated design process.

Moreover, it is typically desired that the metasurface be lossless and passive. To this end, we need to ensure that the normal real power density entering each unit cell of the metasurface is equal to the normal real power density leaving that unit cell. This condition, which is known as local power conservation (LPC)[47], will place a constraint on the stipulated tangential fields. From a mathematical point of view, the inverse problem of inferring tangential fields from desired far-field performance criteria will now have a regularizer that favours LPC [121]. This will therefore limit the solution space. Assuming that the input tangential fields are known and which satisfy LPC, three degrees of freedom are needed in our unit cell topology to enable arbitrary EM field transformations [121]. That is why the unit cells considered in this thesis have three layers.

This chapter presents a macroscopic metasurface design procedure using deep learning techniques under two-dimensional (2D) and three-dimensional (3D) EM wave propagation scenarios.[2] In both 2D and 3D cases, the metasurface lies in the xz plane. For the 2D case, the metasurface is spatially modulated in the x direction and is invariant in the z direction. This enables tailoring the spatial properties of EM waves in the xy plane (i.e, one cut in the far-field zone). On the other hand, in the 3D case, the metasurface is spatially modulated in both x and z directions, thus having the ability to tailor the wave propagation in the entire 3D space. The proposed approach enables us to go directly from the desired far-field performance criteria to the required tangential fields on the metasurface while ensuring that these fields are Maxwellian. In addition, non-conventional performance criteria such as individual HPBWs for multi-beam patterns or maximum null levels (MNL) can be considered. Throughout this thesis, the time dependency of $\exp(j\omega t)$ is considered.

[2]In addition, DL techniques have been used in the microscopic design of metasurfaces [162, 163] and reconfigurable metasurfaces [164, 165].

3.2 Problem Statement and the Overall Methodology

Given (i) a set of desired far-field performance criteria, including main beam directions, HPBWs, nulls, SLLs, and MNLs, (ii) a known incident field, and (iii) a metasurface aperture size, we would like to design a reciprocal, reflectionless, lossless, and passive metasurface that transforms the incoming incident EM wave to an outgoing EM wave that meets the user-defined desired performance criteria.

To this end, we develop a macroscopic metasurface design pipeline as follows.

1. A neural network is developed that takes a set of user-defined far-field performance criteria and then synthesizes an equivalent antenna array so as to meet those desired performance criteria. Throughout this thesis, we refer to this antenna array as the *source array*. In the 3D implementation, this source array consists of infinitesimal electric dipoles and in the 2D case, it consists of infinite line sources.

2. The obtained source array is then used to determine the required tangential electric and magnetic fields on the output face of the metasurface. This step can be easily done through analytical formulae, thus resulting in Maxwellian fields. In other words, inferring a source array serves only as an intermediate step, which is used to determine the required tangential fields on the metasurface.

3. Since the incident EM fields are known and the metasurface is desired to be reflectionless, the tangential fields on the input face of the metasurface are set to be the incident fields. Since the LPC condition is not necessarily satisfied, the metasurface may need loss or gain mechanisms to enable the required EM field transformation. Therefore, a procedure needs to be used to redistribute the power among unit cells so as to meet the LPC constraint. To this end, we utilize auxiliary surface waves along the metasurface [116, 114, 119, 118]. While these auxiliary surface waves can redistribute the power among unit cells, they do not radiate any real power normal to the metasurface.

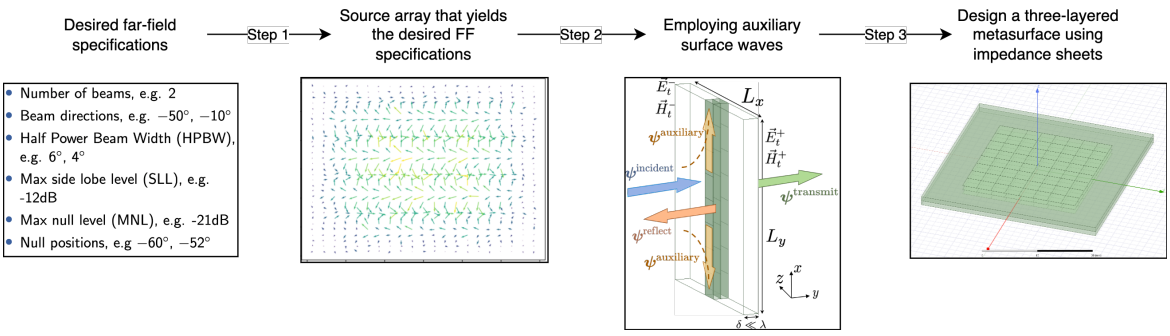


Fig. 3.1 The end-to-end metasurface macroscopic design pipeline utilized in this work. The lengths and directions of the arrows in the source array image depict the amplitude and phase of different infinitesimal dipole antennas respectively.

4. Once LPC is satisfied, the metasurface is implemented using a three-layer unit cell configuration [19, 18] and is simulated in a full-wave solver (Ansys HFSS) to verify how closely it meets the desired far-field performance criteria.

To summarize, a flowchart of the overall design pipeline is shown in Fig. 3.1. Further details will be presented in the following sections.

3.3 Metasurface Background

The fundamental equations to design metasurfaces are the GSTCs. The GSTCs link the field discontinuity supported by the metasurface to its surface susceptibility tensors $\overline{\overline{\chi}}$. Herein, we do not consider any normal susceptibility components (i.e., normal to the metasurface) [137], and thus $\overline{\overline{\chi}}$ is reduced to a 2×2 tensor. The GSTCs can then be written as [21]

$$\hat{n} \times \Delta \boldsymbol{H} = j\omega\epsilon_0 \overline{\overline{\chi}}_{\text{ee}} \boldsymbol{E}_{\text{av}} + jk_0 \overline{\overline{\chi}}_{\text{em}} \boldsymbol{H}_{\text{av}}, \tag{3.1}$$

$$\hat{n} \times \Delta \boldsymbol{E} = -j\omega\mu_0 \overline{\overline{\chi}}_{\text{mm}} \boldsymbol{H}_{\text{av}} - jk_0 \overline{\overline{\chi}}_{\text{me}} \boldsymbol{E}_{\text{av}}. \tag{3.2}$$

In the above two equations, $\hat{n}$ is the normal vector to the metasurface, ϵ_0, μ_0, and k_0 are the permittivity, permeability and wavenumber in free space respectively, and ω is the angular frequency of operation. The "Δ" (difference) and "av" (average) terms in

(3.1) and (3.2) are then expressed as

$$\Delta\psi = \psi^{+} - \psi^{-} \quad , \quad \psi_{\text{av}} = \frac{\psi^{+} + \psi^{-}}{2}, \tag{3.3}$$

where ψ represents the components of $\{\boldsymbol{E}, \boldsymbol{H}\}$ that are tangential to the metasurface and the superscripts "+" and "−" are used to denote the fields on the output and input faces of the metasurface respectively. The subscripts of $\overline{\overline{\chi}}_{\text{ee}}$, $\overline{\overline{\chi}}_{\text{mm}}$, $\overline{\overline{\chi}}_{\text{em}}$, and $\overline{\overline{\chi}}_{\text{me}}$ denote electric-to-electric, magnetic-to-magnetic, magnetic-to-electric, and electric-to-magnetic surface susceptibilities respectively. For example, $\overline{\overline{\chi}}_{\text{em}}$ represents the electric (first subscript) polarization response to a magnetic (second subscript) field excitation.

A reciprocal, lossless and passive metasurface needs to meet the following constraints [21]

$$\overline{\overline{\chi}}_{\text{ee}}^{T} = \overline{\overline{\chi}}_{\text{ee}} \in \mathbb{R}^{2\times 2}, \tag{3.4}$$

$$\overline{\overline{\chi}}_{\text{mm}}^{T} = \overline{\overline{\chi}}_{\text{mm}} \in \mathbb{R}^{2\times 2}, \tag{3.5}$$

$$\overline{\overline{\chi}}_{\text{me}}^{T} = -\overline{\overline{\chi}}_{\text{em}} \in \mathbb{I}^{2\times 2}, \tag{3.6}$$

where $\mathbb{R}$ and $\mathbb{I}$ represent purely real and purely imaginary 2×2 tensors respectively and the superscript "T" denotes the transpose operator. Herein, we further limit ourselves to non-gyrotropic metasurfaces [18] (i.e., no polarization rotation is required), which forces the off-diagonal components of $\overline{\overline{\chi}}_{\text{ee}}$ and $\overline{\overline{\chi}}_{\text{mm}}$ and the diagonal components of $\overline{\overline{\chi}}_{\text{em}}$ and $\overline{\overline{\chi}}_{\text{me}}$ to be zero. Noting that our metasurface is in the xz plane, $\hat{n} = \hat{y}$, and considering the fields tangential to the metasurface as $\{E_z, H_x\}$, the GSTCs will be simplified as

$$-\Delta H_x = j\omega\epsilon_0 \chi_{\text{ee}}^{zz} E_{z,\text{av}} - jk_0 \chi_{\text{me}}^{xz} H_{x,\text{av}}, \tag{3.7}$$

$$-\Delta E_z = j\omega\mu_0 \chi_{\text{mm}}^{xx} H_{x,\text{av}} + jk_0 \chi_{\text{me}}^{xz} E_{z,\text{av}}, \tag{3.8}$$

where the superscripts for χ denote the direction of the field component (second superscript) and the resulting electric and magnetic surface polarization densities

(first superscript). Once the tangential fields on the two faces of the metasurface are obtained, (3.7) and (3.8) can be solved to obtain the required surface susceptibilities. These surface susceptibilities can then be implemented using a three-layer unit cell model. Considering that our metasurface aperture lies in the xz plane for the 2D case ($\partial/\partial z = 0$) under the transverse electric illumination, the electric and magnetic fields tangential to the metasurface each consist of a single vector component, namely, E_z and H_x in (3.7) and (3.8). However, in the 3D case, we need to also include the tangential cross-polarized components, i.e., E_x and H_z components. Similar to [122], we assume that the cross-polarized components of our illuminating antenna are small, and thus we neglect them for simplicity.

3.4 Deep Learning Background and Overall Architecture

As noted in the previous section, the tangential fields on the metasurface $\{E_z, H_x\}$ are needed to calculate the required surface susceptibilities, namely χ_{ee}^{zz}, χ_{mm}^{xx}, and χ_{me}^{xz}. Herein, we develop a neural network based DL algorithm to generate a source array that is able to produce the user-defined far-field performance criteria. This source array is then used to calculate the required tangential fields $\{E_z, H_x\}$ on the output face of the metasurface. To this end, we now discuss some concepts and implementation details of the developed neural network.

The neural network utilized in our work is inspired by the DALL-E network [89], which was originally developed for text-to-image synthesis applications. Although the implementation details of our neural network differ from the original implementation of the DALL-E network, its components aim to achieve similar objectives while being tailored towards our specific application (i.e., far-field pattern synthesis). To better understand the analogy between our application and text-to-image synthesis, let us note that a set of desired far-field performance criteria, similar to text, can be cast as sequential data. Moreover, source arrays can be viewed as multi-dimensional tensors,

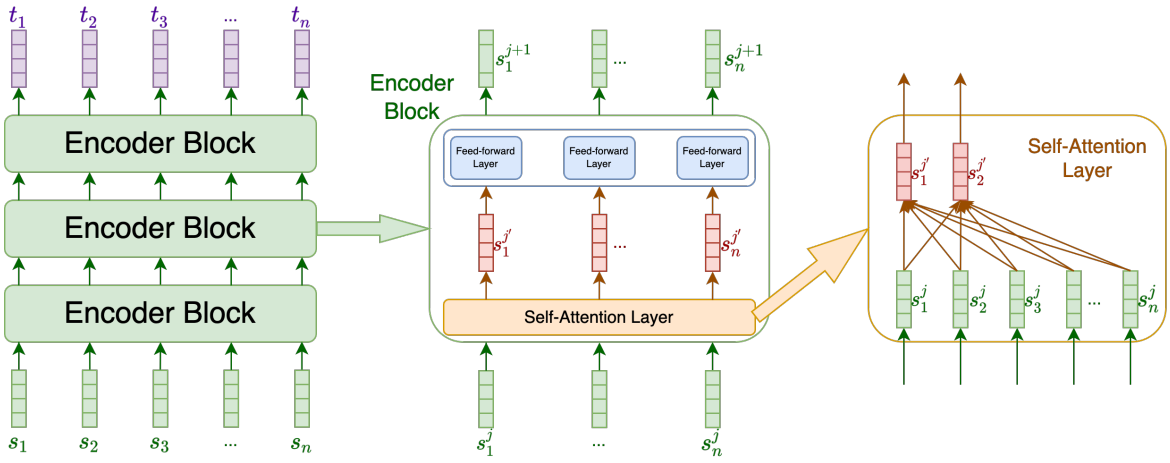


Fig. 3.2 The neural network illustration of the encoder-only transformer. The output of the transformer encoder is normally referred to as the hidden state **H**, but in the case of the encoder-only transformer, **H** is the final output target sequence **T**.

which draw similarities to images. In the remaining parts of this section, we review two fundamental components used in our neural network, namely, encoder-only transformers and discrete variational autoencoders (dVAE).

3.4.1 Encoder-only Transformers

Our transformer component is used to transform a latent representation of user-defined far-field performance criteria to a latent representation of a source array. These latent representations are to be discussed later; in this subsection, we provide background information about transformers [166]. Transformers were originally developed to solve sequence-to-sequence tasks in the field of neural language processing. The sequence-to-sequence task can be interpreted as finding a model that is able to map a discrete source sequence[3] such as words in a sentence to a discrete target sequence which could be words in a different language. These discrete sequences are encoded using embedding layers into sequences of vectors. Thus, the source and target sequences can be represented as $\mathbf{S} = \{\mathbf{s}_1, \ldots, \mathbf{s}_n\}$ and $\mathbf{T} = \{\mathbf{t}_1, \ldots, \mathbf{t}_m\}$ where n and m are the source and target sequence lengths respectively. In summary, sequence-to-sequence problems

[3]Discrete sequences can be represented by non-negative integers (i.e., whole numbers).

can be represented as the mapping of $\mathbf{S}$ to $\mathbf{T}$. In recent years, the transformer based DL models have not only become the state-of-the-art in neural language processing tasks such as question-answering, machine language translation, and text summarizing [167], but have also shown promising results in other areas such as computer vision with applications in text-to-image generation and object detection [168].

The original transformer was designed to be an encoder-decoder based model [166]. The encoder of the transformer encodes the source sequence $\mathbf{S}$ to a sequence of hidden states denoted by $\mathbf{H} = \{\mathbf{h}_1, \ldots, \mathbf{h}_n\}$. The decoder of the transformer then yields the target sequence autoregressively based on the hidden states and all the preceding target sequences. Later, it was shown that an improved performance can be achieved by dropping either the encoder or the decoder of the transformer [169, 170]. Herein, our implementation is based on encoder-only transformers, which are built by stacking encoder blocks as shown in Fig. 3.2 (leftmost) [4]. Each encoder block then consists of a self-attention layer and a feed-forward layer as depicted in Fig. 3.2 (middle). The self-attention layer functions by putting each source vector at the jth encoder block into relation with all input source vectors $\mathbf{S}^j$ as illustrated in Fig. 3.2 (rightmost). The output of the self-attention layer, denoted by $\mathbf{S}^{j'}$, is thus a more refined contextual representation of itself, hence the name self-attention. Each source vector is then passed to a feed-forward layer where information moves from input to output nodes by applying a weighted sum of inputs followed by an activation function. In summary, through incorporating self-attention layers, the transformer is now capable of transforming the original source vector sequence $\mathbf{S}$ from context-independent to context-dependent $\mathbf{H}$. In the case of neural language modelling, context-dependency means that each word in a sentence is aware of its surrounding words, which is critical for language modelling. Similarly, in our application, it means that far-field performance criteria such as nulls are aware of other surrounding far-field features such as sidelobes.

In the case of encoder-only transformers, $\mathbf{H}$ is the final output of the network and should therefore be viewed as the target sequence $\mathbf{T}$ in Fig. 3.2 (leftmost). Consequently,

[4] This Figure is based on [171].

the output $\mathbf{T} = \mathbf{H}$ will have the same length as the input source $\mathbf{S}$ in encoder-only transformers. In our case, $\mathbf{S}$ represents the latent space of far-field performance criteria and $\mathbf{T}$ represents the latent space of source arrays. However, as will be discussed later, the sequence lengths associated with far-field performance criteria and source arrays are fundamentally different. Thus, the issue of non-identical sequence lengths needs to be considered in the development of our entire network.

3.4.2 Discrete Variational Autoencoders (dVAE)

We use discrete variational autoencoders (dVAE) [89] to link far-field performance criteria and source arrays to their own latent representations. This enables the use of our encoder-only transformer which functions in the latent space. This is inspired by the DALL-E network for text-to-image synthesis whose transformer outputs latent representations of images in the form of discrete sequences. These discrete sequences are referred to as image tokens. To yield actual images, dVAEs [89] were developed to decode these tokens. Similarly, in our application, we need to handle tokens associated with far-field performance criteria and source arrays.

The core of the dVAE is an unsupervised learning technique called the autoencoder which consists of two main components. The first component, called the encoder, encodes/compresses the input data x into a latent space z whereas the second component, called the decoder, decodes/decompresses the latent space z to the reconstructed form of the original data $\hat{x}$.[5] Thus, the objective of most antoencoders is to minimize the discrepancy between the original input data x and the reconstructed data $\hat{x}$. It is now easy to see that the training of autoencoder networks is unsupervised since it only needs unlabeled data x. The latent space z can be viewed as the underlying "hidden" representation of the input data x which has a lower dimension compared to x. Therefore, autoencoders may be conceptually viewed as hourglass-shaped structures that

[5]The encoder-decoder component of autoencoders is *not* the same as the encoder-decoder component of transformers. Although both carry similar meanings as they "encode" inputs into hidden/latent space, they are two different components of two different networks.

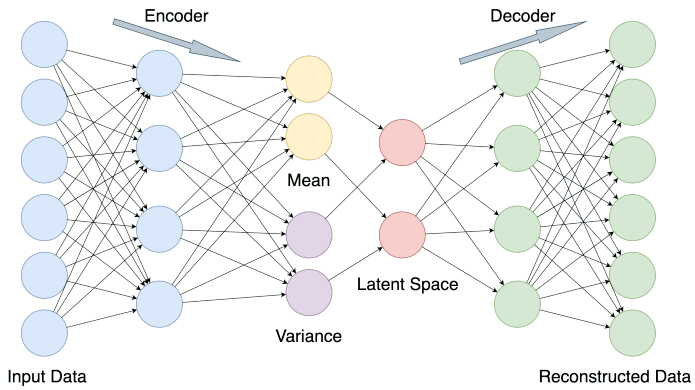


Fig. 3.3 The neural network illustration of variational autoencoder (VAE).

first distill the essential information (through their bottlenecks) and then reconstruct them back.

When using an autoencoder type model, we would ideally want to cluster semantically similar data points in the latent space and to separate them otherwise. However, if autoencoders are implemented based on simple encoder-decoder structures (also known as vanilla autoencoders), they often suffer from not being able to generalize their reconstruction to unseen data [172]. This is due to the fact that vanilla autoencoders tend to isolate individual data points in the latent space, thus, effectively only memorizing every training data point without generalization. The variational autoencoder (VAE) [173] overcomes this issue by enforcing a probabilistic prior on the latent space in a standard Gaussian distribution as illustrated in Fig. 3.3. In other words, instead of producing the latent space directly, the encoder will generate mean and variance vectors from which the latent space is formed using the standard Gaussian distribution.

The autoencoders discussed so far use continuous vector spaces (i.e., vectors of real numbers) to represent the latent space. However, it could be beneficial to represent the latent space in discrete vector spaces (i.e., vectors of whole numbers). In fact, a lot of real-world data favor discrete representations, such as language and objects in an image. In addition to the data representation, there are a number of DL techniques that are specifically designed to work with discrete data such as our utilized encoder-only

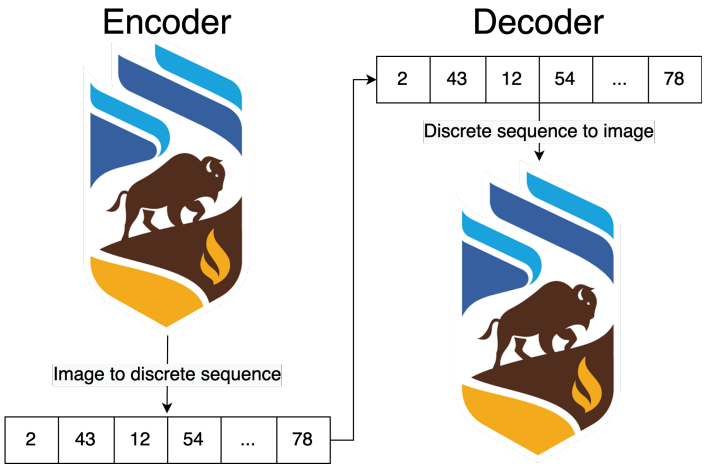


Fig. 3.4 The conceptual illustration of discrete variational autoencoders (dVAE) which shows encoding the University of Manitoba's logo to a discrete sequence and then decoding this sequence back in an attempt to arrive at the original logo.

transformer discussed in Section IV.3.4.1. Consequently, a number of VAE models are designed to distill discrete latent spaces, such as vector quantised VAE [174] and dVAE [89].

Herein, we focus on the dVAE which was originally designed to work with image data as shown in Fig. 3.4. The dVAE network enables having discrete latent spaces by introducing the so-called "codebook" component in its architecture. The codebook behaves like a dictionary where it indexes a list of vectors of real numbers such that each vector is associated with a whole number. Using this approach, the encoder of the dVAE produces a list of discrete numbers as the latent space, and the decoder of the dVAE can then use its codebook to lookup and generate a continuous latent space, thus yielding the final output. As noted earlier, the discrete latent space of dVAE is often referred to as tokens. For example, when the input of the dVAE are images, the latent space of dVAE is referred to as image tokens.

To summarize this section, through the combination of the transformer and dVAE, the DALL-E network transforms text to image tokens using its transformer. It then decodes these image tokens to actual images using its dVAE. In the next section, we are going to adapt the DALL-E network to our own specific application.

Theta (deg)	0.0	1.0	...	59.0	60.0	61.0	...	179.0	180	90.0	...	90.0	90.0	90.0	90.0	90.0	90.0	90.0	...	90.0
Phi (deg)	90.0	90.0	...	90.0	90.0	90.0	...	90.0	90.0	0.0	...	67.0	68.0	69.0	70.0	71.0	72.0	73.0	...	180.0
Beams	0	0	...	0	1	0	...	0	0	0	...	0	0	0	1	0	0	0	...	0
HPBW	0	0	...	1	1	1	...	0	0	0	...	0	1	1	1	1	1	0	...	0
Nulls	0	1	...	0	0	0	...	1	0	0	...	1	0	0	0	0	0	1	...	0
SLL (dB)	-12.3	-12.3	...	-12.3	-12.3	-12.3	...	-12.3	-12.3	-13.8	...	-13.8	-13.8	-13.8	-13.8	-13.8	-13.8	-13.8	...	-13.8
MNL (dB)	-21.5	-21.5	...	-21.5	-21.5	-21.5	...	-21.5	-21.5	-37.6	...	-37.6	-37.6	-37.6	-37.6	-37.6	-37.6	-37.6	...	-37.6

Fig. 3.5 The vectorized far-field performance criteria where the first row lists the theta (θ) angles of the far-field cut. The main beam directions, HPBWs of each beam, and null locations are represented in binary format, where 1 represents the presence and 0 represents the absence of the quantity, and are located at the second, third, and last rows respectively. The side lobe level (SLL) and maximum null levels (MNL) are arrays of constant value and are located at the fourth and fifth rows respectively.

3.5 Source Array Synthesis

Based on the above, we now develop a DL technique to synthesize source arrays from far-field performance criteria. The input to this network is sequential data consisting of far-field performance criteria such as the list of main beam directions, HPBWs, null locations, SLLs, and MNLs. As will be seen, we represent this input as a sequential[6] 2D tensor. The output of the network, depending on the 2D or 3D pattern synthesis, will be linear or planar source arrays which can be represented by 2D or 3D tensors respectively. Thus, similar to text-to-image generation, source array synthesis can be formalized as sequential 2D tensor data to 2D/3D tensor generation. However, there are a few challenges when utilizing a text-to-image model such as the DALL-E architecture for this application. Firstly, the far-field performance criteria are not well-suited for neural network models. For example, the number of HPBWs is a varying quantity depending on the number of main beams whereas the SLL in a given far-field cut is a single quantity representing the maximum level of sidelobes in the entire radiation pattern. In addition, the units are different; e.g., the HPBW is in degrees and the SLL is in dB. Secondly, most text-to-image models such as the DALL-E architecture are designed to handle discrete sequences as their input; however, our input is a continuous sequence with real numbers; e.g., a SLL of -12.3 dB. Lastly, for linear source array

[6]It is "sequential" since the order of angles matters, similar to the order of words in a sentence. Also, the number of angles can vary based on the angular resolution chosen to represent the far-field specifications.

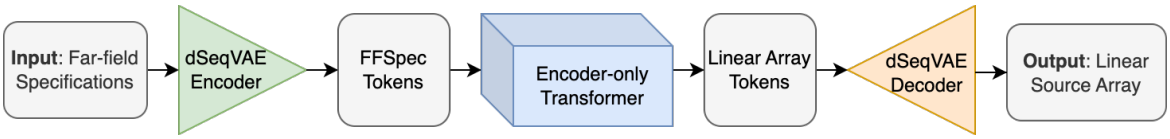
synthesis, the output is a 2D tensor which is not an image-like 3D tensor. In what follows, we address the above challenges.

3.5.1 Formatting Far-Field Specifications to Sequential Data

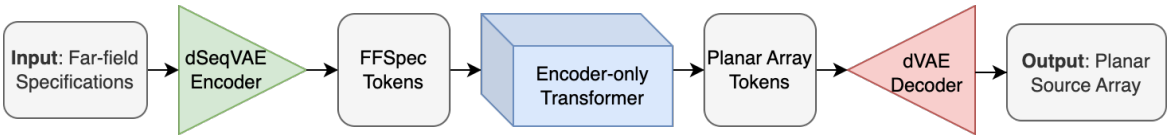
In order to make it easy to train the DL model, we have vectorized the far-field (FF) performance criteria to a sequential 2D tensor, similar to [160], as shown in Fig. 3.5. As can be seen, the first two rows represent the angular locations (θ, φ) to represent the 3D space. In particular, the example shown in Fig. 3.5 considers two FF cuts. They are $\varphi = 90°$ and $\theta = 90°$. The "Beams" row includes binary numbers in which 1 and 0 represent the presence and absence of a main beam respectively. The "HPBW" row also includes binary numbers where the presence of 1 indicates the angles at which the power level is within 3 dB of the main beam level. The "SLL" row indicates the value for the side lobe level. The "Nulls" row is a binary row where 1 indicates the presence of a null at the corresponding angle (θ, φ). Finally, the "MNL" row represents the maximum null level. Each row is then normalized (in the implementation) such that its components lie between 0 and 1 based on the minimum and maximum values of that particular row.

3.5.2 Adapting the dVAE to Source Array Synthesis

As noted earlier, the dVAE network was originally developed to encode images to discrete sequences (image tokens) to be utilized by transformers. Similarly, the dVAE can also be used for planar source array synthesis since a rectangular infinitesimal dipole array can be represented as an image-like two-channel 3D tensor. The two channels are associated with the amplitude and phase of the excitation for each antenna element. This is conceptually similar to color images that have three channels: red, green, and blue (RGB). Moreover, the planar source array is a 3D tensor since it consists of a planar domain whose pixels have a complex-valued excitation coefficient



(a) The proposed neural network architecture of far-field specifications to linear source array synthesis.



(b) The proposed neural network architecture of far-field specifications to planar source array synthesis.

Fig. 3.6 The proposed neural network architecture of far-field specifications to source array synthesis for both (a) 2D and (b) 3D scenarios. It should be noted that in (a) the dSeqVAE encoder and the dSeqVAE decoder on the two sides of the transformer are of two separate models as indicated by the different colors. On the other hand, the dSeqVAE encoders in (a) and (b) are of the same model, and are therefore shown with the same color.

as the third dimension. Herein, we refer to the discrete latent space of the dVAE when applied to planar source arrays as *planar array tokens*.

On the other hand, linear source arrays and desired FF performance criteria on given cuts cannot be represented by an image-like 3D tensor as they lack one dimension. To overcome this issue, we modify the dVAE into the so-called dSeqVAE [175] so that FF performance criteria and linear source arrays can be used by the transformer. (To this end, dSeqVAE replaces the `Conv2D` and `ConvTranspose2D` used in dVAE with `Conv1D` and `ConvTranspose1D` respectively.) Thus, we use the dSeqVAE to encode formatted FF performance criteria (FF specifications) to the so-called *FFSpec tokens* and encode linear source arrays to *linear array tokens*.

3.5.3 Final Network Architecture

The complete neural network architecture developed for this work is shown Fig. 3.6[7], which borrows its general framework from the text-to-image DALL-E network. This architecture has been shown under two scenarios, which are related to linear and planar source array syntheses. These inferred linear and planar source arrays can then be

[7]For further information, see Appendix D.

used to calculate the required tangential fields for metasurface design under 2D and 3D wave propagation scenarios respectively. There are a few unique features associated with our proposed network as compared to the original DALL-E network. Firstly, we employed the dSeqVAE to convert the formatted FF specifications to FFSpec tokens so that they can be utilized by the transformer as its input. Secondly, we used an encoder-only transformer instead of the decoder-only transformer utilized in DALL-E. Using an encoder-only transformer compared to its decoder-only counterpart dramatically reduces the run time in prediction. This is due to the fact that the decoder-only transformer generates the output autoregressively; that is, the network effectively will run multiple times for each prediction. In text-to-image synthesis, this is needed as the network needs to handle various text lengths. On the other hand, our application does not require this autoregressive feature since the lengths of the tokens can be engineered to be fixed. Thus, our proposed network relies on an encoder-only transformer, which will only be executed once during prediction. Thirdly, for linear source array synthesis, we replaced the dVAE network on the output side of the transformer with the dSeqVAE network in order to generate linear source arrays. Finally, encoder-only transformers require the input sequence length to be the same as the output sequence length. With the employment of the dSeqVAE and dVAE on the input and output sides of the transformer, the final network was designed so as to enforce FFSpec tokens and array tokens to have the same length, thus enabling the use of an encoder-only transformer.

3.5.4 Dataset Generation and Training

The data set generation begins by considering aperture sizes for the source arrays. In general, we choose an array size that provides a sufficient number of degrees of freedom in order to meet the specifications, but not so large so as to make the resulting metasurface prohibitively large, i.e., to simplify the design, simulation, and hypothetical manufacturing process. To demonstrate the methodology for planar source array synthesis, we consider a $3\lambda \times 3\lambda$ rectangular array of 256 infinitesimal dipoles

where λ is the wavelength of operation. These infinitesimal dipoles all have the same linear polarization and are evenly spaced on the array aperture which is centered at the origin. The FF requirements can be expressed on multiple cuts in the FF zone, e.g., on the E-plane and H-plane cuts, which are spanned by θ and φ. As for the case of linear source array synthesis, we consider 16 infinite line sources, which are evenly spaced along a line with a total length of 6λ centered at the origin. The FF requirements associated with linear source arrays are expressed on a single cut which is spanned by the single angle φ.

In order to train the neural network properly, we need a large and diverse dataset. In particular, the dataset should have diversity in the numbers of the beams and their locations, variations of the HPBWs for each beam, different SLLs, and so on. To this end, we have constructed a dataset generation pipeline for both linear and planar source array scenarios which is an extension to our previous work [160]. Multiple-beam far-field patterns are generated by superposition of multiple single-beam arrays. Different tapering techniques are randomly applied to arrays so as to vary the side lobe levels such as Dolph-Chebyshev windows, cosine and squared cosine tapering. Random white noise was applied to arrays to further increase the diversity of far-field patterns. Finally, based on the generated far-field pattern, the specifications of interest are captured through signal processing techniques.[8]

As can be seen, the training procedure requires four steps. The first step is to train the dSeqVAE network, shown in green color in Fig. 3.6, using normalized far-field specifications. The second step is to train the dSeqVAE network, shown in orange in Fig. 3.6, using linear source arrays. In the third step, the dVAE network, shown in red in Fig. 3.6, is trained with planar source arrays. Finally, the fourth step involves training

[8]In each epoch, unique datasets for training and testing are generated dynamically, with the generation occurring concurrently with the neural network training process. The size of these datasets scales with the training time. Our neural networks were trained for approximately 2000 epochs, utilizing a dataset of 10 million data points for training and 1 million data points for testing. The training hardware employed was the Tesla V100 GPU, resulting in an overall training period of approximately 4 hours.

the encoder-only transformer with the FFSepc tokens and array tokens obtained from the trained dSeqVAE and dVAE networks for both linear and planar source arrays.

Finally, it should be noted that the above source array aperture dimensions are *a priori* maximum values, which can be decreased by deactivating specific sections of the aperture. The aperture size of the metasurface will then be greater than (or equal to) the effective aperture size of the source array. To understand this, we would like to highlight that the transformation of the source array to the output tangential field profile in this work is conceptually similar to the procedure used in planar near-field antenna measurements [176]. In planar near-field measurements, we characterize the antenna under test by measured tangential fields over a truncated plane, which can then be forward-propagated to obtain the resulting far-field pattern. Similarly, we find the fields of the source array on a truncated plane which will then serve as the metasurface aperture. Thus, our choice of the metasurface aperture is analogous to the choice of the measurement plane in planar near-field antenna measurements. In particular, this plane must be large enough to meet the requirements on the desired critical angle in the far-field [176]. However, we must again take care so as to not make the metasurface prohibitively large. In this work, the metasurface size is chosen in a somewhat *ad hoc* manner, keeping in mind that the size should not be too large so as to make the design, simulation, and hypothetical manufacturing easier, but not too small so as to greatly affect the far-field pattern.

