Research article

Sean Hooten*, Raymond G. Beausoleil and Thomas Van Vaerenbergh

Inverse design of grating couplers using the policy gradient method from reinforcement learning

https://doi.org/10.1515/nanoph-2021-0332
Received July 1, 2021; accepted September 19, 2021;
published online October 7, 2021

Abstract: We present a proof-of-concept technique for the inverse design of electromagnetic devices motivated by the policy gradient method in reinforcement learning, named PHORCED (**PH**otonic **O**ptimization using **R**EINFORCE **C**riteria for **E**nhanced **D**esign). This technique uses a probabilistic generative neural network interfaced with an electromagnetic solver to assist in the design of photonic devices, such as grating couplers. We show that PHORCED obtains better performing grating coupler designs than local gradient-based inverse design via the adjoint method, while potentially providing faster convergence over competing state-of-the-art generative methods. As a further example of the benefits of this method, we implement transfer learning with PHORCED, demonstrating that a neural network trained to optimize 8° grating couplers can then be re-trained on grating couplers with alternate scattering angles while requiring >10× fewer simulations than control cases.

Keywords: adjoint method; deep learning; integrated photonics; inverse design; optimization; reinforcement learning.

***Corresponding author**: **Sean Hooten**, Hewlett Packard Labs, Hewlett Packard Enterprise, Milpitas, CA 95035, USA; and Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA 94720, USA, E-mail: sean.hooten@hpe.com. https://orcid.org/0000-0003-1260-412X
Raymond G. Beausoleil, Hewlett Packard Labs, Hewlett Packard Enterprise, Milpitas, CA 95035, USA,
E-mail: ray.beausoleil@hpe.com
Thomas Van Vaerenbergh, Hewlett Packard Labs, HPE Belgium, B-1831 Diegem, Belgium, E-mail: thomas.vanvaerenbergh@hpe.com. https://orcid.org/0000-0002-7301-8610

1 Introduction

There has been a recent, massive surge in research syncretizing topics in photonics and artificial intelligence/machine learning (AI/ML), including photonic analog accelerators [1–8], physics emulators [9–15], and AI/ML-enhanced inverse electromagnetic design techniques [15–35]. While inverse electromagnetic design via local gradient-based optimization with the adjoint method has been successfully applied to a multitude of design problems throughout the entirety of the optics and photonics communities [36–60], inverse electromagnetic design leveraging AI/ML techniques promise superior computational performance, advanced data analysis and insight, or improved effort towards global optimization. For the lattermost topic in particular, Jiang and Fan recently introduced an unsupervised learning technique called GLOnet which uses a generative neural network interfaced with an electromagnetic solver to design photonic devices such as metasurfaces and distributed Bragg reflectors [24–26]. In this paper we propose a conceptually similar design technique, but with a contrasting theoretical implementation motivated by a concept in reinforcement learning called the policy gradient method – specifically a one-step implementation of the REINFORCE algorithm [61, 62]. We will refer to our technique as PHORCED = **PH**otonic **O**ptimization using **R**EINFORCE **C**riteria for **E**nhanced **D**esign. PHORCED is compatible with any external physics solver including EMopt [63], a versatile electromagnetic optimization package that is employed in this work to perform 2D simulations of single-polarization grating couplers.

In Section 2, we will qualitatively compare and contrast three optimization techniques: local gradient-based optimization (e.g., gradient ascent), GLOnet, and PHORCED. We are specifically interested in a proof-of-concept demonstration of the PHORCED optimization technique applied to grating couplers, which we present in Section 3. We find that both our implementation of the

GLOnet method and PHORCED find better grating coupler designs than local gradient-based optimization, but PHORCED requires fewer electromagnetic simulation evaluations than GLOnet. Finally, in Section 4 we introduce the concept of transfer learning to integrated photonic optimization, where in our application we demonstrate that a neural network trained to design 8° grating couplers with the PHORCED method can be re-trained to design grating couplers that scatter at alternate angles with greatly accelerated time-to-convergence. We speculate that a hierarchical optimization protocol leveraging this technique can be used to design computationally complex devices while minimizing computational overhead.

2 Extending the adjoint method with neural networks

Local gradient-based optimization using the adjoint method has been successfully applied to the design of a plethora of electromagnetic devices. Detailed tutorials of the adjoint method applied to electromagnetic optimization may be found in Refs. [36–39, 43, 46, 60]. Here, we qualitatively illustrate a conventional design loop utilizing the adjoint method in Figure 1(a). This begins with the choice of an initial electromagnetic structure parameterized by vector $\mathbf{p}$, which might represent geometrical degrees-of-freedom like the width and height of the device. This is fed into an electromagnetic solver, denoted by a function g. The resulting electric and magnetic fields, $\mathbf{x} = g(\mathbf{p})$, can be used to evaluate a user-defined electromagnetic *merit function* $f(\mathbf{x})$ – the metric that we are interested in optimizing (e.g., coupling efficiency). Gradient-based optimization seeks to improve the value of the merit function by updating the design parameters, $\mathbf{p}$, in a direction specified by the gradient of the electromagnetic merit function, $\frac{\partial(f \circ g)}{\partial \mathbf{p}}$. The value of the gradient may be obtained very efficiently using the adjoint method, requiring just two electromagnetic simulations regardless of the number of degrees-of-freedom (called the forward simulation and adjoint simulation, respectively). A single iteration of gradient-based optimization is depicted visually in the center of Figure 1(a), where p is a single-dimension point sampled along a toy merit function $(f \circ g)(p)$ representing the optimization landscape (which is unknown, *a priori*). The derivative (gradient) of the merit function is illustrated by the arrow pointing from p in the direction of steepest ascent (assuming this is a maximization problem). During an optimization, we slowly update p in this direction until a local optimum is reached.

The adjoint method chain-rule derivative of the electromagnetic merit function resembles the concept of *backpropagation* in deep learning, where a neural network's weights can be updated efficiently by application of the chain-rule with information from the forward pass. Naturally, we might extend the functionality of the adjoint method by placing a neural network in the design loop. The neural network takes the place of a deterministic update algorithm (such as gradient ascent), potentially learning information or patterns in the design problem that allows it to find a better optimum. Below we present two methods to implement inverse design with neural networks: GLOnet (introduced by Jiang and Fan [24, 25]) and PHORCED (this work). Both methods are qualitatively similar, but differ in the representation of the neural network. In the main text of this manuscript we will qualitatively describe the differences between these techniques; a detailed mathematical discussion may be found in Supplementary Material Section 1.

The GLOnet optimization method is depicted qualitatively in Figure 1(b). The neural network is represented as a deterministic function h_θ that takes in noise $\mathbf{z}$ from some known distribution D and outputs design parameters $\mathbf{p}$. Importantly, the neural network is parameterized by programmable weights $\boldsymbol{\theta}$ that we intend to optimize in order to generate progressively better electromagnetic devices. Similar to regular gradient ascent, we may evaluate the electromagnetic merit function of a device generated by the neural network $(f \circ g)(\mathbf{p})$ using our physics solver, and find its gradient with respect to the design parameters using the adjoint method, $\frac{\partial(f \circ g)}{\partial \mathbf{p}}$. However, the GLOnet design problem is inherently stochastic because of the presence of noise, and therefore the optimization objective becomes the expected value of the electromagnetic merit function, $\mathbb{E}_{\mathbf{z}}[(f \circ g \circ h_\theta)(\mathbf{z})]$ – sometimes called the *reward* in the reinforcement learning literature.[1] In practice this expression can be approximated by taking a simple average over the electromagnetic merit functions of several devices generated by the neural network per iteration. The gradient that is then backpropagated to the neural network is given by the expected value of the chain-rule gradient of the reward function. The first term, $\frac{\partial(f \circ g)}{\partial \mathbf{p}}$, can once again be computed very efficiently using the

1 Note that we have written a generalized version of the reward function defined in the original works by Jiang and Fan [24, 25]. In that case, the reward function is chosen to weight good devices exponentially, i.e. $f \to \exp(f/\sigma)$ where σ is a hyperparameter and f is the electromagnetic quantity of interest. The full function is defined in Eq. (S.18) of Supplementary Materials Section 1.

