Large-scale parameterized metasurface design using adjoint optimization

Mahdad Mansouree, Andrew McClung, Sarath Samudrala, and Amir Arbabi*

Department of Electrical and Computer Engineering, University of Massachusetts Amherst,
151 Holdsworth Way, Amherst, MA 01003, USA

E-mail: arbabi@umass.edu

Abstract

Optical metasurfaces are planar arrangements of subwavelength meta-atoms that implement a wide range of transformations on incident light. The design of efficient metasurfaces requires that the responses of and interactions among meta-atoms are accurately modeled. Conventionally, each meta-atom’s response is approximated by that of a meta-atom located in a periodic array. Although this approximation is accurate for metastructures with slowly varying meta-atoms, it does not accurately model the complex interactions among meta-atoms in more rapidly varying metasurfaces. Optimization-based design techniques that rely on full-wave simulations mitigate this problem but thus far have been mostly applied to topology optimization of small metasurfaces. Here, we describe an adjoint-optimization-based design technique that uses parameterized meta-atoms. Our technique has a lower computational cost than topology optimization approaches, enabling the design of large-scale metasurfaces that can be readily fabricated. As proof of concept, we present the design and experimental demonstration of high numerical aperture metalenses with significantly higher efficiencies than their conventionally-designed counterparts.
Keywords

Adjoint technique, Optimization, Metasurface, Metalens

Introduction

Optical metasurfaces are arrangements of subwavelength meta-atoms that scatter optical waves and generate desirable wavefront, amplitude, and polarization distributions. Metasurface-based designs of numerous optical components have been demonstrated, including lenses, blazed gratings, and holograms. Their planar form factor and the potential for low-cost manufacture have spurred the recent development of complex optical systems made of multiple metasurfaces, or metasystems, such as miniaturized cameras, spectrometers and hyper-spectral imaging systems. However, cascading multiple metasurfaces quickly increases a system’s optical loss and high-performance metasystems require high-efficiency metasurface components.

Currently, most metasurfaces are designed using a unit-cell-based approach in which each meta-atom is simulated as a part of a periodic array. This approach is computationally inexpensive, readily scalable to arbitrarily large structures, and produces designs that can be fabricated easily. However, two implicit assumptions in the unit-cell approach can lead to inefficient metasurface designs: First, the response of a meta-atom with dissimilar neighbors differs from the response of the same meta-atom in an array with identical elements. The ‘local periodicity’ approximation breaks down in structures with rapidly varying meta-atoms, such as high numerical aperture (NA) lenses. This approximation also has reduced accuracy in structures comprising meta-atoms with lower refractive index, in which the response of a meta-atom is more strongly affected by variations of its neighbors. Second, the response map used in unit-cell-based methods records only the normally transmitted response for normally incident light. In an actual metasurface, the transmission angle varies, and hence the true response would deviate from the response map.
can lead to significant mismatches between expected and actual responses of a meta-atom in a metasurface. The effect of such mismatch can be seen in reduction of the efficiency of the device, however, it is hard to exactly distinguish the contribution of each assumption without analytical models.

In the absence of simple and accurate models that capture the individual and collective behaviors of meta-atoms, optimization-based design (i.e., inverse design) is a practical alternative. Optical structures have been designed using a variety of optimization-based approaches. A comprehensive review of these methods is presented in ref. 18. Heuristic optimization algorithms based on random explorations of the optimization space (e.g., particle swarm optimization or genetic algorithms) have been used to design diffraction grating filters, polarization beam splitters and other small structures. Heuristic optimization is well-suited to problems with a small number of degrees of freedom but inefficient for structures with larger design spaces.

In most problems, the gradient of the design objective, necessary for gradient-based algorithms, can be determined using an adjoint technique. Adjoint-based algorithms are suitable for optimization spaces with high dimensionality, and have been used to design high-performance optical devices, including grating couplers, polarization beam splitters, photonic crystals, metagratings and metasurfaces. Adjoint-based optimization is frequently applied to topology optimization problems, in which a structure or its meta-atoms are defined by a large number of pixels or sets of curvilinear patterns. This can lead to efficient metasurface designs, but because even deeply subwavelength changes in meta-atom geometries can significantly alter the scattering response of a design (see Supplementary Note 1 and Fig. S1), the meta-atom geometries should be accurately approximated during simulations. The accurate representation of meta-atoms with arbitrary shapes require high-resolution meshings, practically limiting the structures to 2D or small 3D designs. As a result, the technique has been mostly used for optimizing periodic structures such as gratings and cylindrical structures.
To address this limitation, topology optimization recently has been combined with a local periodicity approximation. In this approach, topology optimization is done on small subdomain of the device whose response within the larger structure is approximated by periodic boundary conditions. These subdomains are subsequently stitched together to form the large-scale structure. This approach enables the optimization of large devices; however, subdomains with periodic boundaries do not accurately model the local response in high-NA metalenses or other rapidly-varying structures, limiting the performance of designs arrived at by this approach.

![Figure 1: Illustration of the parameterized metasurface design process using adjoint optimization. As the structure is updated by the optimization method, the desired output fields start to form.](image)

Instead of designing free-form structures, here we propose and demonstrate an adjoint optimization method based on parameterized rectangular meta-atoms (Fig. 1). Parameterized meta-atoms lack the fine features typical of topology-optimized structures, enabling simulations to converge at relatively low resolution and thus very large metasurfaces to be designed. Confining the design space to simple shapes (e.g., rectangular meta-atoms) also reduces the cost of simulation preprocessing steps like subpixel smoothing (Supplementary Note 2 and Supplementary Fig. S2). More importantly, limiting the optimization to this specific subspace of structures (i.e., rectangular meta-atoms) removes a large number of potential local optima traps without significantly affecting device performance and produces designs that conform to a well-established metasurface platform that can be easily fabricated. Our method also relies on a field interpolation technique for $E_\parallel$ and $D_\perp$ (see methods) and
an efficient time to frequency domain conversion technique to reduce the computational cost of simulating large structures. Our method relies on full-wave, finite difference time domain (FDTD) simulations of the entire structure, and iteratively approaches an optimal design via gradient ascent. A similar parameterized approach based on Mie theory was recently proposed by Zhan et. al.\textsuperscript{40,41} However, the approach is limited to spherical meta-atoms, which are challenging to fabricate, and does not account for the substrate’s effect. The adjoint optimization technique does not rely on the two implicit assumptions used in the unit-cell approach, and thereby achieves higher performing designs. First, the variation in coupling among the meta-atoms caused by the rapid variation of their dimensions is accounted for. Second, no assumption is made about angular dependence of the meta-atom scattering response (i.e., its element factor). In the following, we describe our method, and, as proof of concept, use it to design and fabricate two metalenses with 50 \( \mu \)m diameter. The focusing efficiencies of metalenses designed using this method show experimental improvements of 24\% and 13\% for NAs of 0.78 and 0.95 over counterparts designed by the conventional periodic unit-cell method.

Results

Parameterized adjoint design method

We first describe the metastructure design using the parameterized adjoint optimization method. The design process involves finding a set of meta-atom parameters that generate the desired transformation efficiently. As shown in Fig. 2a, to find the optimal design, we optimize the structure iteratively from a trivial initial design (e.g., a uniform array in which all meta-atoms are identical). In each iteration, the gradient of the objective function with respect to all the design parameters is calculated using only two simulations as conceptually shown in Fig. 2b-c. Based on the computed gradient, each meta-atom is updated, generating a new design that is one step closer to the optimal design. This cycle continues until all the
parameters converge to their final values (Fig. 2d).

