Adjoint Method and Inverse Design for Nonlinear Nanophotonic Devices

Tyler W. Hughes,[‡] Momchil Minkov,[‡] Ian A. D. Williamson, and Shanhui Fan*

Department of Electrical Engineering and Ginzton Laboratory, Stanford University, Stanford, California 94305, United States

S *Supporting Information*

ABSTRACT: The development of inverse design, where computational optimization techniques are used to design devices based on certain specifications, has led to the discovery of many compact, nonintuitive structures with superior performance. Among various methods, large-scale, gradient-based optimization techniques have been one of the most important ways to design a structure containing a vast number of degrees of freedom. These techniques are made possible by the adjoint method, in which the gradient of an objective function with respect to all design degrees of freedom can be computed using only two full-field simulations. However, this approach has so far mostly been applied to linear photonic devices. Here, we present an extension of this method to modeling nonlinear devices in the frequency domain, with the nonlinear response directly included in the gradient computation. As illustrations, we use the method to devise compact photonic switches in a Kerr nonlinear material, in which low-power and high-power pulses are routed in different directions. Our technique may lead to the development of novel compact nonlinear photonic devices.

KEYWORDS: *inverse design, nanophotonics, nonlinear optics, optical switching, integrated photonics*
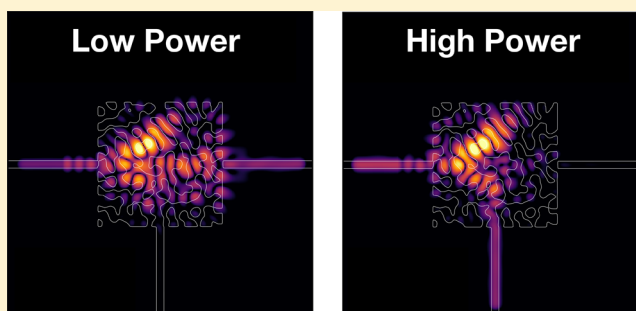
In recent years, there has been significant interest in using computational optimization tools to design novel nano-photonic devices with a wide range of applications.[1−19] Much of this progress[1−12] is made possible by the adjoint method,[18,20] a technique that allows the gradient of an objective function to be computed with respect to an arbitrarily large number of degrees of freedom using only two full-field simulations. Since a very large number of design parameters can then be adjusted simultaneously, the number of structures that need to be evaluated in order to reach a high-performing design can be far smaller compared with the total number of structures in the search space. Therefore, this method makes possible gradient-based design of electromagnetic structures with respect to a large number of free parameters.

Up to now, in photonics the adjoint method has been mostly applied to gradient-based optimization of linear optical devices. The generalization of the adjoint method to nonlinear optical devices would create new possibilities in several exciting fields such as on-chip lasers,[21] frequency combs,[22] spectroscopy,[23] neural computing,[24] and quantum information processing.[25] To this end, several recent works[26−28] have applied adjoint methods to engineer *linear* devices to display favorable properties for nonlinear optical applications, such as high quality factors, small mode volume, or large field overlap between the modes of interest. However, these works do not directly optimize the nonlinear systems.

To solve for the adjoint sensitivity of a nonlinear system, the standard option is to work within a time-domain adjoint formalism, which entails simulating an additional linear system with a time-varying permittivity.[3] However, as this formalism requires the storage of the fields at each time step, it has substantial memory requirements. Furthermore, because in many cases the steady-state behavior of the system is of interest, a frequency-domain approach is preferred as the steady-state response can be obtained directly, without the need for going through a large number of time steps as in a time-domain simulation. The general mathematical formalism for the adjoint method in nonlinear systems is known in the applied mathematics literature.[29] But, with the exception of a very recent preprint that seeks to design a nonlinear element in an optical neural network,[24] such a formalism has not been previously applied to nonlinear photonic device optimizations.

In this work we outline, in detail, how the adjoint method may be used to optimize the steady-state response of a nonlinear optical device in the frequency domain. We first outline a formalism for generalizing adjoint problems to arbitrary nonlinear problems. Then, as a demonstration, we use our method to inverse-design photonic switches with Kerr nonlinearity. Our results may be applied more generally to other objective functions and sources of nonlinearity and provides new possibilities for designing novel nonlinear optical devices.

NONLINEAR ADJOINT METHOD

We first outline the formulation of the adjoint method for the inverse design of nonlinear optical devices. The goal of inverse design is to find a set of real-valued design variables φ that maximize a real-valued objective function $\mathcal{L} = \mathcal{L}(\mathbf{e}, \mathbf{e}^*, \varphi)$, where $\mathcal{L}$ is, most generally, a nonlinear function of its arguments. While for linear systems, $\mathbf{e}$ is a complex-valued vector corresponding to the electric field solution to linear Maxwell's equations, in nonlinear problems, we may write $\mathbf{e}$ as the solution to a nonlinear equation

$$\mathbf{f}(\mathbf{e}, \mathbf{e}^*, \varphi) = 0 \tag{1}$$

For example, eq 1 may represent the steady-state Maxwell's equations with an intensity-dependent permittivity distribution where $\mathbf{e}$ is the electric field distribution. We note that, for the problems we study, the natural choice is to take $\mathbf{e}$ and $\mathbf{e}^*$ as independent parameters, as opposed to treating separately the real and imaginary parts of $\mathbf{e}$.[30,31] The solution to eq 1 may be found with any nonlinear equation solver, such as with the Newton−Raphson method.[32] We further note that the treatment of $\mathbf{e}$ and its complex conjugate as independent variables is necessary for differentiation as will be shown later.

