



Adjoint-based optimization of active nanophotonic devices

JIAHUI WANG,¹ YU SHI,² TYLER HUGHES,¹ ZHEXIN ZHAO,² AND SHANHUI FAN^{2,*}

¹Department of Applied Physics, Stanford University, 348 Via Pueblo, Stanford, CA, 94305, USA

²Department of Electrical Engineering, Stanford University, 350 Serra Mall, Stanford, CA 94305, USA

*shanhui@stanford.edu

Abstract: We show that the adjoint variable method can be combined with the multi-frequency finite-difference frequency-domain method for efficient sensitivity calculations, enabling the systematic optimization of active nanophotonic devices. As a proof of principle demonstration, we have optimized a dynamic isolator structure in two-dimensions, resulting in the reduction of the length of the modulated regions by a factor of two, while retaining good performance in the isolation ratio and insertion loss.

© 2018 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

OCIS codes: (000.3860) Mathematical methods in physics; (130.4110) Modulators; (230.3240) Isolators.

References and links

1. V. J. Sorger, N. D. Lanzillotti-Kimura, R.-M. Ma, and X. Zhang, "Ultra-compact silicon nanophotonic modulator with broadband response," *Nanophotonics* **1**(1), 17–22 (2012).
2. A. Liu, R. Jones, L. Liao, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu, and M. Paniccia, "A high-speed silicon optical modulator based on a metal–oxide–semiconductor capacitor," *Nature* **427**(6975), 615–618 (2004).
3. G. T. Reed, G. Mashanovich, F. Y. Gardes, and D. J. Thomson, "Silicon optical modulators," *Nat. Photonics* **4**(8), 518–526 (2010).
4. W. Cai, J. S. White, and M. L. Brongersma, "Compact, high-speed and power-efficient electrooptic plasmonic modulators," *Nano Lett.* **9**(12), 4403–4411 (2009).
5. P. Dong, S. Liao, D. Feng, H. Liang, D. Zheng, R. Shafiiha, C.-C. Kung, W. Qian, G. Li, X. Zheng, A. V. Krishnamoorthy, and M. Asghari, "Low vpp, ultralow-energy, compact, high-speed silicon electro-optic modulator," *Opt. Express* **17**(25), 22484–22490 (2009).
6. J. Liu, M. Beals, A. Pomerene, S. Bernardis, R. Sun, J. Cheng, L. C. Kimerling, and J. Michel, "Waveguide-integrated, ultralow-energy GeSi electro-absorption modulators," *Nat. Photonics* **2**(7), 433–437 (2008).
7. D. A. Miller, "Attojoule optoelectronics for low-energy information processing and communications," *J. Lightwave Technol.* **35**(3), 346–396 (2017).
8. A. Taflov and S. C. Hagness, *Computational Electrodynamics: The Finite-difference Time-domain Method* (Artech House, 2005).
9. Y. Shi, W. Shin, and S. Fan, "Multi-frequency finite-difference frequency-domain algorithm for active nanophotonic device simulations," *Optica* **3**(11), 1256–1259 (2016).
10. M. Tymchenko, D. Sounas, and A. Alù, "Composite Floquet scattering matrix for the analysis of time-modulated systems," in *IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting (IEEE, 2017)*, pp. 65–66.
11. G. Veronis, R. W. Dutton, and S. Fan, "Method for sensitivity analysis of photonic crystal devices," *Opt. Lett.* **29**(19), 2288–2290 (2004).
12. D. Sell, J. Yang, S. Doshay, R. Yang, and J. A. Fan, "Large-angle, multifunctional metagratings based on freeform multimode geometries," *Nano Lett.* **17**(6), 3752–3757 (2017).
13. T. Hughes, G. Veronis, K. P. Wootton, R. J. England, and S. Fan, "Method for computationally efficient design of dielectric laser accelerator structures," *Opt. Express* **25**(13), 15414–15427 (2017).
14. J. Andkjær, V. E. Johansen, K. S. Friis, and O. Sigmund, "Inverse design of nanostructured surfaces for color effects," *J. Opt. Soc. Am.* **31**(1), 164–174 (2014).
15. 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* **9**(6), 374–377 (2015).
16. L. F. Frellsen, Y. Ding, O. Sigmund, and L. H. Frandsen, "Topology optimized mode multiplexing in silicon-on-insulator photonic wire waveguides," *Opt. Express* **24**(15), 16866–16873 (2016).
17. Z. Yu and S. Fan, "Complete optical isolation created by indirect interband photonic transitions," *Nat. Photonics* **3**(2), 91–94 (2009).

