Physics-Informed Graph Neural Networks for the Inverse Design of GHz Reconfigurable Antenna

Cindy (Hsin) Pan, Naveen Verma, James C. Sturm
Department of Electrical and Computer Engineering
Princeton University
Princeton, NJ, 08540, USA
hp0187@princeton.edu

***Abstract*—Reconfigurable antennas, as a subclass of metasurfaces, offer innovative and dynamic capabilities for wireless communication systems. Specifically, enabling radiation pattern reconfigurability allows for flexible beam steering through reverse-engineering of antenna parameters such as surface current distributions. In this work, we present a physics-informed machine learning model, leveraging fundamental physics such as Kirchhoff's current Law, to predict the switch configurations of 2-dimensional antenna arrays. We utilize a graph neural network (GNN) to effectively capture the spatial relationships between radio-frequency (RF) switches and antenna patches, closely emulating the antenna topology. Simulation results demonstrate that our approach successfully predicts switch configurations needed to generate complex far-field radiation patterns.**

***Index Terms*—inverse design, deep learning, graph neural networks, large area electronics, reconfigurable antenna, metasurfaces, physics-informed machine learning.**

I. Introduction

Recent advancements in deep learning have positioned the inverse design of reconfigurable antennas as a promising approach for antenna design solutions. Traditional design methods rely on iterative simulations and manual optimization [1], [2], which are computationally expensive. State-of-the-art approaches leverage machine learning models, such as deep neural networks and generative models, to accelerate this process by learning the complex relationships between design parameters and performance metrics. Recent studies have demonstrated the use of neural networks for tasks like optimizing beam steering and frequency tuning, showing great promise in improving design efficiency [1], [3].

In this work, we tackle the inverse design of a large-aperture GHz reconfigurable antenna proposed and demonstrated by Can et al. [4], by predicting the switch configurations based on a given current distribution. Since the surface current distribution determines the radiation pattern and the relationship between the far-field pattern and current distribution is well-established [5], [6], numerical methods such as inverse method of moments can be used to analytically solve for the current distributions that produce a desired far-field radiation pattern [7], [8]. On the other hand, the relationship between the switch configuration and the steady state current distribution of a antenna array system is complex and largely unknown. This poses a challenge in antenna inverse-design and applications, as it hinders the ability to dynamically control the radiation pattern through switch manipulation. While deep learning methods such as convolutional neural network could be considered, the graph-based topology of antennas complicates effective feature embedding. To close this gap, we propose a physics-informed GNN framework that predicts the necessary switch configuration to generate a desired current distribution.

II. Inverse Design

A. System Architecture and Feature Embedding

The reconfigurable antenna architecture features RF switches operating in the 2.4 GHz band, interconnecting copper square patches arranged in a 2D array as shown in Fig. 1(a), where the antenna array is mounted on a printed circuit board (PCB) (modeled using "FR4" from CST's material library, with a dielectric constant of 4.5). Beneath the PCB is a layer of alumina dielectric plane modeled using "alumina" from CST's material library, with a dielectric constant of 9.8, followed by an aluminum ground plane. The switch length and the size of the copper patch are both 4.5 mm. A radiation aperture is formed with 7×7 copper patches, covering 5.85cm $\times$ 5.85cm. For simplicity in analysis, we assume the switch configuration below the x-axis is a mirrored version of the configuration above the x-axis. Between each pair of neighboring patches, a switch is placed to distribute current across the copper patches, resulting in 2^{39} possible switch configurations. A reconfigurable antenna of such scale can readily be constructed using Large-Area Electronics [4].