The aim of the optimization is to maximize the objective function with respect to the design variables φ. For this purpose, it is essential to compute the sensitivity of $\mathcal{L}$ with respect to each element of φ. For simplicity, we derive the derivative of the objective function with respect to a single parameter φ, which is written

$$\frac{d\mathcal{L}}{d\varphi} = \frac{\partial\mathcal{L}}{\partial\varphi} + \frac{\partial\mathcal{L}}{\partial\mathbf{e}}\frac{d\mathbf{e}}{d\varphi} + \frac{\partial\mathcal{L}}{\partial\mathbf{e}^*}\frac{d\mathbf{e}^*}{d\varphi} \tag{2}$$

Or, in matrix form as

$$\frac{d\mathcal{L}}{d\varphi} = \frac{\partial\mathcal{L}}{\partial\varphi} + [\partial\mathcal{L}/\partial\mathbf{e} \quad \partial\mathcal{L}/\partial\mathbf{e}^*]\begin{bmatrix} d\mathbf{e}/d\varphi \\ d\mathbf{e}^*/d\varphi \end{bmatrix} \tag{3}$$

To compute $d\mathbf{e}/d\varphi$ and $d\mathbf{e}^*/d\varphi$, we differentiate eq 1:

$$\frac{d\mathbf{f}}{d\varphi} = 0 = \frac{\partial\mathbf{f}}{\partial\varphi} + \frac{\partial\mathbf{f}}{\partial\mathbf{e}}\frac{d\mathbf{e}}{d\varphi} + \frac{\partial\mathbf{f}}{\partial\mathbf{e}^*}\frac{d\mathbf{e}^*}{d\varphi} \tag{4}$$

Equation 4 together with its complex conjugate then yields

$$\begin{bmatrix} \partial\mathbf{f}/\partial\mathbf{e} & \partial\mathbf{f}/\partial\mathbf{e}^* \\ \partial\mathbf{f}^*/\partial\mathbf{e} & \partial\mathbf{f}^*/\partial\mathbf{e}^* \end{bmatrix}\begin{bmatrix} d\mathbf{e}/d\varphi \\ d\mathbf{e}^*/d\varphi \end{bmatrix} = -\begin{bmatrix} \partial\mathbf{f}/\partial\varphi \\ \partial\mathbf{f}^*/\partial\varphi \end{bmatrix} \tag{5}$$

Thus, formally we can rewrite eq 3 as

$$\frac{d\mathcal{L}}{d\varphi} = \frac{\partial\mathcal{L}}{\partial\varphi} - [\partial\mathcal{L}/\partial\mathbf{e} \quad \partial\mathcal{L}/\partial\mathbf{e}^*]\begin{bmatrix} \partial\mathbf{f}/\partial\mathbf{e} & \partial\mathbf{f}/\partial\mathbf{e}^* \\ \partial\mathbf{f}^*/\partial\mathbf{e} & \partial\mathbf{f}^*/\partial\mathbf{e}^* \end{bmatrix}^{-1}\begin{bmatrix} \partial\mathbf{f}/\partial\varphi \\ \partial\mathbf{f}^*/\partial\varphi \end{bmatrix}$$

In analogy with the linear adjoint method, we can now compute the gradient by solving an additional linear system. We define a complex-valued adjoint field $\mathbf{e}_{adj}$ as the solution to

$$(\partial\mathbf{f}/\partial\mathbf{e})^T\mathbf{e}_{adj} + (\partial\mathbf{f}^*/\partial\mathbf{e})^T\mathbf{e}_{adj}^* = -(\partial\mathcal{L}/\partial\mathbf{e})^T \tag{6}$$

We note that, in practice, eq 6 can be computed by solving the system

$$\begin{bmatrix} \partial\mathbf{f}/\partial\mathbf{e} & \partial\mathbf{f}/\partial\mathbf{e}^* \\ \partial\mathbf{f}^*/\partial\mathbf{e} & \partial\mathbf{f}^*/\partial\mathbf{e}^* \end{bmatrix}^T\begin{bmatrix} \mathbf{e}_{adj} \\ \mathbf{e}_{adj}^* \end{bmatrix} = -\begin{bmatrix} \partial\mathcal{L}/\partial\mathbf{e}^T \\ \partial\mathcal{L}/\partial\mathbf{e}^{*T} \end{bmatrix} \tag{7}$$

We note that the adjoint problem, as required to determine the derivative of the objective function, is a *linear* problem, even though the physical problem, as defined by eq 1, is nonlinear.

Finally, the gradient of the objective function is then

$$\frac{d\mathcal{L}}{d\varphi} = \frac{\partial\mathcal{L}}{\partial\varphi} + 2\mathcal{R}\left(\mathbf{e}_{adj}^T\frac{\partial\mathbf{f}}{\partial\varphi}\right) \tag{8}$$

where $\mathcal{R}$ denotes taking the real part. In deriving eq 8, we have used the fact that both $\mathcal{L}$ and φ are real. In the case of multiple parameters φ, we can simply replace $\partial\mathbf{f}/\partial\varphi$ with the matrix $\partial\mathbf{f}/\partial\varphi$. Since $\mathbf{e}_{adj}$ only needs to be solved once regardless of the number of parameters, gradients may be computed with very little marginal cost for an arbitrary number of free parameters, making large-scale, gradient-based optimization possible.