![Diagram](https://via.placeholder.com/150)

Figure 2: Parameterized adjoint optimization. (a) Parameterized optimization: meta-atom dimensions are updated but constrained to a simple shape. (b) Representation of the forward problem. (c) Representation of the adjoint problem. (d) Flow diagram showing steps in the optimization.

The goal of metastructure design is to transform an incident electric field $\mathbf{E}_i$ into a desired transmitted or reflected field distribution $\mathbf{E}_d$. This is schematically illustrated in Fig. 2b, which shows a metasurface transforming a normally incident field into a converging transmitted field. An arbitrary desired output can be achieved by specifying an appropriate transmitted field distribution on a plane $S$ above the metasurface. The metasurface we consider consists of an arrangement of dissimilar meta-atoms positioned on a periodic lattice. Each meta-atom’s shape is described by one or more parameters that are variables in the design process. Thus a design can be expressed as a vector $\mathbf{p}$ containing all the design parameters. In our proposed method, the design is cast as an optimization problem of
maximizing the fraction of the output field in the desired field distribution. Specifically, an optimal design maximizes \( I = |F|^2 \), where

\[
F(p) = \int_S E^d \cdot E^f(p) \, dA. \tag{1}
\]

Here \( E^d \) is the desired field in the frequency domain on the plane \( S \), \( E^f(p) \) is the field realized by a design defined by \( p \) in the forward simulation excited by \( E^i \) (Fig. 2b), and \(^*\) represents the complex conjugate operation. \( F \) is the complex-valued projection of \( E^f \) on \( E^d \).

Optimization starts from an initial design \( p^{(0)} \) and is updated iteratively \( (p^{(1)}, p^{(2)}, \ldots) \) via gradient ascent. This process is illustrated in Fig. 2a: after each iteration, \( p \) approaches its locally optimal value and the performance of the metasurface improves. The gradient \( \nabla_p I \) is used to determine how \( p \) changes in the next step and can be computed using an additional simulation called the adjoint simulation. The adjoint simulation uses the same design \( p \) as the forward simulation, but the structure is instead excited by a surface current density \( J^a_s \equiv E^d^* \) that is placed on the plane \( S \) which generates a backward propagating wave (see Fig. 2c). The electric field in the adjoint simulation is denoted \( E^a(p) \).

In general, the variation of \( F \) with respect to small changes in the boundaries of meta-atoms can be found using the functional derivative of \( F \). An expression for the functional derivative of \( F \) based on symmetries of the Green’s tensor can be found in Ref. 42. Here, we consider the special case of rectangular meta-atoms with square cross-sections (inset of Fig. 2b). For such meta-atoms, \( p = (w_1, w_2, \ldots, w_N) \), where \( w_i \) represents the width of \( i^{th} \) meta-atom. Based on the Lorentz reciprocity theorem\(^{43} \), we show in Supporting Note 3 that the partial derivative of \( F \) with respect to \( w_i \) is given by

\[
\frac{\partial F}{\partial w_i} = \frac{1}{2} j \omega \left( n_m^2 - n_c^2 \right) \int_{\partial \Omega_i} \left( E^f_{\parallel} \cdot E^a_{\parallel} + \frac{1}{n_m^2 n_c^2} D^f_{\perp} D^a_{\perp} \right) \, dA, \tag{2}
\]

where \( \omega \) is the angular frequency of the excitation, \( n_m \) and \( n_c \) are the refractive indices of meta-atom and cladding materials, \( \partial \Omega_i \) represents the four side surfaces of the \( i^{th} \) meta-atom,
$E^f_\parallel$ and $D^f_\perp$ are the tangential components of the electric field and normal component of the displacement field obtained in the forward problem, and $E^a_\parallel$ and $D^a_\perp$ are the corresponding fields in the adjoint problem. The gradient of the objective function necessary to determine the next design is given by $\nabla p I = 2 \text{Re}\{F^* \nabla p F\}$, where $\text{Re}\{\cdot\}$ represents the real part of a complex number. Both forward and adjoint simulations are performed using full-wave simulations of the entire metasurface. Although this is computationally more expensive than techniques that employ local periodicity approximations,\textsuperscript{29,32,34,35} it allows the gradient to be calculated more accurately. The flow diagrams in Fig. 2d and Fig. S3 summarize the optimization procedure.

**Metalens design example**

To demonstrate the parameterized adjoint method, we designed two metalenses with NAs of 0.78 and 0.94 (Fig. 3a). The diameters of both metalenses are 50 μm, yielding focal lengths of 20 μm and 8.3 μm.

The metalenses are composed of 430-nm-tall square $\alpha$Si meta-atoms that are arranged on a rectangular lattice with a period of 320 nm. The meta-atoms rest on a fused silica substrate and are surrounded by air. For these designs, the parameter vector consists of the meta-atom widths, $p = (w_1, w_2, \ldots, w_N)$, where $N \approx 19,200$ is the number of meta-atoms. By imposing symmetries present in the problem we can reduce the design to 4800 independent variables. Still, the large number of independent variables and the long time required for each simulation precludes a detailed study of the design space. Both metalenses are initialized by a uniform array of 140-nm-wide meta-atoms (i.e., $w_i=140$ nm).

Both forward and adjoint simulations were performed using a finite difference time domain (FDTD) solver\textsuperscript{39} with a sinusoidal excitation that was gradually ramped up. In the forward simulations, the metalenses were illuminated by an $x$-polarized, normally incident plane wave (Fig. 3a) with a free-space wavelength of $\lambda_0=850$ nm. The desired output field $E^d$ was selected to be the field of an ideal, spherical-aberration-free flat lens (see Methods)\textsuperscript{44}. 
Figure 3: Simulation and design of the optimized and control metalenses. (a) Schematic of two metalenses with NAs of 0.78 and 0.95. The metalenses are illuminated by normally incident $x$-polarized plane waves. The incident field outside the metalens aperture is blocked by a perfect electric conductor (PEC) layer. (b) The focusing efficiencies of the optimized metalenses during the optimization process. Focusing efficiencies of the control metalenses are shown for comparison.Inset shows color-coded width distributions of meta-atoms at several steps during the optimization process. (c) Snapshot of $E_x$ on the output apertures of the optimized (left) and control metalenses (right). (d) Intensity at the focal planes of the optimized and control metalenses.
To expedite the simulations, symmetric boundary conditions were used along both $x$ and $y$ axes, reducing the simulation volume by a factor of four. The simulations were run until the results converged, and then the fields were converted from time to frequency domains using the method of ref. 45. The fields on the meta-atom side boundaries, necessary to determine $\nabla p F$, were interpolated from points on the Yee grid using a bilinear approach (see Methods and Fig. S4). Further simulation details are described in the Methods section.

In each step of the optimization, the design vector was updated according to $p^{(n+1)} = p^{(n)} + s \nabla p^{(n)} I$, where $s$ is the step size. The step size was chosen to achieve an average meta-atom width change of a few nanometers. As the optimization proceeded, the step size was manually updated, allowing $p$ to converge (see Methods and Fig. S5). To enforce polarization insensitivity, we symmetrized the derivatives along $x = y$ plane (see Methods and Fig. S6).

