DE GRUYTER Nanophotonics 2023; 12(13): 2371–2381

Research Article

Haejun Chung*, Feng Zhang, Hao Li, Owen D. Miller and Henry I. Smith*

Inverse design of high-NA metalens for maskless lithography

https://doi.org/10.1515/nanoph-2022-0761
Received December 13, 2022; accepted February 16, 2023;

Abstract: We demonstrate an axisymmetric inverse-designed metalens to improve the performance of zone-plate-array lithography (ZPAL), one of the maskless lithography approaches, that offer a new paradigm for nanoscale research and industry. First, we derive a computational upper bound for a unit-cell-based axisymmetric metalens. Then, we demonstrate a fabrication-compatible inverse-designed metalens with 85.50% transmission normalized focusing efficiency at 0.6 numerical aperture at 405 nm wavelength; a higher efficiency than a theoretical gradient index lens design (79.98%). We also demonstrate experimental validation for our axisymmetric inverse-designed metalens via electron beam lithography. Metalens-based maskless lithography may open a new way of achieving low-cost, large-area nanofabrication.

Keywords: inverse design; maskless lithography; metalens.

1 Introduction

Zone-plate-array lithography (ZPAL) creates patterns of arbitrary geometry by overlapping focal spots produced by an array of microlenses [1–3]. Such maskless lithography avoids the cost and delay time associated with procuring a photomask [4]. It also avoids troublesome coherence problems associated with mask-based projection photolithography. Figure 1(a) is a schematic of the ZPAL technology. A significant obstacle for ZPAL is the need for a high-efficiency, high-numerical-aperture (NA) microlenses. Recently, the advent of the "metalens" has led to the development of structured, wavelength-scale devices offering superior performance to conventional approaches [5–10]. A standard metalens design approach, called unit-cell design, stitches subwavelength resonators together into a larger device, demonstrating the possibility of focusing [6–10], but it has suffered from low focusing efficiencies or low-NA restrictions. Recently, it has been shown that the unit-cell design approach cannot have high efficiency at high-NA due to a fast oscillation of the field phase and amplitude profiles and a local periodicity violation [11].

In this work, we apply inverse design, a large-scale computational technique [12–17], to discover a high-NA metalens at 405 nm wavelength to realize a high-resolution, high-efficiency ZPAL lithography system. Inverse design enables fast computation of Figure-of-Merit (FoM) gradients with respect to design parameters over the entire design space. Fabrication-constraints are implemented by penalization technique [18, 19] and constrained optimization [20, 21], enabling fast discovery of fabrication-compatible designs.

2 Methods

2.1 Maskless lithography: the microlenses

It is well known that binary π-phase gratings [22], in which the zeroth order is canceled and second and third orders are cut off, have a diffraction efficiency of ∼42% into the +1-order [23]. Based on this, one would expect a maximum focal efficiency of about 40% for a binary π-phase-shifting zone plate [23]. Assuming this, 60% of the transmitted light would constitute background exposure. This background is composed of residual zeroth order as well as a −1-order virtual focus and third order foci, both real and virtual. Because ZPAL's writing scheme consists of overlapping non-coherent focal spots, interference effect are avoided and linear superposition applies. As a consequence, proximity-effect correction (PEC) can be easily implemented if the focal spot's point-spread function, including the broad background, is

---

***Corresponding authors: Haejun Chung**, Department of Electronic Engineering, Hanyang University, Seoul, 04763, South Korea; Department of Artificial Intelligence, Hanyang University, Seoul, 04763, South Korea, E-mail: haejun@hanyang.ac.kr. https://orcid.org/0000-0001-8959-237X; and **Henry I. Smith**, LumArray, Inc., 15 Ward Street, Somerville, MA 02143, USA; and Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA, E-mail: hismith@mit.edu
Feng Zhang, LumArray, Inc., 15 Ward Street, Somerville, MA 02143, USA, E-mail: fengz@lumarray.com
Hao Li and Owen D. Miller, Department of Applied Physics and Energy Sciences Institute, Yale University, New Haven, CT 06511, USA, E-mail: owen.miller@yale.edu

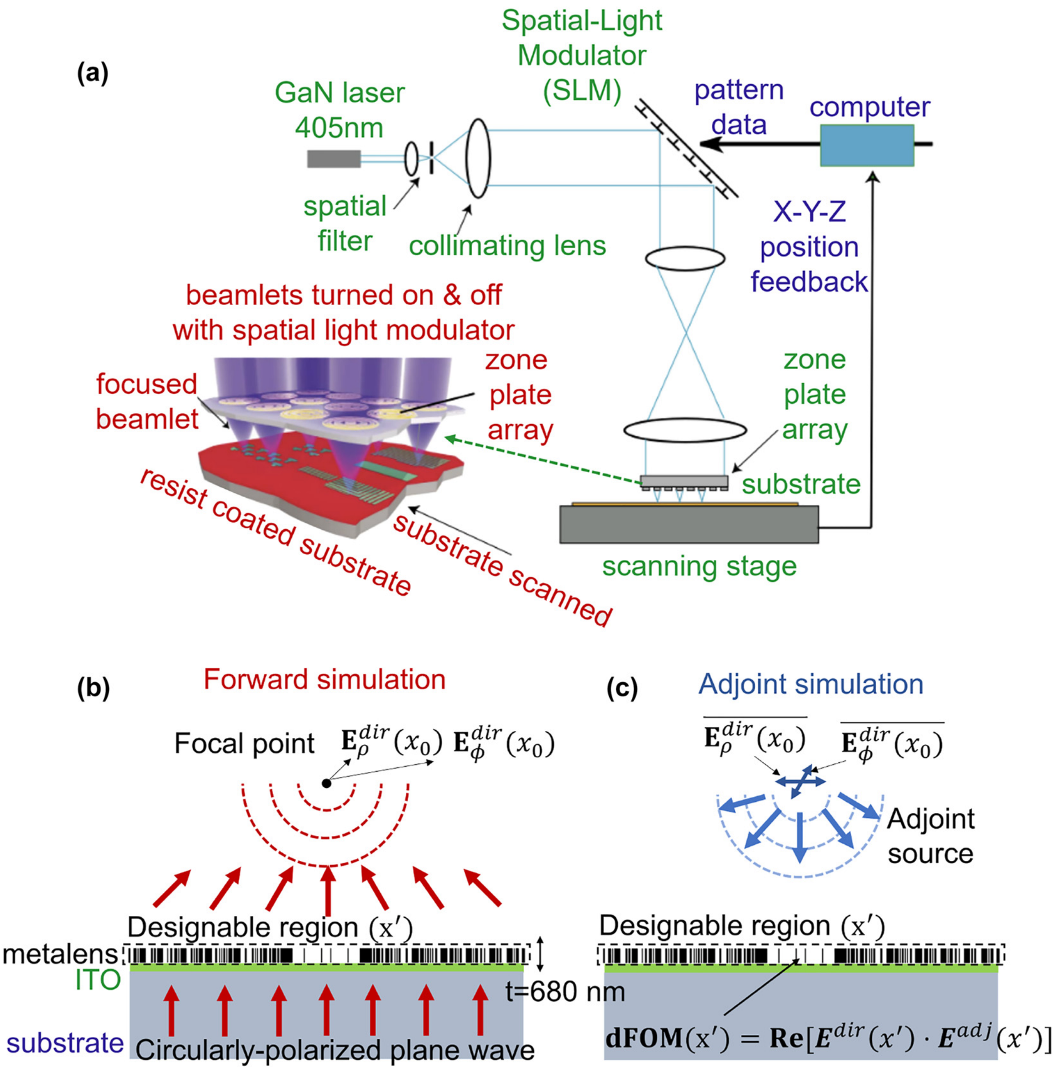

Figure 1: Metalens inverse design. (a) Schematic depiction of zone-plate-array lithography (ZPAL). A CW laser illuminates a spatial-light modulator, each pixel of which controls the light intensity to one zone-plate of the array. By adjusting the focal-spot intensity from zero to maximum in a quasi-continuous manner, linewidth can be controlled and proximity-effects corrected. By moving the stage under computer control, while intelligently modulating focal-spot intensities, patterns of arbitrary geometry can be written. (b) Schematic depiction of forward simulation of inverse design. A circularly-polarized plane wave is excited from the substrate. $E_\rho^{\mathrm{dir}}(x_0)$, $E_\phi^{\mathrm{dir}}(x_0)$ at the focal point are stored for the adjoint simulation. (c) A conjugated $E_\rho^{\mathrm{dir}}(x_0)$, $E_\phi^{\mathrm{dir}}(x_0)$ field now back-propagates to obtain adjoint electric fields at the designable region (x').

known. Although PEC can compensate for a high background level (as is done in scanning-electron-beam lithography where electron backscattering produces a large, broad background) it nevertheless is highly desirable to reduce this background.