In full-wave EM simulations, each copper patch is discretized into hexahedral meshes, whose collective response serves as that of an approximation of the continuous surface. As shown in Fig. 1(b) left, each hexahedral mesh centered at location index (i, j) carries current with surface current density $J_{i,j}$, which is defined as current per unit length along the surface of a conductor:

$$J_{i,j} = (\Re(J_{x_{i,j}}) + i\Im(J_{x_{i,j}}), \Re(J_{y_{i,j}}) + i\Im(J_{y_{i,j}})) \text{ A/m.} \quad (1)$$

Here, the subscripts x and y represent the vector components of the current in the x-y plane. The color code in Fig. 1(b) indicate the euclidean norm magnitude $|J_{i,j}|$ of the current vector $J_{i,j} = (J_{x_{i,j}}, J_{y_{i,j}})$ for each mesh within a single patch such that $|J_{i,j}| = \sqrt{|J_{x_{i,j}}|^2 + |J_{y_{i,j}}|^2}$. Note that the surface current density at each point on a copper patch varies depending on its steady state response based on the state of

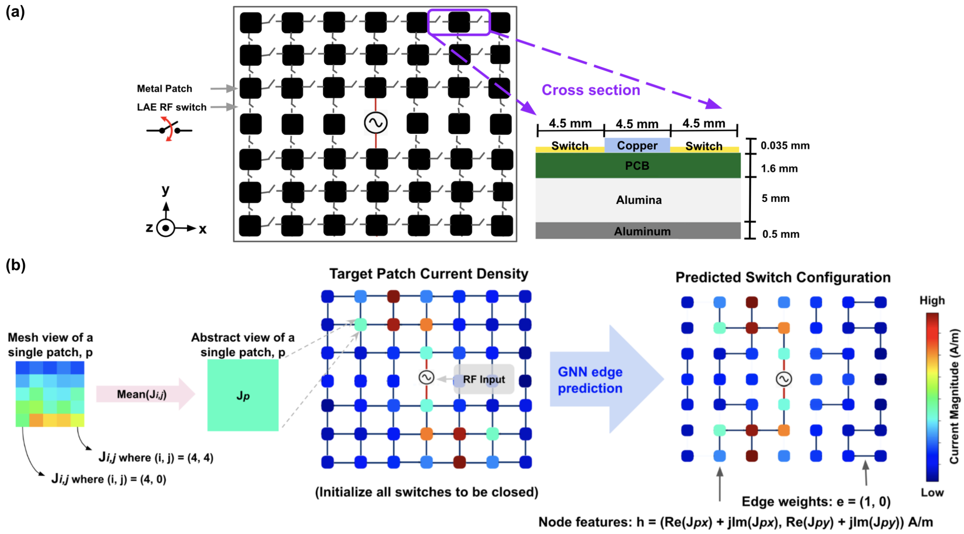

Fig. 1. (a) Left: Top view of the reconfigurable antenna, where a radiation aperture is formed with 7×7 copper patches. Between each neighboring patches, a switch is placed to distribute current across the copper patches. Right: Cross section view of a switch-copper-switch segment of the antenna array. The antenna array is assembled onto a printed circuit board (PCB). Beneath the PCB is a layer of alumina dielectric plane followed by aluminum ground plane. (b) Inverse design of switch configurations based on given surface current distributions of antenna patches (assume symmetry across the x-axis). Left: Data abstraction of the node features is done by averaging the surface current density $J_{i,j}$ within each patch p, yielding a single complex value J_p. Right: Before inverse design, all switches are initialized as closed. After inverse design, predicted switch configurations required to achieve the given surface current distribution is made.

the connecting switches as well as capacitive coupling among neighboring components. By averaging the surface current density $J_{i,j}$ within each patch p, we obtain a single complex value J_p, which will provide an efficient feature embedding representing the surface current density of each patch, where

$$J_p = \text{Mean}(J_{i,j}) = \frac{1}{n*m}\sum_{i=0}^{n-1}\sum_{j=0}^{m-1} J_{i,j} \text{ A/m}. \quad (2)$$

Here, n and m are the number of meshes along the vertical and horizontal axis respectively (in our case $n = 5$ and $m = 5$).

