

Research article

Joeri Lenaerts, Hannah Pinson and Vincent Ginis*

Artificial neural networks for inverse design of resonant nanophotonic components with oscillatory loss landscapes

<https://doi.org/10.1515/nanoph-2020-0379>

Received July 6, 2020; accepted November 3, 2020;

published online November 30, 2020

Abstract: Machine learning offers the potential to revolutionize the inverse design of complex nanophotonic components. Here, we propose a novel variant of this formalism specifically suited for the design of resonant nanophotonic components. Typically, the first step of an inverse design process based on machine learning is training a neural network to approximate the non-linear mapping from a set of input parameters to a given optical system's features. The second step starts from the desired features, e.g. a transmission spectrum, and propagates back through the trained network to find the optimal input parameters. For resonant systems, this second step corresponds to a gradient descent in a highly oscillatory loss landscape. As a result, the algorithm often converges into a local minimum. We significantly improve this method's efficiency by adding the Fourier transform of the desired spectrum to the optimization procedure. We demonstrate our method by retrieving the optimal design parameters for desired transmission and reflection spectra of Fabry–Pérot resonators and Bragg reflectors, two canonical optical components whose functionality is based on wave interference. Our results can be extended to the optimization of more complex nanophotonic components interacting with structured incident fields.

Keywords: artificial neural networks; inverse design; optical resonators.

1 Introduction

Inverse design of optical components is the process of calculating the material properties that yield a specific optical response. Popular computational methods for inverse design are evolutionary algorithms [1] and density topology optimization [2–4].

A prevalent method in density topology optimization is the adjoint method. This technique provides a way to efficiently compute gradients of a loss function with respect to design parameters, as was demonstrated in the design of demultiplexers that separate light of different wavelengths [5, 6]. A similar method underlies the design of tunable metasurfaces [7] and nonlinear nanophotonic devices [8]. Recently, Angeris et al. [9] found a technique to calculate the bounds on the performance of these optical structures.

An interesting alternative formalism was introduced by Peurifoy et al. in 2018 [10]. The authors applied deep learning to the inverse design of multi-layered nanoparticles in two different steps. First, a neural network is trained to predict the particles' scattering cross-section as a function of each shell's thicknesses. This network is trained using a supervised learning algorithm. Second, this trained network is used to perform gradient descent on the design parameters. In other words, after the first training step, all weights and biases of the network are fixed, and only the input parameters can be updated to reduce the distance between the predicted and the desired scattering cross-section. So et al. improved this technique of inverse design of nanoparticles in 2019 [11] by adapting the loss function of the neural network to learn the thickness as well as the material of each layer. More advanced neural network architectures, such as a tandem network [12], Generative Adversarial Networks (GANs) [13–16] and autoencoders [17, 18] have also been successfully used in the field of inverse design for optical components. In addition, deep learning has found use in various vibrant areas of nanophotonics, including integrated photonics [19], holography [20], diamond photonics [21], plasmonic nanoparticles [22]

*Corresponding author: Vincent Ginis, Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, 29 Oxford Street, Cambridge, Massachusetts 02138, USA; and Data Lab/Applied Physics, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium, E-mail: vincent.ginis@vub.be. <https://orcid.org/0000-0003-0063-9608>

Joeri Lenaerts and Hannah Pinson, Data Lab/Applied Physics, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium

and topological photonics [23]. More examples can be found in recent reviews [24, 25].

A problem arises when applying these inverse design methods to optical systems with strong resonances. These resonances are often based on the interference of many partial waves. A highly oscillatory loss function originates because of the periodicity of the phase difference between the interfering waves. The gradient-based inverse design thus often gets stuck in local minima, far away from the global optimum. In this contribution, we propose a solution to this problem, based on Fourier analysis.

2 Predicting the transmission spectra of simple resonators

We first illustrate our technique using a simple one-dimensional Fabry–Pérot resonator: a thin slice of a dielectric medium. When light hits the resonator, it is reflected several times inside the medium. The interference of these partial waves heavily influences the reflected and transmitted light intensities. This leads to a non-trivial dependence of the transmission and reflection spectra on

the wavelength of the incident light and the physical parameters of the resonator, as illustrated in Figure 1(A).

Four parameters determine the interference of the partial waves: the wavelength of the incoming light λ , the angle of incidence θ , the index of refraction n and the width of the resonator l . The wavelength-dependent transmission is then given by

$$T(\lambda, \theta, n, l) = \frac{1}{1 + F \sin^2(\delta_0 / (2\lambda))}, \quad (1)$$

where $F(\theta, n) = 4R / (1 - R)^2$ is the coefficient of Finesse, and $\delta_0(\theta, n, l) = 4\pi n l \cos(\theta_{\text{mat}})$ is a parameter that encodes the phase difference between the different partial waves that are transmitted. Finally, the parameter θ_{mat} is the transmission angle inside the resonator, determined by the incident angle and the index of refraction, in agreement with Snell's law.

