ACS Photonics

pubs.acs.org/journal/apchd5

Article

Inverse Design of Optical Switch Based on Bilevel Optimization Inspired by Meta-Learning

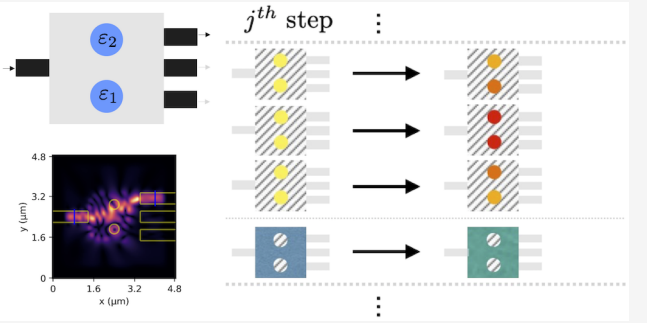
Beicheng Lou,* Jesse Alexander Rodriguez, Benjamin Wang, Mark Cappelli, and Shanhui Fan*

Cite This: *ACS Photonics* 2023, 10, 1806−1812

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: We introduce the concept of meta-learning into the design of active optical switches. An optical switch consists of both tunable and nontunable elements. It has been difficult to apply conventional inverse design methods to optical switches, since the optimal choice of the tunable elements depends on the design of the nontunable elements. Here we show that a bilevel optimization scheme, closely related to the concept of meta-learning, can be used for the design of active optical switches. In this scheme, the inner and outer loops correspond to the optimization of the tunable and nontunable elements, respectively. We illustrate this scheme with two designs of optical switches based on different tuning mechanisms. This approach is generally applicable for the design of optical switches as well as other active and tunable optical devices.



KEYWORDS: *meta-learning, inverse design, tunable, optical switch*

INTRODUCTION

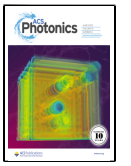
In recent years, inverse design has emerged as a prominent framework for the design of optical devices and has witnessed broad applications.[1−11] In comparison to traditional design methods based on human physical intuition, inverse design relies on computer optimization algorithms and can produce structures that are more compact or that have novel functionalities. Moreover, while most of the initial studies on inverse design in photonics focused on linear and passive structures, there are now emerging interests in applying inverse design to active and nonlinear devices such as optical switches.[12,13]

An important approach in inverse design is to use gradient-based algorithms.[14−18] In this approach, one defines an objective function that mathematically describes the design objective as a function of various design parameters. During the optimization process, one starts with a particular point in the design parameter space and computes the gradient of the objective function with respect to the design parameters at this point. A better choice of design parameters is then produced according to the gradient information. This process continues until a design with sufficiently high performance is reached. These gradient-based algorithms are closely related to the optimization algorithms developed in the machine-learning communities and have been widely used for the inverse design of passive photonic devices.[19−21]

There is, however, an interesting challenge in applying gradient-based algorithms to design optical switches.[12,13,22,23] An optical switch consists of *tunable* elements where certain physical properties such as indices can be varied, as well as *nontunable* elements where the physical properties are fixed. During the operation of the switch, the physical properties of the tunable elements are varied such that the device can switch between performing different functionalities, while the physical properties of the nontunable elements are kept constant. For example, the nontunable elements can be a region of the device where the index is kept constant, whereas the tunable elements can be a region where the index is varied. Alternatively, the nontunable elements can be a background refractive index distribution, and the tunable elements can be a small refractive index modulation of the background.

Unlike the design of a passive structure, the refractive indices of the tunable elements are not fixed during the operation of the device. Moreover, the optimal choices of refractive indices of the tunable elements for the operation of the switches are not known before the design is completed, since these choices depend on the refractive index distribution of the nontunable elements. The gradient-based algorithms therefore must take

Received: January 23, 2023
Published: June 1, 2023



https://doi.org/10.1021/acsphotonics.3c00113
ACS Photonics 2023, 10, 1806−1812

into account the different characteristics of tunable and nontunable elements.

In this paper we show that the inverse design of optical switches can be achieved with the meta-learning algorithm recently developed in the machine-learning community.[24−27] Meta-learning aims to produce a machine-learning model that can perform well on various tasks upon fine-tuning and is typically formulated as a bilevel optimization problem consisting of two nested loops, where the outer loop corresponds to producing the model itself and the inner loop is the fine-tuning. Here we apply the meta-learning concept to the design of an optical switch, where the choice of parameters for the tunable elements correspond to the inner fine-tuning loop. Our result indicates that the concept of meta-learning can be applied to the design of optical switches. This work represents a step forward in the applications of inverse design for active optical devices and highlights a connection between the literature of machine learning and optical design.