18. K. Fang, Z. Yu, and S. Fan, "Photonic Aharonov-Bohm effect based on dynamic modulation," *Phys. Rev. Lett.* **108**(15), 153901 (2012).
19. L. D. Tzuang, K. Fang, P. Nussenzeig, S. Fan, and M. Lipson, "Non-reciprocal phase shift induced by an effective magnetic flux for light," *Nat. Photonics* **8**(9), 701–705 (2014).
20. L. Aplet and J. Carson, "A Faraday effect optical isolator," *Appl. Opt.* **3**(4), 544–545 (1964).
21. Y. Shoji, T. Mizumoto, H. Yokoi, I.-W. Hsieh, and R. M. Osgood Jr, "Magneto-optical isolator with silicon waveguides fabricated by direct bonding," *Appl. Phys. Lett.* **92**(7), 071117 (2008).
22. K. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas. Propag.* **14**(3), 302–307 (1966).
23. J. Lu and J. Vučković, "Nanophotonic computational design," *Opt. Express* **21**(11), 13351–13367 (2013).
24. N. K. Georgieva, S. Glavic, M. H. Bakr, and J. W. Bandler, "Feasible adjoint sensitivity technique for EM design optimization," *IEEE Trans. Microw. Theory Tech.* **50**(12), 2751–2758 (2002).
25. B. Troyanovsky, Z. Yu, and R. W. Dutton, "Physics-based simulation of nonlinear distortion in semiconductor devices using the harmonic balance method," *Comput. Methods in Appl. Mech. Eng.* **181**(4), 467–482 (2000).
26. K. S. Kundert, A. Sangiovanni-Vincentelli, "Simulation of nonlinear circuits in the frequency domain," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **5**(4), 521–535. (1986).
27. S. Esterhazy, D. Liu, M. Liertz, A. Cerjan, L. Ge, K. G. Makris, A. D. Stone, J. M. Melenk, S. G. Johnson, and S. Rotter, "Scalable numerical approach for the steady-state ab initio laser theory," *Phys. Rev. A* **90**(2), 023816 (2014).
28. W. Shin and S. Fan, "Choice of the perfectly matched layer boundary condition for frequency-domain maxwell's equations solvers," *J. Comput. Phys.* **231**(8), 3406–3431 (2012).
29. R. C. Rumpf, "Simple implementation of arbitrarily shaped total-field/scattered-field regions in finite-difference frequency-domain," *Prog. Electromagn. Res. B* **36**, 221–248 (2012).
30. S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University, 2004).
31. F. Callewaert, S. Butun, Z. Li, and K. Aydin, "Inverse design of an ultra-compact broadband optical diode based on asymmetric spatial mode conversion," *Sci. Rep.* **6**, 32577 (2016).
32. F. Wang, J. S. Jensen, and O. Sigmund, "Robust topology optimization of photonic crystal waveguides with tailored dispersion properties," *J. Opt. Soc. Am. B* **28**(3), 387–397 (2011).
33. A. Y. Piggott, J. Petykiewicz, L. Su, J. Vučković, "Fabrication-constrained nanophotonic inverse design," *Sci. Rep.* **7**, 1786 (2017).
34. G. T. Reed, G. Z. Mashanovich, F. Y. Gardes, M. Nedeljkovic, Y. Hu, D. J. Thomson, K. Li, P. R. Wilson, S.-W. Chen, and S. S. Hsu, "Recent breakthroughs in carrier depletion based silicon optical modulators," *Nanophotonics* **3**(4-5), 229–245 (2014).
35. R. W. Boyd, *Nonlinear Optics* (Academic, 2003).
36. K. Inoue, "Four-wave mixing in an optical fiber in zero-dispersion wavelength region," *J. Lightwave Technol.* **10**(11), 1553–1561 (1992).
37. R. V. Roussev, C. Langrock, J. R. Kurz, and M. M. Fejer, "Periodically poled lithium niobate waveguide sum-frequency generator for efficient single-photon detection at communication wavelengths," *Opt. Lett.* **29**(13), 1518–1520 (2004).

1. Introduction

Active nanophotonic devices, such as nanoscale electro-optical modulators [1–6], play a fundamental role in optical communication systems. Improving the modulation efficiency and reducing the footprint of these devices is essential for low-energy information processing and on-chip communication [7]. However, as typical optical materials exhibit relatively low index modulation strength, a long modulation region is generally needed to accomplish many required functionalities. As a long modulation length usually leads to high power consumption, it is important to consider design and optimization methods for reducing the device size and energy consumption for on-chip optical modulators [5–7].

Since an optical modulator is characterized by a time-dependent refractive index, it is natural to consider simulating the electromagnetic properties of an optical modulator using time-domain methods, such as the finite-difference time-domain (FDTD) method [8]. However, in practice, there is a large difference in scale between the modulation frequency and the frequency of optical waves, which makes it very expensive to simulate standard modulators using traditional time-domain methods. The recently developed multi-frequency finite-difference frequency-domain (MF-FDFD) method [9] circumvents this difficulty and can efficiently simulate active nanophotonic devices from first principles. Similar multi-frequency approach for active device simulations has also been implemented in a scattering matrix formalism [10].

In the frequency-domain, the solution of the Maxwell's equations is typically formulated as

solving a linear system of equations. Such a formalism can naturally be combined with many numerical techniques for device optimization. As one example of such numerical techniques, the adjoint variable method has received much attention in recent years as an important tool for performing sensitivity analysis in photonics [11]. With the adjoint variable method, one may efficiently calculate the gradient of an objective function with respect to any arbitrary number of design parameters by the use of only two electromagnetic simulations. This capability allows for efficient, gradient-based, large-scale optimization of optical structures. As a result, the adjoint variable method has been extensively used for the optimization of passive nanophotonic structures [12–16].

In this paper, we combine the MF-FDFD algorithm with the adjoint variable method to design and optimize active nanophotonic devices. We derive relevant equations for the adjoint-based sensitivity analysis and gradient-based optimization in the MF-FDFD framework. As a demonstration, we implement our method to optimize a non-magnetic optical isolator based on dynamic modulation [17–19]. Traditional optical isolators are usually based on magneto-optical materials [20, 21] that are not commonly used for on-chip integrated photonic circuits. In contrast, the dynamic-modulation-based isolators can be designed with standard materials that are compatible with on-chip integration. For the applications of dynamic optical isolators, there is always a need to shrink their footprint so that one can reduce their power consumption. By a combination of the adjoint variable method and the MF-FDFD method, we present a design of an optical isolator that has half the modulation length compared to the structure proposed in [18], while maintaining high isolation ratio and low insertion loss.

This paper is organized as follows. In Section 2, we review the formalism of the MF-FDFD algorithm. We then derive the adjoint variable method in the MF-FDFD framework in Section 3. In Section 4, we describe the dynamic-modulation-based optical isolator introduced in [18], which represents the starting point of our optimization efforts. In Section 5, we provide some details of the algorithms specifically related to the optimization of dynamic optical isolators. In Section 6, we show the optimized structure and its performance. We conclude in Section 7.