As a quantitative measure of performance, we calculated the focusing efficiency of each metalens during the optimization process. Focusing efficiency is directly related to the accurate implementation of the desired field profile, and metalenses with higher focusing efficiencies generally have less undesired scattered light and form higher contrast images close to their optical axes. For a fair comparison with the measured values (see the Experimental demonstration section below), we defined the focusing efficiency as the fraction of the power incident on the metalens aperture that passes through a 7-μm-diameter aperture in its focal plane. Figure 4b shows the focusing efficiencies of the optimized metalenses as their design evolved during the optimization process. Color-coded meta-atom width maps for these metalenses at several steps during the design process are shown as insets in Fig. 3b. At the first step of the optimization, the metalenses were periodic arrays of posts and had low focusing efficiencies. As the design proceeded, patterns similar to Fresnel zones appeared in the metalenses’ width distributions (Fig. 3b, insets), and their focusing efficiencies increased.

The designs were run for 64 iterations, although after only 25 steps their focusing efficiencies reached plateaus. At the last step, the focusing efficiencies of the optimized metalenses
with NAs of 0.78 and 0.94 were 78% and 55%, respectively. For comparison, we designed two control metalenses using the unit-cell approach with NAs, meta-atom heights, lattice constants, and diameters identical to the optimized ones. The simulated focusing efficiencies of the control metalenses are 69% and 43%. The details of the designs and simulations of these control metalenses are presented in Methods. Snapshots of the dominant component of the electric field ($E_x$) at the output apertures of the control and optimized metalenses are presented in Fig. 3c (for $E_y$ distributions see Fig. S7). The field distributions in Fig. 3c show that the optimized metalenses generate the desired fields with smaller phase errors, and consequently produce brighter focal spots than the control metalenses (Fig. 3d). The significantly higher focusing efficiencies of the optimized metalenses compared to their control counterparts demonstrate the efficacy of the parameterized adjoint optimization technique in designing high-performance metasurfaces.

**Experimental demonstration**

For experimental validation, we fabricated and characterized the optimized and control metalenses. The metalenses were fabricated by depositing a layer of aSi on a fused silica substrate and patterning it using electron beam lithography and dry etching (see Methods for details). Figure 5a shows an SEM image of a fabricated metalens. We characterized the metalenses using a setup schematically shown in Fig. 4b. Metalenses were illuminated by a collimated laser beam with a wavelength of $\lambda_0=850$ nm. The light transmitted through the metalens was collected by an objective lens with an NA of 0.95 and reimaged by a tube lens on an image sensor. Images of the focal spots, shown in Fig. 5b, show enhanced peak intensities for the optimized metalenses compared to the control ones.

We measured the focusing efficiencies of the metalenses by measuring the ratio of the optical power focused into a 7-µm-diameter pinhole in the focal plane of the metalenses and the power incident on their apertures (Fig. 5c). The measured focusing efficiencies of the optimized metalenses with NAs of 0.78 and 0.94 are 65% and 49%, respectively, higher than...
Figure 4: Experimental results. (a) Scanning electron beam micrograph of a fabricated metalens. (b) Schematic of the characterization setup for intensity measurements and (c) efficiency measurements. (d) Intensity distributions of optimized and conventional metalenses. (e) Intensity profiles are taken along the dashed lines shown in (d).
values of 52% and 43% obtained for their control counterparts. This represents 24% and 13% relative enhancements for the 0.78 and 0.94 NA lenses, respectively. The smaller increase for the higher NA metalens is attributable to the limitations of our measurement setup (the objective lens used has an NA of 0.95) and to its higher sensitivities to fabrication errors.

To study the sensitivity of our designs, an array of metalenses with a range of constant meta-atom offsets were fabricated alongside those characterized in Fig. 4. The study shows that the optimized metalenses have approximately the same sensitivities as the control ones (see Fig. S8).

**Discussion**

The parameterized adjoint optimization method accurately estimates shape derivatives of parameterized meta-atoms (see Supplementary Note 5 and Fig. S9). In contrast with methods that simulate structures in a dielectric continuum and then discretize to obtain a physically realizable design, meta-atoms designed by our method maintain a dielectric discontinuity at their boundaries throughout the whole design process, i.e., the simulation and design domains are the same. Techniques such as level-set representation can also be used to maintain boundaries with a dielectric discontinuity. We previously demonstrated such a technique in a similar silicon on glass material platform. Compared to the parametrized technique presented in this article, the simulations for the free-form level-set technique require significantly higher resolutions (i.e., much smaller grid size) to converge and the optimization domain has many more local optima. Due to their small features, the optimized metasurfaces obtained using this level-set approach are also significantly more difficult to fabricate. As a result, the application of level-set representation has been limited to small structures.

The parameterized adjoint optimization technique can be easily adapted for designing other types of metasurfaces such as achromatic metasurfaces (see Supplementary Note 4). We have presented the design of achromatic metalenses with parameterized shapes in Figs. S10
and S11. These metasurfaces provide comparable efficiencies to the ones designed using
topology optimization,\textsuperscript{31} and do not pose fabrication challenges similar to those of free-form
structures.

Using simple, parameterized shapes reduces the dimensionality of the metasurface design
space and simplifies the fabrication process. Designs produced by adjoint topology opti-
mization typically require hundreds of steps to converge.\textsuperscript{22,26} Parametrization enables us
to include our knowledge about principles of operation of metasurfaces by selecting proper
arrangement of the meta-atoms and other parameters such as meta-atom height and lattice constant. Our initial design (uniform metasurface comprising identical meta-atoms) although very simple, includes many important characteristics of the final design, so it can converge faster. The metalenses presented in this work evolved to designs with performance superior to the conventionally-designed controls in fewer than 15 steps. The quick convergence enabled us to optimize large-scale (50 \( \mu \)m diameter) metastructures, which, to the best of our knowledge, are currently some of the largest 3D adjoint-optimized metalenses. We previously demonstrated multifunctional multi-layer metasurface ref.\textsuperscript{49} devices using similar methods in approximately the same number of iterations. Furthermore, the number of iterations could be further reduced by implementing an adaptive step size.\textsuperscript{50}

The full-wave simulations employed in this work are computationally expensive. We
employed several techniques to keep the optimization of large devices feasible. The computa-
tional cost of FDTD simulation is directly related to the grid size used. We employed
bi-linear field interpolation, which increases the accuracy of the derivatives without reduc-
ing the grid size, keeping the computation time for each iteration tenable. To convert the
time-domain fields to the frequency domain, we only used two time samples using an effi-
cient harmonic detection method.\textsuperscript{45} This technique enables multi-wavelength optimization
at minimal additional cost (see Supplementary Note 4): wavelengths with independent ob-
jective functions can be incorporated into the simulations by adding appropriate sources and
acquiring a few additional time-domain samples without increasing the number or duration
of the simulations. Though in this work we presented metalenses optimized from a trivial initial state, we could have selected a conventionally designed metasurface (based on a unit-cell approach) as a starting point, which might have positioned the initial and final designs nearer to each other.

Like any other gradient-based optimization method, designs determined by our method represent local optima. However, parameterization allows us to restrict our search to a judiciously selected subspace by using prior knowledge about the problem. For example, information from low NA conventional designs can be useful in determining the appropriate meta-atom height and the lattice constant for a high NA adjoint-optimized design. To improve the chance of finding the global optimum, multiple optimizations starting from initial designs can be run in parallel. Results of such a multiple-seed optimization are shown in Fig. S12. Despite their different starting points, all designs converged to metalenses with similar focusing efficiencies. The observed behavior might not be general, but it seems to be valid at least for optimizing single layer structures with significant practical impact.

Because our method requires little knowledge about the final structure, it allows us to design elements for which conventional techniques fail to produce efficient designs, like multifunctional metasurfaces.\textsuperscript{49,51–53} In multifunctional devices, the interdependence of parameters is significantly more complex than in single function designs and simple models are unable to model meta-atom behavior accurately. In contrast, our method considers all the complex interactions and generates more efficient designs. Our method is also can be easily extended to other kinds of multi-objective optimizations, like robust designs, that are tolerant to fabrication\textsuperscript{54} error.