RESULTS AND DISCUSSION

As the first example that illustrates our approach, we consider the inverse design problem of a compact optical switch with one input port and three output ports. Each port consists of a single-mode waveguide of dielectric constant $\varepsilon_r = 12$. There are 3 different modes of operation where we desire output in the top, middle, and bottom waveguide, respectively. The dielectric constant is only tunable in two small disk regions of radius 0.192 μm, as shown in Figure 1a. The tunable

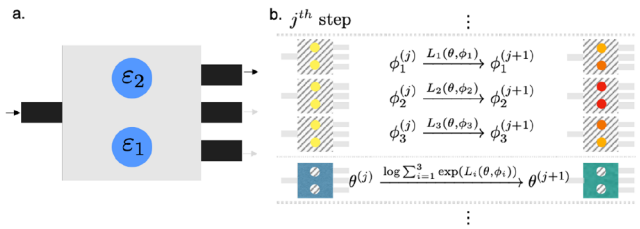


Figure 1. Illustration of (a) the problem of designing an optical switch and (b) the optimization algorithm. In (a), the objective is to design the nontunable element (gray region) such that light can be directed from one input port into one of the three different output ports when the tunable element (blue region) is set to proper values. In (b), $\{\phi_i\}_{i=1}^3$ and θ are alternatingly improved using their corresponding loss functions. During each update step, the colored regions are updated, while the shaded regions are frozen.

elements are assumed to have dielectric constants $\varepsilon_1, \varepsilon_2 \in [1, 12]$, which can be implemented with, e.g., microelectromechanical systems,[28−30] where the effective permittivity in a region can be changed mechanically by varying the fraction of high-index materials in that region. The tunable elements are surrounded by a static structure that is fixed upon fabrication, with the spatial distribution of the dielectric constant binarized at either $\varepsilon_r = 12$ or $\varepsilon_r = 1$. The static structure is the nontunable element of the switch. The design region, which includes both the tunable and nontunable elements, has a dimension of 2.4 μm × 2.4 μm and is located at the center of a simulated region with a size of 4.8 μm × 4.8 μm. The operating frequency corresponds to a free-space wavelength of 1.5 μm.

Denote the dielectric distribution in the design region as

$$\varepsilon_r(x, y) = \varepsilon_\phi(x, y) + \varepsilon_\theta(x, y) \tag{1}$$

where $\varepsilon_\phi(x, y)$ is the dielectric distribution of the tunable elements, parametrized by ϕ and nonzero only in the disk regions. $\varepsilon_\theta(x, y)$ is the dielectric distribution of the static structure, parametrized by θ and nonzero outside the disk regions. The specific form for the parametrization depends on the design constraints.[31] Here we use $\theta \in \mathbb{R}^{61\times 61}$ to define the dielectric distribution pixel by pixel, where $\varepsilon_\theta = \varepsilon_{min} + (\varepsilon_{max} - \varepsilon_{min})(1 + \tanh(\beta\theta))/2$, with $\varepsilon_{min} = 1$ and $\varepsilon_{max} = 12$. β is initialized at 500 and gradually increased toward 1500 to encourage binarization.

For each of the 3 modes of operation, denote the desired mode profile as $P_i(x, y)$, which is nonzero only on a line segment across the i^{th} output waveguide, as indicated by the blue lines in Figure 2. On the line segments, $P_i(x, y)$

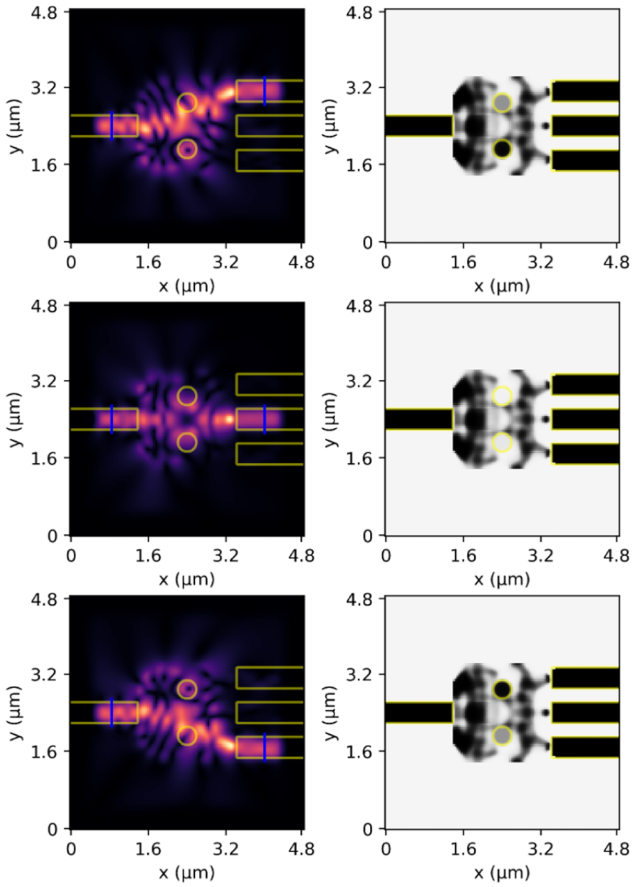


Figure 2. Field distribution (left) in the optimized structure (right) for optical switch that directs incoming light to the bottom (top), middle (middle), and top (bottom) output waveguide. Only the two disk regions are tunable, with their relative permittivities (ε_1, ε_2) being (11.9, 5.8) (top), (1.1, 1.1) (middle), and (5.8, 11.9) (bottom).

corresponds to the fundamental mode of the i^{th} output waveguide. The objective is to maximize the overlap integral between the field $E_z(x, y)$ and the desired mode profile $P_i(x, y)$. The optimal setting of the tunable elements for the i^{th} mode of operation is thus given by

$$\phi_i^*(\theta) = \underset{\phi}{\text{argmin}}\ \mathcal{L}_i(\theta, \phi) \tag{2}$$

$$\text{where}\quad \mathcal{L}_i(\theta, \phi) = -\left|\iint P_i^*(x, y)E_z(x, y)\mathrm{d}x\ \mathrm{d}y\right| \tag{3}$$

https://doi.org/10.1021/acsphotonics.3c00113
ACS Photonics 2023, 10, 1806−1812

The integration in eq 3 is over the whole 4.8 μm × 4.8 μm simulated region, but we recall that $P_i^*(x, y)$ is nonzero only at a line segment across the i^{th} output waveguide.