3.6 Auxiliary Surface Waves

As noted earlier, we use auxiliary surfaces waves along the metasurface to redistribute the power to arrive at lossless and passive metasurfaces [116, 114, 119, 118]. The utilized surface wave optimization algorithm [119] assumes a known incident and transmitted field, and finds the auxiliary surface waves that satisfy the LPC condition when added to the original fields. This is accomplished by minimizing the following

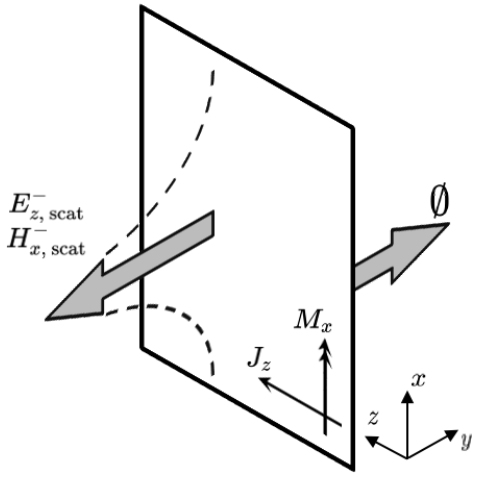


Fig. 3.7 The illustration of the equivalent problem with respect to the input region for the scattered fields, which consist of the auxiliary surface waves used to meet the LPC constraint. Love's condition is assumed for the output region of this equivalent problem and is enforced by its equivalent additive constraint during the optimization process.

expression at each unit cell in a least-squares manner [119]

$$\frac{1}{2}\mathrm{Re}\{E_z^- \times (H_x^-)^*\} - \frac{1}{2}\mathrm{Re}\{E_z^+ \times (H_x^+)^*\}, \tag{3.9}$$

where E_z^- and H_x^- are the *total* tangential electric and magnetic fields at the input face of the metasurface, and E_z^+ and H_x^+ are the *total* tangential electric and magnetic fields at the output face of the metasurface. In addition, the superscript "$*$" denotes the complex conjugate operator. The total field consists of the originally known or desired fields and any added auxiliary surface waves; the latter are the unknowns in the surface wave optimization process. In this work, we assume that the auxiliary waves are added to the input. Thus, the total tangential fields at the input face of the metasurface may be written as

$$E_z^- = E_{z,\mathrm{inc}}^- + E_{z,\mathrm{scat}}^-, \tag{3.10}$$

$$H_x^- = H_{x,\mathrm{inc}}^- + H_{x,\mathrm{scat}}^-, \tag{3.11}$$

where the subscripts "inc" and "scat" represent the incident and scattered fields. In this case, the incident field is the field with which we illuminate the metasurface, and the scattered field is the field produced by the auxiliary surface waves. These scattered fields are optimized to meet the LPC condition at each unit cell in a least-squares manner. In order for these scattered fields to be Maxwellian, we further optimize an additive constraint which enforces Love's condition on the equivalent currents corresponding to the auxiliary fields [177, 119, 178]. These equivalent currents result from an application of the equivalence principle with respect to the input region for the scattered fields, as demonstrated in Fig. 3.7, where the zero-field condition in the output region is enforced by the additive constraint. (Note that under Love's condition, the equivalent electric and magnetic currents will be $J_z = H_x$ and $M_x = E_z$ respectively.) The evanescent quality of these scattered fields, depicted with exponential decay in Fig. 3.7, is obtained by ensuring that the original *total* incident and transmitted powers are the same [119]. Thus, in our approach, we set the tangential EM fields on the input face of the metasurface to (3.10) and (3.11) where $E^-_{z,\text{scat}}$ and $H^-_{x,\text{scat}}$ are associated with auxiliary surface waves traveling along the metasurface for power redistribution. Following this optimization process, we have the complete description of the input and output tangential fields, and thus the metasurface macroscopic design process can now be performed. Finally, we note that the auxiliary surface wave optimization procedure developed in [119] for 2D problems has been extended to 3D for this work.

3.7 Results

The proposed macroscopic metasurface design pipeline is now demonstrated with a few examples for both 2D and 3D cases. For both cases, the design process begins with the user specifying the desired features for the far-field power pattern, namely, beam directions, HPBW for each beam, SLLs, null angles, and MNLs. Herein, for the 3D case, the user provides these specifications for two perpendicular cuts. Then, the neural network will generate a source array so as to meet the user-defined far-field

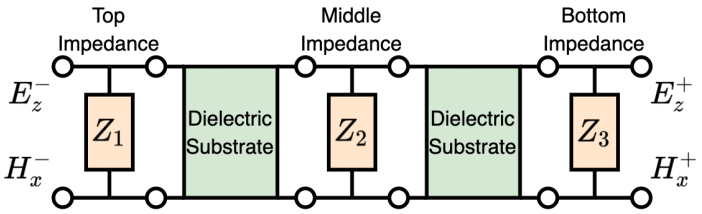


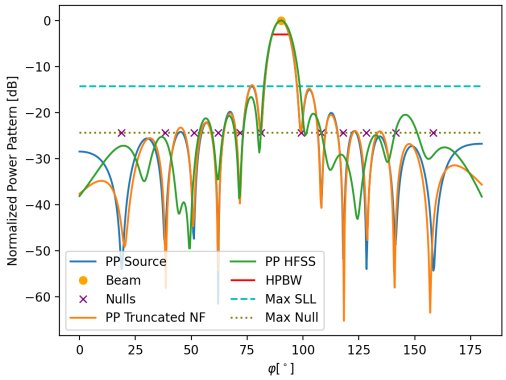
Fig. 3.8 The two-port circuit model for a three-layer metasurface unit cell.

requirements. The tangential electromagnetic fields on the transmitting side of the metasurface can then be analytically computed from this source array. Since the tangential fields on the input face of the metasurface are known (see Section 3.6), the macroscopic properties of the desired metasurface can be computed. Each unit cell of the metasurface is then represented by the widely-used three-layer model shown in Fig. 3.8 [19]. This model consists of three purely reactive impedances, denoted by Z_i where $i = \{1, 2, 3\}$ and two dielectric substrates. Given the substrate and the input/output fields on the two ports of this model, we can find these purely reactive impedances. In the full-wave simulations, these impedances are modeled as impedance boundary conditions in Ansys HFSS.

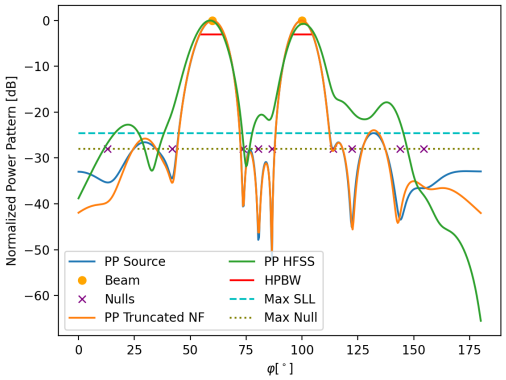
3.7.1 2D Design

Let us first consider 2D design under the transverse electric illumination. The source that illuminates the input face of the metasurface is assumed to be a line source along the z axis. As noted earlier, the metasurface lies in the xz plane and is invariant along the z direction. The length of the metasurface (along the x direction) is assumed to be 10λ where λ is the wavelength of operation at 10.5 GHz. The metasurface is discretized into 61 unit cells. The separation from the line source to the metasurface is $\lambda/2$. The user then provides the desired far-field specifications on the output region of the metasurface which is spanned by φ varying from 0° to 180°.

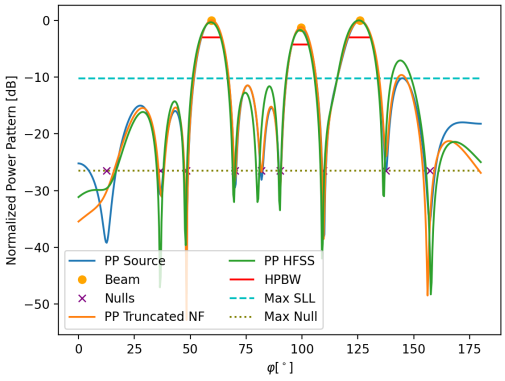
Herein, we have shown three test cases in Fig 3.9(a)-(c). The user-defined desired specifications are shown as “Beam”, “Nulls”, “HPBW”, “Max SLL”, and “Max Null” in these three figures. For each test case, the far-field specifications are used by the neural



(a) Single beam.



(b) Two beams.



(c) Three beams.

Fig. 3.9 Three scenarios for the 2D design cases.

network to yield a linear source array. The far-field power pattern (PP) generated by this source array is shown in blue and is denoted by "PP Source". The tangential electric and magnetic fields close to this array are then obtained on a line with the total length of 10λ (metasurface length). Once these tangential fields are obtained, we can use near-field to far-field (NF-to-FF) transformation to obtain the resulting power pattern. Since the aperture size (10λ) is finite, the NF data along the aperture but outside its 10λ length is assumed to be zero. That is why the resulting power pattern (PP) from this NF-to-FF transformation is denoted by "PP Truncated NF" and is shown in orange. As expected, "PP Truncated NF" matches quite well with the "PP

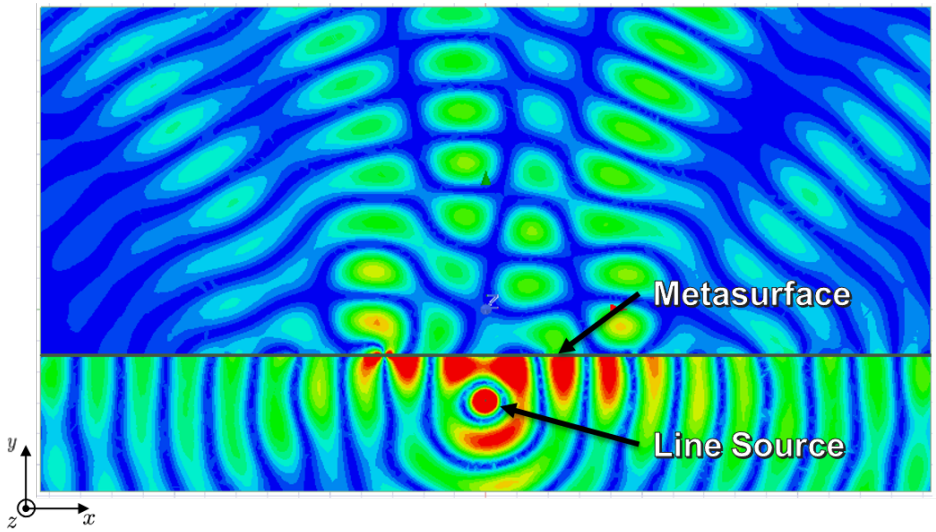


Fig. 3.10 Ansys HFSS simulation of the metasurface under the 2D transverse electric illumination, which plots E_z magnitude at a fixed phase.

Source" with some small discrepancies due to the NF truncation when φ approaches 0° and 180°. Also, these two power patterns meet the desired performance criteria.

In conjunction with the known incident field on the metasurface (i.e., cylindrical waves emanating from the line source), we utilize auxiliary surface waves on the input side of the metasurface to meet the LPC constraint. Once the LPC constraint is met, the three-layer purely reactive impedance sheet model for each unit cell is obtained and used to simulate the metasurface in Ansys HFSS. The utilized dielectric substrate is Rogers RO3010 substrate having a thickness of 0.25 mm (with the relative permittivity of 10.2 and the loss tangent of 0.0035). After performing the full-wave simulation, the tangential fields just after the output face of the metasurface are extracted from Ansys HFSS and are passed to a 2D integral equation solver to find the resulting far-field power pattern. This power pattern is then referred to as "PP HFSS" (green). As can be seen in Fig. 3.9, the HFSS power patterns meet the beam directions and HPBWs generally well. However, meeting the other requirements, in particular nulls, is not as good as these two parameters and gets worse as we move away from main beams. One of the reasons that "PP HFSS" cannot meet the FF performance criteria as well as "PP Source" and "PP Truncated NF" lies in the fact that the transmission line model used

to calculate the impedances in the three-layered unit cell model works based on the local periodicity assumption [179, 123], which is not fully satisfied especially when we employ high spatial field variations associated with the evanescent spectrum (auxiliary surface waves). Moreover, Fig. 3.10 shows the NF simulation of the three-beam example (corresponding to Fig. 3.9(c)) where it can be seen that the fields on the input side of the metasurface have almost maintained their circular wave fronts associated with the line source excitation. Finally, the power transmission efficiency of this metasurface, defined as the ratio of the normal real power leaving the metasurface to the normal real incident power impinging on the metasurface, is 73.1%.[9] The sources of power loss are the reflections from the metasurface (due to numerical and approximation error, e.g., not exactly satisfying the local periodicity assumption) and the loss within the RO3010 dielectric substrates. In addition, it has been noted in [118] that the interaction of strong evanescent waves with the metasurface, described in Section 3.6, results in the further decrease of the power transmission efficiency as compared to metasurfaces that do not utilize these auxiliary surface waves.

3.7.2 3D Design

For the 3D example, the source that illuminates the input face of the metasurface is a horn antenna shown in Fig. 3.11, which is fed by the TE_{10} mode of a rectangular waveguide. Similar to [122], the distance from the horn aperture to the input face of the metasurface is 3λ. The co-polarized component of the horn antenna is E_z. The metasurface's aperture size is $5\lambda \times 5\lambda$ and is uniformly discretized into 30×30 unit cells. In this example, the user defines the desired far-field specifications on two perpendicular cuts (E-plane and H-plane cuts) with the main beam at $(\theta = 70°, \varphi = 90°)$. The desired far-field features on these two cuts are shown in Fig. 3.12, denoted by "Beam", "HPBW", "Nulls", "Max SLL", and "Max Null". Subsequently, the developed neural

[9] Making the unit cell size smaller results in increased transmission efficiency. For example, if the unit cell size is set to about one-tenth of the wavelength, the efficiency increases to 82.0%. In addition, for the original size of the unit cell, if the loss of the Rogers substrate is set to zero, the power transmission efficiency increases from 73.1% to 80.8%.

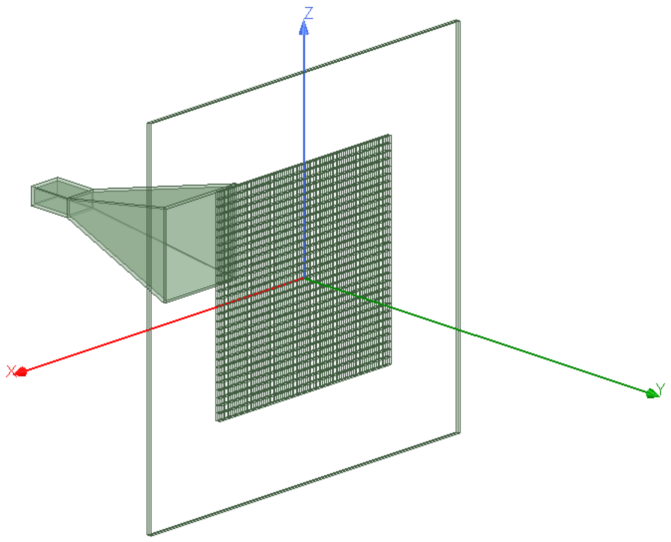
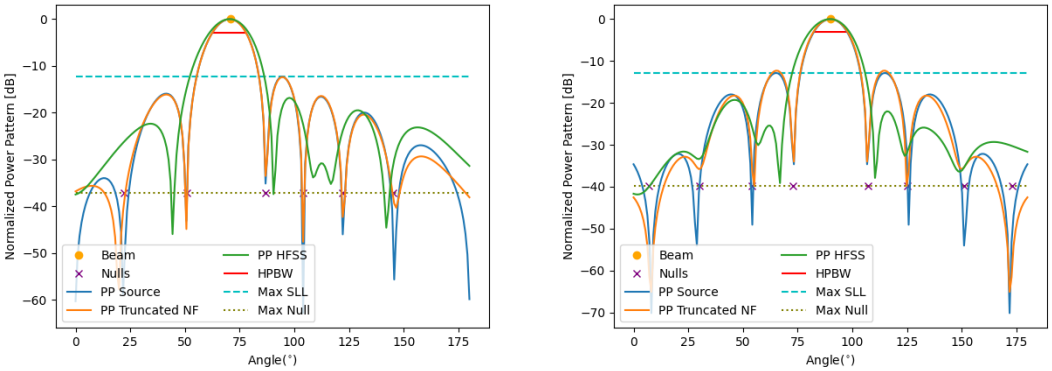


Fig. 3.11 Ansys HFSS metasurface simulation set up for the 3D case with an illuminating horn antenna.

network takes these specifications and outputs a planar source array. The far-field power pattern generated by this planar source array is shown in blue and is indicated by "PP Source". Then, the tangential electric and magnetic fields radiated by the planar source array is found on a plane with the size of $5\lambda \times 5\lambda$, which is the aperture size of the metasurface. NF-to-FF transformation can then be applied to these tangential fields, which generates the power pattern shown in orange and denoted as "PP Truncated NF". Similar to the 2D case, these two power patterns are close to each other and they meet the FF requirements except at a few nulls.

After utilizing auxiliary surface waves to meet the LPC constraint, the three-layered metasurface is constructed. The dielectric substrate has a relative permittivity of 10.2, loss tangent of 0.0035, and thickness of 0.37 mm. In addition, perfect electric conductor (PEC) baffles [123, 122] are used along the x axis to improve the performance of the metasurface. (These baffles cannot be placed along the z axis as the co-polarized electric field component E_z would then be tangential to the PEC baffles.) The metasurface is simulated in Ansys HFSS and the resulting tangential fields close to the output face of the metasurface are collected. The NF-to-FF transformation of this data yields the resulting power pattern shown in green and denoted as "PP HFSS" in Fig. 3.12. As can



(a) E-plane cut (b) H-plane cut

Fig. 3.12 The 3D example where the desired specifications are given on two perpendicular far-field cuts.

be seen, this power pattern satisfies the main beam directions in both cuts. However, the achieved HPBWs are larger than the desired ones. The achieved SLL meets the desired requirement. Moreover, the achieved nulls are generally close to the desired nulls. Since MNLs are not fully satisfied in the "PP Truncated NF", they are not satisfied in "PP HFSS" either. Finally, the near-field simulation of this metasurface in the E-plane cut (yz plane) is shown in Fig. 3.13. The power transmission efficiency of this metasurface is 64.2%. The minor decrease in efficiency when compared to the 2D case can partly be explained by the differences in the assumption of local periodicity. In the 2D case, this assumption is not exactly satisfied along the x direction, but it is satisfied along z due to the nature of the 2D problem (if we were to extend the 2D problem into 3D space); conversely, in the 3D case, both of these axes will contribute error to this assumption, which can result in additional reflections due to this inconsistency.

3.8 Conclusion

We have developed a DL-based macroscopic metasurface design approach. The developed neural network takes user-defined desired far-field performance criteria and then

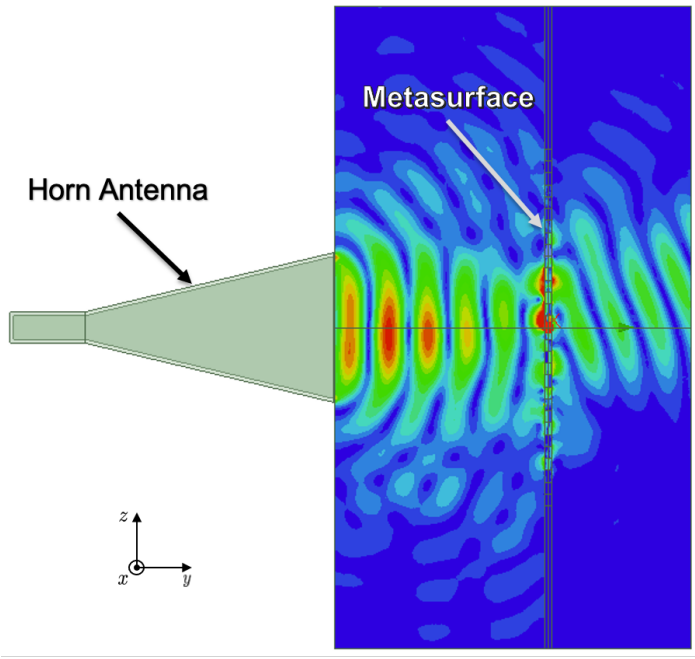


Fig. 3.13 The near-field simulation of the 3D metasurface in the E-plane cut (yz plane), which plots E magnitude at a fixed phase.

outputs a source array so as to meet the desired specifications. This inferred source array, in conjunction with an integral equation based approach, is utilized to find the required tangential fields on the two faces of the metasurface. Finally, a three-layer unit cell topology is used to verify the performance of the metasurface. In all the 2D and 3D examples, the performance of the reconstructed source array is close to the desired specifications. On the other hand, the performance of the resulting metasurface is generally good for the main beam directions, HPBWs and SLLs. However, achieving the desired specifications for nulls and MNLs is typically more challenging. The main advantages of the proposed approach is that (i) the neural network can yield a source array in a fraction of a second, (ii) the user only needs to provide practical performance criteria as opposed to providing full far-field pattern information, and (iii) due to fact that the tangential fields on the output face of the metasurface are obtained from the inferred source array, they are inherently Maxwellian. On the other hand, the developed neural network is designed based on a predefined maximum source array aperture size, which will ultimately limit the achievable directivity.

Chapter 4

Macroscopic Design from 2D Far-Field Masks

Preface

In this chapter, we address one potential drawback of the approach presented in Chapter 3. This is related to the two-step scheme employed in Chapter 3 where the desired far-field (FF) requirements and the LPC constraint are considered separately. Due to this two-step mechanism, the power profile of the aperture field obtained by the DL framework can significantly differ from the power profile of the incident field. This, in turn, will complicate the subsequent LPC optimization step using auxiliary surface waves.

To address this issue, this chapter introduces a single-step metasurface macroscopic design for 2D FF masks, where both the FF requirements and the LPC constraint are considered simultaneously. The materials presented in this chapter are from the journal paper published in *IEEE Antennas and Wireless Propagation Letters* in 2024 [117].[1]

[1]©2024 IEEE. Reprinted, with permission, from **C. Niu**, M. Phaneuf and P. Mojabi, "Transmissive Metasurface Synthesis from Far-Field Masks Using Unsupervised Learning," *IEEE Antennas and Wireless Propagation Letters*, vol. 23, no. 8, pp. 2371-2375, Aug 2024.

Abstract

Designing lossless and passive transmissive metasurfaces requires the knowledge of tangential electromagnetic fields on the metasurface aperture and the enforcement of local power conservation (LPC). In design scenarios where the desired radiation patterns are specified by lower and upper masks, we have developed a deep learning approach to infer the required metasurface aperture fields, while favouring the LPC constraint. In the cases examined here, the aperture fields obtained through the DL approach lead to power patterns that exhibit good alignment with the specified far-field masks, while adhering to the LPC constraint. When simulating the resulting metasurfaces using impedance sheets, this alignment slightly degrades, partly due to the non-zero thickness of the metasurface and the local periodicity assumption used in the unit cell design.

4.1 Introduction

Electromagnetic metasurfaces, consisting of thin subwavelength unit cells, are capable of precise field transformations through the tailoring of their unit cells' properties [15–17]. In pattern synthesis applications, the metasurface design process requires determining the properties of the metasurface from a set of far-field specifications to be achieved. These specifications can take different forms. For instance, they can be in the form of complex (magnitude and phase) field patterns, power patterns, or conventional far-field performance criteria (such as main beam angles, half-power beamwidths, sidelobe levels, and null angles) [75, 121, 71, 114]. In this chapter, the desired far-field specifications are defined by prescribed lower and upper far-field masks within which the resulting power pattern needs to reside. Mask-based design approaches have been developed for antenna array synthesis [180, 72, 181, 112, 182], reflectarray designs [183–185], and metasurfaces [74, 73, 115].

For ease of fabrication, it is desirable for the metasurface to be lossless and passive. In transmissive metasurfaces, which is the focus of this chapter, this necessitates

meeting the local power conservation (LPC) constraint [47]. Given a set of desired far-field specifications, the macroscopic design process [28] for a lossless and passive transmissive metasurface [141, 47, 18] can be broadly classified into (i) two-step and (ii) single-step schemes.

In two-step schemes, an appropriate aperture field is first synthesized to yield the desired far-field specifications. This initial step constitutes an electromagnetic inverse source problem, which can have non-unique solutions [186]. A solution can then be obtained using different methods such as antenna array techniques [110], inversion algorithms [75], and DL approaches [70, 187]. This solution, which is treated as the tangential field on the output face of the metasurface, needs to meet the LPC constraint with respect to the incident field. To this end, the second step incorporates auxiliary surface waves into the metasurface for power redistribution by employing optimization techniques [116, 118, 119] (also, see [114] for a similar approach in reflective metasurfaces). However, the decoupling of these two steps can present the following challenge. The non-uniqueness of the solution in the first step may lead to an aperture field with a power profile significantly different than the incident power profile. Consequently, the optimization algorithm employed in the second step might encounter difficulties in matching these two power profiles through power redistribution via auxiliary surface waves.

On the other hand, single-step schemes directly incorporate the LPC constraint, or an equivalent form of it, into the aperture field synthesis problem. This integration guides the algorithm to single out an aperture field that simultaneously satisfies both the desired specifications and the LPC constraint. Single-step procedures begin by formulating a cost functional and then minimizing it using iterative techniques such as nonlinear conjugate gradient approaches [121, 122]. In most of these formulations, the main unknowns are the equivalent currents of the metasurface, which are related to its tangential aperture fields. (Also, see [73] for a similar approach in metasurface antenna design from far-field masks.) Other formulations of single-step schemes also exist [71, 188]. For example, two sets of unknowns, equivalent currents and metasurface

properties, are considered in [71], which are iteratively updated in an alternating manner to meet convex design goals [71]. This approach has been recently extended to the mask-based design of transmissive metasurfaces [74].

While DL-based approaches have been previously utilized in two-step schemes [70, 187], we are unaware of any DL-based single-step scheme in the design of transmissive metasurfaces. This chapter introduces a DL-based single-step scheme to meet far-field mask-based specifications for two main reasons. Firstly, as will be shown, the proposed approach can outperform a similar DL-based two-step scheme for the reason mentioned above. Secondly, it offers fast computational speed in the design phase.

Finally, it should be noted that this chapter does not consider other metasurface structures such as cavity-based metasurfaces [189] and cascaded pairs [190, 147, 110, 120]. In addition, it does not consider reconfigurable metasurfaces where DL-based approaches can be used to determine the on/off states of diodes from desired far-field masks [115].

4.2 Methodology

We consider a transmissive metasurface that is situated on the xz plane under two-dimensional transverse-electric propagation. The metasurface is invariant in the z direction, and has spatial modulation along the x direction. The input face of the metasurface, $y = 0^-$, is illuminated by the incident field. Since the metasurface should be reflectionless, the tangential fields on its input face are assumed to be the tangential incident fields $(E_z^{\text{inc}}, H_x^{\text{inc}})$. The resulting tangential fields on the output face $(y = 0^+)$ of the metasurface, which are also known as aperture fields, are denoted by (E_z^+, H_x^+). These aperture fields create a far-field pattern which can be described by the azimuth angle $0° \leq \varphi \leq 180°$.

As shown in Fig. 4.1, the neural network of our DL framework requires two inputs: (i) the desired far-field masks and (ii) the incident field impinging on the metasurface. This neural network will then output an antenna array on $y = 0$, referred to as the

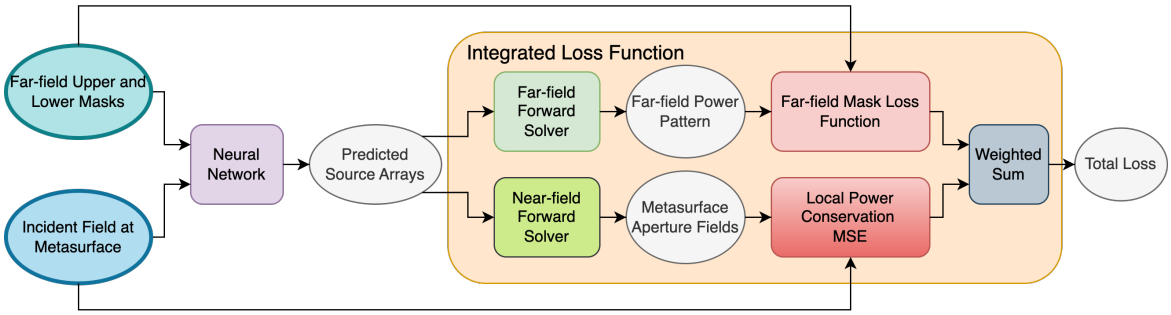


Fig. 4.1 ©2024 IEEE. Forward path of unsupervised learning for the developed mask-based pattern synthesis with LPC (*MBPS-LPC*).

source array herein, which should meet two requirements. Firstly, its far-field power pattern should conform to the prescribed masks. Secondly, its near-field tangential electric and magnetic fields calculated on $y = 0^+$ should meet the LPC constraint. These tangential fields will then be assigned as the tangential fields on the output face of the metasurface, i.e., as (E_z^+, H_x^+), and the source array will be disregarded thereafter. Having found the required tangential fields on the metasurface, three-layer unit cells are designed as in [19].

To this end, we formulate a far-field mask loss function to assess the fitness of predicted power patterns with respect to the given upper and lower masks. This involves the computation of upper (L_U) and lower (L_L) mask losses as

$$L_U \triangleq \frac{1}{n}\sum_{i=1}^{n} \max\Big\{0, P(\varphi_i) - M_U(\varphi_i)\Big\}, \tag{4.1}$$

$$L_L \triangleq \frac{1}{n}\sum_{i=1}^{n} \max\Big\{0, M_L(\varphi_i) - P(\varphi_i)\Big\}, \tag{4.2}$$

where $P(\varphi_i)$ represents the predicted power pattern as a function of azimuth angle φ_i. The index i spans the number of discrete angles ranging from $0°$ to $180°$. Moreover, $M_U(\varphi_i)$ and $M_L(\varphi_i)$ represent the prescribed upper and lower masks as a function of φ_i, respectively. Under an ideal scenario where $M_L(\varphi_i) \leq P(\varphi_i) \leq M_U(\varphi_i)$ for every φ_i, the values of both lower and upper mask losses will be zero. The linear combination of L_U and L_L will then be the output of our mask loss function denoted by L and

calculated as

$$L = \frac{1}{\alpha + \beta}(\alpha L_U + \beta L_L), \tag{4.3}$$

where α and β are the non-negative real numbers providing a relative weight for the upper and lower masks. If one of these variables is set to zero, the enforcement will be either on the lower or upper masks. In our implementation, $\alpha = \beta = 1$.

To integrate the LPC constraint, the neural network is provided with the power profile of the incident field, i.e., $\frac{1}{2}\mathrm{Re}\{E_z^{\mathrm{inc}}(H_x^{\mathrm{inc}})^*\}$ where 'Re' and '$*$' denote the real part and the complex conjugate operators, respectively. This incident power profile will then be compared against the power profile of the metasurface aperture fields, i.e., $\frac{1}{2}\mathrm{Re}\{E_z^+(H_x^+)^*\}$. The discrepancy between these two power profiles is quantified using the mean-squared error (MSE), and is referred to as local power conservation MSE in Fig. 4.1. The total loss is computed using a weight sum over the mask loss and the LPC MSE.

As shown in Fig. 4.1, we have referred to the block containing the forward solvers, mask loss, and the local power conservation MSE as the integrated loss function. However, the naive incorporation of the forward solvers into the integrated loss function fails to provide a backward pass, which is essential for the training process. In the context of deep learning, several strategies can address this issue. The first one is reinforcement learning, where the integrated loss function serves as the so-called environment, and optimization attempts to find the most beneficial action (source array) for the agent (neural network) based on feedback (total loss) from this environment (e.g., similar to [191]). The second approach involves manual gradient calculations, requiring the analytical expressions for the integrated loss function's derivatives, which are then incorporated into the neural network's backward path (e.g., similar to [161, 102]). The third approach employs differentiable approximations, which typically requires training a neural network to approximate and replace the integrated loss function (e.g., similar to [192]). The fourth technique is automatic differentiation, which is widely implemented in deep learning frameworks like TensorFlow and PyTorch. Herein, we implemented the integrated loss function in PyTorch for efficient handling of automatic

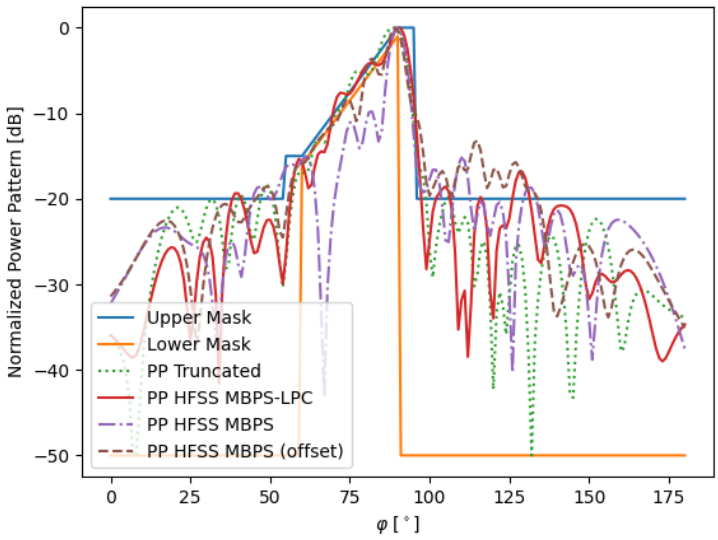


Fig. 4.2 ©2024 IEEE. Desired power pattern (PP) in the shape of a trapezoid and identified by lower and upper masks (first example) along with the achieved far-field power patterns.

differentiation, thus, providing the backward path for the training pipeline of our unsupervised training approach.[2] Finally, we note that our neural network utilizes two one-dimensional convolutional neural networks (CNNs) [193] to process far-field masks and incident fields independently. The outputs of these CNNs are then combined and fed into a multilayer perceptron [150], which predicts a source array.

A training dataset of around 10 million instances was used, comprising a variety of generated masks and incident fields. This includes single-square, multi-square, trapezoidal, and cosecant-squared masks, which are randomly generated in real-time. Concurrently, different incident fields were obtained from up to three line sources placed randomly along the x-axis at one wavelength (λ) separation from the metasurface. The developed neural network is engineered to synthesize source arrays sized at 12λ, consisting of 61 uniformly spaced elements. The fields on the output aperture of the metasurface, denoted by $y = 0^+$ above, are then found via forward propagating the fields of the source array by 0.2λ in the direction perpendicular to the source array.

[2]For further information, see Appendix C.

4.3 Results

We present two examples which consider omega-bianisotropic Huygens' metasurfaces at 10 GHz, with an aperture size of 15λ along the x axis. Each metasurface consists of 91 unit cells, having a unit cell size of about $\lambda/6$. A line source, located λ away from the input face of the metasurface and centered with respect to it, illuminates the metasurface. Each unit cell consists of three layers of purely reactive surface impedances that are connected using two identical dielectric substrates as in [19]. Herein, we do not consider the copper trace implementation of these impedances and model them as impedance boundary conditions in Ansys HFSS. The substrates are assumed to be Rogers RO3010 with relative permittivity of 10.2, loss tangent of 0.0035, and thickness of 0.25 mm. The reported pattern is obtained by the near-field to far-field transformation of the metasurface aperture fields using Ansys HFSS.

The goal of the first example is to design a metasurface that transforms the line source excitation into a far-field power pattern that has a trapezoidal beam shape. This desired trapezoidal shape has been identified with the lower mask (orange) and upper mask (blue) as shown in Fig. 4.2. The upper mask of the trapezoidal beam begins at $\varphi = 95°$ with a 0 dB level and ends at $\varphi = 55°$ with a -15 dB level. The lower mask allows for 1 dB peak to peak ripples from $60° \leq \varphi \leq 90°$. (Note that $\varphi = 90°$ is perpendicular to the metasurface according to the coordinates shown in Fig. 4.3.)

The source array generated by the neural network is then used to find the desired tangential fields on the output face of the metasurface. The power pattern (PP) generated by these tangential fields, denoted by *PP Truncated*[3] in Fig. 4.2, mostly aligns with the upper and lower masks. Having obtained the required tangential fields, the unit cells are designed and the metasurface is simulated in Ansys HFSS. The power pattern of this metasurface is denoted by *PP HFSS MBPS-LPC* in Fig. 4.2, where MBPS-LPC stands for mask-based pattern synthesis with LPC. This achieved power

[3]We assume that the fields outside the 15λ aperture are zero, hence the term "truncated". This *PP Truncated* therefore takes into account the finite aperture size, and serves as an upper bound for the metasurface's performance.