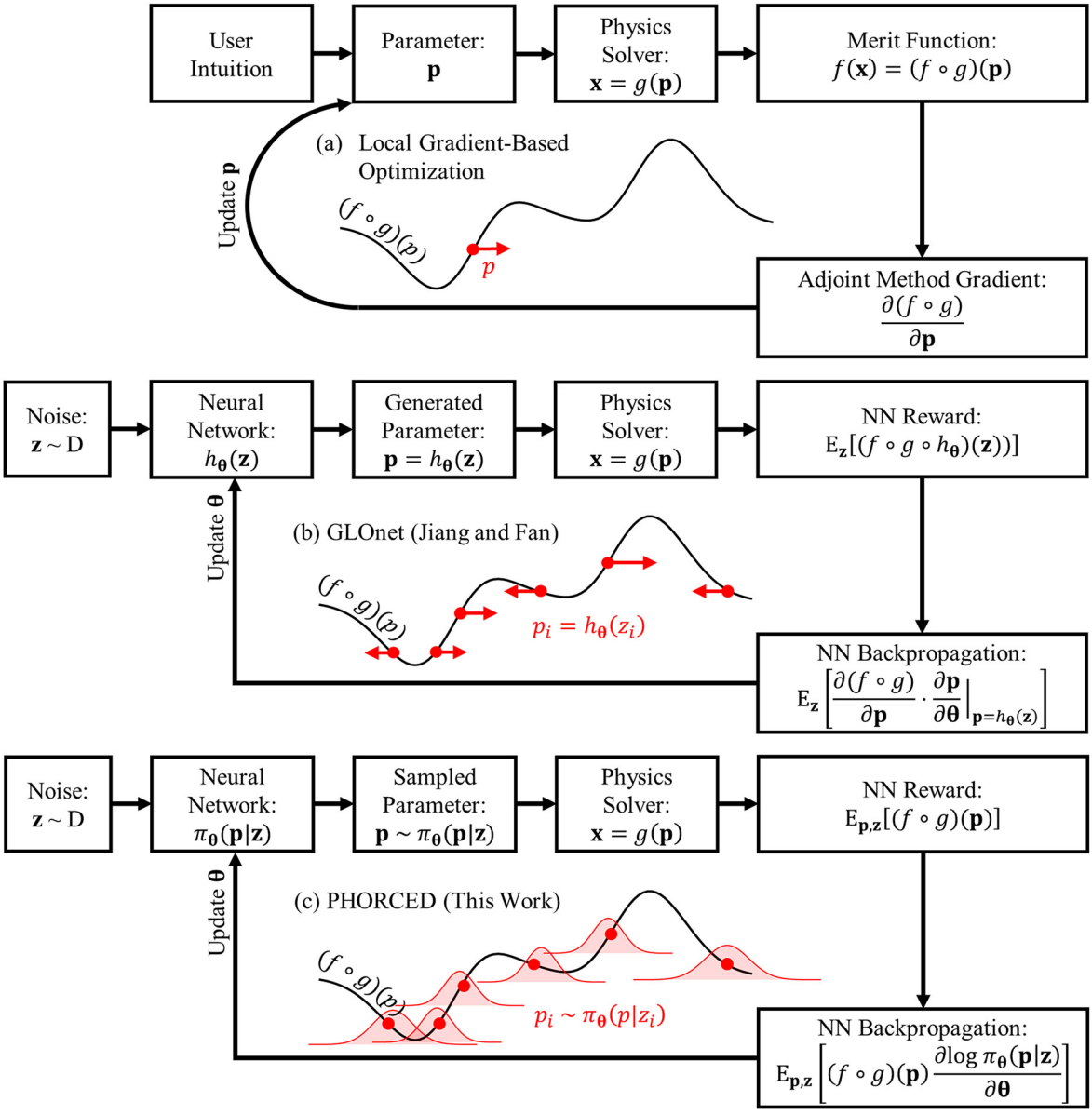

Figure 1: Neural networks provide a natural extension to conventional inverse design via the adjoint method. A typical gradient-based design loop is shown in (a) where the derivatives are calculated using the adjoint method. The GLOnet method (b), originally proposed by Jiang and Fan [24, 25], replaces a conventional gradient-based optimization algorithm with a deterministic neural network. In this work we propose PHORCED (c), which uses a probabilistic neural network to generate devices. (b) and (c) are qualitatively similar, but require different gradients in backpropagation because of the representation of the neural network (deterministic versus probabilistic). In particular, notice that PHORCED does not require an evaluation of the adjoint method gradient of the electromagnetic merit function, $\frac{\partial(f \circ g)}{\partial \mathbf{p}}$.

adjoint method, requiring just two electromagnetic simulations per device. Meanwhile the latter term, $\frac{\partial \mathbf{p}}{\partial \theta}|_{\mathbf{p}=h_\theta(\mathbf{z})}$, can be calculated internal to the neural network using conventional backpropagation with automatic differentiation. Visually, one iteration of the GLOnet method is shown in the center of Figure 1(b). In each iteration, the neural network in the GLOnet method suggests parameters, p_i, which are then individually simulated. Similar to gradient-based optimization from Figure 1(a), we find the gradient of the merit function value with respect to each generated design parameter, represented by the arrows pointing towards the direction of steepest ascent at each point. The net gradient information from many simulated devices effectively tells the neural network where to explore in the next iteration.

With a dense search, the global optimum along the domain of interest can potentially be found.

Our technique, called PHORCED, is provided in Figure 1(c). Qualitatively speaking, it is very similar to GLOnet, but is motivated differently from a mathematical perspective. PHORCED is a special case of the REINFORCE algorithm [61, 62] from the field of reinforcement learning (RL), applied to electromagnetic inverse design. In particular, the neural network is treated as purely probabilistic, defining a continuous conditional probability density function over parameter variables $\mathbf{p}$ conditioned on the input vector $\mathbf{z}$ – denoted by $\pi_\theta(\mathbf{p}|\mathbf{z})$. In other words, instead of outputting $\mathbf{p}$ deterministically given input noise, the neural network outputs probabilities of generating $\mathbf{p}$. We then randomly sample a parameter vector $\mathbf{p}$ for simulation and evaluation of the reward. Note that $\pi_\theta(\mathbf{p}|\mathbf{z})$ is called the *policy* in RL, and in this work is chosen to be a multivariate Gaussian distribution with mean vector and standard deviation of random variable $\mathbf{p}$ as outputs. The reward for PHORCED is qualitatively the same as GLOnet – namely, we intend to optimize the expected value of the electromagnetic merit function. However, because both $\mathbf{p}$ and $\mathbf{z}$ are random variables, we take the joint expected value: $\mathbb{E}_{\mathbf{p},\mathbf{z}}[(f \circ g)(\mathbf{p})]$. Furthermore, because of the probabilistic representation of the neural network, the gradient of the reward with respect to the neural net weights for backpropagation is much different than the corresponding GLOnet case. In particular, we find that the backpropagated chain-rule gradient requires **no evaluation** of the gradient of the electromagnetic merit function, $\frac{\partial(f \circ g)}{\partial \mathbf{p}}$. Consequently, the electromagnetic adjoint simulation is no longer required, implying that fewer simulations are required overall for PHORCED compared to GLOnet under equivalent choices of neural network architecture and hyperparameters.[2] PHORCED is visually illustrated in the center of Figure 1(c). The neural network defines Gaussian probability density functions conditioned on input noise vectors z_i, shown in the light red bell curves representing $\pi_\theta(p|z_i)$, from which we sample points p_i to simulate.[3] Using information from the merit function values, the neural network learns to update the mean and standard deviation of the Gaussians. Consequently, we emphasize that the Gaussian policy distribution is not static because its statistical parameters are adjusted by the trainable weights of the neural network, and is therefore capable of exploring throughout the feasible design space. For adequate choice of distribution and dense enough search, the PHORCED method can potentially find the global optimum in the domain of interest.

Before proceeding it should be remarked that the algorithms implemented by GLOnet and PHORCED have precedent in the literature, with some distinctions that we will outline here. Optimization algorithms similar to GLOnet were suggested in Refs. [64, 65], where the main algorithmic difference appears in the definition of the reward function. In particular, the reward defined in Ref. [64] was the same generalized form that we have presented in Figure 1(b), while Jiang and Fan emphasized the use of an exponentially-weighted reward to enhance global optimization efforts [25]. On the other hand, PHORCED was motivated as a special case of the REINFORCE algorithm [61, 62], but also resembles some versions of evolutionary strategy [64–67]. The main difference between PHORCED and evolutionary strategy (ignoring several heuristics) is the explicit use of a neural network to model the multivariate Gaussian policy distribution, albeit some recent works have used neural networks in their implementations of evolutionary strategy [65, 67] for different applications than those studied here. Furthermore, PHORCED does not require a Gaussian policy; any explicitly-defined probability distribution can be used as an alternative if desired. Beyond evolutionary strategy, a recent work in fluid dynamics [68] uses an algorithm akin to PHORCED called One-Step Proximal Policy Optimization (PPO-1) – a version of REINFORCE with a single policy, π_θ, operating on parallel instances of the optimization problem. The main distinction between PPO-1 and PHORCED is that we have implemented the option to use an input noise vector $\mathbf{z}$ to condition the output policy distribution, $\pi_\theta(\mathbf{x}|\mathbf{z})$, which can effectively instantiate multiple distinct policies acting on parallel instances of the optimization problem. This potentially enables multi-modal exploration of the parameter space, bypassing a known issue of Bayesian optimization with Gaussian probability distributions [65]. However, note that the best results for the applications studied in this work used a constant input vector $\mathbf{z}$, thus

2 However, because the representation of the neural network is different in either case, it would rarely make sense to use equivalent choices of neural network architecture and hyperparameters. Therefore, we make this claim tepidly, emphasizing only that we do not require adjoint simulations in the evaluation of the reward.