We envision that the adaptation of the parameterized adjoint optimization to design of large-scale metasurfaces will enable efficient cascading of multiple metasurfaces to implement compact, complex metasystems with high performance.
Acknowledgement

This work was funded by the Samsung Advanced Institute of Technology, and performed in part at the Center for Nanoscale Systems (CNS) at Harvard University, a member of the National Nanotechnology Coordinated Infrastructure Network (NNCI), which is supported by the National Science Foundation under NSF award no. 1541959.

Methods

Metalens optimization

The two metalenses designed by the adjoint technique and the control metalenses are composed of 430-nm-tall square cross-section aSi meta-atoms ($n_{\text{Si}} = 3.84$) that are positioned on a square lattice with a lattice constant of $\Lambda = 320$ nm. The meta-atoms are on a fused silica substrate ($n_s = 1.45$) and are cladded above by vacuum. One quadrant of each of the metalenses are shown in Fig. S7.

The optimization flowchart is shown in Fig. S3. To reduce the required computational resources, we simulated the fields in a small volume ($52 \, \mu\text{m} \times 52 \, \mu\text{m} \times 1.33 \, \mu\text{m}$) around the metasurface. All metalens optimization simulations were performed using a freely-available, open-source FDTD solver. Time-domain simulations were run until the fields converged (133 fs). The structure is terminated on all sides by a PML boundary condition. Because only the near field of the structure was simulated, fields at the focal plane (Fig. 4) were obtained by Fourier domain propagation. To further expedite the simulations, we exploited symmetries of the structure and fields: even mirror symmetry was specified along the $x$-axis and odd mirror symmetry along the $y$-axis, reducing the simulated volume by a factor of four.

Simulations were done using a workstation with an Intel E5-2680 CPU; 10 cores were used for each simulation. The FDTD grid size and the step size were adjusted manually when a reduction in the rate of improvement was observed. The simulations in each optimization
run began with a grid size of 33 nm (low resolution); after the device efficiency increased, the grid size was reduced to 20 nm (high resolution). Each iteration, consisting of both forward and adjoint simulations, took \sim 15 \text{ min} at low resolution and \sim 97 \text{ min} at high resolution. Color-coded plots of meta-atom widths of the optimized and control lens are shown in Fig. S13.

\textbf{Target field distribution}

For an \( x \)-polarized plane wave input \( \mathbf{E}^i = \hat{x}E_0 \) at wavelength \( \lambda_0 \) originating in a medium with refractive index \( n_c \), the desired field distribution for an ideal metalens with focal length \( f \) is:

\begin{align}
E^d_x &= E_0 t(\theta) \left( \cos \theta \cos^2 \phi + \sin^2 \phi \right) \tag{3} \\
E^d_y &= E_0 t(\theta) \left( 1 + \cos \theta \right) \sin \phi \cos \phi, \tag{4}
\end{align}

where \( t(\theta) = \sqrt{\frac{n_c}{\cos \theta}} \exp\left(-\frac{2\pijf}{\lambda_0 \cos \theta}\right) \), \( \theta = \tan^{-1}\left(\sqrt{x^2 + y^2} / f\right) \)\(^{44}\) is the local deflection angle of the metasurface, and \( \phi = \tan^{-1}(y/x) \) (see Fig. S14).

\textbf{Field interpolation}

The FDTD solver calculates fields on a rectangular grid (Yee grid). However, to determine the gradient, fields on the meta-atom boundaries are required. From the boundary conditions, we know the fields \( D_\perp \) and \( \mathbf{E}_\parallel \) are continuous. To obtain the boundary fields, we interpolated along axes normal to meta-atom boundaries using a two-sided linear fit approach that considers field values at four Yee lattice points (Fig. S4). For each field component \( C \), one linear fit \( C_{\text{in}}(x) \) was determined using two points \((x_{-2}, x_{-1})\) inside the meta-atom, and another, \( C_{\text{out}}(x) \), using two points \((x_1, x_2)\) outside the meta-atom. The field at the boundary \((x_0)\) was found based on the distance-weighted average of these two extrapolated values as
\[ C(x_0) \approx \alpha C_{\text{out}}(x_0) + \beta C_{\text{in}}(x_0), \]  
where \( \alpha \) and \( \beta \) are given by

\[ \alpha = \frac{|x_0 - x_{-1}|}{|x_1 - x_{-1}|}, \beta = \frac{|x_0 - x_1|}{|x_1 - x_{-1}|}. \]

Gradient symmetrization and scaling

To obtain polarization-insensitive metalens designs, in addition to the mirror symmetries along \( x \) and \( y \) axes described above for the simulation domain, we imposed a symmetry along the \( x = y \) line (see Fig. S6). The gradients were first determined for the simulated, \( x \)-polarized field for a quarter of the metalens and then symmetrized according to:

\[ \nabla_p I(x, y) \leftarrow \frac{1}{2} \nabla_p I(x, y) + \frac{1}{2} \nabla_p I(y, x). \]  

This operation is equivalent to computing the gradient for circularly polarized input light and optimizing the metalens using this symmetrized gradient ensures its polarization insensitivity. After determining the symmetrized gradient, the step size \( s \) was selected such that the average of the absolute change of the meta-atom widths \( \nabla_p I \) was equal to a few nanometers (see Fig. S5). The maximum change in the post widths was limited to 10 times the average value to ensure the first order gradient approximation is valid.

At the beginning of the optimization the absolute value of the average change was selected to be equal to 2 nm. Then, as the reduction in the rate of improvement was observed (see Fig. S5), it was reduced to 0.1 nm.

Control metalens designs

To compare the effectiveness of the proposed design method with the conventional unit-cell design approach, we designed two control metalenses using the unit-cell approach. The
control metalenses have the same design parameters as the optimized ones, i.e., with lattice constants of 320 nm, and square cross-section aSi meta-atoms \( (n_{\text{Si}} = 3.84) \) that are 430 nm tall. Simulated transmittance and phase of the transmission coefficient for a periodic array of meta-atoms are shown in Fig. S15a and were used to obtain the design map shown in Fig. S15b.

**Fabrication**

All metalenses were fabricated on the same fused silica substrate. To compensate for systematic errors in lithography, etching and other fabrication processes, an array of offsetted designs were included in the pattern. In each element of this array, widths of square meta-atoms are uniformly changed by a value in a range of \(-15 \text{ nm} \) to \(45 \text{ nm} \) in steps of \(5 \text{ nm} \). Figure S8 shows the measured efficiencies of fabricated metalenses with different offset values.

To pattern the metasurfaces, a 430-nm-thick layer of aSi was deposited on the substrate using plasma-enhanced chemical vapor deposition. Then, an approximately 220-nm-thick layer of electron-beam resist (ZEP520A-7, Zeon) was spin coated on the substrate. To avoid charging effects, a conductive polymer layer (ARPC-5090, Allresist) was spin coated on top of the resist. The patterns were defined using a 125 kV electron-beam lithography system (ELS-F125, Elionix), and then an aluminum oxide hard mask was deposited using an electron-beam evaporator. After lifting off the hard mask in a solvent (Remover PG, Microchem), the sample was etched using an inductively-coupled plasma reactive ion etching tool in an \( \text{SF}_6/\text{C}_4\text{F}_8 \) gas mixture. The hard mask was removed in a heated solution of ammonium hydroxide and hydrogen peroxide.

