



Inverse design in photonics by topology optimization: tutorial

Christiansen, Rasmus E.; Sigmund, Ole

Published in:
Journal of the Optical Society of America - B

Link to article, DOI:
[10.1364/JOSAB.406048](https://doi.org/10.1364/JOSAB.406048)

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Christiansen, R. E., & Sigmund, O. (2021). Inverse design in photonics by topology optimization: tutorial. *Journal of the Optical Society of America - B*, 38(2), 496-509. <https://doi.org/10.1364/JOSAB.406048>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Inverse design in photonics by topology optimization: tutorial

RASMUS E. CHRISTIANSEN^{1,2,*}  AND OLE SIGMUND^{1,2}

¹Department of Mechanical Engineering, Technical University of Denmark, Nils Koppels Allé, Building 404, 2800 Kongens Lyngby, Denmark

²NanoPhoton—Center for Nanophotonics, Technical University of Denmark, Ørstedes Plads 345A, DK-2800 Kgs. Lyngby, Denmark

*Corresponding author: raelch@mek.dtu.dk

Received 20 August 2020; revised 17 November 2020; accepted 2 December 2020; posted 2 December 2020 (Doc. ID 406048); published 22 January 2021

Topology optimization (TopOpt) methods for inverse design of nano-phonic systems have recently become extremely popular and are presented in various forms and under various names. Approaches comprise gradient- and non-gradient-based algorithms combined with more or less systematic ways to improve convergence, discreteness of solutions, and satisfaction of manufacturing constraints. We here provide a tutorial for the systematic and efficient design of nano-phonic structures by TopOpt. The implementation is based on the advanced and systematic approaches developed in TopOpt for structural optimization during the last three decades. The tutorial presents a step-by-step guide for deriving the continuous constrained optimization problem forming the foundation of the TopOpt method, using a cylindrical metalens design problem as an example. It demonstrates the effect and necessity of applying a number of auxiliary tools in the design process to ensure good numerical modeling practice and to achieve physically realizable designs. Application examples also include an optical demultiplexer. © 2021 Optical Society of America

<https://doi.org/10.1364/JOSAB.406048>

1. INTRODUCTION

We provide an introduction and tutorial for using density-based topology optimization (TopOpt) as an inverse-design tool for photonic structures. We use the commercial software package COMSOL Multiphysics (see Ref. [1]) for the numerical implementation, and provide a set of COMSOL models alongside the paper (Code 1, Ref. [2]; also available at <https://www.topopt.mek.dtu.dk>). These models allow for replication of the presented results at the click of a button, as well as provide a starting point for using TopOpt for more advanced photonics applications. After an initial definition of the spatial, temporal, and physics model (Sections 3–5), we demonstrate how to derive the basic continuous constrained optimization problem, which forms the foundation for using TopOpt for inverse design, using a cylindrical metalens design problem as an example (Section 6). The basic optimization problem derived in Section 6 is solved to illustrate how this first naive approach leads to non-physical solutions (Section 8.A). Following this, the problem formulation is extended (step by step, Sections 8.B and 8.C) until it ensures that solutions are physically sensible and support the design of multi-wavelength metalenses (Section 8.D). As a second example, we consider an optical demultiplexer, demonstrating how the TopOpt problem can be extended to ensure that the solution exhibits robustness towards near-uniform geometric perturbations (Section 9).

Finally, we give a brief introduction to more advanced TopOpt tools, which among others, include tools for imposing geometric length-scales in the design; for ensuring the connectivity of the design; and for ensuring that the design conforms to manufacturing constraints (Section 10).

Stated succinctly, density-based TopOpt [3,4] is an inverse-design tool, used to produce highly optimized structures to serve specialized purposes, applicable across most areas of physics [5–9]. A defining feature is that density-based TopOpt uses adjoint analysis for efficient gradient computations. Conceptually, TopOpt offers unparalleled design freedom, as it allows for point-by-point variation in the material distribution constituting the structure under design, with recent work demonstrating the solution to design problems with more than a billion design variables for mechanical problems [5], proving that TopOpt in any practical sense is able to provide unlimited design freedom. Indeed, the main challenge when applying TopOpt is often to limit the design freedom offered by the method, in a way that conforms with fabrication constraints. Within photonics and plasmonics, TopOpt has received increasing attention over the last two decades [8,10] with recent examples of applications including the design of dielectric multiplexers [11], dielectric metalenses [12–14], extreme dielectric confinement structures [15,16], plasmonic nano-antennas [17], and plasmonic nanoparticles for solar cell

applications [18,19] and for enhanced thermal emission [20] and Raman scattering [21], to name but a few.

By now, inverse-design methods go by many names. The basic TopOpt concepts originate from the structural-optimization community, whereas the photonics community has adopted related inverse-design approaches [10] and so-called objective first approaches [22], which differ mainly in the way the optimization problems are defined and solved. An indispensable part of all inverse-design tools with many degrees of freedom is the adjoint sensitivity analysis [8].

In this work, we consider electromagnetism modeled using Maxwell's equations assuming linear, static, homogeneous, isotropic, non-dispersive, and non-magnetic materials. We assume time-harmonic behavior of the field and consider only transverse electric and transverse magnetic problems with material invariance in the polarization direction. All of these assumptions are made for simplicity and are not required to utilize TopOpt in the context of electromagnetism.

For readers who are interested in underlying method development, programming, and software implementation, we have authored a parallel tutorial paper describing a freely available 200 line MATLAB code implementing basic TopOpt for photonics [23]. That paper also includes an example demonstrating the advantages of gradient-based methods over so-called global-optimization methods for the type of inverse-design problems considered here.

2. GOAL

The ultimate goal in any structural design process is to identify the structure that best solves the problem at hand. A more operative formulation of this goal is as follows:

For a given structural design problem, the goal is to identify a structure that maximizes the desired figure(s) of merit, without violating any of the constraints imposed on the problem.

3. SPACE

We assume a Cartesian coordinate system to describe space, i.e., $\mathbf{r} = \{x, y, z\} \in \mathbb{R}^3$ in three dimensions and $\mathbf{r} = \{x, y\} \in \mathbb{R}^2$ in two dimensions, where \mathbb{R} denotes the real numbers. For numerical modeling of the physics, we define a spatially limited modeling domain, Ω , with the interior Ω_I and the boundary Γ , as illustrated in Fig. 1.

4. TIME

We consider problems that are time-harmonic in nature and disregard any transient behavior. The time dependence is therefore modeled using the time-harmonic exponential factor,



Fig. 1. Sketch of model domain $\Omega = \Omega_I \cup \Gamma$ in 2D; Ω_I denotes the interior and Γ the boundary.

$e^{i\omega t}$, where t denotes time, ω the angular frequency, and i the imaginary unit.

5. PHYSICS

Under the aforementioned assumptions, we consider the following field equations for the electric field, \mathcal{E} , and magnetic field, $\mathcal{H} = \frac{1}{\mu_0} \mathcal{B}$ [24]:

$$\begin{aligned} \nabla \cdot \mathcal{E} &= \frac{\rho}{\varepsilon_r \varepsilon_0}, \quad \nabla \cdot \mathcal{H} = 0, \quad \nabla \times \mathcal{E} = -\mu_0 \frac{\partial \mathcal{H}}{\partial t}, \\ \nabla \times \mathcal{H} &= \mathbf{J}_f + \varepsilon_r \varepsilon_0 \frac{\partial \mathcal{E}}{\partial t}, \quad \mathcal{E} = \mathbf{E} e^{i\omega t}, \quad \mathcal{H} = \mathbf{H} e^{i\omega t}. \end{aligned} \quad (1)$$

Here \mathbf{J}_f and ρ denote the free-current and free-charge densities, and ε_0 and μ_0 denote the vacuum electric permittivity and the vacuum magnetic permeability, respectively. Further, ε_r denotes the relative electric permittivity of the medium through which \mathcal{E} and \mathcal{H} propagate, and finally, \mathbf{E} and \mathbf{H} denote the spatially dependent part of the electric and magnetic fields, respectively.

We assume that the current and charge densities are zero in the interior of the model domain, i.e., $\mathbf{J}_f(\mathbf{r}) = \mathbf{0}$, $\rho(\mathbf{r}) = 0$, $\mathbf{r} \in \Omega_I$. Thereby, the following equations for \mathbf{E} and \mathbf{H} in Ω_I are derived:

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}) - \frac{\omega^2}{c^2} \varepsilon_r(\mathbf{r}) \mathbf{E}(\mathbf{r}) = \mathbf{0}, \quad \mathbf{r} \in \Omega_I \subset \mathbb{R}^3, \quad (2)$$

$$\nabla \times \left(\frac{1}{\varepsilon_r(\mathbf{r})} \nabla \times \mathbf{H}(\mathbf{r}) \right) - \frac{\omega^2}{c^2} \mathbf{H}(\mathbf{r}) = \mathbf{0}, \quad \mathbf{r} \in \Omega_I \subset \mathbb{R}^3. \quad (3)$$

Here $c = 1/\sqrt{\mu_0 \varepsilon_0}$ denotes the speed of light in vacuum. In addition to Eqs. (2) and (3), problem specific boundary conditions are imposed on the boundary of the model domain, Γ , to truncate it appropriately and to introduce external fields.

A. Two-Dimensional Model

Assuming material invariance in the out-of-plane direction (z direction) and assuming that either \mathbf{E} or \mathbf{H} is linearly polarized in the z direction allow for the reduction of Eq. (2) [Eq. (3)] to a scalar Helmholtz equation in two dimensions given as

$$\mathcal{L}_{EM}(\phi) = \nabla \cdot (a \nabla \phi) + \frac{\omega^2}{c^2} b \phi = 0, \quad \mathbf{r} \in \Omega_I \subset \mathbb{R}^2. \quad (4)$$

To model a problem for an E_z -polarized field ($E_x = E_y = 0$), denoted TE in the following, one selects $\phi = E_z$, $a = 1$, and $b = \varepsilon_r$. To model a problem for an H_z -polarized field ($H_x = H_y = 0$), denoted TM in the following, one selects $\phi = H_z$, $a = 1/\varepsilon_r$, and $b = 1$. From the solution to Eq. (4), \mathcal{E} and \mathcal{H} (\mathbf{E} and \mathbf{H}) may be computed using Eq. (1).