APPLICATION TO KERR NONLINEARITY

We now apply the general formalism as discussed above to the optimization of nonlinear optical systems. Since the formalism is applicable to linear optical systems as well, for illustration purposes here we use it to treat both the linear and the nonlinear cases, in order to highlight aspects that are unique to nonlinear systems. A schematic outlining the two cases is presented in Figure 1. For a linear system, Maxwell's equations for the steady state at a frequency ω_0 may be written as

$$\mu_0^{-1}\nabla \times \nabla \times \mathbf{E}(\mathbf{r}) - \omega_0^2\epsilon_0\epsilon_r(\mathbf{r})\mathbf{E}(\mathbf{r}) = -i\omega_0\mathbf{J}(\mathbf{r}) \tag{9}$$

where $\mathbf{E}(\mathbf{r})$ is the electric field, $\mathbf{J}(\mathbf{r})$ is the electric current source, $\epsilon_r(\mathbf{r})$ is the relative dielectric permittivity, and we have assumed relative permeability $\mu_r = 1$ everywhere. Compactly, and to make connection to the general formalism in the previous section, this can be written in matrix form as

$$\mathbf{f}(\mathbf{e}, \mathbf{e}^*, \varphi) = A(\epsilon_r)\mathbf{e} - \mathbf{b} = 0 \tag{10}$$

where A is a linear operator, vectors $\mathbf{e}$ and ϵ_r now contain the electric fields and the relative permittivity, respectively, and $\mathbf{b}$ is a vector proportional to the current source. The design parameter φ in this case is the permittivity distribution ϵ_r. Equation 10 can be solved to obtain the electric fields $\mathbf{e}$, as diagrammed by Figure 1(a).

We assume an objective function $\mathcal{L}$ that depends on the field solution to eq 10, and we take the linear relative permittivity distribution as the set of design variables. Because $\partial\mathbf{f}/\partial\mathbf{e} = A$ and $\partial\mathbf{f}/\partial\mathbf{e}^* = 0$ for the linear system, from eq 6, the adjoint field may be written simply as the solution to the equation

$$A^T(\epsilon_r)\mathbf{e}_{adj} = -(\partial\mathcal{L}/\partial\mathbf{e})^T \tag{11}$$

as shown in Figure 1(b). For a reciprocal system, $A^T = A$; thus the original and the adjoint fields are solutions to the same linear problem but with different source terms. Note that the source for the adjoint field, $-(\partial\mathcal{L}/\partial\mathbf{e})^T$ depends on both the objective function and the original solution.

Once the adjoint field is computed, the gradient of the objective function with respect to the permittivity distribution is given, through eq 8, by

$$\frac{d\mathcal{L}}{d\epsilon_r} = \frac{\partial\mathcal{L}}{\partial\epsilon_r} + 2\mathcal{R}\left(\mathbf{e}_{adj}^T\frac{\partial A}{\partial\epsilon_r}\mathbf{e}\right) \tag{12}$$

$$= \frac{\partial\mathcal{L}}{\partial\epsilon_r} - 2\omega_0^2\epsilon_0\mathcal{R}(\mathbf{e}_{adj}^T\mathbf{e}) \tag{13}$$

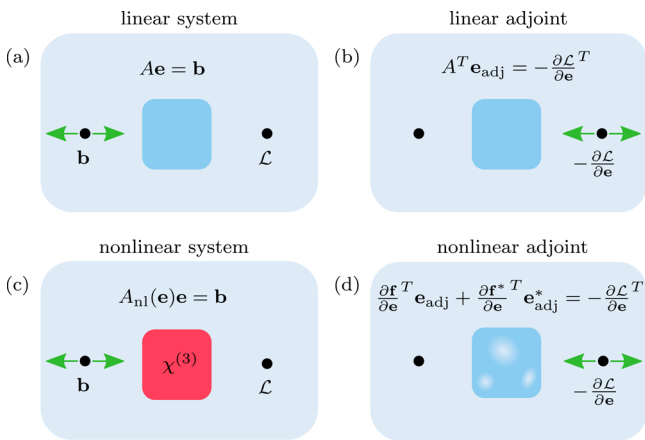**Figure 1.** Illustration of the adjoint field computation for a linear and a nonlinear system. (a) Linear system driven by a point source **b** with an objective function $\mathcal{L}$ given by the field intensity at a measuring point. (b) Adjoint problem for the linear system: the same system driven by a point source given by $-\partial \mathcal{L}/\partial \mathbf{e}$ located at the measuring point. (c) Nonlinear system containing a medium with Kerr nonlinearity (red). The electric fields are the solution to a nonlinear equation. (d) Adjoint problem for the nonlinear system, which is a *linear* system of equations for the adjoint field and its complex conjugate. The Kerr medium is replaced by a linear region whose permittivity depends on the nonlinear fields.