The optimization problem for the static structure parametrized by θ can be written as

$$\theta^* = \underset{\theta}{\text{argmin}}\ \mathcal{L}_{\text{total}}(\theta, \{\phi_i^*(\theta)\}_{i=1}^3) \tag{4}$$

$$\text{where} \quad \mathcal{L}_{\text{total}}(\theta, \{\phi_i\}_{i=1}^3) = \sum_{i=1}^{3} \mathcal{L}_i(\theta, \phi_i) \tag{5}$$

where $\varphi_i^*(\theta)$ is the optimal value of the tunable elements for the i^{th} mode of operation, found from eq 2. Here we want to ensure that all modes function equally well, so we modify eq 5 as

$$\mathcal{L}_{\text{total}}(\theta, \{\phi_i\}_{i=1}^3) = \log \sum_{i=1}^{3} \exp\ \mathcal{L}_i(\theta, \phi_i) \tag{6}$$

where the sum is replaced by the LogSumExp function[32] in order to get a worst-case optimization. The LogSumExp function is a smooth approximation to the maximum function, with $\beta^{-1}\log[\exp(\beta x_1) + \ldots + \exp(\beta x_n)] \to \max\{x_1, \ldots, x_n\}$ as $\beta \to \infty$. Therefore, our optimization objective here is related to worst-case optimizations for optical design.[33−35]

Numerically, the optimization problems in eqs 2−6 can be solved together in an interleaving fashion, as sketched in Figure 1b, by repeating the following algorithm

$$\phi_i \to \underset{\phi}{\text{argmin}}\ \mathcal{L}_i(\theta, \phi) \quad (i = 1, 2, 3) \tag{7}$$

$$\theta \to \underset{\theta}{\text{argmin}}\ \mathcal{L}_{\text{total}}(\theta, \{\phi_i\}_{i=1}^3) \tag{8}$$

In each step, we iteratively improve the parameters in the optimization problems of eqs 7 and 8. At the j^{th} step, the parameters are updated as

$$\phi_i^{(j)} = \phi_i^{(j-1)} - \alpha \cdot \frac{\partial \mathcal{L}_i(\theta^{(j-1)}, \phi_i)}{\partial \phi_i}\bigg|_{\phi_i^{(j-1)}} \tag{9}$$

$$\theta^{(j)} = \theta^{(j-1)} - \alpha \cdot \frac{\partial \mathcal{L}_{\text{total}}(\theta, \{\phi_i^{(j)}\}_{i=1}^3)}{\partial \theta}\bigg|_{\theta^{(j-1)}} \tag{10}$$

As j increases, we expect that $\varphi_i^{(j)}$ asymptotically approaches φ_i^*, and $\theta^{(j)}$ asymptotically approaches θ^*. For numerical demonstration, we optimize for 100 steps with a learning rate of $\alpha = 2 \times 10^{-4}$.

Our approach is closely related to the concept of meta-learning,[24−27] which aims to train a machine learning model to be quickly adaptable across multiple tasks. It can be formulated as a bilevel optimization problem, where for M tasks of interest, the model is parametrized by $\{\theta, \varphi_i\}_{i=1}^M$, with θ being the same across tasks and ϕ_i being different for each task $\mathcal{T}_i \in \{\mathcal{T}_i\}_{i=1}^M$. The training aims to find optimum $\{\theta^*, \varphi_i^*\}$ via

$$\theta^* = \underset{\theta}{\text{argmin}}\ \mathcal{L}_{\text{total}}(\theta, \{\phi_i^*(\theta)\}_{i=1}^3) \tag{11}$$

$$\text{where} \quad \phi_i^*(\theta) = \underset{\phi}{\text{argmin}}\ \mathcal{L}_i(\theta, \phi) \tag{12}$$

The outer-level optimization (eq 11) aims to optimize the meta parameter θ, and the inner-level optimization (eq 12) aims to optimize the task-dependent parameter ϕ_i. The inner-level optimization is typically highly constrained, e.g., in a few-shot fashion[27] where it only takes very few gradient steps and is thus computationally cheap. Our optimization of the tunable and nontunable elements, as discussed in eqs 9 and 10, corresponds closely to the inner- and outer-loop optimization in meta-learning. Despite the analogy, our optimization problem is not data-driven and is formulated as a deterministic one. Our work also has connections to multitask learning.[36,37] Here we draw the parallel to meta-learning because our inverse design problem is inherently bilevel, which formally resembles meta-learning.