In projection photolithography, it is common to express the minimum practical linewidth achievable in dense patterns as,

$$W_{\min} = k_1 \lambda / \mathrm{NA} \tag{1}$$

where λ is the optical wavelength, NA is the lens's numerical aperture, and k_1 is a proportionality factor, obtained empirically. In practice, the lower the background level, the lower one can push k_1 in pursuit of finer linewidths.

In seeking microlenses with higher focal efficiencies than π-phase zone plates, the obvious first approach would be shaping the individual zones, to produce a gradual phase delay across them, a scheme referred to as "blazing" in the case of gratings [24, 25]. Such a microlens is commonly called a Fresnel lens [26, 27]. Aside from the difficulty of performing such blazing with existing nanofabrication tools, it is well known that the focal efficiency of Fresnel lenses drops rapidly with increasing NA [8]. Some improvement can be achieved by so-called effective-index modulation (EIM) [28, 29] in which the thickness of the dielectric is uniform from zone to zone, but across each zone dielectric material is selectively removed to achieve a desired variation of phase delay. Not surprisingly this approach also encounters decreasing focal efficiency with increasing NA [11]. Our intuitive models about how light behaves in transiting and emerging from such complex structures break down when diffraction angles get large.

In recent years, several groups have investigated so-called metasurfaces which includes metalenses [6, 7, 9, 11, 30–32]. The term metasurface was introduced to convey the notion of metal or dielectric structures in which the dimensions of features are below the wavelength of light. In many cases, one can think of these small structures behaving as sub-wavelength antennas.

High-numerical-aperture (NA) metalenses offer the possibility of significantly improving the resolution of the ZPAL maskless lithography system. Here, we study the theoretical efficiency limits of high-NA metalenses and then use inverse design to realize a high-efficiency high-NA metalens. Our inverse-designed metalens shows 85% efficiency, close to the theoretical limit. We also fabricated metalens on SiO_2 substrates to demonstrate the feasibility of our approach.

2.2 Upper bound of unit-cell design

To start, we ask: what is the largest efficiency metalens designed by the unit-cell method could achieve? We apply an analytical technique first developed by some of the authors in Ref. [11]. The idea is as follows. In the unit-cell method, one approximates the response of each metalens unit as though it were part of a periodic structure. The collective plane of outgoing fields should then comprise the stitched-together fields from each individual unit cell. Yet even at this simple level of construction, one can already identify a source of error that must emerge. The stitched-together fields provably are not exact solutions of Maxwell's equations. The most favorable assumption is to take the outgoing fields to be the closest projection of the stitched-together fields onto a complete Maxwell basis. Once one makes this step, however, then it becomes clear that some efficiency must have been lost, as the projected fields will not propagate to the focal point identically to how the original fields would have. The resulting efficiency is an upper bound to the focusing efficiency of a unit-cell design. The detailed mathematical derivations of the various statements above are given in Ref. [11] while we provide a simpler version here. Imagine the best-case unit-cell design scenario, in which the unit cell response has perfect transmission and exactly the right phase in the correct outgoing diffraction order, for every library element. And, let us assume that the non-periodic neighboring cells do not alter this response at all. Then, one could write the (ideal, unit-cell-designed) fields at the outgoing plane of the entire design as

$$\boldsymbol{E}^{\text{unit,ideal}}(\boldsymbol{x}) = \sum_{m=1}^{N_d N_{\text{unit}}} c_{m,\Lambda} \boldsymbol{u}_{m,\Lambda}(\boldsymbol{x}), \tag{2}$$

where the $\boldsymbol{u}_{m,\Lambda}(\boldsymbol{x})$ are unit-cell diffraction basis functions with m denoting a specific diffraction order, N_d is the number of diffraction orders, N_{unit} is the number of unit cells, and Λ is the periodicity of the unit cell. (For subwavelength unit cells, with $\Lambda < \lambda$, there is only one propagating diffraction order, the zeroth order). The coefficients $c_{m,\Lambda}$ are found by time-reversing (phase conjugating) the field of a dipole at the focal point.

Even given all of the ideal assumptions mentioned above, there still must be error that arises in the expression of Eq. (2). The fields given by the collection of basic functions may not even be continuous, let alone proper solutions of Maxwell's equations. We can take the field of Eq. (2) and project it into any complete basis of Maxwell's equations (e.g. plane waves). After project, we compute the focusing efficiency of the projected fields. This focusing efficiency will represent a theoretical upper limit to what is possible via unit-cell design. Any additional deviations from the assumptions above (perfect unit cells, no neighbor interactions, etc.) will lead to further performance degradation.

The results of this procedure are shown in Figure 2. The exit plane fields show how the optimal stitched-together ("ideal") fields are modified, just moderately, into the nearest projection of Maxwell fields. These differences can be more visibly significant in the focal plane, where, for parameters matching those of the later metalenses ($\text{NA} = 0.6$, unit-cell period $= 0.0987\lambda$), the maximum focusing efficiency ($\eta = \frac{P_{\text{3rd,zero}}}{P_{\text{inc}}}$) of a unit-cell design is about 83.8%. This calculation is comparable to any simulations or fabrication results using a 40-nm-width unit cell in this work, while the fabricated EIM metalens uses 170-nm-width unit-cells. In the same theoretical upper limit calculation, the maximum focusing efficiency of a unit-cell design is about 63% for a unit-cell period of 170-nm-width.

2.3 Inverse design of metalenses

Inverse design, with "adjoint"-based optimization, computes gradients with respect to all structural degrees of freedom within forward and backward simulations [11, 12, 14, 33]. Its origin can be traced back to circuit theory, control theory [34], quantum dynamics [35], and deep learning [36–38]. It has been used in nanophotonics [12, 13, 39] for emerging applications such as tunable metasurfaces [16], solar cells [40], waveguide demultiplexers [13, 39], photolithographic mask [41], and CMOS image sensors [42, 43]. Inverse design requires definition of a FoM and its design parameters. In this work, we confine our design space to axisymmetric geometries for faster computations; we can expect this assumption to have little to no effect on the ultimate focusing efficiency, since the ideal focusing functionality can be expected to exhibit axial symmetry as well. We use the full-wave finite difference time domain (FDTD) method [44] to avoid assumptions, potentially ending up with focusing efficiency drops in unit-cell approaches [11], in the design procedure. For the FOM, we use intensity maximization at the focal point, which generally corresponds to the maximization of focusing efficiency in the far field [45, 46]. Our FoM is given by

$$\mathcal{F} = |\mathbf{E}(x_0)|^2 \tag{3}$$

where x_0 is a focal spot at the given lens dimension and NA, as shown in Figure 1(b) and (c). Required computations can be further optimized through near-field to far-field transformation [14] or planewave order decomposition [11]; however, the benefits of applying these techniques are marginal in high-NA metalenses due to a short focal length. Each geometrical parameter is a nanoring with 40 nm width and 860 nm height. The radius of a nanoring is defined as $40 \cdot N$ nm (N is the index of geometry parameter). Our metalens design consists of a SiO_2 substrate (measured refractive index of 1.47 at 405 nm wavelength), 100-nm-thick ITO (measured refractive index of 2.10 at 405 nm wavelength), and 680-nm-thick electron-beam resist CSAR-62 (measured refractive index of 1.59 at 405 nm wavelength). For the geometrical degrees of freedom of nanorings, we confine the density of CSAR-62 between 0 and 1 at every nanoring (40-nm-width, 680-nm-height), and then add penalization functions [18, 19] to the FoM in Eq. (3) to ensure a binary-material constraint. A circularly-polarized plane wave is excited at the substrate, and then $E_\rho^{\text{dir}}(x_0), E_\phi^{\text{dir}}(x_0)$ at the focal point are calculated for the adjoint simulation. A conjugated $E_\rho^{\text{dir}}(x_0), E_\phi^{\text{dir}}(x_0)$ field back-propagates with an adjoint source given by

$$\mathbf{J}_{\text{adj}} = -i\omega \mathbf{P}_{\text{adj}} = -i\omega \partial \mathcal{F} / \partial \mathbf{E} \tag{4}$$