6. OPTIMIZATION PROBLEM

As an example of the derivation of the optimization problem forming the basis of TopOpt, we consider the design problem treated in Section 8.A, which may be stated informally as follows:

Design a cylindrical silicon metalens capable of monochromatic focusing of TM-polarized light at normal incidence into a focal line behind the lens. (A cylindrical lens focuses the incident field into a line rather than a point.)

Neglecting the field behavior at the lens ends, this problem may be modeled in 2D by assuming material invariance in the out-of-plane direction, effectively turning the focal line in 3D into a focal point in 2D.

To solve any structural design problem using TopOpt, it must be formulated as a continuous constrained optimization problem, formally written as

$$\begin{aligned} \max_{\xi} \quad & \Phi(\xi), \quad \Phi: [0, 1]^{\Omega_d} \rightarrow \mathbb{R}, \\ \text{s.t.} \quad & c_i(\xi) = 0, \quad c_i: [0, 1]^{\Omega_d} \rightarrow \mathbb{R}, \quad i \in \{0, 1, \dots, \mathcal{N}_i\}, \quad \mathcal{N}_i \in \mathbb{N}_0, \\ & c_j(\xi) < 0, \quad c_j: [0, 1]^{\Omega_d} \rightarrow \mathbb{R}, \quad j \in \{0, 1, \dots, \mathcal{N}_j\}, \quad \mathcal{N}_j \in \mathbb{N}_0. \end{aligned} \quad (5)$$

Here $\xi(\mathbf{r}) \in [0, 1]$ denotes a continuous field, called the design field, over which the function Φ , called the figure of merit (FOM), is sought maximized. The equations $c_i = 0$ denote \mathcal{N}_i equality constraints, and the inequalities $c_j < 0$ denote \mathcal{N}_j inequality constraints. Looking at Eq. (5), one sees that for a given problem, one must select a FOM, $\Phi(\xi)$, which provides a reliable measure of the performance of the design. Further, one must select a set of functions, c_i and c_j , providing reliable measures of all constraints associated with the design problem. (The function Φ is sometimes called the performance indicator, the objective function, or the objective functional.)

Considering again our baseline example, a simple and reliable measure of how well a metalens focuses TM-polarized light impinging on it at normal incidence is obtained by modeling this process for a fixed input power and evaluating the magnitude of the electromagnetic field intensity at the focal point as

$$\Phi = |\mathbf{E}(\mathbf{r}_{\text{FP}}, \varepsilon_r(\mathbf{r}))|^2 = |H_z(\mathbf{r}_{\text{FP}}, \varepsilon_r(\mathbf{r}))|^2, \quad (6)$$

where $\mathbf{E}(\mathbf{r}_{\text{FP}}, \varepsilon_r(\mathbf{r}))$ denotes the electric field at the focal point \mathbf{r}_{FP} , resulting from a TM-polarized illumination of the material distribution $\varepsilon_r(\mathbf{r})$, constituting the metalens and its surrounding environment modeled by solving Eq. (4) assuming TM polarization.

Alternative FOMs for the design problem could be the power flow through the focal spot, or the integral over the focal plane of the difference between the Airy disc that would be formed by an ideal lens and the field profile formed by the lens under design. All these FOMs can be written as simple functions of the electric and/or magnetic fields evaluated in points, lines, or areas.

When formulating a TopOpt problem, the state equation(s) to be solved [e.g., Eq. (4)] may be thought of as a set of equality constraints written as

$$\mathcal{L}_k(\mathbf{x}_k) = \mathbf{f}_k, \quad k \in \{1, 2, \dots, \mathcal{N}_k\}, \quad \mathcal{N}_k \in \mathbb{N}, \quad (7)$$

where \mathcal{L}_k is an operator applying the effect of the physical system to the state field \mathbf{x}_k for the excitation \mathbf{f}_k .

For our baseline example, having identified our FOM and a way of computing it, by solving Eq. (4), we may now

write the design problem as the following optimization problem:

$$\begin{aligned} \max_{\varepsilon_r(\mathbf{r})} \quad & \Phi(H_z(\mathbf{r}_{\text{FP}}), \varepsilon_r(\mathbf{r})), \\ \text{s.t.} \quad & \mathcal{L}_{EM}(H_z(\mathbf{r}), \varepsilon_r(\mathbf{r})) = f_z(\mathbf{r}). \end{aligned} \quad (8)$$

To solve optimization problems of the form in Eq. (5) using TopOpt, we utilize the continuous design field, $\xi(\mathbf{r})$, to interpolate the material parameters in the state equation between the background material(s) and the material(s) constituting the structure under design. (Depending on the problem at hand, different material interpolation schemes should be used [25].)

In our example, we use a scheme that is linear in ξ to interpolate between air and silicon:

$$\varepsilon_r(\xi(\mathbf{r})) = \varepsilon_{r,Si} + \xi(\mathbf{r}) (\varepsilon_{r,Air} - \varepsilon_{r,Si}), \quad (9)$$

where $\varepsilon_{r,Si}$ and $\varepsilon_{r,Air}$ denote the relative permittivity of silicon and air, respectively. That is, we have that $\xi = 0 \Leftrightarrow \varepsilon_r = \varepsilon_{r,Si}$ and $\xi = 1 \Leftrightarrow \varepsilon_r = \varepsilon_{r,Air}$.

The introduction of the interpolation function adds an equality constraint to the optimization problem, as well as the following two inequality constraints, bounding ξ :

$$0 \leq \xi(\mathbf{r}) \leq 1, \quad \mathbf{r} \in \Omega. \quad (10)$$

Including the material interpolation, the final optimization problem representing our baseline metalens design problem example may now be written as

$$\begin{aligned} \max_{\xi(\mathbf{r})} \quad & \Phi(H_z(\mathbf{r}_{\text{FP}}, \varepsilon_r(\xi(\mathbf{r}))), \\ \text{s.t.} \quad & \mathcal{L}_{EM}(H_z(\mathbf{r}), \varepsilon_r(\xi(\mathbf{r}))) = f_z(\mathbf{r}), \\ & \varepsilon_r(\xi(\mathbf{r})) = \varepsilon_{r,Si} + \xi(\mathbf{r}) (\varepsilon_{r,Air} - \varepsilon_{r,Si}), \\ & 0 \leq \xi(\mathbf{r}) \leq 1. \end{aligned} \quad (11)$$

Note that both equality constraints in Eq. (11) are satisfied implicitly (within the accuracy of the numerical model used to solve/evaluate them), as $H_z(\mathbf{r})$ is computed by solving the first constraint for a given $\varepsilon_r(\mathbf{r})$, which in turn is computed by inserting $\xi(\mathbf{r})$ in the second constraint.

To solve the continuous constrained optimization problem efficiently, TopOpt utilizes gradient-based algorithms. (The method of moving asymptotes [26] is often used, as it is efficient for problems with a large design space and few constraints.) As the name suggests, such algorithms require knowledge of the gradients (sensitivities) of Φ , c_i , and c_j with respect to the design field ξ . These sensitivities may be approximated naively using finite differences, which entails solving the system equations for perturbations of each of the design variables for each design iteration. However, doing so is in most cases prohibitively time consuming due to the large number of equations that must be solved. Rather than using finite differences, it is advisable to use adjoint sensitivity analysis, an approach that requires only solving a single (adjoint) equation for the FOM and one for each constraint in the optimization problem, independent of the size of the design space. Adjoint sensitivity analysis has been used for TopOpt in the context of mechanical engineering for decades, since the earliest works [3,27], and in photonics engineering as detailed in the 2011 review by Jensen and Sigmund [8]. As such, adjoint sensitivity analysis serves as one of the cornerstones of

TopOpt. Detailed derivations of adjoint sensitivity analysis in the context of photonics may be found in Keraly *et al.* [28] and in Niederberger *et al.* [29].

A. Note on Adjoint Sensitivity Analysis

Adjoint sensitivity analysis may be performed for the analytical equations describing the model problem before numerical discretization or for the discretized system used in the numerical model. The former has the advantage that it is often simple to derive and implement the equations, while the latter has the advantage that the gradients obtained from solving the adjoint problems are exact in terms of the discretized model.

B. Note on Optimality

We find it appropriate to stress that no mathematical optimization method, be it global-optimization based like a genetic algorithm [30], artificial-intelligence based [31], or gradient based [32], is able to guarantee global optimality of the solution to a non-convex optimization problem. Hence, there is no guarantee that the final design field constitutes the best possible solution to the design problem at hand, unless all possible design permutations have been tested. At most, a mathematical optimization method is able to guarantee local optimality, by, for example, ensuring that the solution fulfils the Karush–Kuhn–Tucker (KKT) conditions [33]. Therefore, the authors caution the reader against the use of the word optimal when describing a structure designed using any optimization-based design method. When the performance of the structure is discussed, the authors propose using the word optimized, as well as providing a measure of the structure performance relative to theoretical limits or relative to references from the literature, cf. [34–36].

C. Note on Design Uniqueness

For a number of photonic design problems, the authors have found that, depending on the initial guess for the design field, TopOpt is able to identify several qualitatively different designs that perform similarly in terms of the FOM. Hence, the final design geometry may depend strongly on the initial guess without the value of the FOM changing significantly.

D. Note on Parameter Tuning

All optimization-based inverse-design methods use a set of parameters that must be tuned for each design problem. As a result, one cannot expect to get quality results by simply applying a method “out of the box,” without *a priori* knowledge and experience with the physics problem at hand as well as with the inverse-design method itself.

