nature photonics

REVIEW ARTICLE

https://doi.org/10.1038/s41566-018-0246-9

# Inverse design in nanophotonics

Sean Molesky[1], Zin Lin[2], Alexander Y. Piggott[3], Weiliang Jin[1], Jelena Vučković[3] and Alejandro W. Rodriguez[1]*

**Recent advancements in computational inverse-design approaches — algorithmic techniques for discovering optical structures based on desired functional characteristics — have begun to reshape the landscape of structures available to nanophotonics. Here, we outline a cross-section of key developments in this emerging field of photonic optimization: moving from a recap of foundational results to motivation of applications in nonlinear, topological, near-field and on-chip optics.**

The development of devices in nanophotonics[1] — the study of light in structures with features near or below the scale of the electromagnetic wavelength — has historically relied on intuition-based approaches. Impetus for a new device develops from an a priori known physical effect, and then specific features are matched to suitable applications by tuning small sets of characteristic parameters. This approach has had a long track record of success, and has given rise to a rich and widely exploited library of templates (Fig. 1, top). Combining the principles behind index guiding, bandgap engineering and material resonances, this collection of designs enables remarkable manipulation of light over bands of frequencies spanning from the ultraviolet to the mid-infrared: group velocity can be slowed by more than two orders of magnitude[2], light confined to volumes thousands of times smaller than its free-space wavelength[3] and resonances made to persist in micrometre-sized volumes for tens of millions of optical cycles[4].

Yet, as the scope of nanophotonics broadens to include large bandwidth or multi-frequency applications, nonlinear phenomena and dense integration, continuing with this prototypical approach poses a challenge of increasing complexity. For instance, consider the design of a wavelength-scale structure for enhancing nonlinear interactions[5], discussed below. Even in the simplest case, several interdependent characteristics must be simultaneously optimized, among which are large quality factors at each individual wavelength, and nonlinear overlaps, which must be controlled in as small a volume as possible. In such a situation, the standard photonic design library offers no clear way to proceed. There is no definite reason to expect that an optimal design can be found in any of the traditional templates, or that such a design necessarily exists. Moreover, the performance of a given device is likely to be highly dependent on the characteristics of the problem, and as greater demands are placed on functionality, it becomes increasingly doubtful that any one class of structures will have the broad applicability of past devices. The lack of evident strategies for broadband applications also brings to attention the space of structures included in standard designs. Predominately, traditional designs are repetitive mixtures and combinations of highly symmetric shapes described by a small collection of tuning parameters. Since intuition-based optimization is then carried out in terms of these parameters, bearing aside known bounds based on fundamental principles[6–8], typically little is known about how close any one particular device comes to performance limits, or how it compares with modified design possibilities.

The ability to produce and evaluate novel device platforms based on nonlinear and broadband processes, from solar energy conversion[9–13] and thermal energy manipulation[14,15] to on-chip integration[16–19], will objectively impact the future of nanophotonics. If the total design space performance for a given class of problem can be even partially characterized, an immense amount of research effort can be saved, and a new approach for investigating fundamental limits of nanophotonic devices could emerge. In this Review, we bring attention to a collection of recent results showcasing the usefulness of inverse design for both comparing the relative performance of possible structures, and for discovering photonic devices in instances where traditional intuition-based strategies prove difficult to implement. We begin by providing background on inverse design in nanophotonics, highlighting some of the major developments in this field. From this basis of understanding, we then turn to discussion of emerging applications and experimental challenges, motivating ways in which inverse-design techniques have and could be employed in nonlinear, topological, near-field and integrated optics.

## Background

The driving motivations behind inverse design have been present for at least several hundreds of years. They are part of the same family of ideas that led Bernoulli to consider the brachistochrone problem, Maupertuis to propose the principle of least action, and Ambartsumian to question the relation between a set of eigenvalues and its generating differential equation[20]. There are at least two central thrusts: first, to determine the extent that the characteristics of a solution, either actual or desired, determine the system from which they are derived; and second, to find effective algorithms for working from desired characteristics to physical systems.

In the context of nanophotonics, inverse-problem formulations are understandably much more recent[21]. (For an overview of topology optimization in the field of mechanics, predating related ideas in nanophotonics, see Bendsøe and Sigmund[21]. The reader should also be aware that, to focus our review, we will not discuss the closely related development of sensitivity analysis[22] in the microwave community, or the application of inverse design for computer-generated holograms[23], even though conceptually there is almost no difference between these areas. We direct readers interested in this more general history of inverse design to the reviews by Jensen and Sigmund[24] and Sigmund[25], and the articles contained therein.) Nanophotonic optimization, on which we will focus exclusively, began in the late 90s with the work of Spühler et al.[26] and Cox and Dobson[27] (beginning of Fig. 1). In the first article, Spühler et al. designed and fabricated a $SiO_2$/SiON telecom-fibre to ridge-waveguide coupler.

[1]Department of Electrical Engineering, Princeton University, Princeton, NJ, USA. [2]John A. Paulson School of Engineering and Applied Sciences Harvard University, Cambridge, MA, USA. [3]Ginzton Laboratory, Stanford University, Stanford, CA, USA. *e-mail: arod@princeton.edu

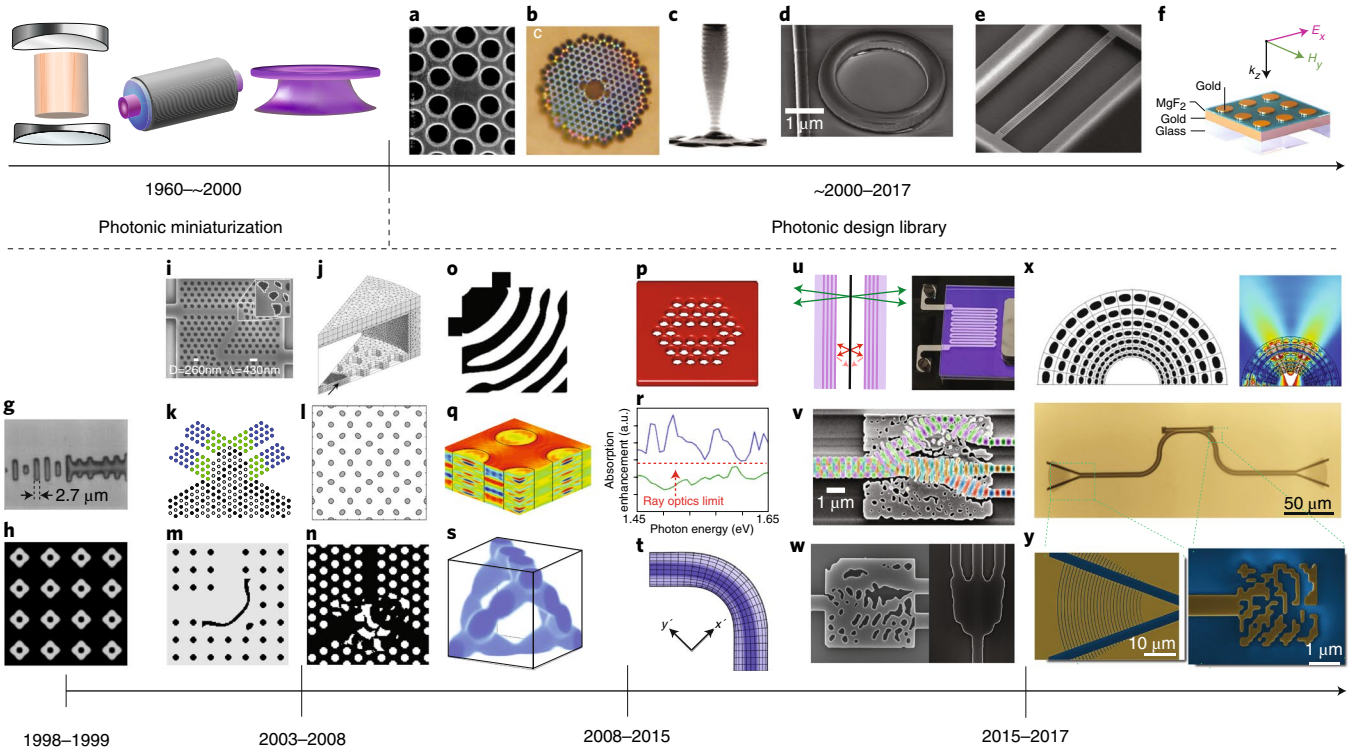



**Fig. 1 | Progression of photonic design templates.** Top: during the second half of the twentieth century, advancement in fabrication capabilities allowed photonic engineering to expand into the microscale and nanoscale. Over the past two decades, this capability has led to the growth of a rich standard library of photonic design templates. From left to right, the examples shown for photonic miniaturization depict a Fabry–Pérot cavity, a fibre cavity and a microdisk resonator. **a**–**f**, The examples for the photonic design library are a PhC defect cavity from Painter et al.[126] (**a**), a PhC fibre from a review by Knight[127] (**b**), a micropost cavity from Pelton et al.[87] (**c**), a microring resonator from Xu et al.[128] (**d**), a nanobeam resonator from Eichenfield et al.[129] (**e**) and a plasmonic sensor from Liu et al.[130], with $E_x$ denoting the electric field, $H_y$ the magnetic field and $k_z$ the wave vector in Cartesian coordinates. (**f**). Bottom: a visual companion to the timeline of developments in photonic optimization described in the text. **g**–**y**, The images are: 1998–1999, a $SiO_2$/SiON telecom-fibre to ridge-waveguide coupler from Spühler et al.[26] (**g**) and an optimized two dimensional PhC from the work of Dobson and Cox[27] (**h**); 2003–2008, a topology optimized 'Z' waveguide bend in a silicon-based PhC from Borel et al.[41] (**i**), a single mode defect cavity from Frei et al.[67] (**j**), a 1.50 × 1.55 μm demultiplexer from Håkansson and Sánchez-Dehesa[48] (**k**), a level-set bandgap optimized PhC from Kao et al.[44] (**l**), a 90° PhC waveguide bend form Jensen and Sigmund[42] (**m**) and a nanoimprinted two-port demultiplexer from Borel et al.[59] (**n**); 2008–2015 an optimized 90° waveguide bend from Tsuji and Hirayama[61] (**o**), a cross-section of an inverse designed holey waveguide fibre from Lu et al.[68] (**p**), a unit cell of a radial junction silicon wire array for solar absorption from Alaeian et al.[9] (**q**), the solar absorption enhancement resulting from a computationally optimized surface texturing from Ganapati et al.[10] (**r**), a three-dimensional fcc bandgap optimized PhC from Men et al.[75] (**s**) and a transformation optics motivated waveguide bend from Liu et al.[16] (**t**); 2015–2017, an optimized layered thermal emitter from Ilic et al.[14] (**u**), a three-port silicon-on-insulator demultiplexer from Frellsen et al. [17] (**v**), a silicon-on-insulator two-port demultiplexer and broadband 1:3 power splitter from Piggott et al.[110,113] (**w**), a topology optimized hyperlens and corresponding magnetic field magnitude profile from Otomori et al.[131] (**x**) and a compact on-chip Fabry–Pérot resonator form Yu et al.[132] (**y**). Panels reproduced from: **a**, ref. [126], AAAS; **b**, ref. [127], Springer Nature Ltd; **c**, ref. [87], APS; **d**, ref. [128], OSA; **e**, ref. [129], Springer Nature Ltd; **f**, ref. [130], American Chemical Society; **g**, ref. [26], IEEE; **h**, ref. [27], Copyright ©1999 Society for Industrial and Applied Mathematics; **i**, ref. [41], OSA; **j**, ref. [67], AIP Publishing; **k**, ref. [48], OSA; **l**, ref. [44], Springer Nature Ltd; **m**, ref. [42], AIP Publishing; **n**, ref. [59], OSA; **o**, ref. [61], IEEE; **p**, ref. [68], OSA; **q**, ref. [9], IOP; **r**, ref. [10], IEEE; **s**, ref. [75], OSA; **t**, ref. [16], OSA; **u**, ref. [14], Springer Nature Ltd; **v**, ref. [17], OSA; **w**, refs [110,113], Springer Nature Ltd; **x**, ref. [131], Springer Nature Ltd; **y**, ref. [132], OSA.

