Physics-guided hierarchical neural networks for Maxwell's equations in plasmonic metamaterials

Sean Lynch,[†] Jacob LaMountain,[*,‡] Bo Fan,[‡] Jie Bu,[¶] Amogh Raju,[§] Daniel Wasserman,[§] Anuj Karpatne,[¶] and Viktor A. Podolskiy[‡]

[†]Miner School of Computer Science
University of Massachusetts Lowell
Lowell, MA 01854, USA
[‡]Department of Physics and Applied Physics
University of Massachusetts Lowell
Lowell, MA 01854, USA
[¶]Department of Computer Science
Virginia Tech
Blacksburg, VA 24061, USA
[§]Department of Electrical and Computer Engineering
University of Texas Austin
Austin, TX 78712, USA

E-mail: jacob_lamountain@student.uml.edu

Abstract

While machine learning (ML) has found multiple applications in photonics, traditional "black box" ML models typically require prohibitively large training data sets. Generation of such data, as well as the training processes themselves, consume significant resources, often limiting practical applications of ML. Here we demonstrate that embedding Maxwell's equations into ML design and training significantly reduces the required amount of data and improves the physics-consistency and generalizability of ML models, opening the road to practical ML tools that do not need extremely large training sets. The proposed physics-guided machine learning (PGML) approach is illustrated on the example of predicting complex field distributions within hyperbolic metamaterial photonic funnels, based on multilayered plasmonic-dielectric composites. The hierarchical network design used in this study enables knowledge transfer and points to the emergence of effective medium theories within neural networks.

Keywords

Machine Learning, Metamaterials, Plasmonics, Physics Guided Machine Learning

1 Introduction

Composite materials with engineered optical properties, metamaterials and metasurfaces, are rapidly advancing as platforms for optical communications, sensing, imaging, and computing.[1–7] The complexity of typical metamaterials makes it almost impossible to understand and optimize their interaction with light based on experimental or analytical theory approaches alone, leaving the problem of light interaction with metamaterials to computational sciences.[1,4,8,9] Currently, finite-difference time domain (FDTD)[10] and finite element methods (FEM)[11,12] represent industry-standard approaches to understanding the optics of nonperiodic composite media.

Machine learning (ML) techniques, particularly neural networks (NNs), have recently been incorporated into the design, evaluation, and measurement of nanophotonic structures.[13–21] Properly trained ML tools can be used as surrogate models that predict the spectral response of composites or—rarely—field distributions within metamaterials.[14,22–24] Since ML does not solve the underlying electromagnetic problem, these predictions are significantly faster than brute-force simulations. However, extensive training sets, often featuring $\sim 10^3 \ldots 10^5$ configurations are required in order to develop high-quality ML models.[14,25–28] The time and computational resources needed to generate these data sets, as well as the time and resources needed for the ML training process, are significant and often serve as the main limitation to ML use in computational photonics.

Embedding physics-based constraints (physics-consistency)[29,30] into the ML training process may be beneficial for the resulting models. ML methods for general solutions of partial differential equations (PDEs) are being developed.[31–38] However, as of now, these techniques are illustrated on convenient "toy" models and cannot be straightforwardly applied to practical electromagnetic problems. Physics-guided machine learning (PGML)[6,7,39] is emerging as a promising platform that can combine data- and physics-driven learning. Notably, previous PGML attempts have been focused on dielectric[39] or relatively simple plasmonic[7] composites. Here we develop PGML models that are capable of predicting electromagnetic

fields within plasmonic metamaterials. We illustrate our technique by analyzing the optical response of metamaterials-based photonic funnels:[40–42] conical structures with strongly anisotropic composite cores that are capable of concentrating light to deep subwavelength areas. We show that physics-based constraints enable training on unlabeled data and significantly improve the accuracy and generalizability of the models. We also attempt to understand the inner workings of the NNs by analyzing the performance of hierarchical models with different data resolutions.
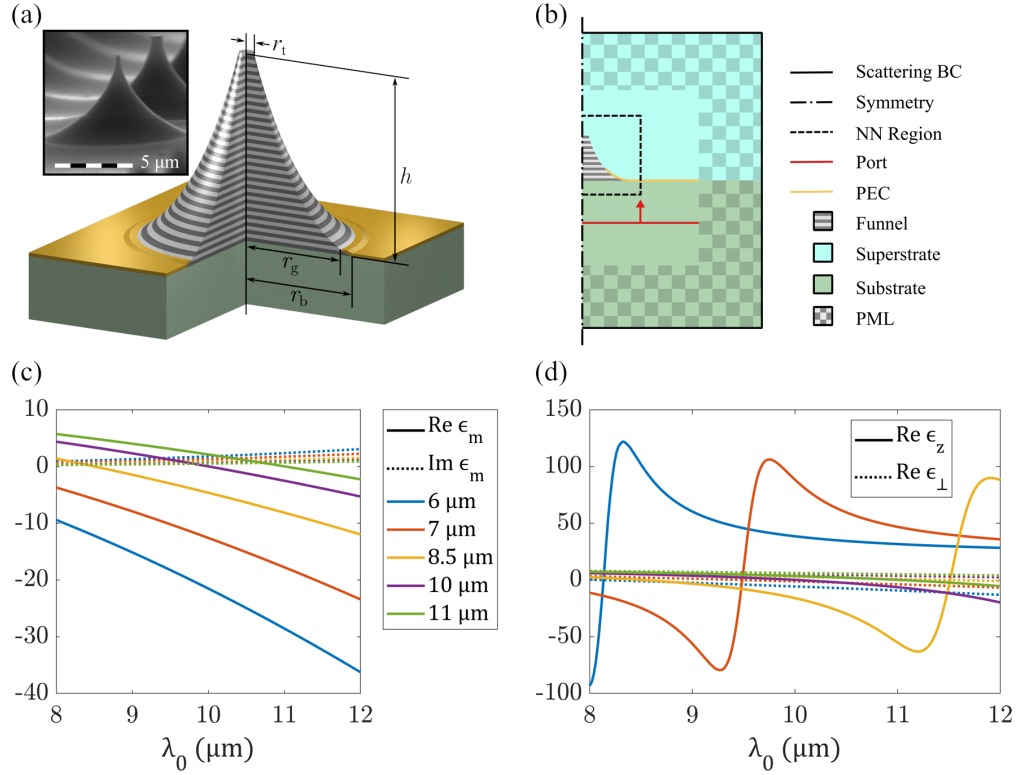

Figure 1: (a) schematic of the photonic funnel with cut-out demonstrating the composite structure of the core; inset shows scanning electron microscope (SEM) image of as-fabricated array of funnels (b) simulation setup used in FEM-based solutions of Maxwell's equations; NNs are trained only on a subregion of the FEM data which contains the funnel; (c,d) wavelength dependence of (c) the permittivity of the highly doped plasmonic components of the funnels and (d) the components of the resulting effective permittivity tensor.

4

2 Hyperbolic Metamaterial-based Photonic Funnels

An electromagnetic composite comprising sufficiently thin alternating layers of non-magnetic materials with permittivities ϵ_1, ϵ_2 and thicknesses d_1, d_2 (see Fig.1) behaves as a uniaxial medium whose optical axis is perpendicular to the layers (direction $\hat{z}$ in this work) and whose diagonal permittivity tensor has components given by: $\epsilon_{xx} = \epsilon_{yy} = \epsilon_\perp = \frac{d_1\epsilon_1 + d_2\epsilon_2}{d_1 + d_2}$ and $\epsilon_{zz} = \frac{(d_1+d_2)\epsilon_1\epsilon_2}{d_1\epsilon_2 + d_2\epsilon_1}$. Such a material supports the propagation of two types of plane waves that differ in their polarization and have fundamentally different dispersions.

The ordinary waves (which have $\vec{E} \perp \hat{z}$) satisfy the dispersion relation $k^2 = \epsilon_\perp \omega^2/c^2$ with $\vec{k}, \omega$, and c representing the wavevector of the wave, operating angular frequency, and speed of light in vacuum, respectively. This dispersion is identical to that of plane waves propagating in a homogeneous isotropic material with permittivity $\epsilon_\perp$. On the other hand, the extraordinary, or transverse-magnetic (TM), waves (with $\vec{H} \perp \hat{z}$) have dispersion $\frac{k_x^2 + k_y^2}{\epsilon_{zz}} + \frac{k_z^2}{\epsilon_\perp} = \frac{\omega^2}{c^2}$. Notably, for anisotropic materials, the dispersion of extraordinary waves is either elliptical or hyperbolic. The topology of the iso-frequency contours strongly depends on the combination of signs of the effective permittivity tensor.

When the components of the permittivity tensor are of opposite signs, the iso-frequency surfaces are hyperboloids. This hyperbolic dispersion has been identified as the enabling mechanism for such unique optical phenomena as negative refraction, strong enhancement of light matter interaction, and for light manipulation in ultrasmall (deep subwavelength) areas.[2,43–46] Hyperbolicity can be achieved by alternating layers of dielectric ($\epsilon_1 = \epsilon_d > 0$) and plasmonic ($\epsilon_2 = \epsilon_m < 0$) layers. In the semiclassical regime (typically, when the layer thickness $\gtrsim 10\,\text{nm}$) the permittivity of the plasmonic layers as a function of angular frequency of light is well described by the Drude model,[47]

$$\epsilon_m(\omega) = \epsilon_\infty \left(1 - \frac{\omega_p^2}{\omega^2 + i\gamma\omega} \right) \tag{1}$$

with $\epsilon_\infty, \omega_p$, and γ being background permittivity, plasma frequency, and scattering rate,

respectively. Here we use $\epsilon_\infty = 12.15, \gamma = 10^{13}s^{-1}$, and parameterize the plasma frequency using the plasma wavelength, λ_p, via $\omega_p = 2\pi c/\lambda_p$.

The most common implementation of these metamaterials leverages a 50/50 composition ($d_1 = d_2 = d$). For such systems topological transitions occur when the permittivity of the plasmonic layers (ϵ_m) and the weighted permittivity of the mixture ($\epsilon_\perp$) change signs.[48,49] The dispersion of TM waves inside the metamaterial is elliptic for shorter wavelengths $\lambda < \lambda_p$. It changes to type-I hyperbolicity ($\epsilon_\perp > 0, \epsilon_{zz} < 0$) for $\lambda_p < \lambda < \tilde{\lambda}_p$, with the renormalized plasma frequency, $\tilde{\lambda}_p$, defined as $\text{Re}(\epsilon_\perp(\tilde{\lambda}_p)) = 0$. Finally, the dispersion of TM waves in the composite becomes type-II hyperbolic ($\epsilon_\perp < 0, \epsilon_{zz} > 0$) for $\tilde{\lambda}_p < \lambda$.