3 Note that while we explored a uniform distribution at the input as well, our best results with PHORCED applied to grating coupler optimization in this work were attained with $\mathbf{z}$ drawn from a Dirac delta distribution, i.e. a constant vector input rather than noise. This has the effect of collapsing the multiple distributions depicted in Figure 1(c) into a single distribution from which we draw multiple samples.

making our implementation similar to PPO-1 in the results below. Regardless of the intricacies mentioned above, we emphasize that both GLOnet and PHORCED are unique in their application to electromagnetic optimization, to the authors' knowledge. In the next section we will compare all three algorithms from Figure 1 applied to grating coupler optimization.

3 Proof-of-concept grating coupler optimization

A grating coupler is a passive photonic device that is capable of efficiently diffracting light from an on-chip waveguide to an external optical fiber. Recent works have leveraged inverse design techniques in the engineering of grating couplers, resulting in state-of-the-art characteristics [21, 42, 44, 52, 56]. In this section we will show how generative neural networks can aid in the design of ultra-efficient grating couplers.

The grating coupler geometry used for our proof-of-concept is depicted in Figure 2(a). We assume a silicon-on-insulator (SOI) wafer platform with a 280 nm-thick silicon waveguide separated from the silicon substrate by a 2 μm buried oxide (BOX) layer. The grating coupler consists of periodically spaced corrugations to the input silicon waveguide with etch depth 190 nm. These grating coupler dimensions are characteristic of a high-efficiency integrated photonics platform operating in the C-band (1550 nm central wavelength) as suggested by a transfer matrix based directionality calculation [55, 69], and are provided as an example in the open-source electromagnetic optimization package EMopt [63] that was used to perform forward and adjoint simulations in this work.

For our optimizations, we will consider 60 total designable parameters that define the grating coupler: the width of and spacing between 30 waveguide corrugations. For a well-designed grating coupler, input light to the waveguide scatters at some angle relative to vertical toward an external optical fiber. In this work we choose to optimize the grating coupler at fixed wavelength and scattering angle assuming specific manufacturing and assembly requirements for a modular optical transceiver application [52, 70]. In our case, the merit function for optimization is the coupling efficiency of scattered light with wavelength $\lambda = 1550$ nm propagating 8° relative to normal, mode-matched to an output Gaussian beam mode field diameter 10.4 μm – characteristic of a typical optical fiber mode. The explicit definition of this electromagnetic merit function may be found in Refs. [44, 52, 71]. Note that we did not include fabrication constraints nor other specifications of interest in grating couplers in our parameterization choice, e.g. the BOX thickness and the Si etch depth, which will be desirable in future optimizations of experimentally-viable devices. Furthermore, the simulation domain is discretized with a $dx = 25$ nm grid step which may result in some inaccuracy for very fine grating coupler features. This simulation discretization was chosen for feasibility of the optimization since individual simulations require about 4 s to compute on a high-performance server with over 30 concurrent MPI processes, and as we will show GLOnet and PHORCED can require as many as 20,000 simulation evaluations for convergence. Nevertheless, we utilized permittivity smoothing [43] to minimize the severity of this effect and obtain physically meaningful results.

We apply the Broyden–Fletcher–Goldfarb–Shanno (BFGS), GLOnet, and PHORCED algorithms to grating coupler optimization with two different initial designs (Initial Seed 1 and Initial Seed 2) in Figure 2(b) and (c). Initial Seed 1 corresponds to the grating depicted in Figure 2(a) and the top of Figure 2(b), where we used a parameter sweep to choose a linear apodization of the etch duty cycle before optimization. Initial Seed 2 corresponds to the grating shown at the top of Figure 2(c) where we use a uniform duty cycle of 90%. Both initial designs have pitch that satisfy the grating equation [44, 52] for 8° scattering. These initial seeds serve to explore the robustness of the optimization algorithms to "good" and "poor" choices of initial condition. Indeed, Initial Seed 1 satisfies physical intuition for a good grating coupler, because chirping the duty cycle is well-known to improve Gaussian beam mode-matching, and thus the initial grating coupler efficiency is already a reasonable value of 56%. Meanwhile, Initial Seed 2 has the correct pitch for 8° scattering, but has a low efficiency of 1% owing to poor mode-matching and directionality. In essence, we are using these two cases as a proxy to explore whether PHORCED and GLOnet can reliably boost electromagnetic performance in a high-dimensional parameter space, even in cases where the intuition about the optimal initial conditions is limited.

BFGS is a conventional gradient-based optimization algorithm similar to that depicted in Figure 1(a), and implemented using default settings from the open-source SciPy optimize module. After optimization with BFGS, the final simulated grating coupler efficiency of Initial Seed 1 and Initial Seed 2 are 86.4 and 69.9% respectively, which are shown in the black dashed lines of Figure 2(b) and (c). Note that the number of simulation calls for BFGS is not shown because it is vastly smaller than that required for