Having reviewed the adjoint formalism for linear optical systems, we now consider nonlinear optical systems. As an example, we introduce Kerr nonlinearity into the system,[33] which corresponds to an intensity-dependent permittivity:

$$\widetilde{\epsilon}_r(\mathbf{r}) = \epsilon_r(\mathbf{r}) + 3\omega_0^2 \epsilon_0 \chi^{(3)}(\mathbf{r})|\mathbf{E}(\mathbf{r})|^2 \tag{14}$$

where $\chi^{(3)}(\mathbf{r})$ is the nonlinear susceptibility distribution. Other types of nonlinear terms can also be treated with the formalism outlined above. Replacing $\epsilon_r(\mathbf{r})$ in eq 9 with $\widetilde{\epsilon}_r(\mathbf{r})$ in eq 14, our system is then described by the equation

$$\mathbf{f}(\mathbf{e}, \mathbf{e}^*, \boldsymbol{\varphi}) = A_{nl}(\epsilon_r, \chi, \mathbf{e})\mathbf{e} - \mathbf{b} = 0 \tag{15}$$

where $A_{nl} \equiv [A(\epsilon_r) - \mathrm{diag}(\chi \odot |\mathbf{e}|^2)]$. Here, $\odot$ is element-wise vector multiplication and $\mathrm{diag}(\mathbf{v})$ represents a diagonal matrix with vector **v** on the main diagonal. The vector χ corresponds to the term $3\omega_0^2 \epsilon_0 \chi^{(3)}(\mathbf{r})$ and $|\mathbf{e}|^2 \equiv \mathbf{e} \odot \mathbf{e}^*$. Again, for concreteness, the design parameters $\boldsymbol{\varphi}$ correspond to the permittivity ϵ_r. The solution to this problem is diagrammed in Figure 1(c).

From eq 15 we may now compute the partial derivatives of **f** with respect to the electric fields **e**, which is needed to construct the adjoint problem.

$$\partial \mathbf{f}/\partial \mathbf{e} = A - 2\mathrm{diag}(\chi \odot |\mathbf{e}|^2) \tag{16}$$

$$\partial \mathbf{f}^*/\partial \mathbf{e} = -\mathrm{diag}(\chi \odot \mathbf{e}^* \odot \mathbf{e}^*) \tag{17}$$

With this, we then express the adjoint field as the solution to eq 6, as diagrammed in Figure 1(d). A technical remark is due: in the linear electromagnetic system, solving for the adjoint field with N grid points requires solving one complex-valued linear problem of size $(N \times N)$, equivalent to two real-valued problems of size $(N \times N)$. For the nonlinear electromagnetic system, because of the coupling between $\mathbf{e}_{adj}$ and $\mathbf{e}_{adj}^*$ and the particular structure of eq 7, the adjoint problem can be written as a real-valued linear system of size $(2N \times 2N)$. This has the same complexity scaling with N as the linear case. We also note that the discretized formulation of the nonlinear adjoint method given here is directly relevant to the numerical results presented below, but in the Supporting Information we also sketch a continuous-variable analogue of the formulation.

Once the adjoint field is computed, the gradient $d\mathcal{L}/d\epsilon_r$ is evaluated from eq 13 as in the linear case. Here for simplicity we do not assume any explicit dependence of the nonlinearity on the design variable. However, the formalism is straightforward to extend to that case, as explained in the Supporting Information.
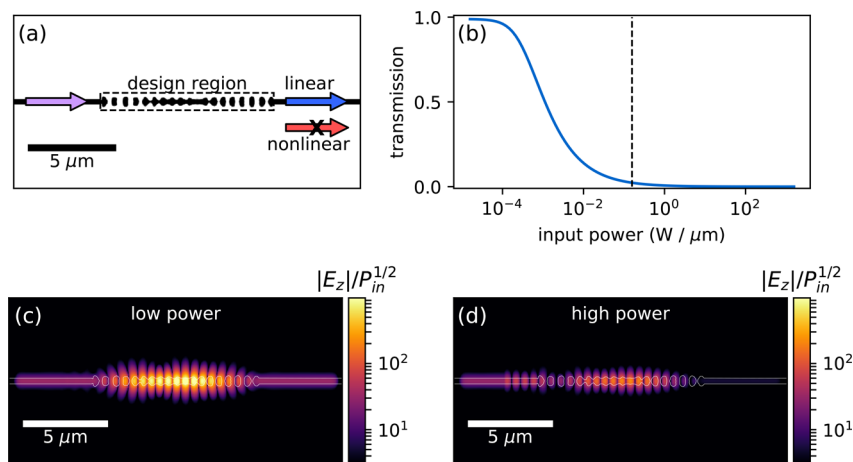