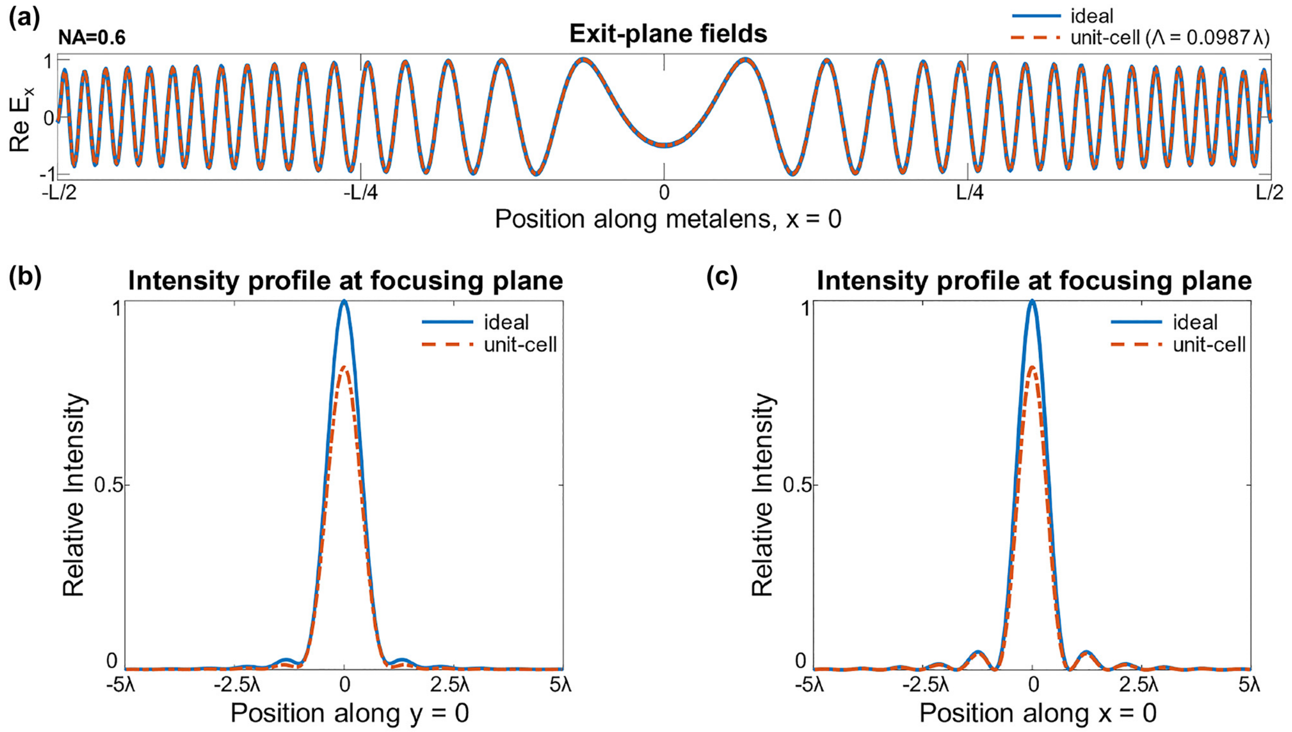

Figure 2: Theoretical upper bounds to the single-frequency metalens designed by a unit-cell approach. (a) The exit fields of a unit cell design (red) cannot exactly match those of the ideal focusing metalens (blue). (b, c) These field differences at the exit planes lead to reduced intensities at the focal point, as plotted along both transverse directions. The theoretical upper limit prediction shows that the 40-nm-width ($\lambda = 0.987$) unit-cell based design can have up to 83.8% focusing efficiency ($\eta = \frac{\mathrm{P_{3rd,zero}}}{\mathrm{P_{inc}}}$).

where **P** is an source dipole density [12]. At the design space, derivatives of the FoM can be calculated via $\partial\mathcal{F}/\partial\varepsilon(\mathbf{x}) = \mathrm{Re}\left[\mathbf{E}_{\mathrm{dir}}(\mathbf{x}) \cdot \mathbf{E}_{\mathrm{adj}}(\mathbf{x})\right]$, where $\mathbf{E}_{\mathrm{dir}}$ is a direct electric field calculated via the forward simulation and $\mathbf{E}_{\mathrm{adj}}$ is an electric field obtained via the adjoint simulation [11, 12]. The transmission normalized focusing efficiency, in this work, is defined as

$$\eta = \frac{P_{\mathrm{3rd,zero}}}{P_{\mathrm{trans}}} \qquad (5)$$

where $P_{\mathrm{3rd,zero}}$ is the electromagnetic power within the third zeroes of the Airy disk area while P_{trans} denotes a transmitted electromagnetic power at the metalens surface. This definition of the transmission normalized focusing efficiency is different than that in an imaging application where focusing efficiency is generally defined as $\eta = \frac{P}{P_{\mathrm{inc}}}$. This difference arises from the nature of the lithography application, where the incidence power can be increased easily; thus, we can ignore reflected power, unlike imaging applications where the incoming light is relevant.

With our approach, a combination of forward and backward simulations takes approximately 330 s, including pre, and post-processing of adjoint derivatives on 64 cores in our workstation (AMD Ryzen Threadripper PRO 3995WX, 2.7 GHz processors). The FoM rapidly increases then converges in about the 25th iteration, after which the penalization [11, 18] transforms the grayscale refractive indices to binary values (i.e., air or CSAR-62), which takes around another 120 iterations as shown in Figure 3. The averaged penalization factor over the designable region ($\sum|\xi(\mathbf{r})|$) is set to a relative value compared to the sum of derivatives of FoM with respect to electric fields ($\sum|\frac{\partial \mathrm{FoM}}{\partial \mathbf{E}(\mathbf{r})}|$). By taking relative penalization factors into account, the design parameters can be updated more systematically (e.g., adjoint derivatives dominant: 1st ∼ 40th iterations, penalization factor dominant: 40th ∼ 150th iterations). The penalization process can be further tuned with Gaussian filters [47], density filters [48, 49] or level-set methods [50, 51].

3 Results

In this section, we design and fabricate effective-medium-based and inverse-designed metalenses. Both approaches assume that we can have control within a fabrication constraint of 40 nm (width) × 680 nm (height) on the size of nanoring. To explore a computational upper limit of the given design problem, the effective index modulation (EIM) approach can have a material degree of freedom in grayscale (gradient refractive index) as shown in Figure 4(a). The inverse design only allows a refractive index of 1.59 (CSAR-62) and 1.00 (air). The refractive indices of each nanoring were chosen by matching a required phase profile $\Delta\phi(r) = \frac{2\pi}{\lambda}(f - \sqrt{r^2 + f^2})$ and transmitted phase $\phi_{\mathrm{trans.}}(r) = \frac{2\pi n(r) h}{\lambda}$ of the gradient index material, where λ is the design wavelength (405 nm), r is the coordinate of each nanoring, f is the focal length (97 um), and $n(r)$ is a refractive index at the radius r. The minimum grid spacing of the