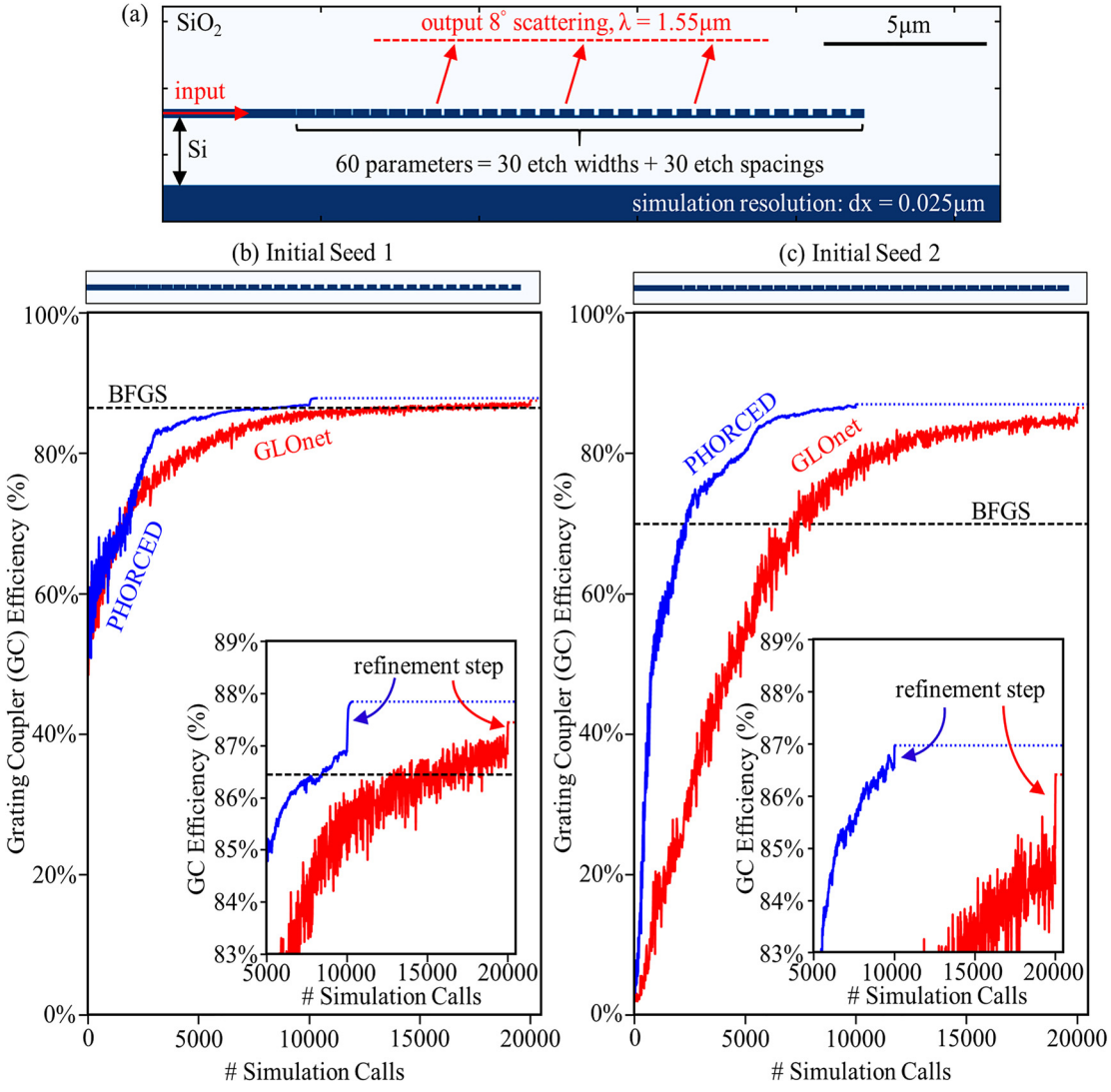

Figure 2: PHORCED and GLOnet outperform conventional gradient-based optimization, with different simulation evaluation requirements. The grating coupler simulation geometry for optimization is shown in (a), consisting of an SOI wafer platform with 280 nm waveguide, 190 nm etch depth, and 2 μm BOX height. We optimized 60 device parameters in total, namely the width and spacing between 30 etch corrugations. The results of the BFGS, GLOnet, and PHORCED algorithms applied to Initial Seed 1 and Initial Seed 2 are presented in (b) and (c) as a function of the number of simulation calls. The initial seeds (grating designs) are illustrated above each optimization plot. The insets depict zoomed-in views of the peak efficiencies attained by PHORCED and GLOnet, along with respective BFGS refinement steps performed on the best design generated by each algorithm.

GLOnet and PHORCED (144 and 214 total simulations for Initial Seed 1 and Initial Seed 2, respectively).

The implementation of GLOnet and PHORCED for the grating coupler optimizations in Figure 2(b) and (c) are described below; other details and specifications, such as a graphical illustration of the neural network models used in either case, may be found in Supplementary Materials Section 2. GLOnet is described qualitatively in Figure 1(b) where we use a deterministic neural network and an exponentially-weighted electromagnetic merit function originally recommended by Jiang and Fan [24, 25] with a chosen hyperparameter $\sigma = 0.6$. PHORCED is described qualitatively in Figure 1(c) where we use a probabilistic neural network modeling a multivariate, isotropic Gaussian output distribution. The electromagnetic merit function used in the reward is just the unweighted grating coupler efficiency, except we used a "baseline" subtraction of the average merit function value in the backpropagated gradient (which is a common tactic in reinforcement learning for reducing model variance [72]). In both methods, we use a stopping criterion of 1000 total optimizer iterations, with 10 devices sampled per

iteration. Note that because GLOnet requires an adjoint simulation for each device and PHORCED does not, the effective stopping criteria are $1000 \times 10 \times 2 = 20,000$ simulation calls for GLOnet and $1000 \times 10 \times 1 = 10,000$ simulation calls for PHORCED. The neural network models for GLOnet and PHORCED were implemented in PyTorch, and are illustrated in Figure S1 of the Supplementary Material. For GLOnet, we used a convolutional neural network with ReLU activations and linear output of the design vector. For PHORCED, we used a simple fully-connected neural network with ReLU activations and linear output defining the statistical parameters (mean and variance) of the Gaussian policy distribution. We applied initial conditions by adding the design vector representing Initial Seed 1 and Initial Seed 2 to either the direct neural network output or the mean of the Gaussian distribution for GLOnet and PHORCED, respectively. Hyperparameters such as the number of weights and learning rate of the neural network optimizer were individually tuned for GLOnet and PHORCED (specifications can be found in Figure S2 of the Supplemental Material). For both neural networks, we used an input vector $\mathbf{z}$ of dimension 5. However, for GLOnet $\mathbf{z}$ was drawn from a uniform distribution $\mathcal{U}(-1, 1)$, while for PHORCED we used a simple Dirac delta distribution centered on 1. In other words, the input for PHORCED was a constant vector of 1's. In this work we achieved our best results with this choice, but similar results (within 1% absolute grating coupler efficiency) could be attained with alternative choices of input distribution. Anecdotally, we found that using a noisy input could improve training stability and performance in toy problems studied outside of this work, but further investigation is required.

We find that PHORCED and GLOnet outperform regular BFGS for both initial conditions studied in Figure 2. For Initial Seed 1 (Figure 2(b)), we generated optimized grating coupler efficiencies of 86.9 and 87.2% for PHORCED and GLOnet, respectively. For Initial Seed 2 (Figure 2(c)) we find optimized grating coupler efficiencies of 86.8 and 85.6% for PHORCED and GLOnet, respectively. Furthermore, as shown in the insets of Figure 2(b) and (c), we were able to marginally improve each of the results by applying a BFGS "refinement step" to the best performing design output from GLOnet and PHORCED. This refinement step was limited to a maximum of 200 iterations, or until another default convergence criterion was met. For Initial Seed 1 we obtained improvements of $\{86.9\% \rightarrow 87.8\%\}/\{87.2\% \rightarrow 87.4\%\}$ for PHORCED/GLOnet, respectively. For Initial Seed 2, we find improvements of $\{86.8\% \rightarrow 87.0\%\}/\{85.6\% \rightarrow 86.4\%\}$ for PHORCED/GLOnet, respectively. Since PHORCED and GLOnet are inherently statistical and noisy, the refinement step is useful for finding the nearest optimum without requiring one to run an exhaustive search of the neural network generator in inference mode.

In summary, we find that the PHORCED + BFGS refinement optimization achieved the best performance for both Initial Seed 1 and Initial Seed 2 with final grating coupler efficiency of 87.8 and 87.0%, respectively. These results agree with a transfer matrix based directionality analysis of these grating coupler dimensions [55, 69], where we find that approximately 88% grating efficiency is possible under perfect mode-matching conditions – meaning that our result for Initial Seed 1 is close to a theoretical global optimum. Notably, GLOnet had better performance than PHORCED in Initial Seed 1 before the refinement step was applied, and it is possible that better results could have been achieved with further iterations of both algorithms (as indicated by the slowly rising slopes of the optimization curves in the insets of Figure 2(b) and (c)). However, we emphasize that PHORCED required approximately 2× fewer simulation than GLOnet with the same number of optimization iterations because of the lack of adjoint gradient calculations.

Perhaps the most important result of these optimizations is that both PHORCED and GLOnet proved to be resilient against our choice of initial condition. Indeed, while BFGS provided a competitive result for Initial Seed 1, it failed to find a favorable local optimum given Initial Seed 2. PHORCED and GLOnet, on the other hand, attained final results within 1% absolute grating coupler efficiency for both initial conditions. This outcome is promising when considering the relative sparsity of each algorithm's search in a high-dimensional design space. Indeed, as mentioned previously, we only sampled 10 devices per iteration of the optimization, meaning that there were fewer samples than dimension of the resulting parameter vectors (60). As an additional reference, we compared our results to CMA-ES (implemented with the open-source package pycma), a popular "blackbox" global optimization algorithm that is known to be effective for high-dimensional optimization [66]. Under the same number of simulation calls as PHORCED (10,000), CMA-ES reached efficiencies of 87.3 and 86.7% for Initial Seed 1 and Initial Seed 2, respectively. Therefore our implementations of PHORCED and GLOnet are competitive with current state-of-the-art blackbox algorithms. While the simulation requirements for convergence of PHORCED and GLOnet remain computationally prohibitive for more complex electromagnetic structures than those studied here, in contrast to local gradient-based

search, our results offer the possibility of global optimization effort in electromagnetic design problems, where we are capable of limiting the tradeoff of performance with search density as well as the need for human intervention in situations where physical intuition is more difficult to ascertain. Furthermore, we believe that our implementations of GLOnet and PHORCED have significant room for improvement, and have advantages that go beyond alternative global optimization algorithms like CMA-ES. In particular, leveraging advanced concepts in deep learning and reinforcement learning can further improve computational efficiency and performance. For example, whereas our implementations of PHORCED and GLOnet used simple neural networks, a ResNet architecture can improve neural network generalizability while simultaneously reducing overfitting [27]. Moreover, one could take advantage of complementary deep learning and reinforcement learning based approaches such as importance sampling [65, 72], or "model-based" methods that could utilize an electromagnetic surrogate model or inverse model to reduce the number of full electromagnetic Maxwell simulations needed for training [13, 19, 28, 32]. Alternatively, whereas we optimized our grating for a single objective (single wavelength, 8° scattering) Jiang and Fan showed that generative neural network based optimization can be extremely efficient for multi-objective design problems (e.g., multiple wavelengths and scattering angles in metasurfaces [24]). Along the same vein, in the next section we will show how a technique known as transfer learning can be used to repurpose a neural network trained with PHORCED for an alternative objective, meanwhile boosting computational efficiency and electromagnetic performance dramatically.