We train a neural network to predict the spectral transmission $T(\lambda)$ from given values for F and δ_0 . The network consists of an input layer with two nodes, a number of hidden layers with adjustable weights and non-linear activation functions, and an output layer with 200 nodes, representing the transmission spectrum sampled at different wavelengths. During this training procedure, a batch of values of F and δ_0 , all normalized between -1 and 1 , are

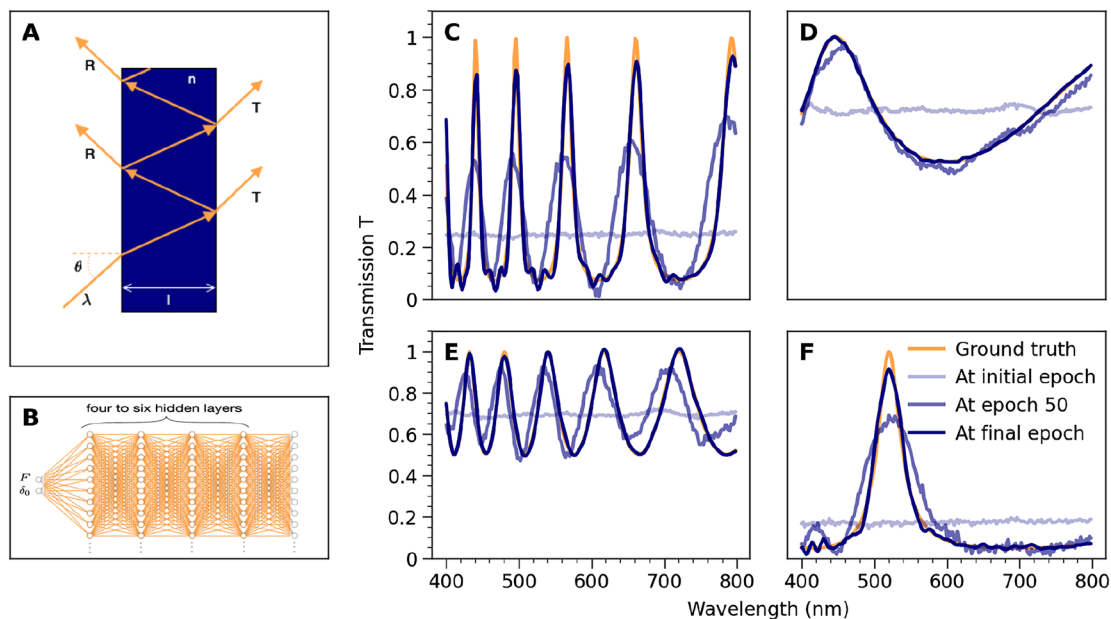


Figure 1: The first step of inverse design using deep learning.

(A) A Fabry–Pérot resonator with refractive index n and width l . Multiple reflections inside the resonator cause interference between different outgoing waves, leading to non-trivial reflection and transmission spectra. (B) A neural network that maps the resonator parameters F and δ_0 to the wavelength-dependent transmission. The networks have four to six hidden layers. The networks we trained had 100 or 200 neurons in the hidden layers and 200 nodes in the output layer, each node corresponding to one wavelength between 400 and 800 nm. The parameters that are updated during training are highlighted in orange. (C)–(F) Predictions of transmissions $T(\lambda)$ for different sets of material parameters. The orange lines show the analytically computed transmissions or ground truth. The blue lines show the predictions during training of the neural network.

provided as input to the network. The network performs a non-linear mapping from this input to an output in the form of a transmission spectrum; the mean squared error (MSE) between this output and the actual transmission spectrum is the loss function. Subsequently, the gradient of this loss function is propagated back through the network to update the values of the weights. This procedure gradually improves the spectral predictions of the network. After a number of epochs—where in each epoch, the entire training data set is used once—predictions of the network closely match the actual spectra. This is visualized in Figure 1.

We performed a search over network architectures and activation functions. The best performing neural network had six hidden layers, 200 nodes in each hidden layer, and the swish activation function. The full details of the meta-search and the resulting network are given in the Supplementary material. The final network model obtains a test mean absolute error (MAE) of $0.485 \pm 0.060\%$. Some examples of predicted spectra $T(\lambda)$ during training are shown in Figure 1(C)–(F). The target spectrum or ground truth, calculated using Eq. (1), is shown in orange; the predictions during different phases of the training are shown in blue. The spectra with sharp peaks are somewhat harder to predict. This can be expected, since we sample the transmission at 200 points. A denser sampling of the transmission should lead to better predictions for transmissions with sharper peaks.

