LASER & PHOTONICS REVIEWS

www.lpr-journal.org

Toward High-Efficiency Ultrahigh Numerical Aperture Freeform Metalens: From Vector Diffraction Theory to Topology Optimization

*Di Sang, Mingfeng Xu, Mingbo Pu, Fei Zhang, Yinghui Guo, Xiong Li, Xiaoliang Ma, Yunqi Fu, and Xiangang Luo**

Optical metasurface is a 2D array of subwavelength meta-atoms that arbitrarily manipulates amplitude, polarization, and wavefront of incident light, offering great advantages in miniaturization and integration. However, the presuppositions hidden in conventional unit-cell-based design approach, including discrete phase sampling, local periodicity approximation, and normal response, imply that the responses and interactions of meta-atoms are almost impossible to be accurately modeled, thus impeding the effective design of high-efficiency ultrahigh numerical aperture (NA) metalens. Here, based on vector diffraction theory and plane wave expansion method, theoretical limitation of metalens efficiency is comprehensively investigated. It is identified that for high-NA metalens, theoretical focusing efficiency is limited by diffraction capability of vector field, while evanescent wave attenuation dominates theoretical efficiency decline for ultrahigh-NA metalens. It is also shown that the efficiency of conventional metalens has a huge gap from theoretical limitation, owing to imperfect diffractive focusing and impedance mismatch reflections. To fill such efficiency gap, the high-efficiency high-NA freeform metalens based on topology optimization is further demonstrated. Particularly, topology geometric constraints are utilized to reduce minimum feature size while keeping high efficiency. The results could shed new light on the understanding and design of ultrahigh-NA metalens and find promising applications in optical imaging, microscopy, and lithography.

1. Introduction

Optical metasurfaces offer the unprecedented possibilities of manipulating light at subwavelength spatial resolution with planar components.[1–4] In comparison with conventional bulk optics, the great advantages in miniaturization and integration of optical metasurfaces have been demonstrated in various optical applications, ranging from deflection,[5] retro-reflection,[6] lens,[7–9] light-field camera,[10,11] holograms,[12] beam-shaping,[13] to polarization conversion.[14] Particularly, owing to the excellent focusing and resolving ability, high numerical aperture (NA) metalens is of great importance to the optical applications requiring high-resolution or large-angle light collection capabilities, such as imaging, microscope, spectroscopy, and lithography.[15–18] However, high-NA metalens suffers from the well-known problem of efficiency loss (i.e., its efficiency declines rapidly with an increase in NA), especially for the ultrahigh-NA scenario,[19,20] thus impeding its practical applications.

Essentially, the low-efficiency issue in high-NA metalens mainly originates from the limitations of unit-cell-based design approach and the influences of the vector field effect. On the one hand, the conventional unit-cell-based design approach contains three implicit assumptions, that is, discrete phase sampling, local periodicity approximation, and normal response,[21–23] which ignore the lattice coupling effect and make it unable to accurately capture the actual response of non-

D. Sang, M. Xu, M. Pu, F. Zhang, Y. Guo, X. Li, X. Ma, X. Luo
State Key Laboratory of Optical Technologies on Nano-Fabrication and Micro-Engineering
Institute of Optics and Electronics
Chinese Academy of Sciences
Chengdu 610209, China
E-mail: lxg@ioe.ac.cn

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/lpor.202200265

DOI: 10.1002/lpor.202200265

D. Sang, M. Xu, M. Pu, F. Zhang, Y. Guo
Research Center on Vector Optical Fields
Institute of Optics and Electronics
Chinese Academy of Sciences
Chengdu 610209, China

D. Sang, Y. Fu
College of Electronic Science and Technology
National University of Defense Technology
Changsha 410073, China

M. Pu, Y. Guo, X. Li, X. Ma, X. Luo
School of Optoelectronics
University of Chinese Academy of Sciences
Beijing 100049, China

normally transmitted light. For the high-NA metalens where the meta-atoms are rapidly varied, such three unfavorable effects make matters even worse, resulting in imperfect focusing and the inevitable reduction of device efficiency. Theoretically, the field errors caused by discrete metasurface sampling can be compensated by using a smaller period or even the continuous phase metasurface structures[24] to increase the sampling density of phase profile. On the other hand, owing to the nonparaxial limit and the nonfarfield approximation, the in-plane field amplitude distributions of high-NA metalens are inhomogeneous and the corresponding polarization distributions are complicated. Therefore, the simplified unit-cell-based design approach supported by the generalized refraction and reflection laws (that is a scalar diffraction theory) is inadequate for the high-NA focusing scenarios, leading to an additional efficiency loss. Instead, vector diffraction theory is a more accurate approach to analyzing the electric field distribution and polarization effect near the exit aperture and the image plane.[25,26] However, the comprehensive analysis of the efficiency loss of high-NA metalens with the rigorous vector diffraction theory has been sorely lacking. More importantly, the efficiency gap between the unit-cell-based design approach and the theoretical limit based on vector diffraction theory is still elusive, even though it is of vital importance for the high-NA metalens design.

