IEEE Open Journal of
Antennas and Propagation

Received 4 May 2023; revised 11 June 2023; accepted 18 June 2023. Date of publication 4 July 2023; date of current version 31 July 2023.

Digital Object Identifier 10.1109/OJAP.2023.3292108

A Deep Learning-Based Approach to Design Metasurfaces From Desired Far-Field Specifications

CHEN NIU (Graduate Student Member, IEEE), MARIO PHANEUF (Graduate Student Member, IEEE), TIANKE QIU (Student Member, IEEE), AND PUYAN MOJABI (Member, IEEE)

Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB R3T 5V6, Canada

CORRESPONDING AUTHOR: C. NIU (e-mail: niuc@myumanitoba.ca)

This work was supported in part by MITACS and in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

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 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.

INDEX TERMS Deep-learning, metasurfaces, pattern synthesis, surface waves.

I. INTRODUCTION

ELECTROMAGNETIC metasurfaces enable the systematic transformation of incoming electromagnetic (EM) waves to desired outgoing ones [1], [2]. These thin devices can tailor the absorption [3], reflection [4], transmission [5] 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 [6]. Moreover, the simultaneous use of temporal and spatial modulations of metasurface properties enables us to control both spatial and spectral attributes of EM waves [7]. These transformations are governed by the generalized sheet transition conditions (GSTCs) [8] 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 [8], polarizabilities [9], or impedances [10]. To design a metasurface, we first need to determine these surface properties. This step, referred to as macroscopic design [11], 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 paper 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 [12]. This can be for achieved, for example, by having non-symmetrical three-layered unit cells [10]. 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 [13], and subsequently, the required surface properties can be determined. However, the tangential EM fields are typically not known in practical situations [14]. 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 [14]. 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 [15].

In the above situation, which is the primary focus of this paper, 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* [14] or deep learning [16] techniques; then, the resulting power pattern can be processed by inversion algorithms to yield the tangential fields on the metasurface [14], [17]. Secondly, the performance criteria can be directly incorporated into an inversion algorithm to infer the required metasurface properties [15]. 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 [18]. 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 deep learning based approach to obtain the required tangential fields from the knowledge of desired far-field performance criteria. Deep learning techniques, which have been previously used for antenna array synthesis, e.g., design based on far-field masks [19], [20], 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) [10], 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 [17]. 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 [17]. That is why the unit cells considered in this paper have three layers.

This paper presents a macroscopic metasurface design procedure using deep learning techniques under two-dimensional (2D) and three-dimensional (3D) EM wave propagation scenarios.[1] 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 paper, the time dependency of $\exp(j\omega t)$ is considered.

1. In addition, deep learning techniques have been used in the microscopic design of metasurfaces [21], [22] and reconfigurable metasurfaces [23], [24].

II. 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 paper, 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 [25], [26], [27], [28]. While these auxiliary surface waves can redistribute the power among unit cells, they do not radiate any real power normal to the metasurface.
4) Once LPC is satisfied, the metasurface is implemented using a three-layer unit cell configuration [13], [29] and is simulated in a full-wave solver (Ansys HFSS) to verify how closely it meets the desired far-field performance criteria.

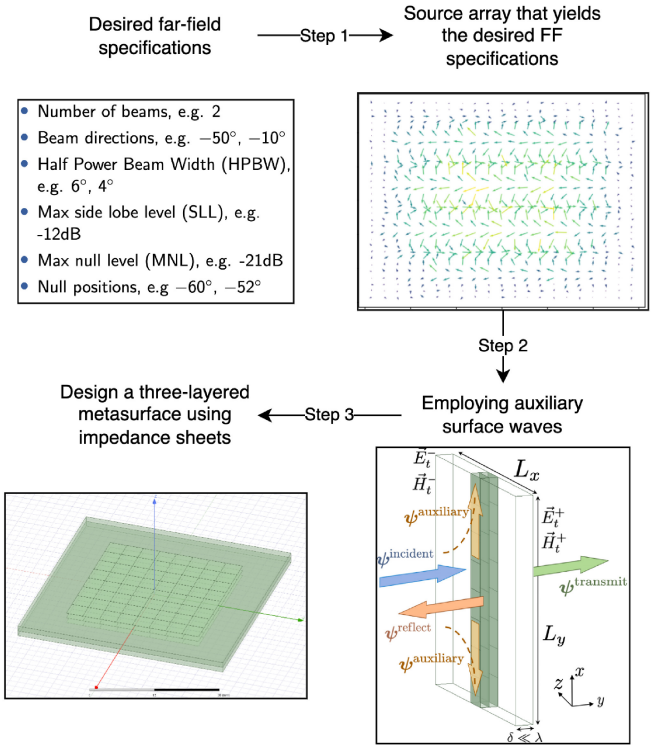


FIGURE 1. The end-to-end metasurface macroscopic design pipeline utilized in this work. The lengths and directions of the arrows in the source array image depict the amplitude and phase of different infinitesimal dipole antennas respectively.