For our inverse design problem as depicted in Figure 1a, the optimized structure and field distributions for the three tasks are shown in Figure 2, where the first row corresponds to task 1, second row corresponds to task 2, and third row corresponds to task 3. The structures for the tasks are shown in the right column and differ only in the two disk regions where their relative permittivities for each task are different. The field distributions on the left show that the input light is completely directed into different output waveguides as the permittivities in the two disk regions are varied. The simulation and the optimization are performed with finite-difference frequency-domain method.[31]

The dynamics of the optimization process is shown in Figure 3. The total loss function for the outer loop (Figure 3a), as

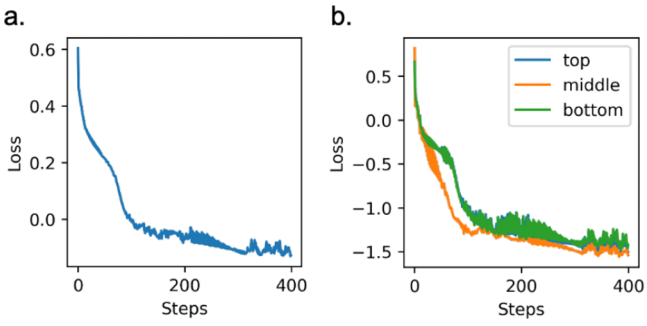


Figure 3. Loss functions evaluated at each step, for the outer loop (a), as defined in eq 6, and the inner loop (b), as defined in eq 3, respectively.

defined in eq 6, improves quickly in the first few steps, but the improvement significantly slows down after around 100 steps. The loss function for the inner loop (Figure 3b), as defined in eq 3, shows that it is much easier to direct the light to the middle waveguide than to the top and bottom waveguides. In the early stage of optimization, especially during steps 50−100, task 2 has the best performance. The gradient of the total loss function in eq 6 is largely dominated by $\mathcal{L}_2(\theta, \varphi_2)$, which corresponds to the middle output waveguide. Since eq 6 assigns more weights to tasks with bad performance, the three loss curves in Figure 3b eventually converge to very similar values. Here we did not encounter any convergence issue, unlike min−max optimizations.[38]

Figure 4 visualizes the expressivity[39] of the structure, namely, how much output variation one can achieve by varying the tunable elements, before and after the optimization of the

https://doi.org/10.1021/acsphotonics.3c00113
ACS Photonics 2023, 10, 1806−1812

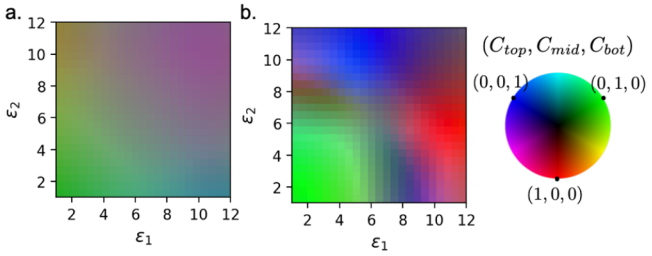


Figure 4. Expressivity of the tunable elements (a) before and (b) after optimization. The color encodes the proportion of mode overlaps with each output waveguide, where red, blue, and green correspond to the coupling efficiency into the top, middle, and bottom waveguides, respectively.

nontunable element. The color encodes the proportion of mode overlaps with each output waveguide, where the red (R), green (G), and blue (B) components in a color correspond to the amount of light directed into the top, middle, and bottom waveguides, respectively. More specifically, the coupling efficiency is defined as

$$C_i = \frac{\left|\iint P_i^*(x, y)E_z(x, y)\mathrm{d}x\ \mathrm{d}y\right|}{\left|\iint P_s^*(x, y)E_z(x, y)\mathrm{d}x\ \mathrm{d}y\right|} \tag{13}$$

where P_i corresponds to the fundamental mode profile of the i^{th} output waveguides, for i = 1, 2, and 3, and P_s corresponds to the fundamental mode profile of the input waveguide at the source location. Figure 4 plots the color $(R, G, B) = (C_{\text{top}}, C_{\text{mid}}, C_{\text{bot}})$ as a function of the dielectric constants of the two disk regions ε_1 and ε_2. Before optimization, the static structure is initialized with a uniform background permittivity of ε_r = 6.5, and the coupling efficiencies into the three output waveguides are relatively low, as shown in Figure 4a, no matter how the tunable elements are adjusted. After optimization, the static structure is parametrized by θ^*, and the coupling efficiencies into the three output waveguides can be higher than 99% when the tunable elements are properly set. Based on Figure 4b, one could initialize the tunable elements at, e.g., $\varepsilon_1 = \varepsilon_2$ = 5.5, which is close to the optimal operating values for all three switch operations, in order to reduce the permittivity variation required for the switching action. Note that, in machine-learning literature, expressivity can adopt various metrics: for a single neural network, it often characterizes sensitivity to input variation;[40] for a class of neural networks, it often characterizes what functions can be performed upon parameter tuning.[39] Here we draw the analogy to the latter.