As shown in Fig. 1(b) right, to embed the features of the reconfigurable antenna into a graph data structure, we define node features h as the mean surface current distributions of each copper patch, J_p, whereas the switch states are embedded as edge weights e.

$$h = (\Re(J_{px}) + i\Im(J_{px}), \Re(J_{py}) + i\Im(J_{py})) \text{ A/m}, \quad (3)$$

$$e = \begin{cases} (1,0) & \text{if a horizontal switch is closed.} \\ (0,1) & \text{if a vertical switch is closed.} \\ (0,0) & \text{if a horizontal/vertical switch is opened.} \end{cases} \quad (4)$$

Note that the only possible embeddings for a horizontal switch are $(1,0)$ and $(0,0)$, and the reverse applies for a vertical switch.

Similar to the mesh coloring scheme, the color coding of the node features in Fig. 1(b) right also indicate the magnitude $|J_p|$ of the current vector $J_p = (J_{px}, J_{py})$ for each patch, where $|Jp| = \sqrt{|J_{px}|^2 + |J_{py}|^2}$. The coloring scheme presented in Fig. 1(b) is applied to all subsequent figures.

To predict switch configurations (edge weights) based on given surface current distributions (node features), we begin by initializing all switches in the closed state. Using inverse design with GNN, we then reconstruct the switch configurations that produce the given current distributions. The data required to train such a GNN model can be acquired with full-wave EM simulations. In this work, CST Microwave Studio is utilized to simulate the current distributions and the resulting far-field radiation patterns. Specifically, we randomly generate 3000 switch configurations as ground truth, input them into CST and use the Finite-difference time-domain method (FTDT) solver to perform forward simulation with respect to their resonant frequency in the 2.4 GHz band, and generate the surface current distributions and radiation patterns associated with each set of switch configurations.

B. Physics informed Modeling

The behavior of the antenna circuits is fundamentally governed by Kirchhoff's Current Law (KCL), which states that the total current entering a node must equal the total current leaving it. The current distribution across the antenna network adheres to KCL, which forms the basis for our proposed physics-informed modeling. In the 2.4 GHz band, the incoming currents from the switches and the capacitive coupling between the patches as well as the ground plane affect the steady state current distribution of the antenna. We aim to train a network to predict if the connection between any two patches is dominated by an active switch or merely due to capacitive coupling. If the network identifies dominating active switch over capacitive coupling, it predicts the switch is closed; otherwise, it predicts the switch is opened. Our physics-informed model is inspired by the interaction network architecture, which performs object relation reasoning on graph inputs [9]. Similar to the interaction network, our

