



**HAL**  
open science

# Inverse Design of Unitary Transmission Matrices in Silicon Photonic Coupled Waveguide Arrays using a Neural Adjoint Model

Thomas W Radford, Peter Wiecha, Alberto Politi, Ioannis Zeimpekis, Otto L Muskens

► **To cite this version:**

Thomas W Radford, Peter Wiecha, Alberto Politi, Ioannis Zeimpekis, Otto L Muskens. Inverse Design of Unitary Transmission Matrices in Silicon Photonic Coupled Waveguide Arrays using a Neural Adjoint Model. 2024. hal-04713884

**HAL Id: hal-04713884**

**<https://laas.hal.science/hal-04713884v1>**

Preprint submitted on 30 Sep 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Inverse Design of Unitary Transmission Matrices in Silicon Photonic Coupled Waveguide Arrays using a Neural Adjoint Model

Thomas W. Radford<sup>1\*</sup>, Peter R. Wiecha<sup>2</sup>, Alberto Politi<sup>1</sup>, Ioannis Zeimpekis<sup>3,4</sup>, and Otto L. Muskens<sup>1</sup>

<sup>1</sup> School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, United Kingdom

<sup>2</sup> LAAS, Université de Toulouse, CNRS, Toulouse, France

<sup>3</sup> School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, United Kingdom

<sup>4</sup> Optoelectronics Research Centre, University of Southampton, Southampton, SO17 1BJ, United Kingdom

\*T.Radford@soton.ac.uk

**Abstract:** The development of low-loss reconfigurable integrated optical devices enables further research into technologies including photonic signal processing, analogue quantum computing, and optical neural networks. Here, we introduce digital patterning of coupled waveguide arrays as a platform capable of implementing unitary matrix operations. Determining the required device geometry for a specific optical output is computationally challenging and requires a robust and versatile inverse design protocol. In this work we present an approach using high speed neural network surrogate based gradient optimization, capable of predicting patterns of refractive index perturbations based on switching of the ultra-low loss chalcogenide phase change material, antimony tri-selenide ( $\text{Sb}_2\text{Se}_3$ ). Results for a  $3 \times 3$  silicon waveguide array are presented, demonstrating control of both amplitude and phase for each transmission matrix element. Network performance is studied using neural network optimization tools such as dataset augmentation and supplementation with random noise, resulting in an average fidelity of 0.94 for unitary matrix targets. Our results show that coupled waveguide arrays with perturbation patterns offer new routes for achieving programmable integrated photonics with a reduced footprint compared to conventional interferometer-mesh technology.

© 2024 The Author(s)

## 1. Introduction

Integrated photonics offers a platform for the miniaturization of optical devices and systems, yielding increased stability, greatly reduced size and complexity compared to traditional optical systems. Integrated photonic technologies have been successfully demonstrated across a wide range of domains including vector-matrix multiplication [1,2], quantum simulation [3,4], signal processing [5] and the development of high speed neural networks [6]. While application-specific photonic integrated circuits require a bespoke design-fabrication cycle for each variation in functionality, there is an emerging interest in platforms that can be re-programmed after fabrication to provide fine-tuning, diversification, or entirely new methods of deployment [7]. Re-programmable devices may also have the ability to dynamically tailor their optical output [8], and use cases were realized across a number of fields including post fabrication device processing [9], optical switches [10] and optical signal compensation [11]. One field in particular looking to utilize reconfigurable technologies is photonic computing [6, 12–14] where the development of next generation optical devices rely on the construction of structures with rapidly configurable output characteristics allowing for high speed information processing.

Commonly used reconfigurable technologies such as micro-heaters [15, 16] and micro electro-mechanical devices (MEMS) [17–19] either introduce additional complex fabrication steps or rely on the coupling of external regulating electronics onto the photonic chip. Such approaches occupy valuable chip space and impose power constraints, while simultaneously introducing undesirable heat and noise into experimental systems. An alternative method for the production of reconfigurable devices involves using non-volatile optical phase change materials (PCMs) such as antimony tri-selenide ( $\text{Sb}_2\text{Se}_3$ ) [20]. PCM based approaches can offer ultra-low loss, compact and passive programming capabilities, while maintaining compatibility with a wide range of optical devices across telecommunications wavelengths [21–25]. PCMs such as  $\text{Sb}_2\text{Se}_3$  are able to exist in multiple non-volatile solid states which can be programmed using optical or electrical pulses. The dielectric functions of amorphous and crys-

talline states of PCMs differ significantly from each other, which depending on the nature of the chemical bonds can result in large order changes in refractive index [26]. An approach using direct optical writing allows the addition of reconfigurability to previously fabricated devices using minimal overheads in fabrication and regulation [27]. Endurance in programmable devices using PCM technologies is seeing improvements [28] with repeated refractive index modulation of  $> 10^6$  switching cycles recently being demonstrated in thin-films of  $\text{Sb}_2\text{Se}_3$  [29,30].

The relationship between input and output modes of any linear optical element can be described by a complex transmission matrix. Through careful design of multi-port devices, we are therefore able to produce an analogue for any given  $n \times m$  matrix so long as a fabrication process can accommodate its production. Integrated approaches using continuously coupled devices offer an interesting platform whereby a device can be specifically designed to implement any arbitrary transmission matrix, confined within the bounds of a micron-scale structure [31]. Interference based devices with the addition of individual scattering sites are a common platform for the implementation of arbitrary optical transmission matrices [32–37], as well as devices consisting of reconfigurable interferometer meshes [8, 12, 13, 38].

In this work, we explore an alternative approach, based on coupled waveguide arrays (CWG) covered with a thin film of PCM programmed with a nano-scale perturbation pattern [39–45]. With focused laser writing, it is possible to reversibly modulate the local refractive index of individual pixels on the device surface with diffraction limited spatial resolution [23, 27]. This pattern facilitates modulation of the coupling coefficients between neighbouring waveguides, allowing full control over the devices transmission properties. As light propagates through a part of the device containing a switched pixel, it develops a phase shift relative to light traveling in an unswitched region due to the local effective index contrast. Using a large enough number of these small effects at each pixel, enables full control over both the waveguide coupling behavior, as well as the phase of the light at output ports across a full  $2\pi$  range.

In contrast to micro-heater or interferometer mesh based devices, the perturbation pixel patterns in our CWG platform are correlated to the device’s transmission matrix in a highly complex manner. A crucial step to develop this technology must therefore be the development of a powerful inverse design protocol, used to find a pixel pattern that implements a specific, complex transmission matrix. Significant research has been undertaken to develop a variety of inverse design processes. Modern techniques can be split into two overarching methods; inverse design through topology optimization [46] and the utilization of deep-neural networks. Topology optimizations are very powerful [47] and have found applications across a wide range of disciplines, from aerospace engineering [48, 49] and medical research [50, 51], to photonics [52–56]. Such optimizations typically result in complex-shaped structures with fine features which can prove difficult to fabricate, in some cases becoming unrealistic to produce at all [57]. Gradient based topology approaches often face practical limitations as gradient optimizations require smooth variables, but their design problem is typically categorized into a binary basis, where material either is present or absent. It should be noted that while there do exist non-gradient based optimization methods, many of these have been demonstrated as intractable for real world practical systems with large degrees of freedom [58, 59]. The most popular current alternative is the integration of artificial intelligence and machine learning into the device design pipeline, with increasing attention being paid to the creation of an AI-topological hybrid approach [60–62].

In this work, we demonstrate an inverse design tool chain using a gradient based optimization with a deep learning surrogate forward model [63], as well as a geometry re-parametrisation using a Wasserstein generative adversarial network (WGAN). Using this high speed technique we build on previous works in which similar methods are employed for the design and development of individual device elements, such as the height, orientation and shape of elliptic meta atom pillars [64, 65], as well as in more complex free-space design tasks [66].

## 2. Neural adjoint inverse design approach

### 2.1. Device Model

To demonstrate the capabilities of our inverse design method, we consider the case of an array of three coupled silicon rib waveguides, implementing an arbitrary  $3 \times 3$  transmission matrix. In our simulations, these waveguides are covered by crystalline  $\text{Sb}_2\text{Se}_3$ , which may then be switched locally, for example experimentally by using direct laser writing [27]. Experimentally it is easier to produce a narrow amorphization laser profile, meaning greater spatial precision can be achieved by using pixel maps with a crystalline background. We consider thin films of PCM, which have been shown experimentally as a feasible geometry [27], allowing a consistent and repeatable index modulation to be applied.

Our coupled waveguide device model is based on a standard silicon on insulator wafer structure, a schematic of the device is shown in Figure 1a. Structural features are etched 120 nm deep into a 200 nm thick silicon overlayer, which lies on top of a 2 $\mu\text{m}$  thick buried oxide layer. The rib waveguides are 500 nm wide, allowing only propagation of the fundamental mode at the considered vacuum wavelength of 1550 nm. The distance between waveguides is 250 nm over a 50 $\mu\text{m}$  long coupled region. To avoid any cross-talk outside of this coupling region,

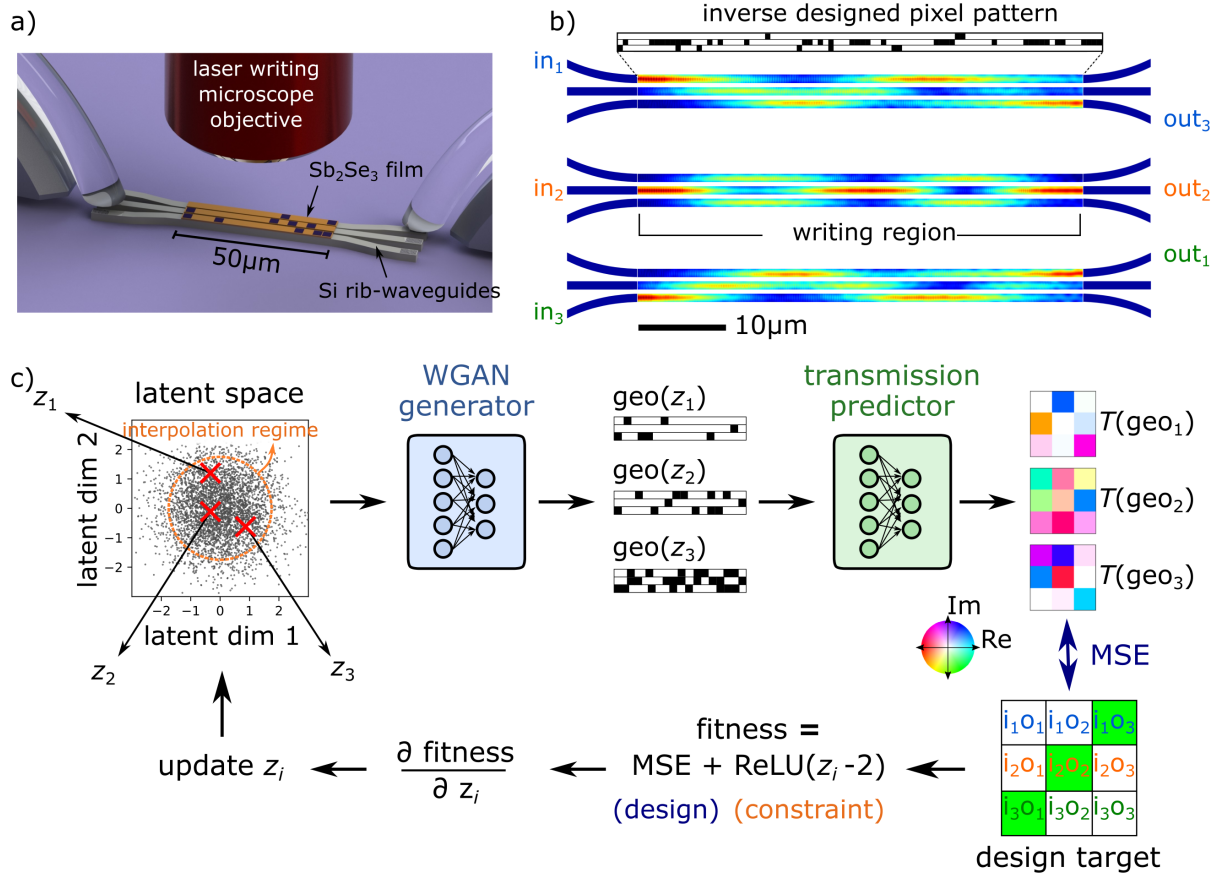


Fig. 1. a) Artistic view of the reconfigurable coupled waveguide device. An array of coupled silicon rib waveguides (here three) is covered by a thin PCM layer (here  $\text{Sb}_2\text{Se}_3$ ), which can be locally switched between crystalline and amorphous state using a laser writing setup. This is parametrized in a pixel pattern, in this figure  $500\text{ nm}$  large pixels along  $50\mu\text{m}$  of the waveguides. The input and output ports of the waveguide array can be individually accessed by optical gratings. b) varFDTD simulations of the electric field intensity maps along the coupled waveguide array for three different input ports. The pattern is designed to correspond to an anti-diagonal exchange matrix with equal phases for each output. c) Full schematic of the inverse design pipeline, used to find the optimal pattern shown in b).

waveguides fan out to a spacing of  $1\mu\text{m}$  at both input and output sides. The waveguides are covered with a  $30\text{ nm}$  thick  $\text{Sb}_2\text{Se}_3$  layer to maintain their single mode performance, displaying an index contrast between crystalline and amorphous states as high as  $\Delta n_{\text{Sb}_2\text{Se}_3} = 0.77$ . Finally, the structures are capped with a semi-infinite cladding layer of  $\text{SiO}_2$ . Pixels are defined on a  $500\text{ nm}$  square grid over the coupling region's area, consisting of a total of 96 pixels for each of the three waveguides (288 pixels in total).

