

Supporting Information for:

# A high aspect ratio inverse-designed holey metalens

*Soon Wei Daniel Lim<sup>1‡\*</sup>, Maryna L. Meretska<sup>1‡</sup>, Federico Capasso<sup>1\*</sup>*

<sup>1</sup>Harvard John A. Paulson School of Engineering and Applied Sciences, 9 Oxford Street,  
Cambridge, MA 02138, USA

\*Corresponding authors: [lim982@g.harvard.edu](mailto:lim982@g.harvard.edu), [capasso@seas.harvard.edu](mailto:capasso@seas.harvard.edu)

‡S.W.D.L. and M.L.M. contributed equally to this work.

## SUPPLEMENTARY METHODS

### **S1. Theory of Bloch modes and effective indices**

Here we study differences and similarities between pillar and hole structures for the same unit cell size and periodicity. Increasing the diameter of pillar nanostructures increases the phase delay (relative to the phase of the incident plane wave) imposed on the transmitted light as more light is confined within the higher refractive index material as compared to the lower index medium. Increasing the diameter of hole nanostructures has the opposite effect whereby the phase delay decreases, since less light passes through the lower refractive index air openings. These transmitted phase behaviors of periodic circular pillar and hole structures as a function of the feature diameter are compared in **Supplementary Figure 7a** (for shorter 0.5  $\mu\text{m}$  structures) and **Supplementary Figure 7c** (for taller 5  $\mu\text{m}$  tall structures). Notably, the transmitted amplitude variation as a function of feature diameter exhibits qualitative differences between pillar and hole arrays (**Supplementary Figure 7b** for 0.5  $\mu\text{m}$  tall structures and **Supplementary Figure 7d** for 5  $\mu\text{m}$  tall structures). For cylindrical pillar structures, the transmitted amplitude (relative to the amplitude of incident light) remains around unity and approximately constant for small diameters. However, for cylindrical holes structures, the transmitted amplitude exhibits large variations with hole diameter, with this variation being the largest for the smallest holes. The approximately constant transmission amplitude for pillar-like meta-atoms allows one to arrange such meta-atoms based on their transmission phase alone. In fact, the transmitted amplitudes of cylindrical pillar arrays also exhibit variations as a function of feature diameter even in the small feature regime; these variations are just shallow, allowing the transmitted amplitude to be treated as effectively constant. The maxima in transmission amplitude for both periodic pillar and hole arrays arise from the

Fabry-Perot resonances of the dominant Bloch eigenmode in the effective medium regime. Specifically, when the effective index  $n_{eff}$  satisfies the resonant transmission condition,  $n_{eff}k_0H=m\pi$ ,  $m \in 1, 2, 3, \dots$ , the relative transmission amplitude is unity. Importantly, in the effective medium regime, the transmitted amplitude is strongly correlated to the transmitted phase: the amplitude maxima in the feature diameter plot occur with a relative phase delay of 0 or  $2\pi$  radians, and the amplitude minima correspond to phase delays of  $\frac{\pi}{2}$  or  $\frac{3\pi}{2}$  radians. Thus, one is not able to select a range of hole-like meta-atoms that span the full  $2\pi$  phase range and have constant transmission efficiency.

The strong correlation of phase and amplitude for hole meta-atoms implies that one will have to settle for low-efficiency hole meta-atoms for required phase delays around  $\frac{\pi}{2}$  or  $\frac{3\pi}{2}$  radians. For phase profiles such as the hyperbolic lens phase profile, which requires phase delays across the full 0 to  $2\pi$  radians range, these low meta-element efficiencies will reduce the overall device efficiency.

The fundamental reason for the amplitude fluctuations with meta-atom diameter can be understood in terms of Bloch eigenmodes, which are sometimes called leaky modes in literature<sup>1</sup>. The electromagnetic field distribution in any lossless structure with transverse periodic boundary conditions can be written as a linear combination of these eigenmodes<sup>1</sup>. Bloch eigenmodes are typically parametrized by the effective index  $n_{eff}$ , which is proportional the rate of phase accumulation  $n_{eff}k_0$  in the direction of propagation orthogonal to the plane with periodicity. Only propagating Bloch eigenmodes contribute significantly to transmissive optical behavior since they are not attenuated upon propagation through the structure. If there is only one propagating Bloch

eigenmode and the layer  $H$  is much thicker than the exponential decay length of the evanescent modes, then we are in the effective medium regime<sup>1</sup> in which the optical properties can be predicted by replacing the structure with a uniform slab of refractive index equal to  $n_{eff}$ . In particular, the complex transmission coefficient  $t_{FP}$  can be written using the Fabry-Perot (FP) formalism as:

$$t_{FP} = \frac{4(n_{eff}/n_0)e_2^{in_{eff}k_0H}}{(n_{eff}/n_0+1)^2 - (n_{eff}/n_0-1)^2 e^{2in_{eff}k_0H}}. \quad (1)$$

For the parameter range we studied, every Bloch mode with an imaginary effective index had a characteristic exponential decay length under 12 nm, thus the imaginary effective index Bloch modes (evanescent modes) contribute minimally to optical transmission properties through thick structures exceeding hundreds of nanometers.

We plot the Bloch mode characteristics of hole and pillar arrays in **Supplementary Figure 8a** and **b**, respectively. We consider square-periodic cylindrical hole and pillar arrays in silicon parametrized by the normalized periodicity  $P$  (to the vacuum wavelength of incident light  $\lambda_0 = 1.55 \mu\text{m}$ ) and the normalized feature diameter  $D$  (to the periodicity). The contour plot details the number of propagating Bloch eigenmodes for each  $(D, P)$  combination for the silicon/air platform. We observe that for small periodicities approximately around  $P < 0.3$  ( $P < 0.28$  for holes and  $P < 0.31$  for pillars), the array only supports one propagating mode, indicating that we are in the effective medium regime. This single-mode regime occupies a larger fraction of the  $(D, P)$  phase space for pillar arrays as compared to hole arrays. As the amount of material (silicon) increases ( $P \rightarrow 1, D \rightarrow 0$  for hole arrays or  $P \rightarrow 1, D \rightarrow 1$  for pillar arrays), the number of propagating modes also increases.