Photonic funnels, conical waveguides with hyperbolic metamaterial cores,[40–42] shown in Fig.1, represent excellent examples of structures capable of manipulating light at a deep subwavelength scale. Recent experimental results[40,42] demonstrate efficient concentration of mid-infrared light with a vacuum wavelength of ~10 µm to spatial areas as small as ~300 nm, $1/30^{\text{th}}$ of the operating wavelength, within an all-semiconductor "designer metal" material platform.[49] Further analysis relates the field concentration near the funnels' tips to the absence of the diffraction limit within the hyperbolic material and to the anomalous internal reflection of light from the funnel sidewall, which forms an interface oblique to the optical axis.[42] Importantly, the optical response of realistic funnels can be engineered at the time of fabrication by controlling the doping of the "designer metal" layers and thereby adjusting the plasma frequency of these layers. The unusual electromagnetic response, strong field confinement, and significant field inhomogeneities make photonic funnels an ideal platform for testing the performance of ML-driven surrogate solvers of Maxwell's equations.

3 Methods

3.1 Data set description and generation

To construct a sufficiently diverse set of labeled configurations, we used FEM to solve for electromagnetic field distributions in photonic funnels with plasmonic layers of different doping concentrations corresponding to plasma wavelengths[50] of $6\,\mu m$, $7\,\mu m$, $8.5\,\mu m$, $10\,\mu m$ and $11\,\mu m$. Fig.1 illustrates the wavelength-dependent permittivity of plasmonic layers with various doping concentrations as well as the corresponding effective medium response of the layered metamaterials. Note the drastic changes of effective medium response as a function of both wavelength and doping.

For each doping level, wavelength-dependent permittivity and electromagnetic field distributions have been calculated with a commercial FEM-based solver[12] (that takes into account that all fields are proportional to $\exp(-i\phi)$ with ϕ being the angular coordinate of the cylindrical reference frame) for free-space wavelengths from $8\,\mu m$ to $12\,\mu m$ with increments of $62.5\,nm$. The FEM model setup is shown schematically in Fig.1a. Electromagnetic waves that are normally incident on the funnel base are generated by the port boundary condition. Perfectly matched layers[11] and scattering boundary conditions are used to make the outside boundaries of the simulation region completely transparent to electromagnetic waves, thereby mimicking the surrounding infinite space. The model, which explicitly incorporates $80\,nm$-thick layers in the funnel cores, is meshed with a resolution of at most $40\,nm$ inside the funnel and $200\,nm$ outside the funnels, with the mesh growth factor set to 1.1 to avoid artifacts related to abrupt changes in mesh size.

For every plasma wavelength and operating frequency the distribution of electromagnetic fields, along with the distributions of permittivities within a small ($5 \times 12\,\mu m$) region of space containing the funnel (see Fig.1) is interpolated onto a rectangular mesh with resolution $12.5\,nm \times 10\,nm$ along the r and z directions, respectively, forming the basis for the data sets used in the study. Note that selecting this internal region of space from the FEM simulations

allows us to (i) implicitly incorporate the proper boundary conditions for both incident as well as scattered electromagnetic fields and (ii) avoid the implementation of perfectly matched layers, ports, and scattering conditions within the physics-based constraints used in training our NNs.

The original FEM-generated data has been then resampled into three separate data sets:

- low-resolution data set, 20×60 pixels with resolution $250 \, \text{nm}$ and $200 \, \text{nm}$ in r and z directions, respectively

- medium-resolution data set, 100×300 pixels with resolution $50 \, \text{nm} \times 40 \, \text{nm}$

- high-resolution data set, 200×600 pixels with resolution of $25 \, \text{nm} \times 20 \, \text{nm}$

4 Neural Network Architecture

On a fundamental level, approximating solutions of Maxwell's equations within metamaterials with ML necessitates a neural network to map the operating frequency and the distribution of permittivity across the composite to the distribution of electromagnetic fields, a problem that is similar to image transformation. Previous analysis[51,52] has demonstrated that convolutional neural networks (CNNs) excel in image transformation. Specifically, encoder-decoder, CNN, and U-net architectures have shown success in electromagnetic problems,[19,21,53–55] presumably due to the cores of the networks learning some low-dimensional representation of the solutions.[34,56] Note, however, that the vast majority of previous ML-driven solvers of Maxwell's equations[21,39,57,58] have analyzed dielectric composites (where electromagnetic fields are relatively smooth) and were trained on relatively large data sets.[14,20,39]

We follow the general approach of constructing U-nets. The design of our networks is summarized in Fig.2. Starting with the pixel resolution of the data set, the proposed CNNs reduce the dimensionality of the problem to 20×10 pixels, and then expand the resulting distributions to their original size.

The linear parts of the network employ standard convolutional and transposed convolutional layers with stride $= 1$ for those parts of the network that preserve pixel size and with stride > 1 for those that perform encoding/downsampling and decoding/upsampling. Hyperbolic tangent activation layers are used to add nonlinearities to the CNN. Combinations of convolutional and tanh layers are marked as thick arrows in Fig.2. In addition, custom layers are introduced to implement skip connections which can propagate the vacuum wavelength and permittivity distributions into the depth of the network for both stability of the resulting NN and to enable evaluation of the physics-consistency of the resulting predictions. These layers operate by directly appending several layers of pixels to the output of a given convolutional layer (thin black arrows in Fig.2) or by first downsampling to the core resolution and then concatenating (thin orange arrows in Fig.2).
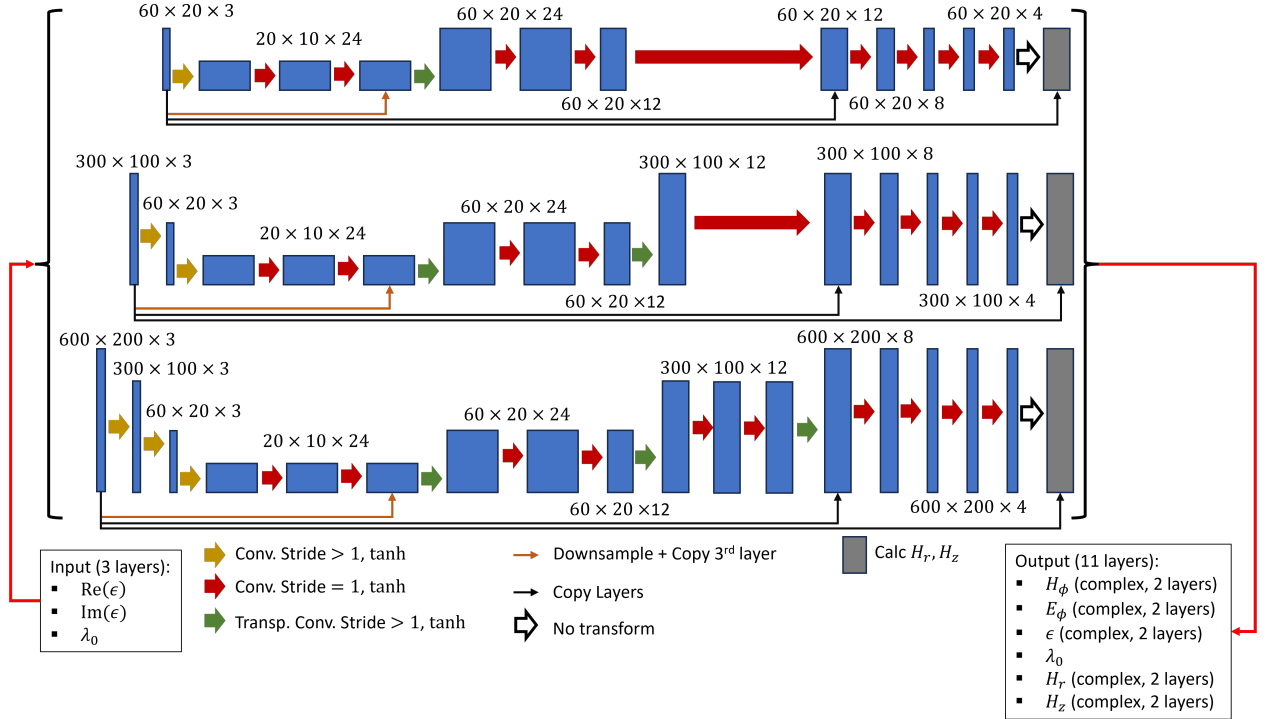

Figure 2: Setup of the CNN used in the study; the three rows represent low-, medium-, and high-resolution networks; boxes represent the size of data as it propagates through the network; arrows represent CNN data operations: each thick solid arrow represents the combination of a (transposed) convolutional layer and a tanh activation layer; thin black and orange arrows represent skip connections; thin red arrows represent input and output.

9

The base part of the NN (blue layers in Fig.2) is designed to learn the distribution of the ϕ components of the electric and magnetic fields. Note that in our hierarchical setup, the core of the networks remains the same, independent of the resolution of the data set, with the outer structure producing encoding/decoding from/to the higher resolution. The inner structure of the network (layer dimensionality and filter size) was optimized using the low-resolution data set. The medium- and high-resolution networks build upon this geometry by adding "hierarchical" downsampling and upsampling layers, implemented via convolutional and transposed convolutional layers in our NNs. Our analysis suggests that it is important to initialize the downsampling layers with unit weights, thereby setting the network for brute-force averaging of permittivity during the initial training iterations.

The physics-agnostic portion of the CNN, which is trained to produce the ϕ components of the magnetic and electric fields, is followed by a physics-informed layer (gray layers in Fig.2) which calculates distributions of the r and z components of the magnetic field based on analytical expressions derived from Maxwell's equations:

$$\vec{E}_{rz} = \frac{-i}{\epsilon\frac{\omega^2}{c^2} - \frac{1}{r^2}} \left(-\frac{1}{r}\vec{D}_{rz}E_\phi - \frac{\omega}{c}\hat{\phi} \times \vec{D}_{rz}H_\phi \right) \tag{2}$$

$$\vec{H}_{rz} = \frac{-i}{\epsilon\frac{\omega^2}{c^2} - \frac{1}{r^2}} \left(-\frac{1}{r}\vec{D}_{rz}H_\phi + \epsilon\frac{\omega}{c}\hat{\phi} \times \vec{D}_{rz}E_\phi \right) \tag{3}$$

where we have introduced the vector differential operator $\vec{D}_{rz}f = \hat{r}\frac{1}{r}\frac{\partial}{\partial r}(rf) + \hat{z}\frac{\partial f}{\partial z}$.