Using a genetic algorithm to determine the width of the SiON core over a distance of 138 μm in 3 μm steps, a 2 dB increase in efficiency was achieved compared with direct coupling. In the second article, Cox and Dobson applied a gradient-search algorithm to the problem of bandgap optimization: starting from a two-dimensional periodic structure composed of two materials, the bandgap was enlarged by symmetric alterations of the dielectric composition, resulting in a 34% increase in bandgap size. The methods used to perform structural optimization in these two early applications of photonic inverse design stand as archetypes for classification, involving either genetic (evolutionary)[28] or gradient-based approaches[29]. Crucially, in genetic algorithms, the sensitivity of the objective to the individual design parameters (derivative information of the objective function) is not necessarily determined. Furthermore, even if gradient information is incorporated into any of the subroutines, it does not deterministically drive the algorithm. This alteration offers both benefits and drawbacks. For complicated non-convex objective functions[30], genetic algorithms may be favoured when gradient information is either not easily accessible or unreliable[31]. However, they tend to be orders of magnitude more computationally expensive for a similar number of design parameters and, depending on the problem, more likely to miss locally optimal designs.

In the five-year period following these initial investigations, notable extensions and contributions were made. Among them, Doosje et al. showed that plane-wave expansions could be used to implement inverse calculations of three-dimensional face-centred cubic (fcc) photonics crystals (PhCs)[32]; Cox and Dobson successfully extended their original work to include in-plane electric

fields[33]; Felici and Heinz considered optimal designs for coupling fibres to adiabatic tapers, making use of cascading algorithms combining coarse- and fine-grained parameterizations[34]; Geremia et al. formulated the design of PhC cavities as a Lagrangian maximization problem involving a generalized cost functional defined in terms of the desired optical characteristics[35]; Jiang et al. used a genetic algorithm to achieve mode matching between PhC and fibre waveguides[36]; and Kizilats et al. applied optimization techniques to improve the design of radiofrequency patch antennas[37]. With few exceptions, most works focused on two classes of problems, involving either bandgap optimization in PhCs or mode coupling in waveguide geometries[34,38]. These problems commonly shared a high degree of symmetry, and typically employed gradient search methods to optimize a small selection of parameters within a predetermined family of periodic designs. Large-scale optimization methods nevertheless began to be simultaneously pursued in the closely related area of sensitivity analysis[39,40], demonstrating huge speed-ups in the characterization of the impact of defects and roughness on photonic devices.

The works of Jensen and Sigmund et al.[41–43], Burger, Kao, Osher and Yablonovitch[44–46], Gerken and Miller[47], Håkansson, Sánchez-Dehesa and Sanchis et al.[48,49], and Preble, M. Lipson and H. Lipson[50] between 2003 and 2005 began to clearly reshape this landscape (second grouping in Fig. 1). First, inverse methods were extended to include a wider range of technologically relevant applications, including PhC waveguide bends showing sub-1% transmission losses over a broad band of frequencies[42], few-wavelength-thick devices capable of acting as frequency demultiplexers[47,48] and more varied PhC configurations for the creation of wide bandgaps[44,50]. Second, the introduction of adjoint density[41,42,47] and level-set topology optimization[44,45,51] vastly broadened the generality and computational efficiency of inverse design.

At a high level, the major benefit of introducing the concepts of level-set and density topology optimization is that they provide a systematic way to organize design possibilities. In the level-set method, a given design is described by splitting the physical optimization domain $D$ into level sets of a solution function $\Phi(\mathbf{x})$ that varies continuously over space $\mathbf{x} \in D$ (defined over each voxel or mesh in a computational cell). Mimicking the description of Burger et al.[46], to consider smooth candidate structures consisting of two materials one can define a partitioning

$$\begin{aligned} \Omega_1 &= \{\Phi(\boldsymbol{x}) < 0\} \\ \Omega_2 &= \{0 < \Phi(\boldsymbol{x})\} \end{aligned} \tag{1}$$

that maps an otherwise continuously varying function to a binary domain. To move towards a device design, $\Phi(\mathbf{x})$ is then evolved either through an equation of motion (such as the Hamilton–Jacobi equation) or via gradients, causing it to settle at local minima. Specifying the optimization domain in this way allows for floating boundaries between material components without the need to provide explicit parameterizations. Through the use topological derivatives[52], this approach also allows for appearance of voids while mitigating conditions conducive to the development of ultrafine (pixel checkerboard) features[53].

In density topology optimization, an even broader design space is considered. Drawing from the finite discretization of the underlying physical problem in a numerical method, each node (line segment, pixel or voxel) within a computational cell is treated as a degree of freedom and 'relaxed' continuously in some range. Mimicking the implementation of Jensen and Sigmund[43] for instance, the permittivity of each node $\epsilon_i$ in structures consisting of two materials is defined as

$$\epsilon_i = \epsilon_1 + \lambda_i(\epsilon_2 - \epsilon_1) \tag{2}$$

where $\epsilon_{1,2}$ denotes the permittivity of the two materials and $\lambda \in [0, 1]$ acts as a relaxation parameter. The problem of finding an optimal structure over the space of all discretized designs then amounts to determining the value of $\lambda$ for each node, while ensuring that the latter takes only its extreme values.

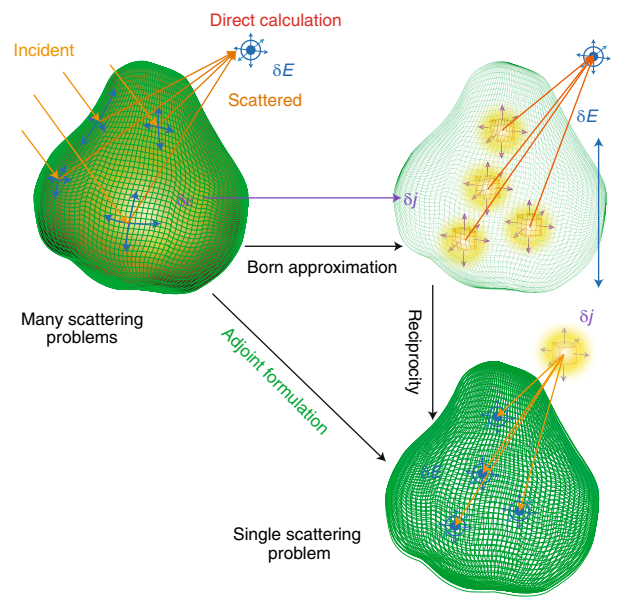
For either approach, the space of possible designs is enormous (on the order of the set of nodes in the optimization domain) and hence to ensure any hope of convergence to a local optimum, iterations based on the gradient of the objective with respect to the design parameters are needed[54]. Here, gradients provide the optimization algorithms with a direction of improvement. While there is no provable guarantee of a globally optimal solution, it is still nevertheless possible to find designs that perform remarkably well (in some cases even globally optimal[55]). The computational effort required to determine and make use of this information is made manageable owing to the adjoint method[56], described in Box 1.

Following these first forays of large-scale optimization methods in photonics, much of the focus moved to the investigation of structures and applications of increasing complexity, including early works in solar energy harvesting[9–11], dispersion engineering[57], subwavelength focusing[58], nanoimprint lithographic methods for fabricating topology optimized structures[59] and nonlinear switching[60]. At the same time, the boost in computational power provided by adjoint techniques naturally led to questions concerning the incorporation of realistic constraints, and computationally workable extensions to larger design domains (third grouping of images Fig. 1). In essence, while increasing generality greatly improved device performance, these gains nevertheless came with new challenges. Without constraints, the feature sizes produced by either level-set or density topology optimization methods are limited only by the size of the chosen domains.

In the density topology optimization approach, the permittivity can vary continuously to make use of gradients. Depending on how material constraints are implemented, this can result in intermediate 'grey' structures, where many iterations are spent fine-tuning designs with graded-index features before settling on binary or piecewise-constant structures. Dealing with such issues amounts to finding the correct point of trade-off in searching the design space. Advancements that took place during the past decade provided foundational understanding of how such a trade-off occurs and how the imposition of penalization filters on either objectives or degrees of freedoms can impact device performance.