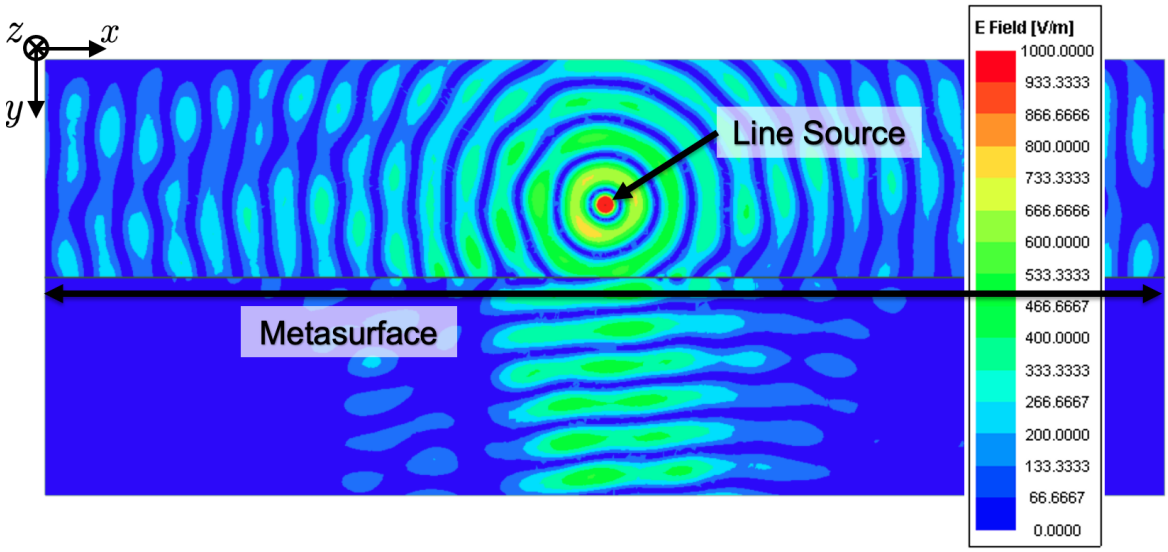


Fig. 4.3 ©2024 IEEE. Plot of the electric field magnitude at a fixed phase in Ansys HFSS for the first example using the MBPS-LPC. The coordinate center is located at the center of the metasurface.

pattern is almost compliant with the specified masks; however, its overall performance is worse than that of *PP Truncated*. This is expected due to the finite thickness of the metasurface and the local periodicity assumption [194] made in the design of unit cells. The near-field simulation for this case is shown in Fig. 4.3, which has a transmission power efficiency of 88.12%.

Finally, to demonstrate the advantage of this single-step DL-based scheme over its two-step counterpart, we also found a metasurface aperture field using the DL approach *without* the LPC constraint. Subsequently, an auxiliary surface wave optimization method [119] was utilized to enforce the LPC constraint, which updated the tangential fields on the metasurface. The simulation of the resulting metasurface led to an unsatisfactory power pattern denoted by *PP HFSS MBPS* in Fig. 4.2. To help this two-step scheme achieve the desired masks, we tuned the location of the line source by shifting it 3.7λ in the positive x direction so that the incident power profile becomes more aligned with the power profile of the aperture fields. Using this new excitation, we were able to use this two-step scheme to design a metasurface with a reasonable power pattern denoted by *PP HFSS MBPS (offset)* in Fig. 4.2. However, this power pattern

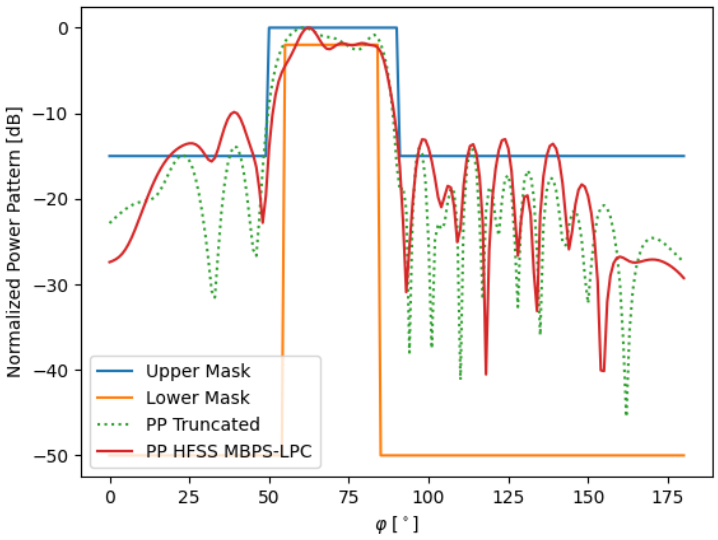


Fig. 4.4 ©2024 IEEE. Desired square-shape power pattern (PP), along with the achieved far-field power patterns for the second example.

is still not as good as that obtained via the proposed single-step scheme (*MBPS-LPC*), and also has a lower transmission power efficiency of 71.25%.

For the second example, we design a metasurface to transform the original centered line source into a power pattern with a square-shape beam while meeting the LPC constraint. This desired shape is defined by upper and lower masks as shown in Fig. 4.4. As can be seen in this figure, *PP Truncated*, i.e., the far-field power pattern due to the reconstructed aperture fields, lies within the prescribed masks. The pattern of the resulting metasurface denoted by *PP HFSS MBPS-LPC* in Fig. 4.4 mostly meets the desired masks except at the left sidelobe. The near-field simulation of this metasurface is shown in Fig. 4.5, which exhibits a transmission power efficiency of 90.23%.

The range of the simulated reactances (absolute values) in these two examples (MBPS-LPC) are from about 0.1 to 184.7 [Ω] as shown in Fig. 4.6. The small reactance values may not be practical for physical implementation. Therefore, approaches such as using four-layered unit cells [118], adding a constant phase to the obtained aperture fields [110], and limiting the range of reactances [73] within the DL framework can be pursued in the future. Alternatively, we can convert the obtained unit cell properties

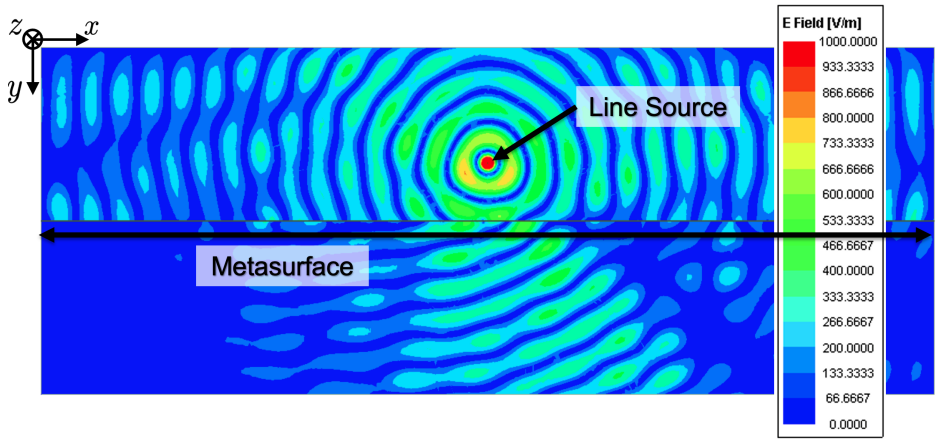


Fig. 4.5 ©2024 IEEE. Plot of the electric field magnitude at a fixed phase in Ansys HFSS for the second example.

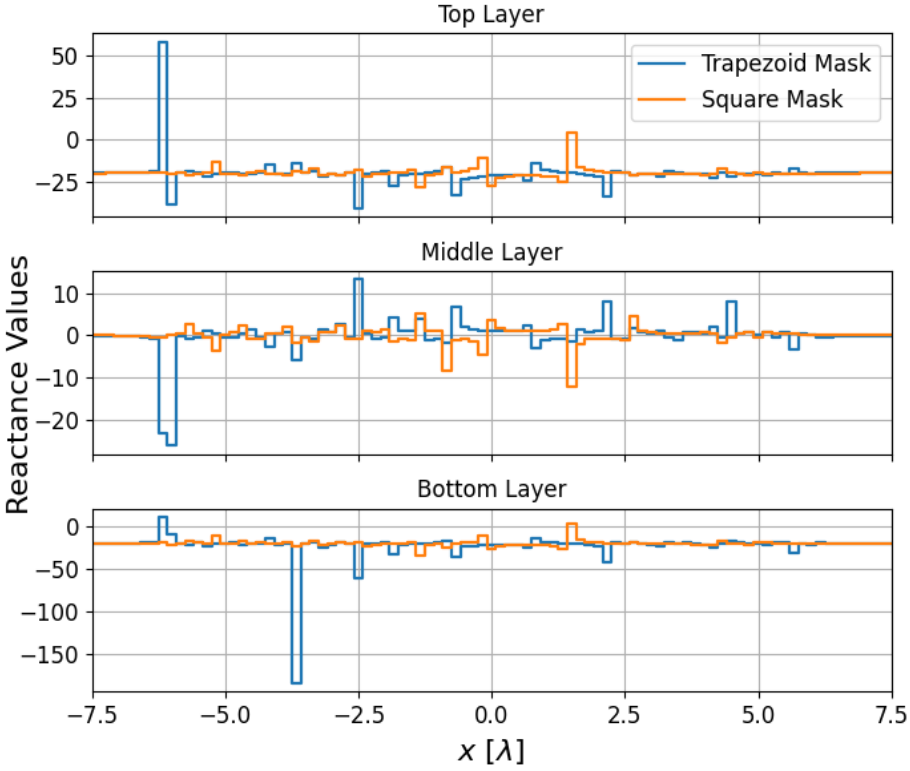


Fig. 4.6 ©2024 IEEE. The reactance values [Ω] of the three-layered metasurface design for the trapezoid and square-shaped masks.

into two-port scattering parameters, and then aim to realize these scattering parameters instead of individual reactance values. In conjunction with the proposed single-step DL-based scheme (MBPS-LPC), a preliminary investigation into this physical design approach was recently pursued in a conference paper [195], where the two-port scattering parameters were passed to a physical unit cell design approach [129] to generate three-layer unit cells featuring dogbone-shaped metallic traces. This preliminary work shows potential in creating physically realizable transmissive metasurfaces to meet mask-based constraints [195]. The mask-based compliance level achieved using three-layer unit cells with ideal impedance sheets represents the upper limit of performance for such physical designs.

4.4 Conclusion

The proposed single-step DL framework is generally able to yield lossless and passive designs with far-field power patterns within the desired masks. The degradation of the results is partly due to the fact that the DL framework does not prioritize smoother field variations to align with the local periodicity assumption used in the unit cell design, nor does it consider the finite thickness of the substrates. The main advantage of the proposed DL framework lies in its computational efficiency and its robustness compared to its two-step counterparts. The neural network takes only a few milliseconds ($\approx$ 2.6 ms on NVIDIA Tesla V100) to generate its output, which can then be easily used to find the required properties of the unit cells.[4]

[4]In contrast to Chapter 3, which considered a two-step scheme, this chapter has focused on a single-step scheme to handle desired far-field masks. If two-step schemes are used for handling far-field masks, both auxiliary surface waves or cascaded metasurfaces may be used to enforce the LPC constraint. Since auxiliary surface waves, described in Section 3.6, has already been used in Chapter 3, we present an example of using cascaded metasurfaces in Appendix A. Note that this footnote and its associated appendix was not included included in our IEEE AWPL paper [117].

Chapter 5

Macroscopic Design from 3D Far-Field Masks

Preface

The previous two chapters considered the macroscopic design of transmissive metasurfaces from desired FF requirements. This chapter, which will be the last chapter dedicated to the macroscopic design aspect, builds upon the work from Chapter 4 to address 3D far-field masks. The materials presented in this chapter are from the conference paper presented in 2024 [196].[1]

Abstract

For transmissive metasurfaces, we develop a deep learning approach that takes user-defined three-dimensional (3D) far-field masks and the incident field impinging on the metasurface as input. This DL approach then aims to infer the required aperture fields of the metasurface such that its resulting power pattern resides within the 3D lower and upper masks and the metasurface remains lossless and passive.

[1] **Chen Niu** and Puyan Mojabi, "Synthesis of Aperture Fields for Transmissive Metasurface Design from 3D Masks Using Deep Learning," *URSI Atlantic Radio Science Meeting (AT-RASC)*, pp. 1-4, Gran Canaria, Spain, May 19-24, 2024.

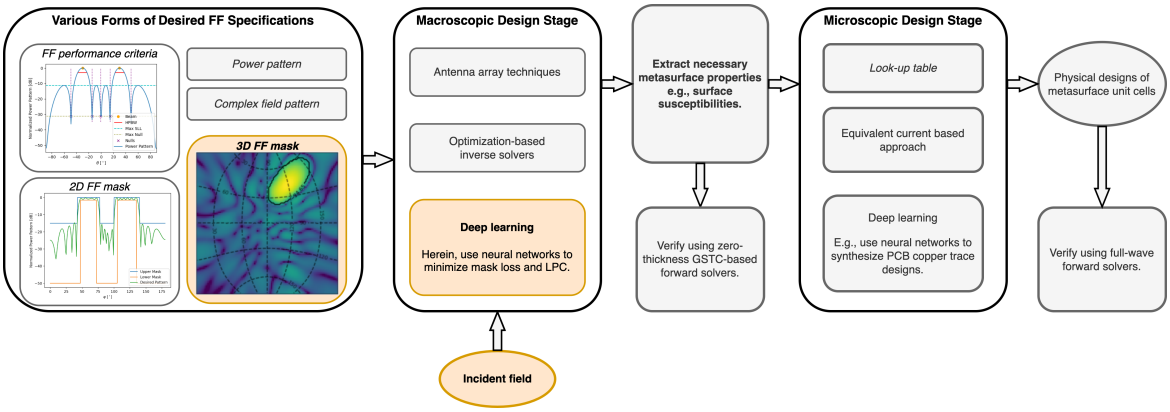


Fig. 5.1 A transmissive metasurface design pipeline. The highlighted parts in orange are the focus of this chapter.

5.1 Introduction

Electromagnetic (EM) metasurfaces are thin engineered structures composed of subwavelength unit cells. Metasurfaces have the capability to systematically manipulate EM wavefronts in terms of their phase, amplitude, and polarization profiles [14]. Their subwavelength unit cells are typically created by (multi-layered) patterned metallic traces supported by dielectric substrates. In this context, the metasurface design problem can be cast as an inverse scattering problem where the goal is to retrieve the shape and location of metallic traces based on the desired response [197]. Due to the significant computational complexity of such an inverse scattering approach, the design process is often divided into two steps: macroscopic and microscopic stages [28].

The macroscopic stage begins by specifying the desired response. For far-field (FF) responses, these can for example include performance criteria such as main beam directions, half-power beamwidths, null locations, and sidelobe levels [70, 71]. Alternatively, the desired response can be expressed through FF lower and upper masks within which the resulting power pattern (PP) must reside [72–74]. Moreover, the desired response may be described by complex field patterns for simultaneous amplitude and phase specification [75]. Given a desired response, techniques such as gradient-based inverse solvers or deep learning approaches can be employed to determine the necessary aperture fields of the metasurface. These techniques often

require to account for metasurface-related constraints. For example, to arrive at lossless and passive transmissive metasurfaces, the aperture fields need to meet the local power conservation (LPC) constraint at each unit cell [47]. Once the aperture fields are determined, they are employed in conjunction with the generalized sheet transition conditions (GSTCs) [21] to calculate the required metasurface properties such as its surface susceptibility profiles.

The microscopic stage involves unit cell design to realize the aforementioned surface susceptibilities. To accomplish this, one may employ techniques such as look-up tables [28] or deep learning approaches [129]. These are typically formed or trained using periodic boundary conditions, which can lead to sub-optimal performance for inhomogeneous (aperiodic) metasurfaces. Thus, new techniques have been proposed which use equivalent currents instead of periodic boundary conditions in unit cell design [79]. To summarize the macroscopic and microscopic stages, the design pipeline has been illustrated in Fig. 5.1.

In [195], we have demonstrated an end-to-end DL framework for designing lossless and passive transmissive metasurfaces from desired 2D FF masks (i.e., masks in a single FF cut). The work in [195] takes 2D FF masks and the incident field impinging on the metasurface as input and then produces physical three-layer unit cells with dogbone-shaped metallic traces. To achieve this, it utilizes unsupervised learning for the macroscopic design [117] and a generative neural network for the microscopic design [129]. Herein, we expand the macroscopic stage of [195, 117] to the 3D realm so that it can handle FF masks in the 3D space.

5.2 Methodology

Based on our 2D work presented in [195, 117], we create a DL-based macroscopic metasurface design pipeline as shown in Fig. 5.2. This framework consists of three key components: (i) 3D FF masks and the incident field as input, (ii) a neural network, and (iii) an integrated loss function that evaluates the fitness of the solution with respect

to the desired 3D mask and the LPC constraint. Let us begin by discussing how we represent our desired 3D masks. Assume that the metasurface is positioned on the xz plane. It is illuminated from $y<0$ region and is then supposed to create a PP in the $y>0$ region. The PP can then be described on a hemisphere in the $y>0$ region with azimuthal angle $0^\circ \leq \varphi \leq 180^\circ$ and elevation angle $0^\circ \leq \theta \leq 180^\circ$. (The equatorial plane of this hemisphere of interest will then be the xz plane.) This PP must reside within the specified 3D lower and upper masks. For 3D mask-based pattern synthesis applications, the uv coordinate system has been commonly used [161].[2] This approach is effective when operating within the central region of the hemisphere where distortion is minimal. However, at the edges of the hemisphere, the use of the uv coordinate system could introduce distortion. To address this, we utilize a global map projection technique known as the Peirce quincuncial projection [198, 199]. This projection was designed to map the globe onto a square, and is also employed for presenting spherical panoramas. Utilizing this technique to represent our 3D FF masks on a hemisphere can be effective in preserving information around the edges of the FF hemisphere. To illustrate this potential advantage, let us consider 27 identical circular cones with the cone half angle of 5° intersecting with our hemisphere of interest. Fig. 5.3 displays the projection of this hemisphere using both the uv projection and the Peirce quincuncial projection. The usable area of the projection is shaded in purple. (For the uv projection, the green-colored corners are not usable, while the Peirce quincuncial projection utilizes the entire square.) As can be seen, in the Peirce quincuncial projection, all 27 circles are distinctly visible with minimal size distortion. However, in the uv projection, only 23 circles are visible, and those at the edges suffer distortion.

Let us now focus on the neural network aspect of this work, which extends [117] to 3D. The neural network is formulated as if we are designing a planar array of infinitesimal dipole antennas, which we refer to as source arrays. The desired source

[2]For a hemisphere located in the $z > 0$ region, we define uv coordinates as $u = \sin(\theta)\cos(\varphi)$ and $v = \sin(\theta)\sin(\varphi)$. In our case, the hemisphere of interest is located at $y > 0$, leading to $u = \cos(\theta)$ and $v = \sin(\theta)\cos(\varphi)$. Essentially, this involves projecting the hemisphere directly onto the equatorial plane, which in our case is the xz plane.

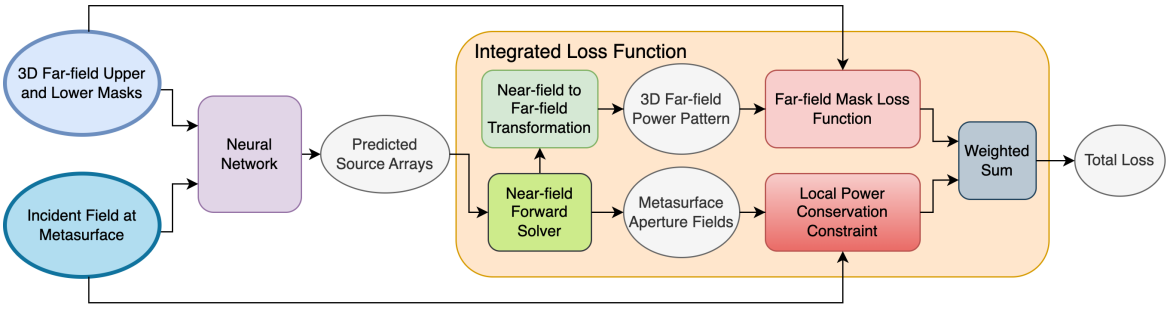
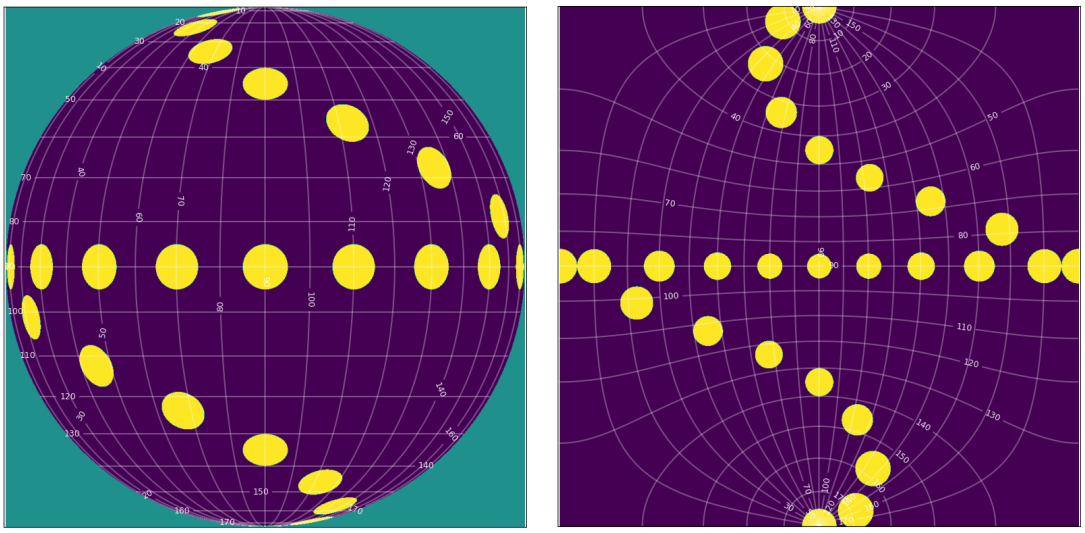


Fig. 5.2 DL-based metasurface macroscopic design pipeline from 3D masks (based on the 2D work described in Chapter 4).

array needs to satisfy two conditions: (i) its FF pattern must reside within the specified 3D mask, and (ii) its tangential near fields must meet the LPC constraint with respect to the incident field. These near fields are calculated on a parallel plane with the same size as the intended metasurface, situated at 0.2λ away from the source array. (The wavelength is denoted by λ.) The incident wave is considered to be a plane wave travelling in the $+\hat{y}$ direction and normally incident on the metasurface. Once this source array is obtained, its tangential near fields are assigned as the metasurface aperture fields and the source array is disregarded thereafter. These tangential fields can be used with the GSTCs to derive the surface susceptibility profiles.

To train our neural network, approximately one million FF masks featuring ellipses with varying sizes and orientations, covering azimuthal angles (φ) within the range of 15° to 165° and elevation angles (θ) between 15° and 165° were randomly generated. To better understand this, let us consider the upper and lower masks depicted by two *solid* elliptical contours in Fig. 5.4. (The dashed contours in this figure denote fixed φ and θ angles.) The outer ellipse represents the upper mask. Since we are dealing with normalized patterns, the maximum power level within the internal region of the outer ellipse is 0 dB. External to this outer ellipse, the power level should be below the value indicated on the ellipse contour, which is −15 dB. On the other hand, the inner ellipse represents the lower mask with the indicated contour value being −1 dB. This implies that the power level is allowed to vary between 0 dB to −1 dB in the internal region of the inner ellipse. The gap between these two ellipses represents the space where



(a) (b)

Fig. 5.3 Projection of the hemisphere of interest with 27 circular cones onto a plane using the (a) uv coordinates and (b) Peirce quincuncial projection coordinates, which we refer to as $x|_{\text{Peirce}}$ and $y|_{\text{Peirce}}$. In both cases, the somewhat vertical contours indicate fixed φ angles, and the somewhat horizontal contours indicate the fixed θ angles.

the PP should transition from its main lobe to a sidelobe level below -15 dB. Herein, our designed neural network creates a $5.5\lambda \times 5.5\lambda$ source array, comprising 34×34 infinitesimal $\hat{z}$-polarized dipoles. This source array is then used to determine the necessary tangential fields of a metasurface with an aperture size of $5\lambda \times 5\lambda$ uniformly discretized into 61×61 unit cells.[3]

5.3 Results

In the first example, the inner ellipse (lower mask) is centered at $\theta = 80°$ and $\varphi = 110°$. Its major and minor axes are chosen to span $40°$ and $36°$ respectively. The outer ellipse (upper mask) extends the inner ellipse by $5°$. These ellipses are then rotated by $45°$

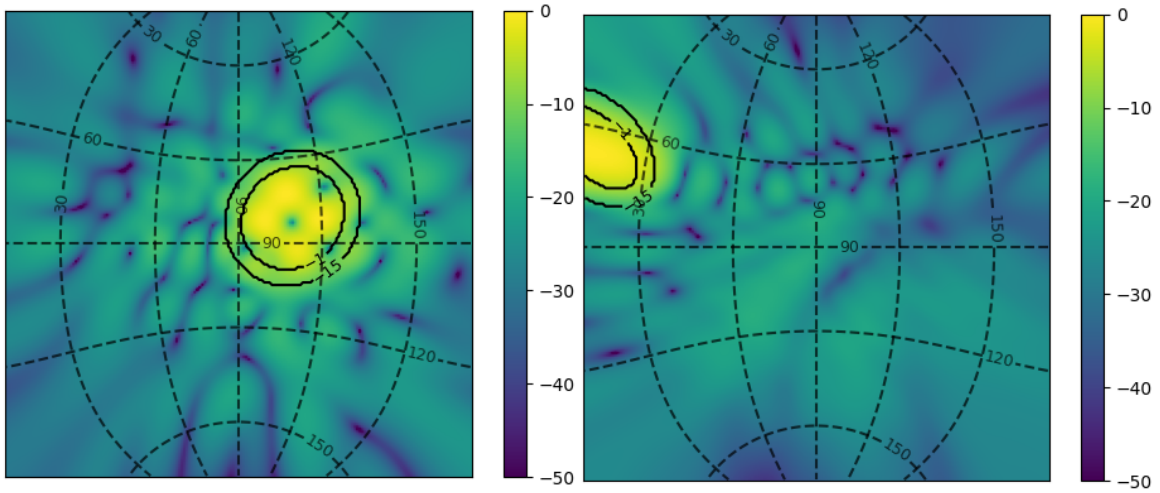
[3]In [117, 195], we set our source array to be smaller than the metasurface aperture. This was due to the fact that our FF pattern was calculated directly from the source array. Thus, the metasurface aperture had to be larger than the source array to capture most of its radiated energy. Consequently, the FF pattern of the metasurface remained close to that of the source array. In this chapter, we have revised our integrated loss function to compute the FF pattern directly from the aperture fields, as illustrated in Fig. 5.2. This has removed the above aperture size constraint. Note that the source array is disregarded after determining the aperture fields.

counterclockwise as depicted in Fig. 5.4(a). In the second example, the inner ellipse is centered at $\theta = 30^\circ$ and $\varphi = 30^\circ$ with major and minor axes spanning 40° and 20° respectively. The outer ellipse is created the same as the first example, and the two ellipses are rotated by 135° counterclockwise. This is shown in Fig. 5.4(b) as two truncated ellipses as parts of them are situated outside the angular range defined in Section 2.

Let us now discuss the PP generated by our neural network. The colorbar in Fig. 5.4 represents the normalized PP in dB. As can be seen, the PP have hot-spots in the intended regions. To better visualize how PP follow the masks, we have also plotted the 3D version of Fig. 5.4 in Fig. 5.5 where the lower and upper masks are identified with blue and green colors respectively. The PP, plotted in red, mostly resides within the upper and lower masks. Finally, we need to check if the achieved metasurface aperture fields meet the LPC constraint. To this end, we plot the normalized LPC mismatch associated with each of the unit cells in Fig. 5.6. (The mismatch is the normalized absolute difference between the incoming and outgoing normal power densities at each unit cell. The mismatch value ranges from zero, i.e., perfect match, to one, i.e., 100% mismatch.) As can be seen, this mismatch is mostly small; however, we have some unit cells with significant LPC mismatch. The average mismatch for these two examples are 21% and 18% respectively which can be handled by employing methods such as auxiliary surface waves for further power redistribution [70].

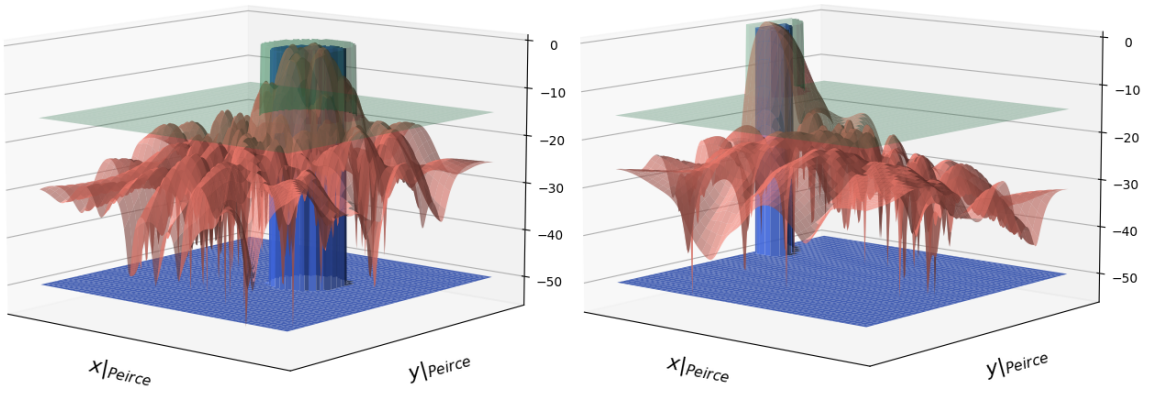
5.4 Conclusion

The macroscopic design of metasurfaces requires the knowledge of the metasurface aperture fields. To this end, we have presented a deep learning approach to infer the necessary aperture fields based on desired 3D far-field masks and the incident field illuminating the metasurface. The power patterns resulting from this technique mostly followed the desired 3D masks. Although the LPC constraint was successfully met at



(a) (b)

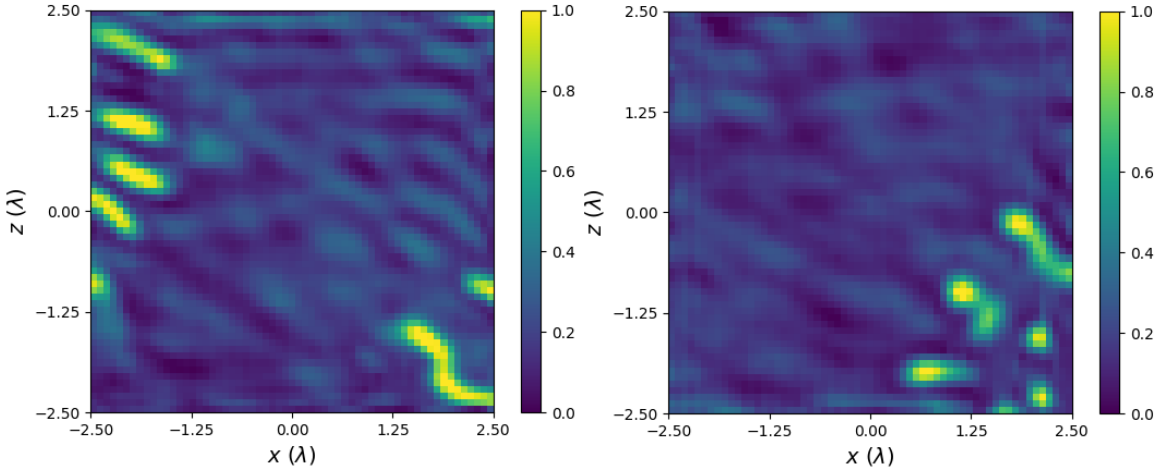
Fig. 5.4 Desired elliptical masks (solid contour) in Peirce quincuncial projected domain ($x|_{\text{Peirce}}$ and $y|_{\text{Peirce}}$ plane) and the achieved normalized power pattern in color for the (a) first and (b) second examples.



(a) (b)

Fig. 5.5 Desired upper and lower masks (green and blue) and the achieved PP (red) shown in Fig. 5.4 but plotted in 3D. (a) First and (b) second examples.

most unit cells, the proposed technique was not able to enforce it at all unit cells. This needs to be treated prior to pursuing physical design.



(a) (b)

Fig. 5.6 Local power conservation mismatch at each unit cell for the (a) first and (b) second examples. The zero value means perfect match and the unity indicates 100% mismatch.

Chapter 6

Microscopic Design from Far-Field Specifications

Preface

In the previous three chapters, we discussed the macroscopic design of metasurfaces. In this chapter, we shift our focus to the microscopic design stage. To this end, we develop a DL-based framework capable of generating three-layer metallic trace designs for unit cells based on the required macroscopic surface properties. (These macroscopic properties can be obtained from the methods discussed in the previous chapters.) We also experimentally evaluate the performance of one of our designs in a compact antenna test range. The materials presented in this chapter are from the journal paper published in *IEEE Open Journal of Antennas and Propagation* in 2023 [129].[1]

Abstract

A deep learning approach based on a diffusion model is proposed to yield metasurface unit cell designs. This method takes desired two-port scattering parameters along with

[1] **C. Niu**, M. Phaneuf, and P. Mojabi, "A Diffusion Model for Multi-Layered Metasurface Unit Cell Synthesis," *IEEE Open Journal of Antennas and Propagation*, vol. 4, pp. 654-666, 2023.

the frequency of operation in an attempt to synthesize three-layered metasurface unit cells. The core of this approach lies in casting the three-layered unit cell synthesis process as conditional three-channel binary image synthesis. The conditions are governed by the desired scattering parameters at a given frequency whereas the binary nature implies the presence and absence of metallic traces. Once synthesized, these unit cells are placed beside each other, without any further tuning, to form the final metasurface. The performance of the whole metasurface, for three different design scenarios, is then tested against full-wave simulation and/or experimental data.

6.1 Introduction

Electromagnetic (EM) metasurfaces are thin artificial materials that can tailor incoming EM waves into desired outgoing EM waves [15]. This tailoring is achieved through the interaction of the incoming EM waves with the subwavelength constituents of the metasurface known as unit cells. The metasurface design procedure can be divided into macroscopic and microscopic steps [28]. In the macroscopic step, the required tangential EM fields on the faces of the metasurface are obtained and then used in conjunction with the generalized sheet transition conditions (GSTCs) [21] to calculate the required surface properties of the metasurface (e.g., surface susceptibilities). This step typically begins by the designer specifying some desired performance criteria such as main beam directions, null angles, and sidelobe levels to be realized by the EM wave transformation performed by the metasurface. These performance criteria will then be used to infer the required tangential fields on the faces of the metasurface using various techniques such as solving an inverse problem [75, 121, 71, 114], antenna array techniques [110], or machine learning approaches [70]. The macroscopic step can be made to be fully automated, which provides a significant advantage for radiation pattern synthesis using metasurfaces.