Coupled waveguide devices are simulated using the finite difference time domain (FDTD) software Lumerical, specifically we use the Variational FDTD (varFDTD) engine, MODE. This engine collapses a 3D device structure into an effective 2D simulation by probing the effective index experienced in slices across the simulation mesh. The resultant simulation achieves comparable accuracy to a full 3D simulation, however requires significantly less computational resources. As this method accounts for the 3D structure of the device model it is commonly referred to as a 2.5D simulation. Examples are shown throughout this work, for example in figure 1b, showing the electric field intensity of a device with a transmission matrix analogous to the inverse identity matrix. After simulation, the device's transmission matrix can be extracted by sweeping across input waveguides, for all input modes the complex field at the peak amplitude of each output is recorded and constructed into a complex matrix.

## 2.2. Inverse design goals and training data generation

The choice of training data is a particular important aspect of the model under study. Similar to previous studies [36], we use pre-conditioned data by specifying a target output profile for a single input port. It is not constructive to use a training set of purely random patterns [36], because random pixel patterns typically do not co-operate in providing an effective state conversion between input and output, as discussed in more detail in Section 3.2. For an acceptable performance of the deep learning based design method, a more directed approach to training data generation is required, one that from the start aims towards the later inverse design task.

To this end, we first implement a brute force iterative optimization algorithm in the training data generation. The process starts by randomly choosing a specific input channel, as well as a random splitting ratio for the output channels. Then, the state of a single, random pixel is switched and the device transmission matrix is evaluated through varFDTD simulations. If, compared to the previous device geometry, the result is closer to the target splitting ratio, the pixel is retained and further simulations are run for each input channel in order to record the full transmission matrix. If not, the pixel is reverted to its initial state, another random pixel is selected, and the process is repeated. After a specified number of iterative cycles the process starts over with a blank device and a new optimization target. At each evaluation step where a perturbation is kept, the transmission of the other input ports is also evaluated and the entire pattern and its associated complex transmission matrix are saved to the dataset. In this way the dataset is populated with a large variety of pixel numbers and sufficient variance in the device performance to build a balanced dataset. A detailed description of this process as well as a specific example can be found in the supplementary information.

Training data generation was carried out in this manner to create an initial dataset containing 21,406 pixel patterns alongside their associated complex transmission matrices. The horizontal plane of symmetry found in the unperturbed waveguide array was furthermore exploited so as to double the size of this training set. In this step we flip both the pixel pattern and transmission matrix along the midpoint of the array, allowing us to extend the training data to contain 42,812 patterns without the need for extra simulation time.

Once a database of pre-optimized pixel patterns and their associated transmission matrices is created, we repeat this process to create a separate validation dataset used to verify that the network is sufficiently generalized and may accurately predict transmission matrices for patterns from outside of the dataset it was trained on. This validation set contains a further 3282 patterns which were in turn doubled using the same method as discussed above.

## 2.3. Gradient Based Inverse Design using a Deep Learning Surrogate

The forward predictor used is a ResNeXt encoder-decoder convolutional neural network [67], the training of which took 1h using a Nvidia RTX3070 GPU. A detailed schematic of this forward network model can be found in the supplemental information figure S1. After training, the forward network is capable of predicting the complex transmission matrix for pixel patterns similar to the training geometries. However, pixel patterns that differ substantially from the training set samples (e.g. many more pixels or strongly clustered pixels), are predicted at very low accuracy.

Our predictor model is trained only on the pre-simulated dataset and has no direct knowledge of the physics in the system. While such purely data-based surrogate models are known to work accurately in the interpolation regime, it is also well known that extrapolation to geometries that differ from those in the dataset is usually very error-prone [68, 69]. In consequence, optimizing directly the pixel patterns using the forward model as a differentiable physics predictor, fails in a remarkably reproducible manner. The gradient based optimization routinely converges towards geometries for which the forward model's predictions are wrong. In other words, optimising the pixel pattern using the neural network's design gradients, leads to geometries with unsatisfactory transmission matrices after varFDTD simulations, but, according to the forward network, should be well optimized.

Thus, a mechanism is necessary to constrain the allowed geometries during optimization to pixel patterns which the forward model is capable of faithfully predicting. Typically this is achieved through the introduction of a boundary loss term in the design optimization [64]. However, in free-form optimizations such a constraint is difficult to formulate.

### *Regularization of the Gradient Based Optimization*

To avoid the optimization converging to the extrapolation regime of the forward network, we use a further deep-learning based approach. Integrating a Wasserstein generative adversarial network with gradient penalty [70] (WGAN-GP), we develop a learned re-parametrization of the pixel patterns geometry from the training dataset. Details of the WGAN-GP architecture are shown in the supplemental information figure S2.

The key idea is, that during training, the WGAN-GP develops a mapping of the pixel pattern geometries into a compact and continuous latent space, in which the pattern geometries of the training data are normally distributed

around the mean value  $\mu_z = 0$  and with a variance of  $\sigma_z = 1$ . The latent vector representations of the geometries in the dataset being normally distributed in the latent space means, that it is possible to interpolate between two latent vectors. Furthermore, every intermediate point in the latent space also corresponds to a pixel pattern that is valid within the distribution of the geometries in the training set.

In other words, the pixel-based representations of the geometries that form a non-convex set in the original parametrization (by pixels; interpolation leads to non-physical patterns with gray-scale pixels), are mapped into a convex set of latent representations of the same geometries. Any interpolation between these latent vectors should still represent a valid perturbation pattern, which can be generated from its latent vector using the WGAN-GP generator network. This also means, that due to the training procedure using normally distributed random sampling, all geometries that are valid interpolations of the dataset samples, lie within said normal distribution (with known  $\mu_z = 0$ ,  $\sigma_z = 1$ ). Therefore, to constrain the inverse design optimization loop to the interpolation regime of the neural network, we have to reformulate the design problem to find an optimum latent vector  $z$  instead of an optimum pixel pattern. If we optimize in the WGAN-GP latent space, we can then simply add a regularization term to the inverse design fitness function (we use mean square error between target and predicted transmission matrix). Since the statistical properties of the WGAN-GP's latent space are known (assuming successful training of the latter), this regularization term has to penalize solutions with latent values far from the mean of the normal distribution. Practically, we add a simple rectified linear unit (ReLU) constraint term to the design fitness function  $\sum_i \text{ReLU}(z_i - 2)$ , that penalizes values of  $z_i$  outside of a  $2\sigma$  range (see figure 1c).

#### *Local Minima during Gradient Descent*

Gradient based optimizations such as used here are prone to become stuck in local minima, and therefore require good initial guesses to reach a globally optimum solution. To ensure that we find a close to ideal solution, we run a large number of concurrent optimizations for each design target (100 in our case). Eventually, the best ranked solution is kept. We note that due to the highly optimized GPU-based parallelization of modern deep learning frameworks such as tensorflow used in this work, the concurrent optimization of many targets is very efficient and fast.

#### *Inverse Design Loop*

A full schematic of our inverse design pipeline is shown in figure 1c. A set of several latent vectors is randomly initialized. The WGAN generator (blue) predicts the corresponding geometries, of which the complex transmission matrices are predicted by the forward network (green). The latter are compared to the design target transmission using the mean square error as a metric. The total fitness comprises an additional constraint term to restrict the optimization to a  $2\sigma$  region of the WGAN latent space, which should roughly correspond to the interpolation regime of the neural networks. The gradients of the fitness are calculated using tensorflow's automatic differentiation capability. The initial latent vectors are updated according to the fitness gradients and the process is repeated. After convergence, the best solution is kept as final design.

This process is able to faithfully create pixel patterns that result in distinctly different optical outputs, depending on which input mode is injected to the device. The varFDTD simulated result of the design target shown in figure 1c is depicted in figure 1b, showing the successful implementation of our example target, the anti-diagonal exchange matrix, with equal phase for light injected in each of the input ports. This method is also capable of addressing the phase of each matrix element, as will be demonstrated in the following sections. Transmission matrices are depicted as colorful sets of  $3 \times 3$  blocks, where the phase and amplitude of each complex matrix element are mapped respectively on the hue and saturation of an HSV color space [37].

### **3. Further optimization of design performance**

#### *3.1. Inverse Design Performance Metrics*

Inverse design performance is assessed using varFDTD simulation of the optimized patterns on a test-set of 1000 Haar random unitary transmission matrix targets. A useful metric to carry out this analysis is the amplitude fidelity. The fidelity compares both the target and resimulated complex transmission matrices, with a perfect match between the two yielding a fidelity of 1. We calculate the amplitude fidelity using the following formula taken from literature [71],  $F = 1/N [\text{Tr}(|U^*| \cdot |U_{\text{sim}}|)]$  where  $U^*$  represents the conjugate transpose of the target complex transmission matrix,  $U_{\text{sim}}$  is the re-simulated result, and  $N$  refers to the number of modes in the given system, for our device model,  $N = 3$ .

Amplitude fidelity as a metric is generally insensitive to the phase agreement between the two matrices under study, as demonstrated in figure S8. While many applications do not require exact phase recreation, or are able to compensate with the addition of phase shifters at input and output ports, in some cases accurate recreation of the correct output phases is crucial. In such applications, it is instead useful to analyze the mean squared error

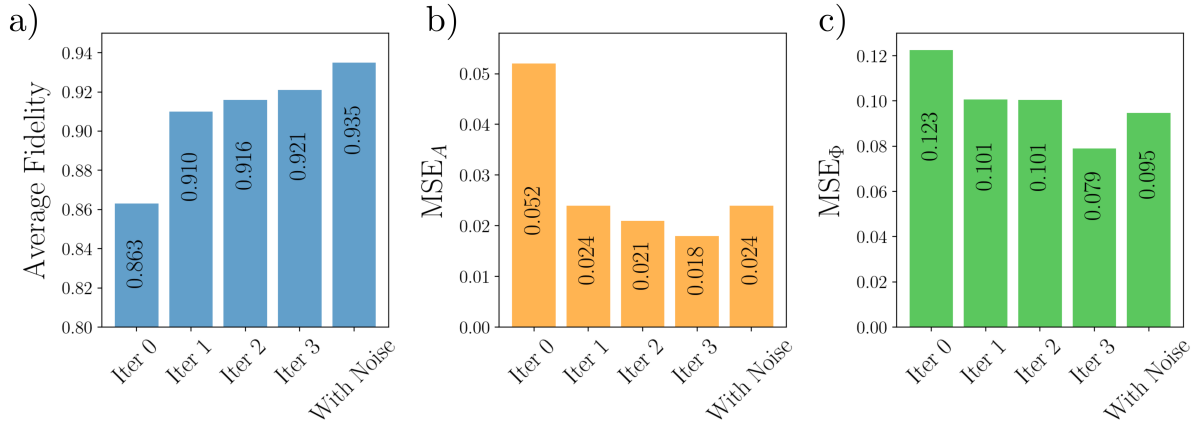


Fig. 2. Average fidelity and MSE values during the process of iterative improvement. Values presented are averages for resimulation of 1000 pixel patterns produced to implement random unitary targets. Each improvement loop contained 2000 patterns which are in turn doubled by y-symmetry.

(MSE) function between target and resimulated phases to assess network performance. The MSE values presented hereafter are calculated across all 9 matrix elements. The phase MSE was furthermore taken using the closest phase difference modulo  $2\pi$  and was normalized by dividing by a factor  $\pi^2$ . In our analysis we consider mainly the amplitude fidelity as a measure of performance for consistency with other works in the field, but will refer to the other metrics where these are most relevant.

