Accelerating Finite-Difference Frequency-Domain Simulations for Inverse Design Problems in Nanophotonics using Deep Learning

LUKAS A. K. SCHULTE,[1,2,3,†] MARCO BUTZ,[1,2,3,†] MARLON BECKER,[4,5] BENJAMIN RISSE[4,5] AND CARSTEN SCHUCK[1,2,3,*]

[1] *Institute of Physics, University of Münster, Wilhelm-Klemm-Str. 10, 48149 Münster, Germany*
[2] *Center for NanoTechnology (CeNTech), Heisenbergstr. 11, 48149 Münster, Germany*
[3] *Center for Soft Nanoscience (SoN), Busso-Peus-Str. 10, 48149 Münster, Germany*
[4] *Institute for Geoinformatics, University of Münster, Heisenbergstr. 2, 48149 Münster, Germany*
[5] *Institute for Computer Science, University of Münster, Einsteinstraße 62, 48149 Münster, Germany*
[†] *Both authors contributed equally to this work*
[*] *carsten.schuck@uni-muenster.de*

Abstract: Inverse design of nanophotonic devices becomes increasingly relevant for the development of complex photonic integrated circuits. Electromagnetic first-order simulations contribute the overwhelming computational cost to the optimization routines in established inverse design algorithms, requiring more efficient methods for enabling improved and more complex design process flows. Here we present such a method to predict the electromagnetic field distribution for pixel-discrete planar inverse designed structures using deep learning. Our model is able to infer accurate predictions used to initialize a conventional Finite Difference Frequency-Domain-algorithm and thus lowers the time required for simulating the electromagnetic response of nanophotonic device layouts by about 50 %. We demonstrate the applicability of our deep learning method for inverse design of photonic integrated powersplitters and mode converters and we highlight the possibility of exploiting previous learning results in subsequent design tasks of novel functionalities via finetuning on reduced data sets, thus improving computational speed further.

1. Introduction

Inverse design plays an increasingly important role in realizing compact and high-performance devices for photonic integrated circuits [1, 2]. The efficiency and complexity of nanophotonic devices with computer generated layouts is increasing steadily, benefiting from recent development of sophisticated optimization algorithms. The range of applications spans from efficient fiber-to-chip coupling [3–5] over complex multi-port sub-wavelength interference-based devices [6] to the realization of well established fundamental circuit components in novel and promising, yet challenging material platforms [7]. Current efforts address demands for higher device performance and more complex functionalities, requiring either larger device footprints [8] or smaller feature sizes [9] to increase the solution space of the optimization problem. The correspondingly larger simulation mesh imposes challenging requirements onto the electromagnetic simulation routines employed in the evaluation stages of iterative design processes. Moreover, the increasing utilization of data driven [10] and machine learning-based [2] approaches leads to a high number of electromagnetic simulations being required during the optimization process for larger problem statements, causing a significantly higher consumption of computational resources. In order to reduce the computational costs of electromagnetic first-order simulations, which demand the biggest share of the employed resources, it becomes imperative to develop new evaluation routines tailored to the specific requirements of iterative nanophotonic inverse design algorithms.

Numerous deep learning based approaches have been demonstrated to accelerate first-order electromagnetic simulations, ranging from an implementation of the perfectly matched layer boundary condition to reduce ill-conditioning of the coefficient matrices [11, 12] over reduction of numeric dispersion [13] to employing surrogate models for the optimization of periodic nanostructures [14, 15]. While the acceleration of simulation algorithms using deep learning in the context of metamaterials has gained great popularity, the application to pixel-discrete inverse designed structures featuring complex and non-periodic geometries poses new challenges.

In this work, we show how the computation time required to solve three dimensional finite-difference frequency-domain (FDFD) problems in the context of iterative nanophotonic inverse design methods can significantly be decreased. We demonstrate a deep learning based pipeline to predict an initial guess of the electromagnetic field, which closely resembles the result of a first-order FDFD-simulation. We use the predicted time-harmonic field to seed commonly used iterative solving methods for large complex-valued sparse matrices, such as the widely-used BiConjugate Gradient Stabilized (BiCGSTAB) method [16]. Our model, based on the U-Net architecture [17], is trained on samples consisting of pairs of inverse designed structures with the corresponding simulated time-harmonic electromagnetic response to the involved input-modes. These samples are picked from intermediate states of an iterative inverse design optimization procedure, which makes our approach independent of specific algorithm choices, thus seamlessly integrating into established inverse design workflows. We here consider the FDFD-residual as a metric for overall performance and find that our deep-learning model is capable of providing initial guesses that closely approach final solutions, therewith vastly outperforming common random or zero-vector initializations. We demonstrate the method on an exemplary optimization procedure of an asymmetric powersplitter where we save 47.3 % of the FDFD-iterations on average, which translate directly into time and resource savings. Once trained on a specific device, the model can quickly adapt to different device functionalities and surrounding nanophotonic geometries. We demonstrate the versatility of our method by finetuning a pretrained model on a very small dataset of unseen devices featuring structures from the optimization of a symmetric powersplitter and a fundamental transverse electric (TE_{00}) to TE_{20} modeconverter. Successful training is quickly achieved, resulting in a decrease of the number of iterations required to solve FDFD-problems on these devices by substantial amounts.