In 2008, Tsuji and Hirayama[61] investigated the issue of losses at a 90° bend, showing that the replacement of the linear relaxation parameter in equation (2) with a smooth function approaching a step discontinuity yields similar convergence properties and structures as those obtained using established penalization methods[62]. In a similar spirit, Wang et al.[63] studied basic trade-offs associated with applications of topology and few-parameter optimization methods to the realization of slow-light PhC waveguides, showing that while the performance of structures generated by density topology optimization is typically superior, producing group velocity indices approaching 300, similar order of magnitude gains can be achieved via simple shape variations. Simultaneously, adaptations of large-scale methods to include fabrication tolerances were pursued by Sigmund generally[64], and by Oskooi et al. in the context of robust waveguide tapers[65], with the goal of producing designs that provide some degree of optimality while remaining robust with respect to structural variations. In the former work, this was done by introducing erosion and dilation operators that alter the optimized structure; in the latter, this was done by considering the performance of the structure relative to alterations in the direction of steepest descent of the objective function. To control feature sizes within the level-set method, a typical approach is to exploit shape parameterizations that automatically satisfy the desired minimum feature constraint, known as a geometry projection method. As an example of such a method, Frei et al. observed that a truncated expansion of the level-set function in a basis of

**Box 1 | Adjoint methods**




Let $\mathcal{F}[\psi(\mathbf{x}),\epsilon(\mathbf{x})]$ be some objective functional, a real number judging the performance of a given design, $\psi(\mathbf{x})$ a field that $\mathcal{F}$ is optimized relative to, $\epsilon(\mathbf{x})$ a controllable set of design parameters, and $\overline{\mathcal{M}}[\psi(\mathbf{x}),\epsilon(\mathbf{x})]=0$ a collection of constraint equations between $\psi(\mathbf{x})$ and $\epsilon(\mathbf{x})$, with $\mathbf{x}$ denoting the spatial coordinate. The relevant derivative (sensitivity) information for locally optimizing $\epsilon(\mathbf{x})$ is then given by the total variation of $\mathcal{F}$ with respect to the design parameters

$$\delta_{\epsilon(\mathbf{x})}\mathcal{F}=\frac{\delta\mathcal{F}}{\delta\epsilon(\mathbf{x})}+\frac{\delta\mathcal{F}}{\delta\psi(\mathbf{x})}\frac{\delta\psi(\mathbf{x})}{\delta\epsilon(\mathbf{x})} \tag{3}$$

Both $\delta\mathcal{F}/\delta\epsilon(\mathbf{x})$ and $\delta\mathcal{F}/\delta\psi(\mathbf{x})$ can be determined directly from the objective functional using analytic methods. The determination of $\delta\psi(\mathbf{x})/\delta\epsilon(\mathbf{x})$, however, is, at face value, problematic. Directly, this quantity is defined by the equation

$$\delta_{\epsilon(\mathbf{x})}\bar{\mathcal{M}}=\frac{\delta\bar{\mathcal{M}}}{\delta\epsilon(\mathbf{x})}+\frac{\delta\bar{\mathcal{M}}}{\delta\psi(\mathbf{x})}\frac{\delta\psi(\mathbf{x})}{\delta\epsilon(\mathbf{x})}=0 \tag{4}$$

giving $\delta\psi(\mathbf{x})/\delta\epsilon(\mathbf{x})=-(\delta\overline{\mathcal{M}}/\delta\psi(\mathbf{x}))^{-1}\delta\overline{\mathcal{M}}/\delta\epsilon(\mathbf{x})$. The combination of $(\delta\overline{\mathcal{M}}/\delta\psi(\mathbf{x}))^{-1}$ and $\delta\overline{\mathcal{M}}/\delta\epsilon(\mathbf{x})$ in this expression is computationally costly to solve. Since $\delta\overline{\mathcal{M}}/\delta\epsilon(\mathbf{x})$ produces a list of variations for each $\epsilon(\mathbf{x})$ (a matrix), solving for this expression requires either as many solves as there are optimization unknowns (iterative algorithms) or pre-computing the matrix inverse, which is typically dense, and then applying the resulting form.

To avoid such a calculation, which would severely limit the type of problems that could be tractably considered, the collection of constraints, $\overline{\mathcal{M}}$ is substituted for the objective $\mathcal{F}$. Inserting equation (4) into equation (3) one obtains

$$\delta_{\epsilon(\mathbf{x})}\mathcal{F}=\frac{\delta\mathcal{F}}{\delta\epsilon(\mathbf{x})}-\frac{\delta\mathcal{F}}{\delta\psi(\mathbf{x})}\left(\frac{\delta\overline{\mathcal{M}}}{\delta\psi(\mathbf{x})}\right)^{-1}\frac{\delta\overline{\mathcal{M}}}{\delta\epsilon(\mathbf{x})} \tag{5}$$

showing that the combination

$$\lambda_{\mathbf{x}}(\cdot)=\delta\mathcal{F}/\delta\psi(\mathbf{x})\,(\delta\overline{\mathcal{M}}/\delta\psi(\mathbf{x}))^{-1} \tag{6}$$

acts as a linear functional on $\delta\overline{\mathcal{M}}/\delta\epsilon(\mathbf{x})$. This gives the adjoint equation

$$\left(\frac{\delta\bar{\mathcal{M}}}{\delta\psi(\mathbf{x})}\right)^{\dagger}\lambda(\mathbf{x})=\frac{\delta\mathcal{F}}{\delta\psi(\mathbf{x})} \tag{7}$$

where † is the adjoint operator. As $\delta\mathcal{F}/\delta\psi(\mathbf{x})$ is simply a number for each $\psi(\mathbf{x})$ (a vector), only a single solve is then required to determine $\lambda(\mathbf{x})$. If the constraint equations are linear in $\psi(\mathbf{x})$, then $\delta\overline{\mathcal{M}}/\delta\psi(\mathbf{x})$ is nearly the original set of constraints, and the same factorization and conditioners used to solve for $\psi(\mathbf{x})$ can be used to reduce the cost of determining $\lambda(\mathbf{x})$.

The above equations show that adjoint methods yield full derivative information from the solution of a problem that is, in every respect, no harder to solve than that of the original field. In certain situations, this can be understood intuitively. Consider for instance the typical electromagnetic problem of optimizing the scattered power from a body due to a given incident field. To iteratively determine the permittivity profile, the sensitivity of the scattered power due to a small change, $\delta\epsilon$, is required at every position in the body, as illustrated in the figure. An analogue of the straightforward approach to obtaining this information as proposed by equation (4) proceeds as follows. Within the Born approximation, the first-order variation in the polarization $\delta j=\delta\epsilon\times(\text{incident field})$ at each position is equivalent to an independent (induced) current source. Therefore, the sensitivity of the scattered power due to any possible $\delta\epsilon$ variation can be found by carrying out a field calculation at each position in the body (requiring as many solutions as there are polarization unknowns/degrees of freedom), and then taking the inner product between the resulting field and the initial scattered field. However, by Lorentz reciprocity the roles of the field and source are interchangeable, and the sensitivity of the scattered power on each permittivity degree of freedom can be equally obtained from a single calculation: the field strength at each position in the body due to a source at the location where the scattered power is computed. From this perspective, the entire sensitivity calculation is recast as the solution of a single 'adjoint' scattering problem, much like equation (7). While this concrete example provides helpful intuition, the true power of adjoint methods is their generality, for example, their applicability to nonreciprocal and nonlinear materials[56]. Even time-domain problems can be treated without altering the above derivation if $\mathbf{x}$ is thought of as a parameterization of the total computational domain (including time).

radial functions implicitly determines the smallest feature size in a given problem[66]. This observation was then applied to demonstrate a tripling of the Purcell factor of a single-defect PhC cavity[67]. Practical evaluation of the relative importance of different facets of topology optimization algorithms also extend to techniques for improving computational efficacy, and key ideas emerged in the works of Lu et al., Men et al., Liang and Johnson, Liu et al. and Elesin et al. These are summarized below.

- In general, the inverse problem of determining both field and structural unknowns can be broken into two subproblems that can be individually solved by either density or level-set optimization methods, explored in Lu et al.[68]. First, one relaxes the absolute constraints imposed by Maxwell's equations and determines the pseudo electric field that simultaneously minimizes both the objective and residual with respect to solution of Maxwell's equations. Next, one considers the electric field found in the first step as a given, and seeks instead to minimize said residual by solving for the correct permittivity. An optimal structure is produced by alternating between these two, less demanding, subproblems[69]. Similar notions were used earlier by Geremia et al.[35] and Englund et al.[70] in studies of defect cavities.
- In cases where it is feasible to determine the dominant mode contributions to the objective function, the total optimization problem can be limited to more tractable subspaces[71]. If the allowed variation between any two iterations is also small, then exploiting an alternating technique as above allows the modes of one structure to be used as an approximation for the modes of another. Combining these two ideas results in a considerably more accommodating system of equations that can be solved by semi-definite programming techniques. (These ideas were explored by Men et al.[71] in the context of bandgap optimization.)
- Along a similar vein, the computational cost and convergence characteristics of problems involving multiple frequency bands, and/or resonant systems, can be dramatically improved using window functions and complex-frequency deformations when the objective function is analytic[72]. First, the objective is multiplied by a meromorphic function peaked around the frequency bands of interest (for example, a Lorentzian). By analytically continuing to the complex plane, the entire objective is then obtained from the spectral residues of the window function, requiring far fewer calculations. In Men et al.[71], this formulation was applied to produce three-dimensional cavities with Purcell factors larger than $10^5$. Another approach for general broadband problems is to move to the time domain (computing the spectral response due to a single broadband source), allowing large frequency ranges to be captured within a single simulation, as explored by Elesin et al.[73]. However, care must be taken to ensure proper conditioning of the objective function, particularly in highly resonant cases.
- Finally, the strengths of inverse design and transformation optics[74] are highly complementary[16]. Coordinate transformations often lead to a simplified understanding of the boundary behaviour that must be achieved for a device to perform efficiently. For example, to keep modes from scattering around a waveguide bend, any permittivity profile that has the same effect as a coordinate transformation producing a 90° rotation will work. However, natural transformations, such as the change to circular coordinates one would be inclined to consider in the above example, tend to produce material profiles that are difficult to fabricate, (unrealistic permittivities, anisotropy and so on). However, this second problem is well suited for inverse design. Knowledge of the exact boundary conditions that must be satisfied means that Maxwell's equations do not need to be solved at each iteration. Instead, the algorithm may focus solely on the objective function, for instance minimizing dielectric anisotropy.

The results stemming from these computational insights are also intriguing. For instance, Men et al.[75] found that even without imposing fabrication constraint, their inverse-design algorithm could not find photonic structures with fractional bandgaps larger than ~30% (for index contrasts smaller than 1:3.6). Perhaps surprisingly, such gaps are only slightly larger than those previously attributed to hand-designed fcc PhCs[76]. Considering the large number of degrees of freedom and initial designs explored, the results indicate that there is likely little room for further bandgap engineering. The suggestion of such a fundamental limitation, while negative, is quite appealing from a theoretical perspective. Instinctively, the size of a bandgap must be, in some way, ultimately bound by material constraints. Regardless of how materials are spatially distributed, a rule like the one discovered by Sigmund and Hougaard for two-dimensional PhC bandgaps[77] (also motivated by inverse design) should also exist in the three-dimensional case. Yet, the existence of an argument proving this fact remains open.

## Current and emerging applications

In this section, we highlight the current and emerging applications of inverse design.