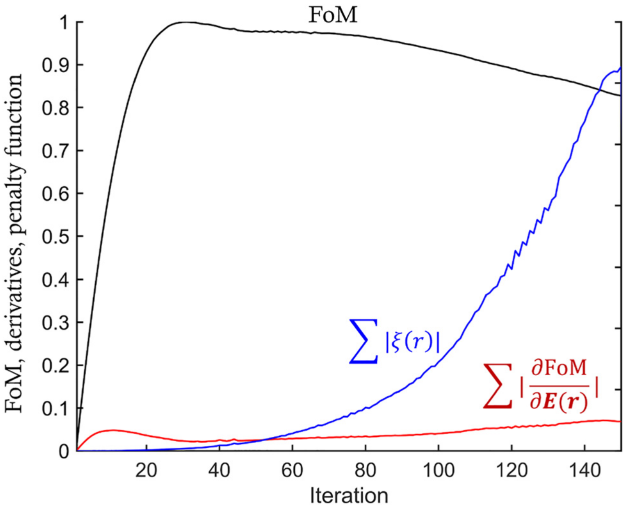

Figure 3: Plot of the Figure-of-Merit (FoM), the sum of derivatives of FoM with respect to electric fields ($\sum|\frac{\partial \mathrm{FoM}}{\partial \mathbf{E}(\mathbf{r})}|$), and the penalization factors ($\sum|\xi(\mathbf{r})|$) versus inverse design iterations. For the first 50 iterations, derivatives of FoM dominate over penalization factors, which mean that the adjoint derivatives are the main driving factors for updating design parameters. After the 50th iteration, penalization factors gradually dominate the adjoint derivatives, making the design parameters converge to the binary level.

simulations was enforced to 40 nm to satisfy fabrication constraints, and then the final design was validated with a finer grid spacing to guarantee the convergence of the simulation. A circularly-polarized planewave was excited at the SiO_2 substrate, as shown in Figure 1(b). The incident power was calculated and then compared against a transmitted power at the plane 2λ above the metalens to compute transmission efficiency shown in Figure 4(e) (transmission normalized focusing efficiency is defined in Eq. (5)). We also verify robustness of our final design by simulating fabrication imperfections, and maintain relatively high efficiency for moderate fabrication imperfections as shown in Figure 6 (Appendix).

3.1 Effective index modulated (EIM) metalens

We designed metalenses for the ZPAL system at 0.57 NA with 135 μm diameter (97 μm focal length). Figure 4(a) shows the optimized metalens with the EIM approach. Since it is infeasible to simulate unit-cell-based metalenses in 3D due to a significant computational burden, we approximate a unit-cell design approach with an effective index-modulated structure with cylindrical symmetry. It has a refractive index gradient varying from 1.00 to 1.59. The effective-medium-based metalens shows 81.63% efficiency at 97 μm focal plane while its transmission normalized focusing efficiency is 79.98%. Compared to the diffraction-limited focusing efficiency (total power within the third zero of the Airy pattern ≈93.8% [52]), our EIM-based metalens shows relatively high efficiency even with rough fabrication constraints (40 nm × 680 nm minimum feature size) as shown in Figure 4(d).

We employ a unit-cell design approach for the realization of EIM metalens. First, metalenses were divided into an integral number of subzone rings, each about 170 nm wide (i.e., small compared to the optical wavelength of 405 nm). Next, each such subzone ring was divided into an integral number of circumferential cells, about 170 nm wide, with their circumferential starting angle randomly varied, as shown in Figure 4(f, g). The width of the unit-cell increases from 40 nm (simulation) to 170 nm (fabrication), which may result in an inevitable efficiency drop. Then, the conversion of grayscale refractive indices (EIM) to a combination of rectangular holes (air) and surrounding materials (CSAR-62) could potentially cause an additional efficiency drop from the simulated result in Figure 4(a). Rectangular holes were programmed into each such cell, as illustrated in Figure 4(f), with their azimuthal angle rotated in accordance with the cell azimuth. Rotational EIM-based metalenses were fabricated on borosilicate-glass samples (obtained from an online vendor) coated with 185 nm of indium-tin-oxide (ITO). The substrates were cleaned using UV ozone and an adhesion promoter spun on at 4000 RPM. 680-nm-thick CSAR resist was achieved at 1250 RPM. Substrates were baked on the hotplate at 150C for 2 min. Finally, the water-soluble conducting film was spun at 4000 RMP. Beamer proximity-effect-correction (PEC) software was used to pattern at 1 nA beam current, 2.5 nm stepping distance, 125 keV beam energy, and 15 ns dwell time. This yields a base dose of 240 $\mu C/cm^2$.

The fabricated metalens shows transmission normalized focusing efficiency of 48%. We assume that the conversion from 40-nm-width unit-cells to 170-nm causes a significant efficiency drop as expected in our previous work [11]. The fabrication imperfection were relatively mild compared to the inverse-designed metalens, discussed in the next section. As in many other publications [6, 7, 9, 53], it is hard to validate unit-cells-based metalens design in full-wave simulations due to a computational expense of a large area metalens. In the next section, we fully utilize both the accuracy of the full-wave simulation and the fast convergence of the inverse design algorithm to discover a 'full-wave validated' inverse-designed metalens.