**Characterization**

We used the setup schematically drawn in Fig. S16a to acquire the focusing efficiency of the metalenses. Each metalens was illuminated by a weakly diverging Gaussian beam with a wavelength of 850 nm that was partially focused by a lens with 5 cm focal length (AC254-
050, Thorlabs). The light passed through the metalens and came into focus at a focal plane. Light in the focal plane was collected by a microscope objective with an NA of 0.95 (UMPlanFI 100×, Olympus), and reimaged by a tube lens (AC254-200, Thorlabs) and a camera (CoolSnap K4, Photometrics).

The focusing efficiency was defined as the ratio of the power focused inside a 7-µm-diameter pinhole in the focal plane of the metalens to the total power incident on the metalens. To measure the efficiency, we measured the power in the reimaged focal plane passing through an aperture with a diameter equivalent to a 7-µm-diameter pinhole in the metalens focal plane. A flip mirror in the imaging system (dashed box in Fig. S16) allowed us to direct the reimaged spot toward an apertured power meter (S120C, Thorlabs) and measure the focusing efficiencies.

The total incident power is measured by redirecting all the power in the partially focused beam (Fig. S16b) to the power-meter. Power incident on the metalens is different from the measured power because of the reflection at the second interface of fused silica to air (Fig. S16b). The measured power was corrected to indicate the actual power incident on the lens.

References

(1) Kamali, S. M.; Arbabi, E.; Arbabi, A.; Faraon, A. A review of dielectric optical metasurfaces for wavefront control. Nanophotonics 2018, 7, 1041–1068.

(2) Chen, F. T.; Craighead, H. G. Diffractive lens fabricated with mostly zeroth-order gratings. Opt. Lett. 1996, 21, 177.

(3) Arbabi, A.; Horie, Y.; Bagheri, M.; Faraon, A. Dielectric metasurfaces for complete control of phase and polarization with subwavelength spatial resolution and high transmission. Nat. Nanotechnol. 2015, 10, 937–943.
(4) Chen, X.; Huang, L.; Mühlenbernd, H.; Li, G.; Bai, B.; Tan, Q.; Jin, G.; Qiu, C.-W.; Zhang, S.; Zentgraf, T. Dual-polarity plasmonic metalens for visible light. *Nat. Commun.* 2012, 3, 1198.

(5) Lalanne, P.; Astilean, S.; Chavel, P.; Cambril, E.; Launois, H. Blazed binary subwavelength gratings with efficiencies larger than those of conventional échelette gratings. *Opt. Lett.* 1998, 23, 1081–1083.

(6) Huang, L.; Chen, X.; Mühlenbernd, H.; Zhang, H.; Chen, S.; Bai, B.; Tan, Q.; Jin, G.; Cheah, K.-W.; Qiu, C.-W.; Li, J.; Zentgraf, T.; Zhang, S. Three-dimensional optical holography using a plasmonic metasurface. *Nat. Commun.* 2013, 4, 2808.

(7) Zhang, X.; Jin, J.; Wang, Y.; Pu, M.; Li, X.; Zhao, Z.; Gao, P.; Wang, C.; Luo, X. Metasurface-based broadband hologram with high tolerance to fabrication errors. *Sci. Rep.* 2016, 6, 19856.

(8) Zheng, G.; Mühlenbernd, H.; Kenney, M.; Li, G.; Zentgraf, T.; Zhang, S. Metasurface holograms reaching 80% efficiency. *Nat. Nanotechnol.* 2015, 10, 308–312.

(9) Larouche, S.; Tsai, Y.-j.; Tyler, T.; Jokerst, N. M.; Smith, D. R. Infrared metamaterial phase holograms. *Nat. Mater.* 2012, 11, 450–454.

(10) Arbabi, A.; Arbabi, E.; Kamali, S. M.; Horie, Y.; Han, S.; Faraon, A. Miniature optical planar camera based on a wide-angle metasurface doublet corrected for monochromatic aberrations. *Nat. Commun.* 2016, 7, 443–803.

(11) Faraji-Dana, M.; Arbabi, E.; Arbabi, A.; Kamali, S. M.; Kwon, H.; Faraon, A. Compact folded metasurface spectrometer. *Nat. Commun.* 2018, 9, 1–8.

(12) Faraji-Dana, M.; Arbabi, E.; Kwon, H.; Kamali, S. M.; Arbabi, A.; Bartholomew, J. G.; Faraon, A. Hyperspectral imager with folded metasurface optics. *ACS Photonics* 2019, 6, 2161–2167.
(13) Arbabi, A.; Horie, Y.; Ball, A. J.; Bagheri, M.; Faraon, A. Subwavelength-thick lenses with high numerical apertures and large efficiency based on high-contrast transmittarrays. *Nat. Commun.* **2015**, *6*, 7069.

(14) Yu, N.; Genevet, P.; Kats, M. A.; Aieta, F.; Tetienne, J.-P.; Capasso, F.; Gaburro, Z. Light propagation with phase discontinuities: generalized laws of reflection and refraction. *Science* **2011**, *334*, 333–337.

(15) Bayati, E.; Zhan, A.; Colburn, S.; Zhelyeznyakov, M. V.; Majumdar, A. Role of refractive index in metalens performance. *Appl. Opt.* **2019**, *58*, 1460–1466.

(16) Yang, J.; Fan, J. A. Analysis of material selection on dielectric metasurface performance. *Opt. Express* **2017**, *25*, 23899–23909.

(17) Torfeh, M.; Arbabi, A. Modeling Metasurfaces Using Discrete-Space Impulse Response Technique. *ACS Photonics* **2020**, *7*, 941–950.

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

(19) Shokooh-Saremi, M.; Magnusson, R. Particle swarm optimization and its application to the design of diffraction grating filters. *Opt. Lett.* **2007**, *32*, 894–896.

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

(21) Sigmund, O. On the usefulness of non-gradient approaches in topology optimization. *Structural and Multidisciplinary Optimization* **2011**, *43*, 589–596.

(22) Jensen, J. S.; Sigmund, O. Topology optimization for nano-photonic. *Laser and Photonics Reviews* **2011**, *5*, 308–321.
(23) Niederberger, A. C. R.; Fattal, D. A.; Gauger, N. R.; Fan, S.; Beausoleil, R. G. Sensitivity analysis and optimization of sub-wavelength optical gratings using adjoints. Opt. Express 2014, 22, 12971–12981.

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

(25) Burger, M.; Osher, S. J.; Yablonovitch, E. Inverse problem techniques for the design of photonic crystals. IEICE T. Electron. 2004, 87, 258–265.

(26) Sell, D.; Yang, J.; Doshay, S.; Yang, R.; Fan, J. A. Large-angle, multifunctional metagratings based on freeform multimode geometries. Nano Lett. 2017, 17, 3752–3757.

(27) Phan, T.; Sell, D.; Yang, J.; Doshay, S.; Fan, J. A. Metasurface Lenses Based on Topology-Optimized Wavelength-Scale Building Blocks. in Conference on Lasers and Electro-Optics, OSA Technical Digest (Optical Society of America, 2018), paper FF3C.6.

(28) Tsuji, Y.; Hirayama, K.; Nomura, T.; Sato, K.; Nishiwaki, S. Design of optical circuit devices based on topology optimization. IEEE Photonics Technol. Lett. 2006, 18, 850–852.

(29) Lin, Z.; Liu, V.; Pestourie, R.; Johnson, S. G. Topology optimization of freeform large-area metasurfaces. Opt. Express 2019, 27, 15765–15775.

(30) Yang, J.; Fan, J. A. Topology-optimized metasurfaces: impact of initial geometric layout. Opt. Lett. 2017, 42, 3161.