3 Inverse design in an oscillatory loss landscape

In the second step of the inverse design procedure, we harness the fully trained network and the fact that its structure allows us to perform a gradient descent from output to input in a straightforward way. Our goal is to find the material parameters that yield the desired transmission spectrum $T(\lambda)$. The input values are initialized with random values, representing a first guess of the material parameters F and δ_0 . Unlike before, only these input values are updated at this point. The weights of the network remain fixed throughout this step. The neural network is used to compute the transmission spectrum from its input values, and subsequently the MSE between the obtained and desired spectrum yields the loss function given by

$$\text{MSE}_{\text{original}} = \sum_{\lambda} |\hat{T}(\lambda) - T(\lambda)|^2. \quad (2)$$

The gradient of this loss function is propagated back to the input nodes, where the values are adjusted in a

direction that minimizes this loss function. Over time they thus converge to the material parameters that yield a transmission spectrum $\hat{T}(\lambda)$ that is most similar to the desired transmission $T(\lambda)$.

We repeatedly encountered one specific problem in our experiments: convergence to one of the local minima in the oscillatory loss landscape. As a result, without carefully setting the initial values for F and δ_0 , the global minimum is often not found. This is illustrated in Figure 2 and in the left panel of Figure 3.

Figure 2(A) depicts the trained network, while Figure 2(B) shows the desired transmission in orange along with two obtained transmission spectra for two different sets of initial values of F and δ_0 in blue and red. Figure 2(C) shows the mean squared error (MSE) between the desired and obtained transmissions during training. The orange line indicates the loss for predictions with the correct parameters F and δ_0 . The other panels represent the evolution of the parameters F and δ_0 . Here, the orange line shows the true parameters from which the desired transmission $T(\lambda)$ was computed. We distinguish two cases. A first case where inverse design works, converging to the global minimum, and a second case where the gradient descent gets stuck in a local minimum. In this second case, the obtained and the desired transmissions do not look alike. We also obtain a considerably higher loss than the one obtained for the correct parameters for F and δ_0 . In the other case, we choose an initial value sufficiently close to the true δ_0 , such that the gradient descent does converge to a global optimum. We see that the spectra are indeed very similar (orange and blue lines). We notice that the loss is as low as it would be for the true parameters.

In Figure 3(A), we see that the loss landscape appears to have strong oscillations along the direction of δ_0 , which is causing the local minima. We can illustrate this further by looking at cross-sections of the loss landscape. For a section at constant δ_0 , the MSE loss is perfectly suited for gradient descent in the direction of F . There is only one minimum to which we can converge. The story is different for the parameter δ_0 (Figure 3(B)–(C)). The MSE loss shows sharp local minima, from which it is difficult to escape.

In the specific case of the oscillatory loss landscapes, as is found in the inverse design of resonant components, we propose a more efficient solution—one grounded in physical arguments: the key idea is to take the Fourier transform of the desired spectrum $T(\lambda)$ into account. Since δ_0 is linked to the number of oscillations in the transmission spectrum, meaningful information about δ_0 is adequately captured by the Fourier transform of the transmission spectrum. Therefore, we introduce a second

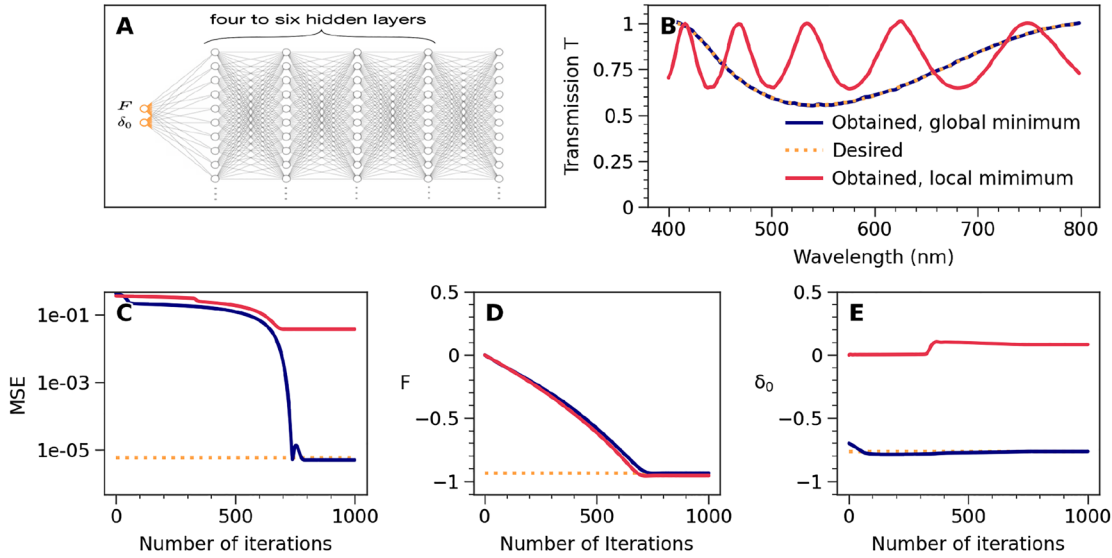


Figure 2: The second step of inverse design using deep learning.