**Nonlinear optics.** The utility of engineered resonances for nonlinear phenomena is well documented[78]. Compared with bulk media, resonators offer both longer interaction timescales and higher field confinements, leading to increased nonlinear interactions. Beginning with the large-etalon cavities initially considered in the mid-60s, and then moving from millimetre- to micrometre-scale whispering gallery mode resonators, and more recently proposed wavelength-scale cavities[79], these ideas have continued to be pushed to realize higher efficiencies (lower pump powers), more compact architectures and faster operating timescales (wider bandwidths). At a finer level of detail, the physics of nonlinear processes in wavelength-scale structures is known to be well described by a small set of parameters that must be simultaneously tuned: the frequencies and decay rates (or quality factors) of each resonance and a set of nonlinear overlap integrals describing interactions between constituent electric fields (generalizing the more commonly known, phase-matching condition associated with propagating waves). Yet, while this statement of required conditions is simple, the search for well-suited structures remains both a technical and conceptual challenge, and none of the standard design principles are readily applicable. For instance, while the creation of bandgaps remains a valuable idea, they typically can only cover one of the active frequencies[75]. Furthermore, even if complete coverage could be achieved, there is no guarantee that it would result in desirable overlap characteristics. Index-guiding structures can have high modal quality factors and operate effectively over large bandwidths[80,81], but require an unideal trade-off between mode confinement and radiative losses. Plasmon polariton resonances can provide excellent confinement and field intensities, but are saddled with the unavoidable presence of material loss[82], limiting the ultimate conversion efficiency that can be achieved. Concurrently, for weak nonlinearities, the distinct resonances of any structure must be orthogonal, which weakens nonlinear overlap integrals and hence interactions[5]. In addition, devices based on $\chi^3$ third harmonic generation and higher-order processes require amplitude-dependent frequency corrections to account for cross- and self-phase modulations that prove difficult to independently tune in few-parameter designs.

The complexity implied in determining structures that simultaneously achieve these various design objectives appears ideally suited to inverse design techniques. As a conformation of this notion, preliminary findings for $\chi^2$ second harmonic generation[79,83], and $\chi^3$ difference frequency generation[84] are presented Fig. 2 (see figure for design descriptions). The three designs depicted are found to have nonlinear figures of merit between one and three orders of magnitude larger than those of previously reported designs up to the millimetre scale. Across the varied design paradigms considered (layered micropillar cavities, multiring structures, metasurfaces and so on), topology optimized structures are observed to have systematically reduced structural symmetry, large quality factors and enhanced nonlinear overlaps. For the nonlinear processes considered, the optimizations consistently find intuitive designs to be overly simplistic, in the sense that they do not make sufficient use of

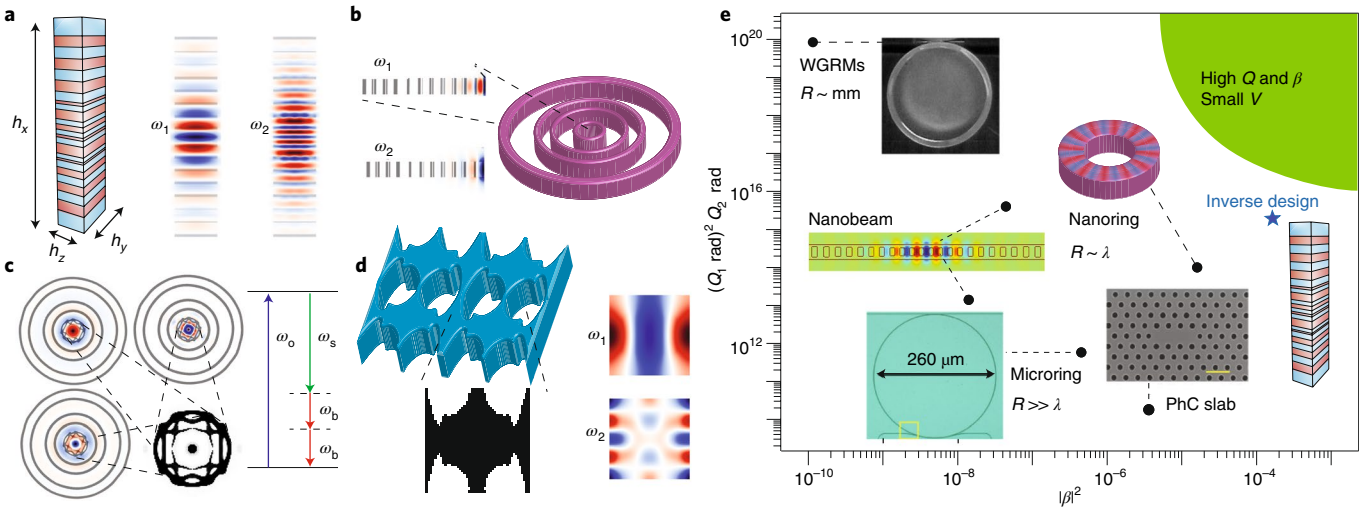



**Fig. 2 | Nonlinear optics.** Nonlinear optical interactions in micro- and nanoscale resonators are regulated by the modal quality factors and nonlinear overlap of the participating modes. Even for the simplest processes, envisioning structures that optimally control and select from this parameter space is challenging. Shown in the figure are three initial applications of inverse design towards this problem. **a**, Left, schematic of a micropost cavity consisting of aperiodically alternating AlGaAs/AlO$_x$ layers designed to enhance the efficiency of $\chi^2$ second harmonic generation[79], with $h_x$, $h_y$ and $h_z$ denoting the lengths 12.60, 5.25 and 1.26 μm respectively. Right, profiles of the overlapping modes at $\omega_1$ and $\omega_2$ participating in this process. **b**,**c**, Density topology optimized gallium arsenide multitrack ring resonators[84] clad in silica for $\chi^2$ second harmonic generation (**b**) and $\chi^3$ difference frequency generation (**c**). Included on the left of **b** is a cross-section of the contributing mode profiles. Similar profiles are shown in conjunction with the discovered structure on the left of **c** and a schematic of the difference frequency generation process is shown on the right. In this image, $\omega_o$, $\omega_s$ and $\omega_b$ stand for the emitted, signal and pump frequencies, comprising the three designed modes of the cavity. **d**, A gallium phosphide metasurface designed for $\chi^2$ second harmonic generation[83] (left), along with the profiles of the two contributing modes (right). All designs are found to have respective nonlinear figures of merit between one and three orders better than previously reported designs. **e**, The figures of merit for the design in **a**, the quality factor product $(Q^{rad}_{\omega_1})^2 Q^{rad}_{\omega_2}$ and modal overlap coefficient $|\beta|^2$, compared with contemporary platforms for second harmonic generation. Here, $\lambda$ denotes the wavelength of the fundamental frequency, $V$ the volume of the device, $R$ the radius (where applicable) and WGRMs is whispering gallery resonator modes.

interference to match the profiles of the interacting modes. (From a practical perspective, stronger overlaps are preferred to higher quality factors as the former are less sensitive to fabrication imperfections and offer greater speed due to larger bandwidths.)

The realization of wavelength-scale nonlinear devices for $\chi^2$ and $\chi^3$ harmonic generation is a necessary step towards the development of a variety of promising on-chip technologies including low threshold lasers[85], supercontinuum[86] and single photon sources[87], and non-reciprocal devices[88]. The recent results illustrated in Fig. 2 suggest the potential usefulness of inverse design in these and related problems in nonlinear optics.

**Exceptional and topological photonics.** Starting from early investigations of two-dimensional bandgaps, dispersion engineering has consistently stood as of one of the strongest motivations for applying inverse design in photonics. Spurred by the currently developing understanding of topological properties in photonic systems[89] and global properties of photonic dispersion bands, this original inspiration has re-emerged in the creation of exceptional points. As in the case of nonlinear phenomena, manipulating the nuanced role that structure plays in determining the exact characteristics of these features seems particularly aligned to inverse approaches.

Exceptional points occur in non-Hermitian problems (macroscopic electromagnetics, acoustics and so on) when two or more of the associated complex eigenvalues coalesce, causing the basis to become incomplete. The associated physical behaviour is markedly different from the more familiar accidental degeneracy encountered in Hermitian systems. First, as modes approach an exceptional point, the remaining eigenmode becomes self-orthogonal, typically quantified in terms of the diverging Petermann factor[90]. Second, the existence of an exceptional point also alters the usual analytic properties of the Green's function, exhibiting a pole of order equal to the degree of coalescence[91]. The ramifications of these altered response characteristics are linked to a long list of interesting optical phenomena, including directional transport[92], anomalous lasing[93] and enhanced spectral selectivity[94]. There are also applications to more conventional photonic processes, and exceptional points have been predicted to enhance the efficiency of spontaneous emission and frequency conversion[95].

Within the past decade, exceptional points have been designed using three primary schemes: geometries involving gain and loss in coupled resonators[96], interacting waveguides[97] and, currently, purely passive PhC lattices[98]. Indirectly, this variety indicates that the subset of systems where exceptional points can occur is quite large and that, with proper design tools, there may be enough freedom to engineer both degree and location. As a promising inroad to the additional physics and design possibilities offered by exceptional points, Fig. 3 illustrates the coalescence of three eigenvalues occurring at the Γ point of an open $C_{2v}$ PhC created by topology optimization[79], leading to an enhancement in the local density of states at certain positions in the crystal.

These early results are also relevant in the burgeoning field of topological photonics. The Dirac band structure that accompanies the creation of an exceptional point is a known precursor to media with non-trivial photonic topologies. Given the potential impact of realizing designer topological properties in practical devices, extending inverse-design methods to deal with other key stepping stones, such as chiral modes and omnidirectional Dirac cones, seems promising.

**Nanoscale optics and metasurfaces.** Over the past several years, large-scale optimization methods have begun to have a significant

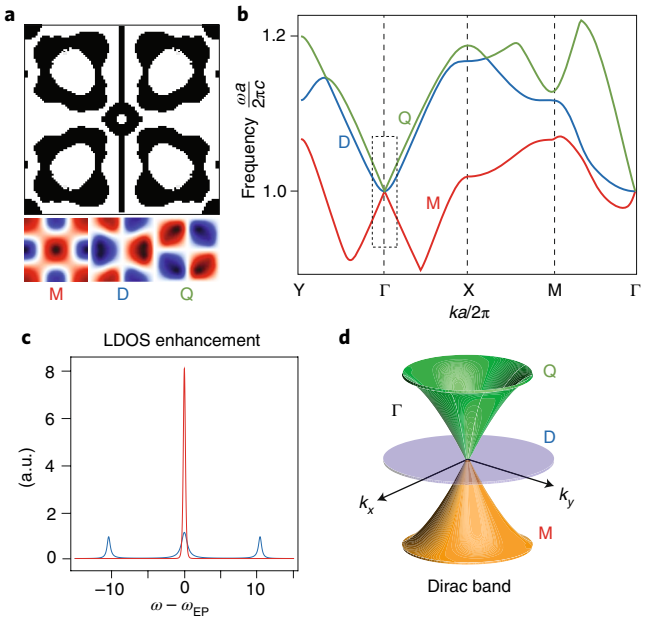



**Fig. 3 | Exceptional and topological photonics. a**,**b**, Modes (**a**, bottom) and band diagram (**b**) of a two-dimensional square lattice discovered by density topology optimization[79] (**a**, top). **c**,**d**, At the Γ point of the Brillouin zone (depicted schematically in **d**), the monopole, dipole and quadrupole modes (labelled as M, D and Q) coalesce creating an exceptional point. The resulting Dirac band structure and self-orthogonality of the modes has been shown to strongly modify the qualitative characteristics of the local density of states (LDOS) (**c**), resulting in enhanced spontaneous emission and nonlinear effects. EP, exceptional point. In this figure, $k$ is used to denote the optical wave vector, $\omega$ the angular frequency, $c$ the speed of light, $a$ the characteristic length of the lattic, and $x/y$ subscripts Cartesian coordinates.

impact on a diverse collection of problems in nano-optics and metasurfaces. Representative selections are depicted in Figs. 4 and 5.