(31) Chung, H.; Miller, O. D. High-NA achromatic metalenses by inverse design. Opt. Express 2020, 28, 6945–6965.
(32) Phan, T.; Sell, D.; Wang, E. W.; Doshay, S.; Edee, K.; Yang, J.; Fan, J. A. High-efficiency, large-area, topology-optimized metasurfaces. *Light: Science & Applications* 2019, 8, 1–9.

(33) Bayati, E.; Pestourie, R.; Colburn, S.; Lin, Z.; Johnson, S. G.; Majumdar, A. Inverse designed metalenses with extended depth of focus. *ACS Photonics* 2020, 7, 873–878.

(34) Lin, Z.; Johnson, S. G. Overlapping domains for topology optimization of large-area metasurfaces. *Opt. express* 2019, 27, 32445–32453.

(35) Pérez-Arancibia, C.; Pestourie, R.; Johnson, S. G. Sideways adiabaticity: beyond ray optics for slowly varying metasurfaces. *Opt. Express* 2018, 26, 30202–30230.

(36) Sigmund, O. Manufacturing tolerant topology optimization. *Acta. Mech. Sinica.* 2009, 25, 227–239.

(37) Wang, F.; Jensen, J. S.; Sigmund, O. Robust topology optimization of photonic crystal waveguides with tailored dispersion properties. *J. Opt. Soc. Am. B* 2011, 28, 387–397.

(38) Farjadpour, A.; Roundy, D.; Rodriguez, A.; Ibanescu, M.; Bermel, P.; Joannopoulos, J.; Johnson, S. G.; Burr, G. Improving accuracy by sub-pixel smoothing in FDTD. Tuning the Optic Response of Photonic Bandgap Structures III. 2006; p 63220G.

(39) Oskooi, A. F.; Roundy, D.; Ibanescu, M.; Bermel, P.; Joannopoulos, J. D.; Johnson, S. G. Meep: A flexible free-software package for electromagnetic simulations by the FDTD method. *Comput. Phys. Commun.* 2010, 181, 687–702.

(40) Zhan, A.; Fryett, T. K.; Colburn, S.; Majumdar, A. Inverse design of optical elements based on arrays of dielectric spheres. *Appl. Opt.* 2018, 57, 1437–1446.

(41) Zhan, A.; Gibson, R.; Whitehead, J.; Smith, E.; Hendrickson, J. R.; Majumdar, A. Controlling three-dimensional optical fields via inverse Mie scattering. *Sci. Adv.* 2019, 5.
(42) Miller, O. D. Photonic Design: From Fundamental Solar Cell Physics to Computational Inverse Design. Ph.D. thesis, UC Berkeley, 2012.

(43) Harrington, R. F. Time-harmonic electromagnetic fields; IEEE Press, 2001; p 480.

(44) McClung, A.; Mansouree, M.; Samudrala, S.; Arbabi, A. Properties of ideal flat metalenses. in Conference on Lasers and Electro-Optics, OSA Technical Digest (Optical Society of America, 2020), paper FM11R.1.

(45) Furse, C. M. Faster than Fourier: Ultra-efficient time-to-frequency-domain conversions for FDTD simulations. IEEE Antenn. Propag. M. 2000, 42, 24–34.

(46) 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 Photonics 2018, 5, 301–305.

(47) Sell, D.; Yang, J.; Doshay, S.; Fan, J. A. Periodic Dielectric Metasurfaces with High-Efficiency, Multiwavelength Functionalities. Adv. Opt. Mater. 2017, 5, 1700645.

(48) Mansouree, M.; Arbabi, A. Metasurface design using level-set and gradient descent optimization techniques. 2019 International Applied Computational Electromagnetics Society Symposium (ACES). 2019; pp 1–2.

(49) Mansouree, M.; Kwon, H.; Arbabi, E.; McClung, A.; Faraon, A.; Arbabi, A. Multifunctional 2.5D metastructures enabled by adjoint optimization. Optica 2020, 7, 77–84.

(50) Johnson, S. G. The NLopt nonlinear-optimization package. http://github.com/stevengj/nlopt.

(51) Arbabi, E.; Arbabi, A.; Kamali, S. M.; Horie, Y.; Faraon, A. Multiwavelength polarization-insensitive lenses based on dielectric metasurfaces with meta-molecules. Optica 2016, 3, 628–633.
(52) Kamali, S. M.; Arbabi, E.; Arbabi, A.; Horie, Y.; Faraji-Dana, M.; Faraon, A. Angle-
multiplexed metasurfaces: Encoding independent wavefronts in a single metasurface 
under different illumination angles. *Phys. Rev. X* 2017, 7, 041056.

(53) Zhou, Y.; Kravchenko, I. I.; Wang, H.; Nolen, J. R.; Gu, G.; Valentine, J. Multi-
layer noninteracting dielectric metasurfaces for multiwavelength metaoptics. *Nano Lett.* 
2018, 18, 7529–7537.

(54) Oskooi, A.; Mutapcic, A.; Noda, S.; Joannopoulos, J. D.; Boyd, S. P.; Johnson, S. G. Ro-
bust optimization of adiabatic tapers for coupling to slow-light photonic-crystal waveg-
uides. *Opt. Express* 2012, 20, 21558–21575.
Supplementary Note 1: Sensitivity of metasurfaces to subwavelength geometric changes

To show that subwavelength features can have a strong effect on metasurface performance, we compare the deflection efficiencies of two very similar beam deflectors. Fig. S1a shows a schematic of the beam deflector: The elements lie on a rectangular lattice with lattice constants of \( a_x = 700 \) nm and \( a_y = 1600 \) nm. The heights of the posts are chosen to be \( h = 800 \) nm and they are made of \( \alpha\)-Si with refractive index of \( n = 3.6 \). The posts are surrounded by air \( (n = 1.0) \) and rest on a fused silica substrate \( (n = 1.456) \). The beam deflectors are designed to deflect a normally-incident, \( x \)-polarized field at a wavelength \( \lambda = 1550 \) nm to an angle of 75°.

Figure S1b and c are both beam deflector designs and very similar to each other: Their boundaries change only by ca. 10 nm (See Fig. S1d). However, the efficiency of these devices is quite different: the design in Fig. S1b has an efficiency of 30%, whereas the design in Fig. S1c has an efficiency of 38%.
Supplementary Note 2: Comparison of subpixel averaging for free-form and parameterized shapes

Methods such as subpixel smoothing [1, 2] can be used to increase the accuracy of simulations without resorting to a high-resolution grid. In most simulations presented in this manuscript we used MEEP [2], which implements subpixel smoothing by default. Subpixel smoothing can be applied to structures with any shape, however, the computational cost of performing subpixel averaging for free-form structures is considerably higher than applying the same method for shapes with simpler boundaries (like the rectangular meta-atoms used in this work).

We performed a study of the time taken to perform subpixel smoothing as a function of resolution for two representative shapes: the first is a nano-post with a rectangular cross-section, and the second is a nano-post with a more complicated polygonal cross-section. The unit cell for both structures is the same, comprising a fused silica substrate with $n = 1.456$ and cladded with air. The meta-atoms are 500 nm tall and are made of $\alpha$-Si with $n = 3.824$ and the simulations wavelength is selected to be 850 nm. The time to initialize each structures using 4 threads as a function of resolution is shown in Fig. S2. As the Figure S2 clearly shows, the time to initialize the polygonal shape is considerably higher than that of the rectangular shape.
Supplementary Note 3: Gradient of complex objective

Here we derive Eq. (2), the expression for the gradient of the complex projection $F$ using the Lorentz reciprocity theorem. The complex projection $F$ for a given design $p$ is determined by Eq. (1). A small change in $p$, understood to be the differential increase $\delta w$ in the size of one of the meta-atoms, produces a related design $p'$ and an associated complex projection $F'$. The difference between these two designs, $\delta F \equiv F' - F$, is given by

$$\delta F = \int_S E^d \cdot \delta E^f \, dA,$$  

(S1)

where $\delta E^f \equiv E^f(p') - E^f(p)$ is the difference between the forward fields realized by these two designs.