(A) A trained neural network mapping the resonator parameters F and δ_0 to the wavelength-dependent transmission. Only the resonator parameters (the input nodes, in orange) are updated. (B) Example of a desired spectrum and spectra obtained through inverse design. The desired transmission is shown in orange and the transmission predicted by the trained neural network for the obtained resonator parameters is shown in blue and red. Two sets of resonator parameters are obtained for different initializations of the input parameters F and δ_0 . (C) Evolution of the loss function during training. (D) Evolution of F during training. The orange line shows the value of F used to create the desired transmission. Values are normalized between -1 and 1 . (E) Evolution of δ_0 during training. The orange line shows the value of δ_0 to create the desired transmission.

loss function, the mean squared error (MSE) on the Fourier power spectrum of the target function $T(\lambda)$, given by

$$\text{MSE}_{\text{Fourier}} = \sum_{f=1}^{10} |\mathcal{F}^2\{\hat{T}(\lambda)\}(f) - \mathcal{F}^2\{T(\lambda)\}(f)|^2 \quad (3)$$

We show the loss landscape of this function in Figure 3(D)–(F). Here, we only consider the first 10 frequency bins, since the transmission spectra do not oscillate more than 10 times in the region of interest. To extract the information on the relative power in the different frequencies, we normalize this total power spectrum to one. The relative power in the different frequencies minimally depends on the oscillation amplitude. This implies that the normalized power spectrum is insensitive to the parameter F , which controls the oscillation amplitude. It also means that every cross-section of the loss landscape along δ_0 with constant F will approximately be the same. We see in Figure 3(D) that this is indeed the case.

We can use this fact to initialize δ_0 sufficiently close to the true value to find the global minimum more quickly in the gradient descent algorithm. Two sections, at $F = 0$ and at the true parameter value $F = -0.932$, are plotted in Figure 3(E)–(F). We can see that these sections reach a global minimum at similar values of δ_0 . Thus, we can compute this value for δ_0 without knowing the true value for F beforehand. Simply computing the minimum of the

obtained section of the Fourier power spectrum at a chosen value of F , e.g. $F = 0$, yields the desired result.

Unfortunately, the assumption that the sections of the MSE on the Fourier power spectra are equal for all values of F is not entirely valid, as we already see in Figure 3. To mediate this problem, we re-initialize δ_0 after a few iterations of gradient descent, when we are closer to the true parameter for F . In our algorithm, we make this update every 100 epochs. For values of F that are closer to the true parameter, the sections will be more similar to the global minimum section. This way, the initial value for δ_0 gets closer to the true parameter every time we re-initialize.

In the end, we are not interested in matching the Fourier transforms of the spectra, but merely in matching the spectra themselves. The Fourier power spectrum we used so far also does not take information on the phase of different frequencies into account. Therefore, our goal is better achieved by a combined loss function that periodically re-initialises δ_0 : the sum of the MSE on the spectra and the MSE on the Fourier power spectra of the spectra. We thus take the MSE loss on the spectra into account in this periodic update:

$$L = \text{MSE}_{\text{original}} + \text{MSE}_{\text{Fourier}}. \quad (4)$$

Our algorithm is thus alternatively minimizing two loss functions. It minimizes the MSE on the spectrum through