As a second example, we consider the design of optical switch based on a different tuning mechanism. Instead of having small regions where the permittivity can be tuned over a wide range as considered above, here we consider the case where the dielectric permittivity is tunable everywhere but only within a small range. This case is typical when exploiting electro-optical effects for modulation.[41−43]

As an illustration, we consider the same design objective as above, where we aim to design a 1-by-3 optical switch, but in this case, we allow the permittivity to be tuned over the entire design region with a smaller tuning range. For this case, one can easily modify the parametrization of the dielectric distribution, by changing eq 1 to

$$\varepsilon_r(x, y) = M_\phi(x, y)^*\varepsilon_\theta(x, y) \tag{14}$$

where $1 - \eta < M_\phi(x, y) < 1 + \eta$ corresponds to a spatially dependent small perturbation of the background dielectric distribution $\epsilon_\theta(x, y)$. Here we choose η = 0.02 for illustration. The algorithm depicted in Figure 1b and the objective functions in eqs 2−6 still apply. The resulting optimized structure and field distributions are shown in Figure 5. Here

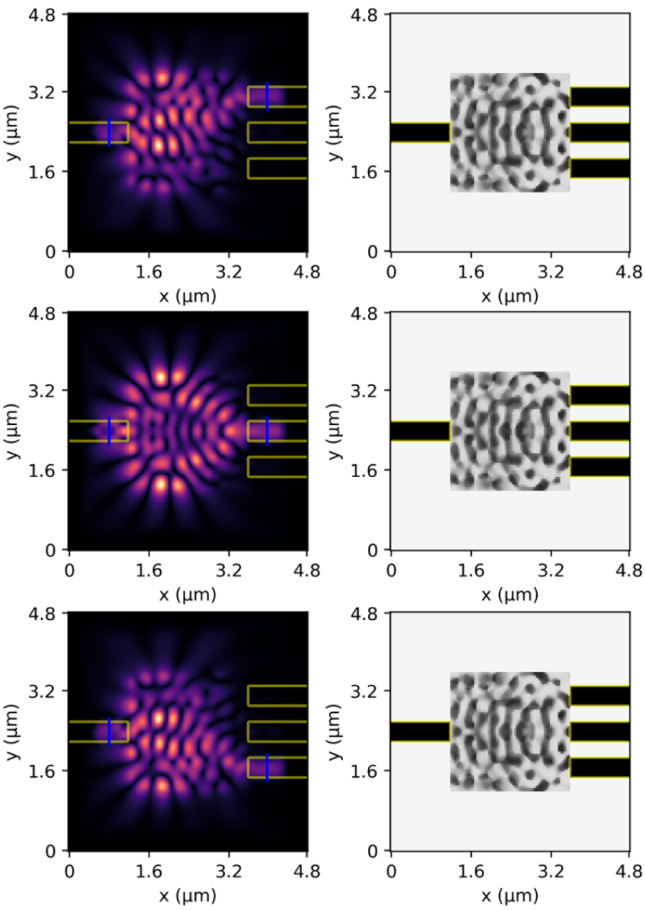


Figure 5. Field distribution (left) in the optimized structure (right) for optical switch that directs incoming light to the bottom (top), middle (middle), and top (bottom) output waveguide. Dielectric distribution in the whole design region is tunable only within 2%.

the difference between the three structures is small since the range of tunability in $M_\phi(x, y)$ is small. However, the field distributions are drastically different and perfectly meet our stated objective of this optical switch.

The dynamics of the optimization process in this case is shown in Figure 6a,b. For better convergence, we set a curriculum for the tunable range parameter η where it is initialized as 0.1 and decrements by 0.002 every 20 steps until it reaches 0.02. The loss curve started increasing after around 500 steps, indicating that the limited tunable range became a bottleneck for the performance. Such a grayscale design with pixel-level modulation is typically prone to error. We evaluated the tolerance of our design to error in Figure 6c, where the dielectric distribution with error takes the form $\varepsilon_r'(x, y) = \varepsilon_r(x, y)^*(1 + s(x, y))$, with $s(x, y)$ sampled from a Gaussian distribution $\mathcal{N}(0, \sigma^2)$ at each pixel. The result is averaged over 100 noise realizations, with the shaded band showing the confidence interval. In conventional methods,[12] the tunable parameters for each task are heuristically

https://doi.org/10.1021/acsphotonics.3c00113
ACS Photonics 2023, 10, 1806−1812

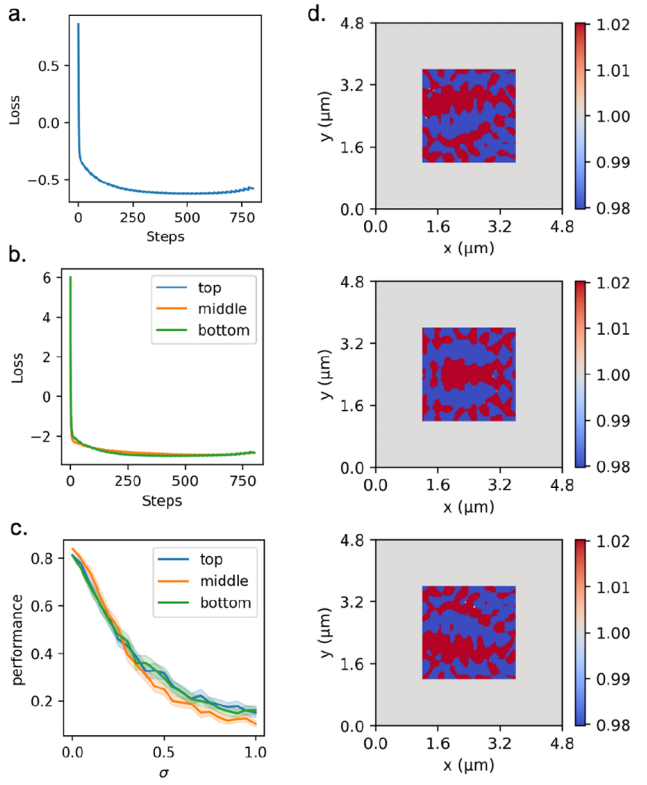


Figure 6. Loss functions evaluated at each step during training, for the outer loop (a), as defined in eq 6, and the inner loop (b), as defined in eq 3. (c) Evaluation of the design's tolerance to fabrication errors. (d) The optimal settings of the tunable part for each task.

prescribed before the optimization and therefore, the optimized structure's performance is highly dependent on the prescription quality; here our tunable part has a huge number of degrees of freedom, which makes conventional methods infeasible. The optimal settings for the tunable part, i.e. $M_\phi(x, y)$, is shown in Figure 6d, which is nontrivial for heuristic-based methodology.