### 3.2. Dataset optimization

After training both networks on the initial dataset (see above), the average amplitude fidelity achieved is around 0.86. While this means in general a good agreement between design target and optimized solution is reached, it also means that the inverse designed patterns still deviate distinctly from the expected results. We therefore invest in optimization of the training data before assessing the performance of the inverse design approach.

#### *Iterative Data Improvement*

As a first measure for dataset optimization, we iteratively extend the training dataset using the results from the inverse design process itself, to improve the design fidelity. Essentially, the idea is to let the neural network predictor learn from its own mistakes [36, 72]. Using the inverse design pipeline, we predict pixel patterns for 2000 random unitary targets. The transmission matrices of the designed patterns are simulated by varFDTD, and these results are appended to the initial training dataset. Subsequently, the models (both, forward predictor and WGAN-GP) are re-trained using the now extended dataset. We repeat this process three times, and benchmark the inverse design quality using 1000 Haar random unitary targets after each iteration. After three iterations, we observe a notable improvement in inverse design performance, with an average amplitude fidelity now of 0.92, as shown in figure 2. This performance is comparable to fidelities obtained by similarly available reconfigurable technologies [71, 73, 74].

#### *Expansion with Random Pixel Patterns*

The CWGs exhibit a well defined small number of modes and randomly generated patterns result in high device throughput, as opposed to for example MMIs where the number of modes inside the device is much higher and therefore high-throughput solutions are more sparse [36]. This property in principle allows us to rapidly add patterns to the training dataset without the need to optimize for device throughput. It may therefore seem appealing to simply append as many random patterns as possible to the training set. However, in our studies we have found a more methodical approach is necessary, as is shown in more detail in supplementary information figure S6 and simply adding more random data does not improve the network performance.

The reason that purely random datasets are found to be ineffective is attributed to the collective action of the pixel patterns. Within the large space of possible solutions, there are many that are not particularly effective in implementing a transmission matrix, as pixels do not sufficiently act in concert or even counteract each other. By including randomly generated data we modify the statistical distribution, which starts containing more and more of these "counterproductive" patterns. Part of the forward model's capacity is then used to fit these useless

noise cases which are not beneficial for the later inverse design task. The WGAN latent space on the other hand will become less efficient in restricting the optimization to useful geometries, since it is also trained on the same data now containing totally random patterns, which are therefore part of its latent space as well. Consequently it is foreseeable that the addition of too many random patterns will result in a decrease of overall inverse design accuracy, which is demonstrated in the supporting information figure S6 and S7.

Adding a small amount of totally random pixel patterns slightly improves the inverse design capacity, yielding decreased validation losses, as well as an increase in the average fidelity from 0.921 to 0.935, as further shown in figure 2 (cases “with noise”). This trend is consistent with other work where a small amount of randomness was found to improve performance [75]. We attribute this improvement to a better prediction of edge case geometries. As pattern optimization happens inside the WGAN latent space, the inverse design is restricted to the statistical distribution of the initial training data. Since the initial data has a bias towards low pixel numbers it may actually be beneficial to also add totally random pixel patterns in order to diversify the training set and make the forward network better deal with patterns at the edge of its latent space

### 3.3. Pattern characteristics

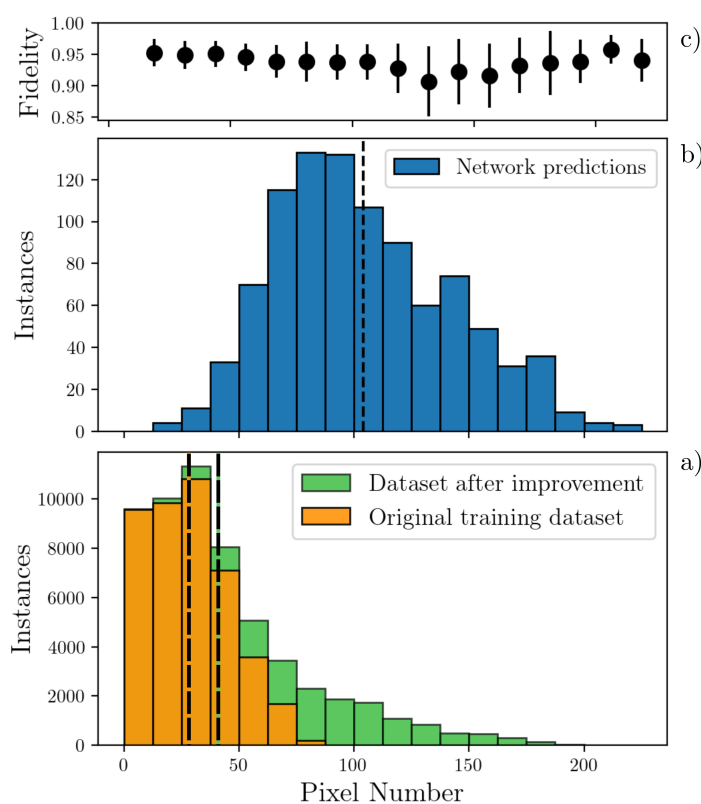


Fig. 3. (a) Distribution of pixel numbers contained in patterns from the initial ‘brute force’ training dataset (orange) and final training dataset after iterative improvement and 2.5% random patterns (green). (b) Pixel histogram of predicted patterns for 1000 random unitary targets. (c) Corresponding average fidelities for each bin of the predicted patterns, with error bars representing the standard deviation from the mean. Dashed lines: average values of 28 pixels for initial training data and 41 pixels for final training data.

The bottom panel in Figure 3 shows histograms of the distribution of active pixels in the initial datasets used for training (orange) and the augmented datasets after three iteration loops (green). The pixel histogram for patterns predicted by the network for a validation dataset of 1000 target transmission matrices are shown in the middle panel (blue). The corresponding average and standard deviation of the fidelity for the predicted patterns is presented in the top panel of the figure. As the training data is generated in an iterative manner and intermediate steps are stored, there is a strong emphasis in the training dataset for patterns contain a low number of pixels. Across the entire database used for training there is an average number of 41 pixels/pattern.

In comparison, the average number of pixels in predicted patterns is 104 and patterns with up to 225 active



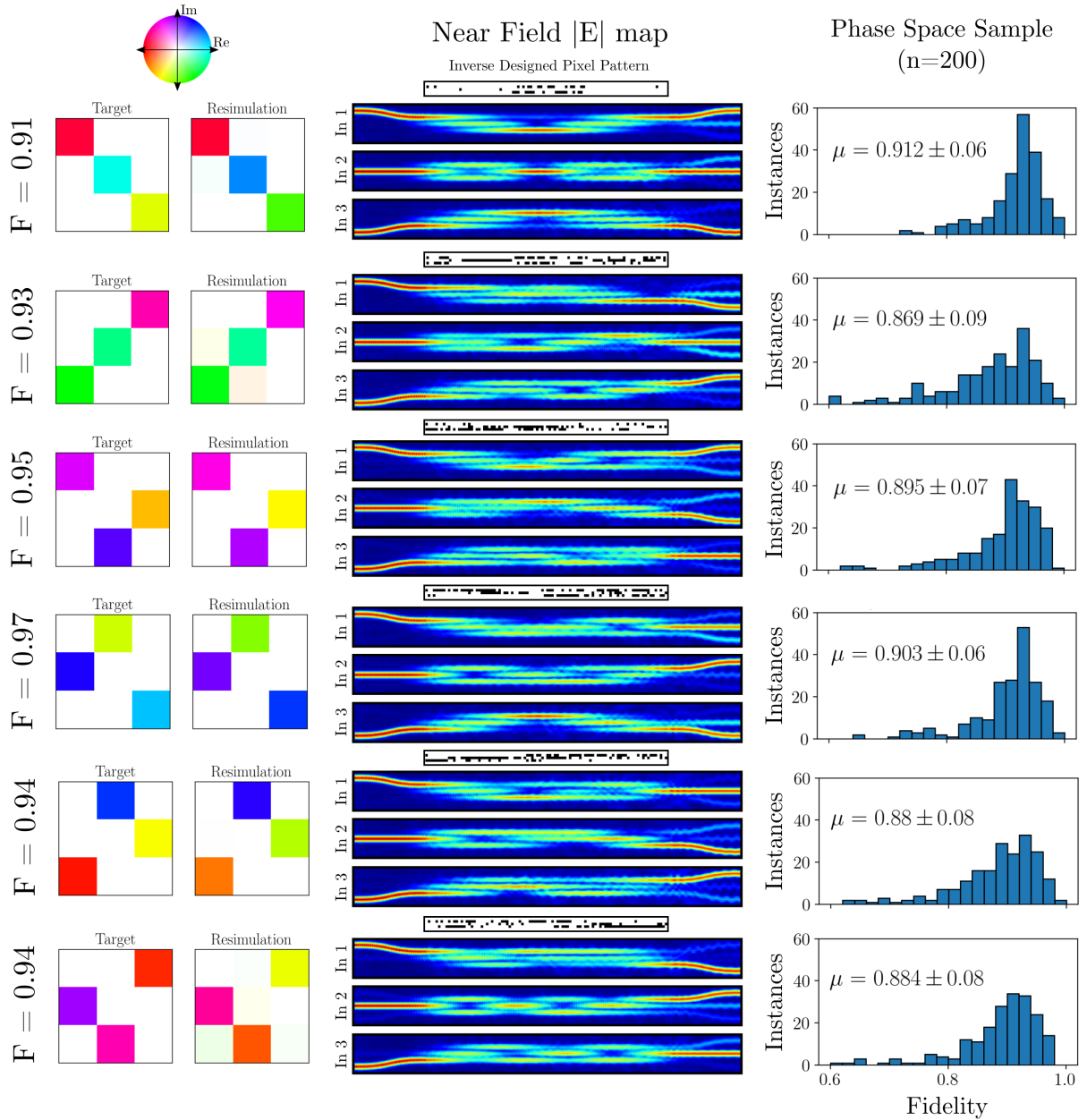


Fig. 4. Performance of predicted patterns for the implementation of all available permutation matrices of the  $n=3$  waveguide array. 200 random phase shifts are applied to each matrix element, allowing for in depth sampling of the available phase space to determine a maximum fidelity for phase insensitive applications. Average fidelities of around 0.9 are achieved for all targets, with peak fidelity values  $> 0.97$  for all targets.

pixels are found. Implementing the target transmission matrices requires predominantly geometries containing more pixels than the average number from the training data. We can observe that as the number of pixels increases above 50, the standard deviation of the fidelity increases quite significantly. An increase in network error rate will lead to a greater spread of results and thus a larger standard deviation. This trend is attributed to the forward network performing less accurately on geometries that are underrepresented in its training data, which could explain some of the errors in the inverse design. More discussion on the influence of forward network accuracy on fidelity will be given below in Section 4.2.

## 4. Results on Transmission Matrix Design

After the iterative dataset expansion, the addition of random patterns, and the training of the network on this final dataset, we proceed to carry out a benchmark of our approach for the inverse design of pixel patterns for a variety of coupled waveguide transmission matrices. These can be broadly split into two groups: permutation matrices, which are a class of orthogonal matrices where individual input and output ports are connected without port mixing, and unitary matrices, providing the most general input-output relationship. In the following we discuss these cases in more detail.

### 4.1. Permutation Matrices

The first class of matrices under study contains those which guide light in a one-to-one fashion, whereby light injection into each input waveguide results in transmission through only one unique output. For a transmission matrix with  $n \times n$  elements, there will be  $n!$  unique permutation matrices, each of which for our device can be seen in figure 4. An example phase shifted matrix is presented for each permutation target as well as its associated varFDTD simulation of the near field electric field intensity. These matrices are a useful test to ensure that the network has generalized to a point where it is able to predict patterns for targets which lie on the very edge of the training geometry space (those confining the output electric field fully to only one port), and to check the ability of inverse designed patterns to produce distinctly different optical outputs for each input mode. In our design challenge, we chose to set the phase of each non-zero matrix element as a free parameter which can be chosen to obtain the best solution in amplitude. This approach appears useful since i) in many use cases, the phase response of an optical switch is not critical and ii) if necessary, an optical phase shifter at each of the outputs is sufficient to rebalance the output phases in the device.

Results shown in figure 4 demonstrate accurate reproduction of both the amplitude and phase distribution for a selection of matrices, demonstrating that a well performing solution can be achieved for all of the orthogonal permutation matrices in the  $3 \times 3$  group. Here solutions were selected on the basis of combining fidelity with a low phase MSE. Pixel patterns as well as local field distributions for each of the input ports are presented for completeness for each of the devices, providing some insight in the internal structure of these solutions. For example, for the identity matrix at the top, we see that the pattern of perturbations is predominantly placed in the middle of the device along its length. This pattern transforms the original device, which had no self-imaging capability, to produce a highly symmetric self-imaging of each of the three inputs onto the output plane.