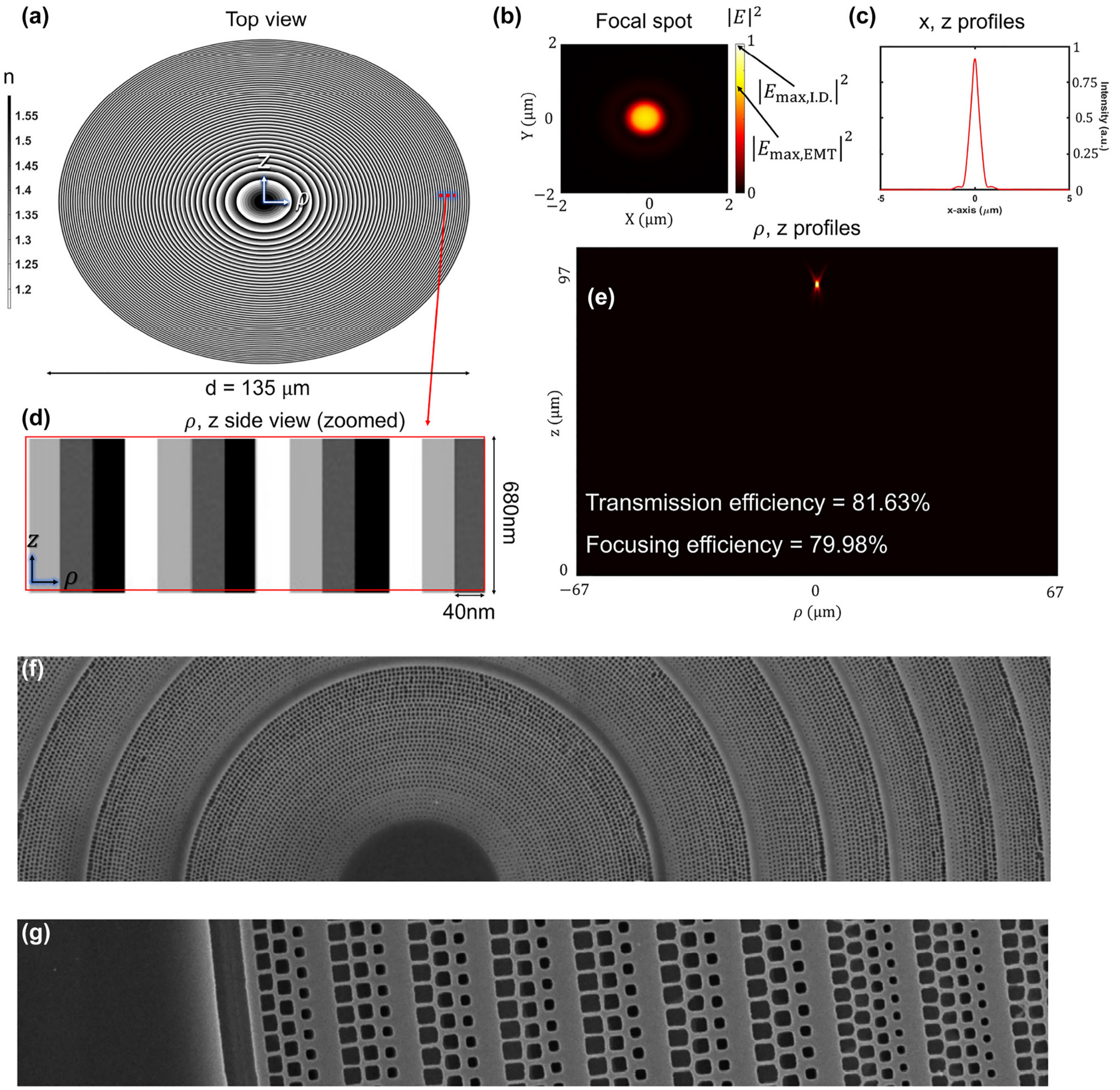

Figure 4: Effective-index modulation (EIM)-based theoretical lens design. (a) Top view of the EIM-based lens for NA = 0.57 ($f = 97$ μm, $d = 135$ μm). (b) Focal spot intensity profile in the transverse (xy) plane at the focal distance of $f = 97$ μm. (c) A transverse cut of the focal spot intensity profile in simulation. The full width at half intensity maximum (FWHM) of the spot is 0.571 μm. (d) ρ, z side view (zoomed) of the EIM-lens. Thickness is 680 nm, covering 2-π phase change with a maximum refractive variation of 1.0–1.59. Each pixel occupies 40 nm radial space in cylindrical symmetry, corresponding to our fabrication resolution. (e) Longitudinal (ρ, z) intensity profile where transmission efficiency and transmission normalized focusing efficiency are 81.63% and 79.98%, respectively. (f) SEM of the EIM metalens which experimentally had a focal efficiency of about 48%. (g) SEM of the outer zones of the EIM metalens.

3.2 Inverse-designed metalens

In this subsection we optimize and fabricate an axisymmetric inverse-designed metalens for ZPAL. To avoid the efficiency drops on: (1) the unit-cell conversions (40-nm-width to 170-nm-width) and (2) gradient refractive indices to EIM structure conversions, we choose two material system (air and CSAR-62), and fabrication-compatible minimum feature size (40 nm × 680 nm). In this way, the result of the full-wave simulation does not suffer efficiency drops on such conversions. Fabrication perfection is required to maintain the simulated efficiency.

As shown in Figure 5(d), the inverse-designed axisymmetric metalens only has refractive index values of 1.0 and 1.59 at 405 nm wavelength. Figure 5(e) shows longitudinal (ρ, z) intensity profile where transmission efficiency and transmission normalized focusing efficiency are 87.46% and 85.50%, respectively. Surprisingly, the transmission normalized focusing efficiency of our inverse-designed metalens exceeds that of the gradient refractive indices lens by 5.2% with only a single material. We analyze the main reasons for high efficiency in two ways. First, unit-cell-based design approaches are relatively vulnerable to rapidly oscillating phase and amplitude profiles because the rapid change of phase/amplitude between neighboring unit cells may fail to create a desired continuous wavefront [11]. In the worst case scenario, only one or two unit cells are placed to model continuous variation of 2-π phase shift in our 0.57 NA EIM-based metalens design. Second, the inverse design technique inherently converges to the local optimum quickly. Therefore, inverse design may often get stuck on bad local optima because of complicated optimization

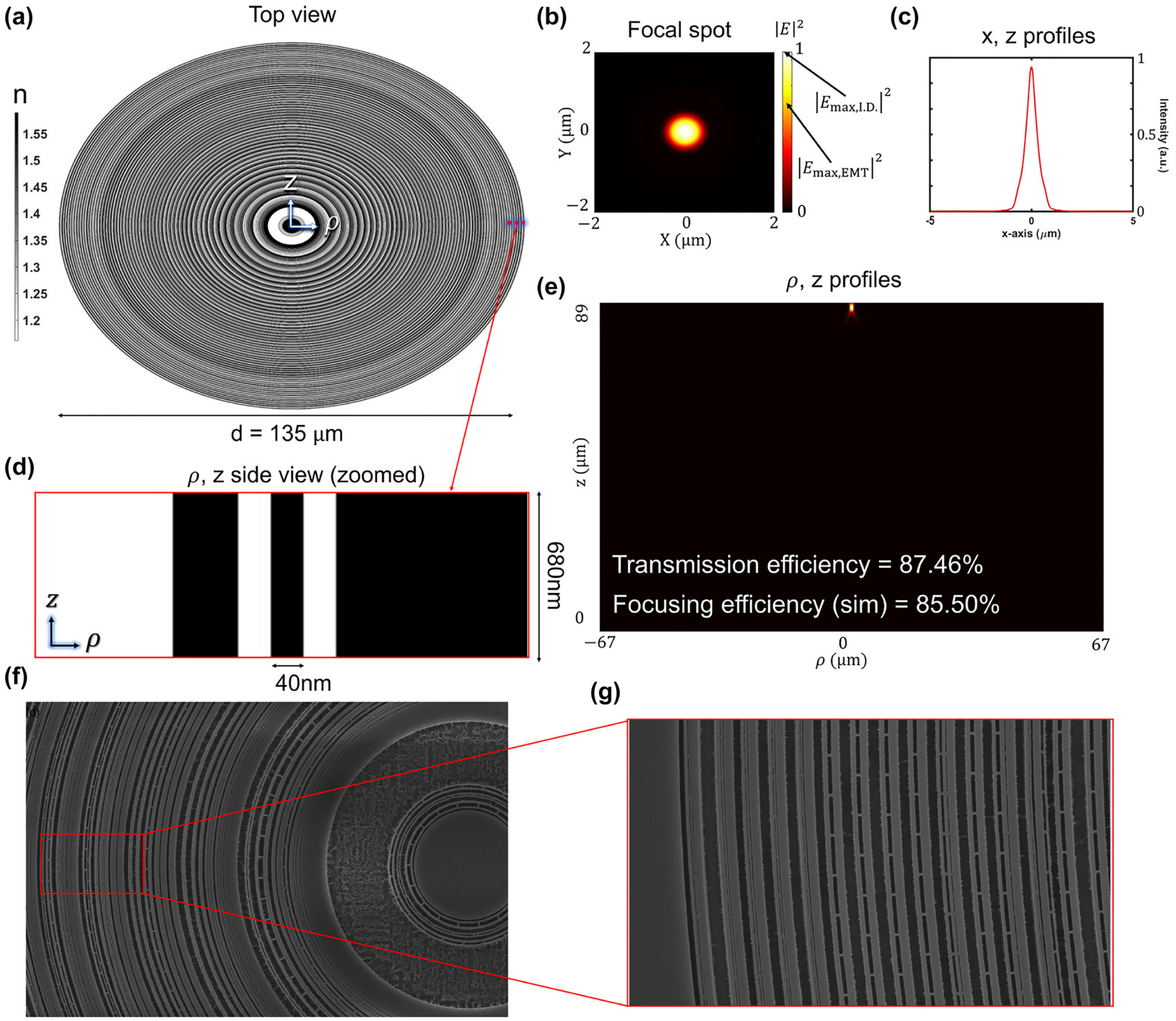

Figure 5: Inverse-designed metalens. (a) Top view of the inverse-designed metalens for NA = 0.60 ($f = 89$ µm, $d = 135$ µm). (b) Focal spot intensity profile in the transverse (xy) plane at the focal distance of $f = 89$ µm. (c) A transverse cut of the focal spot intensity profile in simulation. The full width at half intensity maximum (FWHM) of the spot is 0.611 µm. (d) Zoomed ρ, z side view of the metalens. Thickness is 680 nm, and the minimum design/fabrication resolution is 40 nm in the radial direction with a material refractive index of 1.59. (e) Longitudinal (ρ, z) intensity profile where transmission efficiency and transmission normalized focusing efficiency are 87.46%, 85.50%, respectively. (f) Top view of the SEM image for a fabricated inverse-designed metalens. (g) Close-up image of the SEM, showing tiny bridges to support the high-aspect-ratio (680/40 ≈ 17) nanoring-structure.

issues [54]. However, our design problem, 0.6 NA metalens with 135 μm diameter, has relatively strict fabrication constraints leading to simple optimization problems in local areas. At the same time, it is still a large-area optimization with a total of 6750 design parameters. In other words, the design parameters are strongly correlated in local areas but nearly uncorrelated among largely-spaced parameters. Therefore, the summation of the local solutions found by our inverse design could be close to the upper limit.