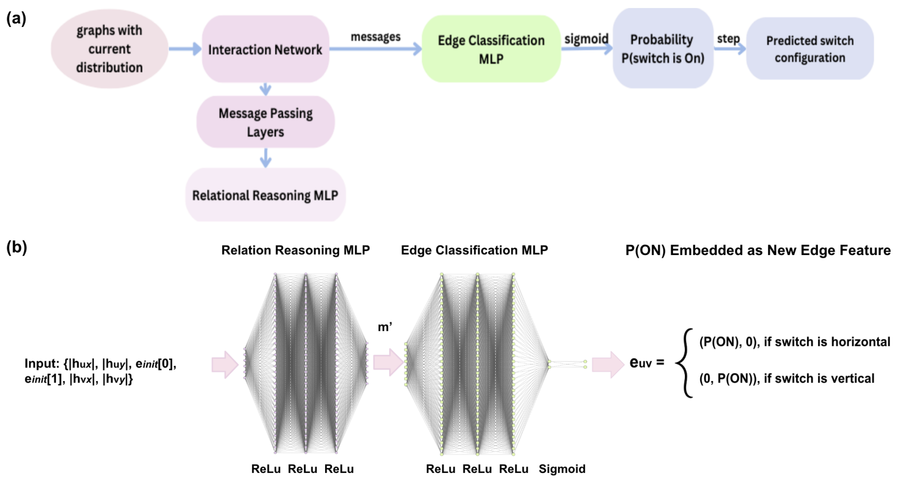

Fig. 2. The proposed network architecture comprises (a) an interaction network and an edge classification multi-layer perceptron (MLP), where the interaction network includes message passing layers followed by a relational reasoning MLP. (b) The relational reasoning MLP has an input dimension of 6. The output of the relational reasoning MLP m' represents a high-dimensional message generated through the message-passing mechanism, which is then used as input for the edge classification MLP. The relational reasoning MLP and edge classification MLP each comprise three fully connected hidden layers with a dimension of 32, utilizing rectified linear unit activation functions for the hidden layers. The output layer of the edge classification MLP employs a sigmoid activation function to indicate the probability of a switch being ON.

proposed network features specialized message-passing layers, implemented in a manner similar to traditional neural network layers in deep learning.

The message passing layers takes an initialized graph as input, $\mathbf{G} = (V, E)$ with node features h_v for each node $v \in V$ and edge features e_{uv} for each $(u, v) \in E$, where $u \in V$ and $v \in V$ are neighboring nodes. The message-passing forward pass is expressed as follows,

$$h_u^{(k+1)} = UPDATE^{(k)}(h_u^{(k)}, m_{\mathcal{N}(u)}^{(k)}), \tag{5}$$

where the subscript k denotes the current number of iterations and $h_u^{(k+1)}$ represents the updated feature of node u, where $u \in V$. The notation $\mathcal{N}(u)$ refers to the one-hop neighborhood of node u, while $m_{\mathcal{N}(u)}^{(k)}$ indicates the message aggregated from node u's neighborhood:

$$m_{\mathcal{N}(u)}^{(k)} = AGGREGATE^{(k)}(h_v^{(k)}, \forall v \in \mathcal{N}(u)). \tag{6}$$

In our proposed model, we aim to emulate KCL. Hence, we use summation as the aggregation function in each message-passing iteration, which ensures that the aggregated messages from a node's neighbors reflect the total incoming and outgoing currents, aligning the message-passing mechanism with the physics governing current flow in the antenna network.

C. Network Architecture

The proposed network architecture is illustrated in Fig. 2 and is built using PyTorch and PyTorch Geometric [10], [11]. The architecture includes an interaction network (IN) with message passing layers and a relational reasoning relational reasoning multi-layer perceptron (MLP), followed by an edge classification MLP. The original IN was developed using basic matrix operations, which were interpreted as series of physical interactions and effects [9]. The primary goal of the proposed architecture is to process graphs where the current distributions are embedded in the nodes and to emulate KCL by the edges.

In the interaction network, 10 message passing layers are implemented to capture information from a 10-hop neighborhood, effectively covering the entire graph. The first layer takes an initialized graph $\mathbf{G} = (V, E)$ as input (example initialized graph is illustrated on the left in Fig. 1(b)) and propagates information between nodes. Following the message passing layers is the relational MLP that learns to infer the message representing the relationship between the nodes. In this context, the message is analogous to the current flowing between two patches of the antenna. Once the messages are computed, the edge classification MLP takes the messages as input to perform binary classification of the switches.

The relational reasoning MLP has an input dimension of 6 and is expressed as $\phi_r(|h_{ux}|, |h_{uy}|, e_{init}[0], e_{init}[1], |h_{vx}|, |h_{vy}|) = m'$. Here, h_u and h_v are the features of the nodes at the ends of an edge, $u \in V$ and $v \in V$, respectively, while e_{init} denotes the initialized edge features. The subscripts x and y, again, correspond to the vector components of the mean surface current density in the x-y plane. The output of the relational reasoning MLP m' represents a high-dimensional message on the edge, which serves as input to the edge classification MLP and can be expressed as $\phi_e(m')$. Both ϕ_r and ϕ_e consist of three fully-connected hidden layers, each with a dimension of 32, connected by rectified linear unit activation function (ReLU) [13], [14]. By activating the final layer of the ϕ_e with a sigmoid function, the network outputs the probability of a switch being in the ON state. This classification outputs the predicted switch configuration of a graph, $W_p(G)$, based on the input current distribution. The forward pass of the proposed network outputs can be expressed as follows,

$$W_p(G) = \begin{cases} 1 & \text{if } P(On) \geq 0.5 \ , \\ 0 & \text{if } P(On) < 0.5 \ , \end{cases} \tag{7}$$

where $P(On) = \phi_e(IN(G))$ is the probability of a switch being closed.