To summarize, a flowchart of the overall design pipeline is shown in Fig. 1. Further details will be presented in the following sections.

III. 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) [30], and thus $\overline{\overline{\chi}}$ is reduced to a 2×2 tensor. The GSTCs can then be written as [8]

$$\hat{n} \times \Delta \boldsymbol{H} = j\omega\epsilon_0 \overline{\overline{\chi}}_{\text{ee}} \boldsymbol{E}_{\text{av}} + jk_0 \overline{\overline{\chi}}_{\text{em}} \boldsymbol{H}_{\text{av}}, \tag{1}$$

$$\hat{n} \times \Delta \boldsymbol{E} = -j\omega\mu_0 \overline{\overline{\chi}}_{\text{mm}} \boldsymbol{H}_{\text{av}} - jk_0 \overline{\overline{\chi}}_{\text{me}} \boldsymbol{E}_{\text{av}}. \tag{2}$$

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 (1) and (2) are then expressed as

$$\Delta\psi = \psi^+ - \psi^-, \qquad \psi_{\text{av}} = \frac{\psi^+ + \psi^-}{2} \tag{3}$$

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 [8]

$$\overline{\overline{\chi}}^{T}_{\text{ee}} = \overline{\overline{\chi}}_{\text{ee}} \in \mathbb{R}^{2\times 2}, \tag{4}$$

$$\overline{\overline{\chi}}^{T}_{\text{mm}} = \overline{\overline{\chi}}_{\text{mm}} \in \mathbb{R}^{2\times 2}, \tag{5}$$

$$\overline{\overline{\chi}}^{T}_{\text{me}} = -\overline{\overline{\chi}}_{\text{em}} \in \mathbb{I}^{2\times 2}, \tag{6}$$

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 [29] (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

$$-\Delta H_x = j\omega\epsilon_0 \chi^{zz}_{\text{ee}} E_{z,\text{av}} - jk_0 \chi^{xz}_{\text{me}} H_{x,\text{av}}, \tag{7}$$

$$-\Delta E_z = j\omega\mu_0 \chi^{xx}_{\text{mm}} H_{x,\text{av}} + jk_0 \chi^{xz}_{\text{me}} E_{z,\text{av}}, \tag{8}$$

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, (7) and (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 (7) and (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 [31], we assume that the cross-polarized components of our illuminating antenna are small, and thus we neglect them for simplicity.

IV. 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^{zz}_{ee}, χ^{xx}_{mm}, and χ^{xz}_{me}. Herein, we develop a neural network based deep learning 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 [32], which was originally developed for text-to-image synthesis applications. Although the implementation details of our neural network differ from the