2. Multi-frequency finite-difference frequency-domain method

To start, we first briefly review the MF-FDFD algorithm for active nanophotonic device simulations [9], and highlight aspects that are specifically relevant for isolator design. We consider a modulated photonic device as described by a time-dependent dielectric function distribution in space:

$$\epsilon(\mathbf{r}, t) = \epsilon_s(\mathbf{r}) + \delta(\mathbf{r}) \cos(\Omega t + \phi(\mathbf{r})), \quad (1)$$

where $\epsilon_s(\mathbf{r})$ describes the underlying passive structure, $\delta(\mathbf{r})$ is the modulation strength, $\phi(\mathbf{r})$ is the modulation phase and Ω is the modulation frequency. We assume that the device is excited by a continuous wave input with a frequency ω_0 . Under this modulation, the time-dependent electric field $\tilde{\mathbf{E}}(\mathbf{r}, t)$ in the device can be expressed as the linear combination of multiple sideband components $\mathbf{E}(\mathbf{r}, \omega_n)$, where $\omega_n = \omega_0 + n\Omega$ and n is an integer:

$$\tilde{\mathbf{E}}(\mathbf{r}, t) = \text{Re} \left\{ \sum_n \mathbf{E}(\mathbf{r}, \omega_n) e^{i\omega_n t} \right\}. \quad (2)$$

By substituting Eqs. (1) and (2) into the time-dependent Maxwell's equations, we can obtain an equation for all the sideband components [9]:

$$\nabla \times \mu^{-1} \nabla \times \mathbf{E}(\omega_n) - \omega_n^2 \epsilon_s \mathbf{E}(\omega_n) - \frac{1}{2} \omega_n^2 \delta e^{i\phi} \mathbf{E}(\omega_{n-1}) - \frac{1}{2} \omega_n^2 \delta e^{-i\phi} \mathbf{E}(\omega_{n+1}) = -i\omega_0 \mathbf{J}(\omega_0) \delta_{n0}. \quad (3)$$

Here for compactness of notation, we have suppressed the spatial dependency of various variables. $\mathbf{J}(\omega_0)$ is the excited source at ω_0 , and δ_{n0} is the Kronecker delta function. By keeping only $2N + 1$

sideband components around ω_0 of Eq. (3), the equation can be written in the block-matrix form as

$$\begin{pmatrix} A_{-N} & C_{-N,-N+1} & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & C_{0,-1} & A_0 & C_{0,1} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & C_{N,N-1} & A_N \end{pmatrix} \begin{pmatrix} \mathbf{E}(\omega_{-N}) \\ \vdots \\ \mathbf{E}(\omega_0) \\ \vdots \\ \mathbf{E}(\omega_N) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ -i\omega_0 \mathbf{J}(\omega_0) \\ \vdots \\ 0 \end{pmatrix}, \quad (4)$$

where $A_n = \nabla \times \mu^{-1} \nabla \times - \omega_n^2 \epsilon_s$, $C_{n,n\pm 1} = -\frac{1}{2} \omega_n^2 \delta e^{\mp i\phi}$. Equation (4) can be written compactly as a linear system:

$$\mathbf{A} \mathbf{E} = -i\omega_0 \mathbf{J}. \quad (5)$$

When discretized on a grid, e.g. on a Yee's grid [22], the resulting system matrix A is sparse and can be solved by a variety of standard numerical linear algebra techniques. By solving Eq. (5), all sideband components can be obtained simultaneously. The technique thus circumvents the difficulty associated with the existence of vastly different time-scales in active devices modeling.

For commonly used optoelectronic materials, ϵ_s is a scalar or a symmetric tensor. As a result, the A_n 's are symmetric matrices and the underlying passive structure satisfies Lorentz reciprocity. On the other hand, with an appropriate choice of the modulation phase distribution in space, the system matrix A is not symmetric, and hence the dynamically modulated structure in general is non-reciprocal. The dynamic isolator proposed in [18] builds upon such intrinsic non-reciprocity in the system matrix of dynamically modulated systems.

3. Adjoint variable method

The formulation of Maxwell's equations for an active optical device in terms of a linear system, as outlined in the previous section, naturally lends itself for its combination with the adjoint variable method. The adjoint variable method is an efficient technique for computing the sensitivity of an objective function to a large number of degrees of freedom [23, 24]. To optimize the active nanophotonic devices with time-dependent permittivity, we combine the adjoint variable method with the MF-FDFD simulation technique. The adjoint variable method and the gradient optimization concept, however, are general and can be implemented with other simulation methods, such as the harmonic balance method [25, 26] for nonlinear circuit and laser simulations [27].

For illustration, we consider a general objective function $\mathcal{F}(\mathbf{E}(\gamma))$, where \mathbf{E} can be obtained from solving Eq. (5), and γ represents some design parameters for the system. For our active nanophotonic device, γ can be any of the parameters in Eq. (1), such as the parameters for the underlying passive structure ϵ_s , as well as those for the active modulators such as the modulation strength δ or the modulation phases ϕ .

The aim of the adjoint variable method is to compute the sensitivity $\partial \mathcal{F}(\mathbf{E}(\gamma)) / \partial \gamma$, which can be expressed as

$$\frac{\partial \mathcal{F}(\mathbf{E}(\gamma))}{\partial \gamma} = \frac{\partial \mathcal{F}}{\partial \mathbf{E}} \frac{\partial \mathbf{E}}{\partial \gamma}. \quad (6)$$

Since $\mathcal{F}(\mathbf{E}(\gamma))$ is an explicit function of \mathbf{E} , $\frac{\partial \mathcal{F}}{\partial \mathbf{E}}$ can be found analytically, and therefore we can calculate them with minimal computational cost. To determine $\partial \mathbf{E} / \partial \gamma$, we take derivatives on both sides of Eq. (5) with respect to γ :

$$\frac{\partial A}{\partial \gamma} \mathbf{E} + A \frac{\partial \mathbf{E}}{\partial \gamma} = 0 \quad (7)$$

to obtain

$$\frac{\partial \mathbf{E}}{\partial \gamma} = -A^{-1} \frac{\partial A}{\partial \gamma} \mathbf{E}. \quad (8)$$