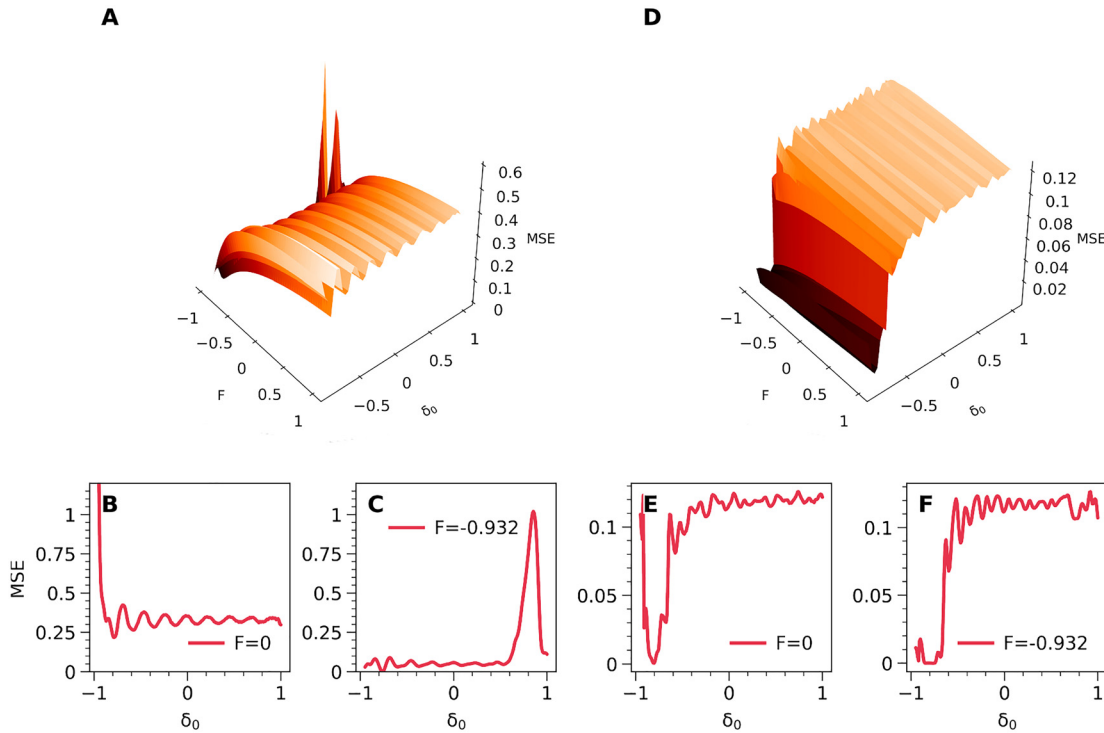


Figure 3: Finding the global minimum in an oscillatory loss landscape.

(A) Mean squared error (MSE) between the desired and predicted transmissions in function of F and δ_0 . The desired transmission is an example $T(\lambda)$ from the test set. The loss function shows many local minima. (B) Cross-section of the MSE on the transmissions as a function of δ_0 for $F = 0$. (C) Cross-section of the MSE as a function of δ_0 for $F = 0.932$. (D) MSE between the Fourier power spectra of the desired and predicted transmissions versus F and δ_0 . Note the different scale on the y-axis with respect to (A–C). (E) Cross-section of the MSE on the Fourier spectra as a function of δ_0 for $F = 0$. (F) Cross-section of the MSE on the Fourier spectra as a function of δ_0 for $F = -0.932$.

gradient descent, yielding better values for F and δ_0 . We avoid getting stuck in local minima during this process by periodically minimizing a second, combined loss function L given above. The latter minimization procedure consists of a straightforward computation of the value for δ_0 corresponding to a minimum of the loss at the current value for F . The other device parameter F is initialized at 0.0. To verify whether a different initialization would influence the results, we repeated the inverse design of one transmission from the test set for 100 different initializations of F . We demonstrate that all simulations converged to the optimal solution. In the following we explain how this method considerably improves the efficiency of the search in the parameter space.

4 Comparison with other search algorithms

Several solutions exist to overcome the premature convergence on local minima. One solution consists of increasing

the stochastic features of the gradient descent. In this way, updates to higher loss regions can occur, allowing the search algorithm to escape from the local minima. Another solution consists of abandoning the neural network gradient descent altogether: we could use the trained network only in the forward direction, to compute the loss for any given set of input parameters, and then combine this with a stochastic local search algorithm such as simulated annealing. Both these approaches potentially visit a vast area of the parameter space before finding the global minimum, thus decreasing the inverse design procedure's efficiency.

To validate our algorithm, we compare it to a classic local search algorithm in the form of simulated annealing (SA) [26]. We use SA to minimize the loss function obtained by the trained neural network's predictions. For a fair comparison, we use both the classic SA as the more advanced dual annealing (DA) algorithm [27]. This DA algorithm combines a larger neighbourhood function with an additional local search and a restart mechanism, resulting in a highly performant search process.

We evaluate our inverse design algorithm and the two simulated annealing algorithms on 1000 transmission

spectra from the Fabry–Pérot test set. The metric is the mean absolute error (MAE) between the obtained and desired spectra. We determine that a global minimum is found when we have an MAE less than 0.01.