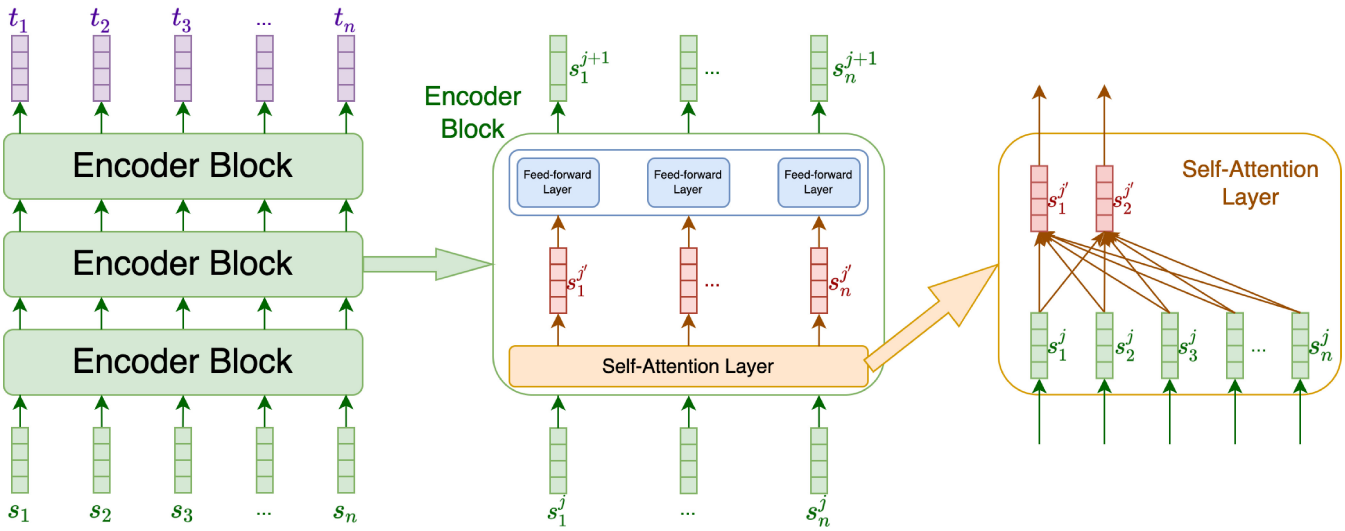


FIGURE 2. The neural network illustration of the encoder-only transformer. The output of the transformer encoder is normally referred to as the hidden state H, but in the case of the encoder-only transformer, H is the final output target sequence T. This Figure is based on [33].

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, 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).

A. 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 [34]. 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[2] 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 can be represented as the mapping of $\mathbf{S}$ to $\mathbf{T}$. In recent years, the transformer based deep learning 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 [35], but have also shown promising results in other areas such as computer vision with applications in text-to-image generation and object detection [36].

The original transformer was designed to be an encoder-decoder based model [34]. 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 [37], [38]. Herein, our implementation is based on encoder-only transformers, which are built by stacking encoder blocks as shown in Fig. 2 (leftmost). Each encoder block then consists of a self-attention layer and a feed-forward layer as depicted in Fig. 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. 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

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

as nulls are aware of other surrounding far-field features such as sidelobes.

In the case of encoder-only transformers, **H** is the final output of the network and should therefore be viewed as the target sequence **T** in Fig. 2 (leftmost). Consequently, the output **T** = **H** will have the same length as the input source **S** in encoder-only transformers. In our case, **S** represents the latent space of far-field performance criteria and **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.

B. DISCRETE VARIATIONAL AUTOENCODERS (DVAE)

We use discrete variational autoencoders (dVAE) [32] 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 [32] 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 $\hat{x}$.[3] 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 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 the unseen data [39]. This is due to the fact that

3. 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.

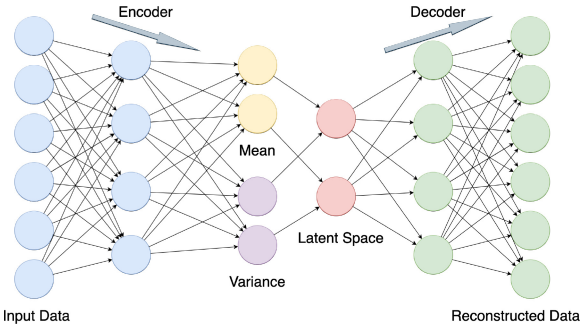


FIGURE 3. The neural network illustration of variational autoencoder (VAE).

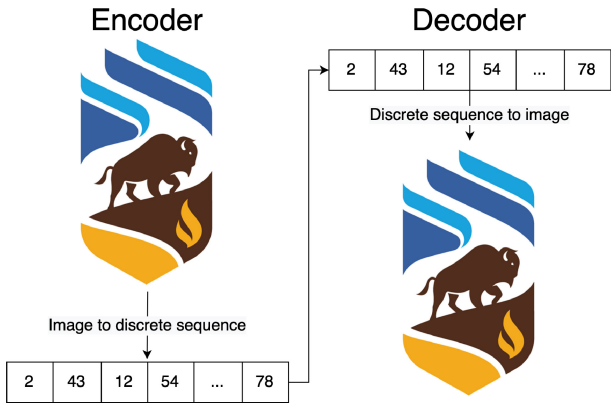


FIGURE 4. The conceptual illustration of discrete variational autoencoders (dVAE) which shows encoding the University of Manitoba's logo to a discrete sequence and then decoding this sequence back in an attempt to arrive at the original logo.