Because the fields are discretized on a regular rectangular grid, all derivatives are approximated with finite difference schemes. Forward and backward differences are used at the edges of the computational domain, while central differences are used within it. Our implementation of the CNNs used in this work and the data sets used in training are available on GitHub and Figshare, respectively.[59,60]

As seen from Eqs.(3), predictions for H_r and H_z may diverge when $\epsilon r^2\omega^2/c^2 \approx 1$. This instability is a direct consequence of applying differential operators in cylindrical geometry. Here, we address the related issues by introducing a regularizing function (see below and SI).

Our approach, illustrated here on an example of cylindrical geometry may be generalized to other curvilinear coordinates.

4.1 Knowledge transfer between different NN

As mentioned above, in the limit of ultrathin layers, the optics of multilayer metamaterials can be adequately described by the effective medium theory. In a related but separate scope, the U-shaped NNs are hypothesized to learn low-dimensional representations of the underlying phenomena. These considerations motivate the hierarchical design of the NNs used in this work.

To explore whether the learning outcomes of the NNs are consistent with the effective medium description, we performed a series of experiments where pretrained lower-resolution networks were used as pretrained cores of higher-resolution *transfer-learning* (TL) networks. In these studies, the learning parameters of the pretrained "core" layers were frozen, with only the averaging and transposed convolution peripheral layers of the higher-resolution NN being trained.

At the implementation level, we drew inspiration from the ResNet[61] architecture's approach of organizing layers into "residual blocks." Specifically, we grouped the frozen layers into a single block, with the internal layer weights corresponding to those of the selected pretrained network. The forward function was designed to perform training within the layers of the block; however, during backpropagation, the weight updates bypass the internal layers of the block, passing directly to the previous layer.

We explored knowledge transfer from low- to medium-resolution networks as well as from medium- to high-resolution networks.

4.2 Training protocols

To assess the benefits of the physics-based constraints, three different regimes of training the CNN are explored. In the base-case *black-box* (BB) scenario, the model min-

imizes only the radially-weighted mean-squared error of the ϕ components of the electric and magnetic fields (directly produced by the physics-agnostic part of the network) $L_\phi = \left\langle w(r) \left[\left| H_\phi^Y - H_\phi^T \right|^2 + \left| E_\phi^Y - E_\phi^T \right|^2 \right] \right\rangle$. Here, the superscripts Y and T correspond to the predicted and ground-truth fields, respectively, the angled brackets, $\langle \cdots \rangle$, represent an arithmetic mean over the simulation region, and the radial weight function, $w(r)$, is used to emphasize the region of small radii where the funnel is located.

The second, *field-enhanced* (FE) model utilizes a hybrid loss that combines the above-described L_ϕ with its analog for the remaining components of the magnetic field,

$$L_{FE} = L_\phi + L_{rz} \tag{4}$$

with $L_{rz} = \left\langle w(r) \left| R \right|^2 \left[\left| H_r^Y - H_r^T \right|^2 + \left| H_z^Y - H_z^T \right|^2 \right] \right\rangle$ and the rz components of the magnetic field being produced by the physics layer of the CNN.

In order to prevent the instability of Eq.(3) from dominating the overall loss, we introduce the regularization function, $R(r, z)$, such that $R(r, z) \to 0$ when $r \to c/(\sqrt{\epsilon(r, z)}\omega)$ (see the supplementary information for details).

Because calculation of the r and z field components requires differentiating the ϕ components, the addition of L_{rz} allows the CNN to learn the relationships between the spatial field distributions and the distributions of their derivatives. Importantly, evaluation of both L_ϕ and L_{rz} terms requires the training set to contain the solutions of Maxwell's equations (labeled data).

Finally, *physics-guided* (PG) training combines the above labeled-data–dependent terms, L_ϕ and L_{rz}, with the physics loss,

$$L_{ph} = \frac{1}{\max \left| H_\phi^Y \right|} \left\langle \left| \frac{\partial (H_z^Y R^2)}{\partial r} - \frac{\partial (H_r^Y R^2)}{\partial z} + i \frac{\omega}{c} \epsilon \, E_\phi^Y R^2 - 2 \left(H_z^Y R \frac{\partial R}{\partial r} - H_r^Y R \frac{\partial R}{\partial z} \right) \right| \right\rangle \tag{5}$$

which represents the (regularized) residual of Maxwell's equations for the H_ϕ component of the field (see the supporting information). Therefore, PG training aims to enforce consis-

tency of the solutions that are generated by the NN with Maxwell's equations. Notably, evaluation of the physics loss does not require labeled data. As a result, *unlabeled-trained* (UL) networks can utilize a combination of labeled and unlabeled data, with the former inherently incorporating the boundary conditions, and the latter allowing the expansion of the training set without computing additional PDE solutions. This UL loss was also used in training the TL networks described in the preceding section

Previous analysis[6] demonstrated that BB- and PG-loss often compete with each other. Here, this competition reflects the different differentiation schemes used by FEM and the PG-loss as well as the existence of multiple solutions to Maxwell's equations (for example, the trivial solution $\vec{E} = \vec{H} = 0$) that do not necessarily satisfy the boundary conditions that are implicitly enforced by labeled data. To guide the network towards the correct implementation of boundary conditions, the weight of the physics-loss, w_{ph}, is dynamically adjusted during training,[6] resulting in the dynamic PG loss,

$$L_{PG} = L_\phi + L_{rz} + w_{ph}L_{ph}. \tag{6}$$

In order to assess the ability of the networks to interpolate and extrapolate between data sets having plasmonic layers with different plasma wavelengths, we train the networks on 50% of the data with plasma wavelengths of 6 and 11 µm *or* with plasma wavelengths of 7 and 10 µm, and add up to 10% of the labeled data from other data sets to the training. The UL models are also provided configurations from the remaining data sets as unlabeled data. The training scenarios are summarized in Table 1, which gives the percent of each data set that was used as labeled and unlabeled data in each network type. Each training scenario has been used to train at least 10 different networks of each resolution and loss type, with the dynamics of their training and validation loss presented in the supplementary information, and their averaged performance summarized below.

Table 1: Labeled and Unlabeled Training Data Composition of Each Network

Network	Labeled % λ_p (µm)					Unlabeled % λ_p (µm)				
	6	7	8.5	10	11	6	7	8.5	10	11
BB_i, FE_i, PG_i	50	0	10	0	50			none		
UL_i, TL_i	50	0	10	0	50	0	0	40	0	0
BB_e, FE_e, PG_e	0	50	10	50	0			none		
UL_e	0	50	10	50	0	25	0	40	0	25
BB_x, FE_x, PG_x	10	50	10	50	10			none		
UL_x	10	50	10	50	10	25	0	40	0	25

5   Results

To demonstrate the impact of physics-based constraints on the accuracy and consistency of NN-predicted fields, we analyze the dependence of the three average losses introduced above (L_ϕ, L_{rz}, and L_{ph}) both on the enforcement of physics-consistency and on the presence of unlabeled data during training. Sample field distributions are presented to illustrate the models' performance. Finally, we analyze the generalizability of the models by evaluating their performance across the plasma wavelengths of the plasmonic layers.

The three components of the loss, L_ϕ, L_{rz} and L_{ph}, are arranged in increasing degree of physics consistency and – simultaneously – decreasing reliance on data. Indeed, L_ϕ, which analyzes only the physics-agnostic output of the networks, relies exclusively on data. L_{rz}, which primarily relies on the output of the physics layer enforces the relationships between the fields at neighboring points [see Eq.(3)]. Lastly, L_{ph} exclusively analyzes physics-consistency and pays no regard to data consistency. Our analysis (see below) illustrates that training with L_{ph} not only improves the consistency with Eq.(5) but also improves other metrics that are related to Maxwell's equations, such as energy conservation – as analyzed through the Poynting theorem (see SI).
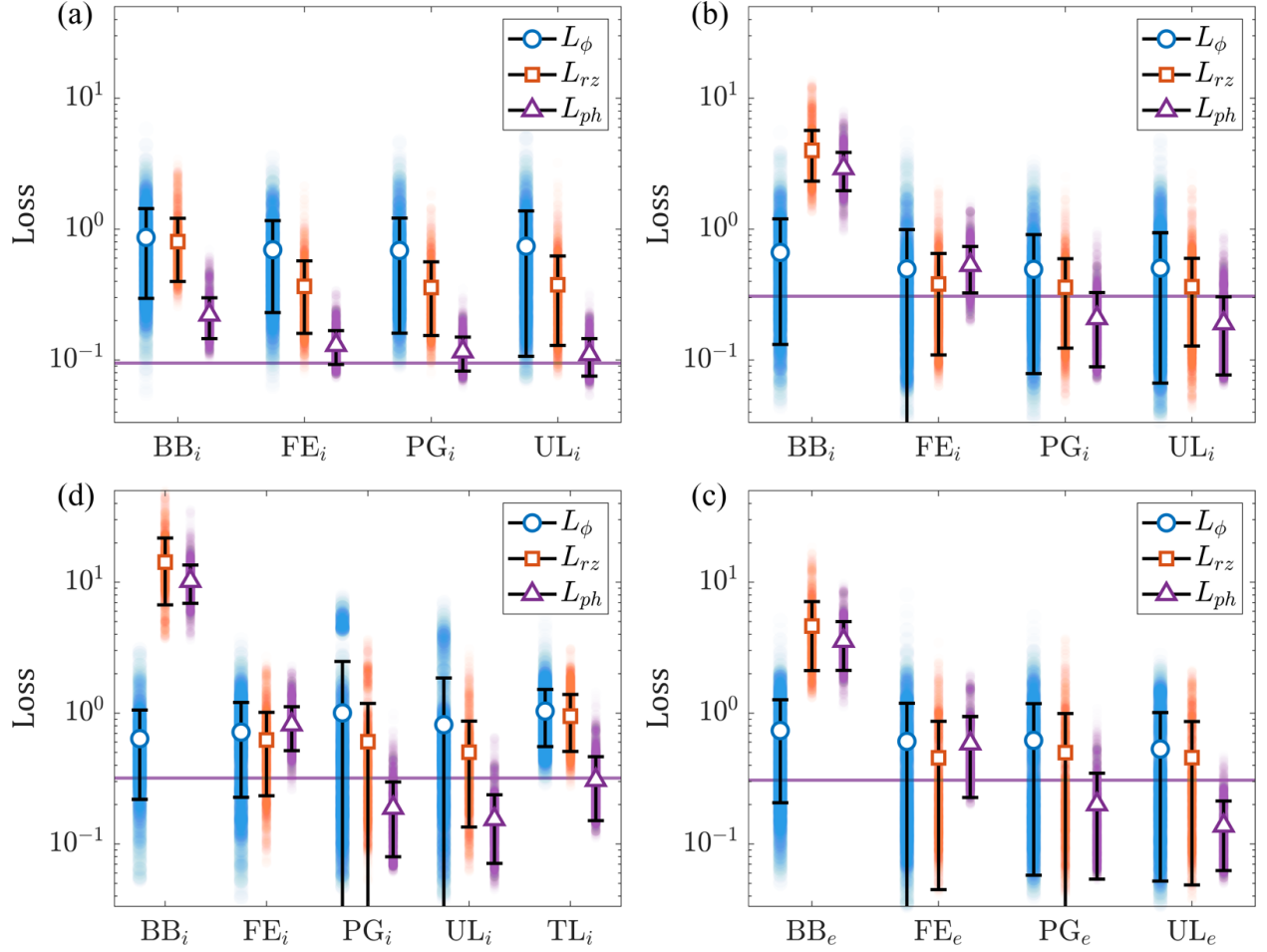
Figure 3: Performance of NNs with different architectures and training protocols, evaluated on the data that was not used in training; panels (a,b,c,d) represent low-resolution (a), medium resolution (b,c), and high-resolution (d) networks (see Table 1 for network labels); loss metrics of individual predictions are represented as filled semitransparent circles, resulting in the color-coded distributions; solid white markers and black bars represent the mean and standard deviation of these distributions; the purple horizontal lines show the average L_{ph} of all interpolated FEM solutions.