7. SOFTWARE

To help the reader start using TopOpt for photonic applications, as well as reproduce the results presented in this paper, a set of COMSOL Multiphysics models (v 5.5) is made available along with this paper. (Note that a COMSOL license is needed to use the software.) Executing the studies in these models without any modifications will reproduce the data used to create Figs. 3–8.

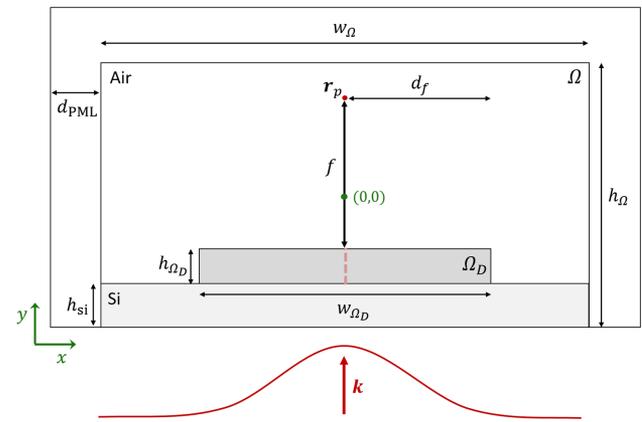


Fig. 2. Model domain, Ω , used for the examples in Sections 8.A–8.D. Values for all lengths defined in the sketch are found in Table 1.

Readers are invited to use these models as the starting point for their own applications. A brief description of the model MetalensCase1.mph may be found in Appendix A.

8. MODEL PROBLEM: THE METALENS

As the first design example, we consider a focusing problem. In particular, we show how to apply TopOpt to design monochromatic and polychromatic cylindrical metalenses capable of focusing Gaussian-enveloped, TM-polarized planewaves at normal incident into a point. The metalenses consist of a region of silicon and air, placed on top of a massive block of silicon in an air background. We start by considering monochromatic focusing (Section 8.A) and demonstrate the beneficial effects of introducing artificial attenuation (Section 8.B) as well as filtering and thresholding operations (Section 8.C) in the design procedure. We then expand the formulation to consider the case of polychromatic focusing (Section 8.D), hereby showing how simple it can be to design broadband cylindrical metalenses using TopOpt.

The model problem is sketched in Fig. 2 and consists of the model domain Ω of height h_Ω and width w_Ω . The domain is thresholded by a perfectly matched layer (PML) [37] of depth d_{PML} on three of four sides, with first-order scattering boundary conditions imposed on the outside of the PML region. On the lower boundary, a TM-polarized planewave, localized using a Gaussian envelope, is introduced using a first-order scattering boundary condition. The choices of boundary conditions made here are made for simplicity, and more elaborate boundary conditions may be applied for a number of reasons, such as to obtain the most accurate modeling of the operating conditions of the lens. The incident TM-polarized field has its H_z component on the boundary given as

$$H_z = \frac{i}{\mu_0 \omega} \cdot \nabla_y E_x,$$

$$E_x = \exp(-x^2/\Delta x_g^2) \exp(-i\mathbf{k} \cdot \mathbf{r}), \quad \mathbf{k} = \frac{2\pi}{\lambda} \langle 0, 1 \rangle. \quad (12)$$

Table 1. Values for Quantities in Fig. 2 for the Cases Considered

	w_Ω [nm]	h_Ω [nm]	d_{PML} [nm]	w_{Ω_D} [nm]	h_{Ω_D} [nm]	h_{Si} [nm]	d_f [nm]	f [nm]
Cases 1–4	4800	1816.2	550	3000	250	125	1500	726.5

A designable region Ω_D of height h_{Ω_D} and width w_{Ω_D} is placed on top of a silicon slab of height h_{Si} and width $w_{\text{Si}} = w_\Omega$. The material distribution in Ω_D is sought tailored to focus the TM-polarized planewave impinging on the region from the silicon into the focal point, \mathbf{r}_p , situated a distance f (focal distance) above the top of Ω_D and a distance d_f from the right boundary of Ω_D . By selecting f and w_{Ω_D} , one determines the numerical aperture of the lens, NA, as

$$\text{NA} = \sin \left(\arctan \left(\frac{w_{\Omega_D}}{2f} \right) \right). \quad (13)$$

Table 1 lists the values of the quantities defined in Fig. 2 for the cases considered in the following. The physical, material, and discretization parameters used in the cases are listed in Table 2. Finally, Table 3 lists parameters related to manipulating the design field and solving the optimization problem.

A. Case 1: Naive Approach

A first naive approach to designing a metalens using TopOpt is to solve Eq. (14), derived in Section 6, using the parameter choices listed in Tables 1–3:

Table 2. Values for the Physical, Material, and Discretization Parameters for the Cases Considered^a

	λ [nm]	n_{Si}	n_{BG}	Δx_g [nm]	NA	h_e [nm]	h_{Ω_D} [nm]
Cases 1–4	550	3.48	1.00	1500	0.9	$\lambda / (10n(\mathbf{r}))$	5

^aWavelength λ ; refractive indices for Si and air, n_{Si} and n_{BG} ; element size used to discretize the model away from the design h_e and in the design domain h_{Ω_D} , h_ϕ ; width of the Gaussian envelope Δx_g ; numerical aperture NA.

Table 3. Values for Quantities Related to Manipulating the Design Field and Solving the Optimization Problem^a

	ξ_{ini}	α_a	β_{ini}	η	r_f	n_{iter}
Case 1	0.5	N/A	N/A	N/A	N/A	1000
Case 2	0.5	1.0	N/A	N/A	N/A	1000
Cases 3–4	0.5	1.0	1.0	0.5	80 [nm]	1000

^aInitial value for the design field ξ_{ini} ; (initial) filter strength β_{ini} ; and threshold level η ; filter radius r_f ; fixed number of inner iterations taken to solve the optimization problem n_{iter} .

$$\begin{aligned} \max_{\xi(\mathbf{r})} \quad & \Phi(H_z(\mathbf{r}), \varepsilon_r(\xi(\mathbf{r}))), \\ \text{s.t.} \quad & \mathcal{L}_{EM}(H_z(\mathbf{r}), \varepsilon_r(\xi(\mathbf{r}))) = f_z(\mathbf{r}), \\ & \varepsilon_r(\xi(\mathbf{r})) = \varepsilon_{r,\text{Si}} + \xi(\mathbf{r})(\varepsilon_{r,\text{Air}} - \varepsilon_{r,\text{Si}}), \\ & 0 \leq \xi(\mathbf{r}) \leq 1. \end{aligned} \quad (14)$$

The reader may solve this problem by executing the study named Optimization in the COMSOL model MetalensCase1.mph. The final design field, $\xi(\mathbf{r})$, obtained by solving the optimization problem, is shown in Fig. 3(a), with black corresponding to silicon and white corresponding to air.

An immediate observation is that $\xi(\mathbf{r})$ contains several gray regions of intermediate values, i.e., $\xi(\mathbf{r}) \in]0, 1[$. These regions appear during the design process, because they allow the metalens to manipulate the wavelength and phase of the

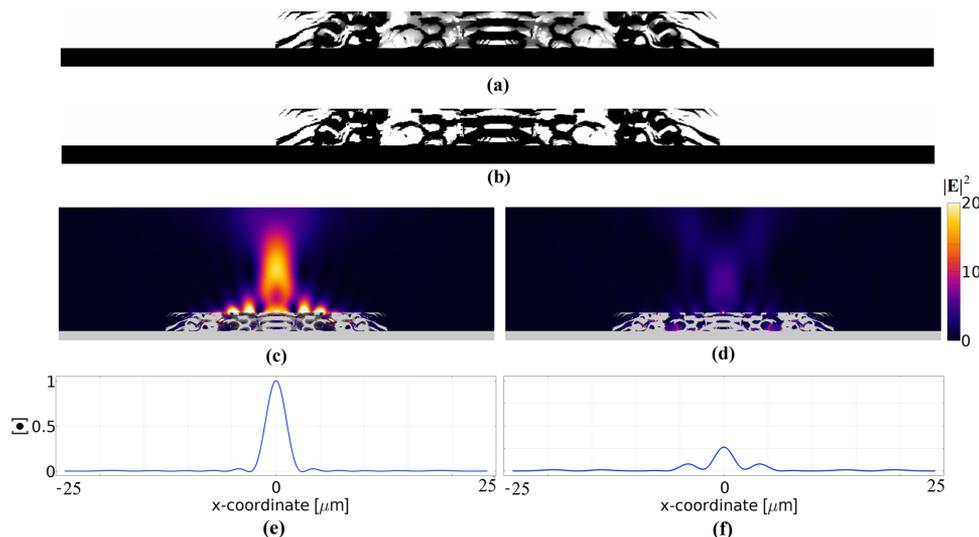


Fig. 3. (a) Design field, $\xi(\mathbf{r})$, obtained by solving the optimization problem in Eq. (14), black (white) corresponds to $\xi = 1$ ($\xi = 0$). (b) Thresholded design field, $\xi(\mathbf{r}) > 0.5$. (c), (d) Electric field intensity $|\mathbf{E}|^2$ (thermal) with the metalens overlaid (gray) for the (c) optimized design in (a) and (d) thresholded design in (b). (e), (f) Max-normalized power flow, $\langle \mathbf{P} \rangle$, through the focal plane for the (e) optimized design in (a) and (f) thresholded design in (b).

electromagnetic field locally with high precision, thus enabling enhancement of the focusing efficiency. While these regions may be beneficial to the performance of the design, they are non-physical, or at least impractical to realize. To fabricate a metalens from the design, one would have to post-process it by removing all intermediate values of $\xi(\mathbf{r})$, or do deep sub-wavelength perforations of the silicon to approximate the refractive index in the gray regions, to obtain a design consisting solely of silicon (black) and air (white).