Anti-diagonal and more complex permutations are seen in the other five diagrams, where the common denominator is that the field profiles typically show only a few bounces between input and output, indicative of the coupling length of the CWG of around  $10 \mu\text{m}$  and the fact that the modal basis is very small. Weak perturbations therefore tend to couple modes in an adiabatic way, without giving rise to strong scattering events or reflections [34].

A sample of fidelities across the available phase space for each permutation matrix is presented in the right column in figure 4, in which 200 random phase shifts are applied to each element in the target matrix. Average fidelities are achieved of around 90% across the phase space sample for all 6 target matrices. When output phases can be freely adjusted to maximize the fidelity, the phase space sampling technique permits significant improvements to the fidelity, resulting in fidelities in excess of 0.97 for all targets.

### 4.2. General Unitary Transmission Matrices

Next to the design of orthogonal permutation matrix targets, which pose a useful design task to ensure that we have precise control over every individual matrix element, in many real world use cases such as optical computing and quantum information processing, mixtures of inputs are required. Unitary matrices represent a particularly useful class of targets representing the most general operations available in the multi-port system. These targets must satisfy the condition  $UU^* = I$ , where  $U^*$  represents the conjugate transpose of the target matrix and  $I$  is the identity matrix, corresponding to a lossless device used for target matrices.

Figure 5a) shows the fidelity distribution for network predicted patterns from a set of 1000 random unitary targets drawn from the Haar distribution. We find that the neural network is able to reach design targets with good fidelity of  $0.935 \pm 0.038$ . This is reflected in the convergence to low values below 0.05 in the amplitude MSE. The distribution of MSE values for phase shows that the prediction of correct phase distributions is more challenging, which will be discussed in Section 5 in more detail. A sub-selection of four matrix targets from across the fidelity distribution are presented in figure 5b) labelled as i) - iv), as indicated by the colored dots in figure 5b), showing the target matrix, the forward network prediction and varFDTD resimulation result. Pixel patterns and near field maps are present the corresponding microscopic configurations of the four patterned devices under study. It can be seen that for all four examples, the forward network prediction agrees well with the target matrix. Main differences can be seen between the network prediction and the varFDTD result, suggesting that an important factor in the

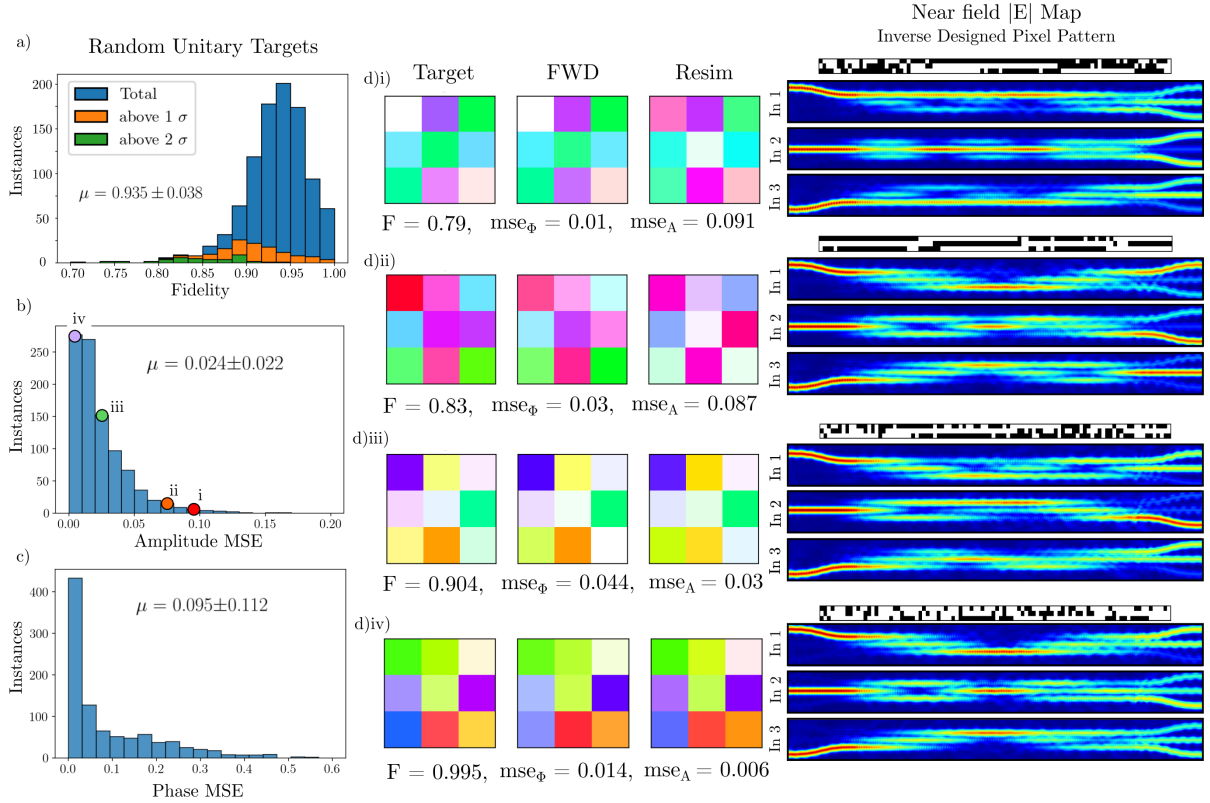


Fig. 5. a) Distribution of fidelities for the reproduction of 1000 random unitary targets. Highlighted are instances which fall outside  $1$  and  $2\sigma$  of the average MSE between forward network prediction and resimulation, indicating a failed forward network prediction. b) and c) show the distribution of MSE values between target and resimulation for these same targets. d) 4 example unitary targets sampled from across the achieved fidelity distribution. We present the target, forward network prediction of the inverse designed pixel pattern, and a resimulated transmission matrix. To the right hand side we plot the respective near field electric field intensity map for a device programmed with the predicted pattern for each target matrix.

fidelity may be the accuracy of the forward network.

To verify the influence of the forward network error, we calculate the MSE between forward network prediction and varFDTD calculation. The fidelity distributions are obtained for parts of the dataset where the MSE lies above one and two standard deviations (labelled as  $1\sigma$  and  $2\sigma$ ), given by respectively the orange and green histograms in figure 5a). This analysis confirms what is seen in the examples, namely that the tail of lower fidelities is correlated with a poor accuracy of the forward network prediction.

The examples furthermore indicate that phase and amplitude errors are not related. Results with low phase error can be found for a low fidelity and *vice versa*. Figure 5d)iii) for example shows visually a good match to the target unitary colors due to low phase error, however the final amplitude fidelity remains below average at 0.9 owing to a disagreement between amplitude values. The near field maps and pixel patterns of random unitaries are not easily understood through intuition, while an underlying phenomenology would be of interest this goes beyond the scope of our study. We do observe the emergence of larger section of connected pixels forming lines, which may be a strategy of the network to achieve large phase shifts in certain matrix elements. Some more discussion on the phase structure of the CWGs is presented in Section 5.

### 4.3. Hadamard and Fourier Matrix

Complex Hadamard matrices play an important role in quantum information theory. They have been used to tackle a number of problems including the development of spin models [76] and analogue quantum simulators [77]. They have also helped establish mathematical frameworks to construct bases of unitary operators and maximally entangled states. Fourier matrices can be used to apply a discrete Fourier transform to a signal through matrix multiplication, making them of particular interest in optical signal processing and computation. For a  $3 \times 3$  matrix, all complex Hadamard matrices are equivalent to the Fourier matrix,  $F_3$  [78]:

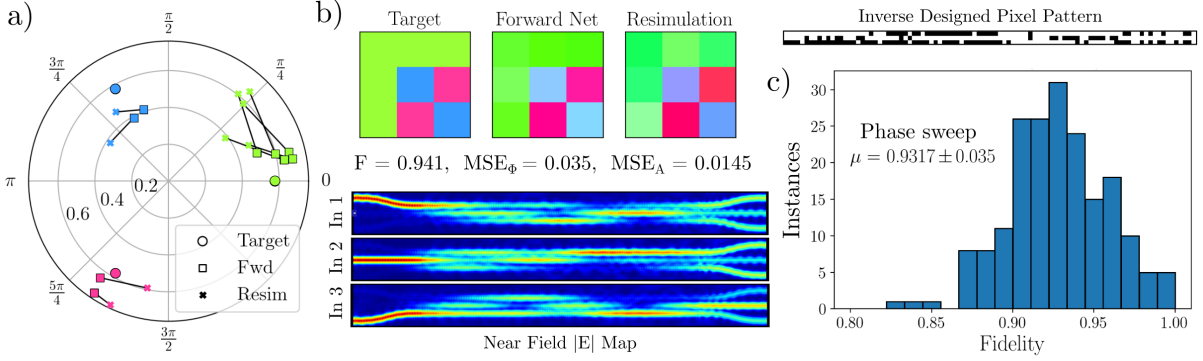


Fig. 6. a) Polar plot comparing each matrix element in the reproduction of the  $3 \times 3$  Hadamard and Fourier matrix (shown in sub figure b). Datapoints are color coded to their associated target value, and black lines indicate the distance on the complex plane between forward network predictions and their resimulated matrix values. c) Shows again a phase sweep, in which a global phase shift is applied to the target matrix, retaining the same intensity distribution resulting in a maximum achievable fidelity in excess of 0.99.

$$F_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{bmatrix}$$

Where  $\omega = \exp(2\pi i/3)$ . It is of interest to check the existence of a solution for this specific matrix operator. As  $F_3$  is a unique matrix, we again give the network some more flexibility in the boundary conditions by allowing a single global phase factor within all matrix elements. A global phase factor can easily be factorized out and compensated in an optical system. It makes sense to include this degree of freedom in the matrix to find the best working point of the CWG under study, taking into account both amplitude and phase MSE. Figure 6a) and b) shows the forward network prediction and resimulated result when attempting to implement the  $F_3$  matrix. An amplitude fidelity of 0.94 is achieved for the selected solution, with a phase MSE of 0.035 and amplitude MSE of 0.0145. The polar plot in figure 6a) is used as a graphical depiction of the accuracy for each individual matrix element, data points are colour coded corresponding with the matrix element they represent, with black lines between associated network predictions and resimulated points representing their separation within the complex plane.

The full histogram of fidelities for a range of global phase values between  $-\pi$  and  $\pi$  is shown in figure 6c) and allows to identify an average fidelity across this phase sample of 0.93, while we again are able to retrieve a peak fidelity of  $> 0.99$  but at a higher phase MSE.

## 5. Discussion

Our work shown here presents coupled waveguide arrays as an alternative to other commonly used reconfigurable platforms such as interferometer meshes or multi-mode based devices for programming arbitrary unitary operators. Approaches for using such devices are well established, with practical realizations demonstrated across a number of fields as discussed in the preamble to this work. Interferometer meshes in particular provide a high degree of control, allowing for the introduction of individual phase delays at single unit cells across the mesh for highly accurate matrix reproduction. There exist several mathematical models [79, 80] which allow for the calculation of where these phase shifts must be applied for arbitrary matrix decomposition. Despite promising results, small insertion losses at each interferometer will sum together, leading to non-negligible losses across the whole device. Furthermore, when targeting large matrices, the total device footprint becomes large when compared to alternatives. Adding to this, many approaches use thermo-optic modulators to introduce the required phase shifts, which add further to this size, introduce electronic overhead and regulating circuits, and produce thermal instability across the chip.

Multi-mode structures such as MMIs and devices which are built around multi-mode waveguides are the most compact technology discussed here. Extra degrees of freedom introduced by the inclusion of different optical modes allow for highly compact matrix decomposition, however, this is not without its drawbacks. Coupling out of the device becomes challenging as any modal-mismatch can introduce significant losses, often requiring some signal pre-processing through techniques such as beam shaping. The need for accurate coupling between modes

also now imposes a requirement for high degrees of spatial accuracy in the programming of such a device. Coupled waveguide arrays strike a midpoint in device footprint, necessitating only slightly more space in the direction of light propagation to allow a round trip between outermost waveguides if aiming for the smallest device possible. In the plane perpendicular to the light injection axis their size remains comparable, if not smaller than MMI type devices. This compact form-factor is significantly smaller than comparable interferometer meshes, and remains within a consistent modal basis, introducing no losses from coupling between modes at the output waveguides making it an attractive option for space-limited applications.