Implementing a boundary inclusion optimization to determine the characteristics of a slab waveguide, Bhargava and Yablonovitch[99] proposed a near-field transducer for heat-assisted magnetic recording. The design has 50% less self-heating than the current standard employed in industry (Fig. 4a). Making use of the boundary element method, Lee et al.[100] investigated the optimization of electromagnetic torques arising from incident optical fields on arbitrary nanostructures. For the example triangular nanoparticle shown in Fig. 4b, the torque generated on the quadrupole mode was increased by a factor of 20. As an example of their edge element method for three-dimensional volume optimizations, Deng and Korvnik[101] applied density topology optimization towards the design of a single-material cloak[102] for a perfect spherical conductor, leading to an order of magnitude reduction in scattered power for waves propagating along the central axis of the discovered structure. Each of these examples exemplifies a technologically relevant area of photonics where complexity hampers direct application of standard design principles. Moreover, while in some cases there is guidance on expected performance from existence of fundamental limits (typically derived from physical constraints such as energy conservation or reciprocity[6]), there are yet many situations (such as in near-field or metasurface applications) where no such bounds exist or are only beginning to emerge[7,8]. In these situations, inverse-design approaches may help determine what sort of performance can be achieved.

Various applications of inverse design have also been reported for more traditional optical problems, such as diffraction, structural colour[103], coupling, polarization control and absorption enhancement in constrained volumes (Figs. 4d–f and 5). In particular, a substantial number of promising results have already been obtained in the context of metasurfaces designed via adjoint topology optimization. The works of Sell et al.[104], Callewaert et al.[105] and Shen et al.[18] have connected inverse design to the larger pursuit of flat optical systems capable of reproducing the functionality of conventional optical components. Each of the three works presents a general scheme and experimental realization for either highly efficient diffraction (Fig. 5a,b) or polarization control (Fig. 5d) that can be applied to an assortment of problems. Much as in the case of band structure, the findings of these studies open broader questions about the breadth of phase, polarization and coherence control[106] that can occur per unit thickness in a structured medium (or a single structured layer).

The metasurface concept also ties in to the pressing demand to improve solar energy capture. Two primary aspects that limit the efficiency of traditional (simple) pn-junction designs are light trapping within the volume where photovoltaic conversion occurs[107] and the width of the solar spectrum, which fundamentally limits the potential conversion efficiency of any single bandgap. Any device design offering improvement in either aspect is notable, especially if it does not impose extreme fabrication difficulties and can be implemented in silicon systems. While there have been many recent works on inverse design for solar cell applications[10,12], we highlight only three examples that address these two issues. The first issue was studied by Lee et al.[108] in 2017 with respect to the quasi-random features that can be imposed on an amorphous silicon surface by wrinkle lithography. Conducting Fourier-based inverse design (also known as concurrent design), the authors were able to enhance light-trapping by a reported factor of five over the spectral range of 400 to 1,200 nm. The second issue has been examined by Kim et al.[109] and Xiao et al.[13], resulting in experimentally realized splitters that physically separate optical and infrared wavelengths with ~70% efficiency. By partitioning the solar spectrum in this way, photovoltaics with different bandgaps can be placed in a side-by-side configuration allowing for multiple bands of high-efficiency operation.

Simultaneously, significant progress has been made on variations of the question of modal multiplexing. For dense chip-scale integration, there is an obvious need to limit the total optical footprint by handling multifrequency bands on a single waveguide. In opposition, there is also a need to access the information stored on these different frequency bands independently. To meet both goals, devices capable of high-fidelity wavelength division multiplexing are required. Adjoint optimized devices from Frellsen et al.[17] and Piggott et al.[110] for implementing this functionality in areas of a few square micrometres at telecom wavelength, with sub-5 dB transmission loss, are shown in the lower part of Fig. 1 and in Fig. 6; Shen et al.[111] and Mak et al.[112] have also reported similar findings. Also shown in Fig. 1w is a three-port power splitter designed using a fabrication tolerant algorithm and measured to have no worse than 23% transmission at any of its three output ports across its 1,400 to 1,700 nm operational bandwidth[113]. Like solar energy capture, any improvement in the components providing these necessary functionalities potentially has far-reaching industrial consequences.

Figure 4e depicts a structure enabling efficient free-space coupling of light into a waveguide from the work of Shen et al.[114]. Figure 4f illustrates a related problem in which light is efficiently coupled between a ring resonator and waveguide at multiple frequencies using wavelength-scale elements. Both problems are routinely dealt with in experimental settings. However, there are surprisingly few options to efficiently couple broadband light either into nanophotonic structures from free space, or from a waveguide into a multimode cavity beyond adiabatic tapers[115].

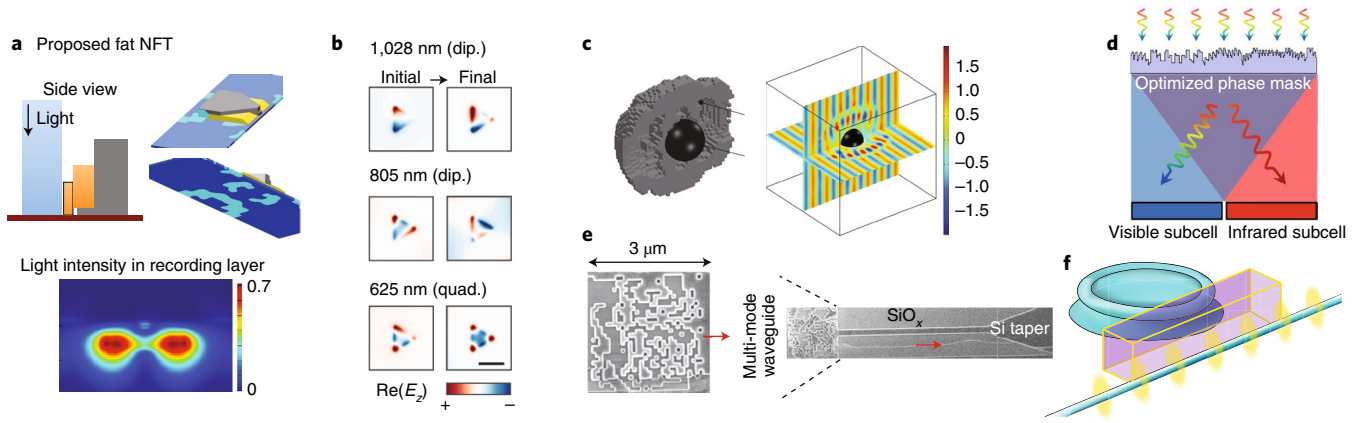



**Fig. 4 | Growth of applications.** The past three years have seen remarkable growth in the variety of systems treated with computational adjoint methods. **a**, A near-field transducer (Fat NFT) for heat-assisted magnetic recording, offering a 50% reduction in self-heating compared with industry standards[99]. **b**, Illumination of an arbitrary nanoscale structure (a triangle) with a structured beam, before and after employing inverse design to increase optical torque. Inverse optimization of the field is observed to produce a 20-fold enhancement in the applied angular force[100]. In this panel, dip. and quad. refer to dipole and quadrapole modes, while $E_z$ stands for the magnitude of the electric field in the perpendicular, *z*, direction. **c**, Left, a three-dimensional electromagnetic cloak, leading to order of magnitude reduction in total scattering[101]. Right, the nearly unaltered electric field profile for a wave incident along the central axis of the cloak. **d**, A schematic of an experimentally realized optimized structure for spectral splitting that achieves 69.5% separation of the optical and infrared spectra, opening new directions for multi-bandgap photovoltaics[13]. **e**, A compact integrated silicon/silicon-dioxide design for implementing free-space-to-waveguide coupling[114] showing a larger bandwidth than traditional approaches. **f**, A conceived application of inverse-design structures to optimize multi-mode coupling between a ring resonator, or more generally any cavity, and a multi-mode waveguide. Panels reproduced from: **a**, ref. [99], IEEE; **c**, ref. [101], The Royal Society; **d**, ref. [13], American Chemical Society. Panels adapted from: **b**, ref. [100], OSA; **e**, ref. [114], OSA.

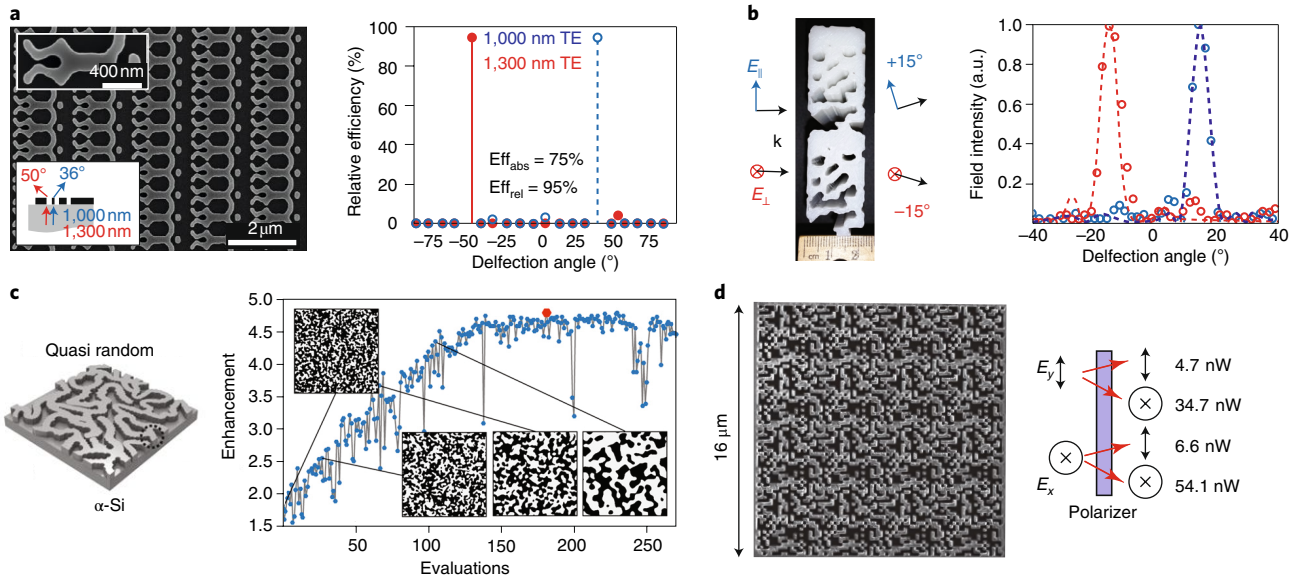



**Fig. 5 | Metasurface photonics.** The figure highlights four recent applications of inverse design to metaphotonics. **a**, A metagrating capable of angularly separating 1,000 nm and 1,300 nm transverse-electric (TE) polarized light with 75% absolute efficiency ($Eff_{abs}$) and 95% relative efficiency ($Eff_{rel}$) by Sell et al.[104]. **b**, A three-dimensional polarization (electric field *E* in and out of plane) splitter designed for microwave applications (~26–33 GHz) by Callewaert et al.[105]. **c**, Left, schematic optimized quasi-random amorphous silicon structure for enhancing absorption of the solar spectrum created by Lee et al.[108]. Right, relative enhancement of light trapping by the structure as function of the optimization iteration. The insets show the corresponding structure at various steps of the optimization over a 3 μm by 3 μm area. **d**, A topology optimized polarizer with ~90% conversion efficiency conceived by Shen et al.[18]. Panels adapted from: **a**, ref. [104], American Chemical Society; **b**, ref. [105], Springer Nature Ltd; **c**, ref. [108], PNAS; **d**, ref. [18], OSA.