4 Transfer learning with the PHORCED method

Transfer learning is a concept in machine learning encompassing any method that reuses a model (e.g. a neural network) trained on one task for a different but related task. Qualitatively speaking and verified by real-world applications, we might expect the retraining of a neural network to occur faster than training a new model from scratch. Transfer learning has been extensively applied in classical machine learning tasks such as classification and regression, but has only recently been mentioned in the optics/photonics research domains [17, 27, 28, 33]. In this work we apply transfer learning to the inverse design of integrated photonics for the first time (to the authors' knowledge), revealing that a neural network trained using PHORCED for the design of 8° grating couplers can be retrained to design grating couplers with varied scattering angle and increased rate of convergence.

Transfer learning applied to grating coupler optimization is qualitatively illustrated in Figure 3(a) and (b). Figure 3(a) shows a shorthand version of the PHORCED optimization of Initial Seed 1 that was performed in Figure 2(b), where a neural network was specifically trained to design an 8° grating coupler. In the case of transfer learning in Figure 3(b), we reuse the trained neural network from Figure 3(a) but now exchange the 8° angle in the grating coupler efficiency merit function with an alternate scattering angle. In particular, we retrain the neural network on six alternative grating coupler angles: $\{2°, 4°, 6°, 10°, 12°, 14°\}$. Note that we maintained the exact same neural network architecture and optimization hyperparameters during these exchanges, including the optimization stopping criterion of 10,000 total simulation calls per training session; the only change in a given optimization was the grating coupler angle. As depicted in Figure 3(c), we show the optimization progressions of these transfer learning sessions (blue/red curves) in comparison to the original PHORCED optimization of the 8° grating coupler from Figure 2(b) (reproduced in the black curve in the middle panel). Also shown are control optimizations for each grating coupler angle using the PHORCED method without transfer learning (in gray). We find that transfer learning to grating couplers with nearby scattering angles (e.g. 6° and 10°) exhibit extremely accelerated rate of convergence relative to the original optimization and control cases. However, transfer learning is less effective or ineffective for more distant angles (e.g. 2°, 4°, and 14°). This observation is shown more clearly in Figure 3(d) where we plot the number of simulations required to reach 80% efficiency in optimization versus the scattering angle for retraining.[4] While the original optimization and control optimizations (black star and gray diamonds) required several thousand simulation calls before reaching this threshold, the 6° and 10° transfer learning optimizations required only about 100 simulations a piece – a >10× reduction in simulation calls, making transfer learning comparable with local gradient-based optimization in terms of computation requirements. On the other hand, the distant grating coupler angle transfer learning optimizations (2°, 4°, and 14°) required similar simulation call requirements to reach the same threshold as the original optimization. Evidently,

4 80% grating coupler efficiency was chosen because it equates to roughly 1 dB insertion loss – an optimistic target for state-of-the-art silicon photonic devices.

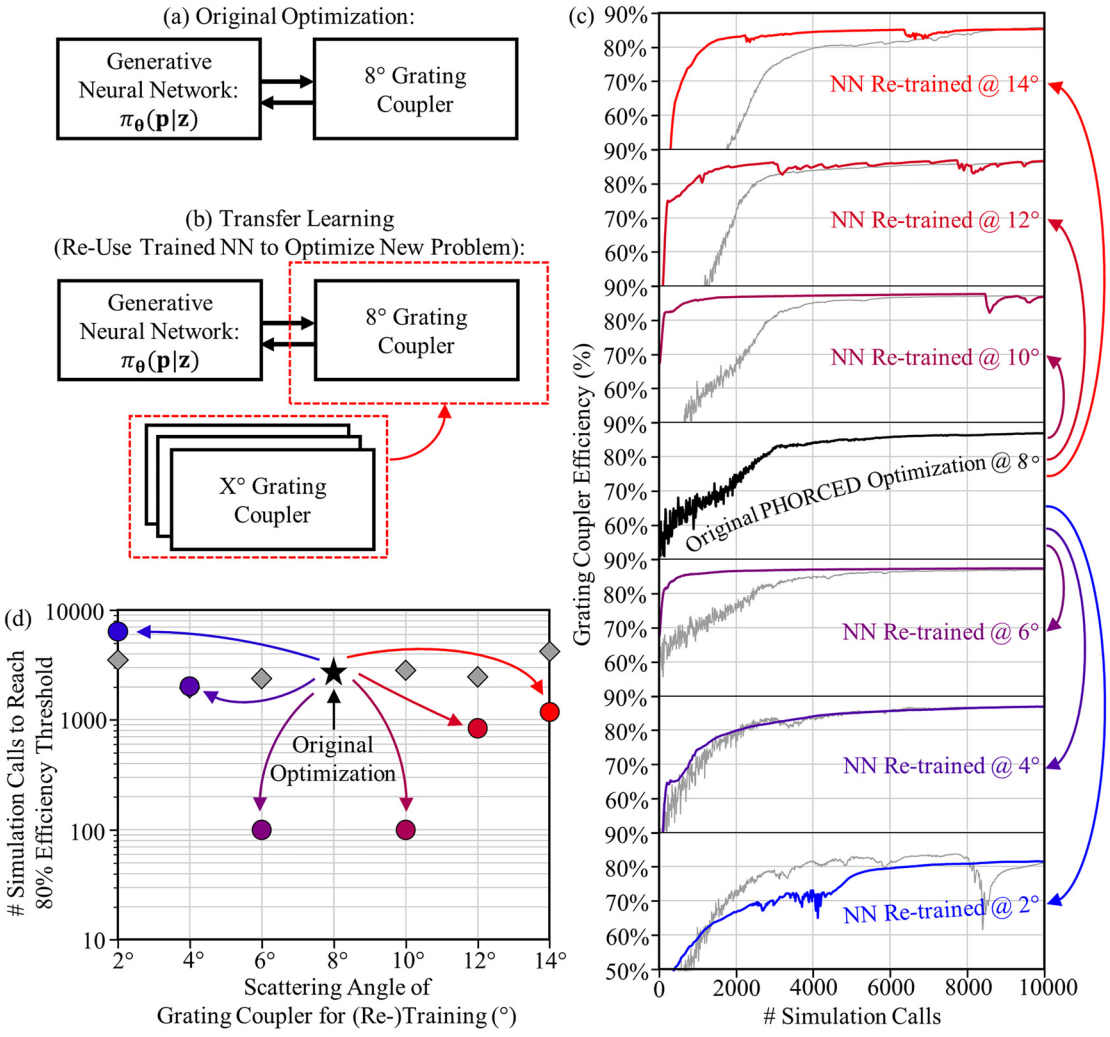

Figure 3: Transfer learning applied to grating coupler design yields accelerated convergence rate in optimization. The original PHORCED optimization from Figure 2(b) is qualitatively depicted as a block diagram in (a) for comparison with the transfer learning approach in (b). Here, we exchange the 8° grating coupler merit function with an alternative grating coupler angle for retraining. Optimization progressions as a function of the number of simulation calls for each of the retraining sessions are shown in the blue/red curves of (c). Control optimizations where transfer learning was not applied are plotted in gray for comparison. In (d) we plot the number of simulation calls for each optimization from (c) to reach 80% grating coupler scattering efficiency, with blue/red colored arrows and dots indicating applications of transfer learning and gray diamonds indicating the control cases.

there is a bias towards less effective transfer learning for small scattering angles. Grating couplers become plagued by parasitic back-reflection for small diffraction angles relative to normal [44, 52], and thus it is possible that the neural network has difficulty adapting to the new physics that were not previously encountered. We conclude that the transfer learning approach is most effective for devices with very similar physics to the device originally optimized by the neural network.