While in general the network performance in our work is strong, there are some examples where the resimulated results are not a good match to the target matrix. By comparing the forward network predictions to resimulated results we may check if forward network predictions remain valid. Using a simple MSE discriminator, instances which fall outside one standard deviation from the mean can be identified and rejected, highlighting cases where the forward network has failed to accurately predict the transmission matrix for a given pattern. The fidelity distribution with cases above  $1$  and  $2\sigma$  highlighted were shown in figure 5. For these designs, the network appears to have strayed from the interpolation regime of the forward network, resulting in large errors in the predicted transmission matrix. Rather unsurprisingly, almost every fidelity below  $0.85$  is a result of a failed network prediction. Excluding these from our statistical analysis the average fidelity now raises to  $0.94 \pm 0.029$ . In the majority of cases however, it appears that the optimization of the geometry in the WGAN's latent space and the addition of the latent constraint successfully limits designs to the surrogate network's validity region. Iterative data improvement may also be continued beyond the three cycles in this study, which apart from making the network learn from its own mistakes allows for augmenting the training data toward larger pixel numbers. The performance gain has to be traded off against the effort needed as the additional improvement for each cycle is expected to saturate.

The network is trained on the full complex field, therefore one might expect similar performance for phase and amplitude, study of our respective average MSE values shows this is however, not the case. One hypothesis for the cause of this discrepancy originates in how brute force optimization targets are generated. During the data generation stage, the phase of light at the outputs is unconstrained and given no optimization weighting, we assess the success of each pixel iteration solely on the intensity at the output ports. Intensity optimization targets are randomly selected from a uniform distribution, resulting in the retrieval of a normal distribution of possible amplitude values centered around  $0.33$  as we expect for a 3-waveguide system. Conversely, analysis of the phase distribution for the same data (shown in figure S10) shows distinctly different behaviour. Unsurprisingly we record a peak for phase values matching to those of an unperturbed device, however, because the optimization is not pushed to suggest pixel patterns with "extreme" phase delays approaching  $2\pi$ , the dataset becomes highly biased towards smaller phase shifts. Although these phases can be accessed by switching extra pixels, there is no incentive for the optimization in doing so, as excess pixels come at the cost of increased scattering losses and lower transmissions, making these patterns more likely to be rejected. Because of this, the final database of pixel patterns represents the available intensity space well, but may fall short when describing the phase space relating to longer phase delays. Consequently it follows that the network's ability to accurately predict patterns to implement such phase delays will be inhibited, explaining the differing MSE results for phase and amplitude we observe.

The major limiting factor in the final generation of valid pixel patterns for user-defined matrix targets is therefore the accuracy with which the forward network can make predictions. Assessment of this will give some insight as to the expected upper limit on network performance when presented with an external design target. Figure S9 shows the distribution of fidelities from forward network predictions on the validation dataset discussed in section 2.3. This represents a "best case" scenario, as these patterns are generated in precisely the same way as the training data. Across the 6564 validation patterns we record an average fidelity of  $0.92 \pm 0.04$ , indicating that for our current predictor model, the achieved average amplitude fidelity of  $0.94$  constitutes what is likely a near maximal achievable value, which is limited by the network performance rather than the physical CWG system. More conventional topology optimizations starting from the endpoint of the neural-adjoint inverse design may be of interest to further converge solutions, which could be a topic of future work.

## 6. Conclusions

In conclusion, our study addresses the potential of low-loss optical phase change materials combined with coupled waveguide arrays as a promising and exciting avenue for the production of next-generation reconfigurable technologies across fields such as quantum simulation, photonic computing and optical data processing. In comparison with other approaches such as integrated circuits based around interferometer meshes or MMIs, coupled waveguide arrays can offer ultra-low-loss reversible modulation requiring no active regulation within a highly compact device footprint. The development of a robust inverse design pipeline using neural network surrogate models allows for rapid, near-real-time prediction of the complex pixel patterns, required to implement a wide range of transmission matrices within a single device model. In our work, the introduction of a Wasserstein generative adversarial network provides a crucial constraint on the gradient based optimization, limiting predicted geometries

to those within the interpolation region of the forward surrogate model. Network performance is enhanced by augmenting and expanding the training dataset, initially through a process of iterative improvement, and in a final step, with the introduction of a small percentage of noisy random data. Although training of the networks is time consuming, taking a few hours with standard consumer level computer hardware, the entire inverse design process remains significantly faster than alternatives such as topology optimizations, allowing predictions on a millisecond timescale once trained.

Presented results demonstrate a high level of control over both intensity and phase of individual matrix elements. While some extreme phase relations may suggest the need to expand the waveguide geometries to longer device, therefore allowing multiple vertical passes of the light, generally performance is strong. Average fidelities of  $0.935 \pm 0.04$  are reported, demonstrating comparable performance to other state of the art, commercially available reconfigurable photonic technologies and further reinforcing the validity of this approach. The fidelity obtained in our study is currently limited by the neural network performance rather than the coupled waveguide system itself and further improvement may be possible in future work. The neural-adjoint design platform introduced is highly versatile and requires minimal overheads to allow functionality across different devices and geometries, showing promise for integration with a range of future optical technologies.

## References

1. H. Zhou, J. Dong, J. Cheng, W. Dong, C. Huang, Y. Shen, Q. Zhang, M. Gu, C. Qian, H. Chen, et al., "Photonic matrix multiplication lights up photonic accelerator and beyond," *Light: Science & Applications*, vol. 11, no. 1, p. 30, 2022.
2. G. Giamougiannis, A. Tsakyridis, M. Moralis-Pegios, G. Mourgiyas-Alexandris, A. R. Totovic, G. Dabos, M. Kirtas, N. Passalis, A. Tefas, D. Kalavrouziotis, et al., "Neuromorphic silicon photonics with 50 ghz tiled matrix multiplication for deep-learning applications," *Advanced Photonics*, vol. 5, no. 1, pp. 016004–016004, 2023.
3. A. Aspuru-Guzik and P. Walther, "Photonic quantum simulators," *Nature physics*, vol. 8, no. 4, pp. 285–291, 2012.
4. J. Wang, F. Sciarrino, A. Laing, and M. G. Thompson, "Integrated photonic quantum technologies," *Nature Photonics*, vol. 14, no. 5, pp. 273–284, 2020.
5. K. Choutagunta, I. Roberts, D. A. Miller, and J. M. Kahn, "Adapting mach–zehnder mesh equalizers in direct-detection mode-division-multiplexed links," *Journal of Lightwave Technology*, vol. 38, no. 4, pp. 723–735, 2019.
6. B. J. Shastri, A. N. Tait, T. Ferreira de Lima, W. H. Pernice, H. Bhaskaran, C. D. Wright, and P. R. Prucnal, "Photonics for artificial intelligence and neuromorphic computing," *Nature Photonics*, vol. 15, no. 2, pp. 102–114, 2021.
7. W. Bogaerts and A. Rahim, "Programmable photonics: An opportunity for an accessible large-volume pic ecosystem," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 5, pp. 1–17, 2020.
8. W. Bogaerts, D. Pérez, J. Capmany, D. A. B. Miller, J. Poon, D. Englund, F. Morichetti, and A. Melloni, "Programmable photonic circuits," *Nature*, vol. 586, pp. 207–216, 2020.
9. X. Chen, M. M. Milosevic, D. J. Thomson, A. Z. Khokhar, Y. Franz, A. F. Runge, S. Mailis, A. C. Peacock, and G. T. Reed, "Post-fabrication phase trimming of mach–zehnder interferometers by laser annealing of germanium implanted waveguides," *Photonics Research*, vol. 5, no. 6, pp. 578–582, 2017.
10. J. Zheng, A. Khanolkar, P. Xu, S. Colburn, S. Deshmukh, J. Myers, J. Frantz, E. Pop, J. Hendrickson, J. Doylend, et al., "Gst-on-silicon hybrid nanophotonic integrated circuits: a non-volatile quasi-continuously reprogrammable platform," *Optical Materials Express*, vol. 8, no. 6, pp. 1551–1561, 2018.
11. E. Staffoli, M. Mancinelli, P. Bettotti, and L. Pavesi, "Equalization of a 10 gbps imdd signal by a small silicon photonics time delayed neural network," *Photonics Research*, vol. 11, no. 5, pp. 878–886, 2023.
12. Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, et al., "Deep learning with coherent nanophotonic circuits," *Nature photonics*, vol. 11, no. 7, pp. 441–446, 2017.
13. D. Pérez, I. Gasulla, L. Crudgington, D. J. Thomson, A. Z. Khokhar, K. Li, W. Cao, G. Z. Mashanovich, and J. Capmany, "Multipurpose silicon photonics signal processor core," *Nature communications*, vol. 8, no. 1, p. 636, 2017.
14. F. Ashtiani, A. J. Geers, and F. Aflatouni, "An on-chip photonic deep neural network for image classification," *Nature*, vol. 606, no. 7914, pp. 501–506, 2022.
15. D. Wu, Y. Wu, Y. Wang, J. An, and X. Hu, "Reconfigurable optical add-drop multiplexer based on thermally tunable micro-ring resonators," *Optics Communications*, vol. 367, pp. 44–49, 2016.
16. A. H. Atabaki, A. A. Eftekhari, S. Yegnanarayanan, and A. Adibi, "Sub-100-nanosecond thermal reconfiguration of silicon photonic devices," *Optics express*, vol. 21, no. 13, pp. 15706–15718, 2013.
17. C. Errando-Herranz, F. Niklaus, G. Stemme, and K. B. Gylfason, "A low-power mems tunable photonic ring resonator for reconfigurable optical networks," in *2015 28th IEEE International Conference on Micro Electro Mechanical Systems (MEMS)*, pp. 53–56, IEEE, 2015.
18. M. Oshita, H. Takahashi, Y. Ajiki, and T. Kan, "Reconfigurable surface plasmon resonance photodetector with a mems deformable cantilever," *ACS Photonics*, vol. 7, no. 3, pp. 673–679, 2020.
19. N. Quack, A. Y. Takabayashi, H. Sattari, P. Edinger, G. Jo, S. J. Bleiker, C. Errando-Herranz, K. B. Gylfason, F. Niklaus, U. Khan, P. Verheyen, A. K. Mallik, J. S. Lee, M. Jezzini, I. Zand, P. Morrissey, C. Antony, P. O'Brien, and W. Bogaerts, "Integrated silicon photonic mems," *Microsystems & Nanoengineering*, vol. 9, p. 27, 2023.
20. M. Delaney, I. Zeimpekis, D. Lawson, D. W. Hewak, and O. L. Muskens, "A new family of ultralow loss reversible