**Experimental challenges.** Since 2004, a variety of designs have been experimentally demonstrated to illustrate the practical viability of inverse methods. Ranging from bends and splitters for PhC waveguides[24,41], to passive components for silicon photonic circuits[17,19,110,113,116] and metasurfaces[104,105], operational devices exist in essentially every major domain of photonics where inverse design has been applied. Yet, to date, none of these structures have found broad industrial application. The primary cause of this incongruity is simple: nearly every device has been fabricated using electron-beam lithography due to the small features that occur naturally in

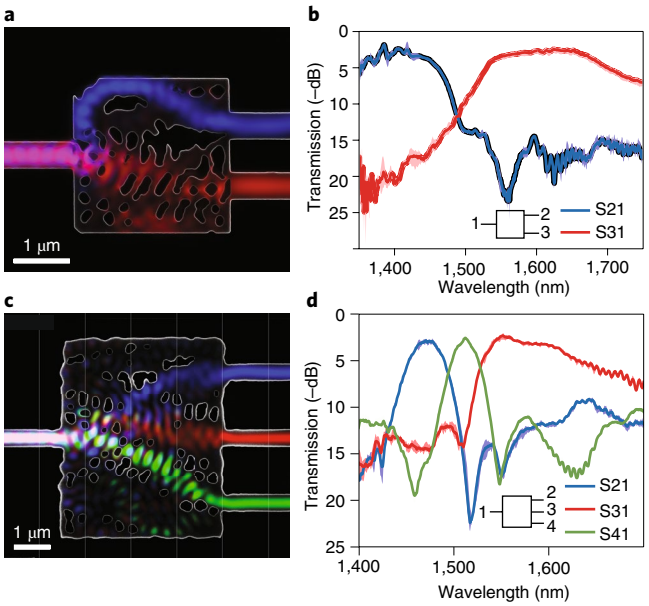



**Fig. 6 | Experimental inverse design. a**,**c**, Two scanning electron microscopy images overlaid with accompanying field profiles for broadband two-channel (**a**) and three-channel (**c**) wavelength splitters, from Piggott et al.[110] and Su et al.[19]. The two-band splitter is designed to separate 1,300 nm (blue) and 1,550 nm (red) signals. The three-band splitter designed to separate 1,500 nm (blue), 1,540 nm (green) and 1,580 nm (red) signals. **b**,**d**, Measured transmission spectra, validating the functionality of these devices, respectively.

current inverse algorithms. For industrial applications, limiting fabrication time, and hence cost, requires compatibility with photolithography; and while explicit constraints imposing a minimal feature size can be implemented in one-dimensional designs with relative simplicity[117], such approaches are considerably more difficult in higher dimensions.

A conceptually simple solution to this challenge is to subdivide the design region into pixels that are larger than the smallest achievable feature size. After removing any intermediate grey structures, the design is then assured to be fabricable[111,112]. However, most current integrated photonic designs make extensive use of smooth curves, such as in waveguide bends, which cannot be represented using this pixel-based paradigm. A variety of approaches for designing easily fabricable structures with smooth curves have thus been proposed. For topology algorithms, this is accomplished by using convolution filters to smear out small features, and image dilation and erosion operations to mimic fabrication imperfections[17,104,118]. For boundary parameterized optimizations, simultaneously limiting the minimum radius of curvature[119], and eliminating any gaps or bridges narrower than a threshold width[19,113], has been shown to increase fabrication tolerance. These methods have been experimentally tested for electron-beam lithography, and simulations indicated that they should also work reasonably well for photolithography when using optical proximity correction. Still, as promising as these results are, they are not robust to process variations in photolithography, such as defocusing and dosage errors, and finding techniques to cope with these additional complications remains an open problem. A potential solution is to incorporate robustness to expected process variations into the optimization problem[120,121], as is currently done in optical proximity correction[122].

The physical size of practical aperiodic devices that can be currently treated with inverse-design methods is limited by the computational cost of the fully vectoral three-dimensional simulations needed to accurately model their performance. Dozens to hundreds of simulations are required to design a single device, which becomes prohibitively expensive as the design domain expands. This limits the type of questions that can be meaningfully treated, and makes it difficult to inverse design interfaces with large structures, such as single-mode optical fibres. In this light, further improvements in the computational machinery (such as iterative solvers) that underpin current inverse methods have the potential to vastly expand the boundaries of application.

## Summary outlook

Applications have always served as the spark for progress in inverse design, and from this fact alone the outlook for these principles in nanophotonics is positive. There are both well-formulated problems in areas such as chip-scale integration and cavity design that remain open as well as areas of application such as energy capture and nonlinear device design where only promising preliminary work has been done. Beyond the areas we have outlined, inverse-design principles appear to offer a new perspective for understanding fluctuation physics and near-field optics. Although currently limited to one dimension[15], the application of topology optimization to find optimally efficient heat-transfer systems[99] promises to have practical and theoretical impact in building further understanding of the practical limits of heat transfer[7]. Extending inverse design to active devices such as modulators and lasers, which are the performance limiting components of many optical systems, would also be extremely useful.

Several key improvements would enable the widespread use of inverse-design methods in practical applications. First and foremost is improving the robustness of design methods to handle process variations in photolithography, which would enable high-throughput fabrication. Parallel to this computational focused tract, advancements in nanoscale lithography[123] appear poised to enlarge the landscape of fabricable structures to include a larger subset of the intricate multiscale features and permittivity gradients[124] that ubiquitously appear in inverse algorithms. Inverse-design methods can, at least in principle, explore the full space of fabricable devices. This capability opens the possibility of reasonably asking what the maximum theoretical performance of an optical device is. More specifically, for a given a functionality, design area, minimum feature size and selection of materials, what is the ultimate achievable performance of an optical device? Establishing such theoretical bounds on the performance would help guide future work in all of photonics. Finally, improvements to the underlying simulations and optimization algorithms could enable design of larger devices, greatly improving the breadth and scope of problems that can be tackled by inverse design. Along these lines, several recent works have begun exploring applications of machine learning in nanophotonics[125], paving the way, potentially, for improvements in fast iterative Maxwell solvers. In the enduring quest for optimal photonic designs, the widespread integration of inverse-design tools seems not only sensible but unavoidable.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41566-018-0246-9.

Received: 6 February 2018; Accepted: 2 August 2018;
Published online: 26 October 2018

## References

1. Koenderink, A. F., Alù, A. & Polman, A. Nanophotonics: shrinking light-based technology. *Science* **348**, 516–521 (2015).
2. Baba, T. Slow light in photonic crystals. *Nat. Photon.* **2**, 465–473 (2008).