Figure 2. Inverse design demonstration of a $1 \rightarrow 1$ port switch. (a) Optical power is input into the left port (purple arrow). The goal of optimizing the design region (blue square) is to maximize power transmission in the linear regime (blue arrow) and minimize transmission in the nonlinear regime (red arrow). The final permittivity distribution after optimizing is also shown. The black regions are chalcogenide with a relative permittivity of 5.95 and a $\chi^{(3)}$ of 4.1×10^{-19} m²/V². The waveguide regions outside the design region have a width of 0.3 μm. The operating wavelength is 2 μm. (b) Transmission as a function of input power, demonstrating the switching behavior at around 10^{-3} W/μm. The dashed black line indicates the input power used for the high-power regime in the optimization. (c, d) Amplitude of the simulated electric field of the final structure, in the linear (c) and nonlinear (d) regimes, respectively, with an input power of 10^{-9} and 0.157 W/μm, respectively. E_z corresponds to the out-of-plane electric field in the 2D simulation.
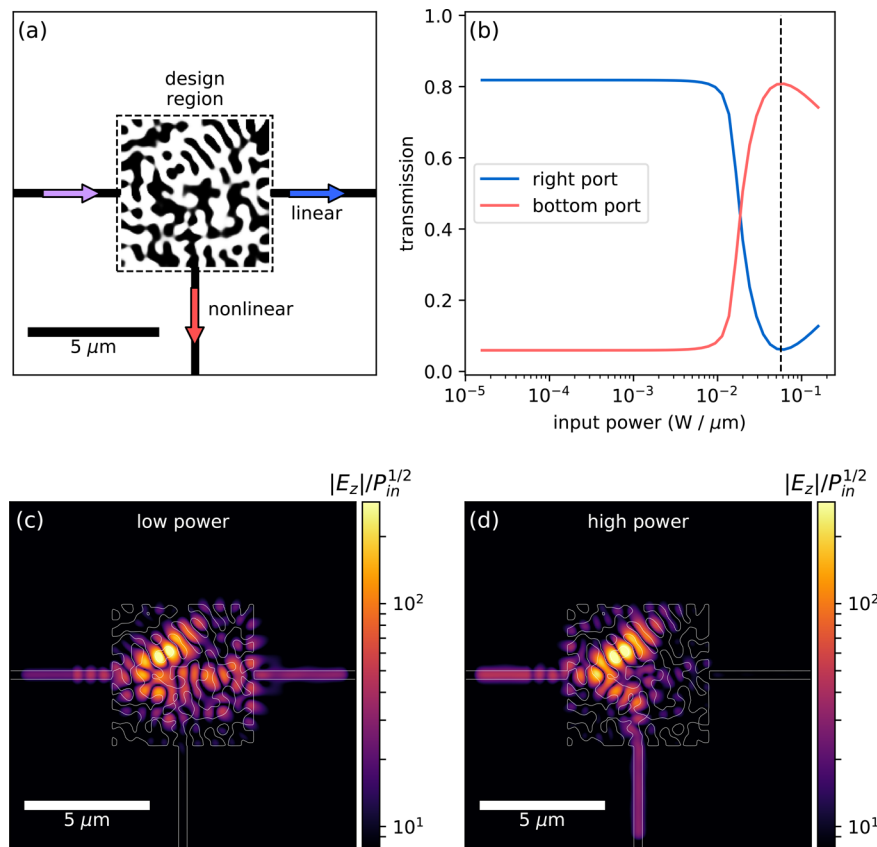
Figure 3. Inverse design demonstration of a $1 \rightarrow 2$ port switch. (a) Optical power is input into the left port (purple arrow). The goal of optimizing the design region (blue square) is to maximize the power transmission to the right port (blue arrow) in the linear regime and maximize transmission to the bottom port (red arrow) in the nonlinear regime. The final permittivity distribution after optimizing is also shown. (b) Transmission in the right (blue) and bottom (red) ports as a function of input power, demonstrating the switching behavior at around 2×10^{-2} W/ μm. The dashed black line indicates the input power used for the high-power regime in the optimization. (c, d) Amplitude of the simulated electric field of the final structure, in the linear (c) and nonlinear (d) regimes, respectively, with an input power of 10^{-9} and 0.057 W/μm, respectively.

■ INVERSE DESIGN OF OPTICAL SWITCHES

We now demonstrate the use of this nonlinear adjoint formalism to inverse design optical switches with desired power-dependent performance characteristics. In Figures 2 and 3, we show the optimization procedures and performance characteristics of a $1 \rightarrow 1$ and $1 \rightarrow 2$ port device, respectively. The operating frequency for both devices corresponds to a free-space wavelength of 2 μm.

For each device, we seek to maximize the corresponding objective function with respect to the permittivity distribution within a fixed design region. To perform the numerical optimization of the structure, we use the finite-difference frequency-domain method (FDFD),[34] where the fields and operators of eq 9 are spatially discretized using a Yee lattice.[35] For simplicity, we restrict our study to two-dimensional structures (i.e., structures with infinite extent in the third dimension) and transverse-magnetic polarization, which has only nonzero out-of-plane electric field components. In the optimization process, we start with an initial relative permittivity in the design region. We solve the electric field distribution in the structure by solving the nonlinear eq 15. Then, we compute the gradient of $\mathcal{L}$ with respect to the relative permittivity at every point in the design region using eq 13. With the gradient information, we perform updates of the design variables using the limited-memory BFGS[36] algorithm, although a simple gradient ascent algorithm would also suffice.

This procedure is repeated until convergence on a final structure.

We choose optimization parameters corresponding to a device made from chalcogenide glass (Al_2S_3), which exhibits a strong $\chi^{(3)}$ response and high damage threshold.[33,37,38] During the optimization, the relative permittivity was constrained to lie between 1 (air) and 5.95 (Al_2S_3). We further assume that the materials exhibit nonlinearity only within the design regions outlined in Figures 2 and 3.

To create a more realistic final structure, the strength of the nonlinear susceptibility was assumed to be proportional to the "density" of the material, ρ, defined as

$$\rho(\mathbf{r}) = \frac{\epsilon_r(\mathbf{r}) - 1}{\epsilon_m - 1} \tag{18}$$

where ϵ_m is the permittivity of the material. This assumption ensures that air regions do not exhibit a nonlinear refractive index. Equation 18 adds an ϵ_r dependence in the nonlinear susceptibility, which is straightforwardly treated in the adjoint method, as discussed in the Supplementary Information. Low-pass spatial filtering and projection techniques[39] were applied during optimization to create binarized (air and chalcogenide) final structures with large, smoothed features. Additional details on this are described in the Supporting Information.