2. Concept and integration

A nanophotonic inverse design problem is typically defined by an area of variable permittivity surrounded by a fixed waveguide geometry together with physical constants such as the refractive indices of the materials involved and the objective function which maps an external excitation to an arbitrary figure of merit. The inverse design algorithm then usually processes information derived from the electromagnetic response of the given structure to an excitation such as an incident waveguide mode to acquire insights into how the variable part of the examined permittivity distribution needs to be modified to achieve the design objectives. This process, illustrated in the upper closed loop show of Figure 1, is performed in an iterative manner with the ultimate goal of maximizing the objective function. However, calculating the electromagnetic response of a given structure is a time-consuming and resource-intensive task because the accurate simulation of structures exhibiting sub-wavelength sized features prohibits the use of approximations and thus requires first-order approaches, such as fine grained finite-element or finite-difference methods, where the latter are commonly encountered in the context of pixel-discrete nanophotonic inverse design employing Yee-grid discretization [18]. The finite-difference equations are often solved in the frequency domain by calculating the time-harmonic field using the FDFD method, because it is straightforward to derive gradient information from steady state fields and due to faster convergence behavior for resonant structures. To derive the FDFD problem from Maxwell's equations we can, without loss of generality, assume a time dependence of $\exp(-i\omega t)$, which yields the time-harmonic Maxwell equation for the electric field E,

$$\nabla \times \mu_0^{-1} \nabla \times \vec{E} - \omega^2 \epsilon \vec{E} = -i\omega \vec{J}, \tag{1}$$

where μ_0 and ϵ are the magnetic permeability and electric permittivity, respectively, ω is the angular frequency and $\vec{J}$ is the current distribution of the source. Using the Yee-discretized forms of the operators and field components, Equation 1 can be rewritten as a system of coupled equations using the relations $\mathbf{A} = \nabla \times \mu_0^{-1} \times -\omega^2\epsilon$, $\vec{x} = E$ and $b = -i\omega\vec{J}$ yielding

$$\mathbf{A} \cdot \vec{x} - \vec{b} = \vec{0}. \tag{2}$$

Here, an exact solution vector satisfies Equation 2 such that the right hand side (RHS) is zero. An inexact solution leads to a RHS different from zero, where we refer to

$$r = ||\mathbf{A} \cdot \vec{x} - \vec{b}||_2 \tag{3}$$

as the FDFD-residual, which can be understood as a metric for how closely $\vec{x}$ resembles the analytically exact solution. $\mathbf{A}$ is a complex valued, ill-conditioned, sparse Matrix, which grows rapidly in size for three dimensional simulations. Therefore, the common approach to solve for $\vec{x}$ is iterative refinement using methods such as BiCGSTAB. These methods iteratively modify $\vec{x}$, such that r is reduced. As soon as r falls below a certain threshold, r_{thresh}, the solution is considered sufficiently accurate and the simulation routine terminates. In general, initializing the simulation with a starting vector $\vec{x}_0$ that is similar to a solution $\vec{x}$ satisfying the residual related convergence criterion leads to a decrease in the number of required FDFD-iterations.