To examine the validity of this effective medium regime approximation over the  $(D, P)$  phase space, we compute the predicted complex transmission coefficient using the largest Bloch mode  $n_{eff}$  for a slab thickness of 5  $\mu\text{m}$  and compare this value with the true value  $t_{RCWA}$  (for the zeroth transmitted order) obtained with 256 plane waves in a RCWA computation. The surface plots in **Supplementary Figure 8a** and **b** show the deviation value of  $|t_{RCWA} - t_{FP}|$  over the  $(D, P)$  space for hole and pillar arrays, respectively. We observe that this deviation is not only small in the few-propagating-mode regime, but also when the small-hole diameters are considered, despite the number of propagating Bloch modes being large for large periodicities. This is indicative of the small-hole arrays having only one dominant Bloch eigenmode into which most of the incident energy is coupled, so these arrays can also be treated using the effective medium formalism.

We plot the effective index of the hole and pillar modes as a function of the feature diameter for the case where  $P = 0.32\lambda = 500 \text{ nm}$  in **Supplementary Figure 8c**. This periodicity corresponds to that used in this holey metalens design for  $\lambda_0 = 1.55 \mu\text{m}$ . In both hole and pillar cases, the dominant Bloch mode corresponds to the band with the highest effective index displayed. We also display the electric field intensity  $|\vec{E}|^2$  distribution for each mode within the figure insets.

The dominant Bloch mode effective index can be used to predict the transmission phase and amplitude for the hole array using Eq. 1 with  $P = 0.32\lambda_0 = 500 \text{ nm}$ . In **Supplementary Figure 8d**, the RCWA-computed transmission phase and amplitude is plotted against the hole diameter, and compared with those predicted based on the dominant Bloch mode effective index substituted into the FP approximation. The close correspondence between the two methods indicates that the complex transmission amplitude can be well-approximated by the FP transmission formula, despite the hole array exhibiting more than one propagating eigenmode for small diameters. The

single real value of the dominant Bloch mode effective index captures the combined behavior of both transmitted phase and amplitude simultaneously.

The insight that we can replace meta-atoms with their corresponding effective index allows us to engineer devices that require a spatially variant refractive index. As described in the main text, we can identify the optimal spatially variant refractive index profile for a flat lens that produces a diffraction limited spot. This is a shift in viewing these hole-like meta-atoms. Instead of being inferior elements for enforcing a phase delay (due to their low transmission efficiencies at specific phase delays), they function precisely as intended as a means for enforcing a local effective index.

## **S2. Electromagnetic simulation and numerical optimization**

The transmission behavior and Bloch eigenmodes of pillar and hole arrays are computed using the Rigorous Coupled Wave Analysis (RCWA) software RETICOLO<sup>2</sup>. The phase of the meta-atom is the complex argument of the transmitted zeroth order amplitude, using the  $e^{-i\omega t}$  convention for time harmonicity. The focal spot profiles for a given diameter profile are computed using the locally periodic approximation, where each hole or pillar meta-atom is replaced with the zeroth-order complex transmission amplitude calculated for an infinite periodic array of that meta-element, and the resulting complex amplitude profile is propagated into free space using a fully vectorial diffraction integral<sup>3</sup>. Specifically, for a radial meta-element diameter profile of the form  $D(r_i), r_i = ip, i = 0, 1, \dots, N_{max}$ , where  $i$  is the radial index of a meta-element,  $p$  is the constant center-to-center pitch of the meta-elements, and  $N_{max}$  is the maximal radial index of the outermost meta-element, we first compute the zeroth-order complex transmission amplitude  $E(D(r_i))$  for each value of  $D(r_i)$ , then propagate this complex, cylindrically symmetric wavefront into the forward domain using the full vectorial Green's function for one transverse polarization.

For numerical optimization, the objective function is the electric field intensity at the focal position and the optimization variables are the effective indices at each radial position on the lens  $n_{eff}(r_i)$ . Since there are 2001 radial positions over the 2 mm diameter lens, there are 2001 real-valued free parameters for optimization. This is a high dimensional optimization problem with local optima that can be identified using iterative gradient descent techniques. The computationally expensive step for such an optimization is obtaining the objective function gradient, which is a vector of partial derivatives of the objective function value with respect to each free parameter. Finite difference estimates of this gradient are inefficient due to the large number of free parameters involved, since at least one objective function calculation needs to be performed for each free parameter partial derivative required. In our study, we employ automatic differentiation, which provides a computationally efficient and exact means of obtaining this gradient, taking no more than the steps required to compute the objective function calculation five times, regardless of the number of free parameters involved<sup>4</sup>. While performing a calculation on an automatic differentiation platform, every mathematical operation is recorded on a computational tree, allowing the tree to be traversed using the chain rule for derivatives to obtain the exact numerical derivative of any parameter with respect to any number of other parameters. We thus perform the electromagnetic field propagation step on the automatic differentiation platform Tensorflow<sup>5</sup>, and employ the limited memory Broyden–Fletcher–Goldfarb–Shanno bound constrained optimization algorithm (L-BFGS-B)<sup>6,7</sup> for gradient descent optimization. The starting point for the optimization is the effective index profile that produces a hyperbolic phase profile. We truncate the optimization

when the gradient norm decreases below a pre-defined tolerance, indicating that the system has converged to a local optimum.