By substituting Eq. (8) into Eq. (6), we obtain

$$\begin{aligned}\frac{\partial \mathcal{F}(\mathbf{E}(\gamma))}{\partial \gamma} &= -\left(\frac{\partial \mathcal{F}}{\partial \mathbf{E}} A^{-1}\right) \frac{\partial A}{\partial \gamma} \mathbf{E} \\ &= -\hat{\mathbf{E}}^T \frac{\partial A}{\partial \gamma} \mathbf{E},\end{aligned}\quad (9)$$

where $\hat{\mathbf{E}}$ can be obtained by solving

$$A^T \hat{\mathbf{E}} = \left(\frac{\partial \mathcal{F}}{\partial \mathbf{E}}\right)^T. \quad (10)$$

To evaluate the sensitivity using Eq. (9), we note that, A is explicit in γ , and thus $\frac{\partial A}{\partial \gamma}$ can also be analytically calculated. Therefore, the most computationally intensive step is in calculating the physical field \mathbf{E} and the adjoint field $\hat{\mathbf{E}}$. Once these two fields are computed, the sensitivity of the system to any arbitrary number of parameters can then be calculated with little additional computational cost. We have therefore shown that the adjoint variable method can be combined with the MF-FDFD method. In contrast to the adjoint variable method formalism in the regular FDFD for passive structure simulations, where the system matrix is symmetric, and hence the original system in Eq. (5) and the reciprocal system in Eq. (10) have the same system matrix, here $A \neq A^T$ as commented in Section 2. Nevertheless, if we solve the original linear system of Eq. (5) with an LU decomposition of A , the solution to the reciprocal system of Eq. (10) can still be straightforwardly obtained. In this case, the computational cost of the adjoint variable method can be further reduced to essentially the cost of solving the original linear system.

For the optimization of isolators, in general, we will be considering the contrast ratio between the scenarios of forward and backward propagations. Each of these scenarios is described by Eq. (5), but with a different source term \mathbf{J} . We denote the physical fields of each scenarios as \mathbf{E}_f and \mathbf{E}_b , respectively. The objective function then has the form $\mathcal{F}(\mathbf{E}_f, \mathbf{E}_b)$, and its sensitivity can be determined as:

$$\frac{\partial \mathcal{F}(\mathbf{E}_f, \mathbf{E}_b)}{\partial \gamma} = -\hat{\mathbf{E}}_f^T \frac{\partial A}{\partial \gamma} \mathbf{E}_f - \hat{\mathbf{E}}_b^T \frac{\partial A}{\partial \gamma} \mathbf{E}_b, \quad (11)$$

where the adjoint fields $\hat{\mathbf{E}}_f, \hat{\mathbf{E}}_b$ can be obtained by solving Eq. (10). In general, to obtain the sensitivity in Eq. (11) therefore requires four full simulations, two for the physical fields, and two for the adjoint fields. However, if we perform an LU decomposition of the system matrix A , all these fields can be obtained with far less additional computational cost afterwards with just back-substitutions. Therefore, we can determine the sensitivity of the performance of a dynamic isolator to system parameters with very little additional computational cost beyond what is needed to simulate the isolator.

4. Dynamic-modulation-based optical isolator

One important emerging application of dynamic modulation is to realize non-magnetic optical isolation [17–19]. In this section, we briefly review the concept of a dynamic-modulation-based optical isolator introduced in [18]. In the next sections, we will show that by combining the adjoint variable method with the MF-FDFD method as discussed in previous sections, we can achieve an optical isolator with half the modulation length as compared to that of [18], with the same maximum modulation strength and similar isolation performance.

We consider a slab waveguide that supports two transverse electric modes with the band structure shown in Fig. 1(a). Under the modulation described by Eq. (1), where $\Omega = \omega_1 - \omega_0$, a direct photonic transition takes place from an even mode at frequency ω_0 to an odd mode at frequency ω_1 as wave propagates through the waveguide.

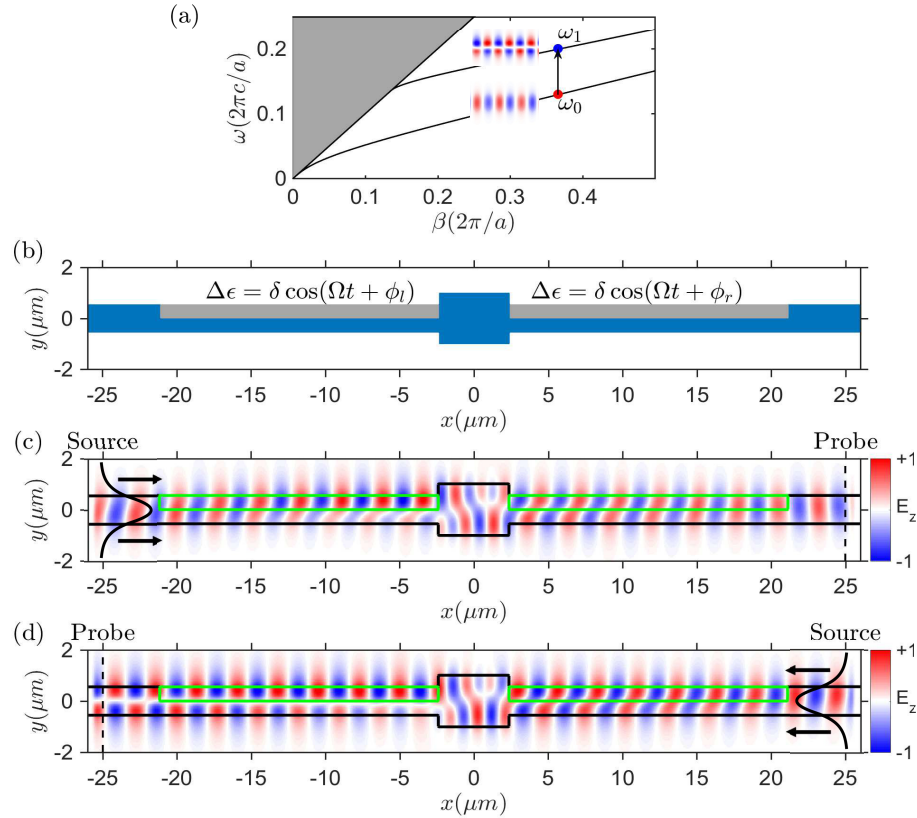


Fig. 1. Design of a dynamic optical isolator. (a) Band structure of a silicon slab waveguide. The width of the slab waveguide is $1.1\mu\text{m}$. The frequency and wavevector are normalized with respect to $a = 1.0\mu\text{m}$. The red and blue dots represent the lowest-order even and odd modes, respectively. The shaded region represents the radiative modes. (b) Optical isolator design. The width of the central waveguide is $2.0\mu\text{m}$. Modulation is applied to the upper half of the waveguide and is marked as the gray regions. The modulation phase difference between the left and right sides is $\pi/2$. (c) Forward propagation. An even mode is injected from the left side of the device. The electric field remains in the even mode after passing through the device. (d) Backward propagation. The even mode is injected from the right side of the device and is converted to the odd mode upon passing through the device.