Without fine tuning of the hyperparameters, both our algorithm and the DA reached the global minimum in 95% of the cases, while the SA algorithm never found the global minimum. But the actual improvement lies in the number of steps taken before the global minimum is reached: DA covers a large part of the parameter space, while our approach by design follows a simple path, almost straight to the global minimum. The path taken by the DA algorithm is 84 times longer than the path of our optimization algorithm. This is further illustrated in the Supplementary material.

5 More general resonant devices

To test if our method works for problems more complex than the Fabry–Pérot resonator, we also applied it to Bragg reflectors. The Bragg reflector consists of two alternating layers of materials with different refractive indices. The interference of many internal reflections again leads to a non-trivial wavelength-dependent reflection spectrum $R(\lambda)$. Now four parameters determine the spectra: the difference Δn between the two refractive indices of the materials, the mean refractive index \bar{n} of the whole structure, the total length Λ of a unit of two alternating layers and the number of these units that are stacked p . The underlying analytical formula for $R(\lambda)$ is given in the Supplemental material.

We trained a neural network to learn the mapping from the four resonator parameters to the reflection $R(\lambda)$. Then we evaluated our inverse design approach on three spectra in the test set. We used the second loss function to initialize the four parameters. We re-initialized them every 300 epochs for Λ and Δn and every 600 epochs for \bar{n} and p . The results of these experiments are shown in Figure 4(A)–(C). These figures show that our inverse design is capable of scaling to a problem with four parameters. In the Supplemental Material, we also show that the performance of the inverse design is not influenced by the number of stacked layers p . The dual annealing algorithm is also able to tackle this problem well. It however takes five times longer to converge, illustrating our algorithm's superior efficiency.

To show that we can also approximate arbitrary reflection spectra with our inverse design method, we borrowed a famous example of a one-dimensional curve from the novel “Le petit prince” by Antoine de Saint-Exupéry. We made a piecewise approximation of an elephant's shape

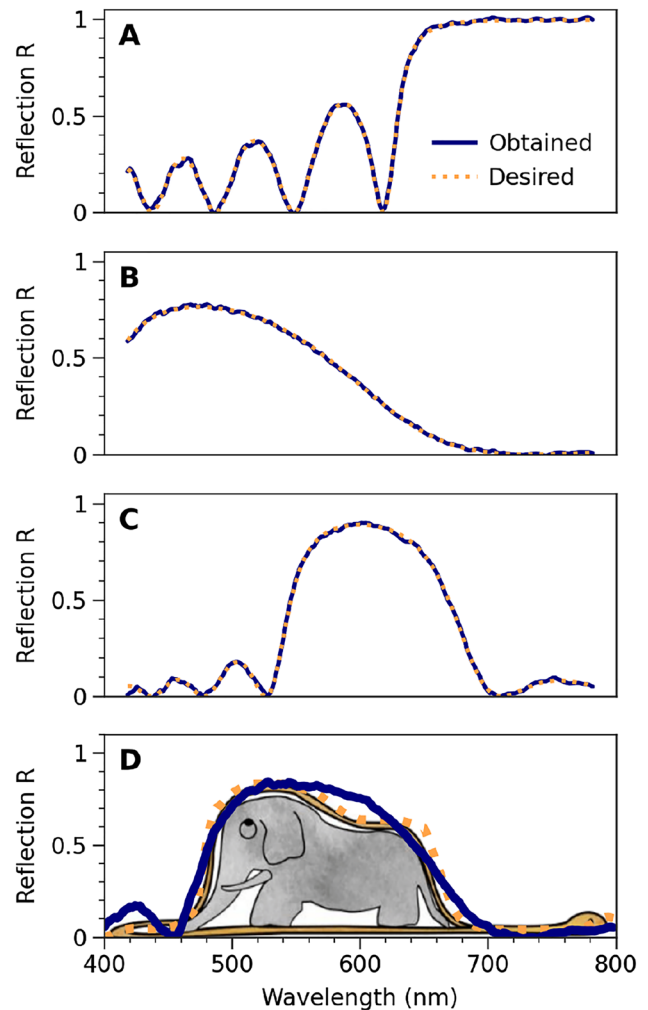


Figure 4: Inverse design of a general Bragg reflector.

(A–C) examples of desired reflection spectra from the test set in combination with the spectra obtained through inverse design. The desired spectrum is shown in orange and the spectrum corresponding to the obtained resonator parameters is shown in blue. (D) Fitting an elephant. The spectrum, shown in orange, approximates the shape of the elephant. The spectrum that corresponds with the obtained resonator parameters is shown in blue.

and performed inverse design on this arbitrary “reflection” spectrum. The result, shown in Figure 4(D), illustrates that our inverse design method is capable of approximating a broad range of reflection spectra.