On the other hand, the microscopic step deals with the physical design of unit cells so as to realize the required surface properties obtained in the macroscopic step. At

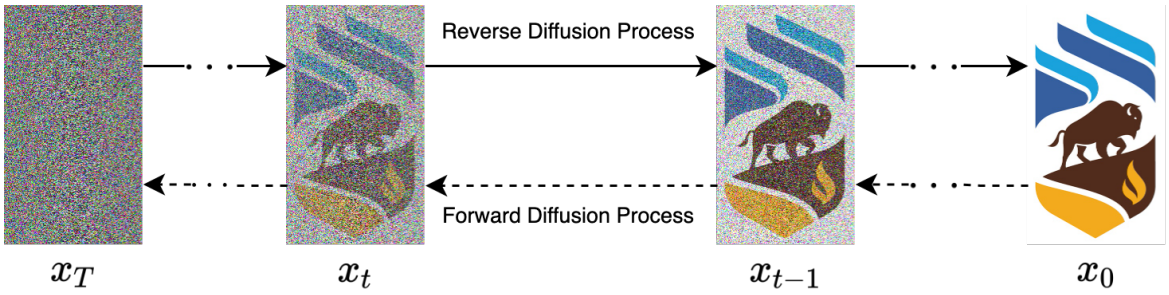


Fig. 6.1 An illustration of the diffusion process for diffusion models. The dashed and solid arrows represent the forward and reverse diffusion processes, respectively.

microwave and mm-wave frequencies, this is mostly achieved by patterned metallic claddings supported by dielectric substrates [141, 18, 47]. To this end, look up tables are often formed through running a large number of full-wave unit cell simulations, which record the scattering response of subwavelength scatterers (e.g., copper dogbone shaped inclusions [123]) of varying dimensions. However, lookup tables are typically formed for single-layer metallic scatterers, and thus do not take into account the near-field coupling in multi-layered metasurface design scenarios. We note that multi-layered designs (e.g., employing three-layer [18] or four-layer [118] unit cells) are important as they provide more degrees of freedom to enable functionalities such as reflectionless wide-angle refraction [45]. To handle this issue, methods such as [125] are proposed to incorporate the near-field coupling effect in the unit cell design procedure. In particular, the method presented in [125] augments the basic unit cell circuit model [141] by current-controlled voltage sources in an attempt to take into account mutual coupling. Although these techniques have shown promising results, they may require further optimization and fine tuning in a full-wave forward solver, which can be time consuming. Thus, as opposed to the macroscopic design, the microscopic design step is not yet fully automated. Therefore, it is useful to provide the microscopic design step with an appropriate initial guess, which is the focus of this chapter.

To provide an appropriate initial guess for the microscopic design, we employ a deep learning approach. deep learning has emerged as the state-of-the-art solution for numerous challenges in the computer science domain, including computer vision [200], natural

language processing [167], and generative tasks such as text-to-image synthesis [89]. The application of DL techniques, particularly generative modeling, to metasurface design may effectively enable the end-to-end metasurface design process, i.e., beginning the design from desired performance criteria and arriving at the required patterned metallic claddings. In fact, generative DL techniques have been already utilized in the synthesis of single- and multi-layered metasurfaces at both microwave [201–205], and visible/near infrared spectra [206, 207]. The existing generative DL techniques for metasurface design mainly consist of generative adversarial networks (GAN) [91, 126, 108], variational auto-encoders (VAE) [208, 78, 209], and combinations of DL techniques with conventional optimization techniques [210, 211].

Herein, we employ a DL approach based on diffusion models [212] for multi-layered unit cell designs. To the best of the authors' knowledge, this is the first time that diffusion models have been used for this purpose. Currently, diffusion models are state-of-the-art techniques for image synthesis applications. Casting multi-layered unit cells as multi-channel images, we investigate the performance of diffusion models to generate unit cells from desired scattering parameters. (Note that the desired scattering parameters can be easily calculated from the required surface properties of the metasurface [131, Ch. 4], which are obtained in the macroscopic design step.) The main advantage for using diffusion models as compared to GANs is to avoid training instabilities [127]. In addition, diffusion models offer more model complexities compared to the VAE, thus, having the potential for better performance. The performance of the proposed deep learning approach is evaluated using both full-wave simulations and experimental data. To this end, we designed three metasurfaces under three different design scenarios by utilizing the unit cells obtained directly from the proposed deep learning approach; i.e., no further optimization has been applied to these unit cells. Thus, the deep learning method presented in this chapter can be viewed as a means to generate initial guess designs, which can be further tuned using full-wave solvers. In what follows, we implicitly assume a time dependency of $\exp(j\omega t)$ where ω and t

denote the angular frequency and time respectively. Finally, we note that this chapter is the extension of our conference paper [213].

6.2 Scope and Problem Statement

Herein, we focus on the design of omega-bianisotropic transmitting Huygens' metasurfaces [214]. These metasurfaces can be implemented by employing three-layered unit cells. The unit cells considered herein employ dogbone shaped metallic traces [18, 123] in the X-band frequency range (8.2 GHz to 12.4 GHz). These metasurfaces can offer electric-to-electric, magnetic-to-magnetic, and magnetoelectric surface susceptibilities [18], thus enabling important functionalities such as wide-angle perfect refraction [45]. We further limit ourselves to lossless and passive unit cells, except for the inevitable loss of the dielectric substrate and the fabricated copper traces. It should be noted that the above scope is merely used to limit the computational complexity associated with the data generation process for neural network training purposes. The proposed DL approach can therefore be extended to go beyond this scope.

The objective of this chapter can be described as follows. Given a desired EM wave transformation and an EM incident field, we obtain the required two-port scattering parameters for each metasurface unit cell. These scattering parameters along with the frequency of operation is fed to the proposed DL diffusion model, which then yields dogbone-shaped three-layered unit cell designs. These unit cell designs are then combined to form the whole metasurface so as to evaluate its performance against the desired transformation. These metasurfaces have spatial modulations in one direction along the metasurface plane and are fully periodic in the other direction; thus, they will be evaluated under two-dimensional (2D) transverse magnetic (TM) or transverse electric (TE) wave propagation.

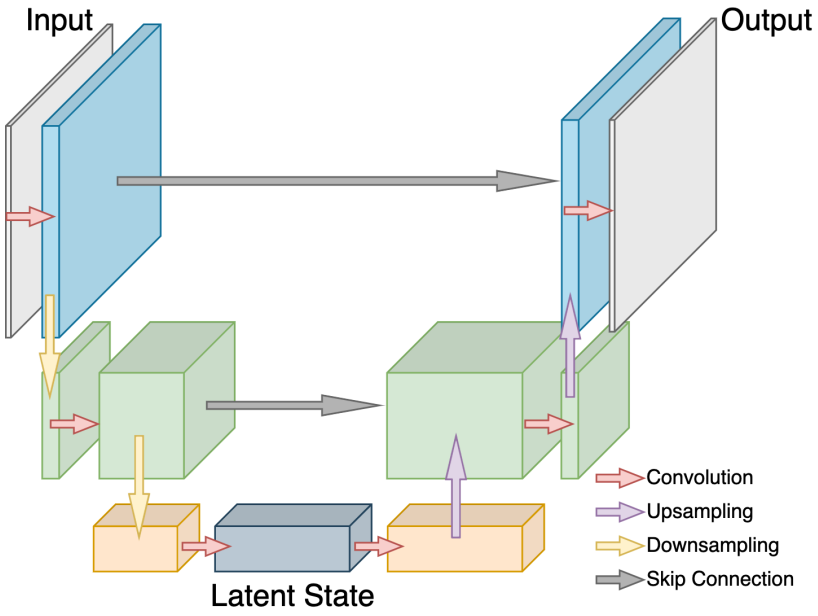


Fig. 6.2 An illustration of the U-Net architecture.

6.3 Background and Methodology

The DL models developed in this work are largely inspired by the conditional image synthesis in the field of computer science. In particular, we have modified the latent diffusion model (LDM) [215] to be able to synthesize metasurface unit cells. The structure of these unit cells includes three-layered patterned metallic claddings supported by thin dielectric substrates. These metallic claddings are modelled as perfect electric conductors (PEC) in the data generation process, and thus are referred to as PEC traces in this chapter. These three-layered PEC traces are to be designed from the knowledge of required scattering parameters. Since the unit cells are treated as two-port networks, the desired scattering parameters will be a 2×2 complex matrix. In the context of conditional image synthesis, these scattering parameters (along with the frequency of operation) can be viewed as the conditions for synthesizing images, which are the unit cell's PEC traces. The LDM belongs to a family of generative techniques called diffusion models [216, 212, 127]. In this section, we discuss the background and implementation of the diffusion models and in particular LDM.

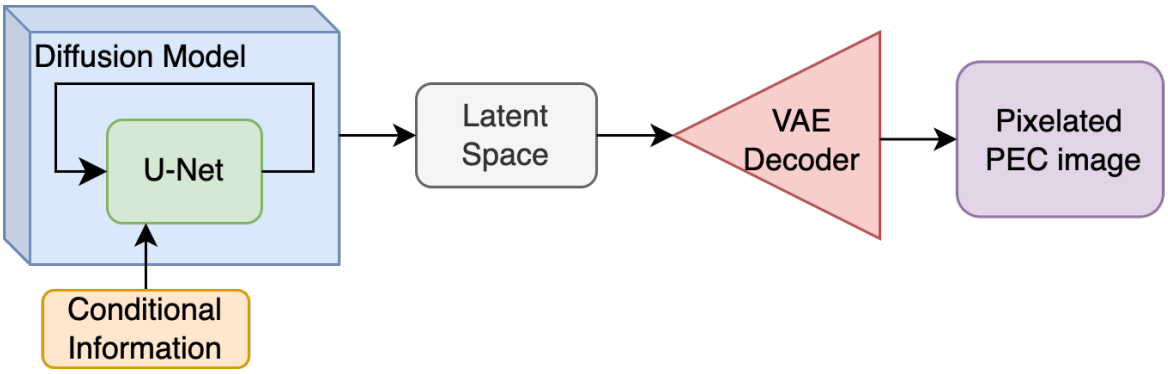


Fig. 6.3 An illustration of the latent diffusion model (LDM). The conditional information, in our case, is the desired two-port scattering parameters along with the frequency of operation.

6.3.1 Diffusion Models (DM)

Diffusion models are a type of generative models that are inspired by non-equilibrium thermodynamics [216]: molecules diffuse from high-density to low-density areas due to entropy. This can be viewed as losing the information associated with their initial state. This is conceptually similar to an aspect of information theory where information within the data can be lost due to gradual intervention of noise. The key concept behind diffusion models is to build a machine learning system that is capable of recovering the information back from noise by learning the systematic information decay. Diffusion models consist of forward and reverse processes. The forward diffusion process defines a Markov chain of diffusion steps to slowly add random noise to data. More importantly, the reverse diffusion process aims to learn to gradually remove noise so as to yield the data of interest. In the forward diffusion process, given the data x_0 and a noising mechanism, we progressively generate the intermediate noisy data, say from x_1 to x_T, through a recursive process [216]. The generative capability of the diffusion models comes from the reverse diffusion process that attempts to recreate the original data x_0 from x_T. Therefore, the goal of the diffusion models is to train a neural network that is able to "denoise" back in time, i.e., from x_T to x_0. The forward and reverse processes of the diffusion models are shown in Fig. 6.1[2].

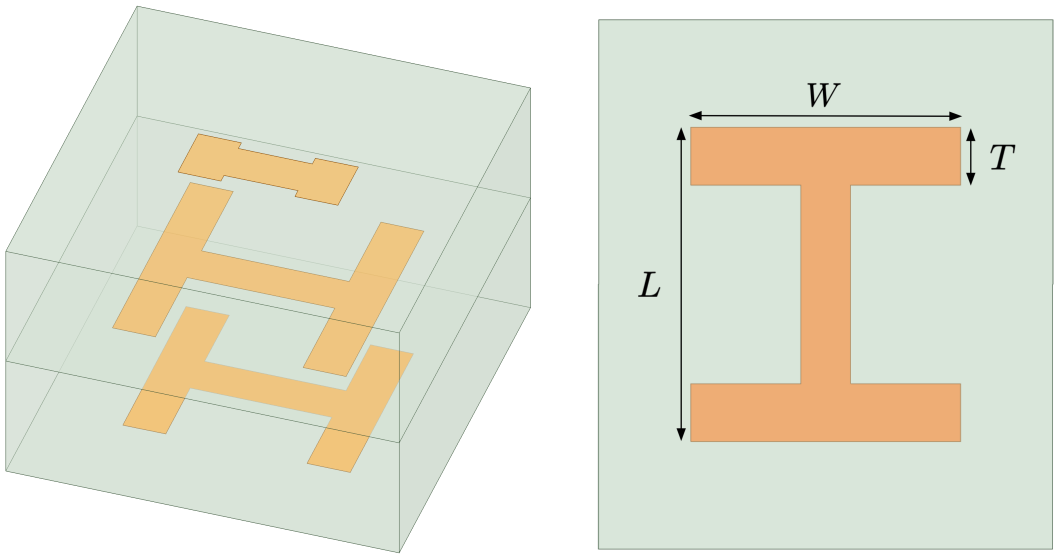
[2]This figure is created based on [127].

In general, the diffusion model is designed to generate random data from noise, e.g., generating random human faces. However, this is not useful for our unit cell design application since the generation of PEC traces needs to be associated with the desired two-port scattering parameters. To this end, with minor modifications, diffusion models are able to synthesize data samples based on conditional information such as class labels (e.g., white male human face) or in our case based on the desired scattering parameters and the frequency of operation. This is achieved by injecting this additional information within each reverse diffusion step. The reverse diffusion process can be implemented by different neural networks. Herein, we utilize the U-Net [217] as the neural network of choice.

6.3.2 U-Net

U-Net [217] was originally designed for image segmentation tasks for medical imaging where the goal is to classify each pixel of the input image to produce a segmentation map. A basic structure of the U-Net is shown in Fig. 6.2[3]. The U-Net distills the input image to a latent state through a series of downsampling steps and constructs the segmentation map (the output) through a series of upsampling steps. The key innovation behind the U-Net is the skip connections that link each stage of downsampling to its upsampling counterpart. These skip connections allow the gradients to have the chance to flow directly from upsampling stages to the corresponding downsampling ones. This reduces the effect of vanishing gradients [217], thus, improving the performance of the neural network. In a typical implementation of the U-Net, the sizes of the input and output data are identical. This feature is suitable for diffusion models as the sizes of x_0 to x_T remain the same.

[3]This figure is created based on a similar figure in [217].



(a) Three-layered dogbone unit cell. (b) Dogbone dimensions.

Fig. 6.4 The parameters L, W, and T are the length, width and trace width of the dogbone. As shown in the left figure, these three dogbone layers are supported by two dielectric substrates (Rogers RO3010).

6.3.3 Latent Diffusion Model (LDM)

Although diffusion models have achieved the state-of-the-art performance in many image synthesis tasks [127], the original implementation of diffusion models often suffers from computational cost in order to achieve high-quality generation. The latent diffusion model (LDM) was developed to address this issue by compressing the raw input image before the diffusion process [215]. This separates the overall image synthesis into two steps: compression and generation. For the compression stage, a variational autoencoder (VAE) [173] is utilized while a diffusion model (in our case, the U-Net) is used for the generation stage. As a result, the raw input image is first compressed into a latent space by VAE. In the generation step, the diffusion model only operates (i.e., denoises) in the latent space, hence the name latent diffusion model. By operating in the compressed latent space as opposed to the raw image space, the LDM offers improved memory efficiency and significantly reduces the overall computational complexity.

6.4 Unit Cell Design

A neural network model based on the LDM was developed in order to synthesize unit cells from desired two-port scattering parameters at a given frequency of operation. We consider three-layered unit cells where the PEC traces are in the shape of dogbones as shown in Fig. 6.4. The size of each unit cell is one-sixth of the wavelength at 12.4 GHz, which is about 4 mm $\times$ 4 mm. These PEC dogbones are supported by two dielectric substrates which are assumed to be Rogers RO3010 substrates with the relative permittivity of 10.2, thickness of 1.28 mm, and the loss tangent of 0.0035.[4] There are three parameters that govern the shape of each dogbone: width, length, and trace width. In our implementation, similar to [123], we have fixed the trace width[5] while varying width and length values to generate different surface properties. Thus, a given three-layered unit cell can be represented by six different W and L (i.e., two per layer).

The input to this network, in our implementation, is a real-valued vector of length seven ($\mathbb{R}^{7\times 1}$) that consists of the operating frequency as well as the real and imaginary parts of the following scattering parameters: S_{11}, S_{12}, and S_{22}. (Since our unit cells are reciprocal, $S_{12} = S_{21}$.) We refer to this $\mathbb{R}^{7\times 1}$ vector as the input vector. Given this input vector, the generated output of the neural network is a three-channel binary image of the pixelated three-layered dogbone unit cell design, see Fig. 6.5. The presence and absence of PEC pixels represents the binary aspect of this image. In our implementation, this will be a $3 \times 128 \times 128$ tensor, which means 128×128 pixels per layer. (The number of pixels should be chosen based on the possible fabrication precision.) We refer to this tensor as *pixelated PEC image*. Finally, we note that for our design parameters, there exist 7.62×10^{11} possible designs.[6]

[4]For fabrication purposes, see Section 6.5.A, these two Rogers RO3010 substrates are bonded by a 1.5 mil ($\approx$ 0.04 mm) Rogers 2929 bondply. Due to the small thickness of this bondply, its presence has been ignored in the unit cell synthesis process.

[5]In our design, we have fixed the trace width to $T = 0.441$mm, which is equivalent to 14 pixels in the pixelated PEC image.

[6]Each dogbone is generated on a 128×128 canvas with a trace width fixed at 14 pixels. Thus, the width and length of the dogbone varies from 20 to 124 and from 38 to 124 pixels respectively.

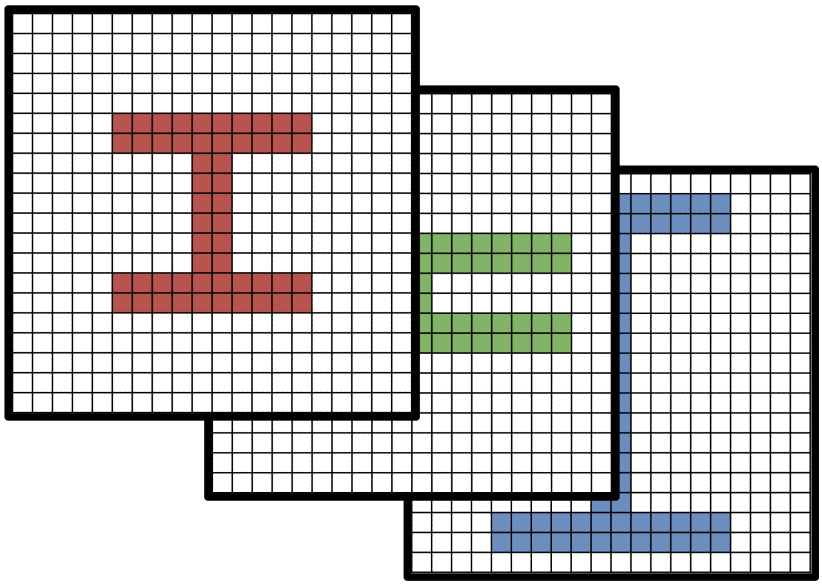


Fig. 6.5 The three-channel pixelated binary image of dogbones (not to scale).

As noted earlier, we utilize the LDM to generate unit cells from given input scattering parameters. The LDM consists of two main components that need to be trained separately. Firstly, the VAE is trained to encode pixelated PEC images to a latent space and vice versa for the decoding. Secondly, the diffusion model, implemented in the U-Net, is trained to synthesize latent tensors based on desired input vectors. In the following two subsections, these two training procedures are described.

6.4.1 VAE Dataset Generation and Training

We employ a standard VAE that is constructed using a convolutional neural network to distill pixelated PEC images into latent tensors for the training and prediction of the LDM. Since the VAE is an unsupervised network (i.e., it only requires pixelated PEC images without their corresponding scattering parameters), we do not need to perform full-wave simulations in its training process. To fully take advantage of this feature such that the VAE network can explore the problem space as much as possible, pixelated PEC images are randomly generated. In fact, in order to speed up the

Consequently, for one layer there are $(124 - 20 + 1) \times (124 - 38 + 1) = 9135$ possible designs. With three layers there will be $9135^3 = 7.62 \times 10^{11}$ designs.

training process and avoid wasting hard-drive space, random pixelated PEC images are generated "on-the-fly" in parallel with the training process. The resulting latent tensor has the size of $4 \times 8 \times 8$, which can be compared to a pixelated PEC image having the dimension of $3 \times 128 \times 128$.

6.4.2 DM Dataset Generation and Training

The DM that was developed in this work closely follows the original network developed in [215]. The backbone of this diffusion network is the U-Net with convolutional layers. To synthesize unit cells based on desired scattering parameters and the frequency of operation, the input vector is injected into each layer of the U-Net through a cross-attention mechanism. It should be noted that the original LDM implementation introduced domain-specific encoders to handle various types of conditional information (e.g., generating images from class labels or text prompts). However, since our network is specific to unit cell synthesis from a single set of particular conditional information (i.e., scattering parameters and the frequency of operation), we omitted the implementation of domain-specific encoders.

To generate a large and diverse dataset for the training of the DM, we have developed an automated dataset generator to produce a labeled dataset. This dataset generator uses Ansys HFSS as its full-wave simulation engine. Each unit cell was simulated under periodic boundary conditions excited by Floquet ports.[7] The size of the training dataset is about $500,000$ and the testing dataset has a size of about $50,000$ which was hidden during the training process. A primary challenge encountered during the training process of the developed DM pertains to its premature overfitting. That is, at the early stage of training, we observed a rising loss in the testing dataset while the training dataset loss continued to decrease. This is likely associated with our small dataset size compared to the vast possibilities for unit cell designs (7.62×10^{11} possible

[7]Due to the use of periodic boundary conditions to characterize the two-port scattering parameters of unit cells, the final metasurface design is still under the local periodicity approximation [218], which can degrade the performance of aperiodic metasurfaces considered herein. Recently, techniques such as [219] have been proposed to tackle this approximation.

designs as noted in Section 6.4). To handle this issue, we developed a neural network model based on EfficientNet [220], which predicts the two-port scattering parameters for given pixelated PEC images. This forward solver was trained by the aforementioned Ansys HFSS dataset. Once trained, it was applied to randomly generated PEC images so as to augment the original Ansys HFSS dataset. In particular, the labeled datasets generated by Ansys HFSS and our forward solver contribute to 95% and 5% of the overall labeled dataset respectively. This enriched dataset was then used to train the DM where the 5% synthetic data was refreshed at each epoch. This approach successfully mitigated the overfitting issue, thereby enhancing the overall performance of the DM.

During the generation process, for a given metasurface design, a list of desired two-port scattering parameters associated with each metasurface unit cell and the operational frequency is fed to the developed LDM. For given two-port scattering parameters, the neural network will on average generate ten unique pixelated PEC images. To minimize the simulation setup in Ansys HFSS, an image processing technique is applied to these uniquely generated pixelated PEC images to find its corresponding dogbone design parameters (width and length for each layer), which can then be easily modeled in Ansys HFSS. These unique dogbone unit cell designs along with their achieved two-port scattering parameters are then recorded to a dataset for the final metasurface design. When constructing the entire metasurface, for each unit cell, we will find the dogbone design whose achieved scattering parameters are "closest" to the desired values. It should be noted that we can employ a few different strategies to find the "closest" dogbone designs. For example, we can optimize for both S_{11} and S_{21} with an equal weight, optimize for S_{11} individually, or merely S_{21} by minimizing the following cost functions

$$\mathcal{C}_1 \triangleq 0.5(|S_{11}^{\text{desired}} - S_{11}^{\text{achieved}}| + |S_{21}^{\text{desired}} - S_{21}^{\text{achieved}}|), \tag{6.1}$$

$$\mathcal{C}_2 \triangleq (|S_{11}^{\text{desired}} - S_{11}^{\text{achieved}}|), \tag{6.2}$$

$$\mathcal{C}_3 \triangleq (|S_{21}^{\text{desired}} - S_{21}^{\text{achieved}}|), \tag{6.3}$$

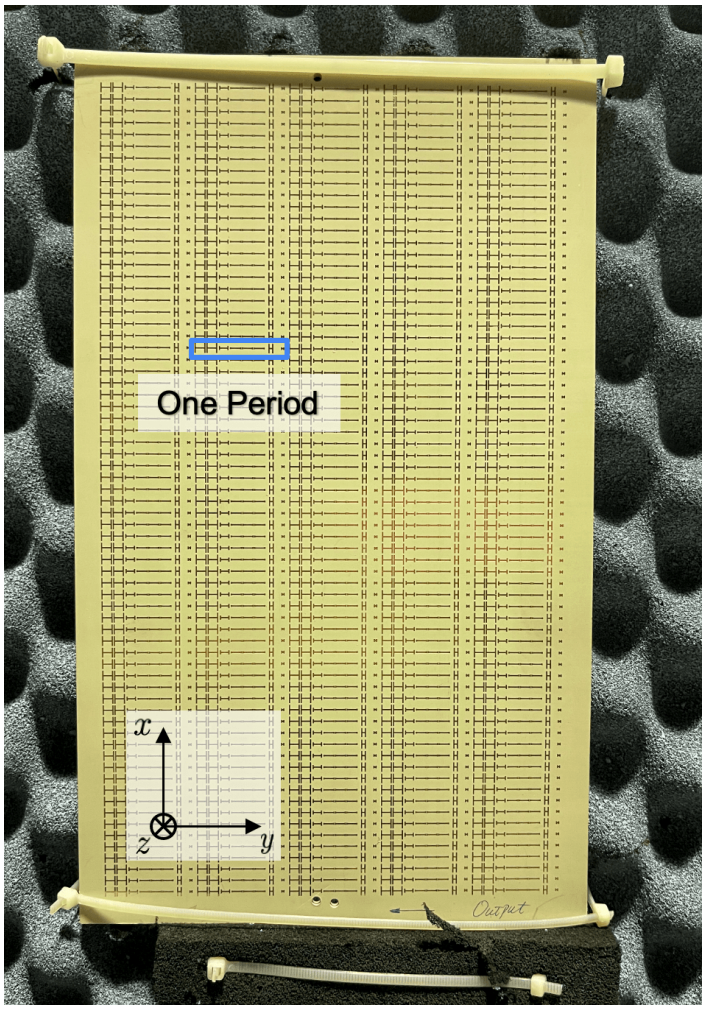


Fig. 6.6 The fabricated metasurface (for the wide-angle refraction example). The output face of the metasurface is visible in this figure. The spatial modulation is along the y axis, with no variation along the x axis. Five periods have been used along the y axis, with eight unit cells in one period.

respectively. Finally, we note that the generation time for the LDM to create unit cell designs (on average, 10 unit cell designs for a given set of scattering parameters) is approximately 7 seconds with the neural network running on an Nvidia Tesla V100. In addition, the HFSS simulation time to evaluate these dogbone unit cells typically ranges from 1 to 8 minutes per unit cell. Thus, the main computational cost of this generation approach is associated with the HFSS run time to single out one unit cell design for a given set of scattering parameters.

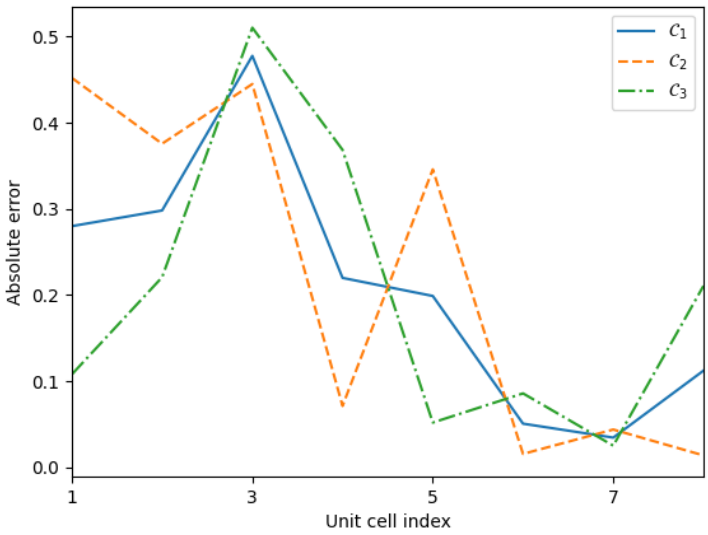


Fig. 6.7 The absolute error of the achieved scattering parameters for each unit cell with respect to the desired ones for the plane wave refraction example. Each period consists of eight unit cells. The final unit cell design is based on $\mathcal{C}_1$. The values of $\mathcal{C}_2$ and $\mathcal{C}_3$ are also reported for the sake of completeness.

6.5 Results

In order to evaluate the performance of the developed neural network for the synthesis of electromagnetic metasurface unit cells, we consider three design examples. In each example, we first briefly talk about the procedure we used to obtain the required two-port scattering parameters for each unit cell (i.e., the macroscopic design step). We then pass these scattering parameters to the LDM to design the unit cells. The performance of the LDM is then reported based on the cost functionals in (6.1)-(6.3) as well as the quality of the required wave transformation achieved by the whole metasurface. The dimensions of the unit cells for each design example can be found in the Appendix B. We emphasize that no further optimizations have been applied on these unit cells and that they are taken directly from the LDM.

6.5.1 Plane wave refraction

For the first example, we design and evaluate a refracting metasurface operating under the 2D TM case. As shown in Fig. 6.6, the metasurface lies in the xy plane, centered at the origin, with the spatial modulation along the y axis. The metasurface has no variation along the x axis, hence the 2D assumption. Noting that the design is under the 2D TM assumption with the propagation plane being the yz plane, the EM field components will be H_x, E_y, and E_z. It has been shown in [42] that Snell's law at the interface of two dielectric media can be generalized when a phase profile is engineered along the interface. A properly designed metasurface when placed at the interface can support the required phase profile. Assuming free space on both sides of the metasurface, refracting a plane wave incident at the angle θ_i (with respect to the normal vector) on the metasurface to the angle θ_t requires a periodic metasurface with the periodicity of

$$P = \frac{\lambda_0}{\sin\theta_t - \sin\theta_i}, \tag{6.4}$$

where λ_0 is the wavelength of operation in free space [42]. In our case, the incident plane wave impinges normally on the metasurface ($\theta_i = 0$); thus, $P = \lambda_0/\sin\theta_t$. Assuming the desired refraction angle to be 69.6° and the frequency of operation to be 10 GHz, the periodicity will be about 32 mm. Noting that the size of each unit cell is about 4 mm, this requires eight unit cells per period. To enable perfect refraction, the scattering parameters associated with each unit cell need to be properly designed to satisfy the local power conservation (LPC) constraint [47], which can also be justified using the generalized scattering parameters [45]. It has been shown in [45, 18, 46, 19] that perfect refraction can be achieved by omega bianisotropic metasurfaces. The unit cells considered herein can provide omega bianisotropy due to the possibility of having an asymmetric structure (i.e., the top and bottom dogbones being different) [47], and can thus be used for this purpose.

To begin the design process, we first assume that the tangential fields on the input face of the metasurface due to a normally incident plane wave (travelling in the $-\hat{z}$

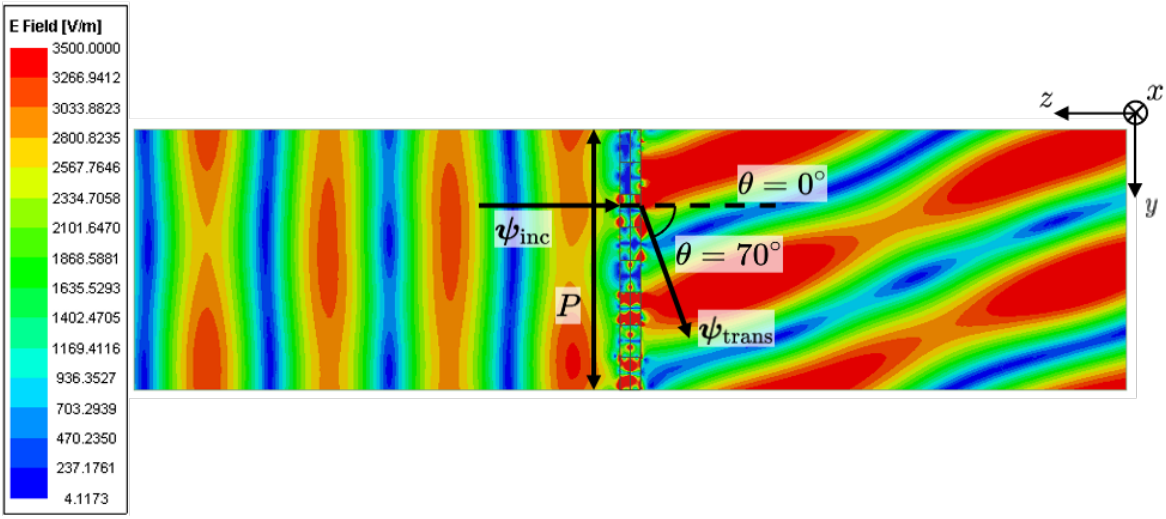


Fig. 6.8 Plot of the electric field magnitude at a fixed phase in Ansys HFSS. In this simulation, one period (P) of the metasurface has been used and the metasurface is enclosed by the periodic boundary conditions. In addition, PEC baffles have been used for this simulation. The left side of the metasurface is its input side (illuminated by ψ_{inc} where ψ denotes $\{E, H\}$) and the right side of the metasurface is its output side yielding the refracted wave ψ_{trans}.

direction) are

$$E_y^{\text{input}} = E_{\text{i}} \quad , \quad H_x^{\text{input}} = E_{\text{i}}/\eta_0, \tag{6.5}$$

where η_0 is the intrinsic impedance of free space and E_{i} is the amplitude of the incident plane wave. (Due to the assumption of a reflectionless metasurface, the input tangential fields will only be affected by the incident field.) Assuming that the output wave is plane wave refracted by θ_t, the tangential fields on the output face of the metasurface will be

$$E_y^{\text{output}} = E_{\text{t}} \cos\theta_t e^{-jk_0 y \sin\theta_t}, \tag{6.6}$$

$$H_x^{\text{output}} = \frac{E_{\text{t}}}{\eta_0} e^{-jk_0 y \sin\theta_t}, \tag{6.7}$$

where E_t is the amplitude of the refracted plane wave. As noted in [46, 221], if E_t is chosen as

$$E_t = \frac{|E_i|}{\sqrt{\cos\theta_t}} e^{j\varphi}, \tag{6.8}$$

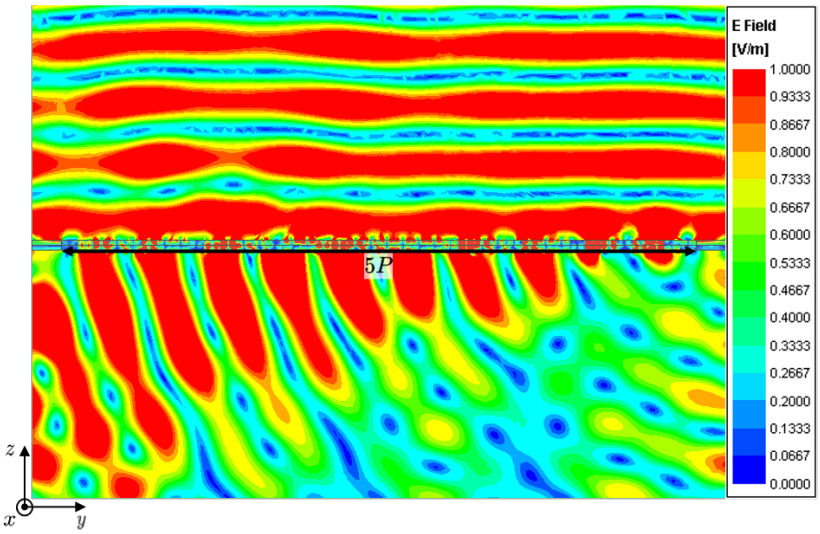


Fig. 6.9 Plot of the electric field magnitude at a fixed phase for the truncated metasurface where five periods ($5P$) are used to form the truncated metasurface (comprising of 40 unit cells and no PEC baffles). From each side, the metasurface is extended by quarter-wavelength absorbers.

where φ is an arbitrary constant phase, the LPC constraint is met and perfect refraction using a lossless and passive metasurface can be obtained. Now that we know the input and output tangential fields, we discretize y in (6.6) and (6.7) into eight equally spaced unit cells so as to cover the whole period P. Knowing the properties of the Rogers RO3010 substrate, we can obtain the required impedance values for each unit cell similar to the procedure explained in [19]. These unit cell properties can then be easily represented by two-port ABCD parameters. The ABCD representations are then converted to the two-port scattering parameter representation [142].[8] These two-port scattering parameters are then fed to the LDM to yield dogbone-based unit cell designs.