As shown in [18], based on the concept of photonic transition as discussed above, one can construct a dynamic optical isolator using the structure shown schematically in Fig. 1(b). The structure consists of two sections of waveguides with the band structure shown in Fig. 1(a), each of them containing a modulator inducing a photonic transition as discussed above. We denote the phases of the modulation in the two modulators as ϕ_l and ϕ_r , respectively. Reciprocity is broken when $\phi_l \neq \phi_r$. And we typically choose $\phi_l - \phi_r = \pi/2$ to maximize the non-reciprocal response. The two sections of the waveguides are separated by a passive waveguide section without a modulator, but with a different width. This section enables one to convert the non-reciprocal phase response in a modulator to a non-reciprocal intensity response.

Based on the discussions above, as a baseline for our optimization effort, we simulate a waveguide isolator design similar to that in [18] with the MF-FDFD algorithm. The waveguide consists of silicon with a dielectric constant $\epsilon_{\text{Si}} = 12.25\epsilon_0$. The width of the two waveguide

sections under modulation is $w = 1.1\mu\text{m}$. We modulate these waveguide sections over half the waveguide width, shown as the gray regions in Fig. 1(b). The frequencies of the lowest-order even and odd modes are $\omega_0 = 2\pi \times 243\text{THz}$ and $\omega_1 = 2\pi \times 375\text{THz}$, respectively, and the modulation frequency is $\Omega = 2\pi \times 132\text{THz}$. We choose a uniform modulation strength $\delta/\epsilon_{\text{Si}} = 1/12.25$ within the modulation regions. The length of the modulated regions is $L_{\text{mod}} = 19\mu\text{m}$. The center waveguide section has a width $w_c = 2.0\mu\text{m}$, and a length $L_{\text{mid}} = 4.75\mu\text{m}$. These parameters are chosen such that the even and odd modes acquire a phase difference of $\pi/2$ as they propagate through the center waveguide section. In the simulation, the resolution of the Yee's grid is set to be $\Delta x = \Delta y = 50\text{ nm}$. The simulation region is surrounded by 20 layers of stretched-coordinate perfectly matched layer (SC-PML) [28]. In order to determine reflections from the device, we use the total-field scattered-field (TF/SF) technique introduced in [29] for the waveguide simulation. In our setup, the total field region is a box that encloses both the modulated regions and the center waveguide section. We place a current source on the boundary of the box to simulate a uni-directional injection of a waveguide mode.

In the MF-FDFD simulation, we only need to consider three frequency components, $\omega_0 - \Omega$, ω_0 , $\omega_0 + \Omega = \omega_1$, since all other field components are non-resonant and have very small amplitudes. Figures 1(c) and 1(d) show the simulation results. An even mode injected from the left is near-completely transmitted while maintaining the modal profile. In contrast, the even mode injected from the right becomes an odd mode upon passing through the structure. The response illustrated in Figs. 1(c) and 1(d) indicates a strong non-reciprocal effect. Based on this effect, one can construct an isolator by combining the structure here with a reciprocal modal filter.

In order to evaluate the performance of the isolator, we define the isolation ratio and insertion loss. We use $T_{m \rightarrow n}$ and $T_{n \leftarrow m}$ to represent the transmission coefficients to the n -th outgoing waveguide mode from the m -th incoming waveguide mode. Here \rightarrow and \leftarrow denote the forward and backward propagations respectively. In the simulations, these transmission coefficients are determined by integrating along the probe lines as shown in Figs. 1(c) and 1(d), for the forward and backward propagations, respectively. As an example,

$$T_{0 \rightarrow n} = \frac{1}{2\omega_n} \text{Re} \int_{-\infty}^{\infty} dy \mathbf{E}(x_p, y, \omega_n) \times \mathbf{H}^*(x_p, y, \omega_n) \cdot \hat{x}, \quad (12)$$

when the injected photon flux in the 0-th mode is unity. Here, x_p is the position of the probe line. The notation for the reflection coefficients $R_{m \rightarrow n}$ and $R_{n \leftarrow m}$ are defined similarly and are obtained by similar integration of the reflected fields at a probe line near the source. The isolation ratio in the unit of dB is defined as

$$C = 10 \log_{10} \frac{T_{0 \rightarrow 0}}{T_{0 \leftarrow 0}}. \quad (13)$$

And the insertion loss in the unit of dB is defined as

$$L = 10 \log_{10} T_{0 \rightarrow 0}. \quad (14)$$

With these definitions, the isolation ratio and insertion loss of the structure Fig. 1(b) are 27.5 dB and -0.16 dB at the operating frequency, respectively.

5. Optimization algorithms for dynamic optical isolators

In this section, we demonstrate that the adjoint variable method can be applied with the MF-FDFD method to optimize actively modulated photonic devices. As an illustration, we show that one can improve the dynamic optical isolator concept as discussed in the previous section.

The setup for the optimization is shown in Fig. 2. The starting point of our design is the same passive structure as shown in Fig. 1(b). Our objective will be to reduce the length of the

modulation regions by a factor of 2, while maintaining high isolation ratio and low insertion loss. For the structure shown in Fig. 1(b), if we reduce the length of the modulation regions by a factor of 2 without making any other change, the isolation ratio is only 2.50 dB. In our optimization, we will design the dielectric structure in the modulation regions with reduced length in order to improve the performance of the isolator. In this section, we discuss some of the details of the optimization algorithms.