5.1   Impact of Physics Information on Accuracy

The performance of the different models is summarized in Fig. 3. With the comparatively simple low-resolution model, adding the physics-based layer to the network and adding the L_{rz} component to the loss function provides enough additional information to adequately represent the coarsely sampled data. Providing the network additional physics-based information (by implementing PG loss) does not quantitatively boost the performance of the

model – due to a combination of the model's simplicity and the mesh being too coarse to resolve the composite structure.

As the resolution and complexity of the model grow, increasing physics-based constraints and adding unlabeled data yields measurable improvements in model performance. Interestingly, the extra physics consistency (as demonstrated by the improving L_{ph} metric) sometimes comes at the cost of a small increase of L_ϕ. This apparent contradiction results from the fact that the data used in training was generated by reinterpolating FEM solutions from a triangular mesh to a rectangular mesh. As a result, the "ground truth" does not yield vanishing L_{ph}. As seen in Fig. 3, predictions of the neural net tend to be closer solutions to Maxwell's equations on the rectangular mesh than the FEM-sourced data.

A more granular look at the NN predictions is shown in Fig. 4 where representative examples of model predictions are compared with FEM solutions. Note that in contrast to their BB counterparts, PG networks predict smoother fields and resolve individual layers of the structure.

Our results are in agreement with previous studies[21,39] that focused on predictions of field distributions in dielectric structures trained on relatively large ($\sim 10^4$ configurations) data sets. Incorporation of physics loss in these NNs resulted in substantial (but limited) improvements in physics consistency (by a factor of $\lesssim 2$). Here, we see similar dynamics for low-resolution networks that require few labeled-data training inputs to achieve their top performance. At the same time, the physics consistency of our medium- and high-resolution networks, which are trained in the data-poor regime, is improved by an order of magnitude as a result of the incorporation of physics-based constraints.

5.2 Knowledge Transfer

As mentioned above, we have attempted knowledge transfer from a pretrained low-resolution network to a medium resolution network and from a pretrained medium resolution network to its high-resolution counterpart. In both cases, a single average-performing lower-resolution
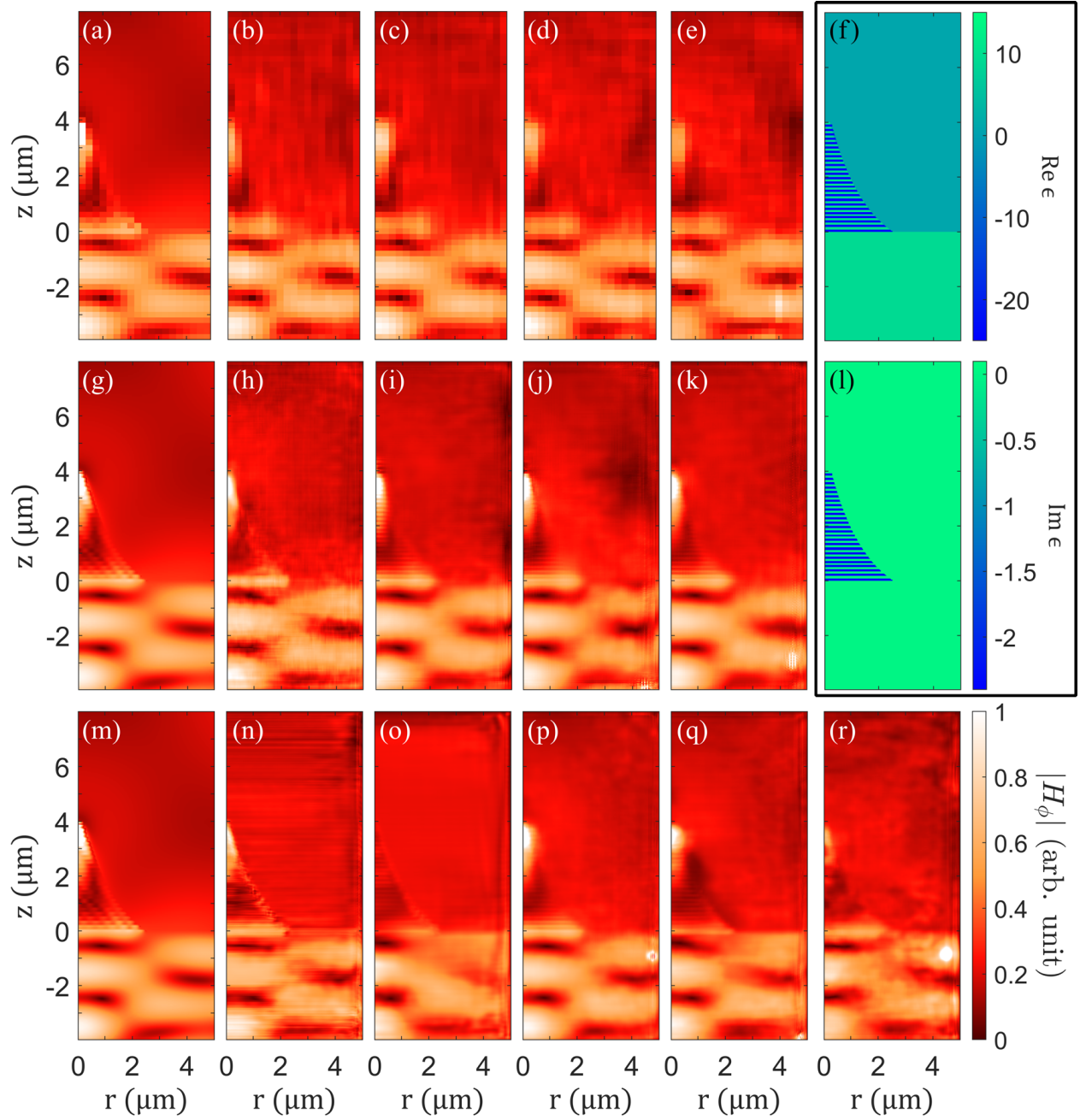
Figure 4: Representative predictions of the NNs with (a..e) low- , (g...k) medium-, and (m...r) high-resolution; input permittivity is shown in panels (f,l); panels (a, g, m) represent ground truth; panels (b, h, n) - predictions of BB_i NNs, panels (c, i, o) - predictions of FE_i networks, panels (d, j, p) - predictions of PG_i networks, and panels (e, k, q) - predictions of UL_i networks; panel (r) illustrates the performance of the TL_i network. Note that higher-performing networks resolve field oscillations on the scale of individual layers within the composite and field concentration near the tip of the funnel.

UL network was chosen as the source of the frozen core of the higher-resolution TL networks. Notably, the low-resolution network poorly resolves the individual layers within the composite. Consistent with this design, implementation of PG loss does not substantially improve network performance (see above), and using a pretrained low-resolution network as a learning-free core of the medium-resolution counterpart does not yield adequate performance of the resulting NN.

In contrast, using a pretrained medium resolution network as a (fixed) core of a high-resolution NN provided reasonable performance. As seen in Figs.3,4, the accuracy of TL_i networks falls between the fully-trained high-resolution FE_i and PG_i NNs.

The physics of finely stratified composites is *analytically* described by effective medium theories (EMT). In the EMT formalism, the spatial distribution of homogenized (averaged over the scale of the inclusion $\sim d$) electromagnetic fields is given by effective parameters (here, $\epsilon_\perp$ and ϵ_{zz}). These homogenized fields, along with equations that relate the effective medium parameters to microscopic distributions of permittivity, can then be used to recover fine-scale field distributions.[1]

The analytical procedure described above is somewhat similar to the operation of the hierarchical TL CNN reported in this work. Indeed, the CNN-based U-nets are known to learn a low-dimensional representation of the underlying phenomena. From this standpoint, while we do not analyze the neural operation of the CNN in detail, the medium-resolution network is likely to learn some form of materials averaging/field recovery by analyzing the transition between the scale of individual layers (resolved at the entrance and exit of the network) and compact representations in its core. The transfer-learning high-resolution wraparound parts of the network likely learn the averaging and upscaling procedures. We reserve the analysis of the relationship between the analytical EMT and the operation of TL-based hierarchial CNNs for future work.

By freezing the inner core of the CNN within knowledge transfer networks we significantly reduce the number of training parameters. Therefore, we expect smaller variability and

faster learning in the TL_i networks as compared with their fully-trained high resolution PG_i counterparts. However, in our implementation, the time required to calculate one training epoch of a TL_i network is almost identical to the time required for one epoch of a PG_i network, indicating that the time spent updating the learnable NN parameters is significantly less than the time spent executing forward and backward propagation steps. Different implementation and optimization settings may affect this result.

At the same time, further analysis (see SI) suggests that TL_i networks converge over a smaller number of epochs than their PG_i counterparts. In addition, in our studies, variation between the performance of the best and the worst TL_i networks was significantly smaller than the variation between the best and the worst PG_i networks.

