Supplementary Information: 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, CA 94305, USA

E-mail: shanhui@stanford.edu

Continuous-variable formulation

In the main text, we have derived the nonlinear adjoint method directly for the spatially-discretized version of Maxwell's equations, as this relates straightforwardly to the numerical implementation. The same results can also be written starting from a continuous-variable definition of the adjoint, as presented e.g. in.[1] Specifically, given an objective functional $\mathcal{L}\left(\boldsymbol{\psi}(\mathbf{r}), \varphi\right)$, where $\boldsymbol{\psi}(\mathbf{r})$ now generally contains all vector components of $\mathbf{e}$ and $\mathbf{e}^*$, and a functional defining the (nonlinear) electromagnetic problem $f(\mathbf{r}, \boldsymbol{\psi}(\mathbf{r}), \varphi)$, the variation in $\mathcal{L}$ with φ is given, in terms of functional derivatives, as

$$\delta_\varphi \mathcal{L} = \frac{\delta \mathcal{L}}{\delta \varphi} + \frac{\delta \mathcal{L}}{\delta \boldsymbol{\psi}} \frac{\delta \boldsymbol{\psi}}{\delta \varphi}. \tag{S1}$$

Furthermore,

$$-\frac{\delta f}{\delta \varphi} = \frac{\delta f}{\delta \boldsymbol{\psi}} \frac{\delta \boldsymbol{\psi}}{\delta \varphi}, \tag{S2}$$

and so, formally,

$$\delta_\varphi \mathcal{L} = \frac{\delta \mathcal{L}}{\delta \varphi} - \frac{\delta \mathcal{L}}{\delta \boldsymbol{\psi}} \left(\frac{\delta f}{\delta \boldsymbol{\psi}}\right)^{-1} \frac{\delta f}{\delta \varphi}, \tag{S3}$$

and the adjoint fields can be defined as the solution to

$$\boldsymbol{\psi}_{\text{adj}} \frac{\delta f}{\delta \boldsymbol{\psi}} = -\frac{\delta \mathcal{L}}{\delta \boldsymbol{\psi}}. \tag{S4}$$

For a linear problem, $f(\mathbf{r}, \boldsymbol{\psi}(\mathbf{r}), \varphi) = F(\mathbf{r}, \varphi)\boldsymbol{\psi}(\mathbf{r}) - b(\mathbf{r})$, the functional derivative $\frac{\delta f}{\delta \boldsymbol{\psi}}$ is just $F(\mathbf{r}, \varphi)$, and the solution to eq. (S4) can be written in terms of the Green's function for that problem, defined for a fixed set of parameters φ as $G_{\text{adj}}(\mathbf{r}, \mathbf{r}'; \varphi)F(\mathbf{r}, \varphi) = \delta(\mathbf{r} - \mathbf{r}')$:

$$\boldsymbol{\psi}_{\text{adj}}(\mathbf{r}) = -\int \mathrm{d}\mathbf{r}' G_{\text{adj}}(\mathbf{r}, \mathbf{r}'; \varphi) \frac{\delta \mathcal{L}}{\delta \boldsymbol{\psi}} \tag{S5}$$

If the system is also reciprocal, the Green's function for the adjoint problem is related to the Green's function of the original problem, defined through $F(\mathbf{r}, \varphi)G_{\text{adj}}(\mathbf{r}, \mathbf{r}'; \varphi) = \delta(\mathbf{r} - \mathbf{r}')$, sim-

ply by $G(\mathbf{r}, \mathbf{r}'; \varphi) = G_{\text{adj}}(\mathbf{r}', \mathbf{r}; \varphi)$. This leads to some intuition behind the adjoint formalism, as discussed in.[1,2] For the case of a nonlinear system, the original problem $f(\mathbf{r}, \boldsymbol{\psi}(\mathbf{r}), \varphi) = 0$ and the adjoint problem of eq. (S4) are of a different nature. Specifically, while the forward problem is nonlinear, the important feature of eq. (S4) is that it is a linear problem in ψ_{adj}, and can be easily solved.

Optimization Details

Table S1 contains the parameters used in the inverse design demonstrations of Fig. 2 and Fig. 3 of the main text. The values of the objective function vs. iteration are shown in Fig. S1 for both the 2-port and 3-port devices.