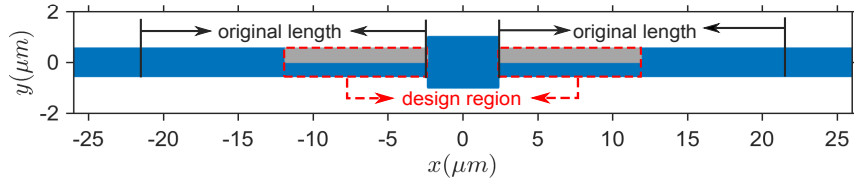


Fig. 2. Optimization Setup. The objective of the optimization is to design the dielectric structure in the regions as indicated by the red dashed lines, such that the lengths of the modulated regions are reduced by a factor of 2 as compared with the original structure shown in Fig. 1(b).

5.1. Objective function

To perform a systematic optimization of the isolator, we seek to minimize an objective function:

$$\mathcal{F}(\mathbf{E}_f, \mathbf{E}_b, \epsilon_s) = \tilde{\mathcal{F}}(\mathbf{E}_f, \mathbf{E}_b) + \kappa \sum_{(i,j)} (\epsilon_{si} - \epsilon_{ij})(\epsilon_{ij} - \epsilon_0). \quad (15)$$

In Eq. (15),

$$\begin{aligned} \tilde{\mathcal{F}}(\mathbf{E}_f, \mathbf{E}_b) = & W_1 T_{0 \leftarrow 0}^2 + W_2 (1 - T_{0 \rightarrow 0})^2 + W_3 (1 - T_{1 \leftarrow 0})^2 + W_4 T_{0 \rightarrow 1}^2 \\ & + W_5 \left(\sum_n R_{0 \rightarrow n} \right)^2 + W_6 \left(\sum_n R_{n \leftarrow 0} \right)^2 \end{aligned} \quad (16)$$

is explicit in \mathbf{E}_f and \mathbf{E}_b which are the fields in the forward and backward propagation scenarios. The weights W_n 's are chosen to balance the various requirements for isolator performance. Specifically, minimizing the first term in Eq. (16) will suppress the backward propagation of the even mode, and minimizing the second term will result in the high transmission of the even mode in the forward direction. Minimizing the sum of the first and second terms therefore will result in high contrast ratio for the even mode. Similarly, minimizing the sum of the third and the fourth terms will result in a high contrast ratio for the odd mode. The fifth and sixth terms are introduced to suppress the reflections in both directions in the optimized structures. In our simulation, W_1 is set to be 100 and all other W_n 's are set to be 1. In the objective function of Eq. (15), the second term $\kappa \sum_{(i,j)} (\epsilon_{si} - \epsilon_{ij})(\epsilon_{ij} - \epsilon_0)$ is added since we hope to get a structure with binary permittivity values corresponding to those of silicon and air after the optimization is concluded. ϵ_{ij} represents the permittivity of the underlying passive structure at the (i, j) -th grid point on the two-dimensional Yee's grid. κ is chosen after calculating $\partial \tilde{\mathcal{F}} / \partial \epsilon_{ij}$ at the first iteration step such that the gradient of this term has the same order of magnitude as that of $\tilde{\mathcal{F}}$.

With the objective function defined above, our optimization problem then becomes

$$\begin{aligned} & \underset{\epsilon_s}{\text{minimize}} \quad \mathcal{F}(\mathbf{E}_f, \mathbf{E}_b, \epsilon_s) \\ & \text{subject to} \quad \epsilon_0 \leq \epsilon_{ij} \leq \epsilon_{si}. \end{aligned} \quad (17)$$

In our simulation, we focus on the dielectric structure design, where the design parameters are the permittivity of the underlying passive structure in the design regions. However, one can similarly optimize the objective function with respect to the modulation phase and the modulation strength as well.

5.2. Gradient-based optimization

Using the method as detailed in Section 3, in each step of the optimization, we calculate the gradient of the objective function $\partial\mathcal{F}/\partial\epsilon_{ij}$ for all the grid points by the adjoint variable method. We then apply the gradient descent with momentum [13, 30] to reduce the value of the objective function. The gradient descent modifies the underlying passive structure at each grid point by

$$\begin{aligned}\Delta\epsilon_{ij} &:= \alpha \frac{\partial\mathcal{F}}{\partial\epsilon_{ij}} + \eta\Delta\epsilon_{ij}, \\ \epsilon_{ij} &:= \epsilon_{ij} - \Delta\epsilon_{ij},\end{aligned}\tag{18}$$

where α is the step size and η controls the influence of the gradient direction from the previous step. Both α and η can be tuned to speed up the convergence rate. Here, we choose the initial step size α such that the largest relative permittivity change $\max_{(i,j)} \Delta\epsilon_{ij} = 0.1$, and we adjust it with the change of $\partial\mathcal{F}/\partial\epsilon_{ij}$ for each subsequent iteration. η is set to be a constant 0.7 during the iterations. To limit the ϵ_{ij} so that it stays within the range between ϵ_0 and ϵ_{Si} , after each iteration, we set

$$\epsilon_{ij} = \begin{cases} \epsilon_0 & \epsilon_{ij} \leq \epsilon_0 \\ \epsilon_{Si} & \epsilon_{ij} \geq \epsilon_{Si} \end{cases}.\tag{19}$$

Within the modulation regions, the modulated permittivity is described by Eq. (1) with a uniform modulation strength δ , except when ϵ_{ij} at a particular grid point decreases to ϵ_0 , in which case the modulation strength is then also set to zero at such grid point. We repeat the iteration step described above until convergence on where \mathcal{F} no longer changes, or when we reach the maximum number of iteration steps.

After the optimization is concluded, we discretize the remaining non-binary values with

$$\epsilon_{ij} = \begin{cases} \epsilon_0 & \epsilon_{ij} \leq (\epsilon_0 + \epsilon_{Si})/2 \\ \epsilon_{Si} & \text{otherwise} \end{cases},\tag{20}$$

and we update the modulation strength profile accordingly. As we will see in the specific example below, using the objective function as discussed above, the optimization tends to produce structures that are almost completely binary, consisting only of silicon and air. Therefore, the modification in the last step is typically minimal.