Our first device, as shown in Figure 2, consists of a waveguide-fed $1 \rightarrow 1$ port geometry with a central design

region. We optimize this design region to maximize power transmission in the linear regime when the incident power is sufficiently low such that the nonlinear terms do not affect the transmission, and minimize transmission in the nonlinear regime when the incident power is at a specific high value such that the nonlinearity plays a significant role. This corresponds to an objective function of the form

$$\mathcal{L}(\mathbf{e}_{\text{low}}, \mathbf{e}_{\text{high}}) = |\mathbf{m}^T \mathbf{e}_{\text{low}}|^2 - |\mathbf{m}^T \mathbf{e}_{\text{high}}|^2 \tag{19}$$

where $\mathbf{e}_{\text{low}}$ and $\mathbf{e}_{\text{high}}$ are the simulated fields with a low and a high input power, respectively, $\mathbf{m}$ is the modal profile of the electric field for the waveguide in the output port, and the objective function is normalized such that its maximum value is 1. The optimization setup and the optimized structure are diagrammed in Figure 2(a). The final structure resembles a resonator between two Bragg mirrors, effectively acting like a bistable switch.[40,41] Figure 2(b) shows the transmission as a function of the input power, and it clearly switches from high to low as the input power increases. This is also illustrated in panels (c) and (d), where we plot the field amplitude distributions in the low-power (high-transmission) regime and in the high-power (low-transmission) regime, respectively. The computed power transmission coefficients for these two panels are 98.2% and 3.1%, respectively. The value of the input power used in the optimization and in panel (d) is shown by a dashed line in panel (b). At this input power, the device exhibits a maximum nonlinear refractive index shift of 4.0×10^{-3}, which is below the damage threshold for Al_2S_3 using sub-nanosecond pulses[42] (see Supporting Information). The transmission spectrum of this structure gives a resonance peak with a full-width at half-maximum of 38 GHz (see Supporting Information). In the Supporting Information, we also list the specific optimization parameters and show the value of the objective function during the optimization process. Reaching the final optimized structure shown in Figure 2 required the evaluation of 2000 structures, but a reasonably high-performing structure is already reached after only a few hundred iterations. This relatively small number of iterations should be contrasted with the vast size of the parameter space spanned by the 12 250 points contained in the design region.

We also use the same technique for the inverse design of a 1 → 2 port switch where light is guided to the right port in the linear regime and to the bottom port in the nonlinear regime. To achieve this design, we define the objective function as

$$\mathcal{L}(\mathbf{e}_{\text{low}}, \mathbf{e}_{\text{high}}) = |\mathbf{m}_{\text{r}}^T \mathbf{e}_{\text{low}}|^2 - |\mathbf{m}_{\text{r}}^T \mathbf{e}_{\text{high}}|^2$$
$$- |\mathbf{m}_{\text{b}}^T \mathbf{e}_{\text{low}}|^2 + |\mathbf{m}_{\text{b}}^T \mathbf{e}_{\text{high}}|^2 \tag{20}$$

where $\mathbf{m}_{\text{r}}$ and $\mathbf{m}_{\text{b}}$ denote the mode profiles of the waveguides in the right and bottom ports, respectively. These are normalized such that the objective function has a maximum value of 1 for a perfect switching operation. The setup of the optimization problem and the final design are diagrammed in Figure 3(a). We note that the device displays a nonintuitive geometry while retaining large features and good binarization.

In Figure 3(b), we plot the transmission through the right and through the bottom ports as a function of input power. This clearly shows the switching of power from the right port to the bottom port in the linear and nonlinear regimes, respectively. Specifically, in the linear regime, the device has a power transmission of 81.8% and 5.9% to the right and bottom ports, respectively, while in the nonlinear regime, at the input

power marked by the dashed line in Figure 3(b), these values are 6.1% and 80.8%, respectively. The electric field amplitudes for linear and nonlinear regimes are displayed in Figure 3(c,d). The operational bandwidth for this device is approximately 90 GHz (see Supporting Information). The device exhibits a maximum nonlinear refractive index shift of 3.9×10^{-3}, which is also below the acceptable damage threshold for Al_2S_3 using sub-nanosecond pulses.[42] In this optimization, the parameter space is spanned by the permittivity of 22 500 points in the design region. A full list of optimization parameters and a plot of the objective function versus iteration number is shown in the Supporting Information.

■ DISCUSSION

We have presented an extension to the adjoint variable method applied to the optimization of an electromagnetic system with Kerr nonlinearity. Our approach can be straightforwardly applied to other types of nonlinearities that do not mix frequencies, such as saturable gain or absorption. Moreover, the methods here should be straightforwardly generalizable to treat nonlinear problems involving frequency mixing. For example, one can imagine implementing a similar adjoint method in combination with the multifrequency finite-difference frequency-domain implementations for nonlinear wave interactions.[43]

In addition to the design of optical switches, our formalism may prove useful for many other interesting problems in nonlinear photonics. For example, one could apply our approach to design nonlinear elements in optical neural networks[44] with specific forms of activation functions. Another interesting application is power regulation in photonic networks. For example, as photonic networks for laser-driven particle accelerators[45] must be able to handle large input powers, it may be of interest to use our approach to design compact optical limiters in these networks. For the purposes of exploring these and many other potential applications, we have made publicly available a software package that implements the algorithms discussed here.[46]

To summarize this paper, we have developed an adjoint method, which enables gradient optimization of nonlinear photonic devices. Our work broadens the capability of inverse design for producing novel nonlinear devices.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsphotonics.8b01522.