The results of Figure 3 lead to a natural follow-up query: can we apply transfer learning multiple times progressively in order to maintain the convergence rate advantage for optimizations at more distant grating coupler angles? We explore this question of *sequential* transfer learning in Figure 4. In Figure 4(a) and (b) we qualitatively compare sequential transfer learning to the original PHORCED optimization from Figure 2(b). As indicated, we replace the original 8° grating coupler scattering angle in the electromagnetic merit function with an alternative scattering angle in the same manner discussed in Figure 3(b). Then, after that optimization has completed, we continue to iterate and exchange the grating coupler angle again. By sequentially applying transfer learning, we hope to slowly introduce new physics

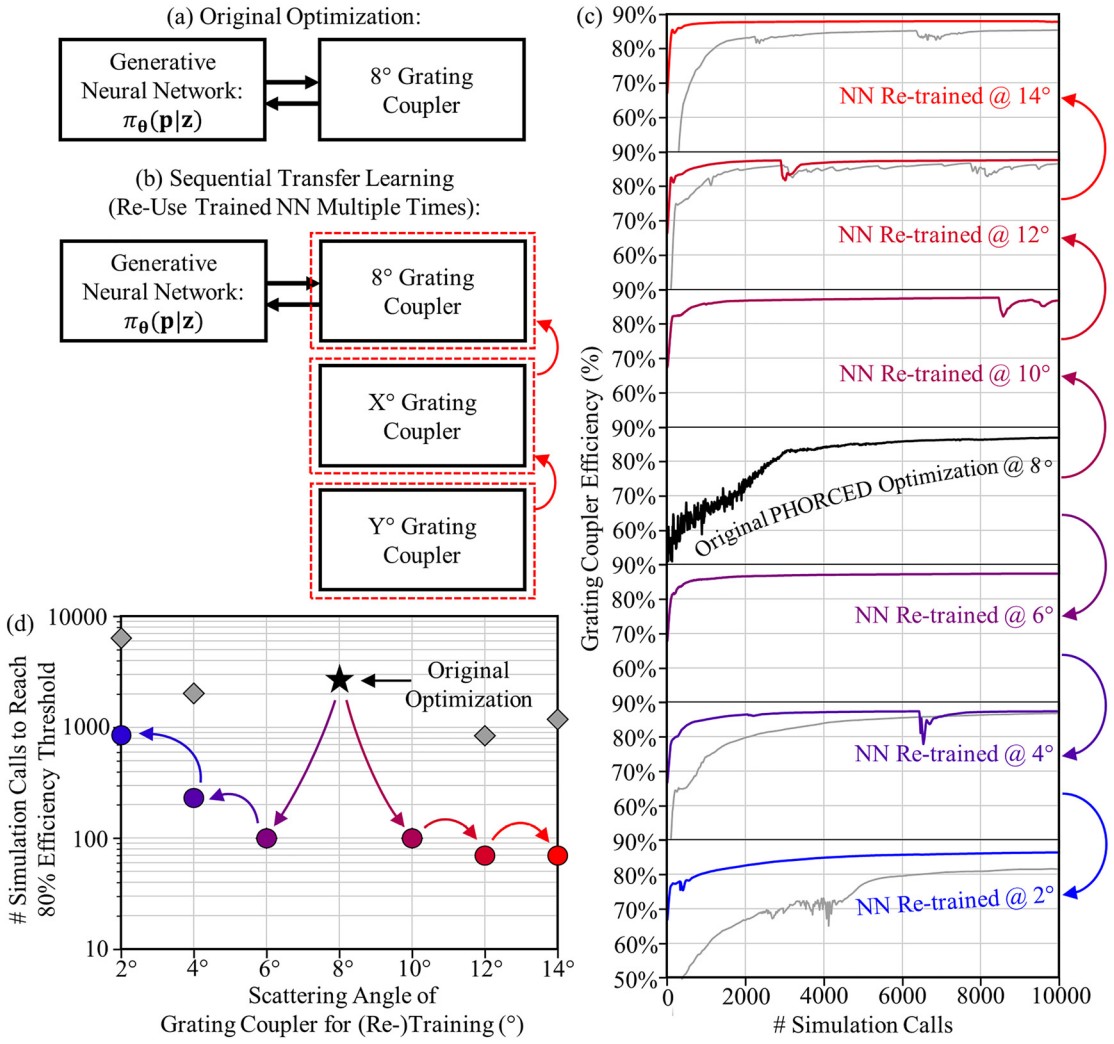

Figure 4: Transfer learning applied sequentially improves convergence for grating coupler scattering angles that are distant relative to the original optimization's scattering angle. The original PHORCED optimization from Figure 2(b) is qualitatively depicted in (a) for comparison with the "sequential" transfer learning approach in (b). Here, we sequentially re-train the neural network (originally trained to generate 8° grating couplers) with progressively different scattering angles, with the intention of slowly changing the physics seen by the neural network. Grating coupler efficiency as a function of the number of simulation calls for each of the retraining sessions are shown in the blue/red curves of (c), where the arrows on the right-hand side show the sequence of each application of transfer learning. The results of the one-shot (non-sequential) transfer learning approach from Figure 3(c) are shown in gray for comparison. In (d) we plot the number of simulation calls needed for the grating coupler optimizations from (c) to reach an 80% efficiency using the sequential transfer learning approach, similar to the corresponding plot in Figure 3(d). The blue/red arrows and dots indicate applications of sequential transfer learning, and the gray diamonds correspond to non-sequential transfer learning cases.

to the neural network such that we can maintain faster convergence at more physically distant problems from the initial optimization. We conduct two sequential transfer learning sessions where we evolve the grating coupler scattering angle in the following steps: $\{8° \to 10° \to 12° \to 14°\}$ and $\{8° \to 6° \to 4° \to 2°\}$. The results of these progressions are shown in Figure 4(c) where we plot the grating coupler efficiency as a function of the number of simulation calls in each (re-)training session. The 6°, 8°, and 10° cases are the same as those shown previously in Figure 3(c); the new results may be seen in the 2°, 4°, 12°, and 14° cases, where blue/red lines indicate the new optimization data and gray lines indicate the non-sequential transfer learning cases from Figure 3(c) for comparison. We observe that sequential transfer learning improves the optimization convergence rate for the distant grating coupler scattering angles, in accordance with our initial prediction. This observation is made more explicit in Figure 4(d) where

we plot the simulation call requirement to reach an 80% grating coupler efficiency threshold for each of the transfer learning optimizations. Blue/red arrows and points indicate sequential transfer learning progressions, while gray diamonds indicate the one-shot transfer learning cases reproduced from Figure 3(d). Evidently, sequential transfer learning improves simulation call requirements by approximately an order of magnitude relative to the single-step cases. While we find that there is still a noticeable bias towards longer (but less severely long) training at smaller scattering angles, sequential transfer learning had the added benefit of providing the most efficient devices overall at every scattering angle in the progression. Note that the plotted simulation call requirements do not include the simulation calls from the previous iteration in the sequential transfer learning progression (each application of transfer learning used 10,000 simulation calls, where the final neural network weightset after those 10,000 simulation calls was used as the initial weightset for the next iteration of transfer learning). Furthermore, note that each optimization in these sequential transfer learning cases used the same neural network architecture and hyperparameters as the original PHORCED optimization (8° case), except for the $\{12° \rightarrow 14°\}$ and $\{4° \rightarrow 2°\}$ cases which required a slightly smaller learning rate in optimization for better performance. The smaller learning rate negligibly affected the 80% efficiency simulation call requirement shown in Figure 4(d).

5 Conclusions

In this work we introduced PHORCED, a photonic optimization package leveraging the policy gradient method from reinforcement learning. This method interfaces a probabilistic neural network with an electromagnetic solver for enhanced inverse design capabilities. PHORCED does not require an evaluation of adjoint method gradient of the electromagnetic merit function with respect to the design parameters, therefore eliminating the need to perform adjoint simulations over the course of an optimization. We anticipate that this fact can be particularly advantageous for multifrequency electromagnetic merit functions, where multiple adjoint simulations would normally be required in a simple frequency-domain implementation of the adjoint method (e.g. see Ref. [52]).

We applied both PHORCED and the GLOnet method to the proof-of-concept optimization of grating couplers. We found that both algorithms could outperform conventional gradient-based BFGS optimization, resulting in state-of-the-art simulated insertion loss for single-etch c-Si grating couplers and resilience against poor choices of initial condition. In future work we intend to implement fabrication constraints, alternative choices of geometrical parameterization, and other criteria to guarantee feasibility and robustness of experimental devices.

As an additional contribution we introduced the concept of transfer learning to integrated photonic optimization, revealing that a trained neural network using PHORCED could be re-trained on alternative problems with accelerated convergence. In particular, we showed that transfer learning could be applied to the design of grating couplers with varied scattering angle. Transfer learning was extremely effective for grating coupler scattering angles within $\approx \pm 4°$ to the original optimization angle, improving the convergence rate by $>10\times$ in some cases. However, this range could be effectively extended to $\approx \pm 6°$ or more using a sequential transfer learning approach, where transfer learning was applied multiple times progressively to slowly change the angle seen by the neural network. Because neural network based design methods such as PHORCED are generally data-hungry, we believe that transfer learning could greatly reduce the electromagnetic simulation and compute time that would otherwise be required by these techniques in the design of complex electromagnetic structures. For example, transfer learning could be used in multiple hierarchical stages to evolve an optimization from a two-dimensional structure to a three-dimensional structure, or from a surrogate model (e.g., the grating coupler model in Ref. [10]) to real physics.