5.3. Summary of the optimization procedure

In Fig. 3, we summarize the optimization procedure for our isolator design. We start with the reduced-length structure, and compute the physical fields \mathbf{E}_f and \mathbf{E}_b , which corresponds to the forward and backward propagation scenarios, by solving Eq. (3). We then compute the corresponding adjoint fields $\hat{\mathbf{E}}_f$ and $\hat{\mathbf{E}}_b$ by solving Eq. (10). Using the physical and the adjoint fields, we then compute the sensitivity of the objective function to the permittivity change using Eq. (11). Using such sensitivity, we update the dielectric structure using the method of gradient descent with momentum, as shown in Eq. (18). We repeat the process until either the objective function converges, or we have reached a maximum number of iteration steps. At this point, we convert the resulting structure to a binary structure with only two permittivities, using the procedure described in Eq. (20). Below, we will show the numerical results obtained using this optimization procedure.

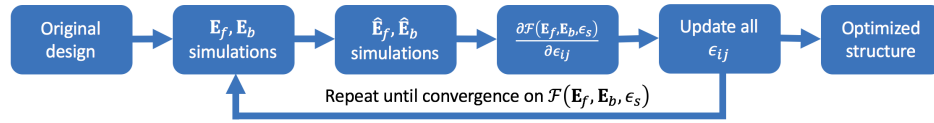


Fig. 3. The flowchart of the optimization process

6. Optimization results

With the algorithms described above, we can obtain the final optimized structure shown in Fig. 4(a). Different from the original structure, the optimized device has an aperiodic arrangement of air holes having irregular shapes and sizes, generated by the optimization algorithms in the design regions. To show that the optimized structure functions as an optical isolator, we plot the steady state field distributions in Figs. 4(b) and 4(c). In the forward direction, the even mode incident from the left passes through the device without modal conversion. Whereas in the backward direction, an even mode incident from the right is near-completely converted to the odd mode upon passing through the structure. The contrast between the forward and backward directions indicates strong non-reciprocity in this system. At the operating frequency, the isolation ratio and insertion loss of the optimized structure are 21.25 dB and -0.05 dB, respectively. The slight deviation of the field from an expected even mode profile at the source location as shown in Fig. 4(c) comes from the interference of the source and the small back reflections from the isolator. Such distortion can also be observed in the original structure as shown in Figs. 1(c) and 1(d). For the original structure, the back reflections are both 7.0% for forward and backward propagations. For the optimized structure, the back reflections are 0.2% and 4.0% for forward and backward propagations, respectively. For the optimized structure, the back reflections are reduced, since there is a term in our objective function that seeks to minimize reflections.

We obtain the optimized structure as discussed above after 500 iterations of Eq. (18). In Figs. 4(d) and 4(e), we provide additional information about the intermediate structures before we reach the optimized structure. Figure 4(d) shows the convergence of the objective function \mathcal{F} . The objective function value decreases rapidly in the first 200 iterations and converges to a constant below 0.01. Since we hope to achieve a binary structure that has only silicon and air as the constituent materials, to characterize how close the intermediate structures are to being binary, we define a binary characterization value B [31]:

$$B = \text{mean}_{(i,j)} \left| 2 \frac{\epsilon_{ij} - \epsilon_0}{\epsilon_{\text{Si}} - \epsilon_0} - 1 \right|. \quad (21)$$

B is 1 when all ϵ_{ij} is equal to either ϵ_0 or ϵ_{Si} and is 0 when all $\epsilon_{ij} = (\epsilon_0 + \epsilon_{\text{Si}})/2$. At the starting point, the design regions are filled with silicon, thus $B = 1$. During the optimization, B first drops to near 0.9, thus the intermediate structures contain regions with permittivity between that of silicon and air. However, after 500 iterations, B converges back to 1, thus the final structure after the iterations and before we apply the final step as described in Eq. (20) already is nearly binary having mostly silicon and air as its constituent materials. The convergence of B indicates the effectiveness of the objective function in Eq. (15) for selecting binary structures with good device performance.

However, we most likely have not reached the global optimum, due to the large number of degrees of freedom in the structure. We have repeated the same optimization procedure, but with a set of starting structures consisting of a random dielectric distribution in the design region. The resulting structures generally performs worse as compared to the optimal structure presented above. It appears that the uniform waveguide structure with a reduced modulation length, which