As noted earlier, the LDM outputs a few designs for a given two-port scattering matrix. We then need to single out one design from these multiple potential solutions. To this end, for each unit cell, each of the potential solutions is evaluated by calculating $\mathcal{C}_1$ in (6.1), with the chosen dogbone unit cell being the one that minimizes $\mathcal{C}_1$. (For the dimensions of these unit cells, see the Appendix B.) The achieved scattering parameters

[8]For further verification, each of these two-port scattering matrices is converted to the two-port generalized scattering matrix [45] to make sure that the absolute of the generalized forward scattering parameters is unity.

of these dogbones are compared to the desired scattering parameters in Fig. 6.7 by plotting $\mathcal{C}_1$ for each unit cell. Although not used for this example, the values of $\mathcal{C}_2$ and $\mathcal{C}_3$ have also been reported in the same figure. Noting $\mathcal{C}_1$, it can be seen that the performance of the first three unit cells is worse than the remaining unit cells. This is partly due to the fact the phase of the desired S_{21} for the first unit cell is close to zero degrees. As noted in [222], this unit cell will then be close to resonance, and meeting its desired scattering parameters can be quite challenging. The presence of the resonance for this unit cell was further indicated by the lossy behaviour of this unit cell as also observed in [222]. These unit cells are then used to form the metasurface. The near-field simulation of this metasurface under the periodic boundary condition and perfect electric conductor baffles [123] is shown in Fig. 6.8, which shows small reflections on the input (left) side and also a refracted plane wave in the output (right) side. The achieved refraction angle, measured geometrically, is about 70° which is very close to the desired refraction angle. In addition, according to the Floquet mode analysis by Ansys HFSS, the *refraction* power efficiency (power contained within the desired Floquet mode with respect to the total power) of this metasurface is about 75%.

For manufacturing purposes, we now limit ourselves to only five periods of this metasurface and also remove the baffles for the ease of fabrication. The near-field simulation of this *truncated* metasurface is shown in Fig. 6.9. The far-field power pattern obtained from this metasurface is shown in Fig. 6.10 (orange). The angle of the main beam is at 64°. The reason that this achieved main beam angle is not closer to the desired refraction angle of 69.6° can be justified by noting that the design process assumes a fully periodic boundary condition which effectively means that the metasurface aperture is of infinite length. In addition, this fully periodic assumption means that the effective aperture size of the metasurface remains the same from any angles. However, this is not the case for the truncated metasurface, which has an aperture size of five periods. This finite aperture size results in a reduced effective aperture size (by a factor of $\cos\theta$) when radiating at $\theta \neq 0$. As noted in [19], this

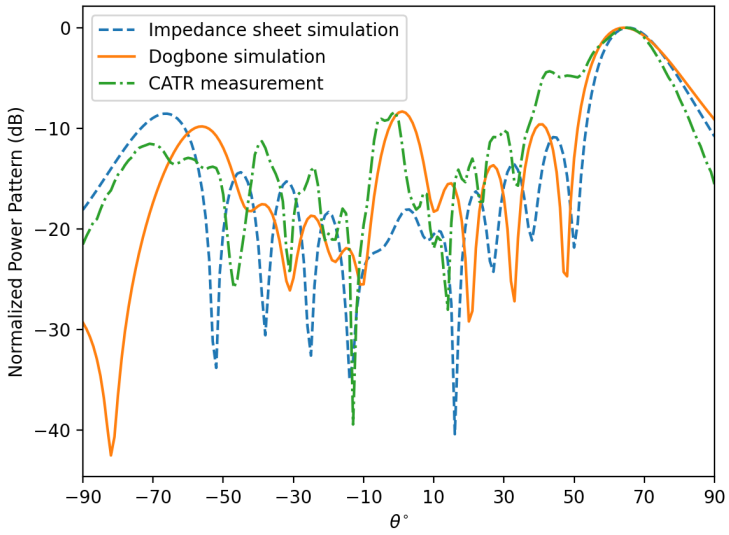


Fig. 6.10 The normalized far-field power pattern plots of the truncated metasurface ($5P$ aperture size) for three cases: (i) Ansys HFSS simulation of the metasurface consisting of dogbone-shaped unit cells, (ii) Ansys HFSS simulation of the metasurface consisting of ideal impedance sheet unit cells, and (iii) the measured far-field power pattern of the fabricated metasurface. (The "correction factor" of $\cos\theta$ has not been applied to any of these plots.)

discrepancy between a fully periodic aperture and a truncated one can be numerically compensated to yield the expected refraction angle, i.e., the main beam angle that would have been obtained if the metasurface had many more periods. Applying this correction factor to the whole pattern, the expected refraction angle becomes 68°, which is closer to the desired 69.6°. To compare this with a more ideal situation, Fig. 6.10 (blue) shows the far-field power pattern of this metasurface when the unit cells are modeled using impedance boundary conditions instead of dogbones. As can be seen, this ideal power pattern exhibits a lower side lobe level around $\theta = 0$ with a main beam angle at 65°. Applying the above correction factor to the whole pattern, the expected refraction angle becomes 68°.

We have also fabricated this metasurface as shown in Fig. 6.6. As can be seen, along the y axis, we have used five periods, and along the x axis, we have repeated the dogbones 70 times to mimic a 2D condition (i.e., $\partial/\partial x = 0$). This fabricated

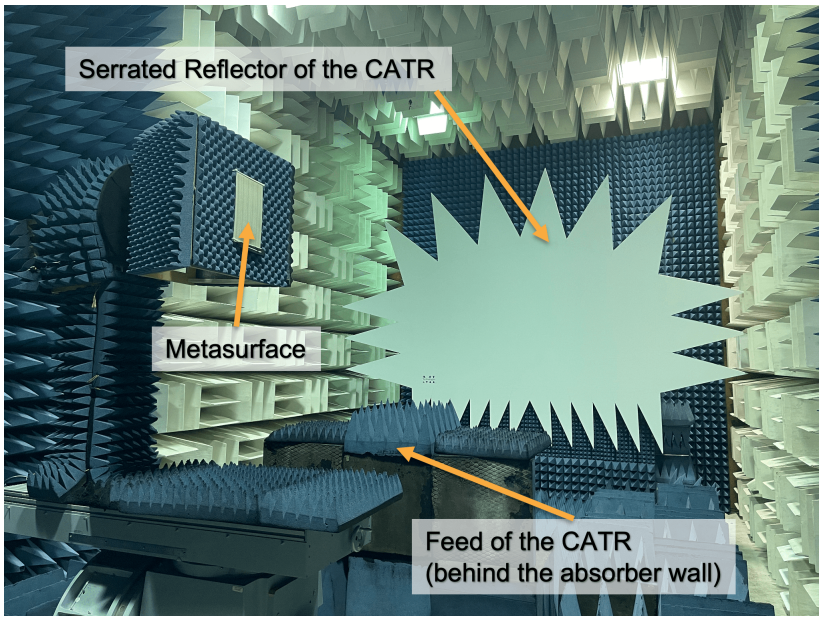


Fig. 6.11 The far-field measurements of the fabricated metasurface in a compact antenna test range (CATR). The output face of the metasurface, i.e., the face that is supposed to interact with a plane wave at 69.6°, is illuminated by the serrated reflector.

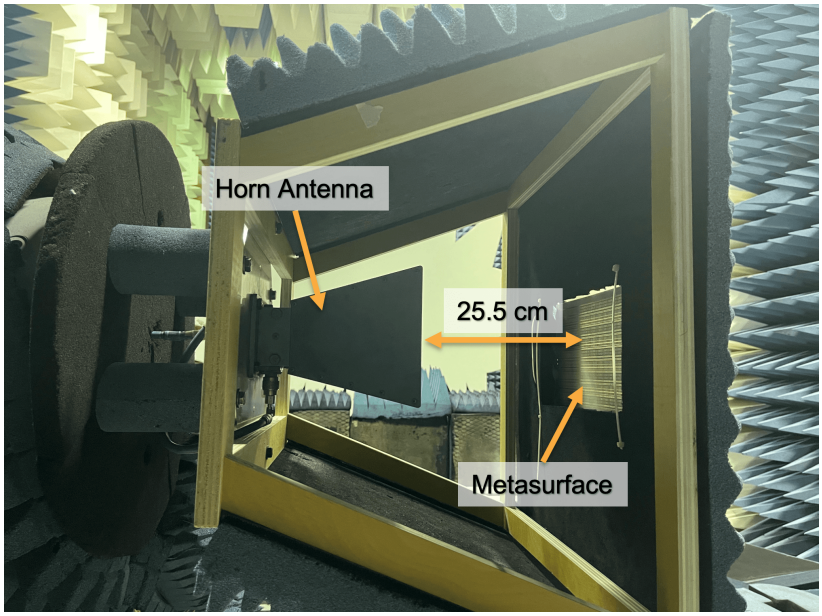


Fig. 6.12 The metasurface mounting and the receiving horn antenna. The horn antenna is located on the input side of the metasurface (i.e., the side that is supposed to interact with a normally incident plane wave).

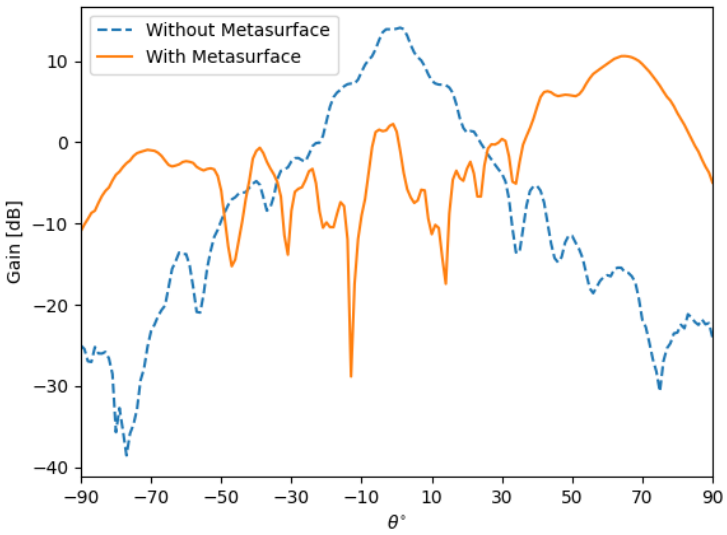


Fig. 6.13 The gain pattern obtained from the CATR measurements with and without the metasurface. The maximum gain of the horn antenna structure with and without the metasurface are 14.1 dB and 10.6 dB respectively.

metasurface has then been tested in a compact antenna test range (CATR) as shown in Fig. 6.11. Once the feed of the CATR illuminates the serrated reflector, a plane wave condition is created in the quiet zone of the CATR. The metasurface under test is then placed in this quiet zone such that its output face is looking toward the reflector. On the input side of the metasurface, a horn antenna is placed to receive the electromagnetic wave passing through the metasurface. This horn antenna is shown in Fig. 6.12 where the distance between the input face of the metasurface and the horn aperture is 25.5 cm (equivalent to about $8.5\lambda_0$). The horn-metasurface structure is then rotated from $-90°$ to $+90°$ with respect to the boresight of the serrated reflector. The recorded power pattern of the horn-metasurface structure at the frequency of operation (10 GHz) is then shown in Fig. 6.10 (green). The main beam angle is at 65°. After applying the correction factor, the expected refraction angle will be at 67°.

Finally, the transmission power efficiency of the truncated metasurface, shown in Fig. 6.9 and obtained from HFSS simulation, is 71.07%. (This efficiency is defined as the ratio between the normal real power leaving the metasurface and the normal real

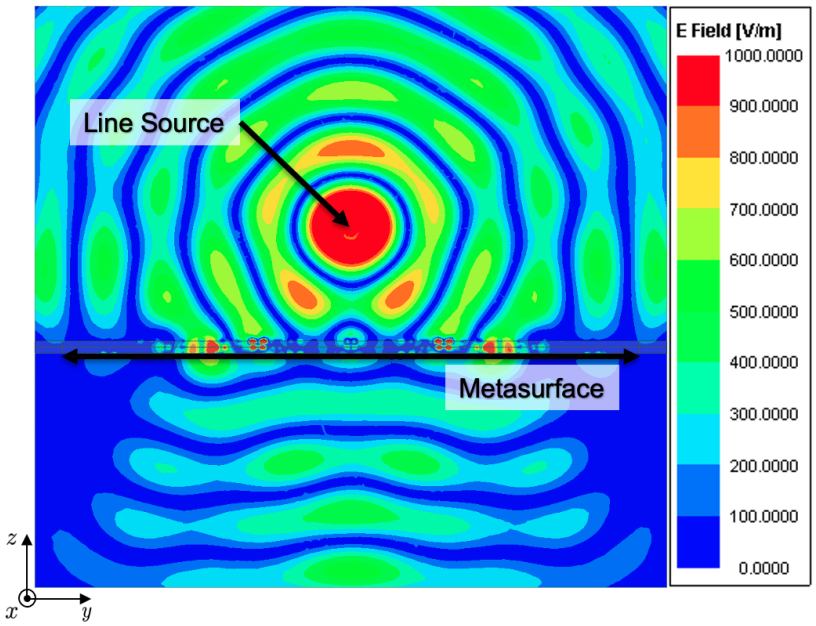


Fig. 6.14 Plot of the electric field magnitude at a fixed phase. The metasurface is illuminated by a line source and then transforms the incident cylindrical wave to a truncated plane wave.

power incident to the metasurface.) We have not obtained the transmission power efficiency from measurements. However, we have compared the maximum gain of the horn-metasurface structure with the gain obtained when the metasurface is removed (while the absorbing structure remains around the horn) as shown in Fig. 6.13. As can be seen, the maximum gain has dropped from 14.1 dB to 10.6 dB when the metasurface is added to the horn-absorbing structure.

6.5.2 Cylindrical wave to plane wave transformation

This design example aims to transform the cylindrical wave generated by a line source into a plane wave; see Fig. 6.14. The distance from the line source to the input face of the metasurface is one wavelength at the operational frequency of 12.4 GHz. This example is under the 2D TE assumption; thus, since the plane of propagation is the yz plane, we have three field components: E_x, H_y, and H_z. The metasurface consists of 31 unit cells and has an aperture size of about $5.2\lambda_0$. Due to the difference between the power profiles associated with the line source and plane wave in the input and

output faces of the metasurface, LPC is not satisfied. To meet the LPC constraint, auxiliary surface waves [116, 114, 119, 118] are employed according to the procedure explained in [119]. In this approach, an auxiliary surface wave is added to the input face of the metasurface that redistributes the power between the unit cells while being evanescent (exponentially decaying) in the direction normal to the metasurface; thus, no real power is reflected (or lost) in the ideal scenario. After augmenting the input tangential fields (resulting from the line source) with the fields due to this auxiliary surface wave, the desired two-port scattering parameters are obtained similar to the procedure explained for the plane wave refraction example. These two-port scattering matrices are then fed to the LDM.

As noted earlier, the LDM yields a few potential unit-cell solutions per given two-port scattering parameters. To single out a unit cell, we evaluate $\mathcal{C}_3$ in (6.3) for these potential solutions, and choose the one that corresponds to the minimum value. The values of $\mathcal{C}_3$ for the chosen unit cells are plotted in Fig. 6.15. Although we have not used $\mathcal{C}_1$ and $\mathcal{C}_2$ to choose the final unit cells, their values have been plotted for the final unit cells. Since the unit cells are chosen based on $\mathcal{C}_3$ (focused on S_{21} performance), the error associated with S_{21} is good, but the other two errors are large. These unit cells are then simulated in Ansys HFSS as shown in Fig. 6.14. Although the resulting plane wave is not of high quality, we do see fields with planes of constant phase emanating from the metasurface, and a certain uniformity in the magnitude. The power transmission efficiency of this metasurface is 66.85%. Similar designs using conventional techniques (i.e., lookup tables followed by further iterative dimension tuning) have shown slightly better performance as compared to this example, e.g., see [223]. However, the main advantage of our design is that it has not gone under any further optimizations. Thus, it can be used as an initial guess for further tuning.

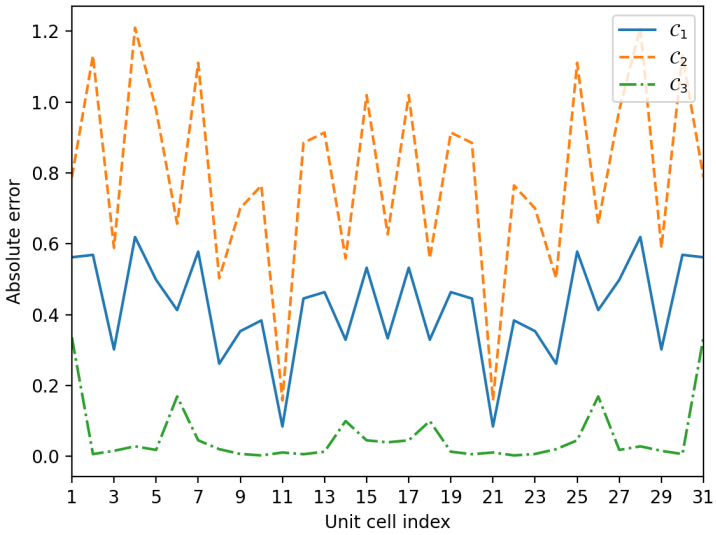


Fig. 6.15 The absolute error of the achieved scattering parameters for each unit cell with respect to the desired ones for the circular wave to plane wave transformation example. The metasurface consists of 31 unit cells. The final unit cell design is based on $\mathcal{C}_3$. The values of $\mathcal{C}_1$ and $\mathcal{C}_2$ are also reported for the sake of completeness.

6.5.3 From far-field performance criteria to metasurface design

For the last example, we consider an end-to-end design procedure where the designer identifies the following: (i) desired far-field performance criteria to be achieved, (ii) the line source excitation and its distance from the metasurface ($1\lambda_0$), (iii) the aperture size of the metasurface ($6.5\lambda_0$) along with the frequency of operation (12.4 GHz) employing 39 unit cells. Similar to the previous example, this design is pursued under the 2D TE assumption. The desired far-field performance criteria are shown in Fig. 6.16 where "•" shows the desired two main beam angles, "×" denotes the desired null angles, and "—" represents the half-power beamwidths (HPBWs). Finally, the two horizontal lines in Fig. 6.16 show the desired sidelobe levels (SLLs) and maximum null levels.

To find the required two-port scattering parameters for each unit cell, we need to know the required tangential fields on the metasurface aperture. To this end, we can employ inversion methods such as [75, 121, 71, 70] to infer these tangential fields

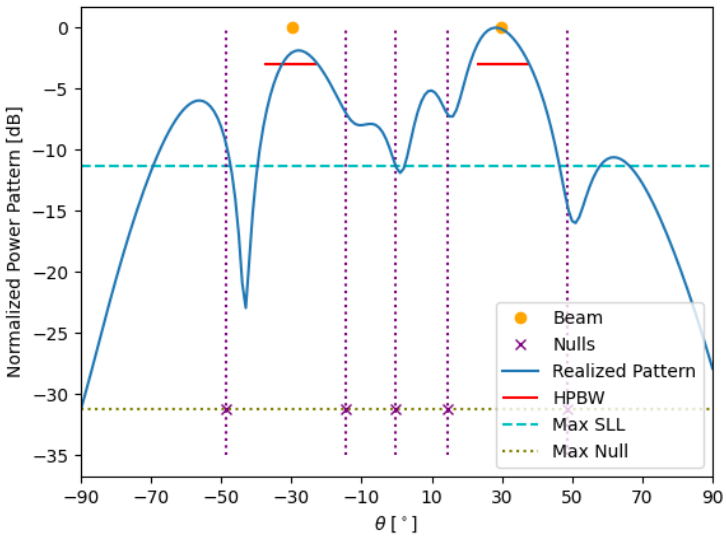


Fig. 6.16 The achieved far-field power pattern from the metasurface designed to meet the desired far-field performance criteria.

from the knowledge of desired far-field performance criteria. Herein, the procedure developed in [70] was used to find these tangential fields and subsequently the required unit cells' two-port scattering matrices. Once the required two-port scattering matrices are found, they are fed to the LDM. The potential solutions for each unit cells are singled out by evaluating $\mathcal{C}_3$ in (6.3). The values of $\mathcal{C}_3$ for the metasurface's 39 unit cells are shown in Fig. 6.17. Similar to the previous example, the values of $\mathcal{C}_1$ and $\mathcal{C}_2$, although not used in this design process, are reported in this figure for the sake of completeness. The obtained unit cells are then placed together to form the final metasurface.

As shown in Fig. 6.18, the metasurface is simulated in Ansys HFSS, which shows the presence of the two beams on its output side. In addition, the cylindrical incident wave on the input side is somewhat distorted, which shows the presence of undesired reflections on the input side of the metasurface. The resulting far-field power pattern is obtained and shown in Fig. 6.16 (solid blue curve). For this example, the realized main beam angles are close to the desired ones, although the realized left beam has not reached the desired level. Similarly, the HPBW requirement for the right beam has

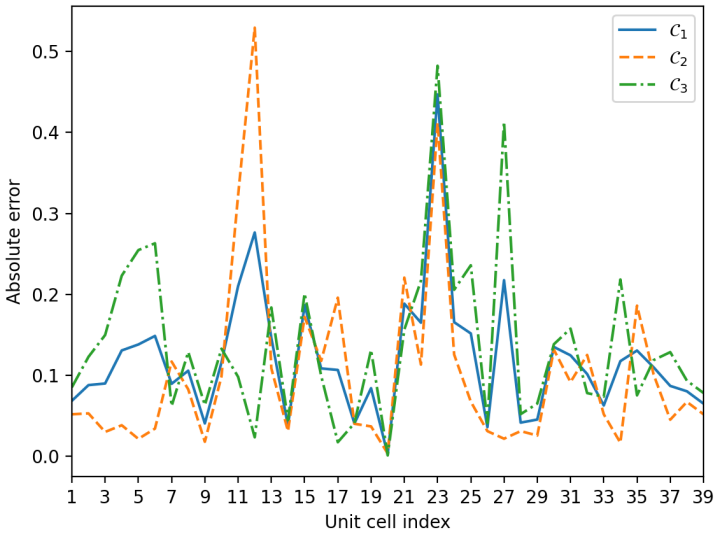


Fig. 6.17 The absolute error of the achieved scattering parameters for each unit cell with respect to the desired ones for the performance criteria design example. The metasurface consists of 39 unit cells. The final unit cell design is based on $\mathcal{C}_3$. The values of $\mathcal{C}_1$ and $\mathcal{C}_2$ are also reported for the sake of completeness.

almost been met while that for the left beam has suffered. Although most of the the null angles have been realized, they are not as deep as required by the desired maximum null level. Also, the achieved SLL is higher than the desired one. In general, this is a difficult design task as the auxiliary surface waves required for power redistribution (in order to satisfy the LPC constraint for lossless and passive design) are heavily dependent on $\mathcal{C}_2$, which was not considered when selecting the unit cells. However, noting that the reported result does not involve any fine tuning, it can be regarded as a promising initial guess. Finally, the power transmission efficiency of this metasurface was found to be 66.95%

6.6 Conclusion

EM metasurfaces enable systematic wave transformations and can thus be utilized for the synthesis of arbitrary radiation patterns and other applications involving tailoring EM waves. The design process consists of the macroscopic and microscopic steps.

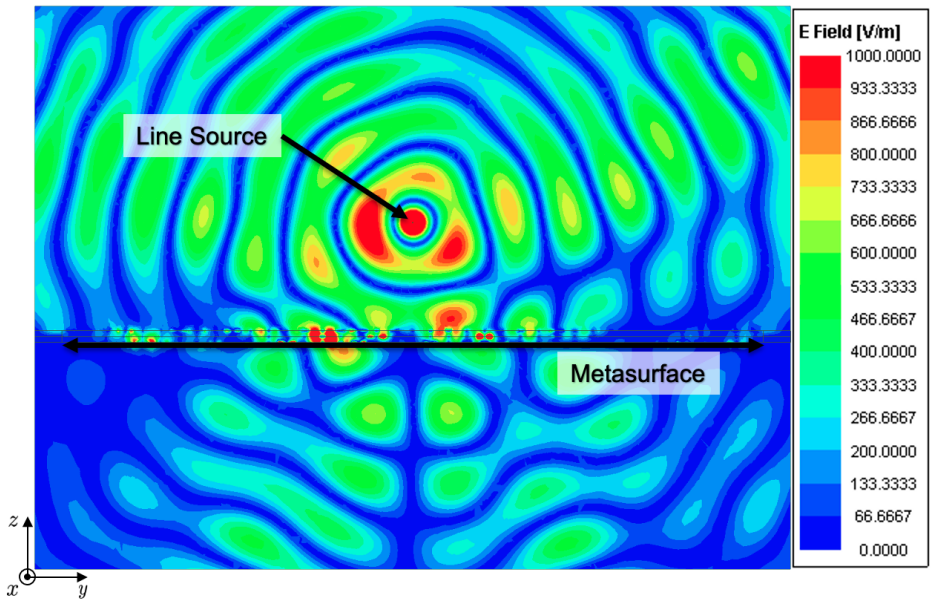


Fig. 6.18 Plot of the electric field magnitude at a fixed phase. The metasurface is illuminated by a line source and then transforms this incident field to an output aperture fields in an attempt to meet the desired far-field performance criteria.

Although the macroscopic design step can be made to be automated, the microscopic step can be iterative and time-consuming. To shorten the design cycle, DL approaches have been investigated for the microscopic design step. Herein, we have proposed an LDM-based deep learning microscopic design approach. This is based on casting multi-layered unit cell synthesis as conditional multi-channel binary image synthesis (one channel per unit cell layer). The performance of this LDM-based unit cell synthesis technique was evaluated against full-wave simulations and experimental data. Although the performance of the metasurfaces formed by the unit cells generated by the LDM is not fully satisfactory, the results show that they can still provide a good initial guess for the microscopic design step.

Chapter 7

Microscopic Design from Far-Field Masks

Preface

In Chapter 6, we focused on the microscopic design stage and also demonstrated an end-to-end metasurface design example based on far-field specifications such as main beam angles and HPBWs. In this chapter, we investigate a complete end-to-end design scheme for a metasurface design based on desired far-field masks. In addition, using this design approach, we fabricate a metasurface and present its experimental evaluation using a planar near-field antenna range. The macroscopic design stage utilizes the DL framework detailed in Chapter 4, while the microscopic design stage employs the DL framework from Chapter 6.

The contents of this chapter, up to the end of Section 7.3 and the first part of the conclusion in Section 7.5, are from a conference paper presented in 2024 [195].[1] On the other hand, the experimental verification reported in Section 7.4 is yet to be submitted for publication.

[1]©2024 IEEE. Reprinted, with permission, from **C. Niu**, M. Phaneuf, and P. Mojabi, "End-to-End Transmissive Metasurface Design from Far-Field Masks Using Deep Learning," *IEEE International Symposium on Antennas and Propagation and ITNC-USNC-URSI Radio Science Meeting*, pp. 1-2, Florence, Italy, 2024.

Abstract

We present a deep-learning based framework for the design of lossless and passive transmissive metasurfaces. The DL framework takes the user-defined lower and upper far-field masks along with the incident field data. It then outputs a metallic trace design for the metasurface such that the resulting power pattern resides within the specified masks.

7.1 Introduction

Electromagnetic (EM) metasurfaces can precisely control the phase, amplitude, and polarization of EM waves in a systematic manner [16]. The metasurface design process is typically divided into two steps: macroscopic and microscopic [28]. The macroscopic stage begins by outlining specific performance criteria to be achieved. In this work, these are prescribed lower and upper far-field (FF) masks within which the resulting power pattern (PP) must resides, e.g., see [224, 73, 187]. These performance criteria are utilized to infer the necessary tangential fields on the metasurface aperture by employing various techniques such as EM inversion [75, 121, 71, 114], array techniques [110], or deep-learning approaches [70]. Subsequently, these tangential fields are employed along with the generalized sheet transition conditions (GSTCs) [21] to compute the necessary metasurface properties. Since our focus is on designing lossless and passive transmissive metasurfaces, the macroscopic stage needs to account for local power conservation (LPC) [47].

The microscopic step involves the physical design of unit cells to realize the required surface properties. This can be accomplished through the use of patterned metallic claddings on multi-layer dielectric substrates [18, 47, 129, 225]. This often involves creating an extensive look-up table of single-layer metallic traces and cascading them properly. Achieving the final design typically requires additional optimization and fine-tuning using a full-wave forward solver. In this chapter, we aim to showcase a DL-based end-to-end lossless and passive transmissive metasurface design pipeline.

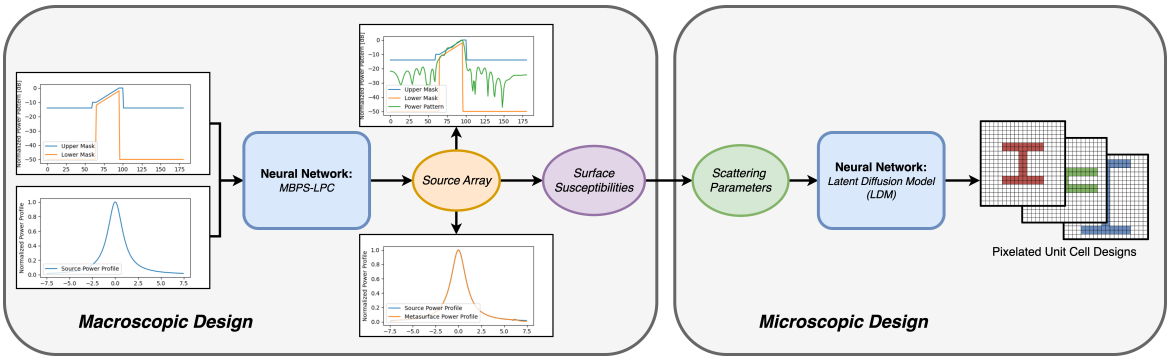


Fig. 7.1 ©2024 IEEE. DL-based end-to-end design pipeline. (The macroscopic and microscopic stages are based on work described in Chapter 4 and Chapter 6 respectively.) The first neural network is termed MBPS-LPC, which indicates a mask-based pattern synthesis (MBPS) approach with LPC.

This approach integrates the DL macroscopic stage presented in [117] with the DL microscopic stage developed in [129] to create an end-to-end design pipeline.

7.2 Methodology

The metasurface is positioned on the xz plane, with no change in the z direction while having spatial modulation along the x direction. The wave propagation is assumed to be two-dimensional transverse electric. The transmissive metasurface is illuminated from $y < 0$ region and is then supposed to create a target PP in the $y > 0$ region. The user provides upper and lower FF masks, along with the power profile of the incident wave impinging on the metasurface. Utilizing this information, the neural network developed in [117] outputs an antenna array, termed as the source array. The emanating fields of this source array, when calculated at the metasurface aperture, must satisfy two conditions: (i) create a PP that resides within the specified masks and (ii) meet the LPC constraint with respect to the power profile of the incident wave. These fields are then adopted as the metasurface aperture fields, and the source array is disregarded thereafter. These aperture fields along with the incident fields and GSTCs are employed to determine the required surface susceptibilities for each unit cell. The two-port scattering parameters for each unit cell are subsequently calculated based

on these surface susceptibilities. These scattering parameters are then fed into the generative neural network developed in [129] to create pixelated images representing the three-layer metallic trace designs (in the shape of dogbones) supported by two substrates. These unit cells are then assembled into the overall metasurface design with no tuning. The overall design pipeline is illustrated in Fig. 7.1.

7.3 Results

To evaluate the performance of the developed DL-based end-to-end metasurface design pipeline, we present an example wherein the target PP needs to conform to the trapezoidal-shaped FF lower and upper masks as shown in the top leftmost of Fig. 7.1. (This is similar to an example considered in [117], but our emphasis now lies on its dogbone-based design rather than solely the impedance sheet design as was the case in [117].) We assume that the metasurface is illuminated by a line source at 10 GHz whose incident power profile is shown in the bottom leftmost of Fig. 7.1. The metasurface aperture, spanning about 15λ where λ is the wavelength, consists of 111 unit cells, each with a size of approximately $\lambda/7.4$ (4.03 mm). The line source is positioned 1λ away from the input face of the metasurface and is centered with respect to it. We perform the design using Rogers RO3010 substrates with the relative permittivity of 10.2, loss tangent of 0.0035, and thickness of 1.28 mm, and evaluate the results using Ansys HFSS. The achieved PP from the dogbone-based design is shown in Fig. 7.2. While the overall trend appears reasonable, there are issues with increased sidelobes and adherence to the trapezoidal shape. Considering that no tuning has been applied to the dimensions of the dogbones, it is expected that this PP can be improved by optimizing these dimensions. To assess this, we replace the dogbones with ideal impedance sheets on the same substrates. As shown in Fig. 7.2, this results in enhancement across all aspects. To investigate another potential avenue for PP improvement, we utilize a thinner RO3010 substrate (0.25 mm, similar to [117]) with

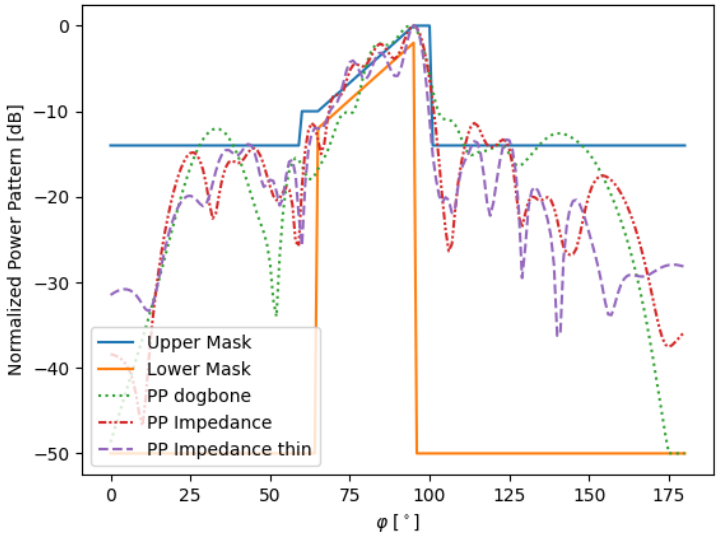


Fig. 7.2 ©2024 IEEE. Trapezoidal-shaped far-field (FF) mask and power patterns (PP) obtained from Ansys HFSS under three scenarios.

impedance sheets. This results in a new PP which nearly meets all the requirements as illustrated in Fig. 7.2.