### **S3. Nanofabrication protocol**

The fabrication protocol for holey metalenses is summarized in **Supplementary Figure 4**. The holey metalens is fabricated on a 12.7 mm x 12.7 mm SOI wafer chip with a  $5\pm0.5$   $\mu\text{m}$  thick crystalline  $\langle 100 \rangle$  phosphorus-doped silicon device layer, 1  $\mu\text{m}$  thick buried oxide layer, and a 500  $\mu\text{m}$   $\langle 100 \rangle$  phosphorus-doped silicon handle layer. 400 nm of  $\text{SiO}_2$  is deposited onto the device layer using plasma enhanced chemical vapor deposition (PECVD, Surface Technology Systems). The  $\text{SiO}_2$  layer is then patterned through electron beam lithography (electron beam resist 950PMMA 7A and Elionix HS-50 lithography system) to introduce the circular meta-atom patterns. The pattern is then transferred onto the  $\text{SiO}_2$  layer with inductively coupled plasma reactive ion etching (ICP RIE, Surface Technology Systems) using the electron beam resist as an etch mask. We use argon/tetrafluoromethane/trifluoromethane/hydrogen ( $\text{Ar}/\text{CF}_4/\text{CHF}_3/\text{H}_2$ ) for this etch step.

After etching the  $\text{SiO}_2$  mask, the deep meta-atom via-holes are etched using a Bosch process with alternating cycles of deposition with  $\text{C}_4\text{F}_8$ , dry etching in oxygen, and dry etching in  $\text{SF}_6$ . The patterned  $\text{SiO}_2$  is used as a hard mask for this Bosch process. During the etching process, the  $\text{SiO}_2$  mask is also slowly etched away from the substrate, leaving approximately 250 nm of  $\text{SiO}_2$  upon completion. The hard mask is then removed by dipping the sample for 3 min in buffered hydrofluoric acid (BHF) 7:1 solution. A layer of PMMA resist is then applied on top of the

patterned surface. This resist layer protects the patterned surface from damage during further processing of the sample.

The chip is then flipped over so that the device layer is on the bottom and a maskless aligner (Heidelberg MLA150) is used to write a 4 mm diameter circle onto a positive photoresist layer (AZ4620, developed in AZ 400K). Next, the photoresist layer is used as the etch mask in a deep reactive ion etch (DRIE, SPTS Technologies, Omega LPX Rapier model) to remove the silicon within that 4 mm circle. The process time for this dry etch is tightly controlled so that the etch stops at the buried oxide layer. The exposed buried oxide layer is then removed with a 10-minute wet etch in BHF 7:1 solution. The sample is then thoroughly cleaned using Piranha solution and oxygen plasma to remove the protective PMMA layer.

The fabricated devices are characterized using focused ion beam (FIB) cross-sectional milling and scanning electron microscopy (SEM) on the Helios Nanolab 660 platform.

The Bosch process is a time-multiplexed etching process that alternates steps between etching and passivation. In the first step the structure is etched using  $\text{SF}_6$  plasma (Si etch) for a short period of time. In the next step the exposed surfaces are passivated with octafluorocyclobutane,  $\text{C}_4\text{F}_8$ , yielding a protective surface that has a much slower etch rate than Si during the  $\text{SF}_6$  cycle. During the following  $\text{SF}_6$  cycle, the etch sputters the passivation layer away from the bottom of the sample, allowing the Si etch to proceed downwards. The passivation layer on the sidewalls protects the sidewalls from lateral etching. The process is repeated multiple times to yield nearly vertical structures with distinctive scalloped sidewalls that represent the small amount of lateral etching produced during each cycle. These scalloped structures are visible in the top part of the hole FIB cross section in **Figure 2d**.

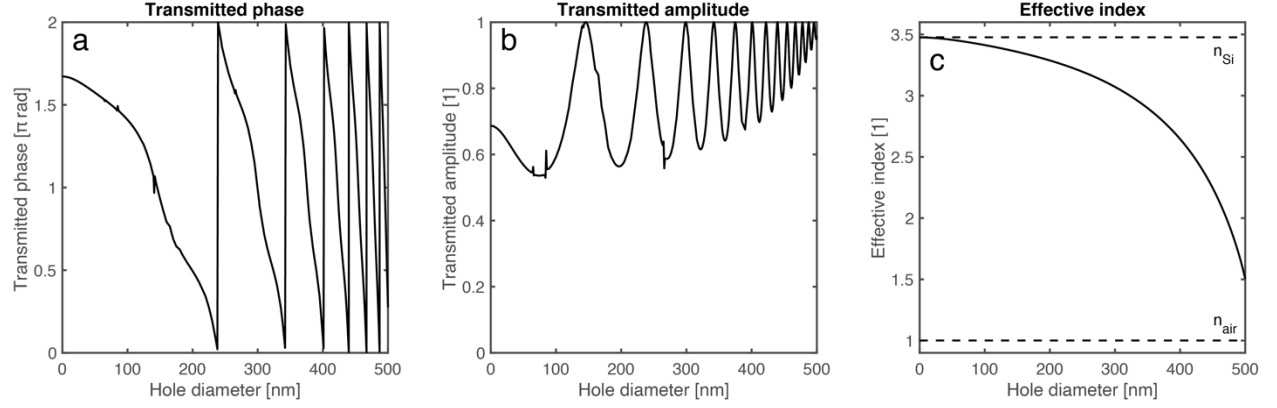
#### **S4. Optical characterization**

For focal spot measurements, we use the setup shown in **Supplementary Figure 6a**. A supercontinuum laser (NKT Photonics SuperK) with a laser-line tunable bandpass filter (NKT Photonics LLTF Contrast, bandwidth  $<5$  nm) is used for illumination. The expanded and collimated laser output from a single-mode fiber was passed through a 2 mm diameter graduated iris diaphragm (pinhole 1) and is normally incident on side II of the holey metalens. The metalens focal plane is imaged using a microscope objective (50x Mitutoyo Plan Apo NIR HR), tube lens (Thorlabs TTL200-A), and a near-infrared camera (Raptor Photonics OW1.7-VS-CL-640, 15  $\mu\text{m}$  pixel pitch). The spatial coordinates of the microscope imaging system are calibrated by imaging the divisions on a microscope stage calibration slide at the same wavelength. We extract four focal spot intensity cross sections and average the Strehl ratios and FWHMs obtained give the mean parameter values with their uncertainties.