5.3 Interpolation vs Extrapolation within the models

As described above (see Table 1), the NNs have been trained on multiple subsets of the data derived from FEM solutions, aiming to assess both correctness and generalizability of the proposed PGML networks. Here we are particularly interested in the ability of the NN to generalize the results between different plasma wavelengths of the doped components of the funnels' cores.

In the "interpolating" models (subscripted i), 50% of the data from the sets representing the lowest and the highest plasma frequencies, and an additional 10% from the data set representing the central plasma wavelength were used as labeled training data. The unlabeled networks further included 40% of the central plasma wavelength data set as un-labeled data. Therefore, the CNN would have to deduce the behavior of the composites with $\lambda_p = 7, 10\,\mu m$. For the "extrapolating" (subscripted e) and "extended extrapolating" (subscripted x) networks, a similar approach was used except with the bulk of labeled train-ing data coming from the 7 and 10 µm plasma wavelengths, having the CNN deduce the behavior of the metamaterials with $\lambda_p = 6, 11\,\mu m$

Typically, data interpolation is a much simpler problem than data extrapolation. How-

ever, this general rule does not hold for our analysis. As seen in Fig.3(b,c), the average performance of the two classes of medium-resolution networks is almost identical to each other, indicating that both interpolation and extrapolation tasks (in terms of λ_p) in our study represent similar difficulties to the NNs.

Fig. 5 provides a more in-depth look at this behavior. In general, as characterized by L_ϕ loss, the networks perform their best in predicting the fields within the metamaterials for the same plasma wavelength that comprises the majority of their labeled training set. Indeed, L_ϕ is ~ 2 times lower for the data that has a plasma wavelength that is well-represented in the training set than for the configurations with plasma wavelengths that contribute few or no instances to the labeled training data. Incorporation of physics-based constraints improves the physics-consistency of the results for all values of λ_p by an order of magnitude, indicating that the CNNs learn the general properties of the field distributions but miss the particular boundary conditions that are encoded in the labeled data.

By comparing the performance of extrapolating networks to their "extended" counterparts [Fig.5(c,e)] it is seen that adding very little labeled data can somewhat address this issue of underspecified boundary conditions: introduction of ~ 20 labeled distributions (total) for $\lambda_p = 6, 11\mu m$ reduces the λ_p-specific L_ϕ by $\sim 20\%$ with almost no effect on L_{ph}.

Interestingly, in all scenarios L_{ph} decreases as a function of λ_p. This behavior traces the strength of the resonance in ϵ_{zz} that decays and moves out of the spectral range of the study as λ_p increases (see Fig.1d).
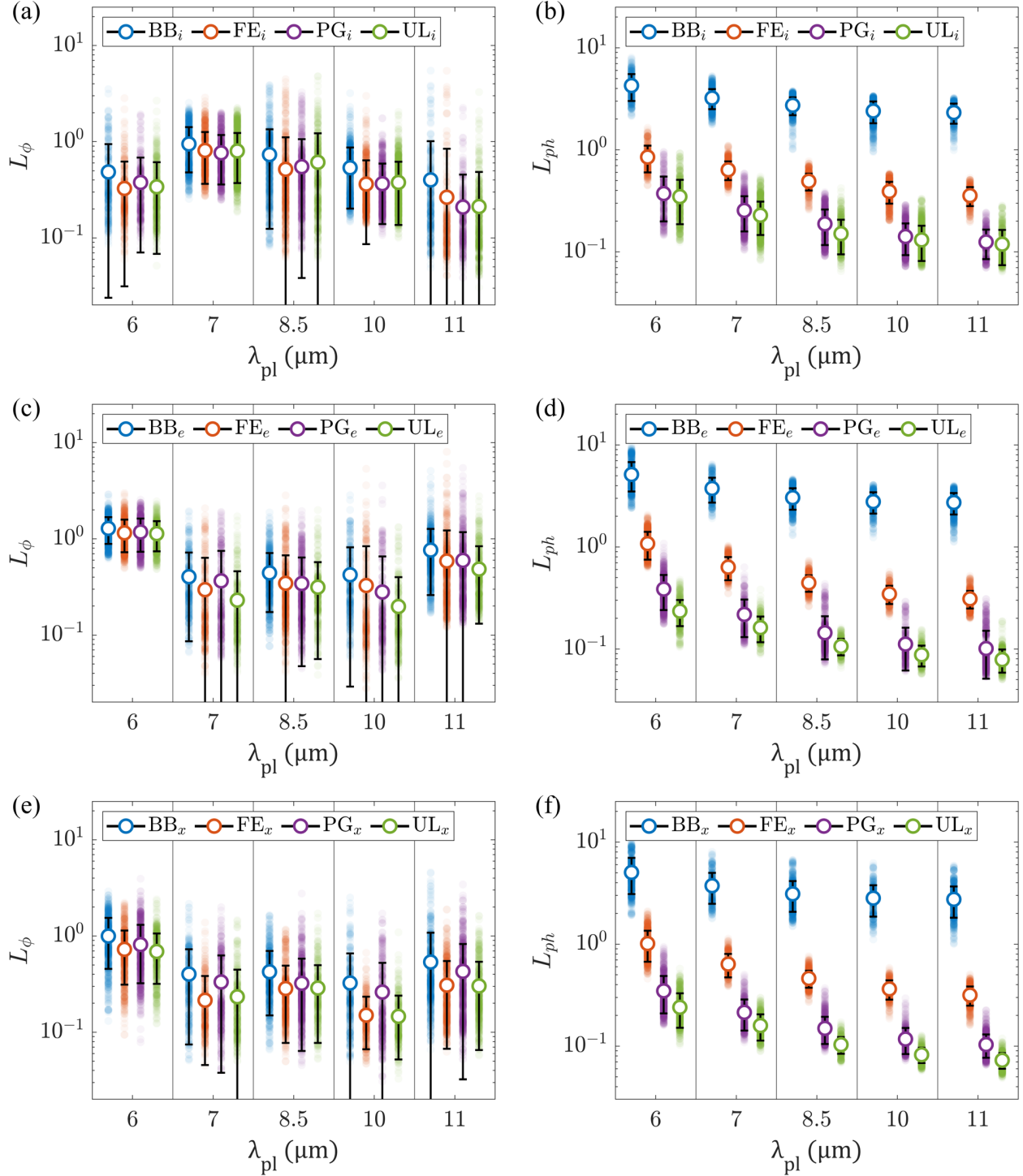
Figure 5: Performance of the medium-resolution networks for predicting the field distribution of composites with given plasma wavelengths; panels (a, c, e) and (b, d, f) represent L_ϕ and L_{ph}, respectively, for (a, b) interpolating, (c, d) extrapolating, and (e, f) extended extrapolating networks; colors represent training protocols; individual predictions are represented as filled semitransparent circles, resulting in the color-coded distributions; solid white circle markers and black bars represent the mean and standard deviations of these distributions.

6 Conclusions

We have presented a hierarchical design of PG neural network surrogate solvers of Maxwell's equations and demonstrated the proposed formalism by predicting the field distributions in hyperbolic metamaterial-based photonic funnels. We have demonstrated that embedding physics information into the ML process, by enforcing the physics-based constraints and by adding unlabeled training configurations, improves the quality of ML predictions in the regime of limited training data. In particular, physics-guided ML predictions are almost two orders of magnitude more physics-consistent than their BB-ML counterparts, even near wavelengths where the layered composite undergoes topological transitions. Separately, we have demonstrated that a hierarchical network architecture enables knowledge transfer from existing pretrained models to higher-resolution NN implementations.

The approach presented can be directly applied to the analysis of complex rotationally-symmetric electromagnetic systems. The technique can be straightforwardly extended to quasi-2D geometries with inclusions of various sizes and shapes by using the appropriate coordinate-representations of Maxwell's equations. The formalism can be further extended to 3D geometries, although we anticipate that such extensions will require significantly larger computational resources.

Funding Sources

Supporting Information Available

The following files are available free of charge.

- supplementary.pdf: an analytical derivation of the regularized physics loss L_{ph}, an

elaboration on our chosen regularization and radial weight functions, a demonstration of the effect of physics consistency on energy conservation, an analysis of the training and validation loss dynamics of our networks

Associated Content

A preprint of this work is available on arXiv.[62]

References

(1) Noginov, M.; Podolskiy, V. A. *Tutorials in Metamaterials*; CRC Press, 2012.

(2) Kildishev, A. V.; Boltasseva, A.; Shalaev, V. M. Planar Photonics with Metasurfaces. *Science* **2013**, *339*, 1232009.

(3) Lin, D.; Fan, P.; Hasman, E.; Brongersma, M. L. Dielectric gradient metasurface optical elements. *Science* **2014**, *345*, 298–302.

(4) Engheta, N.; Ziolkowski, R. *Metamaterials: physics and engineering explorations*; Willey-IEEE press, 2006.

(5) Epstein, A.; Wong, J. P. S.; Eleftheriades, G. V. Cavity-excited Huygens' metasurface antennas for near-unity aperture illumination efficiency from arbitrarily large apertures. *Nature Communications* **2016**, *7*, 10360.

(6) Elhamod, M.; Bu, J.; Singh, C.; Redell, M.; Ghosh, A.; Podolskiy, V.; Lee, W.-C.; Karpatne, A. CoPhy-PGNN: Learning Physics-Guided Neural Networks with Competing Loss Functions for Solving Eigenvalue Problems. *ACM Trans. Intell. Syst. Technol.* **2022**, *13*, 113933.

(7) Ghosh, A.; Elhamod, M.; Bu, J.; Lee, W.-C.; Karpatne, A.; Podolskiy, V. A. Physics-Informed Machine Learning for Optical Modes in Composites. *Advanced Photonics Research* **2022**, *3*, 2200073.

(8) Yu, N.; Genevet, P.; Kats, M. A.; Aieta, F.; Tetienne, J.-P.; Capasso, F.; Gaburro, Z. Light Propagation with Phase Discontinuities: Generalized Laws of Reflection and Refraction. *Science* **2011**, *334*, 333–337.

(9) Banks, J. W.; Henshaw, W. D.; Kildishev, A. V.; Kovačič, G.; Prokopeva, L. J.; Schwendeman, D. W. Solving Maxwell's Equations with a Generalized Dispersive Material Madel on Overset Grids. 2019 International Applied Computational Electromagnetics Society Symposium (ACES). 2019; pp 1–2.

(10) Taflove, A.; Hagness, S. *Computational Electrodynamics: The Finite-Difference Time-Domain Method*; Artech House, 2005.