- phase-change materials for photonic integrated circuits: Sb<sub>2</sub>S<sub>3</sub> and Sb<sub>2</sub>Se<sub>3</sub>,” *Advanced Functional Materials*, vol. 30, no. 36, p. 2002447, 2020.
21. K. J. Miller, R. F. Haglund, and S. M. Weiss, “Optical phase change materials in integrated silicon photonic devices: review,” *Optical Materials Express*, Vol. 8, Issue 8, pp. 2415–2429, vol. 8, pp. 2415–2429, 8 2018.
  22. R. Chen, Z. Fang, F. Miller, H. Rarick, J. E. Fröch, and A. Majumdar, “Opportunities and challenges for large-scale phase-change material integrated electro-photonics,” *ACS Photonics*, vol. 9, no. 10, pp. 3181–3195, 2022.
  23. C. Wu, H. Deng, Y.-S. Huang, H. Yu, I. Takeuchi, C. A. R. Ocampo, and M. Li, “Freeform direct-write and rewritable photonic integrated circuits in phase-change thin films,” *Science Advances*, vol. 10, no. 1, p. eadk1361, 2024.
  24. P. Prabhathan, K. V. Sreekanth, J. Teng, J. H. Ko, Y. J. Yoo, H.-H. Jeong, Y. Lee, S. Zhang, T. Cao, C.-C. Popescu, B. Mills, T. Gu, Z. Fang, R. Chen, H. Tong, Y. Wang, Q. He, Y. Lu, Z. Liu, H. Yu, A. Mandal, Y. Cui, A. S. Ansari, V. Bhingardive, M. Kang, C. K. Lai, M. Merklein, M. J. Müller, Y. M. Song, Z. Tian, J. Hu, M. Losurdo, A. Majumdar, X. Miao, X. Chen, B. Gholipour, K. A. Richardson, B. J. Eggleton, M. Wuttig, and R. Singh, “Roadmap for phase change materials in photonics and beyond,” *iScience*, vol. 26, no. 10, p. 107946, 2023.
  25. D. Tripathi, H. S. Vyas, S. Kumar, S. S. Panda, and R. Hegde, “Recent developments in chalcogenide phase change material-based nanophotonics,” *Nanotechnology*, vol. 34, p. 502001, sep 2023.
  26. M. Wuttig, H. Bhaskaran, and T. Taubner, “Phase-change materials for non-volatile photonic applications,” *Nature photonics*, vol. 11, no. 8, pp. 465–476, 2017.
  27. M. Delaney, I. Zeimpekis, H. Du, X. Yan, M. Banakar, D. J. Thomson, D. W. Hewak, and O. L. Muskens, “Nonvolatile programmable silicon photonics using an ultralow-loss Sb<sub>2</sub>Se<sub>3</sub> phase change material,” *Science Advances*, vol. 7, pp. 3500–3516, 6 2021.
  28. L. Martin-Monier, C. C. Popescu, L. Ranno, B. Mills, S. Geiger, D. Callahan, M. Moebius, and J. Hu, “Endurance of chalcogenide optical phase change materials: a review,” *Opt. Mater. Express*, vol. 12, pp. 2145–2167, Jun 2022.
  29. D. Lawson, S. Blundell, M. Ebert, O. L. Muskens, and I. Zeimpekis, “Optical switching beyond a million cycles of low-loss phase change material Sb<sub>2</sub>Se<sub>3</sub>,” *Optical Materials Express*, vol. 14, no. 1, pp. 22–38, 2024.
  30. M. S. Alam, R. Laing, Z. Bolatbek, R. Heenkenda, R. Gnawali, T. E. Payne, A. Sarangan, J. R. Hendrickson, and I. Agha, “Fast cycling speed with multimillion cycling endurance of ultra-low loss phase change material (Sb<sub>2</sub>Se<sub>3</sub>) by engineered laser pulse irradiation,” *Advanced Functional Materials*, vol. 34, no. 14, p. 2310306, 2024.
  31. D. A. B. Miller, “Self-configuring universal linear optical component [Invited],” *Photonics Research*, vol. 1, p. 1, jun 2013.
  32. V. Liu, Y. Jiao, D. A. Miller, and S. Fan, “Design methodology for compact photonic-crystal-based wavelength division multiplexers,” *Optics letters*, vol. 36, no. 4, pp. 591–593, 2011.
  33. J. Lu and J. Vučković, “Nanophotonic computational design,” *Optics express*, vol. 21, no. 11, pp. 13351–13367, 2013.
  34. R. Bruck, K. Vynck, P. Lalanne, B. Mills, D. J. Thomson, G. Z. Mashanovich, G. T. Reed, and O. L. Muskens, “All-optical spatial light modulator for reconfigurable silicon photonic circuits,” *Optica*, vol. 3, p. 396, Apr. 2016.
  35. L. F. Frellsen, Y. Ding, O. Sigmund, and L. H. Frandsen, “Topology optimized mode multiplexing in silicon-on-insulator photonic wire waveguides,” *Optics express*, vol. 24, no. 15, pp. 16866–16873, 2016.
  36. N. J. Dinsdale, P. R. Wiecha, M. Delaney, J. Reynolds, M. Ebert, I. Zeimpekis, D. J. Thomson, G. T. Reed, P. Lalanne, K. Vynck, et al., “Deep learning enabled design of complex transmission matrices for universal optical components,” *ACS photonics*, vol. 8, no. 1, pp. 283–295, 2021.
  37. V. Nikkhah, A. Pirmoradi, F. Ashtiani, B. Edwards, F. Aflatouni, and N. Engheta, “Inverse-designed low-index-contrast structures on a silicon photonics platform for vector–matrix multiplication,” *Nature Photonics*, pp. 1–8, 2024.
  38. C. Taballione, T. A. Wolterink, J. Lugani, A. Eckstein, B. A. Bell, R. Grootjans, I. Visscher, D. Geskus, C. G. Roelofzen, J. J. Renema, et al., “8×8 reconfigurable quantum photonic processor based on silicon nitride waveguides,” *Optics express*, vol. 27, no. 19, pp. 26842–26857, 2019.
  39. H. B. Perets, Y. Lahini, F. Pozzi, M. Sorel, R. Morandotti, and Y. Silberberg, “Realization of quantum walks with negligible decoherence in waveguide lattices,” *Phys. Rev. Lett.*, vol. 100, p. 170506, May 2008.
  40. A. Peruzzo, M. Lobino, J. C. F. Matthews, N. Matsuda, A. Politi, K. Poulios, X.-Q. Zhou, Y. Lahini, N. Ismail, K. Wörhoff, Y. Bromberg, Y. Silberberg, M. G. Thompson, and J. L. O’Brien, “Quantum walks of correlated photons,” *Science*, vol. 329, no. 5998, pp. 1500–1503, 2010.
  41. I. L. Garanovich, S. Longhi, A. A. Sukhorukov, and Y. S. Kivshar, “Light propagation and localization in modulated photonic lattices and waveguides,” *Physics Reports*, vol. 518, no. 1, pp. 1–79, 2012. Light propagation and localization in modulated photonic lattices and waveguides.
  42. Y. Lahini, G. R. Steinbrecher, A. D. Bookatz, and D. Englund, “Realization of quantum walks with negligible decoherence in waveguide lattices,” *njp Quant. Inf.*, vol. 4, p. 2, 2018.
  43. J. Petrovic, “Multiport waveguide couplers with periodic energy exchange,” *Opt. Lett.*, vol. 40, pp. 139–142, Jan 2015.
  44. J. Petrovic, J. Kršić, A. Maluckov, and J. Veerman, “High-density optical interconnects based on self-imaging in coupled waveguide arrays,” *Optics & Laser Technology*, vol. 163, p. 109381, 2023.
  45. Y. Yang, R. J. Chapman, B. Haylock, F. Lenzini, Y. N. Joglekar, M. Lobino, and A. Peruzzo, “Programmable high-dimensional hamiltonian in a photonic waveguide array,” *Nat. Comm.*, vol. 15, p. 50, 2024.
  46. M. P. Bendsøe, “Optimal shape design as a material distribution problem,” *Structural optimization*, vol. 1, pp. 193–202, 1989.
  47. J. S. Jensen and O. Sigmund, “Topology optimization for nano-photonics,” *Laser & Photonics Reviews*, vol. 5, no. 2,

- pp. 308–321, 2011.
48. L. Martinelli and A. Jameson, “Computational aerodynamics: solvers and shape optimization,” *Journal of heat transfer*, vol. 135, no. 1, p. 011002, 2013.
  49. M. Seabra, J. Azevedo, A. Araújo, L. Reis, E. Pinto, N. Alves, R. Santos, and J. P. Mortágua, “Selective laser melting (slm) and topology optimization for lighter aerospace componentes,” *Procedia Structural Integrity*, vol. 1, pp. 289–296, 2016.
  50. X. Wang, S. Xu, S. Zhou, W. Xu, M. Leary, P. Choong, M. Qian, M. Brandt, and Y. M. Xie, “Topological design and additive manufacturing of porous metals for bone scaffolds and orthopaedic implants: A review,” *Biomaterials*, vol. 83, pp. 127–141, 2016.
  51. A. A. Al-Tamimi, H. Almeida, and P. Bartolo, “Structural optimisation for medical implants through additive manufacturing,” *Progress in Additive Manufacturing*, vol. 5, no. 2, pp. 95–110, 2020.
  52. L. Lu, J. D. Joannopoulos, and M. Soljačić, “Topological photonics,” *Nature photonics*, vol. 8, no. 11, pp. 821–829, 2014.
  53. A. Y. Piggott, J. Lu, K. G. Lagoudakis, J. Petykiewicz, T. M. Babinec, and J. Vučković, “Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer,” *Nature photonics*, vol. 9, no. 6, pp. 374–377, 2015.
  54. Z. Lin, X. Liang, M. Lončar, S. G. Johnson, and A. W. Rodriguez, “Cavity-enhanced second-harmonic generation via nonlinear-overlap optimization,” *Optica*, vol. 3, no. 3, pp. 233–238, 2016.
  55. D. Sell, J. Yang, S. Doshay, R. Yang, and J. A. Fan, “Large-angle, multifunctional metagratings based on freeform multimode geometries,” *Nano letters*, vol. 17, no. 6, pp. 3752–3757, 2017.
  56. R. E. Christiansen, Z. Lin, C. Roques-Carmes, Y. Salamin, S. E. Kooi, J. D. Joannopoulos, M. Soljačić, and S. G. Johnson, “Fullwave maxwell inverse design of axisymmetric, tunable, and multi-scale multi-wavelength metalenses,” *Optics Express*, vol. 28, no. 23, pp. 33854–33868, 2020.
  57. Y. Augenstein and C. Rockstuhl, “Inverse design of nanophotonic devices with structural integrity,” *ACS photonics*, vol. 7, no. 8, pp. 2190–2196, 2020.
  58. O. Sigmund, “On the usefulness of non-gradient approaches in topology optimization,” *Structural and Multidisciplinary Optimization*, vol. 43, pp. 589–596, 2011.
  59. P. Bennet, D. Langevin, C. Essoual, A. Khairah-Walieh, O. Teytaud, P. Wiecha, and A. Moreau, “Illustrated tutorial on global optimization in nanophotonics,” *JOSA B*, vol. 41, pp. A126–A145, Feb. 2024.
  60. R. S. Hegde, “Photonics inverse design: pairing deep neural networks with evolutionary algorithms,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 1, pp. 1–8, 2019.
  61. Y. Ren, L. Zhang, W. Wang, X. Wang, Y. Lei, Y. Xue, X. Sun, and W. Zhang, “Genetic-algorithm-based deep neural networks for highly efficient photonic device design,” *Photonics Research*, vol. 9, no. 6, pp. B247–B252, 2021.
  62. R. V. Woldseth, N. Aage, J. A. Barentzen, and O. Sigmund, “On the use of artificial neural networks in topology optimisation,” *Structural and Multidisciplinary Optimization*, vol. 65, no. 10, p. 294, 2022.
  63. J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, “Nanophotonic particle simulation and inverse design using artificial neural networks,” *Science advances*, vol. 4, no. 6, p. eaar4206, 2018.
  64. Y. Deng, S. Ren, K. Fan, J. M. Malof, and W. J. Padilla, “Neural-adjoint method for the inverse design of all-dielectric metasurfaces,” *Optics Express*, vol. 29, no. 5, pp. 7526–7534, 2021.
  65. S. Ren, W. Padilla, and J. Malof, “Benchmarking deep inverse models over time, and the neural-adjoint method,” 2021.
  66. Y. Augenstein, T. Repan, and C. Rockstuhl, “Neural operator-based surrogate solver for free-form electromagnetic inverse design,” *ACS Photonics*, vol. 10, no. 5, pp. 1547–1557, 2023.
  67. S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, “Aggregated residual transformations for deep neural networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1492–1500, 2017.
  68. P. R. Wiecha and O. L. Muskens, “Deep Learning Meets Nanophotonics: A Generalized Accurate Predictor for Near Fields and Far Fields of Arbitrary 3D Nanostructures,” *Nano Letters*, vol. 20, pp. 329–338, Jan. 2020.
  69. A. Khairah-Walieh, D. Langevin, P. Bennet, O. Teytaud, A. Moreau, and P. R. Wiecha, “A newcomer’s guide to deep learning for inverse design in nano-photonics,” *Nanophotonics*, vol. 12, pp. 4387–4414, Nov. 2023.
  70. I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, “Improved training of wasserstein gans,” *Advances in neural information processing systems*, vol. 30, 2017.
  71. C. Taballione, R. van der Meer, H. J. Snijders, P. Hooijschuur, J. P. Epping, M. de Goede, B. Kassenberg, P. Venderbosch, C. Toebes, H. van den Vlekkert, et al., “A universal fully reconfigurable 12-mode quantum photonic processor,” *Materials for Quantum Technology*, vol. 1, no. 3, p. 035002, 2021.
  72. A.-P. Blanchard-Dionne and O. J. F. Martin, “Successive training of a generative adversarial network for the design of an optical cloak,” *OSA Continuum*, vol. 4, pp. 87–95, Jan. 2021.
  73. M. de Goede, H. Snijders, P. Venderbosch, B. Kassenberg, N. Kannan, D. Smith, C. Taballione, J. P. Epping, H. van den Vlekkert, and J. J. Renema, “High fidelity 12-mode quantum photonic processor operating at ingaas quantum dot wavelength,” in *Integrated Photonics Research, Silicon and Nanophotonics*, pp. ITu4B–3, Optica Publishing Group, 2022.
  74. C. Taballione, M. C. Anguita, M. de Goede, P. Venderbosch, B. Kassenberg, H. Snijders, N. Kannan, W. L. Vleeshouwers, D. Smith, J. P. Epping, et al., “20-mode universal quantum photonic processor,” *Quantum*, vol. 7, p. 1071, 2023.