7.4 Measurements

The three-layered dogbone unit cell designs of the metasurface was fabricated using two sheets of Rogers RO3010 substrates which were bonded by a very thin Rogers 2929 bondply. The Rogers RO3010 substrate has a dimension of 24 inches by 18 inches and a thickness of 1.28 mm, as illustrated in Fig. 7.3. Thus, the total thickness of the metasurface is about 2.6 mm. This metasurface is composed on 111 aperiodic unit cells which are arranged along the x-axis and repeated 145 times along the z-axis to cover the entire substrate area. To mount the metasurface on the measurement apparatus, it is framed with plywood and secured by nylon screws.

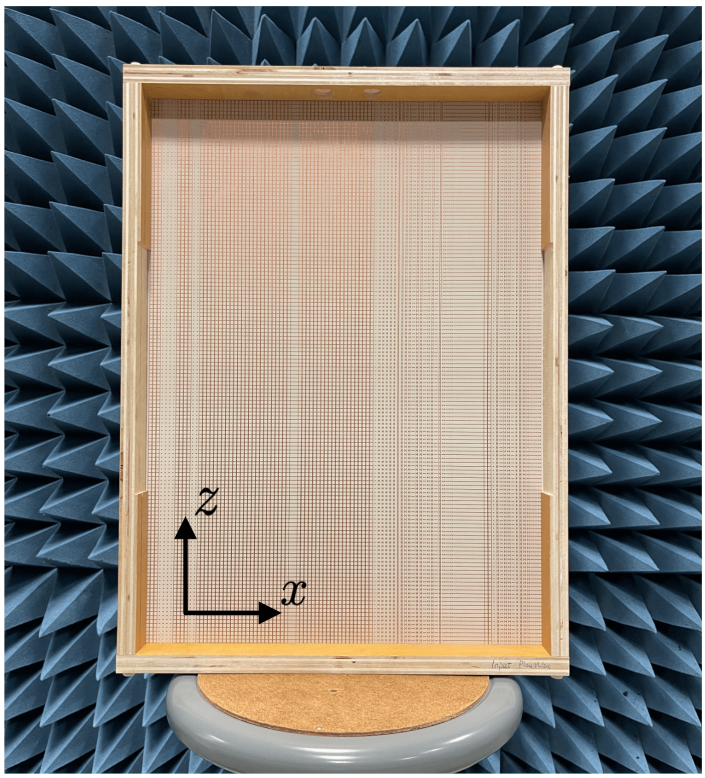


Fig. 7.3 The fabricated metasurface in a wooden frame fastened by nylon screws.

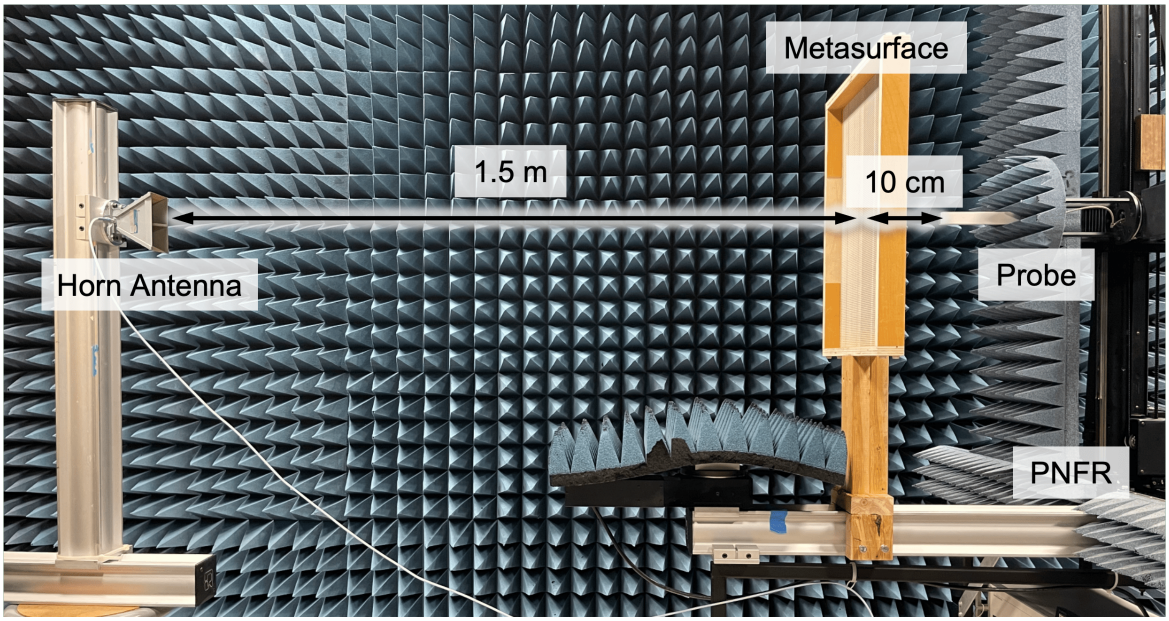


Fig. 7.4 The planar near-field range measurement setup for the fabricated metasurface. (Some absorbers are removed for a better view of the system.)

7.4.1 Planar near-field range (PNFR)

The fabricated metasurface was then characterized using a commercial planar near-field antenna range (PNFR) to evaluate its far-field performance compared to the simulation and the desired FF mask. The PNFR, manufactured by NSI-MI, utilizes the so-called "inverted T" structure and provides a rectilinear grid of near-field planar measurements. These measurements effectively capture the electric field components tangential to the measurement plane, which are then used to calculate the far-field pattern. Fig. 7.4 shows the planar near-field measurement setup. The metasurface was positioned 10 cm away from the near-field measurement plane. A horn antenna, placed 1.5 m away from the input side of the metasurface, was used to mimic a normal plane wave illumination. This is estimated to provide far-field results within $\pm 60^\circ$ of the boresight to the metasurface. To minimize reflections, absorbers were placed around the metasurface during measurements. Near-field data were collected at 21 different frequencies, ranging from 9 GHz to 11 GHz in steps of 0.1 GHz, to account for the potential drift in operating frequency that can occur due to fabrication error. Having collected the near-field data, the far-field (FF) power patterns were generated using the built-in near-field to far-field algorithm of the NSI-MI 2000 software which takes into account probe correction. The power pattern at each measured frequency was analyzed to determine the operating frequency based on the match to the desired masks.

For the fabricated metasurface, the closest match to the simulated power pattern was observed at 9.9 GHz. The power patterns at 9.9 GHz and 10 GHz as well as the simulation result are shown in Fig. 7.5. As illustrated, the fabricated metasurface effectively generates a trapezoidal-shaped far-field pattern at the desired location while maintaining sidelobes at a reasonable level, in accordance with the desired far-field masks. However, there are more fluctuations within the trapezoidal shape compared to the simulation. Several factors can contribute to this. First, the incident wave generated by a horn antenna illuminating the metasurface is not a perfect plane wave as required. Even with the horn antenna placed 1.5 meters (about 50 wavelengths) away from the metasurface (limited by the RF cable length and the anechoic chamber size), the

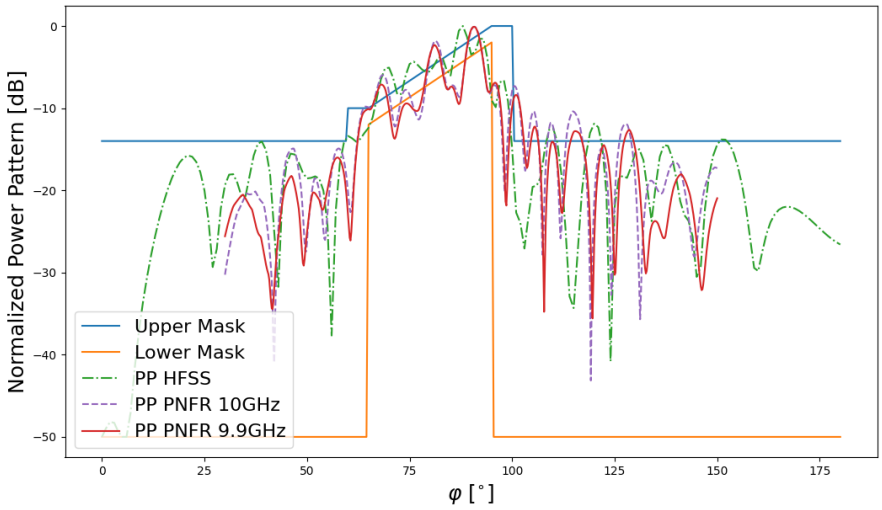
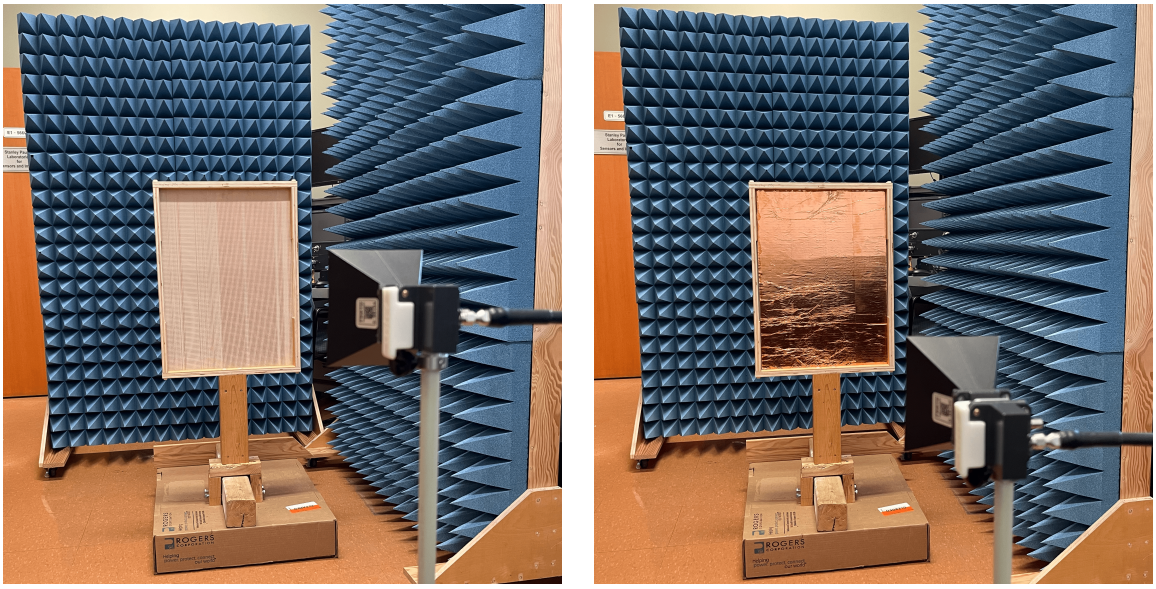


Fig. 7.5 Trapezoidal-shaped far-field (FF) mask and power patterns (PP) obtained from Ansys HFSS simulation and PNFR measurements at 9.9 GHz and 10 GHz.

incident fields that reach the metasurface still exhibit amplitude and phase deviations from an ideal plane wave. These deviations can excite the unit cells differently than designed. Secondly, the designed metasurface is assumed to be 2D, extending infinitely in the periodic direction. This is impossible to fabricate, resulting in some errors associated with this truncation. Thirdly, there could also be errors associated with the manufacturing process. Finally, the most important error in PNFR measurements is related to multiple reflections between the probe (X-band open ended waveguide antenna) and the metasurface under test. The probe and the metasurface are close to each other (about 3 wavelengths away). Due to the limitation of our PNFR system size and the measurement plane, we were not able to increase their separation.

7.4.2 Evaluation of the Reflection

Since the designed metasurface is supposed to be transmissive, the reflection from the metasurface needs to be ideally zero. One way to evaluate the reflection is to utilize the time-domain gating of a vector network analyzer (VNA); e.g., see [53]. To perform this measurement, the metasurface is irradiated by a source (horn) antenna,



(a) (b)

Fig. 7.6 (a) The metasurface under test illuminated by a horn antenna; (b) calibration measurement where the face of the metasurface is covered by a copper tape.

which is connected to Port 1 of a VNA. First, the time-domain response is captured to confirm the location of the metasurface relative to the calibrated feed location. Time gating is then enabled to exclusively consider the reflection from the metasurface. Having applied the time gating, the analysis then switches to the frequency domain to record the reflection response of the metasurface. To calibrate the results, the same measurement procedure is repeated with the metasurface covered in copper tape, as the reflection coefficient from a perfect electric conductor should be 0 dB. The final reflection response is obtained by subtracting the reflection response of the metasurface with the copper tape from the original reflection response of the actual metasurface.

For this measurement, we placed the metasurface 155 cm away from the horn's aperture. This measurement setup is shown in Fig. 7.6(a). The distance from the feeding point (the VNA calibrated reference plane) to the horn aperture is approximately 13 cm. The horn antenna was connected to the Keysight PNA Network Analyzer N5224B. A one-port calibration was performed from 8 GHz to 12 GHz with 6401 points, a 0 dBm power level, and an IF bandwidth of 10 kHz. This calibration was performed using

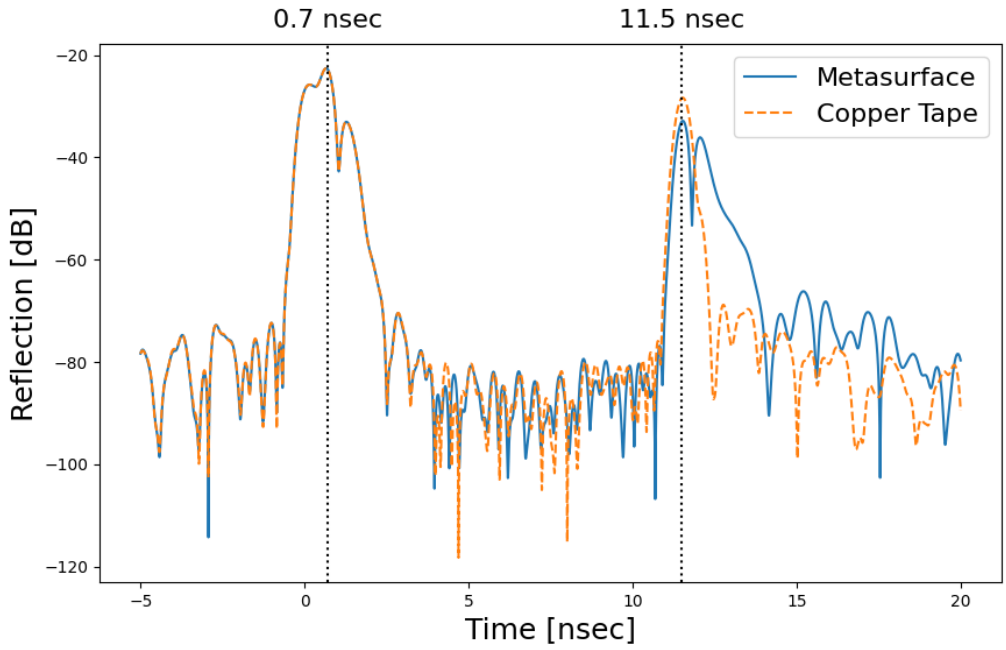


Fig. 7.7 The time-domain reflected signal captured by the VNA for two cases: metasurface under test, and the same metasurface covered by a copper tape.

the Keysight Electronic Calibration (ECal) Module N7555A to establish the reference plane at the feeding point of the horn antenna. (This was similar to the procedure used in [53].)

When the VNA is switched to the time-domain option and the range is set from −5 nsec (before the reference plane) to 20 nsec, we observe two peaks as shown in Fig. 7.7. The first peak occurs at 0.7 nsec, indicating a reflection round trip of 21 cm and a one-way trip of 10.5 cm, which corresponds to the reflection from the horn antenna itself. The second peak occurs at 11.5 nsec, indicating a reflection round trip of 345 cm and a one-way trip of 172.5 cm, which corresponds to the reflection from the metasurface. This result is expected, as the distance from the reference plane to the metasurface is approximately (155 cm + 13 cm = 168 cm). After applying time gating from 10 nsec to 15 nsec such that we only capture the reflected wave from the metasurface, we arrive at the plot shown in Fig. 7.8.

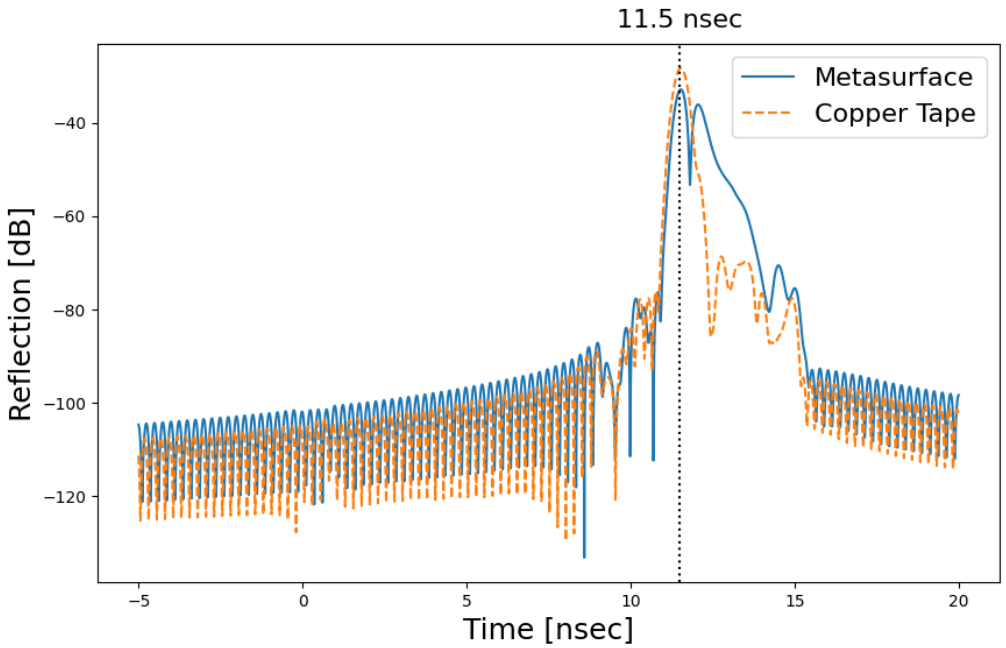


Fig. 7.8 The reflected signal captured by the VNA after applying time gating for two cases: the metasurface under test and the same metasurface covered by a copper tape.

We repeated the same measurement procedure with the metasurface covered by copper tapes where the measurement setup is shown in Fig. 7.6(b) and time-domain responses shown in the dashed lines of Fig. 7.7 and Fig. 7.8. As can be seen, the second peak corresponding to the copper tape occurs at the same time/location as the one with the metasurface. However, the peak magnitude of the reflection is higher for the copper tape, due to its greater reflectivity compared to the metasurface, which indicates that some energy are transmitted through the metasurface. Additionally, the shape of the reflection curve differs from that of the metasurface because the metasurface has three layers, resulting in a different reflective mechanism compared to the copper. Finally, the frequency response with time-gating for both metasurface and copper tapes are shown in Fig. 7.9. As mentioned above, to calibrate the results, the frequency response of the metasurface was subtracted from that of the metasurface with the copper tape, as depicted in Fig. 7.10. The minimum reflection occurs at 9.9 GHz, which aligns with the measurements obtained using the PNFR. Additionally, based on the -10 dB

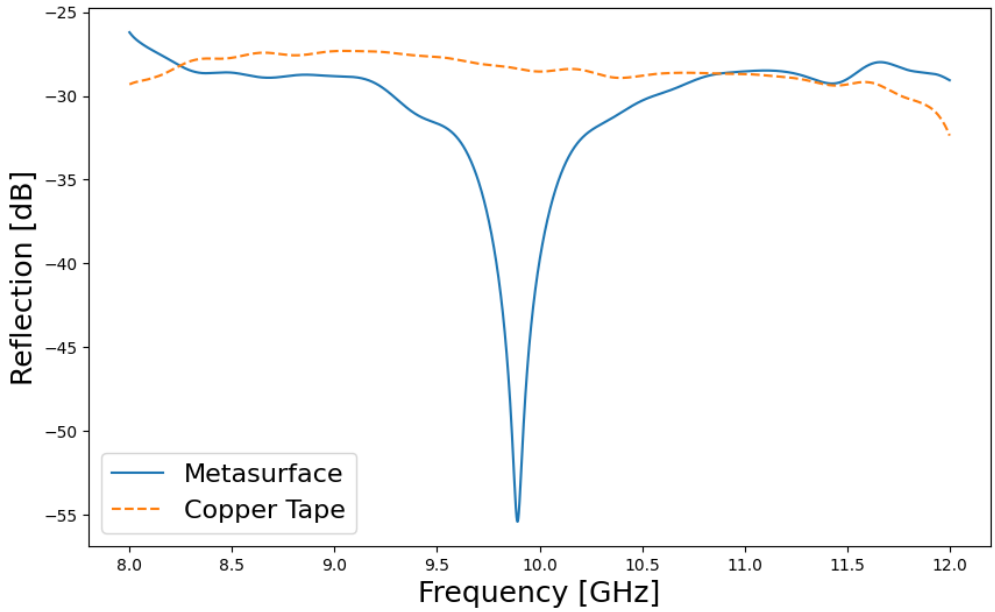


Fig. 7.9 The frequency-domain reflection response for the metasurface and the metasurface with a copper tape after applying time gating.

convention, the metasurface exhibits a bandwidth of 0.25 GHz. We also note that due to our simplistic calibration procedure, the final result suffers from calibration errors which for example show itself as greater than 0 dB reflection (i.e., gain) at around 8 GHz and 12 GHz in Fig. 7.10.

7.5 Conclusions

We have combined two previously developed neural networks to present an end-to-end DL-based approach that outputs a metallic trace transmissive metasurface design from the user-defined FF masks. When illuminated by the given incident EM wave, this metasurface can yield power patterns that mostly adhere to the FF masks. Through ideal impedance sheet simulations, we have speculated that our dogbone-based design has the potential to be improved by tuning its dimensions and employing thinner substrates.

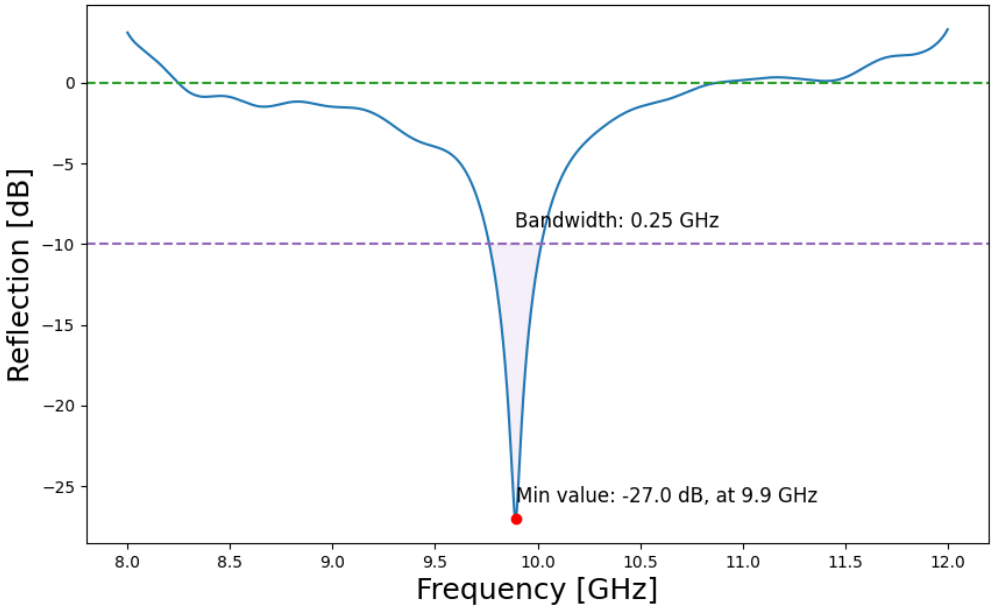


Fig. 7.10 The frequency-domain reflection response for the metasurface after applying time gating and calibration. Note that due to calibration errors, some parts of the reflection are erroneously greater than 0 dB.

We also fabricated a metasurface and experimentally evaluated its performance against a desired FF mask in a PNFR. The experimental result showed that the fabricated metasurface effectively follows the desired mask; however, its pattern was subject to more ripples as compared to the simulated one.

Chapter 8

Conclusions and Future Work

The primary objective of this research was to address challenges in designing radiating systems that can meet user-defined far-field (FF) specifications. This is, in particular, important when the desired specifications do not fit conventional FF patterns. In such cases, the design process can become complex and require many iterations. Additionally, we aimed to develop an affordable solution, which led us to exclude phased array antennas where each radiating element requires separate radio-frequency feeding. Consequently, we chose transmissive metasurfaces where all the unit cells are illuminated with a single antenna, i.e., using one radio-frequency port. Furthermore, to be energy efficient, we focused on lossless, passive, and reflectionless transmissive metasurfaces. As a result, the objective of this thesis was to design lossless, passive, reflectionless, and reciprocal metasurfaces to realize user-defined FF specifications.

To accomplish this, several methodologies are currently being investigated. In this thesis, we decided to focus on deep learning (DL) techniques for the following three reasons. Firstly, our vision is to eventually expand this research to smart (reconfigurable) metasurfaces. This requires reconfiguring the properties of the metasurface (e.g., via PIN diodes) in real time. The DL techniques are suitable for this purpose due to their real-time response. We emphasize that this thesis was focused on fixed functionality (static) metasurfaces, but the methodology developed can be adapted to reconfigurable metasurfaces. Secondly, the copper trace design of metasurfaces typically rely on unit

cell look-up tables. However, these look-up tables are often limited to single layers with specific shapes. Consequently, using these approaches, multi-layered metasurfaces are created by cascading the layers, often ignoring the mutual coupling effect between the layers. We decided to use DL techniques for the copper design step to better capture the effect of mutual coupling into our design by training the neural network using multi-layered unit cells. Lastly, DL techniques can incorporate complicated constraints into its framework. For example, this was used in this thesis to directly incorporate the local power conservation constraint within the DL framework. Based on these three reasons, the DL framework was chosen as the main methodology in this thesis.

8.1 Main Conclusions

Here are the main six conclusions of this research.

1. Our developed DL frameworks are effective in determining the required aperture fields (in our case, the aperture of the metasurface) to meet various forms of desired FF specifications.

2. Our developed DL frameworks, which infers the necessary aperture fields, can be combined with methods incorporating auxiliary surface waves to achieve lossless and passive metasurface designs. However, we have observed that this two-step approach may encounter difficulties in more challenging design scenarios.

3. Our developed DL framework can directly incorporate the lossless and passivity condition (i.e., local power conservation constraint) into its framework, resulting in our single-step scheme. Based on our experience, the single-step scheme outperforms the two-step one.

4. Our developed DL framework is capable of synthesizing the physical design of three-layered unit cells for transmissive metasurfaces. From our experience, the output of the DL technique provides the designer with at least a solid starting point for the design.

5. Integrating the developed DL frameworks for the macroscopic and microscopic design stages, we can achieve an systematic end-to-end DL-based transmissive metasurface design pipeline.

6. Finally, the experimental evaluation of our fabricated metasurfaces show promising results, particularly in achieving the main features of the desired FF specifications. However, increased sidelobe levels and undesired ripples are also observed in the experimental results. To address these issues, modelling errors, such as the local periodicity assumption used in our DL approach, need to be improved.

8.2 Summary and Limitations

Having outlined the primary conclusions of this thesis, let us now review the overall work and key aspects. As noted earlier, our DL frameworks facilitate the design of metasurfaces by guiding antenna designers from user-defined far-field specifications to the physical realization of the metasurfaces. In particular, it focuses on omega-type bianisotropic, lossless, passive, and reciprocal transmissive metasurfaces under linear polarization with the following key user inputs:

- Frequency of operation (within the X-band).
- Desired far-field specifications.
- Electrical size of the metasurface (e.g., 5 to 15 wavelengths).
- Dielectric substrate (e.g., Rogers RO3010) of the metasurface.
- Incident field characteristics.

To meet the overall objective of the thesis, we focused on the following sub-objectives.

8.2.1 Two-step DL based macroscopic metasurface design

- In Chapter 3, we introduced the development of a transformer based neural network that takes user-defined far-field performance criteria and outputs a

source array to meet these specifications. The LPC constraint of the metasurface design was then addressed separately by modifying the resulting field generated by the source array using auxiliary surface waves. (When the tangential fields on both faces of the metasurfaces were determined, the source array was discarded.)

- Multiple metasurfaces with unit cells formed by impedance sheets were designed using these methods for both 2D and 3D scenarios. The performance of these metasurfaces was evaluated using full-wave simulations (by Ansys HFSS). Overall, the metasurfaces performed well in terms of main beam directions, half-power beam widths (HPBWs), and side lobe levels (SLLs). However, achieving the desired specifications for nulls and maximum null levels (MNLs) proved to be more challenging, as discussed in Chapter 3.

- This neural network provides designers with additional freedom to request specifications, such as maximum null levels, that are difficult to achieve with conventional antenna array techniques.

- The main drawback of this approach is the following. Since the LPC constraint is not considered during the initial neural network-generated source array, the power profile of the resulting aperture field can significantly differ from that of the incident field. This can make the subsequent LPC optimization step difficult. However, the main advantage of this two-step approach is that the LPC constraint can be addressed by other techniques such as cascaded metasurface pairs.

8.2.2 Single-step DL based macroscopic metasurface design

- In Chapter 4, we presented the development of a neural network that simultaneously optimize for far-field performance criteria in the form of lower and upper far-field masks and LPC constraint given a specified incident fields.

- The developed DL framework also supports unsupervised learning, allowing the neural network to adapt to arbitrary far-field masks and incident field

requirements. This capability enables more flexible and versatile metasurface design, accommodating a wide range of user specifications without the need of labeled data.

- This single-step macroscopic design is not only faster than its two-step counter parts but also offers greater robustness in the design process.
- Various metasurfaces constructed using impedance sheets were designed to meet different types of far-field masks and evaluated using full-wave simulations. The far-field patterns generated by these metasurfaces generally adhered to the specified masks.
- We further extended this unsupervised learning enabled DL framework for single-step macroscopic metasurface design to 3D far-field masks, as discussed in Chapter 5.
- Although the patterns produced by the metasurfaces generally follow the desired masks, imperfections remain around the main beam features and sidelobes. These issues are partly due to the fact that the DL framework does not prioritize smoother field variations to align with the local periodicity assumption used in unit cell design and does not account for the finite thickness of the substrates.

8.2.3 DL based microscopic metasurface design

- In Chapter 6, we discussed the development of a diffusion model based neural network for the microscopic design of metasurfaces. Given the required macroscopic surface properties as conditional input, the neural network was able to generate multi-channel binary images, which are then translated into multi-layered unit cell metallic trace designs for metasurfaces.
- To evaluate the performance of the microscopic design stage, various metasurfaces were first synthesized using different macroscopic methods followed by full-wave simulations. Their macroscopic properties were then translated into

metallic traces. Metasurfaces designed with macroscopic approaches based on an analytical solution and the two-step DL-based method were discussed in Chapter 6. Metasurfaces designed with macroscopic properties obtained through the single-step DL-based method were discussed in Chapter 7.

- As detailed in Chapter 6, a refracting metasurface was fabricated and its performance was verified using the University of Manitoba's Compact Antenna Test Range (CATR). In Chapter 7, we designed and fabricated a metasurface to produce a trapezoidal-shaped far-field pattern, which was experimentally validated using the Planar Near-Field Antenna Range (PNFR) at the University of Manitoba.

- While this developed neural network offers designers a fast initial metasurface design with reasonable performance and adaptability to produce different shapes of metallic traces, it has some limitations. The trained neural network is quite rigid and can only produce unit cell designs with the same properties it was trained on. These properties include operational frequency range, unit cell dimensions, substrate type, and thickness. Any changes to these properties would require regenerating the dataset and retraining the neural network, which can be time-consuming.

8.3 Future Works

Considering the limitations discussed in the previous section, several improvements and extensions can be made to the work presented in this thesis. These potential developments can be categorized into (i) end-to-end metasurface design and experimental verification for 3D far-field masks, (ii) improve the flexibility of the microscopic design process, (iii) relax the locally periodicity assumption.

8.3.1 Complete metasurface designs from 3D masks

In Chapter 5, we presented a DL framework to infer the necessary aperture fields based on desired 3D far-field masks and the incident field illuminating the metasurface. The next step involves completing the microscopic design by simulating metasurfaces constructed using impedance sheets. Following this, we will proceed to the microscopic stage to design the metallic trace for each unit cell of the metasurface, which can then be fabricated and experimentally verified.

8.3.2 Improve the flexibility of the microscopic design

As we discussed previously, the current deep learning framework for microscopic design is rigid and cannot easily adapt to changes in unit cell dimensions, substrate type, and thickness. To address this limitation, we can modify the current DL framework to use unsupervised learning. This requires incorporating the forward process (i.e. from unit cell metallic trace design to its macroscopic properties) into the training process. There are several ways to achieve this, such as developing a fast, specialized forward solver or training a DL-based surrogate forward model.

8.3.3 Relax the locally periodicity assumption

As discussed several times in this thesis, one of the limitations of the utilized approach is the local periodicity assumption. This assumption implies that for a unit cell to achieve its intended macroscopic properties, it must be surrounded by similar (or ideally identical) unit cells. However, this is not true when designing an aperiodic metasurface, leading to performance degradation. There are several ways to approach this problem. For example, one approach is to use equivalent surface currents [79] as the boundary condition to mimic a more realistic environment when characterizing a given unit cell. Alternatively, we can develop a reconfigurable metasurface that can adjust and fine-tune its inherent macroscopic properties on the fly during operation.

8.4 Closing Remarks

In conclusion, this research has contributed to the design of metasurfaces using deep learning techniques. The methodologies developed offer a structured approach to bridge the gap between the theoretical design and the practical implementation of metasurfaces. Although challenges remain, particularly in enhancing the flexibility and robustness of these designs, the work presented here provides a systematic end-to-end procedure for transmissive metasurfaces design. We sincerely hope that future research can build upon these findings, further advancing the field of electromagnetic metasurface design and applications.

Appendix A

Cascaded Design in a Two-Step Scheme

Cascaded Metasurface

As described in Chapter 4, two techniques may be used in conjunction with the two-step scheme for power redistribution: auxiliary surface waves and cascaded metasurfaces. Since the use of auxiliary surface waves was presented in Chapter 3, herein we show an example using cascaded metasurfaces, see Section 2.1.4, for power redistribution. In particular, we transform a line source to a power pattern with a cosecant-squared shape. This desired shape is defined by upper and lower masks as shown in Fig. A.1. In the cascaded design (similar to Fig. 2.5), the source (line source) is located λ_0 away from the first metasurface and is centered with respect to the metasurface system, as shown in Fig. A.2. The separation between the two metasurfaces is assumed to be $0.75\lambda_0$. In addition, we consider omega-bianisotropic Huygens' metasurfaces at 10 GHz under the two-dimensional TE scenario, and with an aperture size of $15\lambda_0$ along the x axis ($\partial/\partial z = 0$). Each metasurface consists of 151 unit cells. Each unit cell consists of three layers of purely reactive (i.e., passive and lossless) surface impedances that are connected using two identical dielectric substrates. In contrast to Fig. 4.1 that incorporates the LPC constraint into its single-step scheme, the utilized DL approach

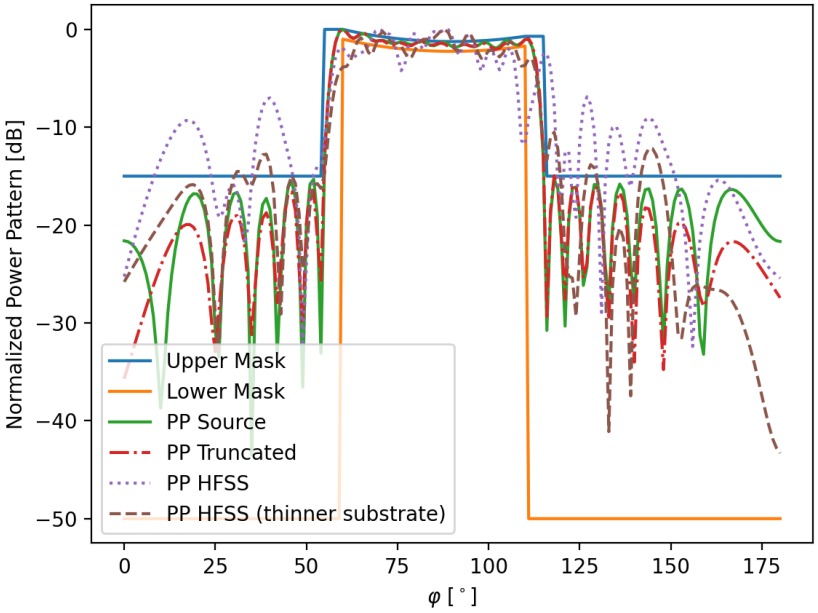


Fig. A.1 Desired power pattern (PP) with a cosecant-squared shape, which is identified by lower and upper masks. The performances of the neural network are denoted by *PP Source* and *PP Truncated* whereas the performances of the metasurface (simulated by Ansys HFSS) employing dielectric substrates with different thickness values are denoted by *PP HFSS* and *PP HFSS (thinner substrate)*.

in this appendix does not incorporate the LPC constraint into its framework, which can be seen in Fig. A.3. Thus, since this DL framework does not incorporate the LPC constraint, the cascaded approach is used to enforce this constraint (i.e, a two-step scheme).