> Additional details including derivation of the method for continuous variables; detailed parameters used in the optimizations; details on implementing permittivity-dependent susceptibility as well as feature size and binarization constraints; nonlinear index shift estimation; and transmission spectra of the optimized designs (PDF)

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: shanhui@stanford.edu.

ORCID ⓘ
Tyler W. Hughes: 0000-0001-7989-0891
Momchil Minkov: 0000-0003-0665-8412

Ian A. D. Williamson: 0000-0002-6699-1973

Author Contributions

‡T. W. Hughes and M. Minkov contributed equally to this work.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Lalau-Keraly, C. M.; Bhargava, S.; Miller, O. D.; Yablonovitch, E. Adjoint shape optimization applied to electromagnetic design. *Opt. Express* **2013**, *21*, 21693−21701.

(2) Wang, J.; Shi, Y.; Hughes, T.; Zhao, Z.; Fan, S. Adjoint-based optimization of active nanophotonic devices. *Opt. Express* **2018**, *26*, 3236−3248.

(3) Elesin, Y.; Lazarov, B. S.; Jensen, J. S.; Sigmund, O. Design of robust and efficient photonic switches using topology optimization. *Photonics and Nanostructures - Fundamentals and Applications* **2012**, *10*, 153−165.

(4) Piggott, A. Y.; Petykiewicz, J.; Su, L.; Vučković, J. Fabrication-constrained nanophotonic inverse design. *Sci. Rep.* **2017**, *7*, 1786.

(5) Piggott, A. Y.; Lu, J.; Lagoudakis, K. G.; Petykiewicz, J.; Babinec, T. M.; Vučković, J. Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer. *Nat. Photonics* **2015**, *9*, 374−377.

(6) Kao, C. Y.; Osher, S.; Yablonovitch, E. Maximizing band gaps in two-dimensional photonic crystals by using level set methods. *Appl. Phys. B: Lasers Opt.* **2005**, *81*, 235−244.

(7) Hughes, T.; Veronis, G.; Wootton, K. P.; England, R. J.; Fan, S. Method for computationally efficient design of dielectric laser accelerator structures. *Opt. Express* **2017**, *25*, 15414−15427.

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

(9) Sigmund, O.; Sondergaard Jensen, J. Systematic design of phononic band-gap materials and structures by topology optimization. *Philos. Trans. R. Soc., A* **2003**, *361*, 1001−1019.

(10) Matzen, R.; Jensen, J. S.; Sigmund, O. Systematic design of slow-light photonic waveguides. *J. Opt. Soc. Am. B* **2011**, *28*, 2374−2382.

(11) Jensen, J. S.; Sigmund, O. Topology optimization of photonic crystal structures: a high-bandwidth low-loss T-junction waveguide. *J. Opt. Soc. Am. B* **2005**, *22*, 1191.

(12) Frellsen, L. F.; Ding, Y.; Sigmund, O.; Frandsen, L. H. Topology optimized mode multiplexing in silicon-on-insulator photonic wire waveguides. *Opt. Express* **2016**, *24*, 16866−16873.

(13) Shen, B.; Wang, P.; Polson, R.; Menon, R. An integrated-nanophotonics polarization beamsplitter with 2.4 × 2.4 μm^2 footprint. *Nat. Photonics* **2015**, *9*, 378−382.

(14) Shen, L.; Ye, Z.; He, S. Design of two-dimensional photonic crystals with large absolute band gaps using a genetic algorithm. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2003**, *68*, 035109.

(15) Minkov, M.; Savona, V. Automated optimization of photonic crystal slab cavities. *Sci. Rep.* **2015**, *4*, 5124.

(16) Minkov, M.; Savona, V. Wide-band slow light in compact photonic crystal coupled-cavity waveguides. *Optica* **2015**, *2*, 631−634.

(17) Shi, Y.; Li, W.; Raman, A.; Fan, S. Optimization of Multilayer Optical Films with a Memetic Algorithm and Mixed Integer Programming. *ACS Photonics* **2018**, *5*, 684−691.

(18) Veronis, G.; Dutton, R. W.; Fan, S. Method for sensitivity analysis of photonic crystal devices. *Opt. Lett.* **2004**, *29*, 2288−2290.

(19) Peurifoy, J.; Shen, Y.; Jing, L.; Yang, Y.; Cano-Renteria, F.; DeLacy, B. G.; Joannopoulos, J. D.; Tegmark, M.; Soljačić, M. Nanophotonic particle simulation and inverse design using artificial neural networks. *Science Advances* **2018**, *4*, No. eaar4206.

(20) Giles, M. B.; Pierce, N. A. An introduction to the adjoint approach to design. *Flow, Turbul. Combust.* **2000**, *65*, 393−415.

(21) Yamashita, D.; Takahashi, Y.; Asano, T.; Noda, S. Raman shift and strain effect in high-Q photonic crystal silicon nanocavity. *Opt. Express* **2015**, *23*, 3951.