6 Conclusion

Finding the global minimum in a parameter space where there are many local minima is a significant problem in inverse design. In this work, we show that, in the context of resonant photonic components, this problem can be

addressed by constructing a combined loss function that also fits the Fourier representation of the target function. Physically, this idea is grounded in the fact that the Fourier representation contains information about the target function's oscillatory behaviour. For many resonant photonic components, these oscillations straightforwardly correspond to specific parameters in the input space. We demonstrated our technique with Fabry–Pérot and Bragg resonators and quantified the superior search efficiency compared to other optimization methods. Our technique will be particularly valuable for inverse design of optical components based on many nanophotonic resonators, such as coupled or multi-layered resonators [28–30], lasers [31], dispersion-engineered metasurfaces [32] and (non-linear) resonators in the context of integrated optics [33, 34]. Finally, because this method can incorporate several parameters of the incident light, we think it will be interesting to study and optimize complex materials interacting with structured light beams [35–39].

Acknowledgement: H.P. and J.L. acknowledge fellowships from the Research Foundation Flanders (FWO-Vlaanderen) under Grant No. 11A6819N and Grant No. 11G1621N. Work at VUB was partially supported by the Research Council of the VUB.

Author contribution: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: H.P. and J.L. acknowledge fellowships from the Research Foundation Flanders (FWO-Vlaanderen) under Grant No. 11A6819N and Grant No. 11G1621N. Work at VUB was partially supported by the Research Council of the VUB.

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References

- [1] P. R. Wiecha, C. Majorel, C. Girard, et al., “Design of plasmonic directional antennas via evolutionary optimization,” *Opt. Express*, vol. 27, no. 20, p. 29069, 2019.
- [2] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, “Inverse design in nanophotonics,” *Nat. Photonics*, vol. 12, pp. 659–670, 2018, arXiv: 1801.06715.
- [3] Z. Lin, B. Groever, F. Capasso, A. W. Rodriguez, and M. Lončar, “Topology-optimized multilayered metaoptics,” *Phys. Rev. Appl.*, vol. 9, no. 4, p. 044030, 2018.
- [4] R. Pestourie, C. Pérez-Arancibia, Z. Lin, W. Shin, F. Capasso, and S. G. Johnson, “Inverse design of large-area metasurfaces,” *Opt. Express*, vol. 26, no. 26, pp. 33732–33747, 2018.
- [5] 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.
- [6] L. Su, A. Y. Piggott, N. V. Sapra, J. Petykiewicz, and J. Vuckovic, “Inverse design and demonstration of a compact on-chip narrowband three-channel wavelength demultiplexer,” *ACS Photonics*, vol. 5, no. 2, pp. 301–305, 2018.
- [7] H. Chung and O. Miller, “Tunable metasurface inverse design for 80% switching efficiencies and 144° angular steering,” *ACS Photonics*, vol. 7, no. 8, pp. 2236–2243, 2020, arXiv: 1910.03132 [physics.optics].
- [8] T. W. Hughes, M. Minkov, I. A. Williamson, and S. Fan, “Adjoint method and inverse design for nonlinear nanophotonic devices,” *ACS Photonics*, vol. 5, no. 12, pp. 4781–4787, 2018.
- [9] G. Angeris, J. Vuckovic, and S. P. Boyd, “Computational bounds for photonic design,” *ACS Photonics*, vol. 6, no. 5, pp. 1232–1239, 2019.
- [10] J. Peurifoy, Y. Shen, L. Jing, et al., “Nanophotonic particle simulation and inverse design using artificial neural networks,” *Sci. Adv.*, vol. 4, no. 6, 2017.
- [11] S. So, J. Mun, and J. Rho, “Simultaneous inverse design of materials and parameters of core-shell nanoparticle via deep-learning: demonstration of dipole resonance engineering,” 2019, arXiv: 1904.02848 [physics.optics].
- [12] D. Liu, Y. Tan, E. Khoram, and Z. Yu, “Training deep neural networks for the inverse design of nanophotonic structures,” *ACS Photonics*, vol. 5, no. 4, pp. 1365–1369, 2018.
- [13] J. Jiang, D. Sell, S. Hoyer, J. Hickey, J. Yang, and J. A. Fan, “Free-form diffractive metagrating design based on generative adversarial networks,” *ACS Nano*, vol. 13, no. 8, pp. 8872–8878, 2019.
- [14] 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.
- [15] 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, 2019.
- [16] S. So and J. Rho, “Designing nanophotonic structures using conditional deep convolutional generative adversarial networks,” *Nanophotonics*, vol. 8, no. 7, pp. 1255–1261, 2019.
- [17] Y. Kiarashinejad, S. Abdollahramezani, and A. Adibi, “Deep learning approach based on dimensionality reduction for designing electromagnetic nanostructures,” *Npj Comput. Mater.*, vol. 6, no. 1, 2020, <https://doi.org/10.1038/s41524-020-0276-y>.
- [18] W. Ma, F. Cheng, Y. Xu, Q. Wen, and Y. Liu, “Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy,” *Adv. Mater.*, vol. 31, no. 35, 2019, <https://doi.org/10.1002/adma.201901111>.
- [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, pp. 1–9, 2019.
- [20] H. Ren, W. Shao, Y. Li, F. Salim, and M. Gu, “Three-dimensional vectorial holography based on machine learning inverse design,” *Sci. Adv.*, vol. 6, no. 16, p. eaaz4261, 2020.
- [21] C. Dory, D. Vercruysse, K. Y. Yang, et al., “Inverse-designed diamond photonics,” *Nat. Commun.*, vol. 10, no. 1, pp. 1–7, 2019.
- [22] I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, “Plasmonic nanostructure design and