Fig. A.1 demonstrates that both *PP Source*[1] and *PP Truncated*[2] lie within the prescribed masks. Employing Rogers RO3010 dielectric substrate with a thickness of 0.25 mm, relative permittivity of 10.2, and loss tangent of 0.0035 results in a power pattern, denoted by *PP HFSS* in Fig. A.1 that fails to meet the prescribed masks.

[1]The source array generated by the neural network has a power pattern, denoted by *PP Source*.

[2]Using this source array, we find the tangential fields on an aperture with a size of $15\lambda_0$, positioned at a distance of λ_0 away from the source array. This serves as the desired tangential fields on the output face of the metasurface. The power pattern generated by these tangential fields, denoted by *PP Truncated*. We assume that the fields outside the $15\lambda_0$ aperture are zero, hence the term "truncated". This *PP Truncated* therefore takes into account the finite aperture size, and serves as an upper bound for the metasurface's performance.

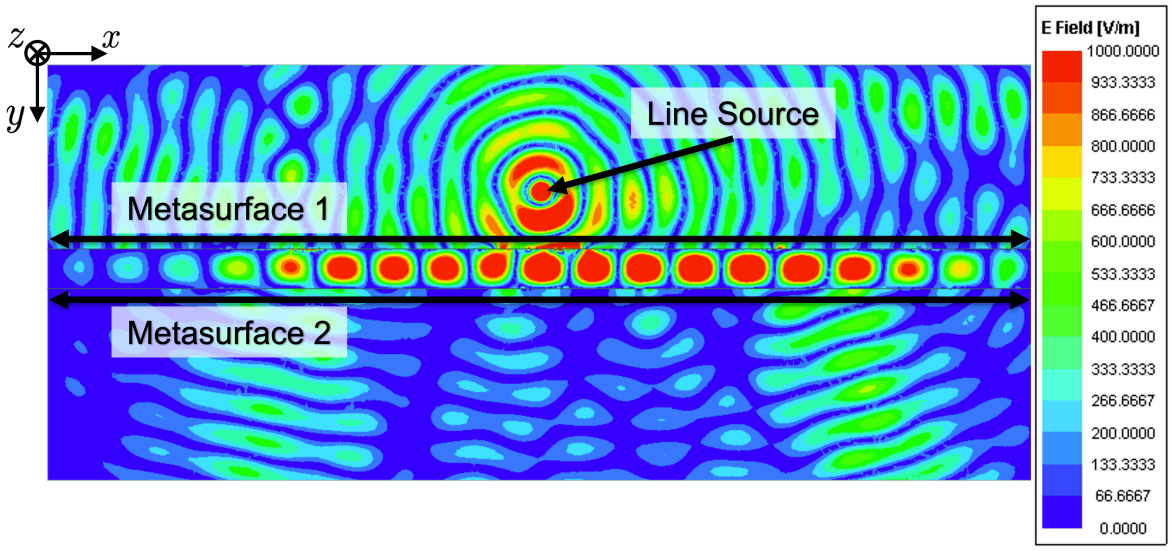


Fig. A.2 Plot of the electric field magnitude at a fixed phase in Ansys HFSS for Example II where a cascaded metasurface system employing a thinner substrate is used. The coordinate center is located at the center of Metasurface 1.

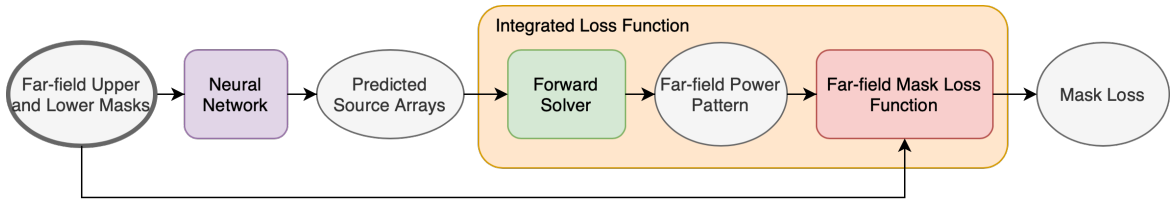


Fig. A.3 The forward path of unsupervised learning for the developed mask-based pattern synthesis (MBPS).

In particular, note the increased side lobe level outside the main beam region for *PP HFSS*. To address this, we reduce the thickness of the substrate by a factor of ten, and re-calculate the impedances for the three-layer unit cells. This adjustment yields a new power pattern, denoted by *PP HFSS (thinner substrate)* in Fig. A.1, which is now much closer to the desired masks. The near-field simulation of this cascaded metasurface system is shown in Fig. A.2.

The poor performance of the original cascaded metasurface system may be attributed in part to the design approach assuming zero thickness metasurfaces and disregarding the finite thickness of the substrate. This then requires further tuning of the final design or utilizing a thinner dielectric substrate.

Appendix B

Additional Data for Chapter 6

In this appendix, we have tabulated the dimensions (in mm) of the dogbones for all the unit cells used for Chapter 6; see Tables B.1-B.3. The top layer is the closest layer to the incident wave (i.e., the input side of the metasurface) and the bottom layer is the closest layer to the generated electromagnetic wave (i.e., the output side of the metasurface).

Table B.1 Unit cell dimensions for the first design example. All dimensions are in mm.

Unit Cell Index	Top W, L	Middle W, L	Bottom W, L
#1	1.133, 1.637	2.267, 3.274	1.007, 1.322
#2	3.778, 3.841	2.833, 1.448	3.400, 3.715
#3	0.881, 3.841	2.833, 1.952	3.211, 1.826
#4	3.904, 1.763	1.322, 2.833	0.944, 3.778
#5	3.904, 2.204	2.078, 1.952	3.841, 1.322
#6	3.715, 2.141	0.881, 2.267	1.511, 2.078
#7	1.574, 2.330	3.715, 3.337	3.211, 1.700
#8	0.630, 3.022	3.904, 1.700	1.574, 2.204

Table B.2 Unit cell dimensions for the second design example. All dimensions are in mm.

Unit Cell Index	Top W, L	Middle W, L	Bottom W, L
#1	2.267, 3.715	0.630, 1.259	2.518, 3.463
#2	0.756, 2.330	1.637, 2.455	1.259, 2.015
#3	0.693, 2.267	1.385, 3.022	2.770, 1.511
#4	1.511, 1.826	0.693, 3.022	1.322, 2.518
#5	3.337, 1.448	0.693, 2.959	1.448, 2.644
#6	1.385, 2.833	1.889, 1.952	0.944, 3.148
#7	1.700, 3.778	1.070, 1.511	3.085, 2.959
#8	0.693, 1.385	0.881, 3.463	0.630, 1.385
#9	1.448, 2.015	2.078, 2.204	1.952, 1.196
#10	0.630, 1.259	0.881, 1.637	1.322, 2.518
#11	3.400, 1.511	2.644, 1.385	0.944, 1.196
#12	1.385, 2.267	1.763, 1.763	1.133, 2.896
#13	2.833, 1.826	0.693, 2.959	3.274, 1.511
#14	1.259, 2.518	0.881, 2.896	1.889, 2.267
#15	1.259, 2.518	0.881, 2.896	1.889, 2.267
#16	1.889, 2.392	1.700, 2.078	0.818, 2.833
#17	1.259, 2.518	0.881, 2.896	1.889, 2.267
#18	1.259, 2.518	0.881, 2.896	1.889, 2.267
#19	2.833, 1.826	0.693, 2.959	3.274, 1.511
#20	1.385, 2.267	1.763, 1.763	1.133, 2.896
#21	3.400, 1.511	2.644, 1.385	0.944, 1.196
#22	0.630, 1.259	0.881, 1.637	1.322, 2.518
#23	1.448, 2.015	2.078, 2.204	1.952, 1.196
#24	0.693, 1.385	0.881, 3.463	0.630, 1.385
#25	1.700, 3.778	1.070, 1.511	3.085, 2.959
#26	1.385, 2.833	1.889, 1.952	0.944, 3.148
#27	3.337, 1.448	0.693, 2.959	1.448, 2.644
#28	1.511, 1.826	0.693, 3.022	1.322, 2.518
#29	0.693, 2.267	1.385, 3.022	2.770, 1.511
#30	0.756, 2.330	1.637, 2.455	1.259, 2.015
#31	2.267, 3.715	0.630, 1.259	2.518, 3.463

Table B.3 Unit cell dimensions for the third example. All dimensions are in mm.

Unit Cell Index	Top W, L	Middle W, L	Bottom W, L
#1	1.448, 1.259	1.133, 3.904	0.944, 3.778
#2	1.826, 1.952	0.881, 3.211	1.070, 3.526
#3	1.637, 1.322	1.322, 3.778	0.881, 3.778
#4	1.133, 1.448	0.818, 3.778	1.007, 3.652
#5	0.693, 1.952	0.693, 3.841	3.463, 1.952
#6	1.007, 2.455	0.630, 3.274	2.833, 1.826
#7	3.589, 1.259	1.448, 2.392	1.070, 2.707
#8	3.589, 1.259	1.448, 2.392	1.070, 2.707
#9	0.756, 1.448	0.693, 2.959	2.518, 1.952
#10	3.337, 1.448	0.693, 2.959	1.007, 2.959
#11	0.756, 3.274	0.693, 3.085	1.322, 2.518
#12	2.959, 2.896	0.881, 1.574	1.763, 3.904
#13	0.630, 3.022	0.630, 3.463	3.841, 1.196
#14	1.574, 1.322	1.133, 3.841	0.944, 3.778
#15	3.652, 3.211	1.007, 1.259	3.022, 2.707
#16	2.015, 1.196	0.944, 3.463	0.818, 1.322
#17	3.841, 1.196	1.637, 2.581	2.959, 1.385
#18	0.630, 1.448	3.778, 1.196	0.630, 3.715
#19	1.763, 1.574	0.693, 3.148	1.007, 2.518
#20	2.078, 2.896	1.826, 3.589	0.944, 3.778
#21	0.630, 1.385	1.637, 2.455	2.770, 1.322
#22	1.070, 1.322	1.133, 3.463	3.841, 3.778
#23	1.259, 2.518	0.881, 2.896	1.889, 2.267
#24	0.630, 3.211	2.707, 1.448	2.015, 2.015
#25	0.944, 2.959	1.259, 2.392	2.392, 1.385
#26	1.574, 1.322	1.133, 3.841	0.944, 3.778
#27	0.881, 2.896	0.818, 3.148	2.707, 3.211
#28	3.211, 1.385	2.078, 1.322	1.826, 2.392
#29	1.196, 2.015	1.259, 3.400	0.630, 1.448
#30	2.959, 2.896	1.322, 1.259	2.204, 3.904
#31	3.085, 2.644	0.630, 1.322	2.833, 3.904
#32	1.007, 1.259	1.133, 3.022	0.881, 2.204
#33	0.944, 1.826	2.770, 1.511	1.196, 2.204
#34	1.133, 2.078	0.693, 3.337	0.881, 2.141
#35	0.756, 2.707	1.196, 2.770	2.455, 1.700
#36	0.756, 2.707	1.196, 2.770	2.455, 1.700
#37	1.574, 1.322	1.133, 3.841	0.944, 3.778
#38	0.693, 1.511	2.078, 3.085	3.841, 1.889
#39	1.448, 1.259	1.133, 3.904	0.944, 3.778

Appendix C

Details on Integrated Loss Function

As discussed in Chapter 4, the DL framework developed for the macroscopic design of metasurfaces from FF masks was trained using unsupervised learning. This approach was enabled by integrating the FF mask loss function and LPC MSE into what we referred to as the integrated loss function. For effective neural network training, the integrated loss function must be differentiable to ensure gradient flow during backpropagation. Specifically, during training, the neural network requires the derivative of the total loss with respect to the predicted source array. This derivative can be obtained either by deriving an analytical expression or by leveraging automatic differentiation implemented in the DL library such as PyTorch. In this work, we utilized the automatic differentiation capability of PyTorch to ensure a modular and adaptive implementation. By constructing the integrated loss function entirely with PyTorch's built-in functions and ensuring that each function used is differentiable, PyTorch automatically computes the derivative of the total loss with respect to the source array using the chain rule. Therefore the goal of this appendix is to confirm the differentiability of all the components within the integrated loss function.

To begin, let us remind that the total loss of the integrated loss function is computed through a weighted sum of FF mask loss and LPC MSE. The FF mask loss can be

expressed through the following:

$$E_{\mathrm{FF}} = \mathcal{F}^{\mathrm{FF}}(s), \tag{C.1a}$$

$$P_{\mathrm{FF}} = 20\log_{10}(\frac{|E_{\mathrm{FF}}|}{\max(|E_{\mathrm{FF}}|)}), \tag{C.1b}$$

$$L_{\mathrm{FF}} = \mathcal{L}^{\mathrm{FF}}(P_{\mathrm{FF}}, M_U, M_L), \tag{C.1c}$$

where s is the predicted source array. That is, s contains the magnitude and phase of each element within the source array, or equivalently its real and imaginary parts. Additionally, $\mathcal{F}^{\mathrm{FF}}$ is the operator used to compute the electric field in the FF zone, E_{FF}. (Therefore, this operator is basically the Green's function operator.) The normalized power pattern is denoted as P_{FF}. Moreover, $\mathcal{L}^{\mathrm{FF}}$ represents the FF mask loss function described in (4.2) which computes the FF mask loss L_{FF}. In addition, M_U and M_L represent the upper and lower FF mask respectively. Similarly, we can express the LPC MSE through the following:

$$E_{\mathrm{NF}} = \mathcal{F}^{\mathrm{NF},E}(s), \tag{C.2a}$$

$$H_{\mathrm{NF}} = \mathcal{F}^{\mathrm{NF},H}(s), \tag{C.2b}$$

$$P_{\mathrm{NF}} = 10\log_{10}(\frac{\mathrm{Re}(E_{\mathrm{NF}}H^*_{\mathrm{NF}})}{\max(\mathrm{Re}(E_{\mathrm{NF}}H^*_{\mathrm{NF}}))}), \tag{C.2c}$$

$$L_{\mathrm{NF}} = MSE(P_{\mathrm{NF}}, P_{\mathrm{in}}). \tag{C.2d}$$

In this case, $\mathcal{F}^{\mathrm{NF},E}$ and $\mathcal{F}^{\mathrm{NF},E}$ are the operators that compute the electric field E_{NF} and magnetic field H_{NF} in the NF zone respectively. (Since we are focused on the tangential fields over the metasurface, this is implied that these near-field electric and magnetic fields are the tangential fields for our 2D TE problem.) The LPC MSE L_{NF} is computed using the MSE between the normalized NF power profile P_{NF} produced by the source array and the known incident power profile P_{in}.

For the 2D TE case, the NF and FF operator $\mathcal{F}$ are implemented using the 2D Green's function as follows,

$$G_{\boldsymbol{E}}(\boldsymbol{r},\boldsymbol{r}') = \frac{-\omega\mu_0}{4} H_0^{(2)}\left(k_0\left|\boldsymbol{r}-\boldsymbol{r}'\right|\right), \tag{C.3}$$

$$G_{\boldsymbol{H}}(\boldsymbol{r},\boldsymbol{r}') = \frac{-k_0(y-y')}{4j\left|\boldsymbol{r}-\boldsymbol{r}'\right|} H_1^{(2)}\left(k_0\left|\boldsymbol{r}-\boldsymbol{r}'\right|\right), \tag{C.4}$$

$$E_z(\boldsymbol{r}) = \int G_{\boldsymbol{E}}(\boldsymbol{r},\boldsymbol{r}')\, s(\boldsymbol{r}')\, d\boldsymbol{r}', \tag{C.5}$$

$$H_x(\boldsymbol{r}) = \int G_{\boldsymbol{H}}(\boldsymbol{r},\boldsymbol{r}')\, s(\boldsymbol{r}')\, d\boldsymbol{r}', \tag{C.6}$$

where $H_0^{(2)}$ is the zeroth-order Hankel function of the second kind, $H_1^{(2)}$ is the first-order Hankel function of the second kind, and k_0 is the wavenumber in free space. The observation vector $\boldsymbol{r}$ can be in the FF zone denoted by $\boldsymbol{r}_{\mathrm{FF}}$ or in the NF zone denoted by $\boldsymbol{r}_{\mathrm{NF}}$. On the other hand, $\boldsymbol{r}'$ indicates the location of the synthetic sources (infinite line sources along $\hat{z}$). In what follows, for simplicity, we drop the subscripts z and x for the tangential electric and magnetic fields respectively. In practice, the integral can be approximated as a matrix-vector product, with the Green's function acting as a precomputed operator. As can be seen, the derivative of E_z and H_x with respect to the source array s is straightforward as their relations are governed by a linear integral operator.

The derivative expression for (C.1b) and (C.2c) can be implemented by performing all operations within PyTorch, leveraging its automatic differentiation and the chain rule. However, it is important to note that differentiation involving operations that transform complex numbers into real numbers is not trivial. For instance, consider taking the derivative of $\|x\|^2$ with respect to x where x is a complex number and $\|\cdot\|$ is the L_2 norm (thus, this is a mapping from complex numbers to real numbers). This is because such expressions are *not* (strictly) complex differentiable. To address this, PyTorch utilizes Wirtinger calculus to handle the differentiation of such functions [226]. This is done by treating x and its complex conjugate x^* as two separate variables, and then finding the derivative with respect to x^*, and finally using $\partial\|x\|^2/\partial x^*$ as the

derivative operator in the optimization. This approach has also been used in other gradient-based optimization techniques that involve complex numbers, as demonstrated in [227] and [228, Appendix A].

The normalized power pattern is subsequently fed into the FF mask loss function, while the normalized NF power is input into the LPC MSE. When implemented in PyTorch, all operations within these two loss functions are differentiable. The final loss is then calculated as a weighted sum of these losses. Since each component of the integrated loss function is differentiable, the entire loss function remains differentiable when implemented in PyTorch.

Appendix D

Additional DL Framework Information

In this appendix, we provide additional training details including the training and validation loss for the deep learning frameworks developed in this research. We also provide more implementation details on the neural networks that we have developed in this thesis. In particular, Figs D.1-D.3 show the convergence of the DL training (training loss and validation loss) for 2D MBPS-LPC (Chapter 4), 3D FF mask (Chapter 5), and dogbone design (Chapter 6). As can be seen, all show appropriate convergence.

Finally, the neural network architectures for dSeqVAE, dVAE, and encoder-only transformer used in Chapter 3 are shown in Fig. D.4(a)-(c), whereas the neural network architecture for the MBPS-LPC approach used in Chapter 4 is shown in Fig. D.5.

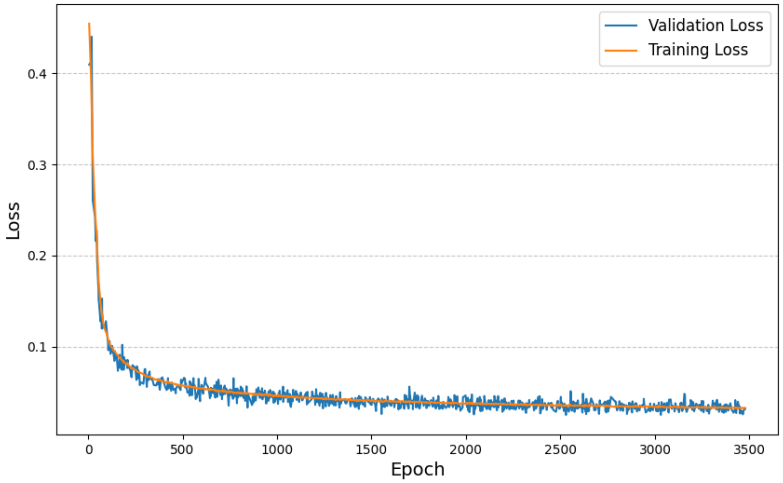


Fig. D.1 Training and validation loss curves for the 2D FF mask used in macroscopic metasurface design as described in Chapter 4. The orange line represents the training loss, while the blue line shows the validation loss over the training epochs.

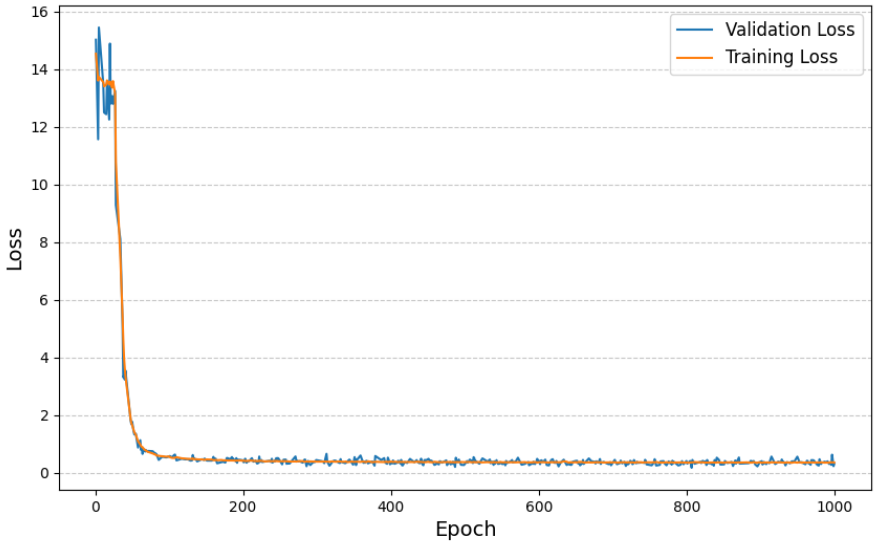


Fig. D.2 Training and validation loss curves for the 3D FF mask used in macroscopic metasurface design as described in Chapter 5. The orange line represents the training loss, while the blue line shows the validation loss over the training epochs.

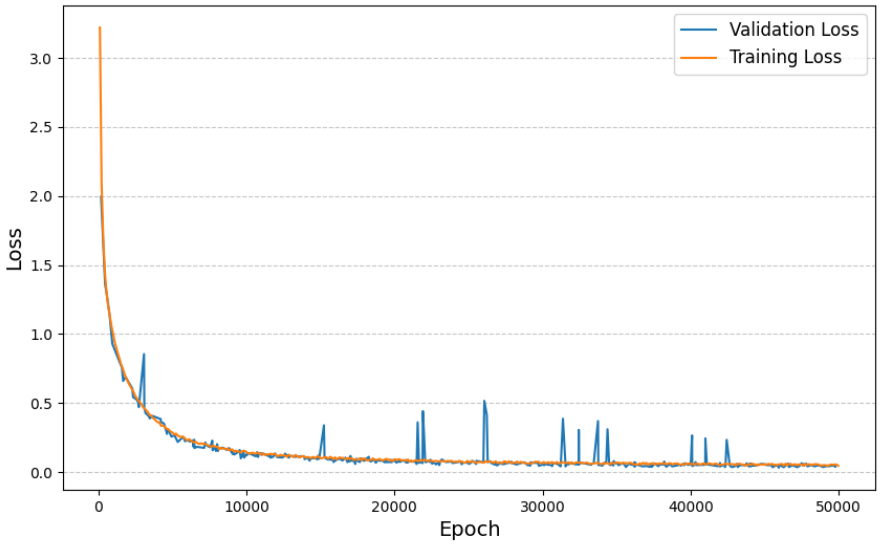
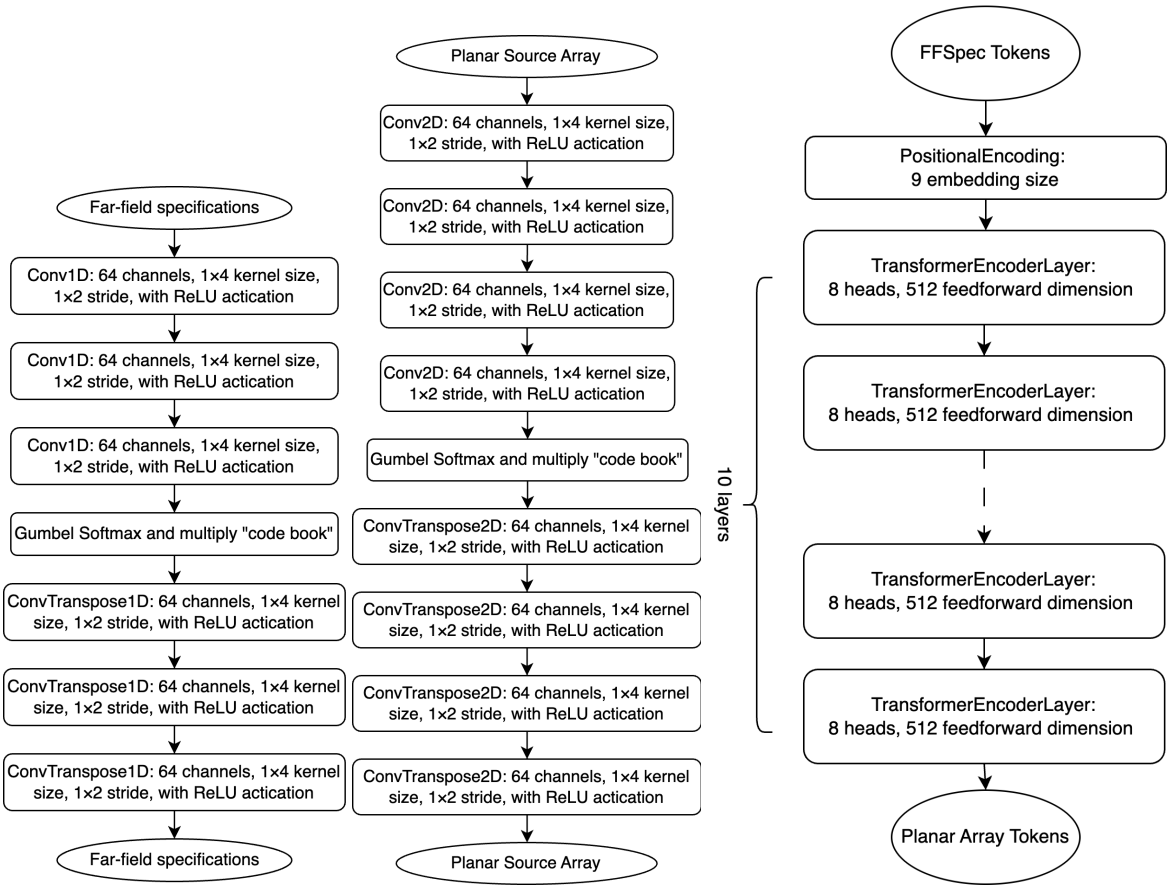


Fig. D.3 Training and validation loss curves for the Diffusion Model used in microscopic metasurface design as described in Chapter 6. The orange line represents the training loss, while the blue line shows the validation loss over the training epochs.



(a) dSeqVAE (b) dVAE (c) Encoder-only transformer

Fig. D.4 The three neural network architectures for the components of the developed DL framework for the far-field specifications to planar source array synthesis as described in Chapter 3, and more specifically in Fig. 3.6.

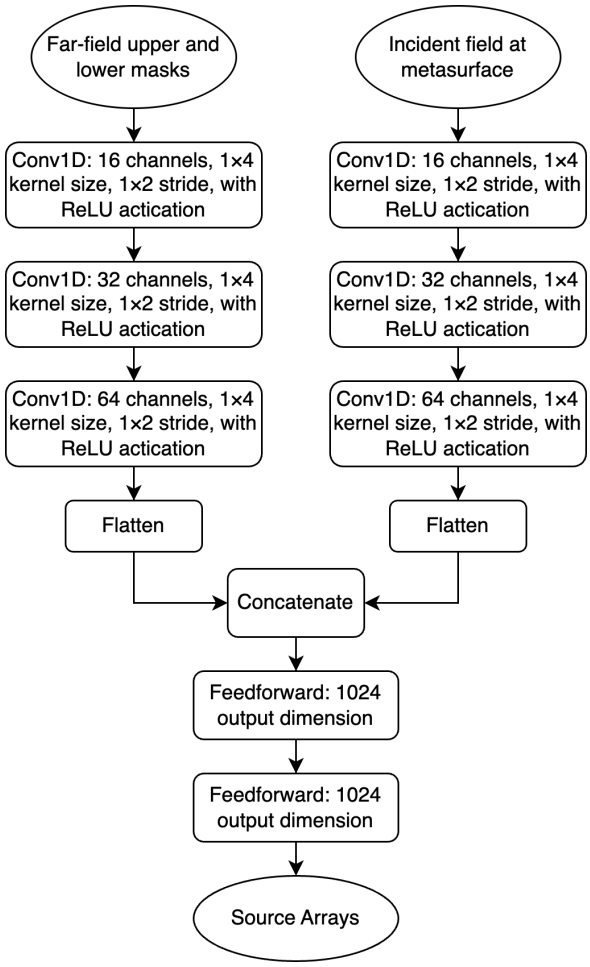


Fig. D.5 The neural network architecture for the DL framework developed for mask-based pattern synthesis with LPC (MBPS-LPC) as describe in Chapter 4, and more specifically in Fig. 4.1.

References

[1] T. K. Sarkar, R. J. Mailloux, A. A. Oliner, M. Salazar-Palma, and D. L. Sengupta, *History of Wireless*. Wiley Interscience, 2006.

[2] B. J. Hunt, "Oliver Heaviside: A first-rate oddity," *Physics Today*, vol. 65, no. 11, pp. 48–54, 11 2012.

[3] "Anniversary of Marconi's first patent," *Nature*, vol. 137, p. 940, 1936.

[4] C. Balanis, *Antenna Theory: Analysis and Design*. New Jersey: John Wiley and Sons, 2005.

[5] R. W. Ziolkowski, "Metamaterials: The early years in the USA," *EPJ Applied Metamaterials*, 2014.

[6] S. Maci, "Electromagnetic metamaterials and metasurfaces: A historical journey," *IEEE Antennas and Propagation Magazine*, pp. 2–19, 2024.

[7] R. M. Walser, "Electromagnetic metamaterials," in *Complex Mediums II: Beyond Linear Isotropic Dielectrics*, vol. 4467, International Society for Optics and Photonics. SPIE, 2001, pp. 1 – 15.

[8] C. Caloz and T. Itoh, *Electromagnetic metamaterials: transmission line theory and microwave applications*. John Wiley & Sons, 2005.

[9] N. Engheta and R. W. Ziolkowski, *Metamaterials: physics and engineering explorations*. John Wiley & Sons, 2006.

[10] S. A. Tretyakov, "A personal view on the origins and developments of the metamaterial concept," *Journal of Optics*, vol. 19, no. 1, p. 013002, 2016.

[11] "Metamorphose vi aisbl," https://www.metamorphose-vi.org/index.php/metamaterials, accessed: 2024-03-14.

[12] 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.

[13] V. G. Veselago, "The electrodynamics of substances with simultaneously negative values of ϵ and μ," *Soviet Physics Uspekhi*, vol. 10, no. 4, pp. 509–514, Apr 1968.

[14] A. Grbic and S. Maci, "EM metasurfaces [Guest Editorial]," *IEEE Antennas and Propagation Magazine*, vol. 64, no. 4, pp. 16–22, 2022.

[15] 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.

[16] C. Pfeiffer and A. Grbic, "Metamaterial Huygens' surfaces: tailoring wave fronts with reflectionless sheets," *Physical review letters*, vol. 110, no. 19, p. 197401, 2013.

[17] G. Minatti, F. Caminita, E. Martini, M. Sabbadini, and S. Maci, "Synthesis of modulated-metasurface antennas with amplitude, phase, and polarization control," *IEEE Transactions on antennas and propagation*, vol. 64, no. 9, pp. 3907–3919, 2016.

[18] 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.

[19] 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.

[20] C. L. Holloway, E. F. Kuester, and A. Dienstfrey, "Characterizing metasurfaces/metafilms: The connection between surface susceptibilities and effective material properties," *IEEE Antennas and Wireless Propagation Letters*, vol. 10, pp. 1507–1511, 2011.

[21] 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.

[22] 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.

[23] 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 the road ahead," *IEEE journal on selected areas in communications*, vol. 38, no. 11, pp. 2450–2525, 2020.

[24] C. Balanis, *Advanced Engineering Electromagnetics*. New York: John Wiley and Sons, 2012.

[25] E. J. Rothwell and M. J. Cloud, *Electromagnetics*, 3rd ed. CRC Press, 2018.

[26] 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.

[27] P.-S. Kildal, *Foundations of Antenna Engineering*. Kildal Anten AB, 2015.

[28] A. Epstein and G. V. Eleftheriades, "Huygens' metasurfaces via the equivalence principle: design and applications," *J. Opt. Soc. Am. B*, vol. 33, no. 2, pp. A31–A50, 2016.

[29] V. G. Ataloglou, M. Chen, M. Kim, and G. V. Eleftheriades, "Microwave Huygens' metasurfaces: Fundamentals and applications," *IEEE Journal of Microwaves*, vol. 1, no. 1, pp. 374–388, 2021.

[30] H. Lamb, "On the reflection and transmission of electric waves by a metallic grating," *Proceedings of the London Mathematical Society*, vol. 29, pp. 523–544, 1898.

[31] R. Mittra, C. H. Chan, and T. Cwik, "Techniques for analyzing frequency selective surfaces-a review," *Proceedings of the IEEE*, vol. 76, no. 12, pp. 1593–1615, 1988.

[32] T.-K. Wu, "Frequency selective surfaces," *Encyclopedia RF Microwave Engineering*, 1995.

[33] J. Huang and J. A. Encinar, *Reflectarray antennas*. John Wiley & Sons, 2007.

[34] A. H. Abdelrahman, F. Yang, A. Z. Elsherbeni, P. Nayeri, and C. A. Balanis, *Analysis and design of transmitarray antennas*. Springer, 2017.

[35] C. Liu, Y. Wu, F. Yang, S. Xu, and M. Li, "Is it possible to design a 1-bit reconfigurable transmitarray element with a single switch?" *IEEE Antennas and Wireless Propagation Letters*, vol. 22, no. 4, pp. 829–833, 2023.