To fabricate an inverse-designed metalens, a dose of 194 μC/cm^2 was applied, and supporting bridges were added to avoid collapsing the high-aspect-ratio nanoring structure. However, despite nano-bridge support, some material adhered to the substrate surface while parts of the upper layers were torn away due to tensile stress. The measured transmission normalized focusing efficiency dropped to 36%. It could be improved by changing the material to a high index material, which reduces the aspect ratio to cover 2π phase range.

4 Conclusions

We designed a full-wave validated axisymmetric metalens for the ZPAL system via inverse design which shows greater focusing efficiency than gradient indices theoretical lenses. The normalized transmission efficiency of the optimized metalens (85.50%) slightly exceeds that of the theoretical upper limit for the given dimension of the unit-cell-based approach as shown in Figure 2, meaning that axisymmetric metalens designs can have higher efficiency than conventional convex microlens and unit-cell designed metalens. ZPAL with metalens-array implementation may open a new way of fabricating micro/nano-pattern at a faster and lower cost compared to nano-patterning electron beam lithography which often suffers high-cost and slow writing time. Furthermore, our design approach can be extended to much larger metalens designs since both the full-wave Maxwell solver (Meep) [44] and adjoint optimization algorithm used in this work scale well with an increasing number of computing cores. Emerging fast-solver techniques such as integral equation-based Maxwell solver combined with adjoint optimization can also scale this approach to macroscopic length scales. Adjoint optimization itself has already proven its scalability in aerodynamic design [55], where more than a billion voxels were simulated to find optimal distributions.

Acknowledgement: The e-beam lithography was done by Mark Mondol at MIT. The computational resource was provided by KISTI (KSC-2022-CRE-0263).

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: This work was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (MSIT) (NRF-2022K1A3A1A91093776) and also under the High-Potential Individuals Global Training Program (IITP-2022-0-0154896), and Artificial Intelligence Graduate School Program (No. 2020-0-01373, Hanyang University) supervised by the IITP (Institute for Information and Communications Technology Planning and Evaluation), and also supported by the MOTIE (Ministry of Trade, Industry & Energy 1415180303 and KSRC (Korea Semiconductor Research Consortium) 20019357 support program for the development of the future semiconductor device.
Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

Appendix A: Robustness of the optimized metalens

The robustness of the optimized metalens was tested by adjusting the permittivity of the finite pixels. The fixed pixel size makes it relatively hard to model a slightly delated or eroded structure in the finite difference scheme. Therefore, we utilize grayscale permittivity to model slightly delated or eroded structures. This method was first used in the subpixel smoothing in the FDTD [44] where interpolating the permittivity over the material surface can significantly improve the stair-case error of the FDTD. As shown in Figure 6, our optimized metalens has relatively robust for the slightly delated or eroded structures.

Appendix B: Optimized metalens data

Geometry parameters for the inverse-designed metalens is listed in this section. Each grid occupies a nanoring of the 40-nm-width, 680-nm-height and the radius corresponding to $40 \cdot N$ nm (N is the index of geometry parameter). The value “1” means permittivity of CSAR-64 at 405 nm, while the value “0” means air: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 1 1 0 0 0 0 0 1 1 0 0 0 0 0 0 1 0 0 1 0 0 0 0 1 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 0 1 1 0 1 1 1 1 1 0 0 1 1 1 0 0 1 1 0 1 0 1 1 1 0 0 1 1 1 1 0 0 0 0 0 1 1 0 1 1 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 1 1 1 1

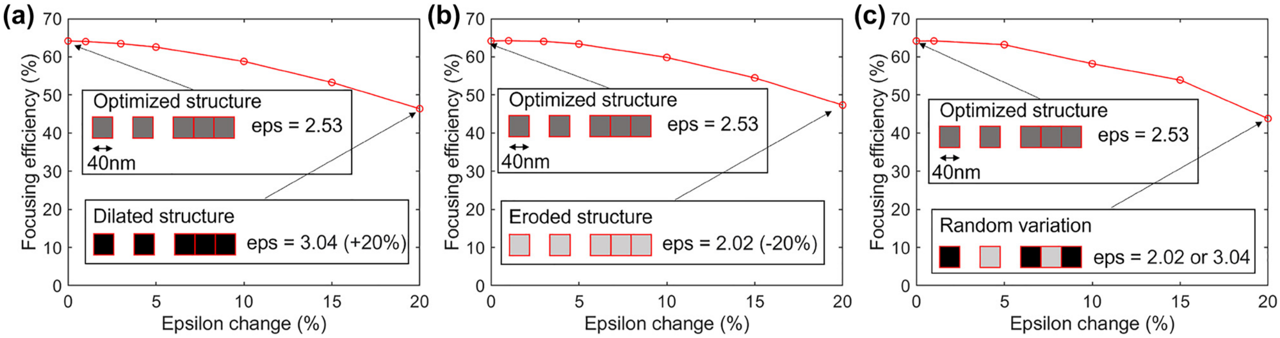

Figure 6: Robustness test with (a) dilated structures, (b) eroded structures, and (c) random variations in the optimized metalens shown in Figure 5. The permittivities of Yee's pixels were modified to introduce relatively small structural variation. The permittivity change effectively represents eroded or dilated structures via subpixel smoothing effect [44]. A ten percent change in the permittivity degrades the transmission normalized focusing efficiency by 5–10 percent. It implies that the optimized metalens maintains its efficiency over slight fabrication error.