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) [40] overcomes this issue by enforcing a probabilistic prior on the latent space in a standard Gaussian distribution as illustrated in Fig. 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 deep learning techniques that are specifically designed to work with discrete data such as our utilized encoder-only transformer discussed in Section IV-A. Consequently, a number of VAE models are designed to distill discrete latent spaces, such as vector quantised VAE [41] and dVAE [32].

Herein, we focus on the dVAE which was originally designed to work with image data as shown in Fig. 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

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

FIGURE 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.

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.

V. SOURCE ARRAY SYNTHESIS

Based on the above, we now develop a deep learning 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[4] 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 synthesis, the output is a 2D tensor which is not an image-like 3D tensor. In what follows, we address the above challenges.

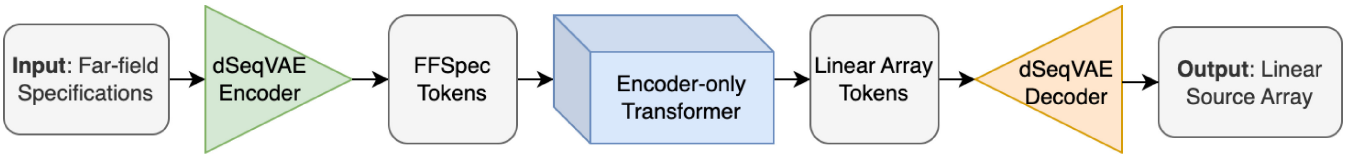
A. FORMATTING FAR-FIELD SPECIFICATIONS TO SEQUENTIAL DATA

In order to make it easy to train the deep learning model, we have vectorized the far-field (FF) performance criteria to a sequential 2D tensor, similar to [16], as shown in Fig. 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. 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.

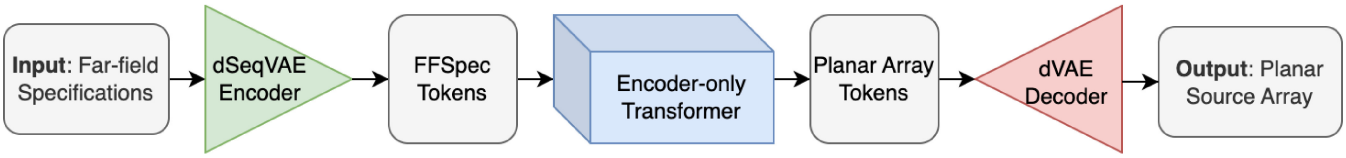
B. ADAPTING THE DVAE TO SOURCE ARRAY SYNTHESIS

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 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*.

4. 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.



(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.

FIGURE 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.

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 [42] 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*.

C. FINAL NETWORK ARCHITECTURE

The complete neural network architecture developed for this work is shown Fig. 6, 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 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.

D. 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 [16].

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.[5]

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. 6, using normalized far-field specifications. The second step is to train the dSeqVAE network, shown in orange in Fig. 6, using linear source arrays. In the third step, the dVAE network, shown in red in Fig. 6, is trained with planar source arrays. Finally, the fourth step involves training 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 [43]. 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 [43]. 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.

VI. 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 [25], [26], [27], [28]. The utilized surface wave optimization algorithm [27] 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 expression at each unit cell in a least-squares manner [27]

$$\frac{1}{2}\mathrm{Re}\left\{E_z^- \times \left(H_x^-\right)^*\right\} - \frac{1}{2}\mathrm{Re}\left\{E_z^+ \times \left(H_x^+\right)^*\right\}, \tag{9}$$

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

$$E_z^- = E_{z,\mathrm{inc}}^- + E_{z,\mathrm{scat}}^-, \tag{10}$$

$$H_x^- = H_{x,\mathrm{inc}}^- + H_{x,\mathrm{scat}}^-, \tag{11}$$

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 [27], [44], [45]. 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. 7, where the zero-field condition in the output region is enforced by the additive constraint. (Note

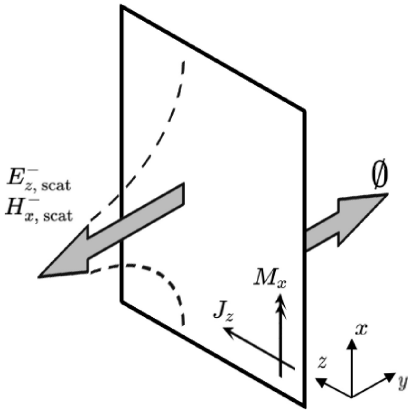


FIGURE 7. The illustration of the equivalent problem with respect to the input region for the scattered fields, which consist of the auxiliary surface waves used to meet the LPC constraint. Love's condition is assumed for the output region of this equivalent problem and is enforced by its equivalent additive constraint during the optimization process.