(11) Jin, J.-M. *The Finite Element Method in Electromagnetics*; John Wiley & Sons, 2014.

(12) Inc., C. COMSOL v. 6.3. `https://www.comsol.com/` (accessed 2025-07-08).

(13) Malkiel, I.; Mrejen, M.; Nagler, A.; Arieli, U.; Wolf, L.; Suchowski, H. Plasmonic nanostructure design and characterization via Deep Learning. *Light: Science & Applications* **2018**, *7*, 60.

(14) Khatib, O.; Ren, S.; Malof, J.; Padilla, W. J. Deep Learning the Electromagnetic Properties of Metamaterials—A Comprehensive Review. *Advanced Functional Materials* **2021**, *31*, 2101748.

(15) Jiang, J.; Chen, M.; Fan, J. A. Deep neural networks for the evaluation and design of photonic devices. *Nature Reviews Materials* **2021**, *6*, 679–700.

(16) Lin, Z.; Roques-Carmes, C.; Pestourie, R.; Soljačić, M.; Majumdar, A.; Johnson, S. G.

End-to-end nanophotonic inverse design for imaging and polarimetry. *Nanophotonics* **2021**, *10*, 1177–1187.

(17) Kudyshev, Z. A.; Sychev, D.; Martin, Z.; Yesilyurt, O.; Bogdanov, S. I.; Xu, X.; Chen, P.-G.; Kildishev, A. V.; Boltasseva, A.; Shalaev, V. M. Machine learning assisted quantum super-resolution microscopy. *Nature communications* **2023**, *14*, 4828.

(18) Kudyshev, Z. A.; Kildishev, A. V.; Shalaev, V. M.; Boltasseva, A. Optimizing startshot lightsail design: A generative network-based approach. *ACS Photonics* **2021**, *9*, 190–196.

(19) Li, W.; Barati Sedeh, H.; Tsvetkov, D.; Padilla, W. J.; Ren, S.; Malof, J.; Litchinitser, N. M. Machine Learning for Engineering Meta-Atoms with Tailored Multipolar Resonances. *Laser & Photonics Reviews* **2024**, 2300855.

(20) Pestourie, R.; Mroueh, Y.; Nguyen, T. V.; Das, P.; Johnson, S. G. Active learning of deep surrogates for PDEs: application to metasurface design. *npj Computational Materials* **2020**, *6*, 164.

(21) Zhelyeznyakov, M.; Fröch, J.; Wirth-Singh, A.; Noh, J.; Rho, J.; Brunton, S.; Majumdar, A. Large area optimization of meta-lens via data-free machine learning. *Communications Engineering* **2023**, *2*, 60.

(22) Ma, W.; Cheng, F.; Xu, Y.; Wen, Q.; Liu, Y. Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy. *Advanced Materials* **2019**, *31*, 1901111.

(23) Ma, W.; Cheng, F.; Liu, Y. Deep-Learning-Enabled On-Demand Design of Chiral Metamaterials. *ACS Nano* **2018**, *12*, 6326–6334.

(24) Lu, L.; Pestourie, R.; Yao, W.; Wang, Z.; Verdugo, F.; Johnson, S. G. Physics-informed

neural networks with hard constraints for inverse design. *SIAM Journal on Scientific Computing* **2021**, *43*, B1105–B1132.

(25) Hestness, J.; Ardalani, N.; Diamos, G. Beyond human-level accuracy: computational challenges in deep learning. Proceedings of the 24th Symposium on Principles and Practice of Parallel Programming. New York, NY, USA, 2019; p 1–14.

(26) Zhao, Y.; Chen, J.; Oymak, S. On the Role of Dataset Quality and Heterogeneity in Model Confidence. 2020; `https://arxiv.org/abs/2002.09831` (accessed 2025-07-09).

(27) Jiang, J.; Fan, J. A. Global optimization of dielectric metasurfaces using a physics-driven neural network. *Nano letters* **2019**, *19*, 5366–5372.

(28) Mao, C.; Lupoiu, R.; Dai, T.; Chen, M.; Fan, J. A. Towards General Neural Surrogate Solvers with Specialized Neural Accelerators. 2024; `https://arxiv.org/abs/2405.02351` (accessed 2025-07-09).

(29) Karniadakis, G.; Kevrekidis, I.; Lu, L.; Perdikaris, P.; Wang, S.; L., Y. Physics-informed machine learning. *Nat. Rev. Phys* **2021**, 422–440.

(30) Huang, L.; Chen, H.; Liu, T.; Ozcan, A. Self-supervised learning of hologram reconstruction using physics consistency. *Nat. Machine Intelligence* **2023**, 895–907.

(31) Bar-Sinai, Y.; Hoyer, S.; Hickey, J.; Brenner, M. P. Learning data-driven discretizations for partial differential equations. *Proceedings of the National Academy of Sciences* **2019**, *116*, 15344–15349.

(32) Wang, S.; Wang, H.; Perdikaris, P. Learning the solution operator of parametric partial differential equations with physics-informed DeepONets. *Science advances* **2021**, *7*, eabi8605.

(33) Lu, L.; Jin, P.; Pang, G.; Zhang, Z.; Karniadakis, G. E. Learning nonlinear operators via

DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence* **2021**, *3*, 218–229.

(34) Seidman, J.; Kissas, G.; Perdikaris, P.; Pappas, G. J. NOMAD: Nonlinear manifold decoders for operator learning. *Advances in Neural Information Processing Systems* **2022**, *35*, 5601–5613.

(35) Goswami, S.; Bora, A.; Yu, Y.; Karniadakis, G. E. In *Machine Learning in Modeling and Simulation: Methods and Applications*; Rabczuk, T., Bathe, K.-J., Eds.; Springer International Publishing: Cham, 2023; pp 219–254.

(36) Li, Z.; Kovachki, N.; Azizzadenesheli, K.; Liu, B.; Bhattacharya, K.; Stuart, A.; Anandkumar, A. Fourier Neural Operator for Parametric Partial Differential Equations. 2021; https://arxiv.org/abs/2010.08895 (accessed 2025-07-09).

(37) Yao, H. M.; Jiang, L.; Sha, W. E. I. Enhanced Deep Learning Approach Based on the Deep Convolutional Encoder–Decoder Architecture for Electromagnetic Inverse Scattering Problems. *IEEE Antennas and Wireless Propagation Letters* **2020**, *19*, 1211–1215.

(38) Raissi, M.; Perdikaris, P.; Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* **2019**, *378*, 686–707.

(39) Chen, M.; Lupoiu, R.; Mao, C.; Huang, D.-H.; Jiang, J.; Lalanne, P.; Fan, J. A. High speed simulation and freeform optimization of nanophotonic devices with physics-augmented deep learning. *ACS Photonics* **2022**, *9*, 3110–3123.

(40) Li, K.; Simmons, E.; Briggs, A.; Nordin, L.; Xu, J.; Podolskiy, V.; Wasserman, D. Sub-diffraction limited photonic funneling of light. *Adv. Opt. Mater* **2020**, *24*, 2001321.

(41) Govyadinov, A. A.; Podolskiy, V. A. Metamaterial photonic funnels for subdiffraction light compression and propagation. *Phys. Rev. B* **2006**, *73*, 155108.

(42) LaMountain, J.; Raju, A.; Wasserman, D.; Podolskiy, V. A. Anomalous reflection for highly efficient subwavelength light concentration and extraction with photonic funnels. *Nanophotonics* **2024**, *13*, 4625–4637.

(43) Drachev, V. P.; Podolskiy, V. A.; Kildishev, A. V. Hyperbolic metamaterials: new physics behind a classical problem. *Opt. Express* **2013**, *21*, 15048–15064.

(44) Prokes, S.; Glembocki, O. J.; Livenere, J. E.; Tumkur, T. U.; Kitur, J. K.; Zhu, G.; Wells, B.; Podolskiy, V. A.; Noginov, M. A. Hyperbolic and plasmonic properties of Silicon/Ag aligned nanowire arrays. *Opt. Express* **2013**, *21*, 14962–14974.

(45) Vasilantonakis, N.; Wurtz, G. A.; Podolskiy, V. A.; Zayats, A. V. Refractive index sensing with hyperbolic metamaterials: strategies for biosensing and nonlinearity enhancement. *Opt. Express* **2015**, *23*, 14329–14343.

(46) Galfsky, T.; Gu, J.; Narimanov, E. E.; Menon, V. M. Photonic hypercrystals for control of light–matter interactions. *Proceedings of the National Academy of Sciences* **2017**,

(47) Jackson, J. *Classical Electrodynamics*; Willey, 1999.

(48) Poddubny, A.; Iorsh, I.; Belov, P.; Kivshar, Y. Hyperbolic metamaterials. *Nature Photonics* **2013**, *7*, 948–957.

(49) Shekhar, P.; Atkinson, J.; Jacob, Z. Hyperbolic metamaterials: fundamentals and applications. *Nano convergence* **2014**, *1*, 1–17.

(50) Law, S.; Adams, D.; Taylor, A.; Wasserman, D. Mid-infrared designer metals. *Optics express* **2012**, *20*, 12155–12165.

(51) Hou, X.; Duan, J.; Qiu, G. Deep Feature Consistent Deep Image Transformations: Downscaling, Decolorization and HDR Tone Mapping. 2017; `https://arxiv.org/abs/1707.09482` (accessed 2025-06-11).

(52) Hou, X.; Gong, Y.; Liu, B.; Sun, K.; Liu, J.; Xu, B.; Duan, J.; Qiu, G. Learning Based Image Transformation Using Convolutional Neural Networks. *IEEE Access* **2018**, *6*, 49779–49792.

(53) Li, H.; Chen, L.; Qiu, J. Convolutional Neural Networks for Multifrequency Electromagnetic Inverse Problems. *IEEE Antennas and Wireless Propagation Letters* **2021**, *20*, 1424–1428.

(54) Huang, G.; Liu, K.; Liang, J.; Cai, C.; Gu, Z. H.; Qi, F.; Li, Y.; Yu, Z. L.; Wu, W. Electromagnetic Source Imaging via a Data-Synthesis-Based Convolutional Encoder–Decoder Network. *IEEE Transactions on Neural Networks and Learning Systems* **2024**, *35*, 6423–6437.

(55) Kudyshev, Z. A.; Kildishev, A. V.; Shalaev, V. M.; Boltasseva, A. Machine learning–assisted global optimization of photonic devices. *Nanophotonics* **2020**, *10*, 371–383.