Note that the switch prediction is effectively a binary classification problem, we aim to optimize the binary cross-entropy (BCE) loss between the ground truth switch configurations $W_t \in (0,1)$ and predicted switch configurations W_p. To account for the varying influence of different edges on the final current distribution, the architecture includes an importance mechanism. Edges are deemed trivial if both connected nodes have low current magnitudes, defined as less than 20% of the maximum current. During training, these trivial edges are assigned reduced importance by lowering the penalty for incorrect predictions. That is, a weighted loss function is employed, where trivial edges weighting 0.1 of loss compared to non-trivial edges. The final loss function $\mathcal{L}$ subject to optimization is expressed as follows,

$$\begin{aligned} \mathcal{L}(W_t, W_p) = -\frac{p_i}{N} \sum_{i=1}^{N} & (W_{t,i} \log(W_{p,i}) \\ & + (1 - W_{t,i}) \log(1 - W_{p,i})) \end{aligned} \tag{8}$$

where N represents the total number of switches evaluated, i is an index that refers to each switch in the graph, and p_i indicates the penalty weighting of the switch $p_i \in (0.1, 1)$. This weighted loss function prevents overfitting and allows the model to focus on the underlying physics of the more critical edges. To facilitate training, the simulated data for the 3000 total switch configurations are used in 9-fold training and cross validation, where the Adam optimizer [15] is employed to optimize the loss function, with 500 epochs, a batch size of 64, and a learning rate of 0.005 with early stopping.

III. Experimental Results

To validate the model, we report the test accuracy, as well as true positive, true negative, false positive, and false negative rates for non-trivial switches (see Tab. I for details). The test accuracy is 93.2%. To further demonstrate the model's ability to replicate complex radiation patterns, we conducted forward simulations using the predicted switch configurations to generate surface current distributions and radiation patterns. Fig. 3(a) outlines the steps taken for simulation-based validation. Our proposed GNN model first takes a graph with the current distribution embedded as node features as input, where all switches are initially closed. The GNN then predicts the switch configuration based on the given current distribution. This predicted switch configuration is input into the full-wave EM simulator, which outputs the corresponding current distribution and radiation pattern. We compare these results with those obtained from the ground truth switch configurations. We compare the current distributions and radiation patterns from the predicted switch configurations with that of the ground truth switch configurations.

TABLE I
Switch prediction performance metrics.

Metric	Accuracy	True Positive	True Negative	False Positive	False Negative
Value	93.2%	87.97%	98.02%	1.98%	12.03%

Fig. 3(b) presents two examples with complex radiation patterns, comparing the simulated current distribution and far-field radiation patterns for both the ground truth and predicted switch configurations. The left two columns and right two columns correspond to example switch configurations 1 and 2, respectively, along with their resulting current distributions and radiation patterns. The top row illustrates the ground truth switch configurations, while the bottom row shows the GNN-predicted switch configurations and their results. In the first example, the ground truth and predicted switch configurations produce nearly identical current distributions, with only two incorrect switch predictions occurring in the third column from the left. These switches were deemed trivial and have minimal impact on the current distribution (error vector magnitude, EVM, of 13.62 A/m), enabling the predicted switch configurations to accurately replicate the three-lobe radiation pattern of the ground truth with a mean-square error (MSE) of 0.001. Similarly, in the second example, a few incorrect switch predictions in the fifth and sixth columns had little impact on the current distribution (EVM of 10.74 A/m), enabling the ground truth radiation pattern to be recreated with a MSE of 0.003. We randomly sampled 10 configurations from the validation test, performed forward simulations, and calculated an average current distribution EVM of 19.85, signal to noise ratio of 35.32, and radiation pattern MSE of 0.0064. In summary, the simulation validates the model's ability to recreate complex radiation patterns from patch current distributions, demonstrating its effectiveness for reconfigurable antenna inverse design through simulations.