1 1 1 1 1 1 0 1 1 1 1 1 1 0 1 1 0 1 1 1 0 1 1 1 0 0 1 1 1 0 1 1 0 1 1 1 0 0
0 1 1 1 1 0 0 0 0 0 1 0 1 1 0 0 0 0 0 0 0 1 1 0 1 0 0 0 0 0 0 0 0 0 1 1 1
1 1 1 1 1 0 0 0 1 1 1 1 1 1 1 0 0 1 0 1 1 1 1 0 0 0 1 1 1 1 0 0 0 1 1 1 0 0
0 0 0 1 1 0 1 0 0 0 0 0 1 0 1 1 0 0 0 0 0 1 1 1 1 1 1 1 1 0 0 1 1 1 1 1 1
1 0 0 1 1 1 1 0 1 0 1 1 1 1 0 0 1 0 1 1 0 0 0 0 1 1 0 0 1 0 0 0 1 1 0 0 0
0 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 0 1 0 1 1 0 1 1 1 0 0 0 0 1 1 0 0 1
0 1 0 0 0 1 0 0 1 0 0 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 1 0 1 1 0 0 1 1 1 1 0
0 1 0 0 1 1 1 0 0 0 0 0 1 1 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 0 1 0 1 1 1 1 0
1 0 1 1 1 0 0 1 1 1 0 1 0 0 0 1 1 0 0 0 0 1 0 0 0 0 1 1 1 1 1 1 1 1 1 0 1 1
0 1 1 1 1 0 1 1 0 0 1 1 1 0 0 1 0 0 0 0 1 1 0 0 0 0 0 0 1 1 1 1 1 1 1 1 0 0
1 1 1 1 0 1 1 0 0 0 0 1 1 0 1 0 0 0 0 0 1 0 1 1 0 0 0 0 1 1 1 1 1 1 1 1 0 0
1 1 1 1 1 0 0 0 1 1 1 0 1 1 0 0 0 0 0 1 1 0 0 0 0 0 1 1 1 1 1 1 1 1 0 1 1 0
1 1 1 1 0 1 0 1 1 1 0 0 0 0 1 1 0 0 1 0 0 0 0 1 1 1 1 1 1 1 1 0 1 1 0 1 1 1
0 1 0 1 1 1 0 0 0 0 1 0 1 1 0 0 0 0 1 1 1 1 1 1 1 1 0 1 1 1 0 1 1 1 0 0 0 1
1 0 1 0 1 0 0 0 1 0 0 0 1 1 1 1 1 1 1 0 0 1 1 1 1 0 0 0 1 1 0 1 0 0 0 0 0
1 1 0 0 0 1 1 1 1 1 1 1 0 0 0 1 1 1 1 0 0 1 1 1 0 0 0 0 0 1 1 0 0 0 1 1 1
1 1 1 1 0 0 1 1 0 1 1 0 0 1 1 0 0 0 1 1 0 0 0 0 0 1 1 1 1 1 1 1 1 0 0 1 1 1
0 1 0 0 0 1 0 0 1 1 0 0 0 0 0 1 1 1 1 1 1 1 0 1 1 1 0 1 0 1 1 0 0 1 0 1
1 0 0 0 0 0 1 1 1 1 1 1 0 0 1 1 1 0 1 0 0 1 1 0 0 1 1 0 0 0 0 0 1 1 1 1 1
1 1 0 1 0 1 1 0 0 1 1 1 0 0 0 1 1 0 0 0 0 1 1 1 1 1 1 1 0 0 1 1 1 0 1 1 0 1
0 0 0 1 1 0 0 0 0 1 1 1 1 1 1 1 0 0 1 1 1 1 0 0 1 0 1 0 1 0 1 0 0 0 1 1 1 1
1 1 0 0 1 1 1 1 0 0 0 0 1 1 1 0 0 0 0 0 0 1 1 1 1 1 1 0 1 0 1 1 0 1 1 0 0
0 1 0 1 0 0 0 0 1 1 1 1 1 1 0 1 0 1 1 1 0 0 1 0 0 1 1 0 0 0 0 1 1 1 1 1 1
0 0 1 1 0 1 0 0 0 1 0 1 0 0 0 0 1 1 1 1 1 1 0 0 1 1 1 0 0 0 0 1 1 0 1 0 0
0 0 1 1 1 1 1 1 0 0 1 1 1 0 0 1 1 0 0 0 1 0 0 0 1 1 1 1 1 1 0 0 1 1 1 0 0
0 0 1 1 0 0 0 0 0 1 1 1 1 1 1 1 0 0 1 1 1 0 0 1 1 0 1 0 0 0 0 1 1 1 1 1 1 0
1 1 1 0 0 0 0 1 1 0 0 0 0 1 1 1 1 1 1 0 1 1 1 0 0 0 0 1 1 1 0 0 0 0 1 1
1 1 1 1 0 1 1 1 0 0 1 0 1 0 0 0 0 0 1 1 1 1 1 1 0 1 1 1 0 1 0 1 0 0 0 1
0 0 1 1 1 1 1 1 0 1 1 1 0 0 1 0 1 0 0 1 0 0 1 1 1 1 1 1 0 1 1 1 0 0 0 1 1 0
0 1 0 0 1 1 1 1 1 1 0 1 1 1 0 1 0 1 0 1 0 0 0 0 1 1 1 1 1 0 1 0 1 1 0 1 1 0
0 0 0 0 0 0 1 1 1 1 1 0 0 1 1 0 1 1 0 0 0 1 0 0 0 0 1 1 1 1 1 0 0 0 1 1 1
0 0 0 0 0 0 0 0 1 1 1 1 1 0 0 0 1 1 1 0 0 0 0 0 1 0 0 1 1 1 1 1 0 0 1 0 1
1 0 0 0 0 0 1 1 1 1 1 1 0 1 1 1 0 0 0 1 1 0 0 0 0 0 1 1 1 1 1 1 0 0 0 1 1
1 0 0 0 0 0 0 1 1 1 1 1 0 1 1 0 0 1 1 0 0 0 0 0 0 1 1 1 1 1 1 0 0 0 1 1 1
0 0 0 0 0 0 1 1 1 1 1 1 0 0 0 1 1 1 0 0 0 0 0 0 0 1 1 1 1 1 0 0 0 1 1 1 0
0 0 0 0 0 0 1 1 1 1 1 0 0 0 1 1 1 0 0 0 0 0 1 1 1 1 1 1 0 0 1 0 1 1 0 0 0
0 0 0 1 1 1 1 1 0 1 0 1 1 0 0 0 0 0 1 1 1 1 1 1 0 1 0 1 1 0 0 0 0 0 1
1 1 1 1 1 0 1 0 1 1 1 0 0 0 0 0 0 0 1 1 1 1 1 0 0 1 0 1 1.

References

[1] H. I. Smith, R. Menon, A. Patel, D. Chao, M. Walsh, and G. Barbastathis, "Zone-plate-array lithography: a low-cost complement or competitor to scanning-electron-beam lithography," *Microelectron. Eng.*, vol. 83, nos. 4–9, pp. 956–961, 2006.

[2] D. Chao, A. Patel, T. Barwicz, H. I. Smith, and R. Menon, "Immersion zone-plate-array lithography," *J. Vac. Sci. Technol., B: Microelectron. Nanometer Struct. Process., Meas., Phenom.*, vol. 23, no. 6, pp. 2657–2661, 2005.

[3] H. I. Smith, "A proposal for maskless, zone-plate-array nanolithography," *J. Vac. Sci. Technol., B: Microelectron. Nanometer Struct. Process., Meas., Phenom.*, vol. 14, no. 6, pp. 4318–4322, 1996.

[4] L. Pang, Y. Liu, and D. Abrams, "Inverse lithography technology (ILT): what is the impact to the photomask industry?," in *Photomask and Next-Generation Lithography Mask Technology XIII*, vol. 6283, Yokohama, Japan, SPIE, 2006, pp. 233–243.

[5] P. Wang, N. Mohammad, and R. Menon, "Chromatic-aberration-corrected diffractive lenses for ultrabroadband focusing," *Sci. Rep.*, vol. 6, no. 1, pp. 1–7, 2016.

[6] M. Khorasaninejad, A. Y. Zhu, C. Roques-Carmes, et al., "Polarization-insensitive metalenses at visible wavelengths," *Nano Lett.*, vol. 16, no. 11, pp. 7229–7234, 2016.

[7] M. Khorasaninejad, W. T. Chen, R. C. Devlin, J. Oh, A. Y. Zhu, and F. Capasso, "Metalenses at visible wavelengths: diffraction-limited focusing and subwavelength resolution imaging," *Science*, vol. 352, no. 6290, pp. 1190–1194, 2016.

[8] P. Lalanne and P. Chavel, "Metalenses at visible wavelengths: past, present, perspectives," *Laser Photonics Rev.*, vol. 11, no. 3, p. 1600295, 2017.

[9] S. Shrestha, A. C. Overvig, M. Lu, A. Stein, and N. Yu, "Broadband achromatic dielectric metalenses," *Light: Sci. Appl.*, vol. 7, no. 1, p. 85, 2018.

[10] W. T. Chen, A. Y. Zhu, V. Sanjeev, et al., "A broadband achromatic metalens for focusing and imaging in the visible," *Nat. Nanotechnol.*, vol. 13, no. 3, p. 220, 2018.

[11] H. Chung and O. D. Miller, “High-NA achromatic metalenses by inverse design,” *Opt. Express*, vol. 28, no. 5, pp. 6945–6965, 2020.

[12] O. D. Miller, “Photonic design: from fundamental solar cell physics to computational inverse design,” Ph.D. thesis, University of California, Berkeley, 2012.

[13] L. Su, A. Y. Piggott, N. V. Sapra, J. Petykiewicz, and J. Vuckovic, “Inverse design and demonstration of a compact on-chip narrowband three-channel wavelength demultiplexer,” *ACS Photonics*, vol. 5, no. 2, pp. 301–305, 2017.

[14] Z. Lin, V. Liu, R. Pestourie, and S. G. Johnson, “Topology optimization of freeform large-area metasurfaces,” *Opt. Express*, vol. 27, no. 11, pp. 15765–15775, 2019.

[15] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vuckovic, and A. W. Rodriguez, “Inverse design in nanophotonics,” *Nat. Photonics*, vol. 12, no. 11, p. 659, 2018.

[16] H. Chung and O. D. Miller, “Tunable metasurface inverse design for 80% switching efficiencies and 144 angular deflection,” *ACS Photonics*, vol. 7, no. 8, pp. 2236–2243, 2020.

[17] H. Chung and S. V. Boriskina, “Inverse design of a single-frequency diffractive biosensor based on the reporter cleavage detection mechanism,” *Opt. Express*, vol. 29, no. 7, pp. 10780–10799, 2021.