Looking forward, we would like to emphasize that PHORCED takes advantage of fundamental concepts in reinforcement learning, but there is a plethora of burgeoning contemporary research in this field, such as advanced policy gradient, off-policy, and model-based approaches. We anticipate that further cross-pollination of the inverse electromagnetic design and reinforcement learning communities could open the floodgates for new research in electromagnetic optimization.

Acknowledgment: The authors thank S. K. Vadlamani for reinforcement learning discussions.
Author contribution: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: Hewlett Packard Enterprise.
Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References

[1] Y. Shen, N. C. Harris, S. Skirlo, et al., "Deep learning with coherent nanophotonic circuits," *Nat. Photonics*, vol. 11, pp. 441–446, 2016.

[2] T. Inagaki, Y. Haribara, K. Igarashi, et al., "A coherent Ising machine for 2000-node optimization problems," *Science*, vol. 354, no. 6312, pp. 603–606, 2016.

[3] N. C. Harris, G. R. Steinbrecher, M. Prabhu, et al., "Quantum transport simulations in a programmable nanophotonic processor," *Nat. Photonics*, vol. 11, no. 7, pp. 447–452, 2017.

[4] Y. Yamamoto, K. Aihara, T. Leleu, et al., "Coherent Ising machines–optical neural networks operating at the quantum limit," *npj Quantum Inf.*, vol. 3, no. 1, pp. 1–15, 2017.

[5] E. Khoram, A. Chen, D. Liu, et al., "Nanophotonic media for artificial neural inference," *Photon. Res.*, vol. 7, no. 8, pp. 823–827, 2019.

[6] T. W. Hughes, I. A. D. Williamson, M. Minkov, and S. Fan, "Wave physics as an analog recurrent neural network," *Sci. Adv.*, vol. 5, no. 12, pp. 1–6, 2019.

[7] B. J. Shastri, A. N. Tait, T. Ferreira de Lima, et al., "Photonics for artificial intelligence and neuromorphic computing," *Nat. Photonics*, vol. 15, no. 2, pp. 102–114, 2021.

[8] X. Xu, M. Tan, B. Corcoran, et al., "11 TOPS photonic convolutional accelerator for optical neural networks," *Nature*, vol. 589, no. 7840, pp. 44–51, 2021.

[9] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.*, vol. 378, pp. 686–707, 2019.

[10] D. Gostimirovic and W. N. Ye, "An open-source artificial neural network model for polarization-insensitive silicon-on-insulator subwavelength grating couplers," *IEEE J. Sel. Top. Quant. Electron.*, vol. 25, no. 3, pp. 1–5, 2019.

[11] Y. Guo, X. Cao, B. Liu, and M. Gao, "Solving partial differential equations using deep learning and physical constraints," *Appl. Sci.*, vol. 10, no. 17, p. 5917, 2020.

[12] Y. Chen, L. Lu, G. E. Karniadakis, and L. D. Negro, "Physics-informed neural networks for inverse problems in nano-optics and metamaterials," *Opt. Express*, vol. 28, no. 8, pp. 11618–11633, 2020.

[13] R. Pestourie, Y. Mroueh, T. V. Nguyen, P. Das, and S. G. Johnson, "Active learning of deep surrogates for PDEs: application to metasurface design," *npj Comput. Mater.*, vol. 6, no. 1, pp. 1–7, 2020.

[14] A. Ghosh, D. J. Roth, L. H. Nicholls, W. P. Wardley, A. V. Zayats, and V. A. Podolskiy, "Machine learning-based diffractive image analysis with subwavelength resolution," *ACS Photonics*, vol. 8, no. 5, pp. 1448–1456, 2021.

[15] L. Lu, R. Pestourie, W. Yao, Z. Wang, F. Verdugo, and S. G. Johnson, "Physics-informed neural networks with hard constraints for inverse design," arXiv:2102.04626 [physics], 2021.

[16] R. Trivedi, L. Su, J. Lu, M. F. Schubert, and J. Vuckovic, "Data-driven acceleration of photonic simulations," *Sci. Rep.*, vol. 9, no. 1, pp. 1–7, 2019.

[17] Y. Qu, L. Jing, Y. Shen, M. Qiu, and M. Soljačić, "Migrating knowledge between physical scenarios based on artificial neural networks," *ACS Photonics*, vol. 6, no. 5, pp. 1168–1174, 2019.

[18] D. Melati, Y. Grinberg, M. Kamandar Dezfouli, et al., "Mapping the global design space of nanophotonic components using machine learning pattern recognition," *Nat. Commun.*, vol. 10, no. 1, pp. 1–9, 2019.

[19] M. H. Tahersima, K. Kojima, T. Koike-Akino, et al., "Deep neural network inverse design of integrated photonic power splitters," *Sci. Rep.*, vol. 9, no. 1, p. 1368, 2019.

[20] A. Demeter-Finzi and S. Ruschin, "S-matrix absolute optimization method for a perfect vertical waveguide grating coupler," *Opt. Express*, vol. 27, no. 12, pp. 16713–16718, 2019.

[21] M. K. Dezfouli, Y. Grinberg, D. Melati, et al., "Design of fully apodized and perfectly vertical surface grating couplers using machine learning optimization," in *Proc. SPIE, Integrated Optics: Devices, Materials, and Technologies XXV*, vol. 11689, San Francisco, USA, International Society for Optics and Photonics, 2021, p. 116890J.

[22] M. M. R. Elsawy, S. Lanteri, R. Duvigneau, G. Briére, M. S. Mohamed, and P. Genevet, "Global optimization of metasurface designs using statistical learning methods," *Sci. Rep.*, vol. 9, no. 1, p. 17918, 2019.

[23] A. M. Hammond and R. M. Camacho, "Designing integrated photonic devices using artificial neural networks," *Opt. Express*, vol. 27, no. 21, pp. 29620–29638, 2019.

[24] J. Jiang and J. A. Fan, "Global optimization of dielectric metasurfaces using a physics-driven neural network," *Nano Lett.*, vol. 19, no. 8, pp. 5366–5372, 2019.

[25] J. Jiang and J. A. Fan, "Simulator-based training of generative neural networks for the inverse design of metasurfaces," *Nanophotonics*, vol. 9, no. 5, pp. 1059–1069, 2020.

[26] J. Jiang and J. A. Fan, "Multiobjective and categorical global optimization of photonic structures based on ResNet generative neural networks," *Nanophotonics*, vol. 10, no. 1, pp. 361–369, 2020.

[27] J. Jiang, M. Chen, and J. A. Fan, "Deep neural networks for the evaluation and design of photonic devices," *Nat. Rev. Mater.*, vol. 6, no. 8, pp. 679–700, 2021.

[28] R. S. Hegde, "Deep learning: a new tool for photonic nanostructure design," *Nanoscale Adv.*, vol. 2, no. 3, pp. 1007–1023, 2020.

[29] M. Minkov, I. A. Williamson, L. C. Andreani, et al., "Inverse design of photonic crystals through automatic differentiation," *ACS Photonics*, vol. 7, no. 7, pp. 1729–1741, 2020.

[30] S. So, T. Badloe, J. Noh, J. Rho, and J. Bravo-Abad, "Deep learning enabled inverse design in nanophotonics," *Nanophotonics*, vol. 2234, pp. 1–17, 2020.

[31] Z. Ma and Y. Li, "Parameter extraction and inverse design of semiconductor lasers based on the deep learning and particle swarm optimization method," *Opt. Express*, vol. 28, no. 15, p. 21971, 2020.

[32] K. Kojima, M. H. Tahersima, T. Koike-Akino, et al., "Deep neural networks for inverse design of nanophotonic devices," *J. Lightwave Technol.*, vol. 39, no. 4, pp. 1010–1019, 2021.

[33] W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nat. Photonics*, vol. 15, no. 2, pp. 77–90, 2021.