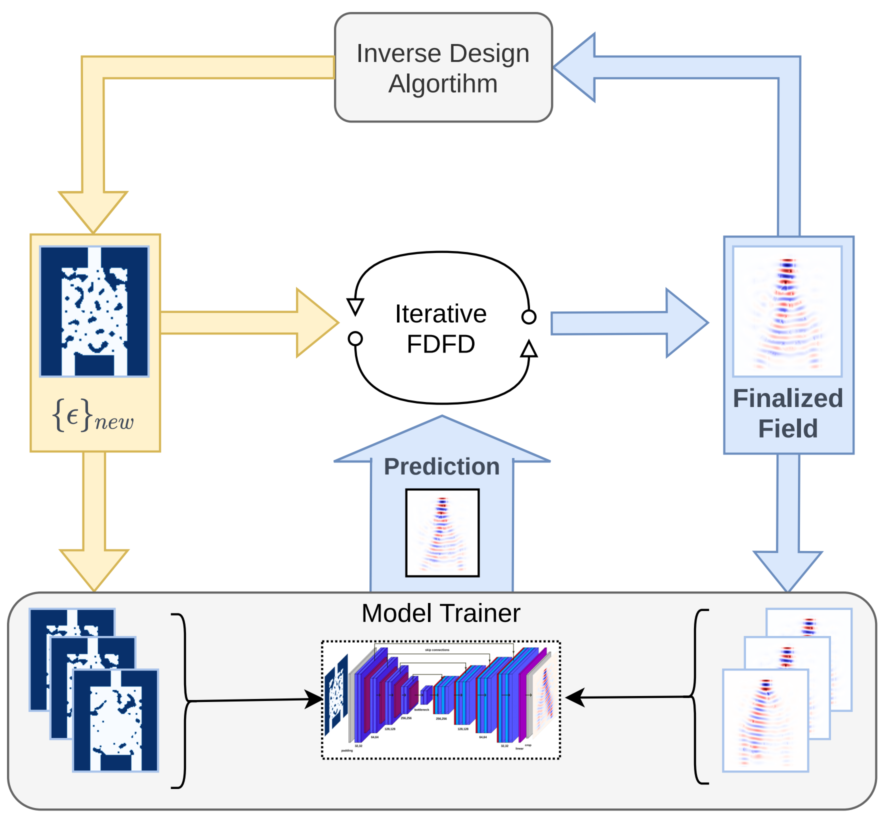

Fig. 1. Flowchart describing the interaction of our deep-learning model (bottom) with arbitrary iterative inverse design workflows based on FDFD simulations.

Inverse design algorithms usually perform a high number of simulations for similar structures that only differ slightly, e.g in a small number of pixels. We here use the structures together with the associated electromagnetic fields to train a deep learning model. The model infers a guess for the electromagnetic field associated with a previously unseen structure based on prior simulations. The dataflow is schematically depicted in Figure 1. We extend the general procedure of iterative inverse design algorithms by the "model trainer" shown in the bottom part of Figure 1. For any new structure that the algorithm produces we query the model for an initial guess serving as a starting point for the FDFD-procedure. The resulting field is evaluated by the inverse design algorithm, as described above, and additionally, paired with the associated structure information, fed into a database of the training- and prediction-module. The module asynchronously trains the model on the acquired data pairs. Thus the prediction gets increasingly more accurate as the optimization proceeds.

3. Deep learning model

Our approach is based on mapping the three dimensional permittivity distribution that represents the simulation cell to a corresponding electric field distribution, consisting of three real valued electric field components and three imaginary field components describing the phase per Yee-grid element. We adapt the V-Net architecture [19], a three dimensional variant of the U-Net [17] commonly employed in image segmentation and reconstruction, to fulfill the special requirements of our application in nanophotonics. The U-Net was shown to be highly successful in handling complex higher dimensional data including irregular shapes and pixel-discrete features [20, 21] while retaining spatial information. This makes it a promising candidate for learning the complex relations between spatial distributions of dielectric material in inverse designed nanophotonic devices and the associated electromagnetic fields.

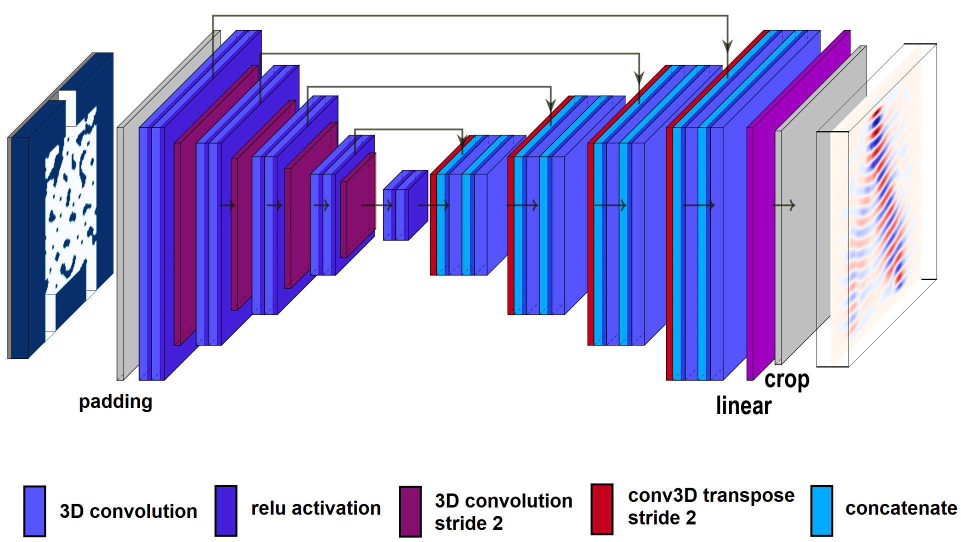

Fig. 2. U-Net-based Neural Network architecture. The permittivity distribution of the inverse designed device including the surrounding waveguide geometry is progressively compressed by the encoder, thus enforcing increased feature abstraction. The highest compression is reached in the network's bottleneck section. Subsequently, the decoder decompresses the feature maps until the spatial dimensions of the original simulation cell are met again. Skip connections between encoder and decoder blocks ascertain information flow.