3. Caldwell, J. D. et al. Sub-diffractional volume-confined polaritons in the natural hyperbolic material hexagonal boron nitride. *Nat. Commun.* **5**, 5221 (2014).
4. Spillane, S. et al. Ultrahigh-*Q* toroidal microresonators for cavity quantum electrodynamics. *Phys. Rev. A* **71**, 013817 (2005).
5. Hashemi, H., Rodriguez, A. W., Joannopoulos, J., Soljačić, M. & Johnson, S. G. Nonlinear harmonic generation and devices in doubly resonant Kerr cavities. *Phys. Rev. A* **79**, 013812 (2009).
6. Yu, Z. Fundamental limit of nanophotonic light trapping in solar cells. *Proc. Natl Acad. Sci. USA* **107**, 17491–17496 (2010).
7. Miller, O. D., Johnson, S. G. & Rodriguez, A. W. Shape-independent limits to near-field radiative heat transfer. *Phys. Rev. Lett.* **115**, 204302 (2015).
8. Arabi, A. & Faraon, A. Fundamental limits of ultrathin metasurfaces. *Sci. Rep.* **7**, 43722 (2017).
9. Alaeian, H., Atre, A. C. & Dionne, J. A. Optimized light absorption in Si wire array solar cells. *J. Opt.* **14**, 024006 (2012).
10. Ganapati, V., Miller, O. D. & Yablonovitch, E. Light trapping textures designed by electromagnetic optimization for subwavelength thick solar cells. *IEEE J. Photovolt.* **4**, 175–182 (2014).
11. Wang, P. & Menon, R. Optimization of periodic nanostructures for enhanced light-trapping in ultra-thin photovoltaics. *Opt. Express* **21**, 6274–6285 (2013).
12. Dühring, M. B. & Sigmund, O. Optimization of extraordinary optical absorption in plasmonic and dielectric structures. *J. Opt. Soc. Am. B* **30**, 1154–1160 (2013).
13. Xiao, T. P. et al. Diffractive spectral-splitting optical element designed by adjoint-based electromagnetic optimization and fabricated by femtosecond 3D direct laser writing. *ACS Photon.* **3**, 886–894 (2016).
14. Ilic, O. et al. Tailoring high-temperature radiation and the resurrection of the incandescent source. *Nat. Nanotech.* **11**, 320–324 (2016).
15. Jin, W., Messina, R. & Rodriguez, A. W. Overcoming limits to near-field radiative heat transfer in uniform planar media through multilayer optimization. *Opt. Express* **25**, 14746–14759 (2017).
16. Liu, D., Gabrielli, L. H., Lipson, M. & Johnson, S. G. Transformation inverse design. *Opt. Express* **21**, 14223–14243 (2013).
17. Frellsen, L. F., Ding, Y., Sigmund, O. & Frandsen, L. H. Topology optimized mode multiplexing in silicon-on-insulator photonic wire waveguides. *Opt. Express* **24**, 16866–16873 (2016).
18. Shen, B., Wang, P., Polson, R. & Menon, R. Ultra-high-efficiency metamaterial polarizer. *Optica* **1**, 356–360 (2014).
19. Su, L., Piggott, A. Y., Sapra, N. V., Petykiewicz, J. & Vuckovic, J. Inverse design and demonstration of a compact on-chip narrowband three-channel wavelength demultiplexer. *ACS Photon* **5**, 301–305 (2017).
20. Chadan, K. & Sabatier, P. C. *Inverse Problems in Quantum Scattering Theory* (Springer Science & Business Media, Berlin, Heidelberg, New York, 2012).
21. Bendsøe, M. P. & Sigmund, O. *Topology Optimization: Theory, Methods and Applications* (Springer, Berlin, Heidelberg, New York, 2003).
22. Georgieva, N. K., Glavic, S., Bakr, M. H. & Bandler, J. W. Feasible adjoint sensitivity technique for EM design optimization. *IEEE Trans. Microw. Theory Tech.* **50**, 2751–2758 (2002).
23. Pratt, R. G. & Worthington, M. Inverse theory applied to multi-source cross-hole tomography. Part 1: acoustic wave-equation method. *Geophys. Prospect.* **38**, 287–310 (1990).
24. Jensen, J. S. & Sigmund, O. Topology optimization for nano-photonics. *Laser Photon. Rev.* **5**, 308–321 (2011).
25. Sigmund, O. On the usefulness of non-gradient approaches in topology optimization. *Struct. Multidiscipl. Optim.* **43**, 589–596 (2011).
26. Spühler, M. M. et al. A very short planar silica spot-size converter using a nonperiodic segmented waveguide. *J. Light. Technol.* **16**, 1680–1685 (1998).
27. Dobson, D. C. & Cox, S. J. Maximizing band gaps in two-dimensional photonic crystals. *SIAM J. Appl. Math.* **59**, 2108–2120 (1999).
28. Back, T., Hammel, U. & Schwefel, H.-P. Evolutionary computation: comments on the history and current state. *IEEE Trans. Evol. Comput.* **1**, 3–17 (1997).
29. Fu, M. C., Glover, F. W. & April, J. Simulation optimization: a review, new developments, and applications. In *2005 Proc. Winter Simulation Conference* 83–95 (IEEE, 2005).
30. Boyd, S. & Vandenberghe, L. *Convex Optimization* (Cambridge Univ. Press, Cambridge, 2004).
31. Baumert, T., Brixner, T., Seyfried, V., Strehle, M. & Gerber, G. Femtosecond pulse shaping by an evolutionary algorithm with feedback. *Appl. Phys. B* **65**, 779–782 (1997).
32. Doosje, M., Hoenders, B. J. & Knoester, J. Photonic bandgap optimization in inverted fcc photonic crystals. *JOSA B* **17**, 600–606 (2000).
33. Cox, S. J. & Dobson, D. C. Band structure optimization of two-dimensional photonic crystals in H-polarization. *J. Comput. Phys.* **158**, 214–224 (2000).
34. Felici, T. & Engl, H. W. On shape optimization of optical waveguides using inverse problem techniques. *Inverse Probl.* **17**, 1141–1162 (2001).
35. Geremia, J., Williams, J. & Mabuchi, H. Inverse-problem approach to designing photonic crystals for cavity QED experiments. *Phys. Rev. E* **66**, 066606 (2002).
36. Jiang, J., Cai, J., Nordin, G. P. & Li, L. Parallel microgenetic algorithm design for photonic crystal and waveguide structures. *Opt. Lett.* **28**, 2381–2383 (2003).
37. Kiziltas, G., Psychoudakis, D., Volakis, J. L. & Kikuchi, N. Topology design optimization of dielectric substrates for bandwidth improvement of a patch antenna. *IEEE Trans. Antennas Propag.* **51**, 2732–2743 (2003).
38. Erni, D. et al. Application of evolutionary optimization algorithms in computational optics. *ACES* **15**, 43–60 (2000).
39. Veronis, G., Dutton, R. W. & Fan, S. Method for sensitivity analysis of photonic crystal devices. *Opt. Lett.* **29**, 2288–2290 (2004).
40. Jiao, Y., Fan, S. & MillYer, D. A. Systematic photonic crystal device design: global and local optimization and sensitivity analysis. *IEEE J. Quantum Electron.* **42**, 266–279 (2006).
41. Borel, P. I. et al. Topology optimization and fabrication of photonic crystal structures. *Opt. Express* **12**, 1996–2001 (2004).
42. Jensen, J. S. & Sigmund, O. Systematic design of photonic crystal structures using topology optimization: low-loss waveguide bends. *Appl. Phys. Lett.* **84**, 2022–2024 (2004).
43. Jensen, J. S. & Sigmund, O. Topology optimization of photonic crystal structures: a high-bandwidth low-loss T-junction waveguide. *JOSA B* **22**, 1191–1198 (2005).
44. Kao, C.-Y., Osher, S. & Yablonovitch, E. Maximizing band gaps in two-dimensional photonic crystals by using level set methods. *Appl. Phys. B* **81**, 235–244 (2005).
45. Burger, M. A framework for the construction of level set methods for shape optimization and reconstruction. *Interfaces Free Bound.* **5**, 301–329 (2003).
46. Burger, M., Osher, S. J. & Yablonovitch, E. Inverse problem techniques for the design of photonic crystals. *IEICE Trans. Electron.* **87**, 258–265 (2004).
47. Gerken, M. & Miller, D. A. Multilayer thin-film structures with high spatial dispersion. *Appl. Opt.* **42**, 1330–1345 (2003).
48. Håkansson, A. & Sánchez-Dehesa, J. Inverse designed photonic crystal de-multiplex waveguide coupler. *Opt. Express* **13**, 5440–5449 (2005).
49. Sanchis, L., Håkansson, A., López-Zanón, D., Bravo-Abad, J. & Sánchez-Dehesa, J. Integrated optical devices design by genetic algorithm. *Appl. Phys. Lett.* **84**, 4460–4462 (2004).
50. Preble, S., Lipson, M. & Lipson, H. Two-dimensional photonic crystals designed by evolutionary algorithms. *Appl. Phys. Lett.* **86**, 061111 (2005).
51. Burger, M. & Osher, S. J. A survey on level set methods for inverse problems and optimal design. *Eur. J. Appl. Math.* **16**, 263–301 (2005).
52. Novotny, A. A. & Sokołowski, J. *Topological Derivatives in Shape Optimization* (Springer Science & Business Media, Heidelberg, New York, 2012).
53. van Dijk, N. P., Maute, K., Langelaar, M. & Van Keulen, F. Level-set methods for structural topology optimization: a review. *Struct. Multidiscipl. Optim.* **48**, 437–472 (2013).
54. Tortorelli, D. A. & Michaleris, P. Design sensitivity analysis: overview and review. *Inverse Probl. Eng.* **1**, 71–105 (1994).
55. Miller, O. D. et al. Fundamental limits to extinction by metallic nanoparticles. *Phys. Rev. Lett.* **112**, 123903 (2014).
56. Giles, M. B. & Pierce, N. A. An introduction to the adjoint approach to design. *Flow Turbul. Combust.* **65**, 393–415 (2000).
57. Riishede, J. & Sigmund, O. Inverse design of dispersion compensating optical fiber using topology optimization. *JOSA B* **25**, 88–97 (2008).
58. Dobson, D. C. & Simeonova, L. B. Optimization of periodic composite structures for sub-wavelength focusing. *Appl. Math. Optim.* **60**, 133–150 (2009).
59. Borel, P. I. et al. Imprinted silicon-based nanophotonics. *Opt. Express* **15**, 1261–1266 (2007).
60. Elesin, Y., Lazarov, B. S., Jensen, J. S. & Sigmund, O. Design of robust and efficient photonic switches using topology optimization. *Photon. Nanostruct.* **10**, 153–165 (2012).
61. Tsuji, Y. & Hirayama, K. Design of optical circuit devices using topology optimization method with function-expansion-based refractive index distribution. *IEEE Photon. Technol. Lett.* **20**, 982–984 (2008).
62. Sigmund, O. & Petersson, J. Numerical instabilities in topology optimization: a survey on procedures dealing with checkerboards, mesh-dependencies and local minima. *Struct. Multidiscipl. Optim.* **16**, 68–75 (1998).
63. Wang, F., Jensen, J. S. & Sigmund, O. High-performance slow light photonic crystal waveguides with topology optimized or circular-hole based material layouts. *Photon. Nanostruct.* **10**, 378–388 (2012).
64. Sigmund, O. Manufacturing tolerant topology optimization. *Acta Mech. Sin.* **25**, 227–239 (2009).
65. Oskooi, A. et al. Robust optimization of adiabatic tapers for coupling to slow-light photonic-crystal waveguides. *Opt. Express* **20**, 21558–21575 (2012).