The Lorentz reciprocity theorem specifies a relationship between two sources at the same frequency $\omega$, $J_1$ and $J_2$, and the fields produced by them, $E_1$ and $E_2$, in the same linear, reciprocal environment [3]:

$$\int E_1 \cdot J_2 \, dV = \int E_2 \cdot J_1 \, dV.$$  

(S2)

Here the integral is understood to extend over all space. Though $p$ and $p'$ represent different media, $E^f(p')$ can be understood as the sum of $E^f(p)$ and fields by an equivalent polarization current

$$J_{eq}(r) = j\omega \left( n_m^2 - n_c^2 \right) E^f(r, p'),$$  

(S3)

where $r$ is in the region formed of $n_m$ in $p'$ that previously consisted of $n_c$ in $p$, and $n_m$ and $n_c$ are the refractive indices of the meta-atom and the cladding material, respectively. We will label this region $\chi$. Thus current $J_{eq}$ and the fields $\delta E^f$ form one of the source–field pairs in Eq. S2. To determine the value of $\delta F$, we can consider a current source $J^a_s \equiv E^d$ on $S$ and its associated response $E^a(p)$. Using Eqs. S2 and S3, we can rewrite Eq. S1 as

$$\delta F = j\omega(n_m^2 - n_c^2) \int_{\chi} E^a(p) \cdot E^f(p') \, dV.$$  

(S4)

We would like to determine $\delta F$ without simulating the perturbed geometry $p'$, so we must approximate $E^f(p')$ using known quantities. We will do this using the dielectric boundary

S3
conditions: The components of the electric field parallel to a dielectric boundary, \( E_\parallel \), are continuous. Similarly, the component of the displacement field perpendicular to a dielectric boundary, \( D_\perp = \epsilon E_\perp \), are continuous. For a small change in the meta-atom geometry, we can make the approximations

\[
E_\parallel(p') \approx E_\parallel(p), \quad \text{(S5)}
\]
\[
D_\perp(p') \approx D_\perp(p). \quad \text{(S6)}
\]

The integration domain \( \chi \) between \( p \) and \( p' \) follows the contour of the meta-atom \( \partial \Omega \) and has thickness \( \delta w \); thus

\[
\delta F = j\omega (n_m^2 - n_c^2) \int_{\partial \Omega} \left[ E_\parallel(p) \cdot E_\parallel^*(p) + \frac{1}{n_m^2 n_c^2} D_\parallel(p) D_\parallel^*(p) \right] dA \delta w / 2, \quad \text{(S7)}
\]
giving the expression in Eq. (2).
Supplementary Note 4: Multi-wavelength optimization

We present a straightforward extension of our parameterized adjoint optimization technique to perform multi-wavelength optimization. Instead of a single harmonic excitation sources, for multi-wavelength optimization, a multi-tone source which includes all the desired wavelengths is used. Thus adding additional wavelength objectives requires almost no additional computational resources.

To calculate the objective function and the boundary fields on the meta-atoms we sample the fields at more time step and calculate the contributions of each distinct wavelengths. At each wavelength, the gradient is calculated similarly to single wavelength case explained in Supplementary Note 3. For multi-wavelength optimization, the total objective function $I_{\text{tot}}$ is chosen to be the product of each wavelength component:

$$I_{\text{tot}} = I_{\lambda_0} I_{\lambda_1} I_{\lambda_2} \cdots$$  \hfill (S1)

The gradient of $I_{\text{tot}}$ is obtained via the product rule:

$$\nabla_p I_{\text{tot}} = (\nabla_p I_{\lambda_0}) I_{\lambda_1} I_{\lambda_2} \cdots + I_{\lambda_0} (\nabla_p I_{\lambda_1}) I_{\lambda_2} \cdots + \cdots$$  \hfill (S2)

We present two multiwavelength cylindrical lenses with materials, layer heights and objectives similar to designs presented in Ref. [4] (Fig. S10 a-b), allowing our technique to be compared with topology-optimized designs. We optimized a 12.5-μm-wide cylindrical metasurface shown in Fig. S10c. The metasurface is composed of an array of titanium dioxide meta-atoms ($n = 2.645$) with height $h = 250$ nm (see inset of Fig. S10a) arranged on a rectangular lattice with a lattice constant of 200 nm.

The design is optimized at 11 wavelengths from 450 nm to 700 nm in steps of 25 nm. The metasurface is excited by an $x$-polarized incident field and designed to focus light at a focal distance $f = 8.2$ μm (i.e., NA=0.6). The optimization proceeds for 99 steps starting from an initial design in which all meta-atoms have equal width $w = 140$ nm. The normalized intensity at the last step is shown in Fig. S10e.

Initially all the meta-atoms have square cross-section. Over the course of the optimization the cross-sections are allowed to change to be rectangular; meta-atoms are also free to move in the $x$ direction. The efficiency of the metasurface at each optimized wavelength over the
course of the optimization is shown in Fig. S10d. Here the efficiency is defined as the ratio of the power between the first zero of an ideal diffraction limited lens on either side of maximum intensity to the power incident on the lens from substrate.

A second design is presented in Fig. S11. In this design we increased the height of the meta-atoms to 500 nm to correspond to the layer height of the freeform design in Ref. [4] (See Fig. S11 a-b). The metasurface and its unit-cell is schematically shown in Fig. S11c. The final efficiencies and the progress of the optimization is shown in Fig. S11d and f.

Material dispersion was not considered in the optimization process and the device was optimized at 11 different wavelengths with a constant refractive index of 2.654 for TiO$_2$. To evaluate the effect of TiO$_2$ dispersion, we simulated the final optimized metalens and found its efficiency while considering the wavelength dependence of TiO$_2$ (see Fig. S17). These efficiency values are shown in Fig. S10d and Fig. S11d. The efficiency values close to the maximum and minimum wavelength range are affected by dispersion, but the overall performance is not significantly affected. The average focusing efficiency values over the wavelength range when ignoring and considering the dispersion are 40.25% and 38.98%, respectively. Both these values are slightly larger than the ones reported in Ref. [4], showing the described method results in devices whose performance are comparable with state-of-the-art topology optimization.
Supplementary Note 5: Gradient accuracy study

The convergence rate of the gradient ascent technique depends on the accuracy of the gradient calculation. To validate the gradient determined by our adjoint method, we compared the derivatives determined by Eq. (2) to a finite difference approximation of the simulated objective function. For this study we used a simple grating, shown in Fig. S9. The grating is a periodic metasurface with a $2 \times 1$ unit-cell composed of 560–nm-tall rectangular amorphous silicon (aSi) meta-atoms resting on a fused silica substrate as shown in the inset of Fig. S9a. The simulation domain is periodic in $x$ and $y$ directions. In the $z$ direction, the structure is terminated above the output plane and below the input plane by perfectly matched layers (PML) with a thickness of 0.425 µm. The design vector for this system consists of the widths of the meta-atoms: $p = (w_1, w_2)$.