75. N. Klokou, J. Gorecki, B. Beddoes, and V. Apostolopoulos, "Deep neural network ensembles for thz-tds refractive index extraction exhibiting resilience to experimental and analytical errors," *Optics Express*, vol. 31, no. 26, pp. 44575–44587, 2023.
76. K. Nomura, "Spin models constructed from hadamard matrices," *Journal of Combinatorial Theory, Series A*, vol. 68, no. 2, pp. 251–261, 1994.
77. D. Leung, "Simulation and reversal of n-qubit hamiltonians using hadamard matrices," *Journal of Modern Optics*, vol. 49, no. 8, pp. 1199–1217, 2002.
78. W. Tadej and K. Życzkowski, "A concise guide to complex hadamard matrices," *Open Systems & Information Dynamics*, vol. 13, no. 2, pp. 133–177, 2006.
79. M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, "Experimental realization of any discrete unitary operator," *Physical review letters*, vol. 73, no. 1, p. 58, 1994.
80. W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, "Optimal design for universal multiport interferometers," *Optica*, vol. 3, no. 12, pp. 1460–1465, 2016.
81. S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated Residual Transformations for Deep Neural Networks," Apr. 2017.
82. S. L. Smith, P.-J. Kindermans, C. Ying, and Q. V. Le, "Don't Decay the Learning Rate, Increase the Batch Size," [arXiv:1711.00489 \[cs, stat\]](https://arxiv.org/abs/1711.00489), Feb. 2018.
83. D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," [arXiv:1412.6980 \[cs\]](https://arxiv.org/abs/1412.6980), Dec. 2014.

## Supplementary Information

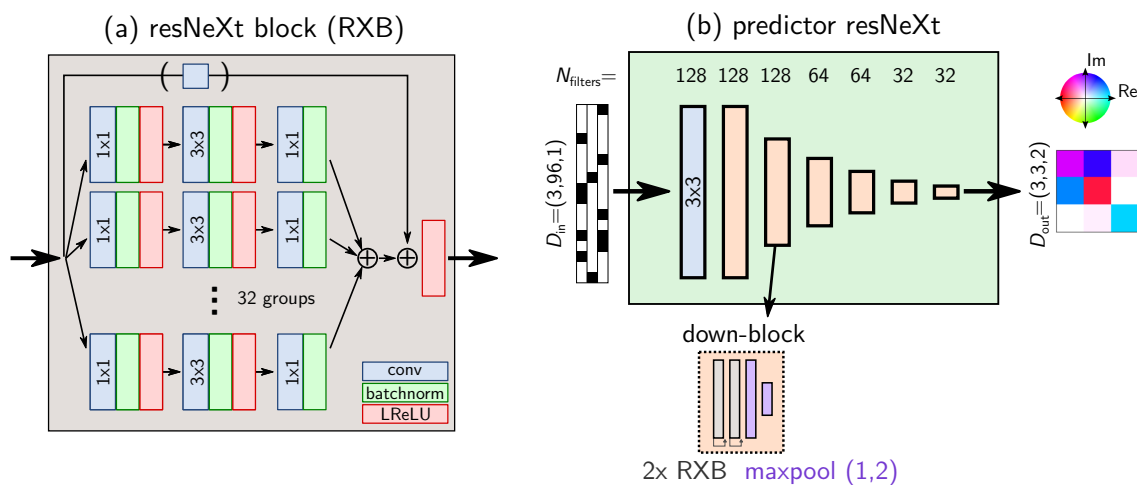


Fig. S1. Detailed sketch of the forward network architecture. The transmission predictor is composed of ResNeXt blocks, using grouped convolutions (a). We use a cardinality of 32 and a bottleneck width of 4. For similar computational cost and network complexity, they are known to be more efficient than classic residual blocks [81]. The forward model (b) is trained for 100 epochs using a simple plateau learning rate reduction with a factor of 0.5. Every 25 epochs we furthermore increase the batchsize by a factor of 2, from a starting value of 16, following the suggestion of [82].

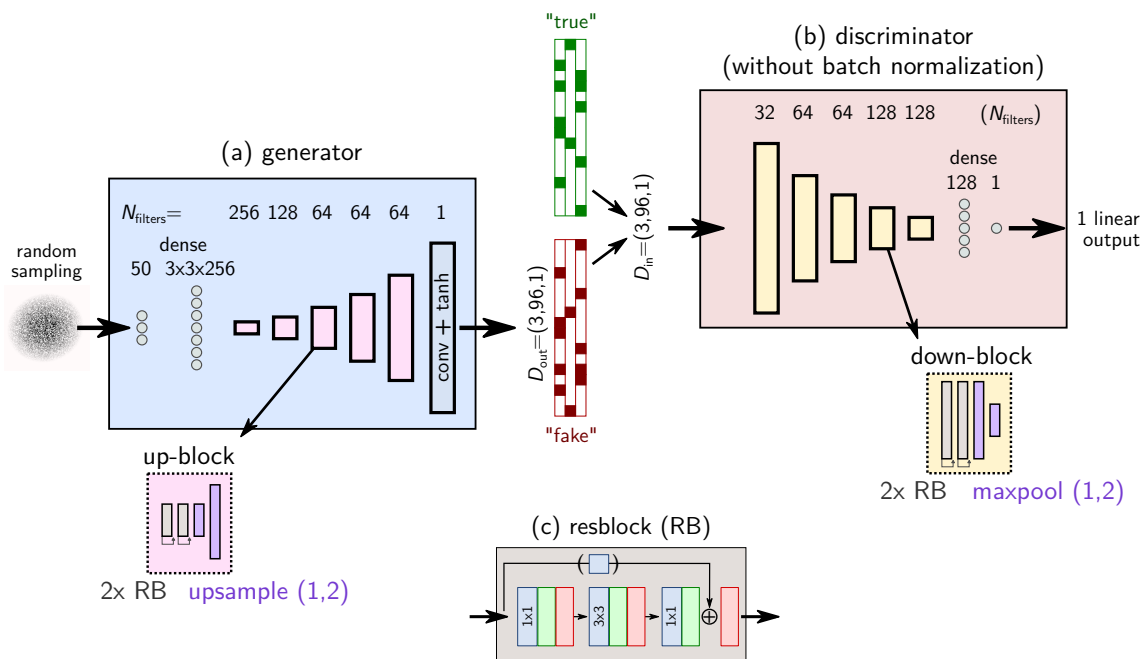


Fig. S2. Detailed sketch of the WGAN-GP generator (a) and discriminator (b) architecture, following closely the layout of the original implementation [70]. In particular, we use conventional resnet blocks (c) and no batch normalization in the discriminator. The latter is trained for 3 steps for a single training step of the generator. We train during 15 epochs first with a batch sizes of 64, followed by 15 additional epochs with a batchsize of 128. Both networks are trained using the Adam optimizer [83] with a fixed learning rate of 0.00005.

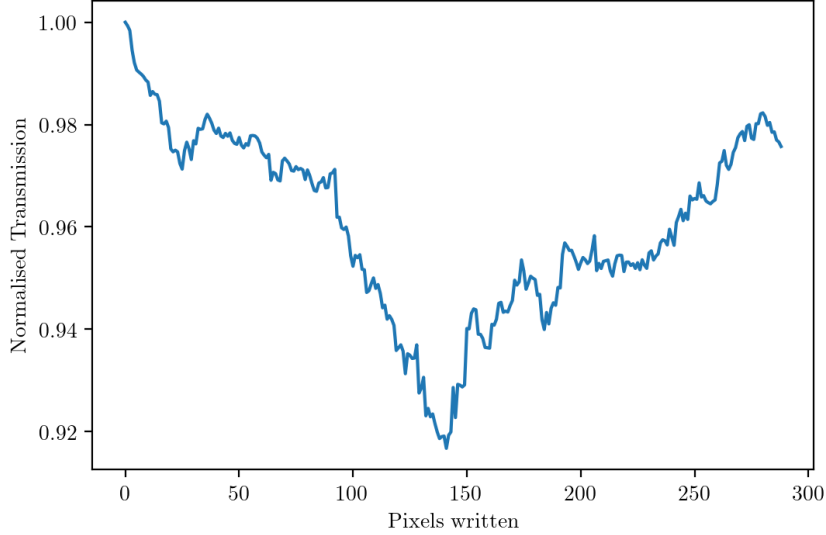


Fig. S3. Total device transmission normalised to the unperturbed throughput as a function of number of pixels switched. Switching pixels introduces reflections and scattering losses at the pixel boundaries, decreasing transmission. As  $> 50\%$  of pixels are switched the number of these boundaries decreases and transmission increases. The maximum recorded drop in transmission is 8%.

#### *Brute force iterative optimization*

A brute force iterative optimization is used to generate training data for training the network presented above. The algorithm is a simple approach which can be applied to any multiport device. To begin the device structure in its unperturbed state is set up using a commonly available varFDTD software. A random splitting ratio and input waveguide is selected before simulating the device to determine the initial output characteristics, a weight is then calculated using a common mathematical approach, for presented training data the weight function is as follows:

$$W = \sum_{i=1}^n 1 - \Delta T_i \quad (1)$$

Where  $n$  is the number of output ports,  $\Delta T$  represents the difference in transmission between the target and current geometry. As the output approaches the target the weight function is maximized and therefore at each iteration we can compare the performance of the pattern on average. This approach is sufficient for a simple single port optimization, however, if a multi-port optimization is preferable then a further summation over the input waveguides would be necessary also.

After the initial weight has been calculated a random pixel is selected and its state switched between crystalline or amorphous. A new weight is now calculated, if the weight has increased we know the transmission has moved closer to the target and the perturbation can be kept, if not the perturbation is reversed. Each time a pixel is saved a sweep across all input waveguides is carried out, allowing for the construction of the full device transmission matrix. Initial testing showed that choosing pixels at random to perturb, decreased simulation time and increased the optimization accuracy for a set number of iterations.

Figure S4 shows an example optimization for a 0%/50%/40% splitting ratio for light injected into input channel 1. This demonstrates both the negligible scattering losses as more and more pixels are written onto the device using our platform, as well as the tendency of such an approach to become stuck in a local minima.

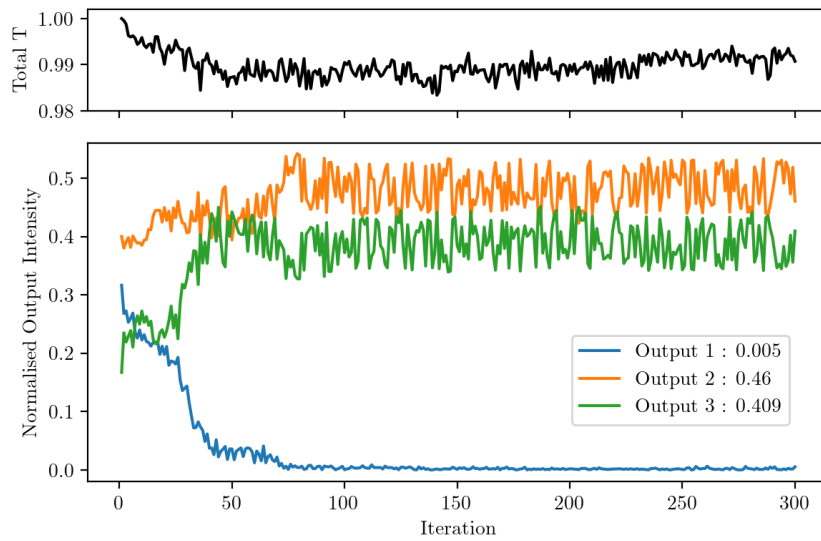


Fig. S4. Example brute force optimization of a single input for a  $3 \times 3$  waveguide array. The target splitting ratio (0%/50%/40%) is approached within 100 iterations before becoming stuck in a local minima. Total device throughput, shown in the top panel remains high as pixels are written, dropping only by 0.6% across the entire optimization.

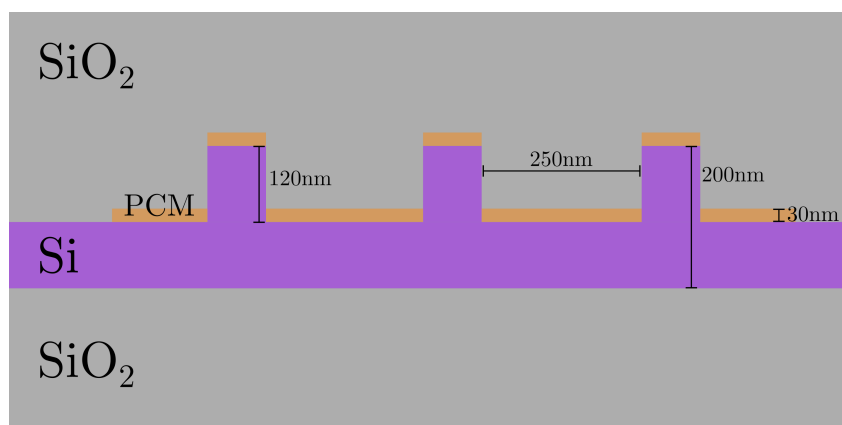


Fig. S5. Cross section schematic of the  $3 \times 3$  Si waveguide array inside the coupling region. Outside of this waveguides fan out to a spacing of  $1 \mu\text{m}$  to avoid cross-talk.