- characterization via deep learning,” *Light Sci. Appl.*, vol. 7, no. 1, pp. 2047–7538, 2018.
- [23] Y. Long, J. Ren, Y. Li, and H. Chen, “Inverse design of photonic topological state via machine learning,” *Appl. Phys. Lett.*, vol. 114, no. 18, p. 181105, 2019.
- [24] S. So, T. Badloe, J. Noh, J. Rho, and J. Bravo-Abad, “Deep learning enabled inverse design in nanophotonics,” *Nanophotonics*, vol. 9, no. 5, pp. 1041–1057, 2020.
- [25] K. Yao, R. Unni, and Y. Zheng, “Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale,” *Nanophotonics*, vol. 8, no. 3, pp. 339–366, 2019.
- [26] K. Scott, C. D. Gelatt, and M. P. Vecchi, “Optimization by simulated annealing,” *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
- [27] Y. Xiang, D. Y. Sun, W. Fan, and X. G. Gong, “Generalized simulated annealing algorithm and its application to the Thomson model,” *Phys. Lett.*, vol. 233, no. 3, pp. 216–220, 1997.
- [28] V. Ginis, M. Piccardo, M. Tamagnone, et al., “Remote structuring of near-field landscapes,” *Science*, vol. 369, no. 6502, pp. 436–440, 2020.
- [29] A. Zubair, M. Zubair, A. Danner, and M. Q. Mehmood, “Engineering multimodal spectrum of Cayley tree fractal meta-resonator supercells for ultrabroadband terahertz light absorption,” *Nanophotonics*, vol. 9, no. 3, pp. 633–644, 2020.
- [30] V. Pacheco-Peña and N. Engheta, “Effective medium concept in temporal metamaterials,” *Nanophotonics*, vol. 9, no. 2, pp. 379–391, 2020.
- [31] H. Sroor, Y. W. Huang, B. Sephton, et al., “High-purity orbital angular momentum states from a visible metasurface laser,” *Nat. Photonics*, pp. 1–6, 2020.
- [32] W. T. Chen, A. Y. Zhu, and F. Capasso, “Flat optics with dispersion-engineered metasurfaces,” *Nat. Rev. Mater.*, pp. 1–17, 2020.
- [33] V. Ginis, P. Tassin, C. M. Soukoulis, and I. Veretennicoff, “Enhancing optical gradient forces with metamaterials,” *Phys. Rev. Lett.*, vol. 110, no. 5, p. 057401, 2013.
- [34] M. Piccardo, B. Schwarz, D. Kazakov, et al., “Frequency combs induced by phase turbulence,” *Nature*, vol. 582, no. 7812, pp. 360–364, 2020.
- [35] H. Rubinsztein-Dunlop, A. Forbes, M. V. Berry, et al., “Roadmap on structured light,” *J. Optic.*, vol. 19, no. 1, p. 013001, 2016.
- [36] L. Liu, A. Di Donato, V. Ginis, S. Kheifets, A. Amirzhan, and F. Capasso, “Three-dimensional measurement of the helicity-dependent forces on a mie particle,” *Phys. Rev. Lett.*, vol. 120, no. 22, p. 223901, 2018.
- [37] A. Forbes, “Structured light from lasers,” *Laser Photon. Rev.*, vol. 13, no. 11, p. 1900140, 2019.
- [38] B. Bhaduri, M. Yessenov, and A. F. Abouraddy, “Anomalous refraction of optical spacetime wave packets,” *Nat. Photonics*, pp. 1–6, 2020.
- [39] V. Ginis, “Refracting spacetime wave packets,” *Nat. Photonics*, vol. 14, no. 7, pp. 405–407, 2020.

Supplementary Material: The online version of this article offers supplementary material (<https://doi.org/10.1515/nanoph-2020-0379>).