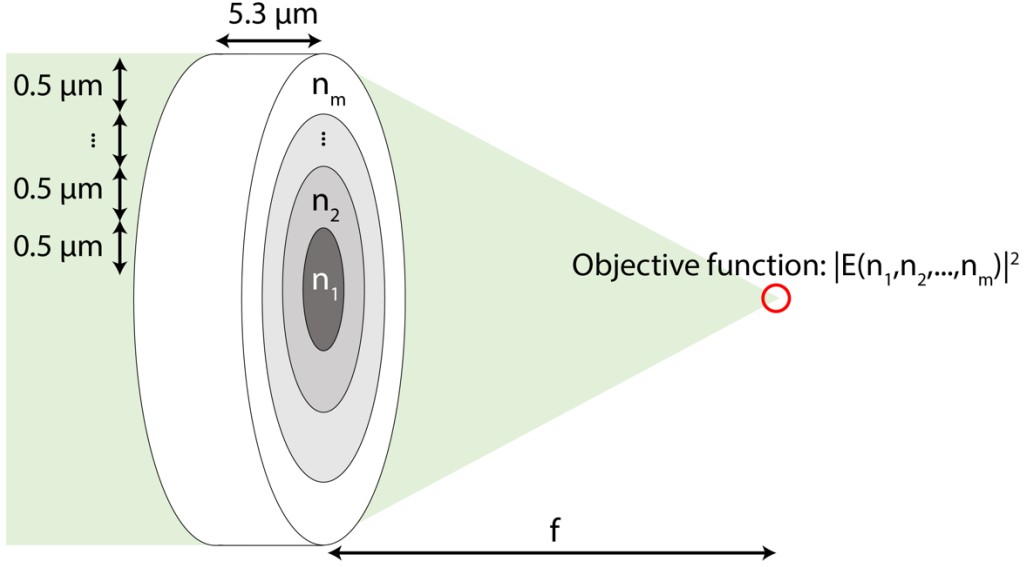
For efficiency measurements, the setup in **Supplementary Figure 6b** is used. The  $\lambda=1.55$   $\mu\text{m}$  laser beam was passed through a 2 mm aperture to illuminate the holey metalens from side II. A 20  $\mu\text{m}$  diameter precision aperture (around 2.6 times the Airy disk diameter of 7.8  $\mu\text{m}$ ) is placed in the focal plane to block out light around the focal spot. The focal spot and incident laser power values are measured directly using a laser power meter (Thorlabs S132C) placed after the 20  $\mu\text{m}$  aperture. The focusing efficiency or absolute efficiency is given by the ratio of the focal spot power divided by the incident laser beam power. The diffraction efficiency or relative efficiency is given by the ratio of the focal spot power divided by the transmitted light power measured when the 20  $\mu\text{m}$  pinhole is removed.

For imaging measurements, the setup in **Supplementary Figure 6c** is used. The  $\lambda=1.55\text{ }\mu\text{m}$  laser beam passed through a rotating diffuser and a negative USAF 1951 resolution target. The holey metalens is used to image the resolution target as the first lens in a 4f system with a  $f=45\text{ mm}$  achromatic doublet (Thorlabs AC254-045-C-ML). The magnified resolution target image is captured by the near-infrared camera (Raptor Photonics OW1.7-VS-CL-640).