5. 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.

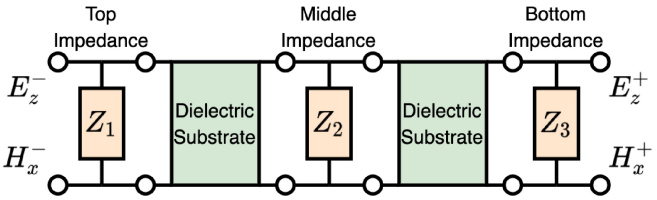


FIGURE 8. The two-port circuit model for a three-layer metasurface unit cell.

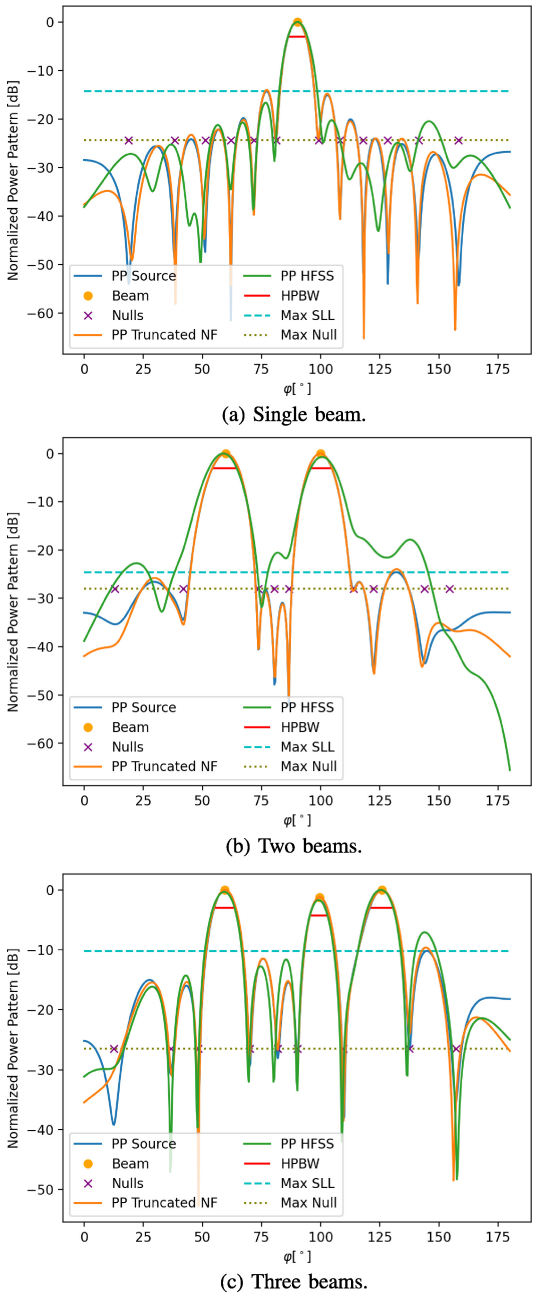
that under Love's condition, the equivalent electric and magnetic currents will be $J_z = H_x$ and $M_x = E_z$ respectively.) The evanescent quality of these scattered fields, depicted with exponential decay in Fig. 7, is obtained by ensuring that the original *total* incident and transmitted powers are the same [27]. Thus, in our approach, we set the tangential EM fields on the input face of the metasurface to (10) and (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 [27] for 2D problems has been extended to 3D for this work.

VII. 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 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 VI), 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. 8 [13]. 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.

A. 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



(a) Single beam.

(b) Two beams.

(c) Three beams.

FIGURE 9. Three scenarios for the 2D design cases.

(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. 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 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 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. 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 [46], [47], which is not fully satisfied especially when we employ high spatial field variations associated with the evanescent spectrum (auxiliary surface waves). Moreover, Fig. 10 shows the NF simulation of the three-beam example (corresponding to Fig. 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%.[6] The sources of power loss are the reflections

6. 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%.