(56) Gong, R.; Tang, Z. Investigation of convolutional neural network U-net under small datasets in transformer magneto-thermal coupled analysis. *COMPEL - The international journal for computation and mathematics in electrical and electronic engineering* **2020**, *39*, 959–970.

(57) Lim, J.; Psaltis, D. MaxwellNet: Physics-driven deep neural network training based on Maxwell's equations. *APL Photonics* **2022**, *7*, 011301.

(58) Zhang, P.; Hu, Y.; Jin, Y.; Deng, S.; Wu, X.; Chen, J. A Maxwell's Equations Based Deep Learning Method for Time Domain Electromagnetic Simulations. 2020 IEEE Texas Symposium on Wireless and Microwave Circuits and Systems (WMCS). 2020; pp 1–4.

(59) Lynch, S.; LaMountain, J.; Fan, B.; Bu, J.; Raju, A.; Wasserman, D.; Karpatne, A.; Podolskiy, V. A. Physics informed machine learning for photonic funnels. 2025; `https://github.com/viktor-podolskiy/` `physics-informed-machine-learning-for-photonic-funnels` (accessed 2025-06-11).

(60) Lynch, S.; LaMountain, J.; Fan, B.; Bu, J.; Raju, A.; Wasserman, D.; Karpatne, A.; Podolskiy, V. A. PGML Dataset. 2025; `https://figshare.com/articles/dataset/` `PGML_Dataset/28439978` (accessed 2025-02-21).

(61) He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016; pp 770–778.

(62) Lynch, S.; LaMountain, J.; Fan, B.; Bu, J.; Raju, A.; Wasserman, D.; Karpatne, A.; Podolskiy, V. A. Physics-guided hierarchical neural networks for Maxwell's equations in plasmonic metamaterials. 2025; `https://arxiv.org/abs/2502.17644` (accessed 2025-07-09).

TOC Graphic
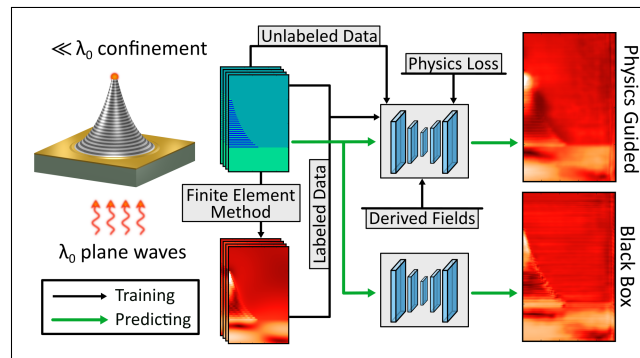
Physics-guided hierarchical neural networks for Maxwell's equations in plasmonic metamaterials; Supplementary Materials

Sean Lynch,[†] Jacob LaMountain,[*,‡] Bo Fan,[‡] Jie Bu,[¶] Amogh Raju,[§] Dan Wasserman,[§] Anuj Karpatne,[¶] and Viktor A. Podolskiy[‡]

[†]Miner School of Computer Science

University of Massachusetts Lowell

Lowell, MA 01854, USA

[‡]Department of Physics and Applied Physics

University of Massachusetts Lowell

Lowell, MA 01854, USA

[¶]Department of Computer Science

Virginia Tech

Blacksburg, VA 24061, USA

[§]Department of Electrical and Computer Engineering

University of Texas Austin

Austin, TX 78712, USA

E-mail: jacob_lamountain@student.uml.edu

Maxwell's Equations in Cylindrical Geometry

As mentioned in the main manuscript, we use the rotational symmetry of our problem to reduce the three-dimensional vectorial Maxwell's equations to equations describing the behavior of the ϕ-components of the electric and magnetic fields (which vary smoothly throughout the geometry). Assuming that all fields are proportional to $\exp(-i\phi)$, once E_ϕ and H_ϕ are known (for example, as predicted by the neural net), the remaining components of the fields can be calculated via

$$\vec{E}_{rz} = \frac{-i}{\epsilon\frac{\omega^2}{c^2} - \frac{1}{r^2}}\left(-\frac{1}{r}\vec{D}_{rz}E_\phi - \frac{\omega}{c}\hat{\phi} \times \vec{D}_{rz}H_\phi\right) \tag{S1}$$

$$\vec{H}_{rz} = \frac{-i}{\epsilon\frac{\omega^2}{c^2} - \frac{1}{r^2}}\left(-\frac{1}{r}\vec{D}_{rz}H_\phi + \epsilon\frac{\omega}{c}\hat{\phi} \times \vec{D}_{rz}E_\phi\right) \tag{S2}$$

where the differential operator $\vec{D}_{rz}$ is defined by $\vec{D}_{rz}f = \hat{r}\frac{1}{r}\frac{\partial}{\partial r}(rf) + \hat{z}\frac{\partial f}{\partial z}$.

Maxwell's equations also provide additional constraints, ensuring self-consistency of the field components:

$$\frac{\partial H_r}{\partial z} - \frac{\partial H_z}{\partial r} = -i\epsilon\frac{\omega}{c}E_\phi \tag{S3}$$

$$\frac{\partial E_r}{\partial z} - \frac{\partial E_z}{\partial r} = i\frac{\omega}{c}H_\phi \tag{S4}$$

the first of which is used as a basis for L_{ph} in the manuscript.

Regularization Function

The neural network directly predicts E_ϕ and H_ϕ, and the physics layer then calculates H_r and H_z using Eq.(S2), above. It is seen, however, that for transparent materials, Eq.(S2) diverges when $r^2\epsilon\omega^2/c^2 = 1$.

In approximate numerical solutions (such as those analyzed in our work), this condition leads to instabilities that – if left unaddressed – would dominate both L_{rz} and L_{ph} loss

functions.

To address these underlying instabilities, we introduce the regularization function $R(r,z)$ in such a way that regularized fields $\vec{\mathcal{E}} = R\vec{E}$ and $\vec{\mathcal{H}} = R\vec{H}$ remain finite within the simulation domain. Explicitly,

$$R(r,z) = 0.1\frac{\epsilon(r,z)r^2\frac{\omega^2}{c^2} - 1}{\epsilon(r,z)r^2\frac{\omega^2}{c^2} + 0.1} \tag{S5}$$

is used in our work (In principle, any $R(r,z)$ that vanishes, at least linearly, when $r^2\epsilon\omega^2/c^2 = 1$ can be used).

To find an appropriate physics loss function, we first recast Eq.(S3) in terms of regularized fields

$$\frac{\partial}{\partial z}\left(\frac{\mathcal{H}_r}{R}\right) - \frac{\partial}{\partial r}\left(\frac{\mathcal{H}_z}{R}\right) = -i\epsilon\frac{\omega}{c}\frac{\mathcal{E}_\phi}{R}. \tag{S6}$$

We then apply the derivatives and rearrange the resulting relationships, arriving at:

$$\frac{\partial}{\partial z}\left(R\mathcal{H}_r\right) - \frac{\partial}{\partial r}\left(R\mathcal{H}_z\right) + 2\left(\mathcal{H}_z\frac{\partial R}{\partial r} - \mathcal{H}_r\frac{\partial R}{\partial z}\right) + i\epsilon\frac{\omega}{c}R\mathcal{E}_\phi = 0. \tag{S7}$$

Recasting the latter equation back to the actual fields yields the physics residual used for physics loss in our work

$$\Lambda_{ph} = \frac{\partial}{\partial z}\left(R^2 H_r\right) - \frac{\partial}{\partial r}\left(R^2 H_z\right) + 2\left(RH_z\frac{\partial R}{\partial r} - RH_r\frac{\partial R}{\partial z}\right) + i\epsilon\frac{\omega}{c}R^2 E_\phi. \tag{S8}$$

Note that physics-consistent solutions should satisfy $\Lambda_{ph}(r,z) \equiv 0$.

Radial Weight Function

The fundamental electromagnetic phenomena enabled by photonic funnels (anomalous reflection and subdiffractive light confinement) are encoded in the field distributions within and in close proximity to the funnels. Since these field distribution features are more important for understanding the electromagnetism of the funnels and at the same time are more com-

plicated than the diffraction-limited field distributions outside the funnels, a weight function is used to improve learning of the fields at small radii. In our networks, this radial weight function is a sigmoid given by

$$w(r) = \frac{5}{1 + 10e^{2(r-3)}} + 0.5, \tag{S9}$$

where r is in µm.

Physics-Consistency and Energy Conservation

An important physical principle, which should hold (at least approximately, given our discretization) for physically consistent fields is energy conservation. We may therefore use the deviation from energy conservation as an additional measure of physics-inconsistency. Furthermore, we can utilize this metric to demonstrate the impact of imposing physics-consistency through the inclusion of L_{ph} in the training loss function.

The conservation of energy for monochromatic fields in linear dispersive media can be written as

$$\frac{\omega}{c}\text{Im}\left\{\epsilon\left|\vec{E}\right|^2 + \mu\left|\vec{H}\right|^2\right\} - \nabla \cdot \text{Re}\{\vec{E} \times \vec{H}^*\} = 0. \tag{S10}$$

The left-hand side of this equation acts as a kind of "Poynting residual" when applied to electromagnetic fields, with nonzero values representing local violations of energy conservation.