A simple post-processing approach to obtain a physically realizable design is to threshold $\xi(\mathbf{r})$ at 0.5. Doing so results in the crisp black and white design presented in Fig. 3(b). However, the threshold operation results in a design that is significantly different from the optimized design. Hence, it is unlikely that the two designs will perform equally well. In fact, the FOM drops from $\Phi \approx 18.2$ [V²/m²] for the design in Fig. 3(a) to $\Phi \approx 4.7$ [V²/m²] for the thresholded design in Fig. 3(b). From a visual investigation of the $|E|$ -field emitted from the original design, seen in Fig. 3(c), and the thresholded design, seen in Fig. 3(d), it is clear that the E-fields are different and that significantly less energy is transmitted through the thresholded design. Considering the max-normalized, time averaged power flow, $\langle \mathbf{P} \rangle$, normal to the focal plane in Figs. 3(e) and 3(f) (a metric often used to determine lens performance), it is clearly seen that the power focused at the focal point drops significantly when thresholding the design.

From this analysis, it is clear that the first, naive, approach to applying TopOpt that does not include any regularization of the design field risks yielding poor results because intermediate values of $\xi(\mathbf{r})$ are allowed in the design process but not in the final physical metalens. In the next iteration, we demonstrate a modification ensuring a final optimized design without intermediate values, namely, a scheme for implicitly penalizing intermediate values of ξ .

B. Case 2: Artificial Attenuation

As we care only about maximizing $|E|^2$ at the focal spot, a simple yet effective approach to eliminate intermediate design values is to introduce artificial attenuation of the electromagnetic field for any intermediate value of ξ (also called penalization damping or pamping [38]) by adding an imaginary part to the interpolation scheme for the relative permittivity as

$$\varepsilon_r(\xi) = 1 + \xi(\varepsilon_r - 1) - i\alpha_a \xi(1 - \xi). \quad (15)$$

This way of physical penalization of intermediate design field values is much preferred to explicit penalization, such as adding $\int_{\Omega_D} \xi(1 - \xi) d\mathbf{r}$ to the FOM, which tends to get the result stuck in bad local minima. The added artificial attenuation means that if ξ takes an intermediate value anywhere in Ω_D where there is a non-zero electric field, it will result in less energy propagating through the metalens, which is ultimately detrimental to maximizing Φ . The modified optimization problem in Eq. (16) may be solved by the reader by executing the study named Optimization in the model MetalensCase2.mph:

$$\begin{aligned} \max_{\xi(\mathbf{r})} \quad & \Phi(H_z(\mathbf{r}), \varepsilon_r(\xi(\mathbf{r}))), \\ \text{s.t.} \quad & \mathcal{L}_{EM}(\varepsilon_r(\mathbf{r}), H_z(\mathbf{r})) = \mathbf{f}_z(\mathbf{r}), \\ & \varepsilon_r(\xi) = 1 + \xi(\varepsilon_r - 1) - i\alpha_a \xi(1 - \xi), \\ & 0 \leq \xi(\mathbf{r}) \leq 1. \end{aligned} \quad (16)$$

Executing this study yields the $\xi(\mathbf{r})$ -field shown in Fig. 4(a). In the figure, it is observed that $\xi(\mathbf{r})$ now contains (almost) no intermediate values. In fact, thresholding $\xi(\mathbf{r})$ at 0.5 results in the (almost) identical black and white design, shown in Fig. 4(b).

Evaluating the performance of the two designs in Figs. 4(a) and 4(b), one obtains near identical values of the FOM, namely, $\Phi \approx 17.84$. Inspecting the $|E|^2$ -fields for the optimized and thresholded designs in Figs. 4(c) and 4(d), they indeed look (nearly) identical. Considering the time averaged power flow

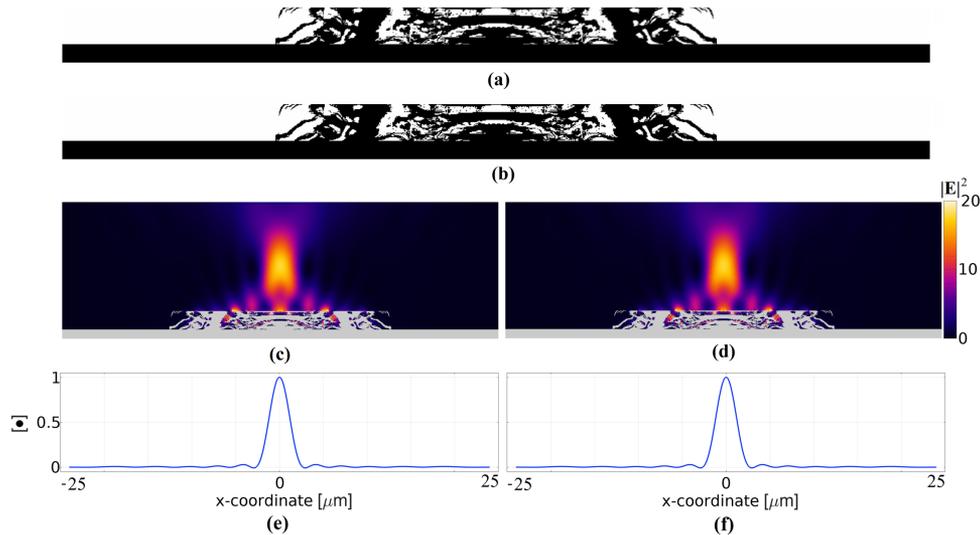


Fig. 4. (a) Design field, $\xi(\mathbf{r})$, obtained by solving the optimization problem in Eq. (16), black (white) corresponds to $\xi = 1$ ($\xi = 0$). (b) Thresholded design field, $\xi(\mathbf{r}) > 0.5$. (c), (d) Electric field intensity $|E|^2$ (thermal) with the metalens overlaid (gray) for the (c) optimized design in (a) and (d) thresholded design in (b). (e), (f) Max-normalized power flow through the focal plane for the (e) optimized design in (a) and (f) thresholded design in (b).

through the focal plane for the two designs, shown in Figs. 4(e) and 4(f), it is seen that the power focused at the focal point does not drop when thresholding the design.

Thus, from a performance perspective, Case 2 yields an optimized monochromatic cylindrical metalens, which is physically realizable, as it consists solely of silicon and air. However, if we look at the design in Fig. 4(b), we observe multiple pixel-by-pixel varying material regions as well as several tiny and narrow features, which raises the following problems: first, the quality of the numerical modeling. Having pixel-by-pixel varying material parameters is in general not good numerical modeling and may, in some cases, result in poor accuracy of the numerical model. Second, it is unlikely that it is possible to accurately manufacture designs with extremely small and rapidly varying features. Third, a design with such features is likely not mechanically stable.

Multiple methods for amending the problem of tiny and single-pixel features have been developed in the context of mechanical engineering [39,40]. In the next iteration we demonstrate a conceptually simple approach.

C. Case 3: Filter and Threshold

This case introduces a well-known filtering and thresholding procedure [41] to control spatial design-field variations. This is a simple, yet effective, way of introducing a weak sense of length-scale into the design and remedies poor numerical modeling with single-pixel features. The smoothing filter is applied by solving the following auxiliary partial differential equation (PDE) in the design domain Ω_D for the filtered design field, $\tilde{\xi}$, with the original design field, ξ , as input and homogeneous Neumann boundary conditions on all other boundaries [42]:

$$-\left(\frac{r_f}{2\sqrt{3}}\right)^2 \nabla^2 \tilde{\xi}(\mathbf{r}) + \tilde{\xi}(\mathbf{r}) = \xi(\mathbf{r}), \quad r_f \geq 0, \quad \mathbf{r} \in \Omega_D. \quad (17)$$

(Note that the filtering procedure is not extended beyond the design domain for simplicity; however, it is trivial to do so.)

Here r_f denotes the desired spatial filtering radius. By varying the filter radius, r_f , it is possible to exert control of the size of the features appearing in the filtered design field.

Next, the filtered field is thresholded using a smoothed approximation of the Heaviside function [39] to recover a nearly discrete design:

$$\tilde{\xi} = \frac{\tanh(\beta \cdot \eta) + \tanh(\beta \cdot (\tilde{\xi} - \eta))}{\tanh(\beta \cdot \eta) + \tanh(\beta \cdot (1 - \eta))}, \quad \beta \in [1, \infty[, \quad \eta \in [0, 1]. \quad (18)$$

Here β controls the threshold sharpness, and η controls the threshold value. At $\beta = 1$, the thresholding has little effect, i.e., $\tilde{\xi} \approx \xi$, whereas for β approaching infinity, the thresholded field takes only values of zero or one, i.e., $\lim_{\beta \rightarrow \infty} \tilde{\xi} \in \{0, 1\}$.

The filter and threshold procedure is applied during the solution to the optimization problem using a continuation scheme, where β is increased from a relatively low starting value to a relatively high stopping value. The low starting value of β allows the design to develop as if no thresholding was performed in the

initial stage of solving the optimization problem, while the high stopping value allows a pure black and white design in the later stage of solving the optimization problem.

The optimization problem being solved in this case is written as

$$\begin{aligned} \max_{\tilde{\xi}(\mathbf{r})} \quad & \Phi(H_z(\mathbf{r}), \varepsilon_r(\tilde{\xi}(\mathbf{r}))), \\ \text{s.t.} \quad & \mathcal{L}_{EM}(\varepsilon_r(\tilde{\xi}(\mathbf{r})), H_z) = \mathbf{f}, \\ & \varepsilon_r(\tilde{\xi}) = 1 + \tilde{\xi}(\varepsilon_r - 1), \\ & \tilde{\xi} = \frac{\tanh(\beta \cdot \eta) + \tanh(\beta \cdot (\tilde{\xi} - \eta))}{\tanh(\beta \cdot \eta) + \tanh(\beta \cdot (1 - \eta))}, \\ & -\left(\frac{r_f}{2\sqrt{3}}\right)^2 \nabla^2 \tilde{\xi}(\mathbf{r}) + \tilde{\xi}(\mathbf{r}) = \xi(\mathbf{r}), \\ & 0 \leq \tilde{\xi}(\mathbf{r}) \leq 1. \end{aligned} \quad (19)$$

The reader may solve the problem in Eq. (19), using a continuation of the β values, by executing the study named Continuation in the model MetalensCase3.mph. Hereby the field, $\tilde{\xi}(\mathbf{r})$, shown in Fig. 5(a), is obtained. It is observed that the final design consists (almost) solely of silicon and air and that this design contains no single-pixel features. Comparing Fig. 4(a) to Fig. 5(a), it is seen that the design consists of significantly fewer and larger features. Thresholding the final design at 0.5 results in the (almost) identical design shown in Fig. 5(b).