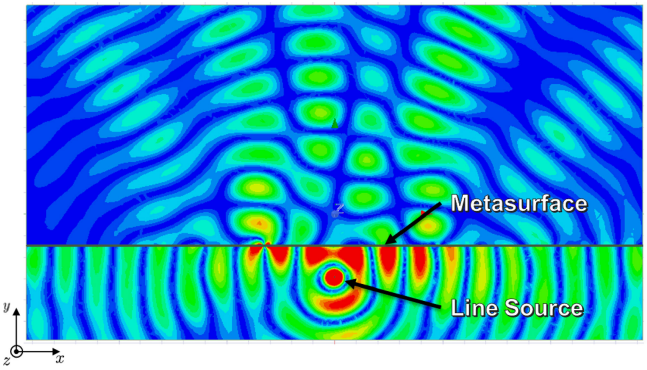


FIGURE 10. Ansys HFSS simulation of the metasurface under the 2D transverse electric illumination, which plots E_z magnitude at a fixed phase.

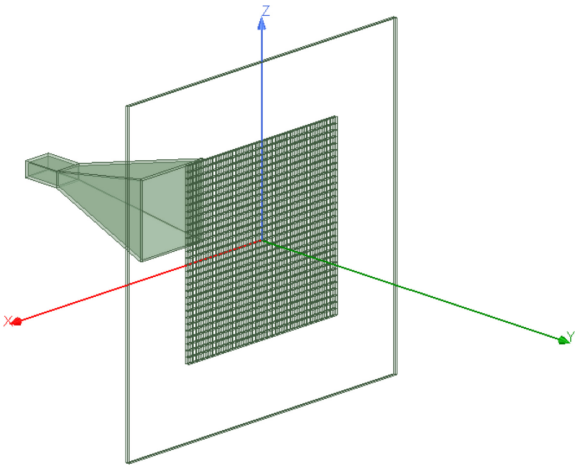
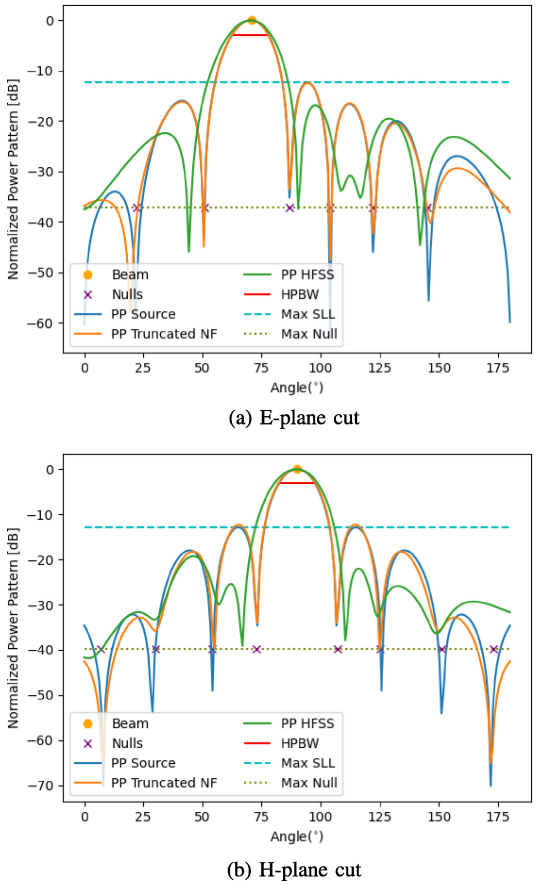


FIGURE 11. Ansys HFSS metasurface simulation set up for the 3D case with an illuminating horn antenna.

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 [28] that the interaction of strong evanescent waves with the metasurface, described in Section VI, results in the further decrease of the power transmission efficiency as compared to metasurfaces that do not utilize these auxiliary surface waves.

B. 3D DESIGN

For the 3D example, the source that illuminates the input face of the metasurface is a horn antenna shown in Fig. 11, which is fed by the TE_{10} mode of a rectangular waveguide. Similar to [31], 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 5λ × 5λ 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. 12, denoted by "Beam", "HPBW", "Nulls", "Max SLL", and "Max Null". Subsequently, the developed neural network



(a) E-plane cut

(b) H-plane cut

FIGURE 12. The 3D example where the desired specifications are given on two perpendicular far-field cuts.