(22) Okawachi, Y.; Saha, K.; Levy, J. S.; Wen, Y. H.; Lipson, M.; Gaeta, A. L. Octave-spanning frequency comb generation in a silicon nitride chip. *Opt. Lett.* **2011**, *36*, 3398.

(23) Moon, J. H.; Kim, J. H.; Kim, K.-j.; Kang, T.-H.; Kim, B.; Kim, C.-H.; Hahn, J. H.; Park, J. W. Absolute Surface Density of the Amine Group of the Aminosilylated Thin Layers: Ultraviolet-Visible Spectroscopy, Second Harmonic Generation, and Synchrotron-Radiation Photoelectron Spectroscopy Study. *Langmuir* **1997**, *13*, 4305−4310.

(24) Khoram, E.; Chen, A.; Liu, D.; Wang, Q.; Yu, Z. Stochastic Optimization of Nonlinear Nanophotonic Media for Artificial Neural Inference. arXiv preprint *arXiv:1810.07815*, **2018**.

(25) Guo, X.; Zou, C.-L.; Jung, H.; Tang, H. X. On-Chip Strong Coupling and Efficient Frequency Conversion between Telecom and Visible Optical Modes. *Phys. Rev. Lett.* **2016**, *117*, 123902.

(26) Lin, Z.; Liang, X.; Lončar, M.; Johnson, S. G.; Rodriguez, A. W. Cavity-enhanced second-harmonic generation via nonlinear-overlap optimization. *Optica* **2016**, *3*, 233.

(27) Lin, Z.; Lončar, M.; Rodriguez, A. W. Topology optimization of multi-track ring resonators and 2D microcavities for nonlinear frequency conversion. *Opt. Lett.* **2017**, *42*, 2818.

(28) Bravo-Abad, J.; Rodriguez, A.; Bermel, P.; Johnson, S. G.; Joannopoulos, J. D.; Soljačić, M. Enhanced nonlinear optics in photonic-crystal microcavities. *Opt. Express* **2007**, *15*, 16161−16176.

(29) Strang, G. *Computational Science and Engineering*; Wellesley-Cambridge Press, 2007; Chapter 8.

(30) Kreutz-Delgado, K. Complex gradient operator and the CR-calculus. arXiv preprint *arXiv:0906.4835*, **2009**.

(31) Remmert, R. *Theory of Complex Functions*; Springer Science & Business Media, 2012.

(32) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes 3rd ed.: The Art of Scientific Computing*; Cambridge University Press, 2007.

(33) Boyd, R. W. *Nonlinear Optics*; Academic Press, 2008.

(34) 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.

(35) Yee, K. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Trans. Antennas Propag.* **1966**, *14*, 302−307.

(36) Byrd, R. H.; Lu, P.; Nocedal, J.; Zhu, C. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* **1995**, *16*, 1190−1208.

(37) White, R. T.; Monro, T. M. Cascaded Raman Shifting of High-Peak-Power Nanosecond Pulses in As_2S_3 and As_2Se_3 Optical Fibers. *Opt. Lett.* **2011**, *36*, 2351.

(38) Lamont, M. R.; Luther-Davies, B.; Choi, D.-Y.; Madden, S.; Eggleton, B. J. Supercontinuum Generation in Dispersion Engineered Highly Nonlinear (γ = 10 W/m) As_2S_3 Chalcogenide Planar Waveguide. *Opt. Express* **2008**, *16*, 14938.

(39) Zhou, M.; Lazarov, B. S.; Wang, F.; Sigmund, O. Minimum length scale in topology optimization by geometric constraints. *Computer Methods in Applied Mechanics and Engineering* **2015**, *293*, 266−282.

(40) Soljačić, M.; Ibanescu, M.; Johnson, S. G.; Fink, Y.; Joannopoulos, J. D. Optimal bistable switching in nonlinear photonic crystals. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2002**, *66*, 055601.

(41) Yanik, M. F.; Fan, S.; Soljačić, M.; Joannopoulos, J. D. All-optical transistor action with bistable switching in a photonic crystal cross-waveguide geometry. *Opt. Lett.* **2003**, *28*, 2506.

(42) Chorel, M.; Lanternier, T.; Lavastre, E.; Bonod, N.; Bousquet, B.; Néauport, J. Robust optimization of the laser induced damage threshold of dielectric mirrors for high power lasers. *Opt. Express* **2018**, *26*, 11764−11774.

(43) Shi, Y.; Shin, W.; Fan, S. Multi-frequency finite-difference frequency-domain algorithm for active nanophotonic device simulations. *Optica* **2016**, *3*, 1256−1259.

(44) Shen, Y.; Harris, N. C.; Skirlo, S.; Prabhu, M.; Baehr-Jones, T.; Hochberg, M.; Sun, X.; Zhao, S.; Larochelle, H.; Englund, D.; Soljačić, M. Deep learning with coherent nanophotonic circuits. *Nat. Photonics* **2017**, *11*, 441.

(45) Hughes, T. W.; et al. On-Chip Laser-Power Delivery System for Dielectric Laser Accelerators. *Phys. Rev. Appl.* **2018**, *9*, 054017.

(46) Hughes, T. W.; Minkov, M.; Williamson, I. A. D. *Angler* − an Adjoint Nonliner Gradient open-source package. https://github.com/fancompute/angler, 2018.