Table S1: Parameters used in the optimization study. Column '2-port' refers to the device from Fig. 2. Column '3-port' refers to the device from Fig. 3

parameter	symbol	value (2-port)	value (3-port)	units
max relative permittivity	ϵ_{m}	5.95	5.95	-
nonlinear susceptibility	$\chi^{(3)}$	4.1×10^{-19}	4.1×10^{-19}	$\text{m}^2\ \text{V}^{-2}$
input power	P_0	157	57	$\text{mW}/\mu\text{m}$
free space wavelength	λ_0	2	2	μm
design region length	L	10	6	μm
design region height	H	1.6	6	μm
waveguide width	w	300	300	nm
grid size	g	40	40	nm
low-pass filter feature size	R	160	200	nm
projection strength	β	100	500	-
projection mid-point	η	0.5	0.5	-

Permittivity-dependent Nonlinear Susceptibility

In the inverse design demonstration of the main text, we made the assumption that the nonlinear susceptibility distribution was proportional to the density of material in the design

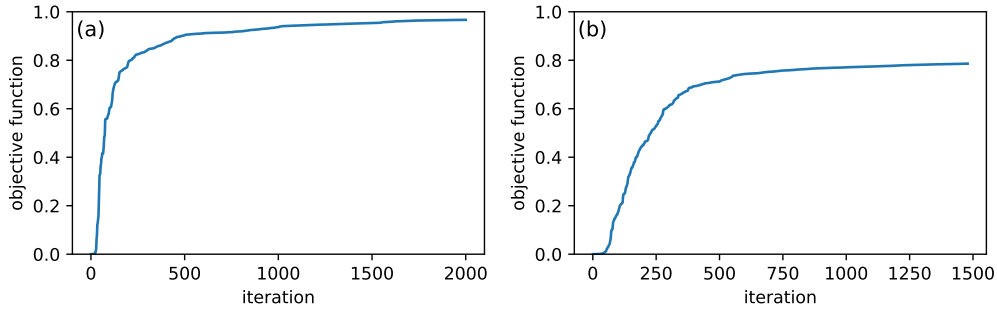

Figure S1: Objective function vs. iteration of the optimization for (a) 2-port device of Fig. 2 and (b) 3-port device of Fig. 3 of the main text.

region. Here we will derive the form of the adjoint sensitivity with this modification.

With our assumption, the nonlinear susceptibility vector $\boldsymbol{\chi}$ may be written in terms of the scalar magnitude of the nonlinear susceptibility $\chi^{(3)}$ and the relative permittivity vector $\boldsymbol{\epsilon}_\mathrm{r}$ explicitly as

$$\boldsymbol{\chi} = 3\omega_0^2\epsilon_0\chi^{(3)}\frac{\boldsymbol{\epsilon}_\mathrm{r}-\mathbf{1}}{\epsilon_\mathrm{m}-1} \tag{S6}$$

where $\mathbf{1}$ is a vector of all ones and ϵ_m is the maximum relative permittivity allowed in the optimization, corresponding to the material relative permittivity.

When choosing to relative permittivity distribution as the set of design variables, $\boldsymbol{\varphi} = \boldsymbol{\epsilon}_\mathrm{r}$, the nonlinear adjoint problem requires the calculation of the partial derivatives $\partial\mathbf{f}/\partial\mathbf{e}$, $\partial\mathbf{f}^*/\partial\mathbf{e}$, and $\partial\mathbf{f}/\partial\boldsymbol{\epsilon}_\mathrm{r}$. When choosing the form of $\boldsymbol{\chi}$ from eq. (S6), the partial derivatives $\partial\mathbf{f}/\partial\mathbf{e}$ and $\partial\mathbf{f}^*/\partial\mathbf{e}$ are the same as derived in the main text. However, the term $\partial\mathbf{f}/\partial\boldsymbol{\epsilon}_\mathrm{r}$ takes on a more complicated form given by

$$\frac{\partial\mathbf{f}}{\partial\boldsymbol{\epsilon}_\mathrm{r}} = \frac{\partial}{\partial\boldsymbol{\epsilon}_\mathrm{r}}\left[A(\boldsymbol{\epsilon}_\mathrm{r})\mathbf{e} - 3\omega_0^2\epsilon_0\chi^{(3)}\frac{\boldsymbol{\epsilon}_\mathrm{r}-\mathbf{1}}{\epsilon_m-1}\odot\mathbf{e}\odot\mathbf{e}\odot\mathbf{e}^* - b\right] \tag{S7}$$

$$= \frac{\partial A}{\partial\boldsymbol{\epsilon}_\mathrm{r}}\mathbf{e} - 3\omega_0^2\epsilon_0\chi^{(3)}\frac{1}{\epsilon_m-1}\mathrm{diag}\left(\mathbf{e}\odot|\mathbf{e}|^2\right) \tag{S8}$$

$$= \mathrm{diag}\left(\mathbf{e}\right) - 3\omega_0^2\epsilon_0\chi^{(3)}\frac{1}{\epsilon_m-1}\mathrm{diag}\left(\mathbf{e}\odot|\mathbf{e}|^2\right), \tag{S9}$$

where we make use of the fact that $\left(\frac{\partial A}{\partial\boldsymbol{\epsilon}_\mathrm{r}}\right)_{ijk} = \delta_{ij}\delta_{jk}$, where δ is the Kronecker delta.