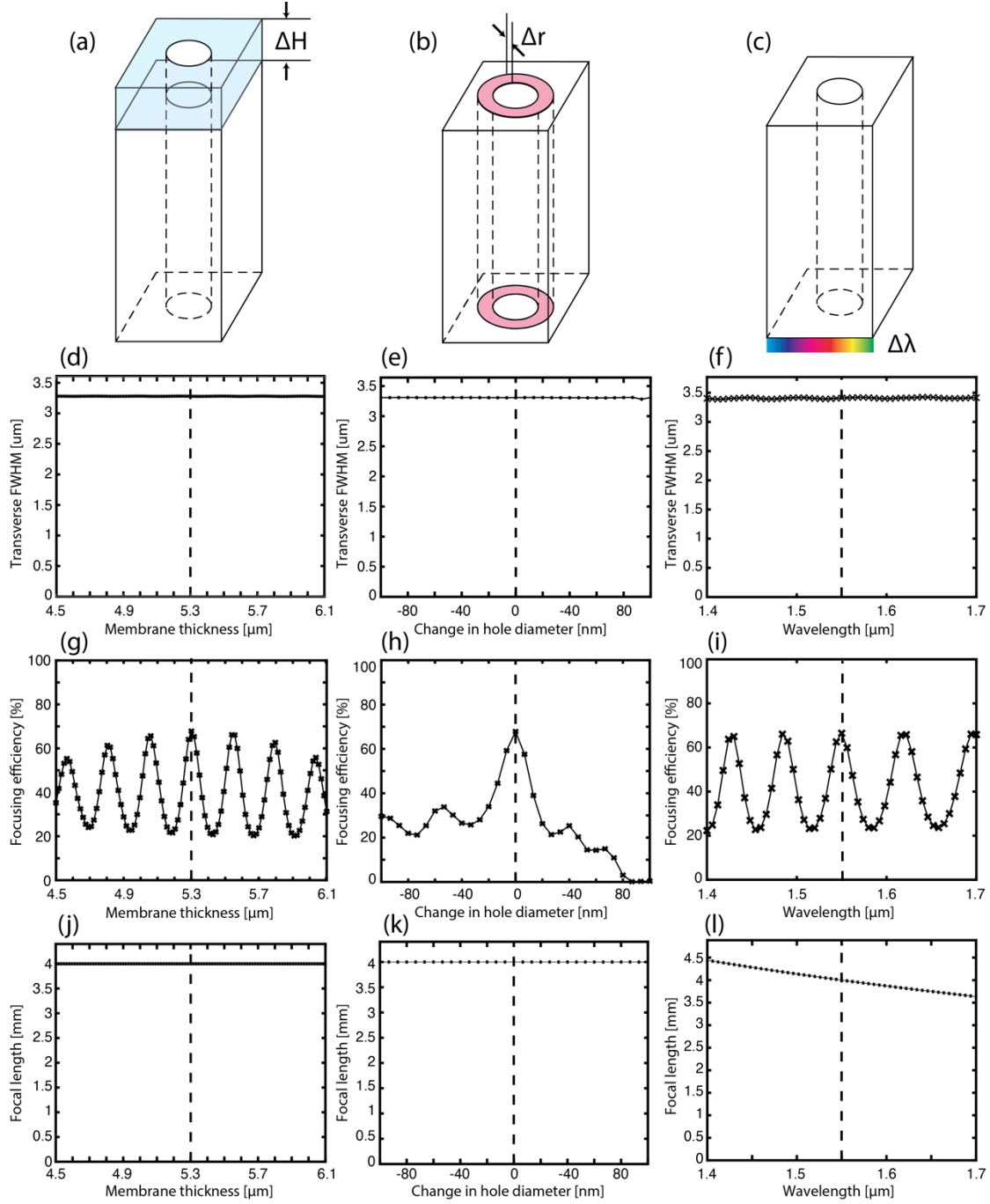
## SUPPLEMENTARY FIGURES



**Supplementary Figure 1.** Transmitted optical properties for the holey metalens library of  $H = 5.3$   $\mu\text{m}$  deep holes in silicon at  $\lambda = 1550$  nm, with periodicity  $P = 500$  nm. (a) Transmitted phase as a function of hole diameter. (b) Transmitted amplitude as a function of hole diameter. (c) Effective index of the dominant Bloch mode as a function of hole diameter. The refractive indices of silicon and air at  $\lambda = 1550$  nm are also shown in dotted lines.



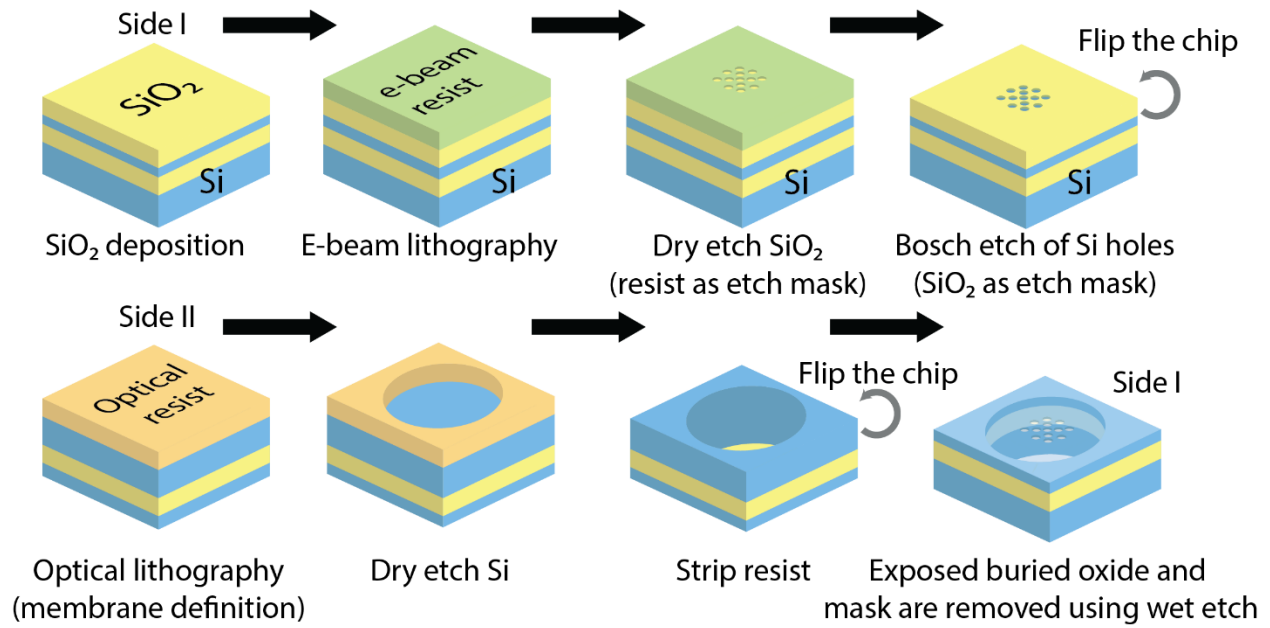
**Supplementary Figure 2.** Diagram indicating the optimization geometry. The thick metasurface is partitioned into  $m = 2001$  thin rings of radial extension  $0.5 \mu\text{m}$ , and each is associated with one effective index value  $n_i$ . The electric field above each ring position is computed using the complex amplitude transmission coefficient associated with each effective index value using the locally periodic assumption, and the electric field is then propagated to the focal plane using the full vectorial propagator integral. The objective function for the optimization is the electric field intensity at the focal position, and the variable parameters are the 2001 effective index values.