Our adaptation of the U-Net architecture is depicted in Figure 2. It preserves the fundamental structure of the original U-Net while incorporating modifications to consider three dimensional complex-valued data. Accommodating for three independent complex channels on both, the input and output data, enables us to map materials with a uniform and non-uniform refractive index with arbitrary loss-coefficients to complex valued three dimensional electric fields where the imaginary part corresponds to the phase, which is needed to seed FDFD-algorithms. We apply symmetric zero-padding in each dimension of the input layer, such that the activation map size can be decreased in the *encoder* for at least four times by a factor of two, enabling variable input sizes. This downsampling operation is implemented by applying stridden $2 \times 2 \times 2$ convolutions to connect the decoder blocks, which has been shown to be superior to pooling operations [22]. The four blocks employ three 3D-convolutional layers with a kernel-size of $3 \times 3 \times 3$ and 2^{l+5} output channels with l being the index of the encoder block starting at $l = 0$. After each layer we apply a ReLU-activation function [23]. The output of the second layer in each block is connected with its counterpart in the decoder section using a skip-connection. After passing through the encoder section, the data is processed in the *bottleneck*. It implements two 3D-convolutional layers with 512 channels, and ReLU-activations. The output of the bottleneck serves as input to the *decoder* section. The decoder section implements the encoder blocks in the reverse order where the blocks are connected using transposed convolution operations to provide additional learnable parameters to reconstruct the input-shape. The skip connections originating from the second layer of each encoder block are connected to the input of the second layer of each decoder block and are thus concatenated with the output of the transposed convolutional layers. The output of the last decoder block is finally processed by a $1 \times 1 \times 1$ convolution with six output channels corresponding to the real and imaginary components of the electric field at each grid point of the simulation cell. We do not apply any final activation function to retain the full dynamic range of the field components. Finally, we crop the output to yield field-data of the same shape as the input before we apply the zero-padding. For the loss function we use the mean-squared-error in the training procedure to minimize the difference between the predicted and simulated electric field. We further employ the Adam optimizer, a stochastic gradient descent method based on adaptive estimation of first-order and second-order moments, which is well suited for large-scale data sets and non-convex optimization problems [24].

4. Application to the Inverse Design of an Asymmetric Powersplitter

In the following, we will apply the above method to structures that were found in the optimization process of an asymmetric powersplitter as an exemplary device. Being substantial components of photonic integrated circuits, the efficiency of these elements has great impact on the system's overall performance, for example in optical communication [25,26], optical phased arrays [27,28] or signal processing in complex chip layouts [29]. The optimization was conducted using an autonomously learning agent [30] where a splitting ratio of 90 by 10 for a wavelength of $\lambda = 775$ nm was cosen as the objective. We here consider the 100 nm tantalum-pentoxide-on-insulator platform, which has attractive properties for nonlinear and quantum photonics [2,31,32]. The simulation extends over 120 by 100 by 50 grid cells of 40 nm side lengths each, while including a perfectly matched layer of 10 pixels in each direction. Our model was trained on $22'750$ samples, while the validation- and test-sets consisted of $3'750$ and 750 samples, respectively, featuring discrete, i.e. fabrication-ready, permittivity distributions.

In Figure 3 (a) we show an exemplary device, which is part of our test-set. We observe a very high agreement of the simulated and predicted electric field, of which we show the dominant field component for a fundamental transverse electric input mode in Figure 3 (b) and Figure 3 (c), respectively. To illustrate the differences between the simulated and predicted field, we show the deviation for each pixel in Figure 3 (d). The largest variations arise when a high intensities occur at complex geometrical features, which can be observed close to the input waveguide at the top of the design area in the depicted example. We conducted the training for