In addition to these two examples, our algorithm is versatile and can be generalized to other scenarios. For example, one could desire a design involving multiple materials, each with a different tunable range of permittivity and occupying a different region, potentially with different constraints. All these complications can be formulated into the parametrization of the dielectric distribution as in eq 14 or incorporated into the objective function as regularization terms. As another example, one may desire a design that performs different functionalities at different frequencies or incident angles, but the frequencies and the incident angles themselves can be adjusted. In these cases, we can use the outer loop to optimize the design, and the optimal adjustments of the frequencies or incident angles can be obtained in the inner loop. Our algorithm is also general across different optimizers, simulation methods, and parametrizations. For example, our optimization process can be applied with 3D finite-difference frequency-domain method[44] or finite element methods[45] as well. Other switching mechanisms, such as phase-change materials,[46−48] and other types of tunable photonic devices[49−55] could also be considered.

CONCLUSION

We introduce the concept of meta-learning into the design of active optical switches. An optical switch consists of both tunable and nontunable elements. There are many limitations in conventional inverse design methods for optical switches, since the optimal choices of the tunable elements depends on the design of the nontunable elements. While it is possible to prescribe heuristic values for tunable elements in simple cases, it is highly suboptimal when more complicated tunable elements are involved. Here we show that a bilevel optimization scheme consisting of two nested loops, closely related to the concept of meta-learning, can be used for the design of active optical switches. In this scheme, the inner and outer loops correspond to the optimization of the tunable and nontunable elements, respectively. We illustrate this scheme with two designs of optical switches based on different tuning mechanisms. This approach is generally applicable for the design of optical switches as well as other active and tunable optical devices. Overall, our work points out the huge potential in exploiting bilevel optimization for inverse design of tunable optical devices.

AUTHOR INFORMATION

Corresponding Authors

Beicheng Lou − *Department of Applied Physics and Ginzton Laboratory, Stanford University, Stanford, California 94305, United States;* orcid.org/0000-0001-7846-5551; Email: beichenglou@stanford.edu

Shanhui Fan − *Department of Electrical Engineering and Ginzton Laboratory, Stanford University, Stanford, California 94305, United States;* orcid.org/0000-0002-0081-9732; Email: shanhui@stanford.edu

Authors

Jesse Alexander Rodriguez − *Department of Mechanical Engineering, Stanford University, Stanford, California 94305, United States*

Benjamin Wang − *Department of Mechanical Engineering, Stanford University, Stanford, California 94305, United States*

Mark Cappelli − *Department of Mechanical Engineering, Stanford University, Stanford, California 94305, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsphotonics.3c00113

Funding

This work is supported by a MURI Project from the U.S. Air Force Office of Scientific Research (Grant No. FA9550-21-1-0244).

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Molesky, S.; Lin, Z.; Piggott, A. Y.; Jin, W.; Vucković, J.; Rodriguez, A. W. Inverse design in nanophotonics. *Nat. Photonics* **2018**, *12*, 659−670.

(2) Liu, Z.; Zhu, D.; Raju, L.; Cai, W. Tackling Photonic Inverse Design with Machine Learning. *Advanced Science* **2021**, *8*, 2002923.

(3) Christiansen, R. E.; Sigmund, O. Inverse design in photonics by topology optimization: tutorial. *J. Opt. Soc. Am. B* **2021**, *38*, 496−509.