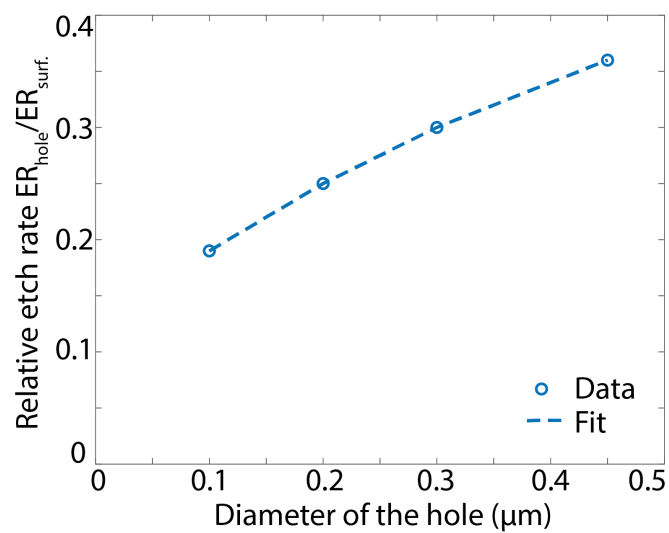
**Supplementary Figure 3.** Robustness of the design to fabrication and spectral imperfections.

Three types of systematic errors are considered (a) membrane thickness variations, (b) the hole diameter variations, and (c) spectral variations. Vertical dotted lines in the following graphs

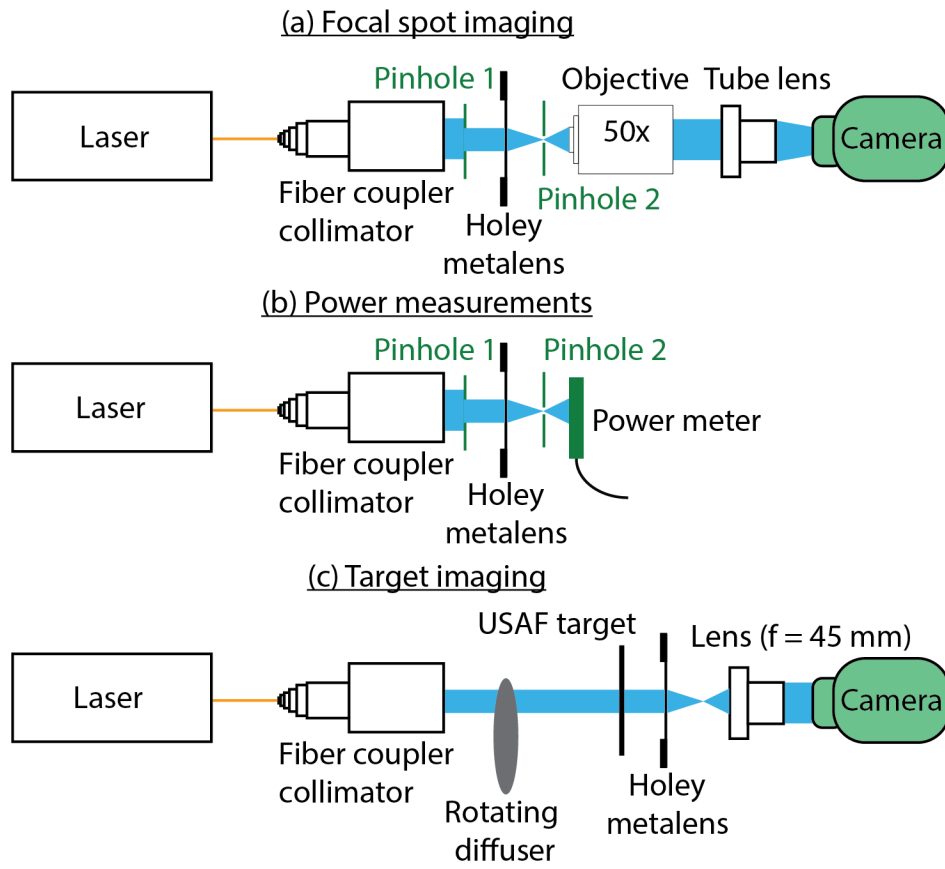
indicate the design value. (d) Plot of the transverse focal spot full-width-at-half-maximum (FWHM) intensity as a function of membrane thickness. (e) Plot of the transverse focal spot FWHM intensity as a function of hole diameter variation. (f) Plot of the transverse focal spot FWHM intensity as a function of incident wavelength. (g) Plot of primary focal spot focusing (absolute) efficiency as a function of membrane thickness. (h) Plot of primary focal spot focusing efficiency as a function of hole diameter variation. (i) Plot of primary focal spot focusing efficiency as a function of incident wavelength. (j) Plot of primary focal length as a function of membrane thickness. (k) Plot of primary focal length as a function of hole diameter variation. (l) Plot of primary focal length as a function of incident wavelength.