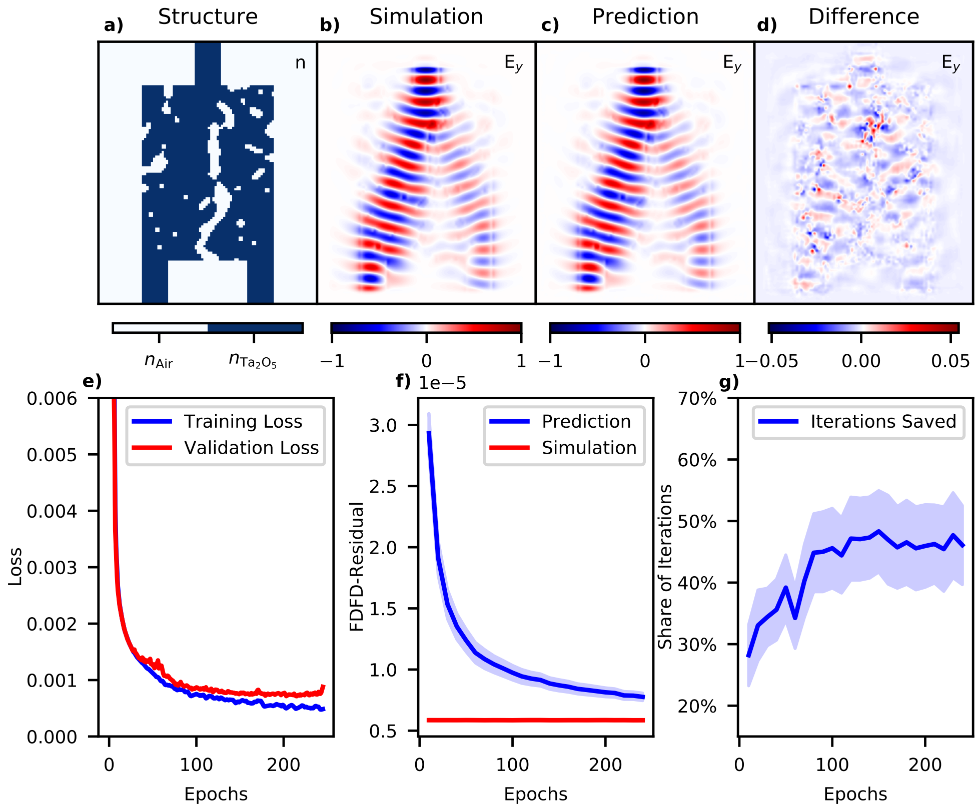

Fig. 3. **(a)** Structure of an exemplary asymmetric powersplitter in the test-data set. The presence of dielectric material and air are depicted in blue and white, respectively. **(b)** The simulated electric field. **(c)** The electric field as predicted by the model. **(d)** The (re-scaled) difference of the simulated and predicted field. All field plots show the dominant E_y-component, which is oriented perpendicular to the direction of propagation. **(e)** The loss-functions for the training- and validation-set. **(f)** The evolution of the FDFD-residuals for the predicted and simulated field during training as calculated for the test-data set. **(g)** The share of FDFD-iterations that were saved on the test-set using the predicted field as a starting point compared to initialization with a zero-vector.

245 epochs, where we found the best performance on the validation-set at epoch 215 based of the smallest mean-square-error deviation between the predicted and the simulated fields, which is depicted in Figure 3 (e). We observe a monotonic decrease of the FDFD-residual calculated every ten epochs for the test-set, which is shown in Figure 3 (f). Until a certain residual is reached, this directly translates to savings in the number of iterations required to finalize the simulation when initializing the FDFD-solver with the predicted field, as shown in Figure 3 (g). We save an average of 47.3 % of the FDFD-iterations in the optimal case, which directly translates into time- and resource savings. However, we note that after a certain number of epochs, the FDFD-residual, surprisingly, does not directly correlate with additional savings in the number of required FDFD-iterations. We attribute this behavior to the non-monotonic characteristics of conjugate-gradient based solvers, despite the BiCGSTAB algorithm employed in this work offering a good trade-off between overall convergence speed and stability [16].

5. Application to Other Device Layouts and Retraining

We further apply our model, which has previously been trained for predicting fields for the asymmetric powersplitter, to other inverse designed structures to demonstrate the versatility of the method. Corresponding exemplary structures found while optimizing a symmetric powersplitter and a fundamental to TE_{20} modeconverter for the 100 nm Ta_2O_5 platform, are shown in Figure 4 (a) and (b), respectively. While the fields predicted by the model trained on the asymmetric powersplitter data already yielded improvements in the number of FDFD-iterations required to finish the simulation when applied directly, i.e. without retraining, a short retraining using a reduced number of structures encountered in the new optimization procedure, leads to a significantly stronger improvement, as shown in Figure 4 (c). The retraining was conducted on a training-set of 2048 and 2560 new samples that were generated during the optimization procedure of the symmetric powersplitter and the modeconverter, respectively. For the symmetric powersplitter, we see an improvement from 39.4 % ±9.9 % to 53.0 % ±15.8 % and for the modeconverter we observe an increase in the share of saved FDFD-iterations from 27.8 % ±9.6 % to 48.4 % ±14.1 %.

Figure 4 (d) and (e) show the simulated, predicted and difference fields before and after retraining for the symmetric powersplitter and modeconverter, respectively. Prior to being retrained, the model is already capable of predicting fields that resemble the simulated fields reasonably well to guarantee appreciable resource savings during FDFD-simulations. Although it has never been exposed to training data featuring surrounding waveguide geometries, such as the wider centralized output waveguide in the modeconverter structures show in Figure 4 (b), the model is able to generalize and correctly predicts a propagating wave in the output waveguide where the majority of light is propagating in the desired TE_{20} target mode. However, one can observe a bias resulting from the asymmetric training data only consisting of fields where the majority of the power is routed to the left output port in the predicted fields for the symmetric powersplitter. After retraining the model, the predictions do no longer exhibit the aforementioned bias and the overall magnitude of the difference field is significantly reduced, thus allowing for additional resource savings.