Evaluating the performance of the designs in Figs. 5(a) and 5(b), one finds that nearly identical values of the FOM are obtained, namely, $\Phi \approx 17.64$. When comparing the value of Φ with the value for the design from Case 2 in Fig. 4(a), where no filtering was imposed, a decrease of merely $\approx 1\%$ is observed. That is, the removal of the pixel-by-pixel variations has little impact on the performance of the design, whereas the design geometry has been simplified significantly.

Looking at the $|\mathbf{E}|^2$ -field for the original design and the thresholded design [Figs. 5(c) and 5(d)], they look (nearly) identical. Looking at the differences between the time averaged power flow through the focal plane in Figs. 5(e) and 5(f), it is clearly seen that the power focused at the focal point is (nearly) identical for the two designs.

In conclusion, Case 3 presents an optimized monochromatic cylindrical metalens that consists solely of silicon and air with significantly larger features than those seen for Case 2, which makes fabrication simpler. Further, from a numerical modeling point of view, the design in Case 3 does not contain rapid pixel-by-pixel material variations, which may jeopardize numerical convergence and precision.

D. Case 4: Multiple State-Equation Optimization

There exist many applications where a structure is sought designed to maximize a set of FOMs simultaneously. A simple way of doing this is to agglomerate these into a single FOM, e.g., using a p -norm. Another way is to use a min/max formulation [43]. In this example, we demonstrate the former.

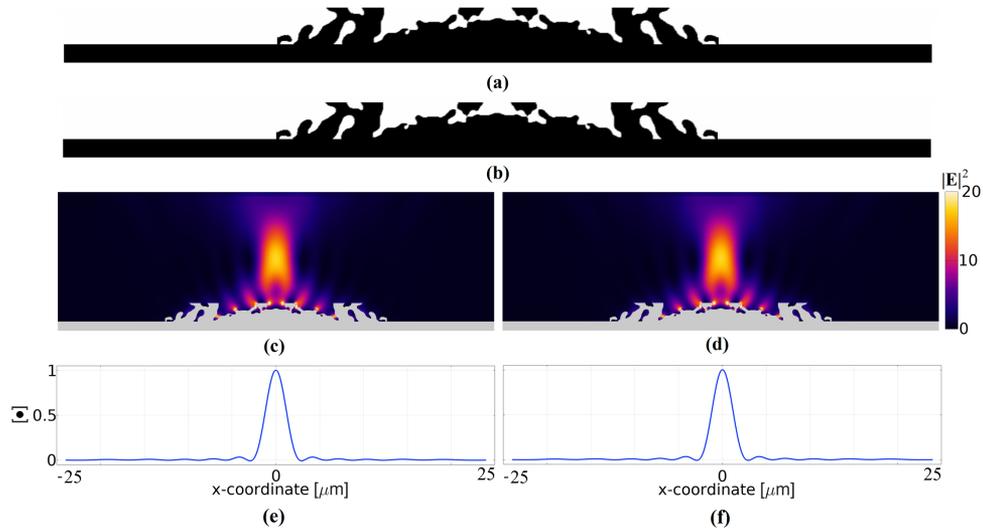


Fig. 5. (a) Design field, $\xi(\mathbf{r})$, obtained by solving the optimization problem in Eq. (19), black (white) corresponds to $\xi = 1(\xi = 0)$. (b) Thresholded design field, $\xi(\mathbf{r}) > 0.5$. (c), (d) Electric field intensity $|\mathbf{E}|^2$ (thermal) with the metalens overlaid (gray) for the (c) optimized design in (a) and (d) thresholded design in (b). (e), (f) Max-normalized power flow through the focal plane for the (e) optimized design in (a) and (f) thresholded design in (b).

We consider the problem of designing a metalens for broadband operation. That is, instead of only maximizing the focusing efficiency at a single wavelength, we target three wavelengths in a 100 nm band simultaneously, i.e., $\lambda \in \{500 \text{ nm}, 550 \text{ nm}, 600 \text{ nm}\}$. To do this, we reformulate the FOM as follows:

$$\tilde{\Phi} = \left(\sum_{i=1}^{N_\lambda=3} \sqrt{\Phi_i(H_z(\lambda_i, \mathbf{r}, \varepsilon_r(\tilde{\xi}(\mathbf{r})))} \right)^2, \quad (20)$$

where Φ_i is defined in Eq. (6). This formulation inherently puts the greatest weight on the Φ_i , which takes the lowest value. To evaluate Eq. (20), a set of i state equations must be solved. Note that the solution to the state equations can in principle be done in parallel and hence does not need to increase wall clock time.

Besides the above change to the FOM, the optimization problem being solved remains the same as in Eq. (19). The reader may solve the new problem by executing the study named Continuation in the model MetalensCase4.mph.

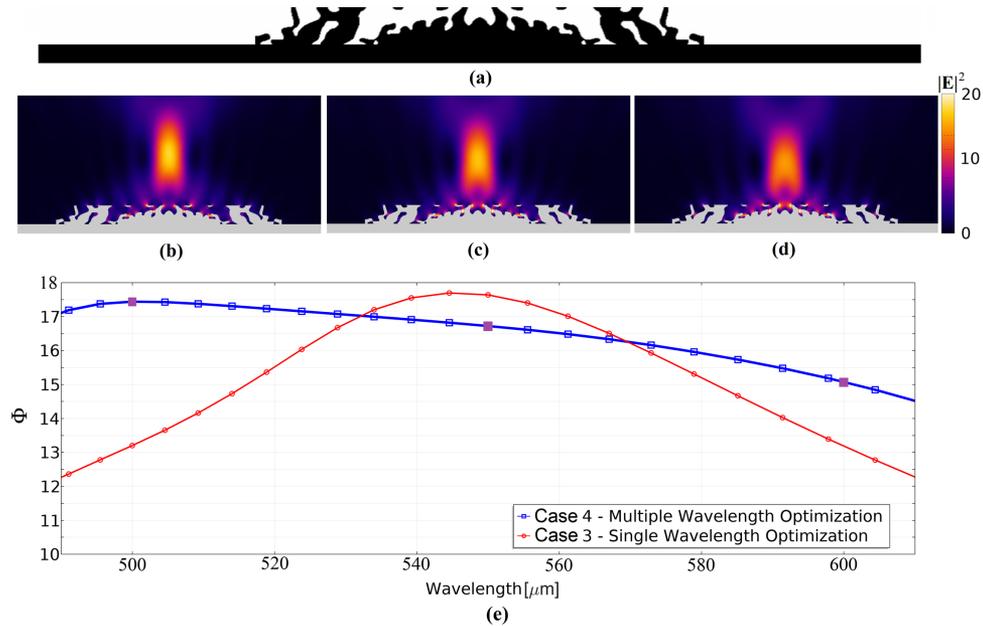


Fig. 6. (a) Design field, $\xi(\mathbf{r})$, obtained by solving the optimization problem in Eq. (14), black (white) corresponds to $\xi = 1(\xi = 0)$. (b)–(d) Electric field intensity $|\mathbf{E}|^2$ (thermal) with the metalens overlaid (gray) for the optimized design in (a) for (b) $\lambda = 500 \text{ nm}$, (c) $\lambda = 550 \text{ nm}$, (d) $\lambda = 600 \text{ nm}$. (e) Φ as a function of wavelength for the design in (a) (blue line) and the design in Fig. 5(a) (red line).

Doing so yields the design field, $\bar{\xi}(\mathbf{r})$, shown in Fig. 6(a). The $|\mathbf{E}|^2$ -fields for the three targeted wavelengths are shown in Figs. 6(b)–6(d). To demonstrate that we have designed a metalens with a better average performance over the targeted wavelengths compared to targeting only a single wavelength, we evaluate the performance of the design in Fig. 6(a) and the design optimized for the single wavelength $\lambda = 550$ nm, shown in Fig. 5(a), for wavelengths in the interval from 480 nm to 620 nm. The value of the FOM in Eq. (6) as a function of wavelength is plotted in Fig. 6(e). In this figure, it is observed that the design optimized for a single wavelength (unsurprisingly) performs best at that wavelength. However, when considering the full wavelength interval, it is clearly observed that the metalens optimized for three wavelengths performs best when averaged over the three wavelengths and also when averaged over the entire interval. Thus, by accepting a performance drop at the central wavelength, it is possible to design a lens with significantly better broadband performance. More constant performance over the frequency interval may be targeted by raising the p -norm value from two to a higher value. However, note that too high values of p (e.g., $p > 10$) make the problem highly nonlinear, which possibly causes ill convergence.

9. MODEL PROBLEM: THE DEMULTIPLEXER

As an illustration of the broad applicability of TopOpt within photonics design, we consider the design of an optically small photonic demultiplexer (device footprint $\approx 2.4\lambda_1^2$), intended to direct light from a single input waveguide to two different output waveguides depending on the wavelength of the incident light. The example also demonstrates how TopOpt may be used to create designs exhibiting geometric robustness towards near-uniform variations in the geometry—a type of variation similar to those associated with sample over (under) exposure (or etching) during various nano-fabrication processes [44–46].