(4) Colburn, S.; Majumdar, A. Inverse design and flexible parameterization of meta-optics using algorithmic differentiation. *Communications Physics* **2021**, *4*, 65.
(5) Schubert, M. F.; Cheung, A. K. C.; Williamson, I. A. D.; Spyra, A.; Alexander, D. H. Inverse Design of Photonic Devices with Strict Foundry Fabrication Constraints. *ACS Photonics* **2022**, *9*, 2327−2336.
(6) Rodríguez, J. A.; Abdalla, A. I.; Wang, B.; Lou, B.; Fan, S.; Cappelli, M. A. Inverse Design of Plasma Metamaterial Devices for Optical Computing. *Phys. Rev. Applied* **2021**, *16*, 014023.
(7) Rodríguez, J. A.; Cappelli, M. A. Inverse design of plasma metamaterial devices with realistic elements. *J. Phys. D: Appl. Phys.* **2022**, *55*, 465203.
(8) Wang, J.; Shi, Y.; Hughes, T.; Zhao, Z.; Fan, S. Adjoint-based optimization of active nanophotonic devices. *Opt. Express* **2018**, *26*, 3236−3248.
(9) Jin, W.; Li, W.; Orenstein, M.; Fan, S. Inverse Design of Lightweight Broadband Reflector for Relativistic Lightsail Propulsion. *ACS Photonics* **2020**, *7*, 2350−2355.
(10) Ahn, G. H.; Yang, K. Y.; Trivedi, R.; White, A. D.; Su, L.; Skarda, J.; Vučković, J. Photonic Inverse Design of On-Chip Microresonators. *ACS Photonics* **2022**, *9*, 1875−1881.
(11) Vercruysse, D.; Sapra, N. V.; Yang, K. Y.; Vučković, J. Inverse-Designed Photonic Crystal Circuits for Optical Beam Steering. *ACS Photonics* **2021**, *8*, 3085−3093.
(12) Chung, H.; Miller, O. D. Tunable Metasurface Inverse Design for 80% Switching Efficiencies and 144°Angular Deflection. *ACS Photonics* **2020**, *7*, 2236−2243.
(13) Di Domenico, G.; Weisman, D.; Panichella, A.; Roitman, D.; Arie, A. Large-Scale Inverse Design of a Planar On-Chip Mode Sorter. *ACS Photonics* **2022**, *9*, 378−382.
(14) Mao, S.; Cheng, L.; Zhao, C.; Khan, F. N.; Li, Q.; Fu, H. Y. Inverse Design for Silicon Photonics: From Iterative Optimization Algorithms to Deep Neural Networks. *Applied Sciences* **2021**, *11*, 3822.
(15) Chen, Y.; Lan, Z.; Su, Z.; Zhu, J. Inverse design of photonic and phononic topological insulators: a review. *Nanophotonics* **2022**, *11*, 4347−4362.
(16) Li, Z.; Pestourie, R.; Lin, Z.; Johnson, S. G.; Capasso, F. Empowering Metasurfaces with Inverse Design: Principles and Applications. *ACS Photonics* **2022**, *9*, 2178−2192.
(17) Minkov, M.; Williamson, I. A. D.; Andreani, L. C.; Gerace, D.; Lou, B.; Song, A. Y.; Hughes, T. W.; Fan, S. Inverse Design of Photonic Crystals through Automatic Differentiation. *ACS Photonics* **2020**, *7*, 1729−1741.
(18) Svanberg, K. A Class of Globally Convergent Optimization Methods Based on Conservative Convex Separable Approximations. *SIAM Journal on Optimization* **2002**, *12*, 555−573.
(19) Liu, D.; Tan, Y.; Khoram, E.; Yu, Z. Training Deep Neural Networks for the Inverse Design of Nanophotonic Structures. *ACS Photonics* **2018**, *5*, 1365−1369.
(20) Jiang, J.; Fan, J. A. Global Optimization of Dielectric Metasurfaces Using a Physics-Driven Neural Network. *Nano Lett.* **2019**, *19*, 5366−5372.
(21) Jiang, J.; Fan, J. A. Simulator-based training of generative neural networks for the inverse design of metasurfaces. *Nanophotonics* **2020**, *9*, 1059−1069.
(22) Agrahari, R.; Ghosh, S. K.; Bhattacharyya, S. *Optical Switching*; John Wiley & Sons, Ltd, 2022; Chapter 1, pp 13−30.
(23) Raja, A. S.; et al. Ultrafast optical circuit switching for data centers using integrated soliton microcombs. *Nat. Commun.* **2021**, *12*, 5867.
(24) Finn, C.; Abbeel, P.; Levine, S. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. ICML. 2017; pp 1126−1135.
(25) Rajeswaran, A.; Finn, C.; Kakade, S. M.; Levine, S. Meta-Learning with Implicit Gradients. *arXiv:1909.04630 [cs.LG]* **2019**, na.
(26) Lou, B.; Zhao, N.; Wang, J. Meta-learning from sparse recovery. Fifth Workshop on Meta-Learning at the Conference on Neural Information Processing Systems, 2021.
(27) Wang, Y.; Yao, Q.; Kwok, J. T.; Ni, L. M. Generalizing from a Few Examples: A Survey on Few-Shot Learning. *ACM Comput. Surv.* **2021**, *53*, 1.
(28) Zhang, X.; Kwon, K.; Henriksson, J.; Luo, J.; Wu, M. C. A large-scale microelectromechanical-systems-based silicon photonics LiDAR. *Nature* **2022**, *603*, 253−258.
(29) Arbabi, E.; Arbabi, A.; Kamali, S. M.; Horie, Y.; Faraji-Dana, M.; Faraon, A. MEMS-tunable dielectric metasurface lens. *Nat. Commun.* **2018**, *9*, 812.
(30) Qiao, Q.; Liu, X.; Ren, Z.; Dong, B.; Xia, J.; Sun, H.; Lee, C.; Zhou, G. MEMS-Enabled On-Chip Computational Mid-Infrared Spectrometer Using Silicon Photonics. *ACS Photonics* **2022**, *9*, 2367−2377.
(31) Hughes, T. W.; Williamson, I. A.; Minkov, M.; Fan, S. Forward-Mode Differentiation of Maxwell's Equations. *ACS Photonics* **2019**, *6*, 3010−3016.
(32) Goodfellow, I.; Bengio, Y.; Courville, A. Deep Learning. MIT Press, 2016; http://www.deeplearningbook.org (accessed 2023−04−01).
(33) Schevenels, M.; Lazarov, B.; Sigmund, O. Robust topology optimization accounting for spatially varying manufacturing errors. *Computer Methods in Applied Mechanics and Engineering* **2011**, *200*, 3613−3627.
(34) Lebbe, N.; Dapogny, C.; Oudet, E.; Hassan, K.; Gliere, A. Robust shape and topology optimization of nanophotonic devices using the level set method. *J. Comput. Phys.* **2019**, *395*, 710−746.
(35) Hammond, A. M.; Oskooi, A.; Johnson, S. G.; Ralph, S. E. Photonic topology optimization with semiconductor-foundry design-rule constraints. *Opt. Express* **2021**, *29*, 23916−23938.
(36) Wang, H.; Zhao, H.; Li, B. Bridging multi-task learning and meta-learning: Towards efficient training and effective adaptation. *International Conference on Machine Learning*, 2021; pp 10991−11002.
(37) Abdollahzadeh, M.; Malekzadeh, T.; Cheung, N.-M. M. Revisit multimodal meta-learning through the lens of multi-task learning. *Advances in Neural Information Processing Systems* **2021**, *34*, 14632−14644.
(38) Hsieh, Y.-P.; Mertikopoulos, P.; Cevher, V. The Limits of Min-Max Optimization Algorithms: Convergence to Spurious Non-Critical Sets. *Proceedings of the 38th International Conference on Machine Learning*, 2021; pp 4337−4348.
(39) Zhang, C.; Bengio, S.; Hardt, M.; Recht, B.; Vinyals, O. Understanding deep learning requires rethinking generalization. *International Conference on Learning Representations*, 2017.
(40) Poole, B.; Lahiri, S.; Raghu, M.; Sohl-Dickstein, J.; Ganguli, S. Exponential expressivity in deep neural networks through transient chaos. *Advances in Neural Information Processing Systems* **2016**, na.
(41) Sinatkas, G.; Christopoulos, T.; Tsilipakos, O.; Kriezis, E. E. Electro-optic modulation in integrated photonics. *J. Appl. Phys.* **2021**, *130*, 010901.
(42) Qiao, L.; Tang, W.; Chu, T. 32 × 32 silicon electro-optic switch with built-in monitors and balanced-status units. *Sci. Rep.* **2017**, *7*, 42306.
(43) Sharkawy, A.; Shi, S.; Prather, D. W.; Soref, R. A. Electro-optical switching using coupled photonic crystal waveguides. *Opt. Express* **2002**, *10*, 1048−1059.
(44) Shin, W.; Fan, S. Choice of the perfectly matched layer boundary condition for frequency-domain Maxwell's equations solvers. *J. Comput. Phys.* **2012**, *231*, 3406−3431.
(45) Jin, J. *The Finite Element Method in Electromagnetics*, 3rd ed.; Wiley-IEEE Press, 2014.
(46) Zhang, Y.; et al. Broadband transparent optical phase change materials for high-performance nonvolatile photonics. *Nat. Commun.* **2019**, *10*, 4279.
(47) Abdollahramezani, S.; Hemmatyar, O.; Taghinejad, H.; Krasnok, A.; Kiarashinejad, Y.; Zandehshahvar, M.; Alu, A.; Adibi, A. Tunable nanophotonics enabled by chalcogenide phase-change materials. *Nanophotonics* **2020**, *9* (5), 1189−1241.