6. Conclusion

We introduced a deep learning-based method to predict the electromagnetic field response for inverse designed nanophotonic structures exposed to a specific input mode. We adapted the U-Net architecture and proved our model's ability to associate sub-wavelength structures with the corresponding electromagnetic fields and deduce knowledge about unseen structures. The method is applicable to any iterative inverse design algorithm based on finite-difference frequency-domain simulations. For the example of an asymmetric powersplitter we find that training on a corresponding data set reduces the time expenditure of the simulation routines by 47.3 % on average. The model is able to generalize and deliver accurate predictions for structures with qualitatively different shapes, waveguide geometries and functionalities, such as the symmetric powersplitter and the modeconverter, shown in Figure 4 (a) and (b), respectively. When using the original model to predict initial field vectors seeding the FDFD-procedure we were able to save 39.4 % and 27, 8 % of the FDFD-iterations, respectively, which directly translates into runtime and resource savings. A short retraining on a limited number of directly related samples significantly improved the resource savings to an average of 53.0 % and 48.4 % for the symmetric powersplitter and modeconverter, respectively. We note that these resource savings improve the previous result for the asymmetric powersplitter, which we attributed to the increased diversity within the training data set, thus enhancing the model's ability to generalize.

It is straightforward to apply our architecture to larger simulation cells, where the computational effort increases orders of magnitude faster than the effort for model inference. We hence expect

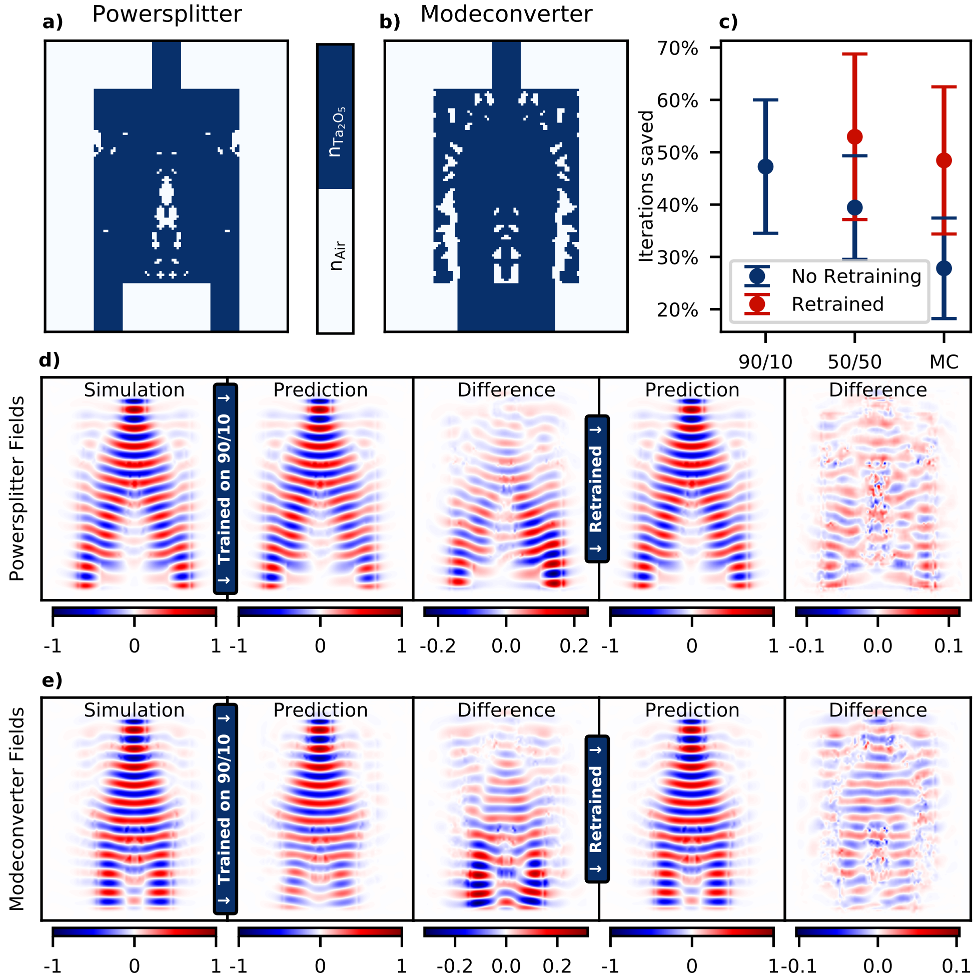

Fig. 4. **(a)** Structure of a symmetric powersplitter. **(b)** Structure of a fundamental to TE_{20} modeconverter. **(c)** FDFD-iterations saved for the original asymmetric powersplitter ("90/10"), the symmetric powersplitter ("50/50") and the modeconverter ("MC"). **(d)** and **(e)** dominant field components for the symmetric powersplitter and modeconverter, respectively. From left to right: The simulated field, the predicted field before retraining, the difference between the predicted field and the simulated field before retraining (rescaled), the predicted field after retraining, the difference between the predicted field and the simulated field after retraining (rescaled).

that the effectiveness of the proposed deep-learning extension improves further when applied to larger problems. We also expect a higher sample-efficiency as well as improved quality of the predictions when the model is trained on device-field combinations, additionally featuring continuous rather than just discrete permittivity distributions. These structures, although not representing realizable devices, are often encountered in established inverse design routines [9] and provide additional insights, for example related to wavelength-permittivity relations. Moreover, analytical inverse design methods often rely on the calculation of adjoint fields, which the model could also be trained on.