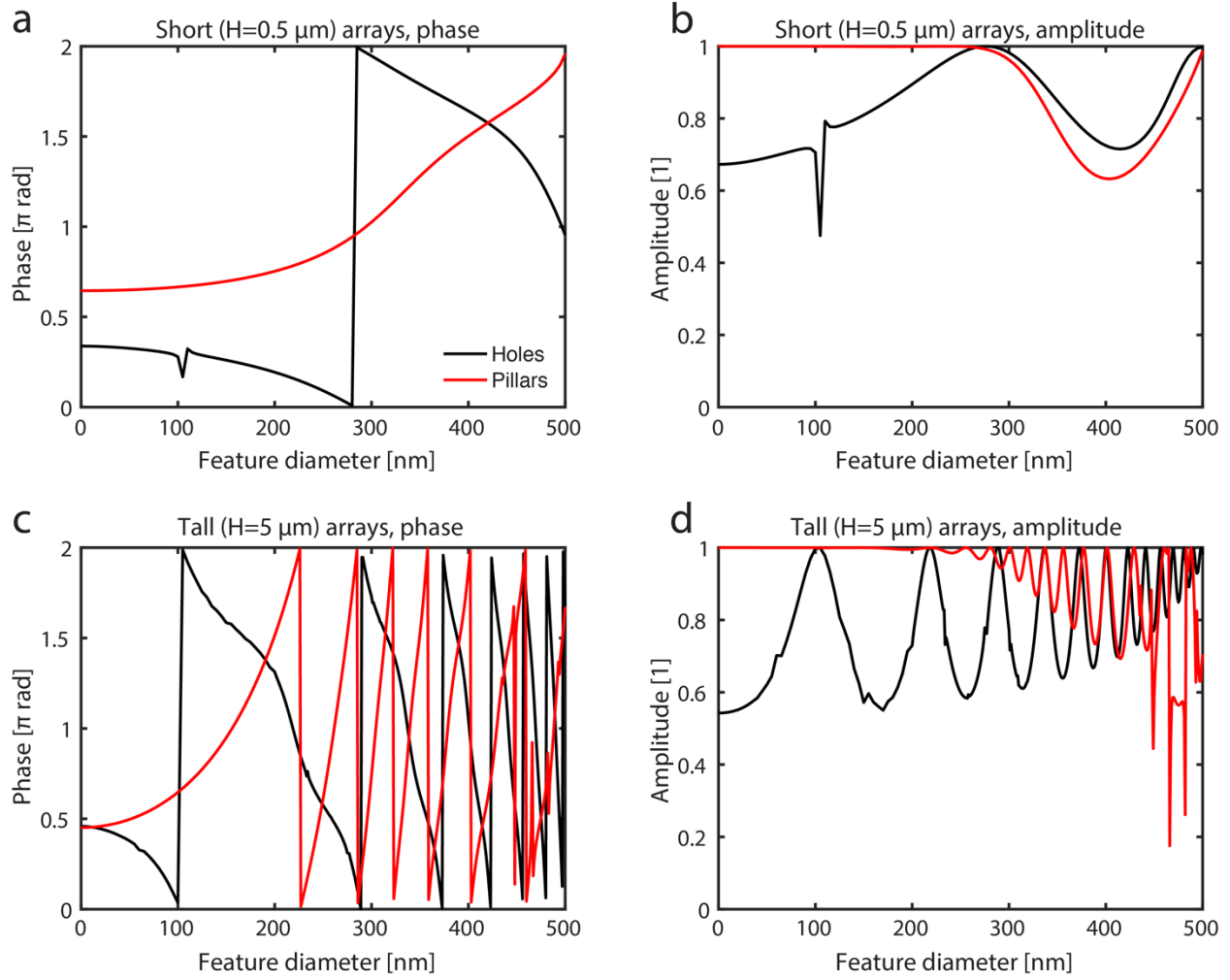
**Supplementary Figure 4.** Fabrication workflow for holey metalenses. From top left to bottom right: First, an SiO<sub>2</sub> hard mask layer and electron beam resist layer is deposited on the device layer of an SOI wafer (side I). The resist layer is patterned with electron-beam lithography and the pattern is transferred to the SiO<sub>2</sub> layer through dry etching using the resist as an etch mask. Next, the electron beam resist is removed, and the deep high-aspect ratio holes are etched into the silicon device layer using deep reactive-ion Bosch etching. The SiO<sub>2</sub> mask is removed using a BHF wet etch, and the surface of the lens is covered with resist for protection. The reverse side of the chip (side II) is coated with optical resist and large circular pattern is written using a maskless aligner. The exposed silicon is dry etched until the buried oxide layer is exposed. The exposed buried oxide is removed using a BHF wet etch to expose the final perforated silicon membrane.



