Neural-adjoint method for the inverse design of all-dielectric metasurfaces: supplement

YANG DENG, SIMIAO REN, KEBIN FAN, JORDAN M. MALOF, AND WILLIE J. PADILLA*

Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina 27708, USA
*willie.padilla@duke.edu

This supplement published with The Optical Society on 24 February 2021 by The Authors under the terms of the Creative Commons Attribution 4.0 License in the format provided by the authors and unedited. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

Supplement DOI: https://doi.org/10.6084/m9.figshare.14036765

Parent Article DOI: https://doi.org/10.1364/OE.419138
Neural-adjoint method for the inverse design of all-dielectric metasurfaces: supplemental document

1. NUMERICAL SIMULATION

The cylindrical resonators’ geometry was previously demonstrated in the THz regime[1, 2]. To prove deep neural work capability with high-dimensional inputs, we increased the geometrical dimensionality by introducing the elliptical structures to previous cylindrical resonators, and each elliptical resonator undergoes a rotation angle ranging from -45 to 45 degrees. Furthermore, governing the fabrication practicality, we fixed all elliptical resonators to have the same height. To migrate from the THz regime to the infrared, materials and geometry sizes are scaled accordingly. We sized down our unit cell volume refers to the ratio of THz frequency and infrared frequency used in the legacy design and current design. Then an optimization on an adequate scale to finalize our geometry boundary listed in Table 1. We chose silicon carbide (SiC) for our simulations, considering its high melting point, high oxidation resistance, and reasonable absorption coefficient in near-infrared. We used experimentally measured relative permittivity data of SiC from 0 to 300 THz ($\epsilon_r = 10.08 + i0.22$), from which the numerical frequency domain solver performed a frequency dependent fit. To implement the unit cell boundary conditions in CST Microwave Studio we used for our simulation, we used finite frequency solvers to perform the numerical simulations, which also take considerations of the coupling effect between the four resonators within the unit cell. To minimize the time cost of the simulations, we lightly comprise the simulation accuracy to simulation speed. We tuned the Floquet mode to have one mode at both ports. The simulation mesh is tetrahedral, and we used a second-order solver with accuracy at $1 \times 10^{-7}$. The resulting spectrum of the simulation has 2001 data points within the range of 100-500 THz.

| Step | h  | p   | $r_{x_1}/r_{y_1}$ | $\theta_n$ |
|------|----|-----|-------------------|------------|
| 1    | 0.3| 1   | 0.1               | -45        |
| 2    | 0.375| 1.125| 0.1125           | -22.5      |
| 3    | 0.45 | 1.25 | 0.125            | 0          |
| 4    | 0.525| 1.375| 0.1375           | 22.5       |
| 5    | 0.6 | 1.5  | 0.15             | 45         |
| 6    | -  | -   | 0.1625           | -          |
| 7    | -  | -   | 0.175            | -          |
| 8    | -  | -   | 0.1875           | -          |
| 9    | -  | -   | 0.25             | -          |

$n$ corresponds to the first to the fourth elliptical resonator in one super cell

The total possible number of geometrical combinations of our $2 \times 2$ metasurface is $8^4 \times 6^5 = 1.04 \times 10^{12}$. It is impossible to use a conventional numerical simulation approach to exploit the entire geometry space to achieve the targeted spectrum. We find that the average simulation time per geometrical configuration per CPU is approximately three minutes. Thus it would take
about 600 million years to finish exploring the entire geometry space with one CPU. The fast forward dictionary search (FFDS) inverse method was shown feasible for THz ADMs absorbers, where all 812 million possible geometries can be computed in a day [3]. To compute our entire geometry space with a size of over a trillion parameters would take FFDS over three and a half years. Thus the NA method is a good choice when the parameter space become too large for a FFDS approach.