The model problem considered in this example is shown in Fig. 7, and Table 4 lists values for the length quantities defined in the figure.

The model problem setup consists of the model domain Ω with height h_Ω and width w_Ω and contains air as the background medium with the input (output) waveguides and a beam splitter consisting of silicon. The designable region Ω_D , constituting the demultiplexer, has height h_{Ω_D} and width w_{Ω_D} and is placed at the center of Ω . The material distribution in Ω_D is sought tailored to maximize the time averaged power flow from the input waveguide into one of the two output waveguides, depending on the wavelength of the light. The model domain is thresholded by a PML of depth d_{PML} on the left side, with a first-order scattering boundary condition imposed on the outside of the PML region. First-order scattering boundary conditions are also imposed on the remaining three sides of Ω . On the left boundary, a TE-polarized planewave, localized using a Gaussian envelope, is introduced into the input waveguide of height $h_{\text{WG},1}$ and width w_{WG} . The incident TE-polarized field has the E_z component given as

$$E_z = \exp(-y^2/\Delta y_g^2) \exp(-i\mathbf{k}_j \cdot \mathbf{r}),$$

$$\mathbf{k}_j = \frac{2\pi}{\lambda_j} \langle 0, 1 \rangle, \quad j \in \{1, 2\}. \quad (21)$$

More advanced boundary conditions may be applied for a number of reasons, and the choices of the boundary conditions made here are made purely for simplicity.

We demonstrate how to optimize the design to achieve performance robustness towards (nearly) uniform geometric perturbations by using the double filtering method [47]. In brief, the method consists of applying the filter and threshold procedure described in Section 8.C twice on the design field, where in the second application three different threshold values are applied to obtain three different realizations of the design fields corresponding to under (over) etching. Six state equations are then solved, two for each of the three realizations of the design fields, corresponding to the two wavelengths targeted by the demultiplexer. The design problem may be formulated as the following constrained optimization problem:

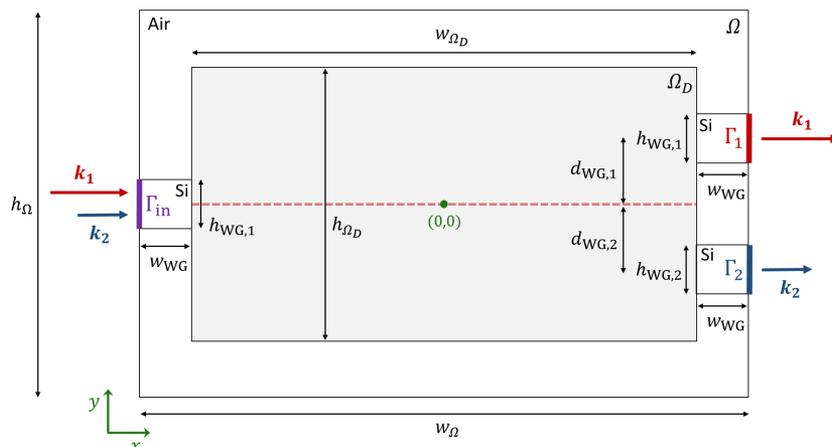


Fig. 7. Model domain Ω considered in Section 9. Values for the lengths are found in Table 4.

Table 4. Values for Lengths in Fig. 7

Parameter	w_Ω	h_Ω	d_{PML}	w_{Ω_D}	h_{Ω_D}	$h_{WG,1}$	$h_{WG,2}$	w_{WG}	$d_{WG,1}$	$d_{WG,2}$
Value [nm]	6000	3000	1000	2000	2000	299	356	2000	511	483

Table 5. Values for the Physical, Material, and Discretization Parameters Used in the Example in Section 9^a

Parameter	λ_1	λ_2	n_{Si}	n_{BG}	h_e	$h_{dx}(h_{\Omega_D})$	$h_{dy}(h_{\Omega_D})$	Δy_g
Value	1300 nm	1550 nm	3.48	1.0	$\frac{\lambda_1}{10n}$	10 nm	10 nm	356 nm

^aWavelength λ ; refractive indices for Si and air, n_{Si} and n_{BG} ; element size used to discretize the model away from the design h_e , and in the design domain h_{dx} , h_{dy} ; width of the Gaussian envelope Δy_g .

Table 6. Values for Quantities Related to Manipulating the Design Field and Solving Optimization Problem in Eq. (22)

Parameter	ξ_{ini}	$\beta_{ini,1}$	$\beta_{ini,2}$	η_1	$\eta_{2,1}$	$\eta_{2,2}$	$\eta_{2,3}$	$r_{f,1}$	$r_{f,2}$	n_{iter}
Value	0.5	5	2.5	0.5	0.3	0.5	0.7	100 [nm]	50 [nm]	1000

$$\begin{aligned}
\max_{\xi(\mathbf{r})} \quad & \Phi = \sum_{j=1}^{N_\lambda=2} \sum_{k=1}^{N_k=3} \sqrt{\left(\int_{\Gamma_j} \mathbf{P} \left(\lambda_j, \tilde{\xi}(\mathbf{r}, \eta_{2,k}) \right) \right) d\mathbf{r}} \quad j \in \{1, 2\}, \quad k \in \{1, 2, 3\}, \\
\text{s.t.} \quad & \mathcal{L}_{EM} \left(\lambda_j, \varepsilon_r \left(\tilde{\xi}(\mathbf{r}, \eta_{2,k}) \right), E_z \right) \mathbf{x} = \mathbf{f}(\lambda_j), \\
& \varepsilon_r \left(\tilde{\xi} \right) = 1 + \tilde{\xi} (\varepsilon_r - 1) + i \tilde{\xi} \left(1 - \tilde{\xi} \right), \\
& \tilde{\xi} = \frac{\tanh(\beta_2 \cdot \eta_{2,k}) + \tanh(\beta_2 \cdot (\tilde{\xi} - \eta_{2,k}))}{\tanh(\beta_2 \cdot \eta_{2,k}) + \tanh(\beta_2 \cdot (1 - \eta_{2,k}))}, \\
& - \left(\frac{r_{f,2}}{2\sqrt{3}} \right)^2 \nabla \tilde{\xi}(\mathbf{r}) + \tilde{\xi}(\mathbf{r}) = \tilde{\xi}(\mathbf{r}), \\
& \tilde{\xi} = \frac{\tanh(\beta_1 \cdot \eta_1) + \tanh(\beta_1 \cdot (\tilde{\xi} - \eta_1))}{\tanh(\beta_1 \cdot \eta_1) + \tanh(\beta_1 \cdot (1 - \eta_1))}, \\
& - \left(\frac{r_{f,1}}{2\sqrt{3}} \right)^2 \nabla \tilde{\xi}(\mathbf{r}) + \tilde{\xi}(\mathbf{r}) = \tilde{\xi}(\mathbf{r}), \\
& 0 \leq \tilde{\xi}(\mathbf{r}) \leq 1.
\end{aligned} \tag{22}$$

The physical, material, and discretization parameters used in the model are listed in Table 5.

Parameters related to manipulating the design field and solving Eq. (22) are listed in Table 6.

Given the choice of parameters, the demultiplexer is designed for near-uniform erosion (dilation) of ± 8 nm around the nominal design, approximating variation that may be experienced in electron beam lithography from over (under) exposure during fabrication.

The optimization problem may be solved by executing the Continuation study in the model DemultiplexerExample.mph. Doing so results in the design fields presented in Figs. 8(a)–8(c). The absolute value of the power flow through the demultiplexer at $\lambda = 1300$ nm is plotted in Figs. 8(d)–8(f) and at $\lambda = 1550$ nm in Figs. 8(g)–8(i).

The absolute and relative transmittances for the designs in Figs. 8(a)–8(c) are listed in Table 7. The absolute transmittance is computed as the power flow through the output waveguides (Γ_1 and Γ_2 in Fig. 7) relative to the power flow through a waveguide of identical width

to the input waveguide, excited identically. The relative transmittance is computed as the power flow through the output waveguides (Γ_1 and Γ_2 in Fig. 7) relative to the power flow through the input waveguide (Γ_{in} in Fig. 7).

The optimization problem is formulated such that identical power flow, through the relevant output waveguide, across the six cases is preferable to maximize the FOM. Looking at the first three columns in Table 6 and scaling these by the total power flow through the reference waveguide, this is exactly what is observed for the optimized demultiplexer design. If desired, one could trivially change the weighting of the individual cases to target larger (smaller) transmittance for a particular case.

10. BRIEF DISCUSSION ON USEFUL TOOLS

Since its inception in the late 1980s, a range of auxiliary tools have been developed for use with density-based TopOpt. While it is outside the scope of this tutorial to demonstrate all these tools, the following subset is discussed in brief.