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 [31], [47] 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"

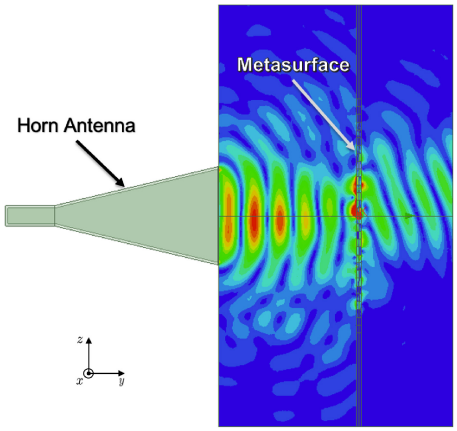


FIGURE 13. The near-field simulation of the 3D metasurface in the E-plane cut (*yz* plane), which plots *E* magnitude at a fixed phase.

in Fig. 12. As can 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. 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.

VIII. CONCLUSION

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

ACKNOWLEDGMENT

The authors would like to thank the Canadian Microelectronics Corporation (CMC) for the provision of ANSYS Campus Solutions.

REFERENCES

[1] 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 Propag. Mag.*, vol. 54, no. 2, pp. 10–35, Apr. 2012.

[2] A. Grbic and S. Maci, "EM metasurfaces [guest editorial]," *IEEE Antennas Propag. Mag.*, vol. 64, no. 4, pp. 16–22, Aug. 2022.

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

[4] Y. Ra'di, V. S. Asadchy, and S. A. Tretyakov, "Tailoring reflections from thin composite metamirrors," *IEEE Trans. Antennas Propag.*, vol. 62, no. 7, pp. 3749–3760, Jul. 2014.

[5] C. Pfeiffer and A. Grbic, "Metamaterial Huygens' surfaces: Tailoring wave fronts with reflectionless sheets," *Phys. Rev. Lett.*, vol. 110, May 2013, Art. no. 197401.

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

[7] Z. Wu, C. Scarborough, and A. Grbic, "Space–time–modulated metasurfaces with spatial discretization: Free-space *N*-path systems," *Phys. Rev. Appl.*, vol. 14, Dec. 2020, Art. no. 064060. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevApplied.14.064060

[8] 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, Jul. 2015.

[9] Y. Ra'Di, C. Simovski, and S. Tretyakov, "Thin perfect absorbers for electromagnetic waves: Theory, design, and realizations," *Phys. Rev. Appl.*, vol. 3, no. 3, 2015, Art. no. 37001.

[10] 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, Sep. 2016.

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

[12] V. G. Ataloglou, M. Chen, M. Kim, and G. V. Eleftheriades, "Microwave Huygens' metasurfaces: Fundamentals and applications," *IEEE J. Microw.*, vol. 1, no. 1, pp. 374–388, Jan. 2021.

[13] 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," *Phys. Rev. B, Condens. Matter, Mater. Phys.*, vol. 97, no. 12, 2018, Art. no. 125433.

[14] 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, Mar. 2020.

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

[16] C. Niu, M. Kelly, and P. Mojabi, "An encoder-only transformer to generate power patterns from far-field performance criteria," in *Proc. 16th Eur. Conf. Antennas Propag. (EuCAP)*, 2022, pp. 1–4.

[17] T. Brown, Y. Vahabzadeh, C. Caloz, and P. Mojabi, "Electromagnetic inversion with local power conservation for metasurface design," *IEEE Antennas Wireless Propag. Lett.*, vol. 19, no. 8, pp. 1291–1295, Aug. 2020.

[18] 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, Nov. 2020.

[19] S. Zhang, D. Huang, B. Niu, and M. Bai, "High-efficient optimisation method of antenna array radiation pattern synthesis based on multilayer perceptron network," *IET Microw. Antennas Propag.*, vol. 16, no. 12, pp. 763–770, 2022.

[20] 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, Oct. 2021.

[21] 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 Trans. Antennas Propag.*, vol. 70, no. 7, pp. 5105–5119, Jul. 2022.

[22] S. Pearson, P. Naseri, and S. V. Hum, "A beam-splitting bianisotropic metasurface designed by optimization and machine learning," *IEEE Open J. Antennas Propag.*, vol. 3, pp. 798–811, 2022.

[23] C. Liu et al., "A programmable diffractive deep neural network based on a digital-coding metasurface array," *Nat. Electron.*, vol. 5, no. 2, pp. 113–122, 2022.