Furthermore, we see room for improving specific aspects of our model architecture. Apart from additional hyperparameter tuning concerning parameters such as the number of layers and filters in the encoder-, decoder- and bottleneck-block, the usage of regularization techniques, such as dropout [33], might increase the quality of the predicted fields. Using dropout also facilitates the possibility to query the model multiple times using the Monte-Carlo dropout scheme [34]. We consider the analysis of the variance of the predicted fields under consideration of certain dropout-connections a promising approach for obtaining a prediction certainty metric, which can be used to decide if a first-order FDFD-simulation following the prediction may be omitted entirely.

Reverted queries, inverting the direction of the data flow, may hold further potential for exploiting the model's ability for generalization. Specifying a target field and inferring a matching structure could, for example, be applied to models trained on powersplitters in order to find devices exhibiting different splitting ratios. Resulting structures may then provide a starting point for further optimization using established inverse design algorithms. The model capabilities for predicting fields for waveguide geometries that significantly differ from the geometries it has been trained on, indicates the possibility to construct a universal backbone for field predictions in arbitrary environments. The accuracy can then again be significantly increased by retraining such models using very small data sets featuring related geometries, underlying the potential of our method for complex inverse design problems.

Funding.

Deutsche Forschungsgemeinschaft (CRC-1459 (C05)); Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen (421-8.03.03.02–130428).

Acknowledgments.

We would like to thank the Münster Nanofabrication Facility (MNF) for their support in nanofabrication related matters. C.S. acknowledges support from the Ministry for Culture and Science of North Rhine-Westphalia (421-8.03.03.02–130428). The authors acknowledge support by the German Research Foundation (DFG, CRC 1459).

Disclosures.

The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," Nat. Photonics **12**, 659–670 (2018).
2. G. Moody, V. J. Sorger, D. J. Blumenthal, P. W. Juodawlkis, W. Loh, C. Sorace-Agaskar, A. E. Jones, K. C. Balram, J. C. F. Matthews, A. Laing, M. Davanco, L. Chang, J. E. Bowers, N. Quack, C. Galland, I. Aharonovich, M. A. Wolff, C. Schuck, N. Sinclair, M. Lončar, T. Komljenovic, D. Weld, S. Mookherjea, S. Buckley, M. Radulaski, S. Reitzenstein, B. Pingault, B. Machielse, D. Mukhopadhyay, A. Akimov, A. Zheltikov, G. S. Agarwal, K. Srinivasan, J. Lu, H. X. Tang, W. Jiang, T. P. McKenna, A. H. Safavi-Naeini, S. Steinhauer, A. W. Elshaari, V. Zwiller, P. S. Davids, N. Martinez, M. Gehl, J. Chiaverini, K. K. Mehta, J. Romero, N. B. Lingaraju, A. M. Weiner, D. Peace, R. Cernansky, M. Lobino, E. Diamanti, L. T. Vidarte, and R. M. Camacho, "Roadmap on integrated quantum photonics," J. Physics: Photonics **4**, 012501 (2022).
3. J. Lu and J. Vučković, "Objective-first design of high-efficiency, small-footprint couplers between arbitrary nanophotonic waveguide modes," Opt. Express **20**, 7221–7236 (2012).
4. A. Michaels and E. Yablonovitch, "Inverse design of near unity efficiency perfectly vertical grating couplers," Opt. Express **26**, 4766–4779 (2018).
5. A. Y. Piggott, J. Lu, T. M. Babinec, K. G. Lagoudakis, J. Petykiewicz, and J. Vučković, "Inverse design and implementation of a wavelength demultiplexing grating coupler," Sci. Reports **4**, 7210 (2014).
6. L. Su, A. Y. Piggott, N. V. Sapra, J. Petykiewicz, and J. Vuckovic, "Inverse design and demonstration of a compact on-chip narrowband three-channel wavelength demultiplexer," Acs Photonics **5**, 301–305 (2018).
7. C. Shang, J. Yang, A. M. Hammond, Z. Chen, M. Chen, Z. Lin, S. G. Johnson, and C. Wang, "Inverse-designed lithium niobate nanophotonics," ACS Photonics **10**, 1019–1026 (2023).