As was the case with Λ_{ph}, to make use of this residual with network-predicted fields, we must suppress the numerical instabilities which arise from calculating the r and z field components via Eq.S1 and Eq.S2. Because this residual expression is proportional to the square of the fields, adequately suppressing the numerical errors introduced in calculating these additional components requires multiplication by R^2 resulting in the regularized Poynting

4

residual:

$$\Lambda_P = R^2 \left(\frac{\omega}{c} \mathrm{Im}\{\epsilon\} \left| \vec{E} \right|^2 - \nabla \cdot \mathrm{Re}\{\vec{E} \times \vec{H}^*\} \right) \tag{S11}$$

for non-magnetic materials. In analogy with how our physics loss is defined in terms of our physics residual, a Poynting loss, L_P, was calculated as the average magnitude of the Poynting residual normalized by the square of the maximum field, $\max(|E_\phi|, |H_\phi|)$ (here, Gaussian units are used for convenience, making amplitudes of electric and magnetic fields of the plane wave comparable to each other).
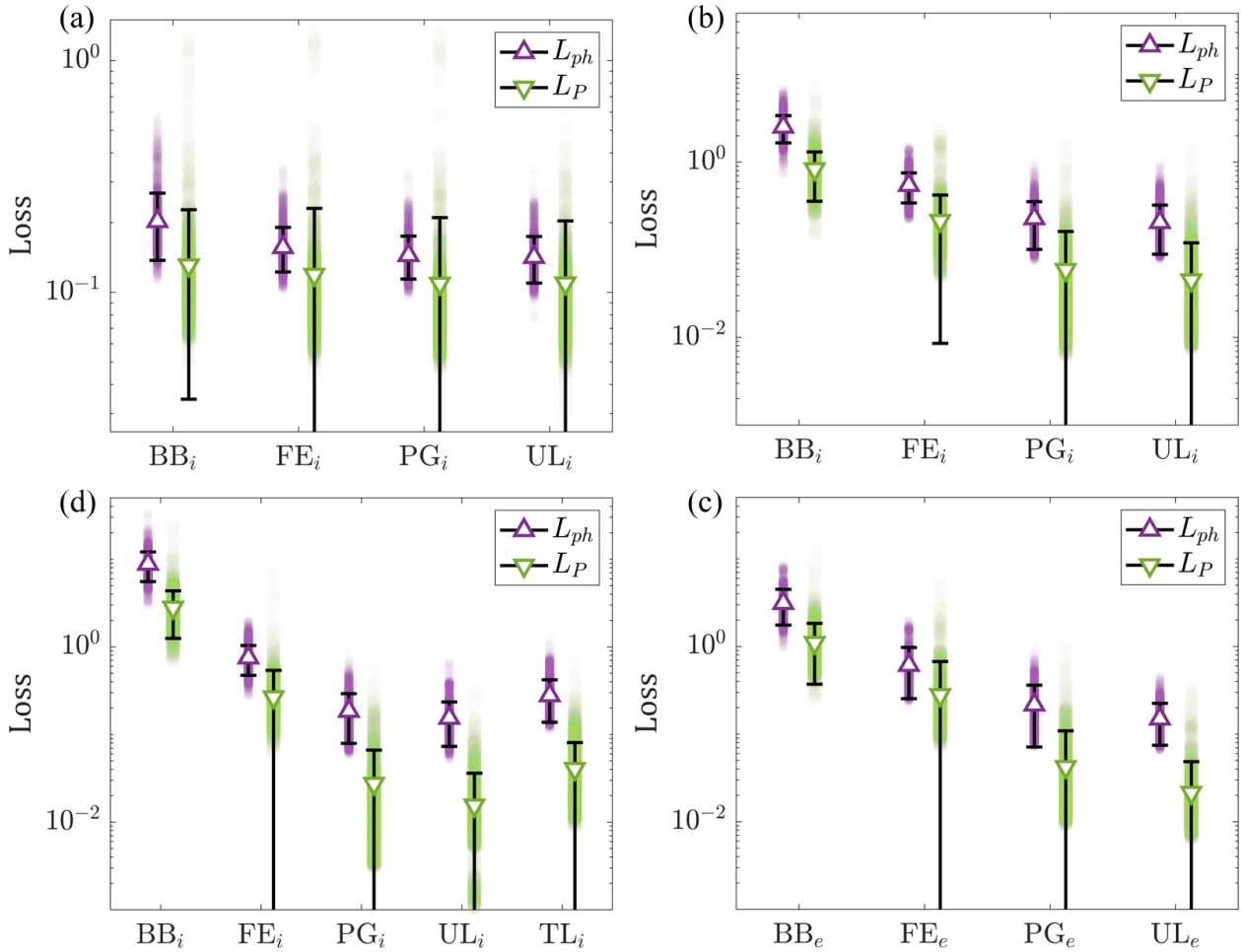

Figure S1: Physics and Poynting losses of NNs with different architectures and training protocols, evaluated on the subset of data that was not used in training sets; panels (a,b,c,d) represent low-resolution (a), medium resolution (b,c), and high-resolution (d) networks; Losses of individual predictions are represented as filled semi-transparent circles; solid white markers and black bars represent the mean and standard deviations of these distributions.

To demonstrate that the physics-consistency enforced by L_{ph} results in fields which better satisfy energy conservation, we calculated the Poynting loss for the predictions of networks of each size across all test data. The results are shown in Fig.S1. It is clearly seen that improving physics-consistency (characterized by lowering L_{ph}) yields improvement in energy conservation (characterized by lowering L_P).
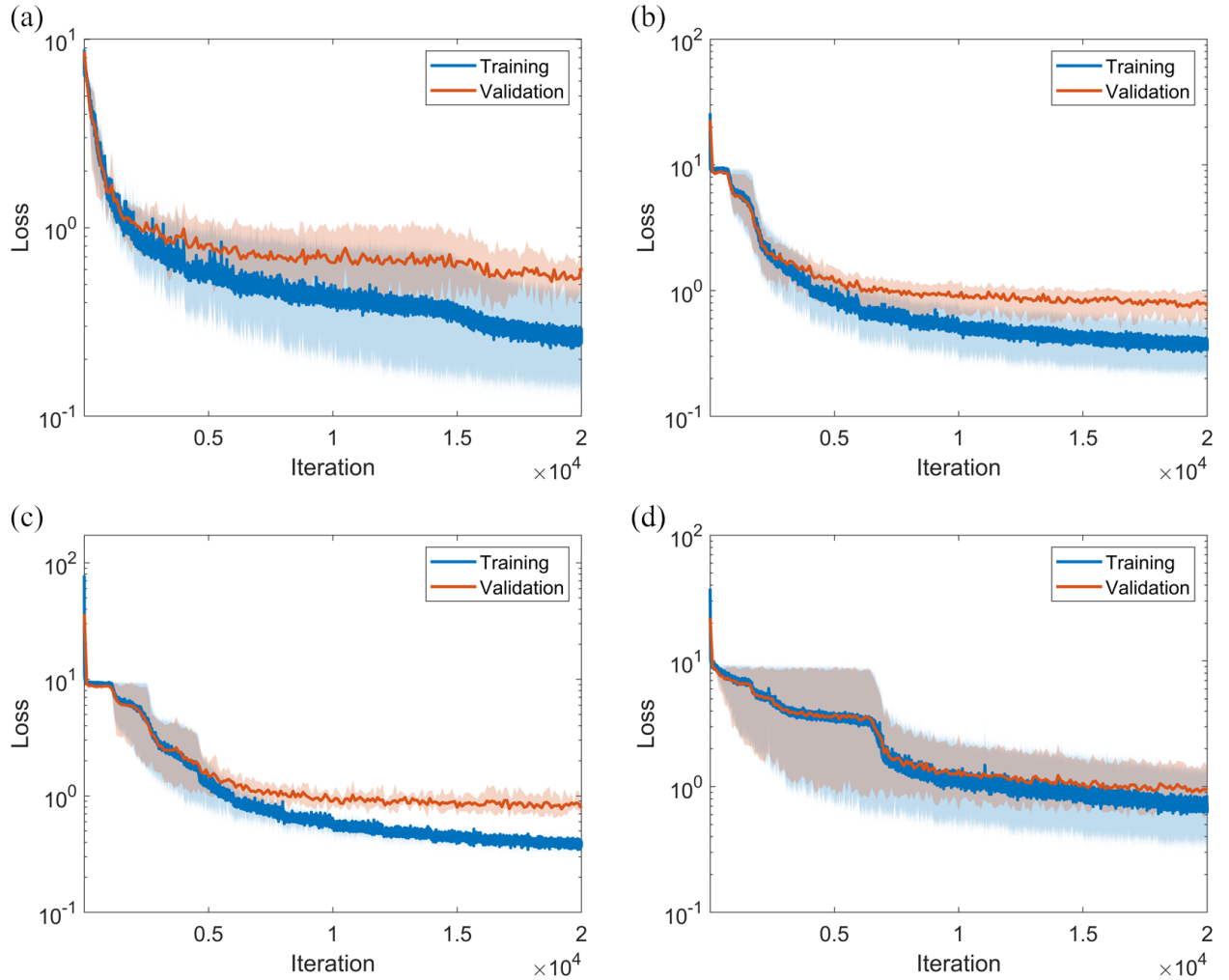
Loss Dynamics

Figure S2: Loss dynamics of medium resolution networks showing training and validation loss against training iteration for (a) black-box, (b) field-enhanced, (c) physics-guided, and (d) unlabeled-trained networks. The solid lines show the losses averaged over networks while the shaded regions are bounded by the maximum and minimum losses at each iteration.

The training and validation loss curves for each of the four medium resolution "interpolating" network configurations are presented in Fig. S2. Across all cases, the validation loss closely follows the trend of the training loss throughout the training process. While the two losses begin at similar values—occasionally with the validation loss slightly lower during the early epochs—they gradually separate as training progresses with the validation loss becoming slightly greater towards convergence. This pattern is expected given the model's exposure to the training data and indicates stable, consistent generalization to unseen data, with no evidence of significant overfitting.

Similar results were seen for high-resolution networks, pointing again to stable training without overfitting. Additionally, we compared the validation loss dynamics between high-resolution UL_i and TL_i networks, summarized in Fig.S3. It can be seen that TL_i networks require fewer training epochs to converge and tend to outperform their UL_i counterparts during early training iterations. However, UL_i networks tend to eventually outperform their TL_i counterparts.

This dynamics reflects the comparatively smaller parameter space of TL networks, which are therefore more readily able to find optimal network configurations and less likely to fall into (and become stuck within) local minima. However, this smaller network dimensionality comes at the price of expressiveness, somewhat limiting the ability of TL networks to fine-tune their predictions.
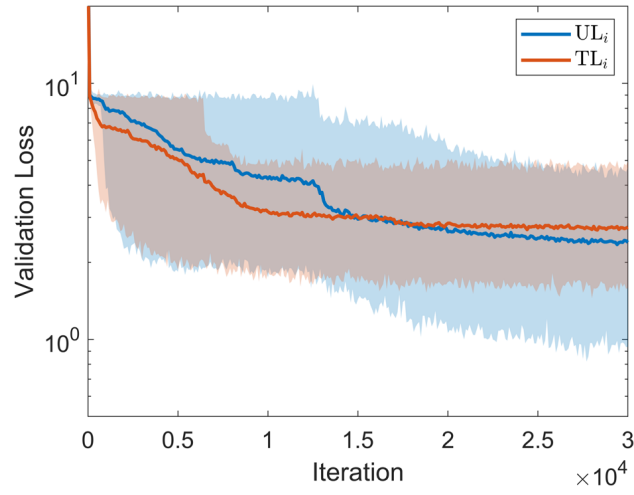
Figure S3: Validation loss dynamics of high-resolution unlabeled-trained and transfer-learning networks. The solid lines show the losses averaged over networks while the shaded regions are bounded by the maximum and minimum losses at each iteration.