IV. Conclusion

In this work, we developed, trained, and validated a physics-informed GNN-based model for reconfigurable antenna inverse design through simulations. By integrating KCL and antenna topology into the network architecture, the model accurately predicted switch configurations capable of recreating surface current distributions and far-field radiation patterns. This demonstrates the effectiveness of leveraging physics-informed modeling for antenna inverse design. While the architecture has been optimized for the current antenna dimensions, future work will explore its scalability across varying dimensions and include real-world experimental validation.

References

[1] R. F. Harrington, Field Computation by Moment Methods, Wiley-IEEE Press, 1993.

[2] M. Steer. Rijeka, Microwave and RF Design: A Systems Approach. Croatia: SciTech, 2013.

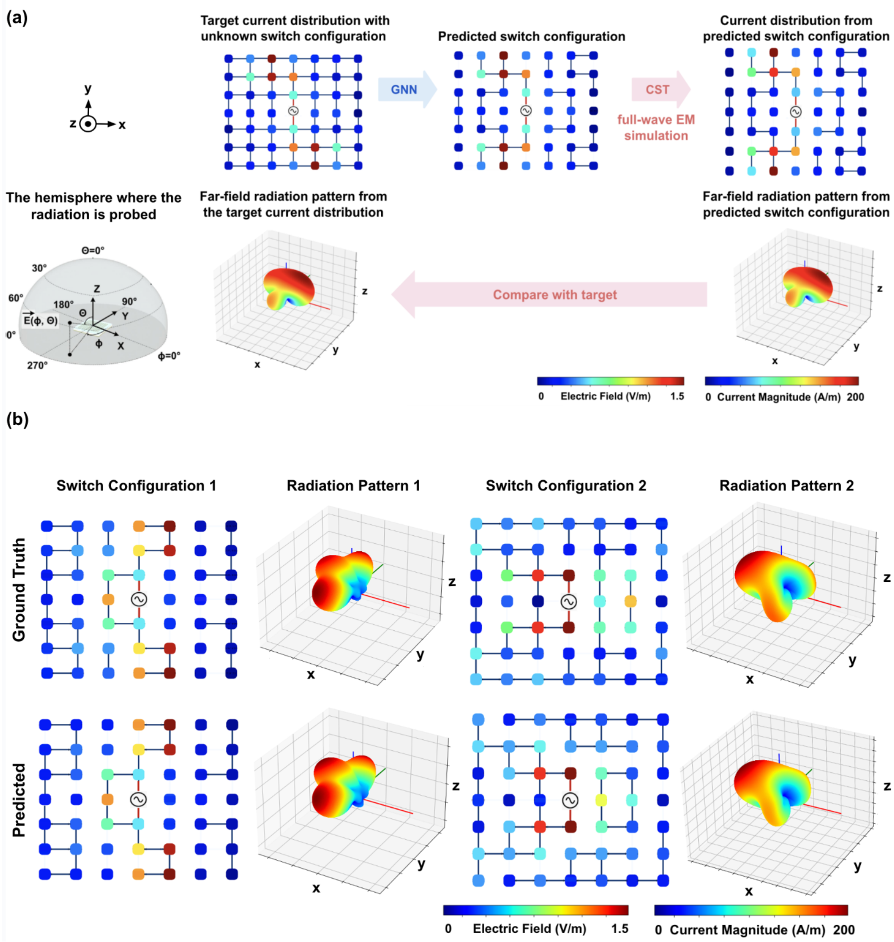

Fig. 3. (a) Left: Coordinate axes definition: The antenna is positioned on the xy-plane at z = 0. The radiation pattern is measured over a hemisphere in the far-field region. Right: Flowchart of simulated experimental validation. (b) Example comparisons of the full-wave EM simulated current distribution and far-field radiation patterns of the reconfigurable antenna for both the ground truth and predicted switch configurations. The left two columns and the right two columns show example switch configuration 1 and example switch configuration 2 respectively along with their results. The top row displays the ground truth switch configurations along with the corresponding current distributions and radiation patterns, while the bottom row presents the GNN-predicted switch configurations and their resulting current distributions and radiation patterns.

[3] Z. Jiao and Y. Zhao, "Design of frequency and beam reconfigurable antenna based on encoded reflectors for Wi-Fi and IoT applications," AIP Advances, vol. 14, no. 6, 2024, p. 065127. https://doi.org/10.1063/5.0213545.

[4] C. Wu, et al., "Gigahertz Large-Area-Electronics RF Switch and its Application to Reconfigurable Antennas," 2020 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 2020, pp. 33.6.1-33.6.4, doi: 10.1109/IEDM13553.2020.9372057.

[5] C. Balanis, Antenna Theory. Wiley, 1997.

[6] S. R. Saunders and A. Aragón-Zavala, Antennas and Propagation for Wireless Communication Systems. Hoboken, NJ, USA: Wiley, 2007.