https://doi.org/10.1021/acsphotonics.3c00113
ACS Photonics 2023, 10, 1806−1812

(48) Ko, J. H.; Yoo, Y. J.; Lee, Y.; Jeong, H.-H.; Song, Y. M. A review of tunable photonics: Optically active materials and applications from visible to terahertz. *IScience* **2022**, *25* (8), 104727.
(49) He, Q.; Sun, S.; Zhou, L. unable/Reconfigurable Metasurfaces: Physics and Applications. *Research* **2019**, *2019*, 1.
(50) Lou, B.; Zhao, N.; Minkov, M.; Guo, C.; Orenstein, M.; Fan, S. Theory for Twisted Bilayer Photonic Crystal Slabs. *Phys. Rev. Lett.* **2021**, *126*, 136101.
(51) Guo, C.; Guo, Y.; Lou, B.; Fan, S. Wide wavelength-tunable narrow-band thermal radiation from moiré patterns. *Applied Physics Letters* **2021**, *118* (13), 131111.
(52) Lou, B.; Fan, S. Tunable Frequency Filter Based on Twisted Bilayer Photonic Crystal Slabs. *ACS Photonics* **2022**, *9* (3), 800−805.
(53) Lou, B.; Wang, B.; Rodríguez, J. A; Cappelli, M.; Fan, S. Tunable guided resonance in twisted bilayer photonic crystal. *Science Advances* **2022**, *8* (48), eadd4339.
(54) Tang, H.; Lou, B.; Du, F.; Zhang, M.; Ni, X.; Xu, W.; Jin, R.; Fan, S.; Mazur, E. On-chip optical twisted bilayer photonic crystal. *arXiv* **2023**, arXiv:2303.02325.
(55) Abdelraouf, O. A. M.; Wang, Z.; Liu, H. Recent Advances in Tunable Metasurfaces: Materials, Design, and Applications. *ACS Nano* **2022**, *16* (9), 13339−13369.

https://doi.org/10.1021/acsphotonics.3c00113
ACS Photonics 2023, 10, 1806−1812