[24] T. Shan, X. Pan, M. Li, S. Xu, and F. Yang, "Coding programmable metasurfaces based on deep learning techniques," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 10, no. 1, pp. 114–125, Mar. 2020.

[25] 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, Dec. 2016, Art. no. 256103.

[26] J. Budhu, L. Szymanski, and A. Grbic, "Design of planar and conformal, passive, lossless metasurfaces that beamform," *IEEE J. Microw.*, vol. 2, no. 3, pp. 401–418, Jul. 2022.

[27] M. Phaneuf and P. Mojabi, "An integral equation approach to optimize surface waves for lossless and passive omega-bianisotropic Huygens' metasurfaces," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (APS/URSI)*, 2022, pp. 1–2.

[28] V. G. Ataloglou and G. V. Eleftheriades, "Arbitrary wave transformations with Huygens' metasurfaces through surface-wave optimization," *IEEE Antennas Wireless Propag. Lett.*, vol. 20, no. 9, pp. 1750–1754, Sep. 2021.

[29] 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 Trans. Antennas Propag.*, vol. 66, no. 3, pp. 1321–1330, Mar. 2018.

[30] K. Achouri and O. J. F. Martin, "Angular scattering properties of metasurfaces," *IEEE Trans. Antennas Propag.*, vol. 68, no. 1, pp. 432–442, Jan. 2020.

[31] T. Qiu, M. Phaneuf, and P. Mojabi, "Inverse metasurface design from 3D desired power patterns on two cuts," *IEEE Antennas Wireless Propag. Lett.*, early access, Apr. 11, 2023, doi: 10.1109/LAWP.2023.3266337.

[32] A. Ramesh et al., "Zero-shot text-to-image generation," OpenAI, San Francisco, CA, USA, Rep. 139, 2021.

[33] J. Alammar. "The illustrated GPT-2 (visualizing transformer language models)." Accessed: May 4, 2023. [Online]. Available: https://jalammar.github.io/illustrated-gpt2

[34] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 5998–6008.

[35] T. Wolf et al., "Transformers: State-of-the-art natural language processing," in *Proc. Conf. Empirical Methods Nat. Lang. Process. Syst. Demonstrations*, Oct. 2020, pp. 38–45. [Online]. Available: https://aclanthology.org/2020.emnlp-demos.6

[36] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, "Transformers in vision: A survey," *ACM Comput. Surveys*, vol. 54, no. 10s, pp. 1–41, 2022.

[37] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*.

[38] A. Radford et al., *Language Models are Unsupervised Multitask Learners*, OpenAI, San Francisco, CA, USA, 2019.

[39] D. Bank, N. Koenigstein, and R. Giryes, "Autoencoders," 2020, *arXiv:2003.05991*.
[40] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in *Proc. 2nd Int. Conf. Learn. Rep. (ICLR)*, 2014, pp. 1–6.
[41] A. van den Oord, O. Vinyals, and K. Kavukcuoglu. "Neural discrete representation learning." 2017. [Online]. Available: http://arxiv.org/abs/1711.00937
[42] C. Niu and P. Mojabi, "Power pattern to planar dipole array synthesis using a text-to-image transformer based model," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Mtg.*, 2022, pp. 493–494.
[43] S. Gregson, J. McCormick, and C. Parini, *Principles of Planar Near-Field Antenna Measurements*. London, U.K.: Inst. Eng. Technol., 2007.
[44] M. Phaneuf and P. Mojabi, "On the formulation and implementation of the Love's condition constraint for the source reconstruction method," *IEEE Trans. Antennas Propag.*, vol. 70, no. 5, pp. 3613–3627, May 2022.
[45] M. Phaneuf, T. Qiu, and P. Mojabi, "On the importance of the Love's condition for inverse equivalent-source metasurface design," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (APS/URSI)*, 2021, pp. 601–602.
[46] J. Y. Lau and S. V. Hum, "Reconfigurable transmitarray design approaches for beamforming applications," *IEEE Trans. Antennas Propag.*, vol. 60, no. 12, pp. 5679–5689, Dec. 2012.
[47] G. Xu, S. V. Hum, and G. V. Eleftheriades, "Augmented huygens' metasurfaces employing baffles for precise control of wave transformations," *IEEE Trans. Antennas Propag.*, vol. 67, no. 11, pp. 6935–6946, Nov. 2019.