[34] D. Melati, M. Kamandar Dezfouli, Y. Grinberg, et al., "Design of compact and efficient silicon photonic micro antennas with

perfectly vertical emission,'' *IEEE J. Sel. Top. Quant. Electron.*, vol. 27, no. 1, pp. 1–10, 2021.
[35] R. Hegde, ''Sample-efficient deep learning for accelerating photonic inverse design,'' *OSA Continuum*, vol. 4, no. 3, pp. 1019–1033, 2021.
[36] J. S. Jensen and O. Sigmund, ''Topology optimization for nano-photonics,'' *Laser Photon. Rev.*, vol. 5, no. 2, pp. 308–321, 2011.
[37] J. Lu and J. Vučković, ''Objective-first design of high efficiency, small-footprint couplers between arbitrary nanophotonic waveguide modes,'' *Opt. Express*, vol. 20, no. 7, pp. 7221–7236, 2012.
[38] C. M. Lalau-Keraly, S. Bhargava, O. D. Miller, and E. Yablonovitch, ''Adjoint shape optimization applied to electromagnetic design,'' *Opt. Express*, vol. 21, no. 18, pp. 21, 2013.
[39] Y. Elesin, B. S. Lazarov, J. S. Jensen, and O. Sigmund, ''Time domain topology optimization of 3D nanophotonic devices,'' *Photonics Nanostructures: Fundam. Appl.*, vol. 12, no. 1, pp. 23–33, 2014.
[40] 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,'' *Nat. Photonics*, vol. 9, no. 6, pp. 374–377, 2015.
[41] L. F. Frellsen, Y. Ding, O. Sigmund, and L. H. Frandsen, ''Topology optimized mode multiplexing in silicon-on-insulator photonic wire waveguides,'' *Opt. Express*, vol. 24, no. 15, pp. 16866–16873, 2016.
[42] L. Su, R. Trivedi, N. V. Sapra, A. Y. Piggott, D. Vercruysse, and J. Vučković, ''Fully-automated optimization of grating couplers,'' *Opt. Express*, vol. 26, no. 4, pp. 2614–2617, 2017.
[43] A. Michaels and E. Yablonovitch, ''Leveraging continuous material averaging for inverse electromagnetic design,'' *Opt. Express*, vol. 26, no. 24, pp. 31717–31737, 2018.
[44] A. Michaels and E. Yablonovitch, ''Inverse design of near unity efficiency perfectly vertical grating couplers,'' *Opt. Express*, vol. 26, no. 4, pp. 4766–4779, 2018.
[45] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, ''Inverse design in nanophotonics,'' *Nat. Photonics*, vol. 12, no. 11, pp. 659–670, 2018.
[46] T. W. Hughes, M. Minkov, I. A. D. Williamson, and S. Fan, ''Adjoint method and inverse design for nonlinear nanophotonic devices,'' *ACS Photonics*, vol. 5, no. 12, pp. 4781–4787, 2018.
[47] Y. Liu, W. Sun, H. Xie, et al., ''Very sharp adiabatic bends based on an inverse design,'' *Opt. Lett.*, vol. 43, no. 11, pp. 2482–2485, 2018.
[48] N. M. Andrade, S. Hooten, S. A. Fortuna, K. Han, E. Yablonovitch, and M. C. Wu, ''Inverse design optimization for efficient coupling of an electrically injected optical antenna-LED to a single-mode waveguide,'' *Opt. Express*, vol. 27, no. 14, pp. 19802–19814, 2019.
[49] D. Vercruysse, N. V. Sapra, L. Su, R. Trivedi, and J. Vučković, ''Analytical level set fabrication constraints for inverse design,'' *Sci. Rep.*, vol. 9, no. 1, pp. 1–7, 2019.
[50] Y. Augenstein and C. Rockstuhl, ''Inverse design of nanophotonic devices with structural integrity,'' *ACS Photonics*, vol. 7, no. 8, pp. 2190–2196, 2020.
[51] E. Bayati, R. Pestourie, S. Colburn, Z. Lin, S. G. Johnson, and A. Majumdar, ''Inverse designed metalenses with extended depth of focus,'' *ACS Photonics*, vol. 7, no. 4, pp. 873–878, 2020.
[52] S. Hooten, T. V. Vaerenbergh, P. Sun, S. Mathai, Z. Huang, and R. G. Beausoleil, ''Adjoint optimization of efficient CMOS-compatible Si-SiN vertical grating couplers for DWDM applications,'' *J. Lightwave Technol.*, vol. 38, no. 13, pp. 3422–3430, 2020.
[53] W. Jin, S. Molesky, Z. Lin, K. M. C. Fu, and A. W. Rodriguez, ''Inverse design of compact multimode cavity couplers,'' *Opt. Express*, vol. 26, no. 20, pp. 26713–26721, 2018.
[54] W. Jin, W. Li, M. Orenstein, and S. Fan, ''Inverse design of lightweight broadband reflector for relativistic lightsail propulsion,'' *ACS Photonics*, vol. 7, no. 9, pp. 2350–2355, 2020.
[55] A. Michaels, M. C. Wu, and E. Yablonovitch, ''Hierarchical design and optimization of silicon photonics,'' *IEEE J. Sel. Top. Quant. Electron.*, vol. 26, no. 2, pp. 1–12, 2020.
[56] P. Sun, T. V. Vaerenbergh, M. Fiorentino, and R. Beausoleil, ''Adjoint-method-inspired grating couplers for CWDM O-band applications,'' *Opt. Express*, vol. 28, no. 3, pp. 3756–3767, 2020.
[57] Z. Lin, C. Roques-Carmes, R. Pestourie, M. Soljačić, A. Majumdar, and S. G. Johnson, ''End-to-end nanophotonic inverse design for imaging and polarimetry,'' *Nanophotonics*, vol. 10, no. 3, pp. 1177–1187, 2021.
[58] Z. Omair, S. M. Hooten, and E. Yablonovitch, ''Broadband mirrors with >99% reflectivity for ultra-efficient thermophotovoltaic power conversion,'' in *Proc. SPIE, Energy Harvesting and Storage: Materials, Devices, and Applications XI*, vol. 11722, Orlando, USA, International Society for Optics and Photonics, 2021, p. 1172208.
[59] D. Vercruysse, N. V. Sapra, K. Y. Yang, and J. Vučković, ''Inverse-designed photonic crystal devices for optical beam steering,'' arXiv:2102.00681 [physics], 2021.
[60] Z. Zeng, P. K. Venuthurumilli, and X. Xu, ''Inverse design of plasmonic structures with FDTD,'' *ACS Photonics*, vol. 8, no. 5, pp. 1489–1496, 2021.
[61] R. J. Williams, ''Simple statistical gradient-following algorithms for connectionist reinforcement learning,'' *Mach. Learn.*, vol. 8, no. 3, pp. 229–256, 1992.
[62] R. S. Sutton, D. McAllester, S. Singh, and Y. Mansour, ''Policy gradient methods for reinforcement learning with function approximation,'' in *Advances in Neural Information Processing Systems*, vol. 12, S. Solla, T. Leen, and K. Müller, Eds., Cambridge, USA, MIT Press, 2000.
[63] A. Michaels, *EMopt*, 2019 [Online]. Available at: https://github.com/anstmichaels/emopt.
[64] L. Faury, F. Vasile, C. Calauzénes, and O. Fercoq, ''Neural generative models for global optimization with gradients,'' arXiv:1805.08594 [cs], 2018.
[65] L. Faury, C. Calauzenes, O. Fercoq, and S. Krichen, ''Improving evolutionary strategies with generative neural networks,'' arXiv:1901.11271 [cs], 2019.
[66] N. Hansen, ''The CMA evolution strategy: a tutorial,'' arXiv:1604.00772 [cs, stat], 2016.
[67] T. Salimans, J. Ho, X. Chen, S. Sidor, and I. Sutskever, ''Evolution strategies as a scalable alternative to reinforcement learning,'' arXiv:1703.03864 [cs, stat], 2017.

[68] H. Ghraieb, J. Viquerat, A. Larcher, P. Meliga, and E. Hachem, "Single-step deep reinforcement learning for openloop control of laminar and turbulent flows," *Phys. Rev. Fluids*, vol. 6, no. 5, p. 053, 2021.
[69] A. Michaels, *GCSlab*, 2019 [Online]. Available at: https://github.com/anstmichaels/gcslab.
[70] S. Mathai, P. Rosenberg, G. Panotopoulos, et al., "Detachable 1x8 single mode optical interface for DWDM microring silicon photonic transceivers," in *Proc. SPIE, Optical Interconnects XX*, vol. 11286, San Francisco, USA, International Society for Optics and Photonics SPIE, 2020, pp. 62–71.
[71] T. Watanabe, M. Ayata, U. Koch, Y. Fedoryshyn, and J. Leuthold, "Perpendicular grating coupler based on a blazed antiback-reflection structure," *J. Lightwave Technol.*, vol. 35, no. 21, pp. 4663–4669, 2017.
[72] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, Cambridge, MA, MIT press, 2018.

Supplementary Material: The online version of this article offers supplementary material (https://doi.org/10.1515/nanoph-2021-0332).