Simulations to verify the accuracy of the gradient vector estimated by the adjoint method were performed using a commercial finite element solver (COMSOL Multiphysics). To find the necessary electric fields in the forward simulation, the structure was excited by a surface electric current located inside the substrate. Using conformal meshing on the boundaries of meta-atoms the exact boundary fields required to calculate the derivatives in forward and adjoint simulations are obtained with minimal discretization errors.

In the forward simulation, the electric current source generates a plane wave excitation hitting the structure from below. In the adjoint problem, the structure was excited by a surface electric current placed on the output plane indicated in Fig. S9 that generates a plane wave hitting the structure exactly opposite to the direction of the first diffraction order.

We selected the desired field distribution $E^d = \frac{1}{A} \exp \left(-j \frac{2\pi}{\lambda_0} \sin(\theta_1)x\right) \hat{x}$, where $A$ is the area of the grating unit-cell, $\theta_1 = \sin^{-1} \frac{\lambda_0}{2a} = 55.61^\circ$ is the diffraction angle of the $+1$ diffraction order, $a = 515$ nm the metasurface lattice constant, and $\lambda_0 = 850$ nm the free-space wavelength of the excitation. The complex projection $F$ is defined as in Eq. (1), where the integral was computed over a grating period, and the objective function was defined as $I = |F|^2$. In the forward simulation, the structure is excited by a normally-incident $x$-polarized plane wave as depicted in Fig. S9a. The first diffraction order efficiency is calculated by $\eta = \frac{I \cos(\theta_1)}{2Z_0 P_0}$, where $I$ is the objective function, $P_0$ is the power of the incident
field to the unit-cell, $Z_0$ is the impedance of free-space.

In the adjoint simulation (Fig. S9b), the structure is excited by a surface current on a plane $S$ normal to the $z$ direction that lies 90 nm above the top surface of the meta-atoms. The adjoint source generates a plane wave that arrives at the metasurface with incidence angle $\theta_1$.

Our choice of $E^d$ makes $I$ proportional to the diffraction efficiency $\eta$ of the $+1$ diffraction order. Holding $w_2$ fixed at 100 nm, forward simulations were run varying $w_1$ between 100 nm and 200 nm in 1 nm increments. The diffraction efficiency as a function of $w_1$ is shown in Fig. S9c; its estimated derivatives, determined by both finite-difference and adjoint methods, are shown in Fig. S9d. The good agreement between the results of adjoint and finite-difference methods confirms the accuracy of the adjoint method.

References

[1] A. Farjadpour, D. Roundy, A. Rodriguez, M. Ibanescu, P. Bermel, J. Joannopoulos, S. G. Johnson, and G. Burr, “Improving accuracy by sub-pixel smoothing in fdtd,” in *Tuning the Optic Response of Photonic Bandgap Structures III*, vol. 6322, p. 63220G, International Society for Optics and Photonics, 2006.

[2] A. F. Oskooi, D. Roundy, M. Ibanescu, P. Bermel, J. D. Joannopoulos, and S. G. Johnson, “Meep: A flexible free-software package for electromagnetic simulations by the FDTD method,” *Comput. Phys. Commun.*, vol. 181, pp. 687–702, 2010.

[3] R. F. Harrington, *Time-harmonic Electromagnetic Fields*. IEEE Press, 2001.

[4] H. Chung and O. D. Miller, “High-na achromatic metalenses by inverse design,” *Opt. Express*, vol. 28, pp. 6945–6965, Mar 2020.
Supplementary Figures

Figure S1: Freeform beam deflectors. (a) Schematic of the beam deflector. (b) Initial beam-deflector. (c) Beam-deflector after applying small boundary changes on the meta-atoms. Red areas in (b) and (c) indicate α-Si and blue areas indicate air. (d) The difference between two designs in (b) and (c). Blue indicates the areas where α-Si is added and red where it is removed. (e) Deflection efficiencies before and after applied changes.
Figure S2: Comparison between subpixel smoothing time for a free-form (orange) and parameterized (blue) meta-atom.
Figure S3: Flowchart of adjoint optimization. In each iteration, two simulations are done and the derivatives are calculated. The cycle is repeated until the design converged.
Figure S4: Interpolation scheme. (a) Yee grid and structural boundary. (b) A line is fit to each pair of grid points inside and outside the meta-atom. Then, the two lines are extrapolated and averaged to estimate the value on the boundary. In the averaging, the value estimated by each line has a weight equal to the inverse distance between the boundary and closest point on the line.
Figure S5: Average change in each iteration.
Figure S6: Symmetry planes in the simulations for each polarization. (a) Symmetries for $x$-polarized incident light. (b) The symmetry for $y$-polarized incident light. The geometry has a mirror symmetry for both axes and symmetry along the $x = y$ line. This symmetry is enforced when updating the structure.
Figure S7: Cross-polarization components of the electric field $E_y$. The metalenses are excited with an $x$-polarized plane wave.
Figure S8: Measured efficiency for different values of bias in width of the meta-atoms. (a) for the metalens with NA=0.78, and (b) for NA=0.94. Adjoint optimized metalens is shown with solid line and control design is indicated by dashed line.
Figure S9: Adjoint derivative accuracy validation. (a) Schematic of a grating’s forward simulation, and (b) its adjoint simulation. (c) Diffraction efficiency of the +1 order as a function of $w_1$. (d) Comparison of derivative values calculated by finite-difference and adjoint methods.
Figure S10: Comparison of 250-nm-thick multi-wavelength cylindrical lenses with NA=0.6 designed with Ref. [4] and the parametrized adjoint optimization techniques. (a) Schematic drawing and (b) focusing efficiencies of topology optimized two-dimensional lens adapted with permission from [4]. (c) Schematic drawing of the cylindrical lens and its unit-cell designed with adjoint optimization method. (d) Focusing efficiencies. (e) Normalized field-intensity profiles at each wavelength. (f) Efficiencies at each step of the optimization.
Figure S11: Comparison of 500-nm-thick multi-wavelength cylindrical lenses with NA=0.6 designed with Ref. [4] and the parametrized adjoint optimization techniques. (a) Schematic drawing and (b) focusing efficiencies of topology optimized two-dimensional lens adapted with permission from [4]. (c) Schematic drawing of the cylindrical lens and its unit-cell designed with adjoint optimization method. (d) Focusing efficiencies. (e) Normalized field-intensity profiles at each wavelength. (f) Efficiencies at each step of the optimization.
Figure S12: Efficiency of metalenses with NA=0.78 and diameter of 50µm at each step of the optimization. Each optimization started with different starting metasurface composed of a uniform array of meta-atom with equal width.
Figure S13: Color-coded plots of the meta-atom widths of the optimized and control metalens. (a) For NA of 0.78 and (b) for NA of 0.94. Because the structure is symmetric with respect to $x$ and $y$ axes, only the meta-atom widths in the first quadrant are shown.
Figure S14: Relative coordinates for the desired output function in (a) for $xy$ plane and (b) in $yz$ plane.
Figure S15: Conventional unit-cell approach design maps. (a) Simulated transmittance and phase of the transmission coefficient for control metalenses. (b) Design map used for the control metalenses.
Figure S16:  (a) Schematic drawing of the characterization setup. The fabricated meta-surfaces were illuminated by a partially focused beam. The light passes through the sample and then is re-imaged on an image sensor using an objective lens (NA=0.95) and a tube lens. To measure the efficiency, a flip mirror was used to redirect the light toward an aperture with a diameter equal to the 7 µm in the focal plane. (b) The setup is used to measure the incident power on the metalens.
Figure S17: Refractive index of titanium dioxide.