[7] H. Rezaei, J. Meiguni, M. Soerensen, J. Fan, and D. Pommerenke, "Source reconstruction in near field scanning using inverse MoM for RFI application," in Proceedings of the 2019 IEEE International Symposium on Electromagnetic Compatibility, Signal Power Integrity (EMC+SIPI), New Orleans, LA, USA, 2019, pp. 584-589, doi: 10.1109/ISEMC.2019.8825241.

[8] S. H. Raad, J. S. Meiguni, and R. Mittra, "Inverse MoM approach to near-field prediction and RFI estimation in electronic devices with multiple radiating elements," IEEE Access, vol. 11, pp. 21313-21325, 2023, doi: 10.1109/ACCESS.2023.3251221.

[9] P. W. Battaglia, et al., "Interaction networks for learning about objects, relations and physics," in Advances in Neural Information Processing Systems, vol. 29, D. Lee, et al., eds., Curran Associates, Inc., 2016, pp. 4502, arXiv:1612.00222.

[10] M. Fey and J. E. Lenssen, "Fast graph representation learning with PyTorch Geometric," in Representation Learning on Graphs and Manifolds Workshop at the 7th International Conference on Learning Representations, 2019, arXiv:1903.02428.

[11] A. Paszke, et al., "Pytorch: An imperative style, high-performance deep learning library," in Advances in Neural Information Processing Systems, vol. 32, H. Wallach, et al., eds., Curran Associates, Inc., 2019, arXiv:1912.01703.

[12] G. DeZoort, S. Thais, J. Duarte, et al., "Charged Particle Tracking via Edge-Classifying Interaction Networks," Comput. Softw. Big Sci., vol. 5, no. 26, 2021. https://doi.org/10.1007/s41781-021-00073-z.

[13] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in Proceedings of the 27th International Conference on Machine Learning, ICML, Madison, WI, USA, 2010, p. 807.

[14] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in Proceedings of the 14th International Conference on Artificial Intelligence and Statistics, G. Gordon, D. Dunson, and M. Dudík, eds., vol. 15, JMLR, Fort Lauderdale, FL, USA, 2011, p. 315.

[15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proceedings of the 3rd International Conference on Learning Representations, Y. Bengio and Y. LeCun, eds., 2015, arXiv:1412.6980.

[16] M. Mashayekhi, P. Kabiri, A. S. Nooramin, et al., "A reconfigurable graphene patch antenna inverse design at terahertz frequencies," Sci. Rep., vol. 13, no. 8369, 2023. https://doi.org/10.1038/s41598-023-35036-4.