66. Frei, W., Tortorelli, D. & Johnson, H. Geometry projection method for optimizing photonic nanostructures. *Opt. Lett.* **32**, 77–79 (2007).
67. Frei, W. R., Johnson, H. & Choquette, K. D. Optimization of a single defect photonic crystal laser cavity. *J. Appl. Phys.* **103**, 033102 (2008).
68. Lu, J., Boyd, S. & Vučković, J. Inverse design of a three-dimensional nanophotonic resonator. *Opt. Express* **19**, 10563–10570 (2011).
69. Boyd, S., Parikh, N., Chu, E., Peleato, B. & Eckstein, J. Distributed optimization and statistical learning via the alternating direction method of multipliers Found. Trend. *Mach. Learn.* **3**, 1–122 (2011).
70. Englund, D., Fushman, I. & Vucković, J. General recipe for designing photonic crystal cavities. *Opt. Express* **13**, 5961–5975 (2005).
71. Men, H., Nguyen, N. C., Freund, R. M., Parrilo, P. A. & Peraire, J. Bandgap optimization of two-dimensional photonic crystals using semidefinite programming and subspace methods. *J. Comput. Phys.* **229**, 3706–3725 (2010).
72. Liang, X. & Johnson, S. G. Formulation for scalable optimization of microcavities via the frequency-averaged local density of states. *Opt. Express* **21**, 30812–30841 (2013).
73. Elesin, Y., Lazarov, B. S., Jensen, J. S. & Sigmund, O. Time domain topology optimization of 3D nanophotonic devices. *Photon. Nanostruct.* **12**, 23–33 (2014).
74. Chen, H., Chan, C. T. & Sheng, P. Transformation optics and metamaterials. *Nat. Mater.* **9**, 387–396 (2010).
75. Men, H., Lee, K. Y., Freund, R. M., Peraire, J. & Johnson, S. G. Robust topology optimization of three-dimensional photonic-crystal band-gap structures. *Opt. Express* **22**, 22632–22648 (2014).
76. Maldovan, M. & Thomas, E. L. Diamond-structured photonic crystals. *Nat. Mater.* **3**, 593–600 (2004).
77. Sigmund, O. & Hougaard, K. Geometric properties of optimal photonic crystals. *Phys. Rev. Lett.* **100**, 153904 (2008).
78. Ou, Z. & Kimble, H. J. Enhanced conversion efficiency for harmonic generation with double resonance. *Opt. Lett.* **18**, 1053–1055 (1993).
79. Lin, Z., Liang, X., Lončar, M., Johnson, S. G. & Rodriguez, A. W. Cavity-enhanced second-harmonic generation via nonlinear-overlap optimization. *Optica* **3**, 233–238 (2016).
80. Fürst, J. et al. Naturally phase-matched second-harmonic generation in a whispering-gallery-mode resonator. *Phys. Rev. Lett.* **104**, 153901 (2010).
81. Bi, Z.-F. et al. High-efficiency second-harmonic generation in doubly-resonant $\chi^2$ microring resonators. *Opt. Express* **20**, 7526–7543 (2012).
82. Khurgin, J. B. How to deal with the loss in plasmonics and metamaterials. *Nat. Nanotech.* **10**, 2–6 (2015).
83. Sitawarin, C., Jin, W., Lin, Z. & Rodriguez, A. W. Inverse-designed photonic fibers and metasurfaces for nonlinear frequency conversion. *Photon. Res.* **6**, B82–B89 (2018).
84. Lin, Z., Lončar, M. & Rodriguez, A. W. Topology optimization of multi-track ring resonators and 2D microcavities for nonlinear frequency conversion. *Opt. Lett.* **42**, 2818–2821 (2017).
85. Takahashi, Y. et al. A micrometre-scale Raman silicon laser with a microwatt threshold. *Nature* **498**, 470–474 (2013).
86. Halir, R. et al. Ultrabroadband supercontinuum generation in a CMOS-compatible platform. *Opt. Lett.* **37**, 1685–1687 (2012).
87. Pelton, M. et al. Efficient source of single photons: a single quantum dot in a micropost microcavity. *Phys. Rev. Lett.* **89**, 233602 (2002).
88. Bi, L. et al. On-chip optical isolation in monolithically integrated non-reciprocal optical resonators. *Nat. Photon.* **5**, 758–762 (2011).
89. Lu, L., Joannopoulos, J. D. & Soljačić, M. Topological photonics. *Nat. Photon.* **8**, 821–829 (2014).
90. Heiss, W. The physics of exceptional points. *J. Phys. Math. Theor.* **45**, 444016 (2012).
91. Pick, A. et al. General theory of spontaneous emission near exceptional points. *Opt. Express* **25**, 12325–12348 (2017).
92. Regensburger, A. et al. Parity-time synthetic photonic lattices. *Nature* **488**, 167–171 (2012).
93. Peng, B. et al. Loss-induced suppression and revival of lasing. *Science* **346**, 328–332 (2014).
94. Hodaei, H. et al. Enhanced sensitivity at higher-order exceptional points. *Nature* **548**, 187–191 (2017).
95. Pick, A., Lin, Z., Jin, W. & Rodriguez, A. W. Enhanced nonlinear frequency conversion and Purcell enhancement at exceptional points. *Phys. Rev. B* **96**, 224303 (2017).
96. Peng, B. et al. Parity-time-symmetric whispering-gallery microcavities. *Nat. Phys.* **10**, 394–398 (2014).
97. Rüter, C. E. et al. Observation of parity–time symmetry in optics. *Nat. Phys.* **6**, 192–195 (2010).
98. Zhen, B. et al. Spawning rings of exceptional points out of Dirac cones. *Nature* **525**, 354–358 (2015)
99. Bhargava, S. & Yablonovitch, E. Lowering HAMR near-field transducer temperature via inverse electromagnetic design. *IEEE Trans. Magn.* **51**, 1–7 (2015).
100. Lee, Y. E., Miller, O. D., Reid, M. H., Johnson, S. G. & Fang, N. X. Computational inverse design of non-intuitive illumination patterns to maximize optical force or torque. *Opt. Express* **25**, 6757–6766 (2017).
101. Deng, Y. & Korvink, J. G. Topology optimization for three-dimensional electromagnetic waves using an edge element-based finite-element method. *Proc. R. Soc. A* **472**, 20150835 (2016).
102. Andkjær, J., Asger Mortensen, N. & Sigmund, O. Towards all-dielectric, polarization-independent optical cloaks. *Appl. Phys. Lett.* **100**, 101106 (2012).
103. Andkjær, J., Johansen, V. E., Friis, K. S. & Sigmund, O. Inverse design of nanostructured surfaces for color effects. *JOSA B* **31**, 164–174 (2014).
104. Sell, D., Yang, J., Doshay, S., Yang, R. & Fan, J. A. Large-angle, multifunctional metagratings based on freeform multimode geometries. *Nano Lett.* **17**, 3752–3757 (2017).
105. Callewaert, F., Velev, V., Kumar, P., Sahakian, A. & Aydin, K. Inverse-designed broadband all-dielectric electromagnetic metadevices. *Sci. Rep.* **8**, 1358 (2018).
106. Okoro, C., Kondakci, H. E., Abouraddy, A. F. & Toussaint, K. C. Demonstration of an optical-coherence converter. *Optica* **4**, 1052–1058 (2017).
107. Garnett, E. & Yang, P. Light trapping in silicon nanowire solar cells. *Nano Lett.* **10**, 1082–1087 (2010).
108. Lee, W.-K. et al. Concurrent design of quasi-random photonic nanostructures. *Proc. Natl Acad. Sci. USA* **114**, 8734–8739 (2017).
109. Kim, G., Dominguez-Caballero, J. A., Lee, H., Friedman, D. J. & Menon, R. Increased photovoltaic power output via diffractive spectrum separation. *Phys. Rev. Lett.* **110**, 123901 (2013).
110. Piggott, A. Y. et al. Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer. *Nat. Photon.* **9**, 374–377 (2015).
111. Shen, B., Wang, P., Polson, R. & Menon, R. An integrated-nanophotonics polarization beamsplitter with 2.4 × 2.4 $\mu^m 2$ footprint. *Nat. Photon.* **9**, 378–382 (2015).
112. Mak, J. C., Sideris, C., Jeong, J., Hajimiri, A. & Poon, J. K. Binary particle swarm optimized 2 × 2 power splitters in a standard foundry silicon photonic platform. *Opt. Lett.* **41**, 3868–3871 (2016).
113. Piggott, A. Y., Petykiewicz, J., Su, L. & Vučković, J. Fabrication-constrained nanophotonic inverse design. *Sci. Rep.* **7**, 1786 (2017).
114. Shen, B., Wang, P., Polson, R. & Menon, R. Integrated metamaterials for efficient and compact free-space-to-waveguide coupling. *Opt. Express* **22**, 27175–27182 (2014).
115. Niederberger, A. C., Fattal, D. A., Gauger, N. R., Fan, S. & Beausoleil, R. G. Sensitivity analysis and optimization of sub-wavelength optical gratings using adjoints. *Opt. Express* **22**, 12971–12981 (2014).
116. Frandsen, L. H. & Sigmund, O. Inverse design engineering of all-silicon polarization beam splitters. *Proc. SPIE* 9756, 97560Y (2016).
117. Michaels, A. & Yablonovitch, E. Inverse design of near unity efficiency perfectly vertical grating couplers. *Opt. Express* **26**, 4766–4779 (2018).
118. Lazarov, B. S., Wang, F. & Sigmund, O. Length scale and manufacturability in density-based topology optimization. *Arch. Appl. Mech.* **86**, 189–218 (2016).
119. Lalau-Keraly, C. M., Bhargava, S., Miller, O. D. & Yablonovitch, E. Adjoint shape optimization applied to electromagnetic design. *Opt. Express* **21**, 21693–21701 (2013).
120. Zhou, M., Lazarov, B. S. & Sigmund, O. Topology optimization for optical projection lithography with manufacturing uncertainties. *Appl. Opt.* **53**, 2720–2729 (2014).
121. Menon, R., Rogge, P. & Tsai, H.-Y. Design of diffractive lenses that generate optical nulls without phase singularities. *J. Opt. Soc. Am. A* **26**, 297–304 (2009).
122. Rosenbluth, A. E. et al. Optimum mask and source patterns to print a given shape. *J. MicroNanolithogr. MEMS MOEMS* **1**, 13–31 (2002).
123. Kasahara, K. et al. Recent progress in nanoparticle photoresists development for EUV lithography. *Proc. SPIE* 9776, (977604 (2016).
124. Urness, A. C., Anderson, K., Ye, C., Wilson, W. L. & McLeod, R. R. Arbitrary GRIN component fabrication in optically driven diffusive photopolymers. *Opt. Express* **23**, 264–273 (2015).
125. Zibar, D., Wymeersch, H. & Lyubomirsky, I. Machine learning under the spotlight. *Nat. Photon.* **11**, 749–751 (2017).
126. Painter, O. et al. Two-dimensional photonic band-gap defect mode laser. *Science* **284**, 1819–1821 (1999).
127. Knight, J. C. Photonic crystal fibres. *Nature* **424**, 847–851 (2003).
128. Xu, Q., Fattal, D. & Beausoleil, R. G. Silicon microring resonators with 1.5-μm radius. *Opt. Express* **16**, 4309–4315 (2008).
129. Eichenfield, M., Chan, J., Camacho, R. M., Vahala, K. J. & Painter, O. Optomechanical crystals. *Nature* **462**, 78–82 (2009).
130. Liu, N., Mesch, M., Weiss, T., Hentschel, M. & Giessen, H. Infrared perfect absorber and its application as plasmonic sensor. *Nano Lett.* **10**, 2342–2348 (2010).

131. Otomori, M., Yamada, T., Izui, K., Nishiwaki, S. & Andkjær, J. Topology optimization of hyperbolic metamaterials for an optical hyperlens. *Struct. Multidiscipl. Optim.* **55**, 913–923 (2017).
132. Yu, Z., Cui, H. & Sun, X. Genetically optimized on-chip wideband ultracompact reflectors and Fabry–Perot cavities. *Photon. Res.* **5**, B15–B19 (2017).

## Acknowledgements

This work was supported by the National Science Foundation under Grant No. DMR-1454836, Grant No. DMR 1420541, and Award EFMA-1640986; and the National Sciences and Engineering Research Council of Canada under PDF-502958-2017. All authors acknowledge helpful comments made during the preparation of this manuscript by D. A. B. Miller.

## Additional information



**Correspondence** should be addressed to A.W.R.