[18] M. Stolpe and K. Svanberg, “On the trajectories of penalization methods for topology optimization,” *Struct. Multidiscip. Optim.*, vol. 21, no. 2, pp. 128–139, 2001.

[19] M. M. Neves, O. Sigmund, and M. P. Bendsøe, “Topology optimization of periodic microstructures with a penalization of highly localized buckling modes,” *Int. J. Numer. Methods Eng.*, vol. 54, no. 6, pp. 809–834, 2002.

[20] J. J. Reuther, A. Jameson, J. J. Alonso, M. J. Rimllnger, and D. Saunders, “Constrained multipoint aerodynamic shape optimization using an adjoint formulation and parallel computers, part 1,” *J. Aircr.*, vol. 36, no. 1, pp. 51–60, 1999.

[21] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, New York, NY, USA, Academic Press, 2014.

[22] M. Li, Z. Li, and J. Yao, “Photonic generation of precisely pi phase-shifted binary phase-coded microwave signal,” *IEEE Photonics Technol. Lett.*, vol. 24, no. 22, pp. 2001–2004, 2012.

[23] R. Menon, D. Gil, and H. I. Smith, “Experimental characterization of focusing by high-numerical-aperture zone plates,” *JOSA A*, vol. 23, no. 3, pp. 567–571, 2006.

[24] D. W. Wilson, P. D. Maker, R. E. Muller, P. Z. Mouroulis, and J. Backlund, “Recent advances in blazed grating fabrication by electron-beam lithography,” in *Current Developments in Lens Design and Optical Engineering IV*, vol. 5173, San Diego, California, USA, Society of Photo-Optical Instrumentation Engineers (SPIE), 2003, pp. 115–126.

[25] E. G. Loewen, M. Neviére, and D. Maystre, “Grating efficiency theory as it applies to blazed and holographic gratings,” *Appl. Opt.*, vol. 16, no. 10, pp. 2711–2721, 1977.

[26] K. Miyamoto, “The phase Fresnel lens,” *JOSA*, vol. 51, no. 1, pp. 17–20, 1961.

[27] N. M. Ceglio and H. I. Smith, “Micro-Fresnel zone plates for coded imaging applications,” *Rev. Sci. Instrum.*, vol. 49, no. 1, pp. 15–20, 1978.

[28] K. Hassan, J. A. Dallery, P. Brianceau, and S. Boutami, “Integrated photonic guided metalens based on a pseudo-graded index distribution,” *Sci. Rep.*, vol. 10, no. 1, pp. 1–7, 2020.

[29] A. Alexopoulos, “Effective-medium theory of surfaces and metasurfaces containing two-dimensional binary inclusions,” *Phys. Rev. E*, vol. 81, no. 4, p. 046607, 2010.

[30] X. Ni, S. Ishii, A. V. Kildishev, and V. M. Shalaev, “Ultra-thin, planar, Babinet-inverted plasmonic metalenses,” *Light: Sci. Appl.*, vol. 2, no. 4, p. e72, 2013.

[31] G. Yoon, K. Kim, D. Huh, H. Lee, and J. Rho, “Single-step manufacturing of hierarchical dielectric metalens in the visible,” *Nat. Commun.*, vol. 11, no. 1, pp. 1–10, 2020.

[32] G. Yoon, K. Kim, S. U. Kim, S. Han, H. Lee, and J. Rho, “Printable nanocomposite metalens for high-contrast near-infrared imaging,” *ACS Nano*, vol. 15, no. 1, pp. 698–706, 2021.

[33] H. Chung, “Computational upper-limit of directional light emission in nano-LED via inverse design,” *Opt. Express*, vol. 30, no. 6, pp. 9008–9020, 2022.

[34] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, New York, NY, USA, John Wiley & Sons, 1962.

[35] M. Demiralp and H. Rabitz, “Optimally controlled quantum molecular dynamics: a perturbation formulation and the existence of multiple solutions,” *Phys. Rev. A*, vol. 47, no. 2, pp. 809–816, 1993.

[36] P. J. Werbos, *The Roots Of Backpropagation*, New York, NY, USA, John Wiley & Sons, 1994.

[37] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, p. 436, 2015.

[38] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[39] A. Y. Piggott, J. Lu, K. G. Lagoudakis, J. Petykiewicz, T. M. Babinec, and J. Vuckovic, “Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer,” *Nat. Photonics*, vol. 9, no. 6, pp. 374–377, 2015.

[40] C. M. Lalau-Keraly, S. Bhargava, O. D. Miller, and E. Yablonovitch, “Adjoint shape optimization applied to electromagnetic design,” *Opt. Express*, vol. 21, no. 18, pp. 21693–21701, 2013.

[41] S. H. Nam, M. Kim, N. Kim, et al., “Photolithographic realization of target nanostructures in 3D space by inverse design of phase modulation,” *Sci. Adv.*, vol. 8, no. 21, p. eabm6310, 2022.

[42] P. B. Catrysse, N. Zhao, W. Jin, and S. Fan, “Subwavelength Bayer RGB color routers with perfect optical efficiency,” *Nanophotonics*, vol. 11, no. 10, pp. 2381–2387, 2022.

[43] N. Zhao, P. B. Catrysse, and S. Fan, “Perfect RGB-IR color routers for sub-wavelength size CMOS image sensor pixels,” *Adv. Photonics Res.*, vol. 2, no. 3, p. 2000048, 2021.

[44] A. F. Oskooi, D. Roundy, M. Ibanescu, P. Bermel, J. Joannopoulos, and S. G. Johnson, “MEEP: a flexible free-software package for electromagnetic simulations by the FDTD method,” *Comput. Phys. Commun.*, vol. 181, no. 3, pp. 687–702, 2010.

[45] H. Shim, H. Chung, and O. D. Miller, “Maximal free-space concentration of electromagnetic waves,” *Phys. Rev. Appl.*, vol. 14, no. 1, p. 014007, 2020.

[46] K. Schab, L. Jelinek, M. Capek, and M. Gustafsson, “Upper bounds on focusing efficiency,” *Opt. Express*, vol. 30, no. 25, pp. 45705–45723, 2022.

[47] D. Sell, J. Yang, S. Doshay, R. Yang, and J. A. Fan, “Large-angle, multifunctional metagratings based on freeform multimode geometries,” *Nano Lett.*, vol. 17, no. 6, pp. 3752–3757, 2017.

[48] E. W. Wang, D. Sell, T. Phan, and J. A. Fan, “Robust design of topology-optimized metasurfaces,” *Opt. Mater. Express*, vol. 9, no. 2, pp. 469—482, 2019.

[49] F. Wang, J. S. Jensen, and O. Sigmund, “Robust topology optimization of photonic crystal waveguides with tailored dispersion properties,” *JOSA B*, vol. 28, no. 3, pp. 387—397, 2011.

[50] D. Vercruysse, N. V. Sapra, L. Su, R. Trivedi, and J. Vuckovic, “Analytical level set fabrication constraints for inverse design,” *Sci. Rep.*, vol. 9, no. 1, pp. 1—7, 2019.

[51] M. Mansouree and A. Arbabi, “Metasurface design using level-set and gradient descent optimization techniques,” in *2019 International Applied Computational Electromagnetics Society Symposium (ACES)*, IEEE, 2019, pp. 1—2.

[52] M. Born and E. Wolf, *Principles Of Optics: Electromagnetic Theory of Propagation*, *Interference and Diffraction of Light*, Oxford, England, Elsevier, 2013.

[53] B. H. Chen, P. C. Wu, V. C. Su, et al., “GaN metalens for pixel-level full-color routing at visible light,” *Nano Lett.*, vol. 17, no. 10, pp. 6345—6352, 2017.

[54] J. Park, S. Kim, D. W. Nam, H. Chung, C. Y. Park, and M. S. Jang, “Free-form optimization of nanophotonic devices: from classical methods to deep learning,” *Nanophotonics*, vol. 11, no. 9, pp. 1809—1845, 2022.

[55] N. Aage, E. Andreassen, B. S. Lazarov, and O. Sigmund, “Giga-voxel computational morphogenesis for structural design,” *Nature*, vol. 550, no. 7674, pp. 84—86, 2017.