Thus, while the adjoint field $\mathbf{e}_{adj}$ will have the same form as in the main text, when

computing the sensitivity as in eq. (8) in the main text, one must insert the form of $\partial \mathbf{f}/\partial \boldsymbol{\epsilon}_\mathrm{r}$ from eq. (S9). This is in contrast to the usual case where the nonlinear susceptibility is fixed, $\partial \mathbf{f}/\partial \boldsymbol{\epsilon}_\mathrm{r} = \mathrm{diag}\,(\mathbf{e})$.

Maintaining minimum feature size and binarization

To create realistic devices with sufficiently large minimum feature sizes and binarized permittivity distributions, we employed filtering and projection schemes during our optimization. These schemes are discussed in great detail in[3] and related works.

Rather than updating the permittivity distribution directly, one may instead choose to update a design density $\boldsymbol{\rho}$, which varies between 0 and 1 within the design region. To create a structure with larger feature sizes, a low pass filter can be applied to $\boldsymbol{\rho}$ to created a filtered density, labelled $\tilde{\boldsymbol{\rho}}$:

$$\tilde{\rho}_i = \frac{\sum_{j\in\mathcal{D}} W_{ij}\rho_j}{\sum_{j\in\mathcal{D}} W_{ij}}, \tag{S10}$$

where $\mathcal{D}$ denotes the design region, and W is the spatial filter, defined for a feature size of R as

$$W_{ij} = \begin{cases} R - |r_i - r_j| & \text{if } |r_i - r_j| \leq R \\ 0, & \text{otherwise} \end{cases} \tag{S11}$$

with $|r_i - r_j|$ being the distance between points i and j. This defines a low-pass spatial filter on $\boldsymbol{\rho}$ with the effect of smoothing out features with length scale below R.

Now, for binarization of the structure, a projection scheme is used to recreate the final permittivity from the filtered density. For this we define $\bar{\boldsymbol{\rho}}$ as the projected density, which is created from $\tilde{\boldsymbol{\rho}}$ as

$$\bar{\rho}_i = \frac{\tanh(\beta\eta) + \tanh(\beta[\tilde{\rho}_i - \eta])}{\tanh(\beta\eta) + \tanh(\beta[1-\eta])}. \tag{S12}$$

Here, η is a parameter between 0 and 1 that controls the mid-point of the projection, typically 0.5, and β controls the strength of the projection, typically around 100.

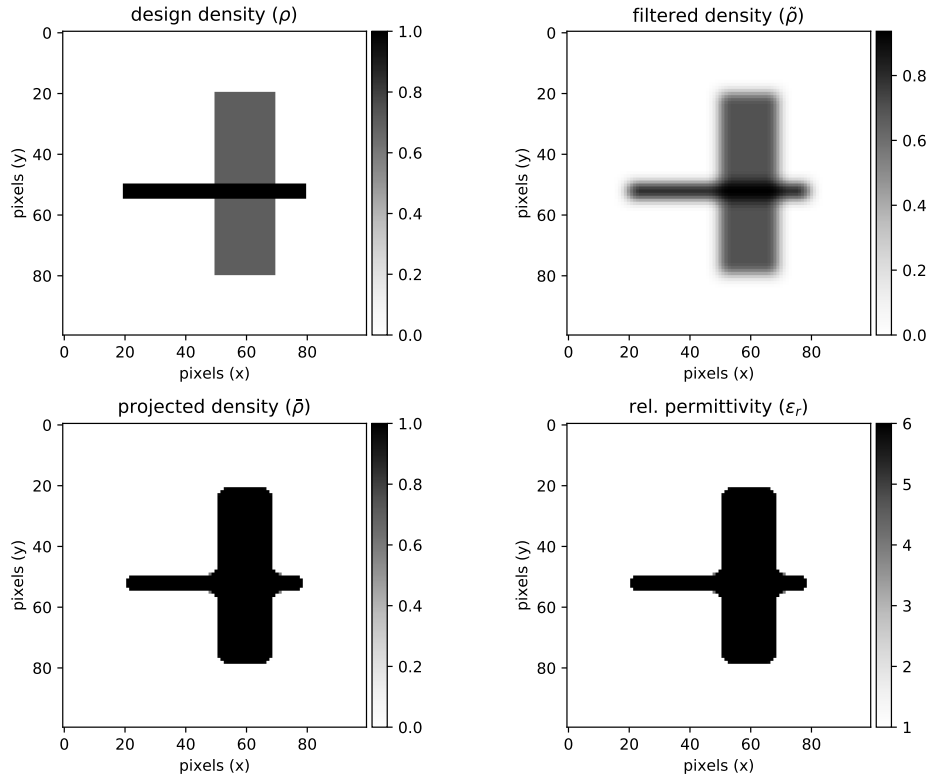

Figure S2: Filtering and projection of an example design density, ρ. (top left) the original density ρ before processing. (top right) the density after applying a low pass filter, $\tilde{\rho}$. (bottom left) the density $\bar{\rho}$ after applying projection. (bottom right) the final relative permittivity distribution ϵ_r. The parameters used are $R = 200$nm, $\beta = 100$, $\eta = 0.5$.