8. P. R. Wiecha, A. Arbouet, C. Girard, and O. L. Muskens, "Deep learning in nano-photonics: inverse design and beyond," Photonics Res. **9**, B182–B200 (2021).
9. M. Butz, A. S. Abazi, R. Ross, B. Risse, and C. Schuck, "Inverse design of nanophotonic devices using dynamic binarization," Opt. Express **31**, 15747–15756 (2023).
10. R. Trivedi, L. Su, J. Lu, M. F. Schubert, and J. Vuckovic, "Data-driven acceleration of photonic simulations," Sci. Reports **9**, 19728 (2019).
11. H. M. Yao and L. Jiang, "Machine-learning-based pml for the fdtd method," IEEE Antennas Wirel. Propag. Lett. **18**, 192–196 (2018).
12. H. M. Yao and L. Jiang, "Enhanced pml based on the long short term memory network for the fdtd method," IEEE Access **8**, 21028–21035 (2020).
13. S. Qi and C. D. Sarris, "Numerical dispersion compensation for fdtd via deep learning," in *2022 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (AP-S/URSI),* (IEEE, 2022), pp. 671–672.
14. W. Kim and J. Seok, "Simulation acceleration for transmittance of electromagnetic waves in 2d slit arrays using deep learning," Sci. reports **10**, 10535 (2020).
15. M. Chen, R. Lupoiu, C. Mao, D.-H. Huang, J. Jiang, P. Lalanne, and J. A. Fan, "High speed simulation and freeform optimization of nanophotonic devices with physics-augmented deep learning," ACS Photonics **9**, 3110–3123 (2022).
16. G. Sleijpen, H. Van der Vorst, and D. Fokkema, "BiCGstab(l) and other hybrid bi-CG methods," Numer. Algorithms **7**, 75–109 (1994).
17. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," (2015). ArXiv:1505.04597.
18. Kane Yee, "Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media," IEEE Trans. on Antennas Propag. **14**, 302–307 (1966).
19. F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," (2016). ArXiv:1606.04797.
20. D. Peng, Y. Zhang, and Guan, "End-to-end change detection for high resolution satellite images using improved UNet++," Remote. Sens. **11**, 1382 (2019).
21. H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, and J. Wu, "UNet 3+: A full-scale connected UNet for medical image segmentation," (2020). ArXiv:2004.08790.
22. J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, "Striving for simplicity: The all convolutional net," (2015). ArXiv:1412.6806.
23. A. F. Agarap, "Deep learning using rectified linear units (ReLU)," (2019). ArXiv:1803.08375.
24. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," (2017). ArXiv:1412.6980.
25. M. H. Tahersima, K. Kojima, T. Koike-Akino, D. Jha, B. Wang, C. Lin, and K. Parsons, "Deep neural network inverse design of integrated photonic power splitters," Sci. reports **9**, 1368 (2019).
26. A. Y. Piggott, J. Petykiewicz, L. Su, and J. Vučković, "Fabrication-constrained nanophotonic inverse design," Sci. reports **7**, 1786 (2017).
27. M. J. Heck, "Highly integrated optical phased arrays: photonic integrated circuits for optical beam shaping and beam steering," Nanophotonics **6**, 93–107 (2017).
28. S. A. Miller, Y.-C. Chang, C. T. Phare, M. C. Shin, M. Zadka, S. P. Roberts, B. Stern, X. Ji, A. Mohanty, O. A. J. Gordillo *et al.*, "Large-scale optical phased array using a low-power multi-pass silicon photonic platform," Optica **7**, 3–6 (2020).
29. M. Teng, A. Honardoost, Y. Alahmadi, S. S. Polkoo, K. Kojima, H. Wen, C. K. Renshaw, P. LiKamWa, G. Li, S. Fathpour *et al.*, "Miniaturized silicon photonics devices for integrated optical signal processors," J. Light. Technol. **38**, 6–17 (2019).
30. M. Butz, A. Leifhelm, M. Becker, B. Risse, and C. Schuck, "A universal approach to nanophotonic inverse design through reinforcement learning," in *CLEO: Science and Innovations,* (Optica Publishing Group, 2023), pp. STh4G–3.
31. L. Splitthoff, M. Wolff, T. Grottke, and C. Schuck, "Tantalum pentoxide nanophotonic circuits for integrated quantum technology," Opt. Express **28**, 11921–11932 (2020).
32. R. Bankwitz, M. Wolff, A. Abazi, P.-M. Piel, L. Jin, W. Pernice, U. Wurstbauer, and C. Schuck, "High-quality factor Ta2O5-on-insulator resonators with ultimate thermal stability," Opt. Open (2023). https://doi.org/10.1364/opticaopen.23619024.v1.
33. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," J. Mach. Learn. Res. **15**, 1929–1958 (2014).
34. R. Camarasa, D. Bos, J. Hendrikse, P. Nederkoorn, E. Kooi, A. Van Der Lugt, and M. De Bruijne, "Quantitative comparison of monte-carlo dropout uncertainty measures for multi-class segmentation," in *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging, and Graphs in Biomedical Image Analysis: Second International Workshop, UNSURE 2020, and Third International Workshop, GRAIL 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 8, 2020, Proceedings 2,* (Springer, 2020), pp. 32–41.