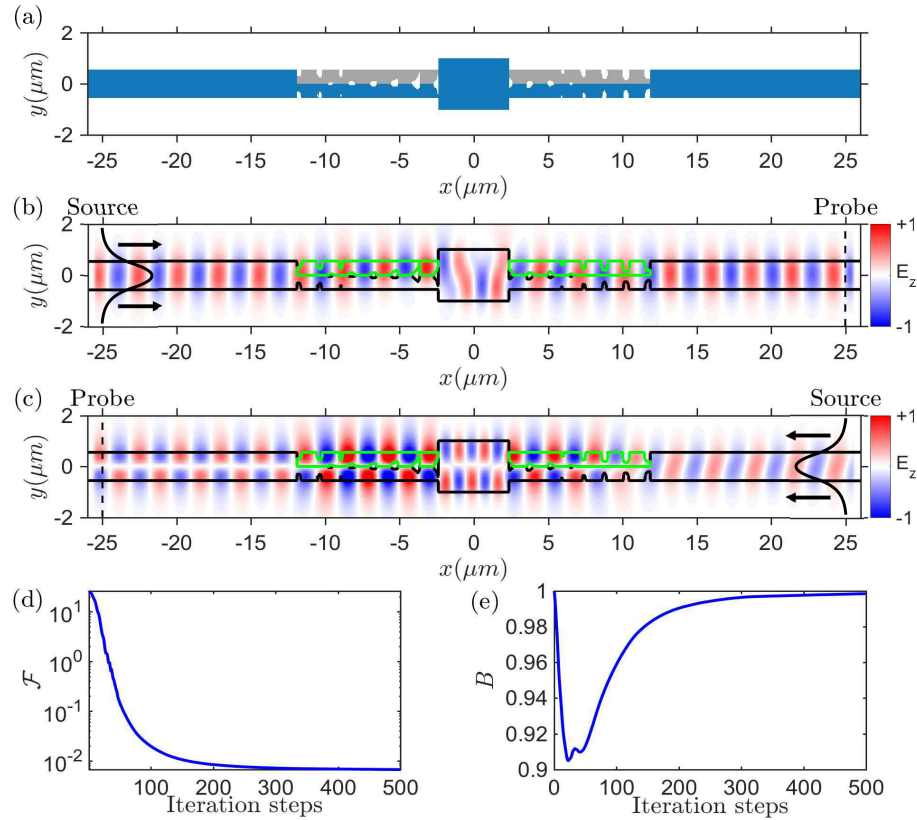


Fig. 4. The optimized optical isolator. (a) Optimized structure. The colored regions (including grays and blue regions) represent the silicon device, which is surrounded by free space. The gray regions indicate the modulated regions. Notice that the length of the modulated region (enclosed by the green lines) is half the length in the original structure as shown in Fig. 1(b). (b) Forward propagation. An even mode is injected from the left side of the device. The electric field remains in the even mode after passing through the device. (c) Backward propagation. The even mode is injected from the right side of the device and is converted to the odd mode upon passing through the device. (d) Objective function values as a function of iteration steps. (e) Binary characterization values as a function of iteration steps (see [Visualization 1](#) for successive convergence of the optimized structure).

is the starting point of our optimization presented in the paper, is "closer" to the optimal device in some sense than a random starting point. The operation of our device relies on the interactions between waveguide modes. A random starting structure strongly perturbs the modal properties of the waveguide.

In Fig. 5, we plot the spectra of the isolation ratio and insertion loss of the original structure (Fig. 1(b)), the reduced-length structure where we reduce the length of the modulation regions by a factor of 2 without optimization (Fig. 2), and the optimized structure (Fig. 4(a)). The original structure has high contrast ratio and low insertion loss over a broad bandwidth. In the reduced-length structure, the isolation ratio decreases drastically, but the insertion loss remains low. Compared with the reduced length structure, the optimized structure, which has the same length of modulation regions, has a much higher isolation ratio while maintaining a low insertion loss. Note that the bandwidth of the optimized structure is narrower than that of the original

full-length structure due to the weak resonance effect introduced by the aperiodic structure, which is essential for the reduction of the length of the modulated regions.

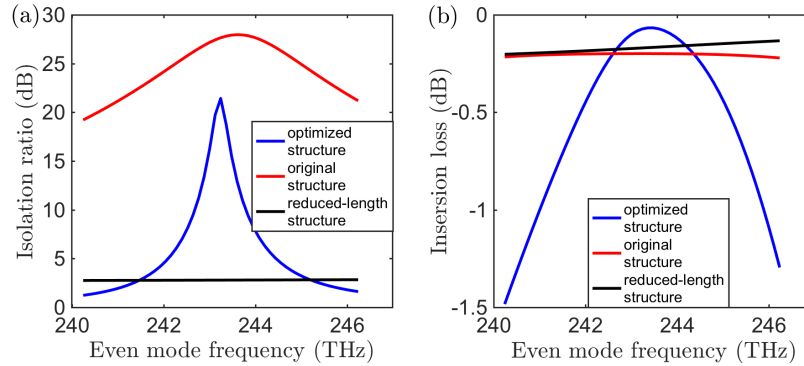


Fig. 5. (a) The spectrum of the isolation ratio for each isolator structure. (b) The spectrum of the insertion loss for each isolator structure. The red lines are the spectra for the original structure as shown in Fig. 1. The black lines are for the reduced-length structure before optimization as shown in Fig. 2. And the blue lines are for the optimized structure as shown in Fig. 4.

Furthermore, to analyze the fabrication tolerance of the optimized structure, we perform two additional calculations where the size of air holes in the optimized structure are increased or reduced by 10 nm. Such perturbations correspond to typical errors in standard fabrication processes [32]. For the structure where the hole sizes are increased, the isolation ratio and insertion loss are 9.8 dB and -0.19 dB, respectively. For the structure where the hole sizes are reduced, the isolation ratio and insertion loss are 15.0 dB and -0.14 dB, respectively. Thus, the optimized device still operates reasonably well even in the presence of such fabrication errors. For future works, in order to make the optimized structure more stable against fabrication errors, we can add structure robustness constraints as one of the minimization terms in the objective function [33].

For a proof of principle demonstration, to reduce the modulation length, we use a high modulation frequency and large modulation strength. The optimization algorithm can be applied for modulators with realistic modulation frequency and modulation strength [34]. In addition, we expect that this adjoint variable method could be applied to the optimization of nonlinear optical isolators based on frequency-mixing, where an optical pump beam couples two signal beam together [35–37]. In the simulation above, we construct the isolator with lossless medium. However, our optimization algorithm can also be used to optimize structures containing lossy materials. Our optimized structure retains its performance characteristics when modest amount of loss is added by setting a non-zero imaginary part of dielectric permittivity.

7. Conclusions

We have shown that the adjoint variable method can be combined with the multi-frequency finite-difference frequency-domain method to enable systematic optimization of active nanophotonic devices. As a proof of principle demonstration, we have optimized a dynamic isolator structure in two-dimensions, where the optimization results in the reduction of the length of the modulated regions by a factor of two, while retaining good performance in isolation ratio and insertion loss. The scheme that we discuss here should be generally applicable for the optimization of low-energy-consumption nanophotonic devices.

Funding

U. S. Defense Advanced Research Project Agency (DARPA) (HR00111720034); U. S. Air Force Office of Scientific Research (FA9550-15-1-0335 and FA9550-17-1-0002).

Acknowledgments

The authors would like to thank Sacha Verweij for fruitful discussions.