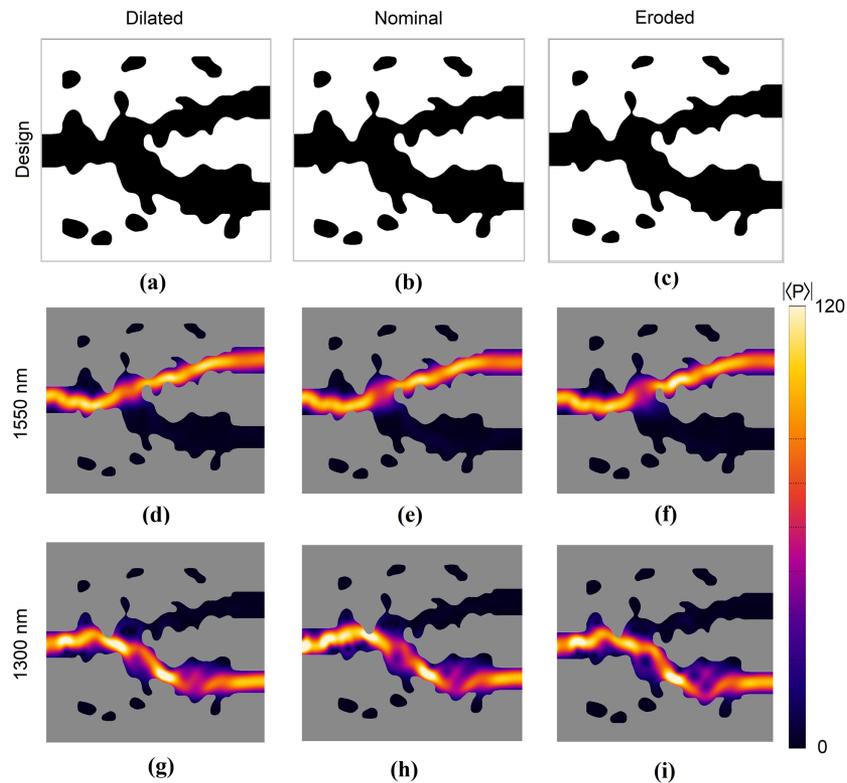


Fig. 8. (a) Dilated, (b) nominal, and (c) eroded demultiplexer design with black (white) corresponding to silicon (air). (d)–(f) Magnitude of the power flow at $\lambda = 1300$ nm for the designs in (a)–(c). (g)–(i) Magnitude of the power flow at $\lambda = 1550$ nm for the designs in (a)–(c).

Table 7. Absolute and Relative Transmittances for the Designs in Fig. 8^a

$\lambda \setminus T$	Fig. 8(a)	Fig. 8(b)	Fig. 8(c)	Fig. 8(a)	Fig. 8(b)	Fig. 8(c)
	T_{Abs}	T_{Abs}	T_{Abs}	T_{Rel}	T_{Rel}	T_{Rel}
1300 nm	≈ 0.32	≈ 0.32	≈ 0.32	≈ 0.86	≈ 0.69	≈ 0.82
1550 nm	≈ 0.31	≈ 0.31	≈ 0.30	≈ 0.83	≈ 0.85	≈ 0.81

^aTotal power flow through the reference waveguide is $\approx 7.06 \cdot 10^{-4}$ W/nm at $\lambda = 1300$ nm and $\approx 7.47 \cdot 10^{-4}$ W/m at $\lambda = 1550$ nm.

The performance and geometry of structures designed using TopOpt have, in some cases, been found to depend strongly on the choice of the material interpolation function [e.g., Eq. (9)]. For this reason, a number of different interpolation schemes have been developed for different applications. It is thus advisable to dedicate time and effort to identifying a good interpolation scheme for a given problem. One example is the design of plasmonic nanoparticles for localized extreme field enhancement, where a nonlinear interpolation scheme was demonstrated to outperform several other interpolations schemes [Eq. (9)] by orders of magnitude in terms of the final design performance [25].

When designing structures for some photonic and plasmonic applications, the optimized geometries have been found to contain features with details on the order of a few nanometers [16,21]. However, even with state-of-the-art fabrication techniques, there is a lower limit to the manufacturable feature size. To ensure that designs are optimized while adhering to fabrication limitations, a number of tools for imposing a minimum length-scale in the design have been developed. If optimizing

for geometric robustness, length-scale may be imposed straightforwardly using the double filter technique [47] (assuming a constant topology across all design realizations). If the design problem is highly sensitive to geometric perturbations, making it impossible to design high-performance geometrically robust structures, or if geometric robustness is not a concern for the design at hand, one may instead use a geometric constraint approach [45] to impose a minimum length-scale.

For some problems, such as the design of photonic membrane structures, physics dictates that all members of the structure must be connected, as free-floating members are impossible to realize. Using TopOpt, it is straight forward to include a connectivity constraint in the design process, e.g., by using a virtual temperature method [48].

For some fabrication techniques, only specific design variations are allowed. As an example, in standard electron beam lithography, the design blueprint must be two-dimensional, as little-to-no variation of the design in the out-of-plane direction is possible. Using TopOpt, it is straightforward to keep the design field constant in a particular spatial dimension by applying a simple mapping operation to the design field and integrate

the design sensitivities in that spatial dimension to attain correct sensitivity information, while maintaining a constant design geometry in that direction. Other fabrication techniques allow for a smooth variation of the design in the out-of-plane direction, while simultaneously limiting these variations through a maximally allowable structural slant angle. Using TopOpt, it is simple to create such designs with varying height and a limited (or fixed) slant angle using a smoothed threshold operation [14].

Finally, when considering the use of the density-based TopOpt approach described here as a design tool for photonic structures, it is worth noting that the method has been demonstrated to be able to eliminate the need for applying proximity-effect-correction (PEC) in both electron-beam and optical-projection lithography [49] by modifying the filtering and threshold procedure and using the design field directly as the exposure dose or fabrication mask, respectively. Further, the filtering and threshold procedure used in TopOpt has been demonstrated to be applicable for performing the PEC step for electron-beam lithography using an optimization-based procedure [46].

Many additional tools and techniques have been developed and explored, such as design variable linking [50] and accounting for random geometric uncertainties using perturbation techniques [51]. An overview of a range of different tools for ensuring lengths-scale and manufacturability may be found in [40].

11. CONCLUSION

We have presented a tutorial for applying TopOpt to photonic structural design using the design of a set of cylindrical metalenses and a demultiplexer as examples of applications. First, a simple naive TopOpt problem formulation was derived, and it was demonstrated that this formulation led to several problems. Iteratively, a number of well-established methods were introduced in the problem formulation, and it was demonstrated how these enabled the design of physically realizable and geometrically robust structures.

While this work for simplicity considers only examples in two spatial dimension, an extension of the method to three spatial dimensions is trivial from the point of view of the TopOpt method. Using the COMSOL-Multiphysics-based software provided with this work, it is only a matter of using a 3D component instead of a 2D component for modeling the physics. The main challenge when extending the method from two to three spatial dimensions is the computational bottleneck associated with solving the electromagnetic state equation(s) for large-scale problems. This is, however, a challenge related to the numerical modeling of the physics rather than to the TopOpt method. One possible approach for treating (some) large-scale three-dimensional TopOpt problems is to consider a time domain model for the physics and using finite difference time domain solvers [52]; another is to use overlapping domains techniques [53,54] to partition the physics problem into computationally tractable sub-problems.

The reader is invited to adapt the software provided with this work to their photonics research, hereby unlocking the power of TopOpt for the design and optimization of structures for their particular applications.

Finally, we invite the more numerically inclined reader to study our accompanying MATLAB tutorial paper [23], which, apart from a 200 line compact and transparent MATLAB implementation of TopOpt problems similar to the ones discussed here, also includes a short comparison to a non-gradient genetic-algorithm-based approach.

APPENDIX A: COMSOL MODEL DESCRIPTION

This appendix provides a brief description of the COMSOL Multiphysics model.

MetalensCase1.mph used to design the first iteration of a metalens in Section 8.A.

In the model, Global Definitions contains the definitions of all model parameters defined in Tables 1–3, such as the lens width, targeted wavelength, and design resolution.

The framework used to set up and model the physics and design problem is the standard 2D component. Under the 2D component, the Definitions node is used to define: the objective function (FOM); operations related to the design field; operations related to plotting the solutions; material interpolation function, Eq. (9); a probe for printing the FOM; all mapping operators used to manipulate the design field; and PML domains. The Geometry node sets up all geometric elements used to build the model domain (Fig. 2). The Materials node contains definitions of all the material parameters for the non-designable regions of the model domain. The Electromagnetic Waves, Frequency Domain node defines and configures the physics model [Eq. (2)]. To enable optimization, the material parameters in the Wave Equation, Electric sub-node is set to User defined, and the relative permittivity is set equal to the material interpolations function [Eq. (9)]. Two Scattering Boundary Condition sub-nodes are defined, where one is used to introduce the incident field into the model domain along the lower boundary and the other is applied to the remaining boundaries of the domain. The Optimization node is used to define the optimization problem, i.e., the FOM and the design variable constraint from Eq. (8). Finally, the Mesh node is used to set up and construct the finite element mesh for the model domain.

Two Studies are included in the model. The first is named Optimization and is used to execute the design procedure. In this study, an Optimization node is added to define the optimization method used, the maximum allowed number of model evaluations, and the type of optimization. The Frequency Domain study step defines the frequencies that are targeted in the optimization and the physics interfaces that are part of the study step. Executing this study executes the optimization. The second study is named Analysis and is used to analyze the final design by performing a narrowband frequency sweep around the targeted frequency.

Finally, in the Results node, five Derived Values are defined to compute the FOM and the power flow through the focal point and focal plane. Further, eight plots are set up to allow easy visualization of the optimized design, the resulting $|\mathbf{E}|^2$ -field, power flow in the focal plane for the optimized structure obtained in the Optimization study, and the structure under analysis in the Analysis study.

Note that a set of auxiliary parameters, not mentioned in the paper, are defined in the COMSOL Multiphysics model(s) for practical purposes, making it easier to set up the model geometry, etc.

Funding. Danmarks Grundforskningsfond (DNRF147); Villum Fonden (8692).

Disclosures. The authors declare that there are no conflicts of interest related to this paper.