[36] F. Yang and Y. Rahmat-Samii, *Electromagnetic band gap structures in antenna engineering*. Cambridge university press Cambridge, UK, 2009.

[37] K. Achouri, "Synthesis and applications of electromagnetic metasurfaces," Ph.D. dissertation, École Polytechnique de Montréal, 2017.

[38] A. Grbic, L. Jiang, and R. Merlin, "Near-field plates: Subdiffraction focusing with patterned surfaces," *Science*, vol. 320, no. 5875, pp. 511–513, 2008.

[39] D. Sievenpiper, J. Colburn, B. Fong, J. Ottusch, and J. Visher, "Holographic artificial impedance surfaces for conformal antennas," in *2005 IEEE Antennas and Propagation Society International Symposium*, vol. 1B, 2005, pp. 256–259 vol. 1B.

[40] S. Maci, G. Minatti, M. Casaletti, and M. Bosiljevac, "Metasurfing: Addressing waves on impenetrable metasurfaces," *IEEE Antennas and Wireless Propagation Letters*, vol. 10, pp. 1499–1502, 2011.

[41] G. Minatti, F. Caminita, E. Martini, M. Sabbadini, and S. Maci, "Synthesis of modulated-metasurface antennas with amplitude, phase, and polarization control," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 9, pp. 3907–3919, 2016.

[42] N. Yu, P. Genevet, M. A. Kats, F. Aieta, J.-P. Tetienne, F. Capasso, and Z. Gaburro, "Light propagation with phase discontinuities: Generalized laws of reflection and refraction," *Science*, vol. 334, no. 6054, pp. 333–337, 2011.

[43] A. H. Dorrah, "Fixed and scannable pencil beam antennas using metasurfaces and metamaterial structures," Ph.D. dissertation, University of Toronto, 2022.

[44] 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.

[45] J. P. S. Wong, A. Epstein, and G. V. Eleftheriades, "Reflectionless wide-angle refracting metasurfaces," *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 1293–1296, 2016.

[46] V. S. Asadchy, M. Albooyeh, S. N. Tcvetkova, 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.

[47] A. Epstein and G. V. Eleftheriades, "Arbitrary power-conserving field transformations with passive lossless omega-type bianisotropic metasurfaces," *IEEE Trans. Antennas Propag.*, vol. 64, no. 9, pp. 3880–3895, 2016.

[48] J. A. Kong, "Theorems of bianisotropic media," *Proceedings of the IEEE*, vol. 60, no. 9, pp. 1036–1046, 1972.

[49] C. Pfeiffer and A. Grbic, "Bianisotropic metasurfaces for optimal polarization control: Analysis and synthesis," *Physical Review Applied*, vol. 2, no. 4, p. 044011, 2014.

[50] 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.

[51] M. Kim and G. V. Eleftheriades, "Design and demonstration of impedance-matched dual-band chiral metasurfaces," *Scientific reports*, vol. 8, no. 1, p. 3449, 2018.

[52] Y. Ra'di, C. R. Simovski, and S. A. Tretyakov, "Thin perfect absorbers for electromagnetic waves: Theory, design, and realizations," *Phys. Rev. Applied*, vol. 3, p. 037001, Mar 2015.

[53] Z. Liu, N. Bayat, and P. Mojabi, "On the use of absorbing metasurfaces in microwave imaging," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 12, pp. 9026–9031, 2021.

[54] 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.

[55] M. Phaneuf and P. Mojabi, "On solving inverse source problems with metasurfaces performing analog computations," in *2024 18th European Conference on Antennas and Propagation (EuCAP)*, 2024, pp. 1–5.

[56] A. Silva, F. Monticone, G. Castaldi, V. Galdi, A. Alù, and N. Engheta, "Performing mathematical operations with metamaterials," *Science*, vol. 343, pp. 160–163, 2014.

[57] M. El Badawe, "Metasurfaces for antennas, energy harvesting and imaging," Ph.D. dissertation, University of Waterloo, 2018.

[58] M. Phaneuf and P. Mojabi, "On the use of reconfigurable space-time metasurface enclosures for microwave imaging," *IEEE Antennas and Wireless Propagation Letters*, vol. 21, no. 11, pp. 2196–2200, 2022.

[59] 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, 2018.

[60] 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.

[61] 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, p. 98, 2019.

[62] X. Wang and C. Caloz, "Spread-spectrum selective camouflaging based on time-modulated metasurface," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 1, pp. 286–295, 2021.

[63] C. Caloz and Z.-L. Deck-Léger, "Spacetime metamaterials—part i: General concepts," *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 3, pp. 1569–1582, 2020.

[64] Z. Wu, C. Scarborough, and A. Grbic, "Space-time-modulated metasurfaces with spatial discretization: Free-space n-path systems," *Phys. Rev. Appl.*, vol. 14, p. 064060, Dec 2020.

[65] P.-S. Kildal, "Definition of artificially soft and hard surfaces for electromagnetic waves," *Electronics Letters*, vol. 24, pp. 168–170, 1988.

[66] ——, "Artificially soft and hard surfaces in electromagnetics," *IEEE Transactions on Antennas and Propagation*, vol. 38, no. 10, pp. 1537–1544, 1990.

[67] P.-S. Kildal, A. Kishk, and S. Maci, "Special issue on artificial magnetic conductors, soft/hard surfaces, and other complex surfaces," *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 1, pp. 2–7, 2005.

[68] M. Dehmollaian, G. Lavigne, and C. Caloz, "Comparison of tensor boundary conditions with generalized sheet transition conditions," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 12, pp. 7396–7406, 2019.

[69] M. Idemen and A. H. Serbest, "Boundary conditions of the electromagnetic field," *Electronics Letters*, vol. 13, no. 23, pp. 704–705, 1987.

[70] C. Niu, M. Phaneuf, T. Qiu, and P. Mojabi, "A deep learning-based approach to design metasurfaces from desired far-field specifications," *IEEE Open J. Antennas Propag.*, vol. 4, pp. 641–653, 2023.

[71] S. Pearson and S. V. Hum, "Optimization of electromagnetic metasurface parameters satisfying far-field criteria," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 5, pp. 3477–3488, 2022.

[72] G. M. Battaglia, G. G. Bellizzi, A. F. Morabito, G. Sorbello, and T. Isernia, "A general effective approach to the synthesis of shaped beams for arbitrary fixed-geometry arrays," *Journal of Electromagnetic Waves and Applications*, vol. 33, no. 18, pp. 2404–2422, 2019.

[73] M. Zucchi, F. Vernì, M. Righero, and G. Vecchi, "Current based automated design of realizable metasurface antennas with arbitrary pattern constraints," *IEEE Transactions on Antennas and Propagation*, vol. 71, no. 6, pp. 4888–4902, 2023.

[74] S. Pearson and S. V. Hum, "Electromagnetic metasurface beam shaping using far field masks," in *IEEE Int. Symp. Antennas Propag. and USNC-URSI Rad. Sci. Mtg.*, 2022, pp. 1878–1879.

[75] T. Brown, C. Narendra, Y. Vahabzadeh, C. Caloz, and P. Mojabi, "On the use of electromagnetic inversion for metasurface design," *IEEE Trans. Antennas Propag.*, vol. 68, no. 3, pp. 1812–1824, 2020.

[76] K. Achouri, M. A. Salem, and C. Caloz, "General metasurface synthesis based on susceptibility tensors," *IEEE Trans. Antennas Propag*, vol. 63, no. 7, pp. 2977–2991, July 2015.

[77] F. S. Cuesta, I. A. Faniayeu, V. S. Asadchy, and S. A. Tretyakov, "Planar broadband huygens' metasurfaces for wave manipulations," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 12, pp. 7117–7127, 2018.

[78] P. Naseri and S. V. Hum, "A generative machine learning-based approach for inverse design of multilayer metasurfaces," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 9, pp. 5725–5739, 2021.

[79] J. Budhu, N. Ventresca, and A. Grbic, "Unit cell design for aperiodic metasurfaces," *IEEE Transactions on Antennas and Propagation*, vol. 71, no. 9, pp. 7387–7394, 2023.

[80] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[81] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks*, vol. 61, pp. 85–117, 2015.

[82] B. Müller, J. Reinhardt, and M. T. Strickland, *Neural networks: an introduction*. Springer Science & Business Media, 2012.

[83] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, pp. 115–133, 1943.

[84] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation, parallel distributed processing, explorations in the microstructure of cognition, ed. de rumelhart and j. mcclelland. vol. 1. 1986," *Biometrika*, vol. 71, pp. 599–607, 1986.

[85] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.

[86] Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE transactions on neural networks and learning systems*, vol. 30, no. 11, pp. 3212–3232, 2019.

[87] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing," *ieee Computational intelligenCe magazine*, vol. 13, no. 3, pp. 55–75, 2018.

[88] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, "Language models are few-shot learners," *arXiv preprint ArXiv:2005.14165*, 2020.

[89] A. Ramesh, M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, M. Chen, and I. Sutskever, "Zero-shot text-to-image generation," OpenAI, San Francisco, California, USA, Tech. Rep., 2021.

[90] Z. Wang, J. Chen, and S. C. Hoi, "Deep learning for image super-resolution: A survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 10, pp. 3365–3387, 2020.

[91] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Q. Weinberger, Eds., vol. 27. Curran Associates, Inc., 2014.

[92] Y. Pu, Z. Gan, R. Henao, X. Yuan, C. Li, A. Stevens, and L. Carin, "Variational autoencoder for deep learning of images, labels and captions," *Advances in neural information processing systems*, vol. 29, 2016.

[93] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko *et al.*, "Highly accurate protein structure prediction with alphafold," *Nature*, vol. 596, no. 7873, pp. 583–589, 2021.

[94] A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon, and E. D. Cubuk, "Scaling deep learning for materials discovery," *Nature*, vol. 624, no. 7990, pp. 80–85, 2023.

[95] A. Tanaka, A. Tomiya, and K. Hashimoto, *Deep learning and physics*. Springer, 2021, vol. 1.

[96] S. D. Campbell and D. H. Werner, *Advances in Electromagnetics Empowered by Artificial Intelligence and Deep Learning*. John Wiley & Sons, 2023.

[97] A. Massa, D. Marcantonio, X. Chen, M. Li, and M. Salucci, "DNNs as applied to electromagnetics, antennas, and propagation—A review," *IEEE Antennas and Wireless Propagation Letters*, vol. 18, no. 11, pp. 2225–2229, 2019.

[98] M. Li and M. Salucci, *Applications of Deep Learning in Electromagnetics: Teaching Maxwell's Equations to Machines*. IET, 2023.

[99] B. Zhang, "Modeling of electromagnetic responses with respect to design variables: Local gradients, surrogate modeling, and deep learning neural network," Ph.D. dissertation, UCLA, 2024.

[100] H. M. Yao and L. J. Jiang, "Machine learning based neural network solving methods for the fdtd method," in *2018 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting*. IEEE, 2018, pp. 2321–2322.

[101] I. Giannakis, A. Giannopoulos, and C. Warren, "A machine learning-based fast-forward solver for ground penetrating radar with application to full-waveform inversion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 7, pp. 4417–4426, 2019.

[102] C. Cui, W. T. Li, X. T. Ye, P. Rocca, Y. Q. Hei, and X. W. Shi, "An effective artificial neural network-based method for linear array beampattern synthesis," *IEEE Trans. Antennas Propag.*, vol. 69, no. 10, pp. 6431–6443, 2021.

[103] M. R. Ghaderi and N. Amiri, "Application of machine learning techniques in phased array antenna synthesis: A comprehensive mini review," *Journal of Communications*, vol. 18, no. 10, 2023.

[104] M. Li, R. Guo, K. Zhang, Z. Lin, F. Yang, S. Xu, X. Chen, A. Massa, and A. Abubakar, "Machine learning in electromagnetics with applications to biomedical imaging: A review," *IEEE Antennas and Propagation Magazine*, vol. 63, no. 3, pp. 39–51, 2021.

[105] K. Edwards, N. Geddert, K. Krakalovich, R. Kruk, M. Asefi, J. Lovetri, C. Gilmore, and I. Jeffrey, "Stored grain inventory management using neural-network-based parametric electromagnetic inversion," *IEEE Access*, vol. 8, pp. 207 182–207 192, 2020.

[106] A. Fedeli, "Microwave tomography with lstm-based processing of the scattered field," *IEEE Open Journal of Antennas and Propagation*, vol. 2, pp. 213–223, 2021.

[107] W. Ma, F. Cheng, Y. Xu, Q. Wen, and Y. Liu, "Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy," *Advanced Materials*, vol. 31, no. 35, p. 1901111, 2019.

[108] J. A. Hodge, K. Vijay Mishra, and A. I. Zaghloul, "Multi-discriminator distributed generative model for multi-layer RF metasurface discovery," *GlobalSIP 2019 - 7th IEEE Global Conference on Signal and Information Processing, Proceedings*, no. January 2020, 2019.

[109] S. An, B. Zheng, H. Tang, M. Y. Shalaginov, L. Zhou, H. Li, M. Kang, K. A. Richardson, T. Gu, J. Hu, C. Fowler, and H. Zhang, "Multifunctional metasurface design with a generative adversarial network," *Advanced Optical Materials*, vol. 9, no. 5, p. 2001433, 2021.

[110] V. G. Ataloglou, A. H. Dorrah, and G. V. Eleftheriades, "Design of compact Huygens' metasurface pairs with multiple reflections for arbitrary wave transformations," *IEEE Trans. Antennas Propag.*, vol. 68, no. 11, pp. 7382–7394, 2020.

[111] J. Gohel, M. Phaneuf, and P. Mojabi, "Design and experimental evaluation of cascaded metasurface pairs," in *2023 IEEE Conference on Antenna Measurements and Applications (CAMA)*, 2023, pp. 985–990.

[112] O. 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, 1994.

[113] T. Brown, C. Narendra, and P. Mojabi, "On the use of the source reconstruction method for metasurface design," in *12th European Conference on Antennas and Propagation (EuCAP 2018)*, 2018, pp. 1–4.

[114] J. Budhu, L. Szymanski, and A. Grbic, "Design of planar and conformal, passive, lossless metasurfaces that beamform," *IEEE Journal of Microwaves*, vol. 2, no. 3, pp. 401–418, 2022.

[115] T. Shan, X. Pan, M. Li, S. Xu, and F. Yang, "Coding programmable metasurfaces based on deep learning techniques," *IEEE J. Emerg. Sel. Top. Circuits Syst.*, vol. 10, no. 1, pp. 114–125, 2020.

[116] A. Epstein and G. V. Eleftheriades, "Synthesis of passive lossless metasurfaces using auxiliary fields for reflectionless beam splitting and perfect reflection," *Phys. Rev. Lett.*, vol. 117, p. 256103, Dec 2016.

[117] C. Niu, M. Phaneuf, and P. Mojabi, "Transmissive metasurface synthesis from far-field masks using unsupervised learning," *IEEE Antennas and Wireless Propagation Letters*, vol. 23, no. 8, pp. 2371–2375, 2024.

[118] V. G. Ataloglou and G. V. Eleftheriades, "Arbitrary wave transformations with Huygens' metasurfaces through surface-wave optimization," *IEEE Antennas Wirel. Propag. Lett.*, vol. 20, no. 9, pp. 1750–1754, 2021.

[119] M. Phaneuf and P. Mojabi, "An integral equation approach to optimize surface waves for lossless and passive omega-bianisotropic Huygens' metasurfaces," in *IEEE Int. Symp. Antennas Propag. and USNC-URSI Radio Science Meeting (APS/URSI)*, 2022, pp. 1–2.

[120] ——, "An integral equation approach towards the design of compact metasurface pairs," in *2022 16th European Conference on Antennas and Propagation (EuCAP)*, 2022, pp. 1–4.

[121] T. Brown, Y. Vahabzadeh, C. Caloz, and P. Mojabi, "Electromagnetic inversion with local power conservation for metasurface design," *IEEE Antennas Wirel. Propag. Lett.*, vol. 19, no. 8, pp. 1291–1295, 2020.

[122] T. Qiu, M. Phaneuf, and P. Mojabi, "Inverse metasurface design from 3D desired power patterns on two cuts," *IEEE Antennas and Wireless Propagation Letters (Early Access)*, pp. 1–5, 2023.

[123] 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, 2019.

[124] E. Abdo-Sánchez, M. Chen, A. Epstein, and G. V. Eleftheriades, "A leaky-wave antenna with controlled radiation using a bianisotropic huygens' metasurface," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 1, pp. 108–120, 2019.

[125] 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.

[126] P. Naseri, G. Goussetis, N. J. Fonseca, and S. V. Hum, "Synthesis of multiband reflective polarizing metasurfaces using a generative adversarial network," *Scientific Reports*, vol. 12, no. 1, p. 17006, 2022.

[127] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in Neural Information Processing Systems*, vol. 33, pp. 6840–6851, 2020.

[128] A. Dosovitskiy and T. Brox, "Generating images with perceptual similarity metrics based on deep networks," *Advances in neural information processing systems*, vol. 29, 2016.

[129] C. Niu, M. Phaneuf, and P. Mojabi, "A diffusion model for multi-layered metasurface unit cell synthesis," *IEEE Open J. Antennas Propag.*, vol. 4, pp. 654–666, 2023.

[130] S. W. Lin, Q. Ma, and T. J. Cui, "A smart metasurface based on q-learning algorithm," in *2021 2nd International Conference on Electronics, Communications and Information Technology (CECIT)*, 2021, pp. 147–150.

[131] K. Achouri and C. Caloz, *Electromagnetic Metasurfaces: Theory and Applications*. John Wiley & IEEE Press, 2021.

[132] J. A. Kong, *Electromagnetic Wave Theory*. EMW Publishing, 2000.

[133] C. Caloz and A. Sihvola, "Electromagnetic chirality, part 2: The macroscopic perspective [electromagnetic perspectives]," *IEEE Antennas and Propagation Magazine*, vol. 62, no. 2, pp. 82–98, 2020.

[134] ——, "Electromagnetic chirality, part 1: The microscopic perspective [electromagnetic perspectives]," *IEEE Antennas and Propagation Magazine*, vol. 62, no. 1, pp. 58–71, 2020.

[135] M. Dehmollaian, Y. Vahabzadeh, K. Achouri, and C. Caloz, "Limitations of the metasurface diluted-slab model," *IEEE Journal on Multiscale and Multiphysics Computational Techniques*, vol. 5, pp. 255–264, 2020.

[136] 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, 2019.

[137] K. Achouri and O. J. F. Martin, "Angular scattering properties of metasurfaces," *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 1, pp. 432–442, 2020.

[138] S. Tretyakov, *Analytical modeling in applied electromagnetics*. Artech House, 2003.

[139] V. S. Asadchy, M. S. Mirmoosa, A. Diaz-Rubio, S. Fan, and S. A. Tretyakov, "Tutorial on electromagnetic nonreciprocity and its origins," *Proceedings of the IEEE*, vol. 108, no. 10, pp. 1684–1727, 2020.

[140] A. Sihvola, A. Viitanen, I. Lindell, and S. Tretyakov, "Electromagnetic waves in chiral and bi-isotropic media (artech house antenna library)," *Norwood, MA, USA: Artech House*, 1994.

[141] C. Pfeiffer and A. Grbic, "Millimeter-wave transmitarrays for wavefront and polarization control," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 12, pp. 4407–4417, 2013.

[142] D. M. Pozar, *Microwave engineering*. John wiley & sons, 2005.

[143] R. E. Collin, *Field Theory of Guided Waves*, 2nd ed. Wiley-IEEE Press, 1991.

[144] D. F. Sievenpiper, "High-impedance electromagnetic surfaces," Ph.D. dissertation, University of California at Los Angeles, 1999.

[145] M. Phaneuf and P. Mojabi, "Inverse equivalent-source techniques for lossless and passive metasurfaces with noncanonical shapes and geometries," *IEEE Antennas and Wireless Propagation Letters*, vol. 23, no. 7, pp. 2160–2164, 2024.

[146] T. Brown and P. Mojabi, "Cascaded metasurface design using electromagnetic inversion with gradient-based optimization," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 3, pp. 2033–2045, 2022.

[147] B. O. Raeker and A. Grbic, "Compound metaoptics for amplitude and phase control of wave fronts," *Physical Review Letters*, vol. 122, no. 11, p. 113901, 2019.

[148] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[149] H. Taud and J.-F. Mas, "Multilayer perceptron (mlp)," *Geomatic approaches for modeling land change scenarios*, pp. 451–455, 2018.

[150] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[151] K. P. Murphy, *Probabilistic Machine Learning: An introduction*. MIT Press, 2022.

[152] R. Pestourie, Y. Mroueh, T. V. Nguyen, P. Das, and S. G. Johnson, "Active learning of deep surrogates for pdes: application to metasurface design," *npj Computational Materials*, vol. 6, no. 1, p. 164, 2020.

[153] M. Caron, P. Bojanowski, A. Joulin, and M. Douze, "Deep clustering for unsupervised learning of visual features," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 132–149.

[154] W. Wang, Y. Huang, Y. Wang, and L. Wang, "Generalized autoencoder: A neural network framework for dimensionality reduction," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2014, pp. 490–497.

[155] K. P. Murphy, *Probabilistic Machine Learning: Advanced Topics*. MIT Press, 2023.

[156] Y. Ra'di, V. S. Asadchy, and S. A. Tretyakov, "Tailoring reflections from thin composite metamirrors," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 7, pp. 3749–3760, 2014.

[157] C. Pfeiffer and A. Grbic, "Metamaterial Huygens' surfaces: Tailoring wave fronts with reflectionless sheets," *Phys. Rev. Lett.*, vol. 110, p. 197401, May 2013.

[158] 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.

[159] A. Epstein and G. V. Eleftheriades, "Huygens' metasurfaces via the equivalence principle: design and applications," *JOSA B*, vol. 33, no. 2, pp. A31–A50, 2016.

[160] C. Niu, M. Kelly, and P. Mojabi, "An encoder-only transformer to generate power patterns from far-field performance criteria," in *2022 16th European Conference on Antennas and Propagation (EuCAP)*, 2022, pp. 1–4.

[161] S. Zhang, D. Huang, B. Niu, and M. Bai, "High-efficient optimisation method of antenna array radiation pattern synthesis based on multi-layer perceptron network," *IET Microwaves, Antennas & Propagation*, vol. 16, no. 12, pp. 763–770, 2022.

[162] P. Naseri, S. Pearson, Z. Wang, and S. V. Hum, "A combined machine-learning/optimization-based approach for inverse design of nonuniform bianisotropic metasurfaces," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 7, pp. 5105–5119, 2022.

[163] S. Pearson, P. Naseri, and S. V. Hum, "A beam-splitting bianisotropic metasurface designed by optimization and machine learning," *IEEE Open Journal of Antennas and Propagation*, vol. 3, pp. 798–811, 2022.

[164] C. Liu, Q. Ma, Z. J. Luo, Q. R. Hong, Q. Xiao, H. C. Zhang, L. Miao, W. M. Yu, Q. Cheng, L. Li *et al.*, "A programmable diffractive deep neural network based on a digital-coding metasurface array," *Nature Electronics*, vol. 5, no. 2, pp. 113–122, 2022.

[165] T. Shan, X. Pan, M. Li, S. Xu, and F. Yang, "Coding programmable metasurfaces based on deep learning techniques," *IEEE journal on emerging and selected topics in circuits and systems*, vol. 10, no. 1, pp. 114–125, 2020.

[166] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in Neural Information Processing Systems*, vol. 30. Curran Associates, Inc., 2017.

[167] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, J. Davison, S. Shleifer, P. von Platen, C. Ma, Y. Jernite, J. Plu, C. Xu, T. Le Scao, S. Gugger, M. Drame, Q. Lhoest, and A. Rush, "Transformers: State-of-the-art natural language processing," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, Oct. 2020, pp. 38–45.

[168] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, "Transformers in vision: A survey," *ACM computing surveys (CSUR)*, vol. 54, no. 10s, pp. 1–41, 2022.

[169] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[170] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever *et al.*, "Language models are unsupervised multitask learners," *OpenAI blog*, vol. 1, no. 8, p. 9, 2019.

[171] J. Alammar, "The illustrated gpt-2 (visualizing transformer language models)." [Online]. Available: https://jalammar.github.io/illustrated-gpt2

[172] D. Bank, N. Koenigstein, and R. Giryes, "Autoencoders," *arXiv preprint arXiv:2003.05991*, 2020.

[173] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," in *2nd Int. Conf. on Learning Representations ICLR, Banff, Canada*, 2014.

[174] A. van den Oord, O. Vinyals, and K. Kavukcuoglu, "Neural discrete representation learning," *CoRR*, vol. abs/1711.00937, 2017.

[175] C. Niu and P. Mojabi, "Power pattern to planar dipole array synthesis using a text-to-image transformer based model," in *2022 IEEE Int. Symp. Antennas Propag. and USNC-URSI Radio Sci. Mtg.*, 2022.

[176] S. Gregson, J. McCormick, and C. Parini, *Principles of Planar Near-Field Antenna Measurements*. United Kingdom: The Institution of Engineering and Technology (IET), 2007.

[177] M. Phaneuf and P. Mojabi, "On the formulation and implementation of the love's condition constraint for the source reconstruction method," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 5, pp. 3613–3627, 2022.

[178] M. Phaneuf, T. Qiu, and P. Mojabi, "On the importance of the Love's condition for inverse equivalent-source metasurface design," in *2021 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (APS/URSI)*, 2021, pp. 601–602.

[179] J. Y. Lau and S. V. Hum, "Reconfigurable transmitarray design approaches for beamforming applications," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 12, pp. 5679–5689, 2012.

[180] T. Isernia, O. Bucci, and N. Fiorentino, "Shaped beam antenna synthesis problems: Feasibility criteria and new strategies," *Journal of Electromagnetic Waves and Applications*, vol. 12, no. 1, pp. 103–138, 1998.

[181] T. Isernia and A. F. Morabito, "Mask-constrained power synthesis of linear arrays with even excitations," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 7, pp. 3212–3217, 2016.

[182] S. G. Hay and G. T. Poulton, "On the existence of nonsuperdirective aperture antennas with directivity patterns within desired bounds," *Radio Science*, vol. 31, no. 6, pp. 1671–1679, 1996.

[183] D. R. Prado, M. Arrebola, M. R. Pino, R. Florencio, R. R. Boix, J. A. Encinar, and F. Las-Heras, "Efficient crosspolar optimization of shaped-beam dual-polarized reflectarrays using full- wave analysis for the antenna element characterization," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 2, pp. 623–635, 2017.

[184] J. A. Zornoza and J. A. Encinar, "Efficient phase-only synthesis of contoured-beam patterns for very large reflectarrays," *International Journal of RF and Microwave Computer-Aided Engineering*, vol. 14, no. 5, pp. 415–423, 2004.

[185] D. Rodríguez Prado, A. Campa, M. Arrebola, M. R. Pino, J. A. Encinar, and F. Las-Heras, "Design, manufacture, and measurement of a low-cost reflectarray for global earth coverage," *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 1418–1421, 2016.

[186] A. J. Devaney and G. C. Sherman, "Nonuniqueness in inverse source and scattering problems," *IEEE Trans. Antennas Propag.*, vol. 30, no. 5, pp. 1034–1042, Sep 1982.

[187] C. Niu, H. P. Schreckenbach, and P. Mojabi, "Deep learning based metasurface synthesis from far-field masks," in *IEEE Int. Symp. Antennas Propag. and USNC-URSI Rad. Sci. Mtg.*, 2023, pp. 1–2.

[188] C. Narendra, T. Brown, and P. Mojabi, "Gradient-based electromagnetic inversion for metasurface design using circuit models," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 3, pp. 2046–2058, 2022.

[189] A. Epstein and G. V. Eleftheriades, "Arbitrary antenna arrays without feed networks based on cavity-excited omega-bianisotropic metasurfaces," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 4, pp. 1749–1756, 2017.

[190] T. Brown and P. Mojabi, "Cascaded metasurface design using electromagnetic inversion with gradient-based optimization," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 3, pp. 2033–2045, 2022.

[191] B. Zhang, C. Jin, K. Cao, Q. Lv, and R. Mittra, "Cognitive conformal antenna array exploiting deep reinforcement learning method," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 5094–5104, 2022.

[192] C. Cui, W. T. Li, X. T. Ye, Y. Q. Hei, P. Rocca, and X. W. Shi, "Synthesis of mask-constrained pattern-reconfigurable nonuniformly spaced linear arrays using artificial neural networks," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 6, pp. 4355–4368, 2022.

[193] Y. LeCun, B. Boser, J. Denker, D. Henderson, R. Howard, W. Hubbard, and L. Jackel, "Handwritten digit recognition with a back-propagation network," *Advances in neural information processing systems*, vol. 2, 1989.

[194] J. Y. Lau and S. V. Hum, "Reconfigurable transmitarray design approaches for beamforming applications," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 12, pp. 5679–5689, 2012.

[195] C. Niu, M. Phaneuf, and P. Mojabi, "End-to-end transmissive metasurface design from far-field masks using deep learning," in *Submitted to IEEE Int. Symp. Antennas Propag. and USNC-URSI Rad. Sci. Mtg.*, 2024, pp. 1–2.

[196] C. Niu and P. Mojabi, "Synthesis of aperture fields for transmissive metasurface design from 3d masks using deep learning," in *URSI Atlantic Radio Science Meeting (URSI AT-RASC*, Gran Canaria, Spain, May 2024, pp. 1–4.

[197] P. Mojabi, M. Phaneuf, C. Niu, T. Qiu, and M. Kelly, "Electromagnetic inversion for metasurface design–A review," in *URSI EMTS Symp (available online at TechRxiv)*, Canada, May 2023.

[198] C. S. Peirce, "A quincuncial projection of the sphere," *American Journal of Mathematics*, vol. 2, no. 4, pp. 394–396, 1879.

[199] L. Solanilla, A. Oostra, and J. P. Yáñez, "Peirce quincuncial projection," *Revista Integración*, vol. 34, no. 1, pp. 23–38, 2016.

[200] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. Springer, 2016, pp. 21–37.

[201] X. Shi, T. Qiu, J. Wang, X. Zhao, and S. Qu, "Metasurface inverse design using machine learning approaches," *Journal of Physics D: Applied Physics*, vol. 53, no. 27, 2020.

[202] Z. Liu, D. Zhu, S. P. Rodrigues, K.-T. Lee, and W. Cai, "Generative model for the inverse design of metasurfaces," *Nano Letters*, vol. 18, no. 10, p. 6570–6576, Sep 2018.

[203] F. Ghorbani, S. Beyraghi, J. Shabanpour, H. Oraizi, H. Soleimani, and M. Soleimani, "Deep neural network-based automatic metasurface design with a wide frequency range," *Scientific Reports*, vol. 11, no. 1, pp. 1–8, 2021.

[204] M. Soltani, F. Ghorbani, S. Beyraghi, Y. Azizi, and M. Soleimani, "Multifunctional automatic metasurface design using deep learning approach," *TechRxiv*, 2022.

[205] J. Noh, Y.-H. Nam, S. So, C. Lee, S.-G. Lee, Y. Kim, T.-H. Kim, J.-H. Lee, and J. Rho, "Design of a transmissive metasurface antenna using deep neural networks," *Opt. Mater. Express*, vol. 11, no. 7, pp. 2310–2317, Jul 2021.

[206] F. Wen, J. Jiang, and J. A. Fan, "Robust freeform metasurface design based on progressively growing generative networks," *ACS Photonics*, vol. 7, no. 8, pp. 2098–2104, 2020.

[207] M. Li, Y. Zhang, and Z. Ma, "Deep-learning-based metasurface design method considering near-field couplings," *IEEE Journal on Multiscale and Multiphysics Computational Techniques*, vol. 8, pp. 40–48, 2023.

[208] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," in *2nd Int. Conf. on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014.

[209] W. Ma, F. Cheng, Y. Xu, Q. Wen, and Y. Liu, "Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy," *Advanced Materials*, vol. 31, no. 35, p. 1901111, 2019.

[210] P. Naseri, S. Pearson, Z. Wang, and S. V. Hum, "A combined machine-learning/optimization-based approach for inverse design of nonuniform bianisotropic metasurfaces," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 7, pp. 5105–5119, 2022.

[211] S. Pearson, P. Naseri, Z. Wang, and S. V. Hum, "Applying GSTCs in electromagnetic surface inverse design using optimization and machine learning," in *2021 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (APS/URSI)*, 2021, pp. 1511–1512.

[212] Y. Song and S. Ermon, "Generative modeling by estimating gradients of the data distribution," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[213] C. Niu and P. Mojabi, "A variational auto-encoder and transformer based approach for metasurface unit cell synthesis," in *2022 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (AP-S/URSI)*, 2022, pp. 1376–1377.

[214] V. G. Ataloglou, G. Egorov, J. Kim, G. Xu, A. H. Dorrah, A. Ohadi, M. Kim, and G. V. Eleftheriades, "Static and reconfigurable Huygens' metasurfaces: Use in antenna beamforming and beam steering," *IEEE Antennas and Propagation Magazine*, vol. 64, no. 4, pp. 73–84, 2022.

[215] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 10 684–10 695.

[216] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, and S. Ganguli, "Deep unsupervised learning using nonequilibrium thermodynamics," in *International Conference on Machine Learning*. PMLR, 2015, pp. 2256–2265.

[217] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[218] J. Y. Lau and S. V. Hum, "Reconfigurable transmitarray design approaches for beamforming applications," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 12, pp. 5679–5689, 2012.

[219] J. Budhu and A. Grbic, "Patterned unit cell design in aperiodic metasurfaces," in *2022 Sixteenth International Congress on Artificial Materials for Novel Wave Phenomena (Metamaterials)*, 2022, pp. 077–079.

[220] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proceedings of the 36th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, K. Chaudhuri and R. Salakhutdinov, Eds., vol. 97. PMLR, 09–15 Jun 2019, pp. 6105–6114.

[221] M. Kelly, "Toward an end-to-end design procedure for electromagnetic huygens' metasurfaces," Master's thesis, University of Manitoba, Winnipeg, Manitoba, Canada, 2023.

[222] M. Chen, A. Epstein, and G. V. Eleftheriades, "Design and experimental verification of a passive Huygens' metasurface lens for gain enhancement of frequency-scanning slotted-waveguide antennas," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 7, pp. 4678–4692, 2019.

[223] M. Kelly, M. Phaneuf, and P. Mojabi, "A physically realizable metasurface incorporating auxiliary surface waves for power redistribution," in *17th European Conf. on Antennas and Propagation*, Florence, Italy, 2023, pp. 1–5.

[224] S. Pearson and S. V. Hum, "Electromagnetic metasurface beam shaping using far field masks," in *IEEE Int. Symp. Antennas Propag. and USNC-URSI Rad. Sci. Mtg.*, 2022, pp. 1878–1879.

[225] M. Phaneuf, M. Kelly, and P. Mojabi, "Experimental evaluation of forward and inverse solvers for metasurface design," in *2023 XXXVth General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)*, 2023, pp. 1–4.

[226] PyTorch, "Autograd for complex numbers," https://pytorch.org/docs/stable/notes/autograd.html#autograd-for-complex-numbers, 2024.

[227] P. Mojabi and J. LoVetri, "Use of Wirtinger calculus in gauss-newton inversion of microwave tomography data," in *2014 XXXIth URSI General Assembly and Scientific Symposium (URSI GASS)*, 2014, pp. 1–4.

[228] C. Narendra and P. Mojabi, "Phaseless Gauss-Newton inversion for microwave imaging," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 1, pp. 443–456, 2021.