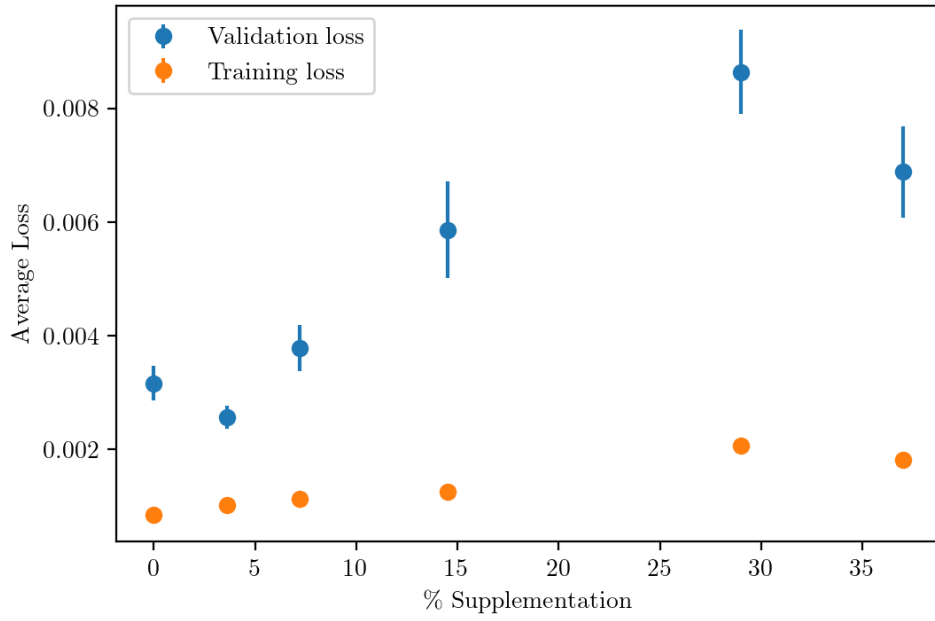


Fig. S6. Effect of supplementing a small amount of randomly generated "noisy" data. Un-optimized patterns are simulated and added into the initial training data set. Here we present average training losses for the forward network, as well as losses when predicting the transmission matrix for a separate validation dataset. Results are the average of 5 full training cycles with the standard deviation represented by the error bars.

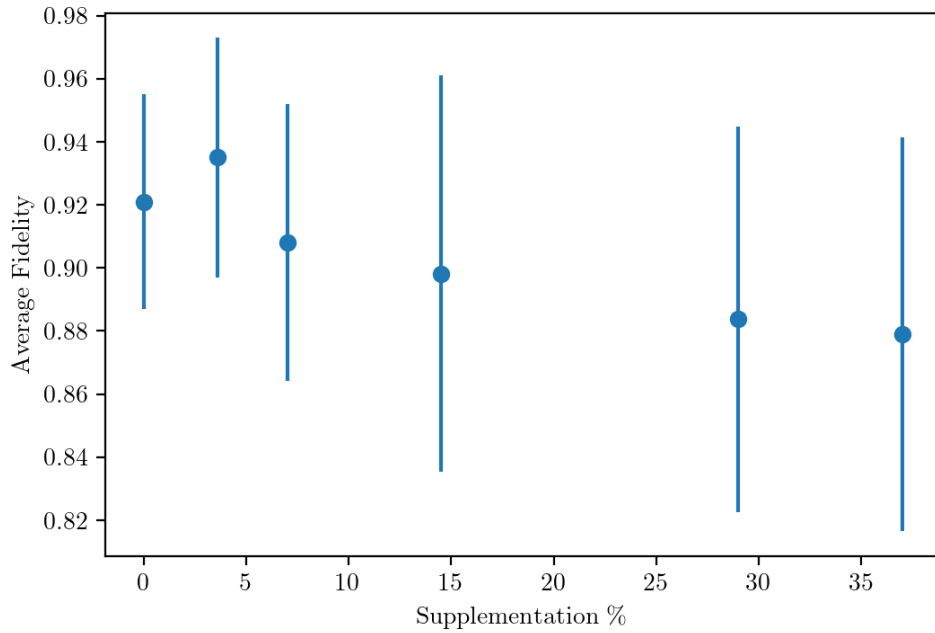


Fig. S7. After dataset augmentation with varying percentages of random noisy data, each network was presented with 1000 random unitary targets for which the average fidelity is plotted above.

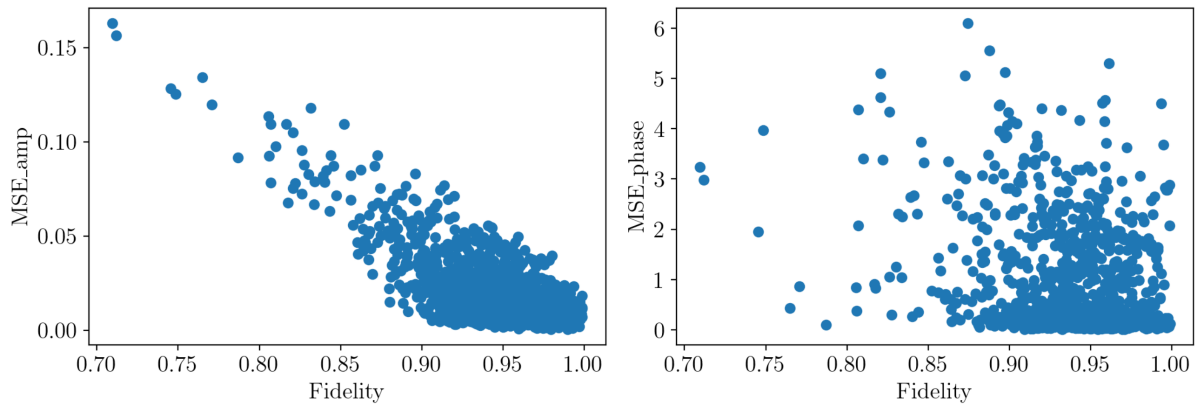


Fig. S8. Correlation between amplitude fidelity and MSE for both the phase and amplitude across 1000 Haar random matrix targets. MSE is calculated between the target and resimulation matrix with the average across all 9 elements presented here. We observe a linear correlation between amplitude MSE and fidelity but a mostly uncorrelated relationship with phase MSE as we would have predicted from our fidelity formulation.

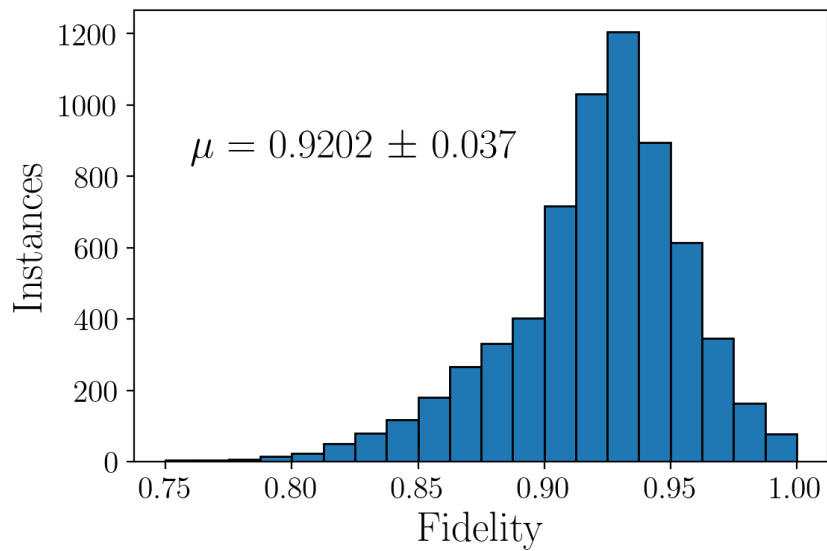


Fig. S9. Amplitude fidelity distribution taken from forward network predictions for the validation dataset of pixel patterns. This prediction error is intrinsic to all presented results, and will become a limiting factor as the network is improved, outlining an upper limit on total inverse design accuracy.

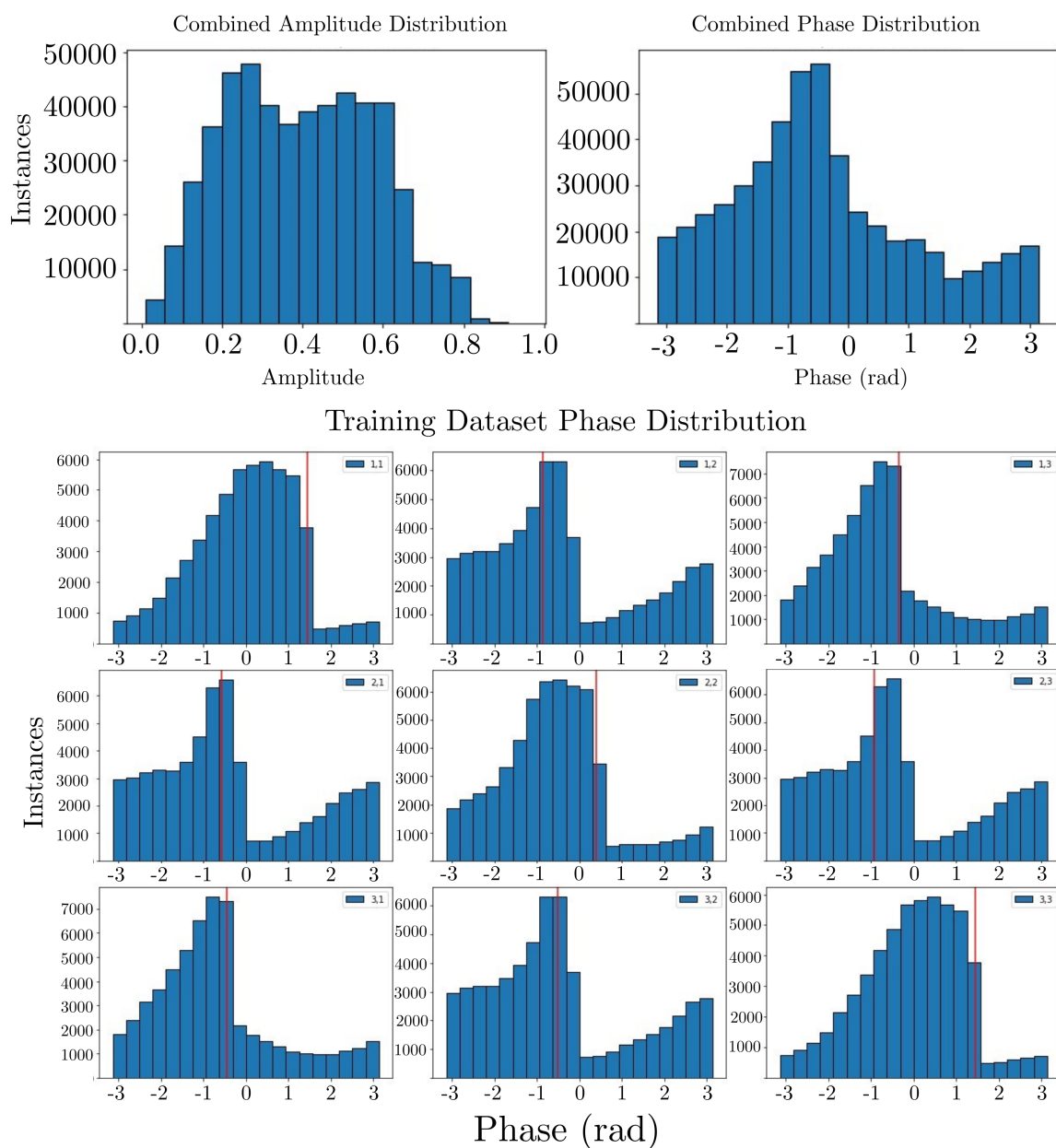


Fig. S10. Phase and intensity distributions across each matrix element in the training dataset. Amplitude values show a mostly normal distribution centered around 0.33 as we would expect for a  $n=3$  waveguide array. Phase values are not addressed in data generation and therefore tend to smaller phase shifts as these will introduce fewer losses. Below we see the phase distribution per matrix element further demonstrating this behaviour. The red line indicated the phase of an unperturbed device, where it should be noted phase is wrapped at  $2\pi$  so the ridge in instances to the right hand side is in fact the end of the tail to the left.