The relative permittivity can then be determined from $\bar{\boldsymbol{\rho}}$ as

$$\boldsymbol{\epsilon}_{\mathrm{r}} = (\epsilon_m - 1)\bar{\boldsymbol{\rho}} + 1, \tag{S13}$$

where ϵ_m is the maximum permittivity.

The effect of these filtering and the binarization techniques on a sample permittivity set is illustrated in Fig. S2. In the optimizations of the main text, these techniques were performed only within the design region and required minimal modifications to the adjoint sensitivity. The determination of $\partial\boldsymbol{\epsilon}_{\mathrm{r}}/\partial\bar{\boldsymbol{\rho}}$, $\partial\bar{\boldsymbol{\rho}}/\partial\tilde{\boldsymbol{\rho}}$, and $\partial\tilde{\boldsymbol{\rho}}/\partial\boldsymbol{\rho}$ were required to compute the derivatives of the objective function with respect to the underlying $\boldsymbol{\rho}$. For more details, see the software package accompanying this work.[4]

Nonlinear Index Shift

Here we estimate the maximum nonlinear index shift of chalcogenide (Al_2S_3) materials. Based on,[5–7] Al_2S_3 has a nonlinear index (n_2) between 3×10^{-14} and 2×10^{-13} cm^2/W. From,[8] the damage threshold is 2.5 J/cm^2. At a pulse duration of 100 ps, this damage threshold corresponds to 2.5×10^{10} W/cm^2. Together with the nonlinear refractive index, the maximum refractive index shift sustainable by the material is approximately

$$\Delta n = n_2 \; I_{damage} \approx 7.5 \times 10^{-4} - 5 \times 10^{-3} \tag{S14}$$

with a corresponding pulse bandwidth (for a bandwidth-limited Gaussian pulse) of $\approx$ 4.5 GHz. Our final structures exhibit maximum refractive index shifts below 5×10^{-3} and their objective functions have FWHM bandwidths above 10 GHz. This suggests that they should exhibit their desired switching effects without damage using pulse durations on the order of 100 ps and input powers on the order of 100 mW/μm.

Transmission Spectra

The transmission vs. frequency for the two devices in the main text is shown in Fig. S3 in the linear regime.

References

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

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

(3) Zhou, M.; Lazarov, B. S.; Wang, F.; Sigmund, O. Minimum length scale in topology

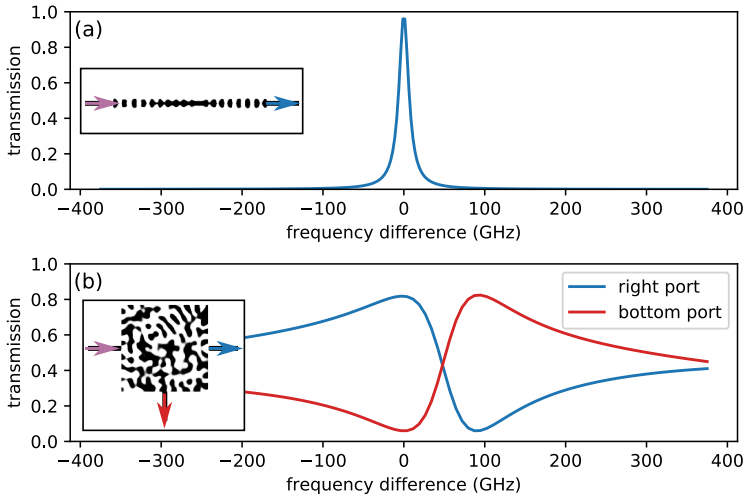

Figure S3: (a) Transmission spectrum through the 2-port device for the low power (linear) regime. (b) Transmission spectrum through the right (blue) and bottom (red) ports of the 3-port device in the low power (linear) regime. The x-axis represents the difference in frequency with respect to the design frequency.

optimization by geometric constraints. *Computer Methods in Applied Mechanics and Engineering* **2015**, *293*, 266–282.

(4) Hughes, T. W.; Minkov, M.; Williamson, I. A. D. fdfdpy. `https://github.com/fancompute/angler`, 2018.

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

(6) 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. *Optics Express* **2008**, *16*, 14938.

(7) 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. *Optics Letters* **2011**, *36*, 2351.

(8) 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. *Optics express* **2018**, *26*, 11764–11774.