REFERENCES

- COMSOL AB, "COMSOL multiphysics v. 5.5," <https://www.comsol.com>.
- R. E. Christiansen and O. Sigmund, "Inverse design in photonics by topology optimization," Code 1, figshare (2021), <https://doi.org/10.6084/m9.figshare.12833777>.
- M. P. Bendsøe and N. Kikuchi, "Generating optimal topologies in structural design using a homogenization method," *Comput. Methods Appl. Mech. Eng.* **71**, 197–224 (1988).
- M. P. Bendsøe and O. Sigmund, *Topology Optimization* (Springer, 2003).
- N. Aage, E. Andreassen, B. S. Lazarov, and O. Sigmund, "Giga-voxel computational morphogenesis for structural design," *Nature* **550**, 84–86 (2017).
- J. Alexandersen and C. S. Andreasen, "A review of topology optimization for fluid-based problems," *Fluids* **5**, 29 (2020).
- J. Alexandersen, N. Aage, C. S. Andreasen, and O. Sigmund, "A density-based topology optimization methodology for thermoelectric energy conversion problems," *Struct. Multidiscip. Optim.* **57**, 1427–1442 (2018).
- J. S. Jensen and O. Sigmund, "Topology optimization for nano-photonics," *Laser Photon. Rev.* **5**, 308–321 (2011).
- R. E. Christiansen and E. Fernandez-Grande, "Design of passive directional acoustic devices using topology optimization—from method to experimental validation," *J. Acoust. Soc. Am.* **140**, 3862–3873 (2016).
- S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vuckovic, and A. W. Rodriguez, "Inverse design in nanophotonics," *Nat. Photonics* **12**, 659–670 (2018).
- 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**, 374–377 (2015).
- Z. Lin, V. Liu, R. Pestourie, and S. G. Johnson, "Topology optimization of freeform large-area metasurfaces," *Opt. Express* **27**, 15765–15775 (2019).
- H. Chung and O. D. Miller, "High-NA achromatic metalenses by inverse design," *Opt. Express* **28**, 6945–6965 (2020).
- R. E. Christiansen, Z. Lin, C. R. Cames, Y. Salamin, S. E. Kooi, J. D. Joannopoulos, M. Soljačić, and S. G. Johnson, "Fullwave Maxwell inverse design of axisymmetric, tunable, and multi-scale multi-wavelength metalenses," *Opt. Express* **28**, 33854–33868 (2020).
- X. Liang and S. G. Johnson, "Formulation for scalable optimization of microcavities via the frequency-averaged local density of states," *Opt. Express* **21**, 30812–30841 (2013).
- F. Wang, R. E. Christiansen, Y. Yu, J. Mørk, and O. Sigmund, "Maximizing the quality factor to mode volume ratio for ultra-small photonic crystal cavities," *Appl. Phys. Lett.* **113**, 241101 (2018).
- E. Wadbro and C. Engström, "Topology and shape optimization of plasmonic nano-antennas," *Comput. Methods Appl. Mech. Eng.* **293**, 155–169 (2015).
- S. P. Madsen, J. Christiansen, R. E. Christiansen, J. Vester-Petersen, S. H. Møller, H. Lakhotiya, E. E. Adnan Nazir, S. Roesgaard, O. Sigmund, J. Lissau, E. Destouesse, M. Madsen, B. Julsgaard, and P. Balling, "Improving the efficiency of upconversion by light concentration using nanoparticle design," *J. Phys. D* **53**, 073001 (2020).
- J. Christiansen, J. Vester-Petersen, S. Roesgaard, S. H. Møller, R. E. Christiansen, O. Sigmund, S. P. Madsen, P. Balling, and B. Julsgaard, "Strongly enhanced upconversion in trivalent erbium ions by tailored gold nanostructures: toward high-efficient silicon-based photovoltaics," *Sol. Energy Mater. Sol. Cells* **208**, 110406 (2020).
- Z. A. Kudyshev, A. V. Kildishev, V. M. Shalaev, and A. Boltasseva, "Machine-learning-assisted metasurface design for high-efficiency thermal emitter optimization," *Appl. Phys. Rev.* **7**, 021407 (2020).
- R. E. Christiansen, J. Michon, M. Benzaouia, O. Sigmund, and S. G. Johnson, "Inverse design of nanoparticles for enhanced Raman scattering," *Opt. Express* **28**, 4444–4462 (2020).
- J. Lu and J. Vučković, "Objective-First design of high-efficiency, small-footprint couplers between arbitrary nanophotonic waveguide modes," *Opt. Express* **20**, 7221–7236 (2012).
- R. E. Christiansen and O. Sigmund, "A 200 line MATLAB code for inverse design in photonics by topology optimization: tutorial," *J. Opt. Soc. Am. B* **38**, 510–520 (2021).
- D. J. Griffiths, *Introduction to Electrodynamics*, 4th ed. (Pearson Education Limited, 2014).
- R. E. Christiansen, J. Vester-Petersen, S. P. Madsen, and O. Sigmund, "A non-linear material interpolation for design of metallic nano-particles using topology optimization," *Comput. Methods Appl. Mech. Eng.* **343**, 23–39 (2019).
- K. Svanberg, "A class of globally convergent optimization methods based on conservative convex separable approximations," *SIAM J. Optim.* **12**, 555–573 (2002).
- D. A. Tortorelli and P. Michaleris, "Design sensitivity analysis: overview and review," *Inverse Prob. Eng.* **1**, 71–105 (1994).
- C. M. Lalau-Keraly, S. Bhargava, O. D. Miller, and E. Yablonovitch, "Adjoint shape optimization applied to electromagnetic design," *Opt. Express* **21**, 21693–21701 (2013).
- A. C. R. Niederberger, D. A. Fattal, N. R. Gauger, S. Fan, and R. G. Beausoleil, "Sensitivity analysis and optimization of sub-wavelength optical gratings using adjoints," *Opt. Express* **22**, 12971–12981 (2014).
- D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Learning* (Addison, 1989).
- S. J. Russell and P. Norvig, *Artificial Intelligence A Modern Approach*, 3rd ed. (Prentice Hall, 2010).
- K. Svanberg, "The method of moving asymptotes—a new method for structural optimization," *SIAM J. Optim.* **12**, 555–566 (2002).
- J. Nocedal and S. J. Wright, *Numerical Optimization Second Edition* (Springer, 2006).
- S. Molesky, P. Chao, W. Jin, and A. W. Rodriguez, "Global T operator bounds on electromagnetic scattering: upper bounds on far-field cross sections," *Phys. Rev. Res.* **2**, 033172 (2020).
- M. Gustafsson, K. Schab, L. Jelinek, and M. Capek, "Upper bounds on absorption and scattering," *New J. Phys.* **22**, 073013 (2020).
- J. Michon, M. Benzaouia, W. Yao, O. D. Miller, and S. G. Johnson, "Limits to surface-enhanced Raman scattering near arbitrary-shape scatterers," *Opt. Express* **27**, 35189–35202 (2019).
- J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Phys.* **114**, 185–200 (1994).
- J. S. Jensen and O. Sigmund, "Topology optimization of photonic crystal structures: a high-bandwidth low-loss T-junction waveguide," *J. Opt. Soc. Am. B* **22**, 1191–1198 (2005).
- F. Wang, B. S. Lazarov, and O. Sigmund, "On projection methods, convergence and robust formulations in topology optimization," *Struct. Multidiscip. Optim.* **43**, 767–784 (2011).
- B. S. Lazarov, F. Wang, and O. Sigmund, "Length scale and manufacturability in density-based topology optimization," *Arch. Appl. Mech.* **86**, 189–218 (2016).
- J. K. Guest, J. H. Prevost, and T. Belytschko, "Achieving minimum length scale in topology optimization using nodal design variables and projection functions," *Int. J. Numer. Methods Eng.* **61**, 238–254 (2004).
- B. S. Lazarov and O. Sigmund, "Filters in topology optimization based on Helmholtz-type differential equations," *Int. J. Numer. Methods Eng.* **86**, 765–781 (2011).

43. O. Sigmund and J. S. Jensen, "Length scale and manufacturability in density-based topology optimization," *Phil. Trans. R. Soc. Lond. A* **361**, 1001–1019 (2003).
44. M. Jansen, B. S. Lazarov, M. Schevenels, and O. Sigmund, "On the similarities between micro/nano lithography and topology optimization projection methods," *Struct. Multidiscip. Optim.* **48**, 717–730 (2013).
45. M. Zhou, B. S. Lazarov, F. Wang, and O. Sigmund, "Minimum length scale in topology optimization by geometric constraints," *Comput. Methods Appl. Mech. Eng.* **293**, 266–282 (2015).
46. E. H. Eriksen, A. Nazir, P. Balling, J. Vester-Petersen, R. E. Christians, O. Sigmund, and S. P. Madsen, "Dose regularization via filtering and projection: an open-source code for optimization-based proximity-effect-correction for nanoscale lithography," *Microelectron. Eng.* **199**, 52–57 (2018).
47. R. E. Christiansen, B. S. Lazarov, J. S. Jensen, and O. Sigmund, "Creating geometrically robust designs for highly sensitive problems using topology optimization—acoustic cavity design," *Struct. Multidiscip. Optim.* **52**, 737–754 (2015).
48. Q. Li, W. Chen, S. Liu, and L. Tong, "Structural topology optimization considering connectivity constraint," *Struct. Multidiscip. Optim.* **54**, 971–984 (2016).
49. M. Zhou, B. S. Lazarov, and O. Sigmund, "Topology optimization for optical projection lithography with manufacturing uncertainties," *Appl. Opt.* **53**, 2720–2729 (2014).
50. R. E. Christiansen and O. Sigmund, "Designing meta material slabs exhibiting negative refraction using topology optimization," *Struct. Multidiscip. Optim.* **54**, 469–482 (2016).
51. B. S. Lazarov, M. Schevenels, and O. Sigmund, "Topology optimization with geometric uncertainties by perturbation techniques," *Int. J. Numer. Methods Eng.* **90**, 1321–1336 (2012).
52. Y. Elesin, B. Lazarov, J. Jensen, and O. Sigmund, "Time domain topology optimization of 3D nanophotonic devices," *Photon. Nanostruct. Fundam. Appl.* **12**, 23–33 (2014).
53. M. J. Gander and H. Zhang, "A class of iterative solvers for the Helmholtz equation: factorizations, sweeping preconditioners, source transfer, single layer potentials, polarized traces, and optimized Schwarz methods," *SIAM Rev.* **61**, 3–76 (2019).
54. Z. Lin and S. G. Johnson, "Overlapping domains for topology optimization of large-area metasurfaces," *Opt. Express* **27**, 32445–32453 (2019).