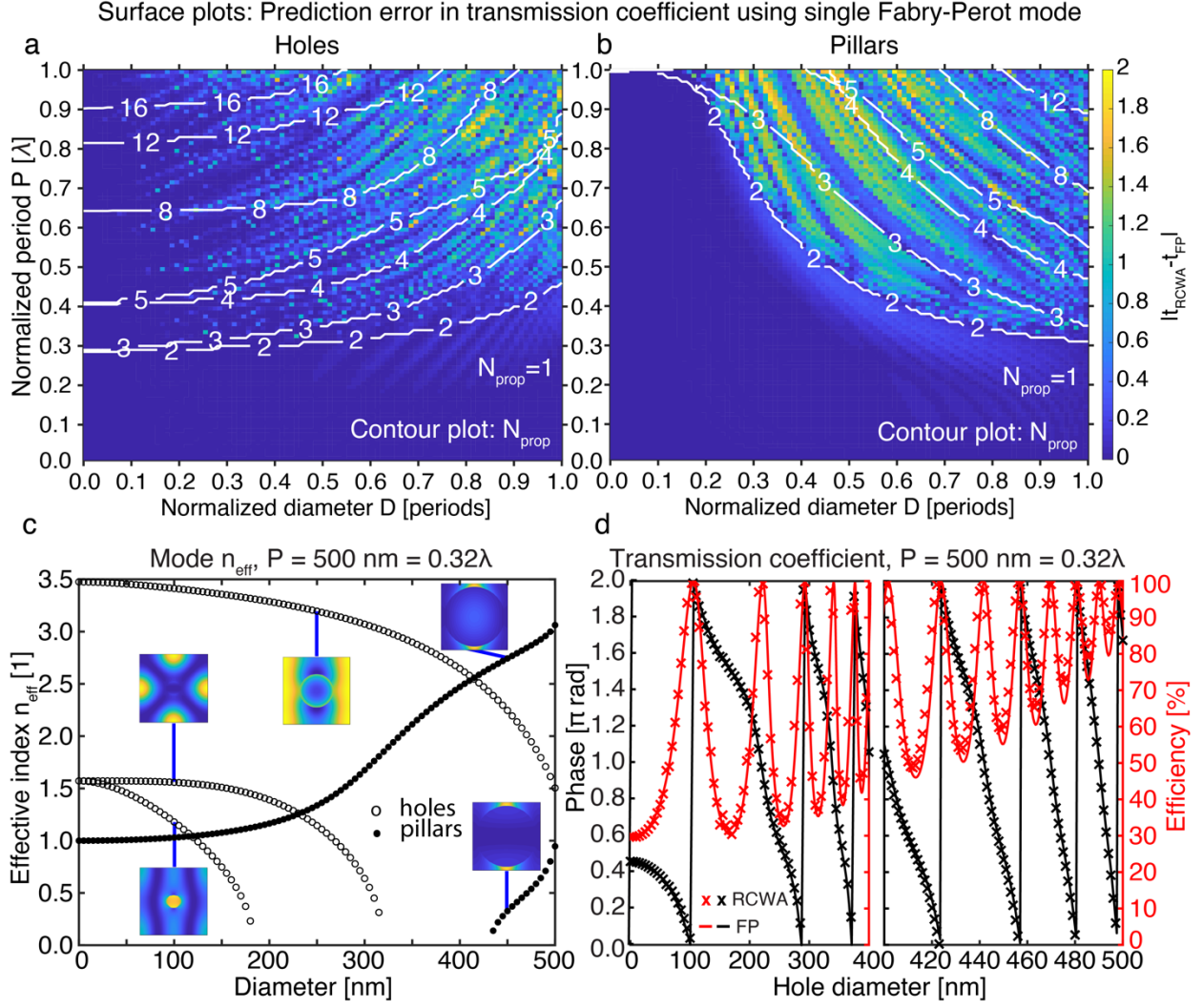
**Supplementary Figure 5.** Relative deep reactive-ion Bosch etch rate measured for various hole diameters within a uniform hole array. The etch rate is normalized to that of an open surface. The fitted line (second order polynomial) is also plotted in the dotted line.



**Supplementary Figure 6.** Experimental setup used for (a) imaging of the focal spot, (b) power measurements, and (c) resolution target imaging.



**Supplementary Figure 7.** Phase and amplitude dependence of hole and pillar meta-atoms for two structure thicknesses. Top row: short ( $0.5 \mu\text{m}$ ) holes and pillars. Bottom row: tall ( $5 \mu\text{m}$ ) holes and pillars. (a) Transmitted phase for pillar and hole arrays of  $500 \text{ nm}$  height. (b) Amplitude (relative to the incident amplitude) for structures of  $500 \text{ nm}$  height. (c) Transmitted phase for pillar and hole arrays of  $5 \mu\text{m}$  height. (d) Amplitude (relative to the incident amplitude) for structures of  $5 \mu\text{m}$  height.



**Supplementary Figure 8.** Eigenmode properties for the (a) periodic hole array and (b) periodic cylindrical pillar array geometries, as a function of the normalized periodicity  $P$  (to the incident vacuum wavelength  $\lambda$ ) and the normalized diameter  $D$  (to  $P$ ). Surface plots: prediction error in the complex transmission coefficient  $t$  when using a single Fabry-Perot (FP) mode as compared to the full calculation from rigorous coupled wave analysis (RCWA). Contour plots: Number of propagating Bloch eigenmodes  $N_{prop}$ . (c) Plot of the propagating Bloch eigenmode effective refractive indices for the case where  $\lambda=1550 \text{ nm}$ ,  $P=500 \text{ nm}=0.32\lambda$ . The inset plots exhibit the

eigenmode electric field intensity distributions  $|\vec{E}|^2$  for the labelled eigenmodes. The dominant Bloch mode corresponds to the branch with the highest refractive index in this regime. (d) Complex transmission coefficient (phase and absolute efficiency) dependence on the hole diameter for the hole array where  $\lambda = 1550$  nm,  $P = 500$  nm  $= 0.32\lambda$ . Crosses represent full RCWA computations with many eigenmodes and lines represent predictions using the dominant Bloch eigenmode effective index alone. The close correspondence between the FP predictions and RCWA calculations indicate that the hole array is well-approximated by the behavior of the dominant Bloch eigenmode.

## SUPPLEMENTARY REFERENCES

- (1) Lalanne, P.; Hugonin, J. P.; Chavel, P. Optical Properties of Deep Lamellar Gratings: A Coupled Bloch-Mode Insight. *Journal of Lightwave Technology* **2006**, *24* (6), 2442–2449.
- (2) Hugonin, J. P.; Lalanne, P. Reticolo Software for Grating Analysis. Institut d’Optique: Orsay, France 2005.
- (3) Marathay, A. S.; McCalmont, J. F. Vector Diffraction Theory for Electromagnetic Waves. *Journal of the Optical Society of America A* **2001**, *18* (10), 2585–2593.
- (4) Griewank, A. On Automatic Differentiation. *Mathematical Programming: Recent Developments and Applications* **1989**, 83–108.
- (5) Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; Kudlur, M.; Levenberg, J.; Monga, R.; Moore, S.; Murray, D. G.; Steiner, B.; Tucker, P.; Vasudevan, V.; Warden, P.; Wicke, M.; Yu, Y.; Zheng, X. TensorFlow: A System for Large-Scale Machine Learning. In *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16)*; 2016; pp 265–283.
- (6) Zhu, C.; Byrd, R. H.; Lu, P.; Nocedal, J. Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization. *ACM Transactions on Mathematical Software* **1997**, *23* (4), 550–560.
- (7) Byrd, R. H.; Lu, P.; Nocedal, J.; Zhu, C. A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM Journal on Scientific Computing* **1995**, *16* (5), 1190–1208.