2. DEEP NEURAL NETWORK ARCHITECTURE

We built the entire network and neural-adjoint method using the PyTorch platform. The DNN used for the neural adjoint method consists of twelve fully connected linear layers, four 1D transpose convolutional layers for upsampling, and one final 1D convolutional layer for spectrum smoothing. The linear fully connected layers have the following structure[14, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 500], and all hidden layers except the last linear layer are batch normalized, activated by Leaky_Relu. The choice of how many hidden layers and their widths is crucial and must be carefully considered. If one uses too few neurons underfitting can occur, whereas with too many neurons overfitting (memorizing the training data set) may result. [4] The transpose convolutional layers have kernel size [16, 16, 33, 33] and filter size [4, 4, 4, 4]. The final convolutional layer has kernel size 1 and stride at 1. The data loader takes geometry inputs and the first 2000 data points of the absorptivity spectrum to generate the training and test datasets. The post-processing truncates the predicted spectra’s first and last fifty points to drop the convolutional layers bump at the edges. We use L2 regularization, batch normalization, and the ADAM optimizer.

3. NA INVERSE METHOD

Because NA a gradient-based procedure, it will only converge to a locally optimal solution. Because the NA method is amenable to parallelization on graphics processing units (standard hardware for deep learning), this entire process can be computed in very little time. In our experiments, we can run the NA method with $T = 1000$ on our desktop computer with an Nvidia 2080ti GPU and complete processing in under 1 minute. Since the NA method finds the globally optimal solution even in its worst-performing cases, our results suggest that the NA always (or nearly always) finds the globally optimal solutions, even for highly complex problems like ours. Interestingly, this suggests that the main obstacle of custom design is no longer the inverse model, but rather the space over which we choose to search for designs i.e., the shapes we consider (cylinder, crosses, etc) and their parameter settings (e.g., radii, height). Although powerful, to use deep learning methods we must necessarily define a range of these settings so that we can collect simulations to train our models, and this space limits where we can search for designs. However, the design needed to realize our targeted scattering parameters may not exist in this initial search space. As we show subsequently, the NA method can also be used to identify where this initial search space can be expanded so that it is most likely to contain the desired solution, providing a solution to this emergent obstacle to complex material design.

4. DATA AUGMENTATION

Our simulations’ unit cell boundary conditions allow us to do four times data augmentation on our dataset because the infinite plane of unit cells consists of four different resonators’ combinations that give almost identical spectra with fluctuations from CST software. However, we learned that the DNN could quickly learn the correlation between four different resonators’ combinations. The forward model will know which input geometries share the same spectrum in high fidelity if the entire dataset is augmented before splitting into the training and validation dataset. Therefore, the forward model will give a false mean square error much lower than actual loss performance on an independent validation set. After applying data augmentation to the training and validation sets after the splits, we observe that the four times augmentation did not significantly improve accuracy. We believe that, with our 60000 simulations (24000 simulations after augmentation), the augmented data points are still too sparse to cover the entire geometry space defined by our geometry boundaries.
5. GEOMETRY SPACE EXPLORATION THROUGH UMAP

We use Uniform Manifold Approximation and Projection (UMAP) to explore our solution geometry space and realize that angles have more random impacts on distributing the best NA solutions. Thus, we plotted the UMAP with ten parameters, excluding rotational angles. The plotted UMAP demonstrates a clear trend that the MSE is decreasing in one direction. To confirm that the decreasing trend matches the increasing of resonators’ height, we further marked the points from maximum and minimum height boundaries, respectively. The clustering of points towards the best MSE performance suggests that the NA method is finding the best local minima.

Fig. S1. Uniform Manifold Approximation and Projection plotted with 10-dimensional geometry inputs indicates a strong correlation between the MSE performance and the increasing height.

REFERENCES

1. X. Liu, K. Fan, I. V. Shadrivov, and W. J. Padilla, “Experimental realization of a terahertz all-dielectric metasurface absorber,” Opt. express 25, 191–201 (2017).
2. K. Fan, J. Y. Suen, X. Liu, and W. J. Padilla, “All-dielectric metasurface absorbers for uncooled terahertz imaging,” Optica 4, 601–604 (2017).
3. C. C. Nadell, B. Huang, J. M. Malof, and W. J. Padilla, “Deep learning for accelerated all-dielectric metasurface design,” Opt. express 27, 27523–27535 (2019).
4. J. Heaton, Introduction to neural networks with Java (Heaton Research, St. Louis, 2008).