Optimization-based inverse design is an alternative scheme to overcome the inefficient designs of the unit-cell-based approach. Particularly, a rigorous full-wave solver capable of modeling near-field coupling is required to accurately capture the collective behaviors of meta-atoms. And the algorithm efficiency should be specifically considered for the metalens structure optimization in high-dimensional parameter spaces. Adjoint-based topology optimization is an efficient algorithm well-suited for high-NA metalens inverse design, which utilizes the Born approximation and the Lorentz reciprocity to sharply reduce the computational complexity of large-scale electromagnetic calculations.[21,23,27–29] Moreover, topology optimization can modify the discrete periodic structure to a continuous freeform structure, offering a high-efficiency solution for a variety of metalens optimization.[21,30–33] However, the unfettered topology variations of meta-atoms occurring even at deep subwavelength scale will introduce sensitive and unpredictable alterations of electromagnetic scattering response. Therefore, it is a great challenge to design a freeform metalens to meet the practical fabrication ability while keeping the actual efficiency almost undegraded.

Here, we theoretically and comprehensively analyze the efficiency loss behaviors of high-NA metalens under the nonparaxial limit and nonfarfield approximations, by using a rigorous vector diffraction theory and plane wave expansion method (PWEM). The theoretical efficiency limit of high-NA metalens bounded by diffractive ability and evanescent wave loss is investigated, and it is shown that the efficiency gap appearing in the unit-cell-based design approach results from many complicated factors, including the insufficient spatial sampling of phase profiles and the field distribution error due to local periodicity approximation and normal response. In order to fill such efficiency gap, a high-NA freeform metalens is demonstrated with adjoint-based topology optimization. Full-wave simulations show that, even for the ultrahigh-NA of up to 0.99, the topology-optimized freeform metalens can still achieve subdiffraction focusing with high transmission efficiency (62.5%) and focusing efficiency (14.7%), demonstrating its great advantage in comparison with unit-cell-based design approach. Furthermore, geometric constraints control are introduced into the optimization process to meet the practical fabrication demand. Specifically, the minimum feature size is greatly optimized from 5 to 100 nm while keeping a relatively high focusing efficiency (12.8%), leading to a high-efficiency high-NA freeform metalens design well-suited for practical fabrication. Our results may provide helpful suggestions for the high-NA metalens design and broaden the general applicability of topology-optimized freeform metasurfaces.

2. Results

2.1. Vector Diffraction Theory of Planar Metalens

The metalens focuses collimated incident light into a spot in transmission mode. Based on the generalized laws of reflection and refraction, the desired phase profile $\phi_d(r)$ is discretely sampled at the center of the periodic unit-cell and defined by

$$\phi_d(r) = -\frac{2\pi}{\lambda_d}\left(\sqrt{r^2 + f_d^2} - f_d\right) \tag{1}$$

where λ_d is the design wavelength, r is the position to the center of metalens, and f_d is the focal length of metalens. Note that λ_d = 633 nm is adopted in all the designs demonstrated below.

As shown in **Figure 1**a, the vector diffraction theory is adopted to analyze the ideal field distribution near exit aperture and the image plane of an idealized planar metalens (see Note S1, Supporting Information for details).[25,26] Here, the planar metalens is assumed to be reflectionless and lossless, offering an approximation model for practical design. It deflects the transmitted wave-vector, and the field perpendicular to the wave-vector needs to be reprojected onto the coordinate axis according to the direction of the deflected wave-vector. With a linear-polarized monochromatic normal incident plane wave ($\mathbf{E^{in}} = \hat{\mathbf{x}}E_0$), the field distribution after passing through a focusing planar metalens is governed by

$$\begin{aligned}\mathbf{E^{out}}(\theta,\varphi) = t(\theta)E_0[&\hat{\mathbf{x}}(\cos\theta\cos^2\varphi \\ &+ \sin^2\varphi) + \hat{\mathbf{y}}(\cos\theta - 1)\sin\varphi\cos\varphi + \hat{\mathbf{z}}\sin\theta\cos\varphi]\end{aligned} \tag{2}$$

where $\theta = \arctan(\sqrt{x^2+y^2}/f_d)$ is the local deflection angle of the metalens, $\varphi = \arctan(y/x)$, and $t(\theta) = 1/\sqrt{\cos\theta}\exp[\phi_d(r)]$ accounts for apodization function (i.e., amplitude factor) and the spatially varying focusing phase profile imparted by the metalens. Note that if the angle or polarization of incident light is varied, the Equation (2) should be modified accordingly.

Based on the above theory, the magnitudes of different electric field components near the exit aperture of metalens for different NA are shown in Figure 1b. Considering rapidly changing spatial profiles together with the feasible computation time, a medium aperture of $32\lambda_d$ (about 20 μm) was chosen as the computation time increases significantly with the increment of metalens size. For the low-NA metalens (NA = 0.1), it has a relatively uniform E_x

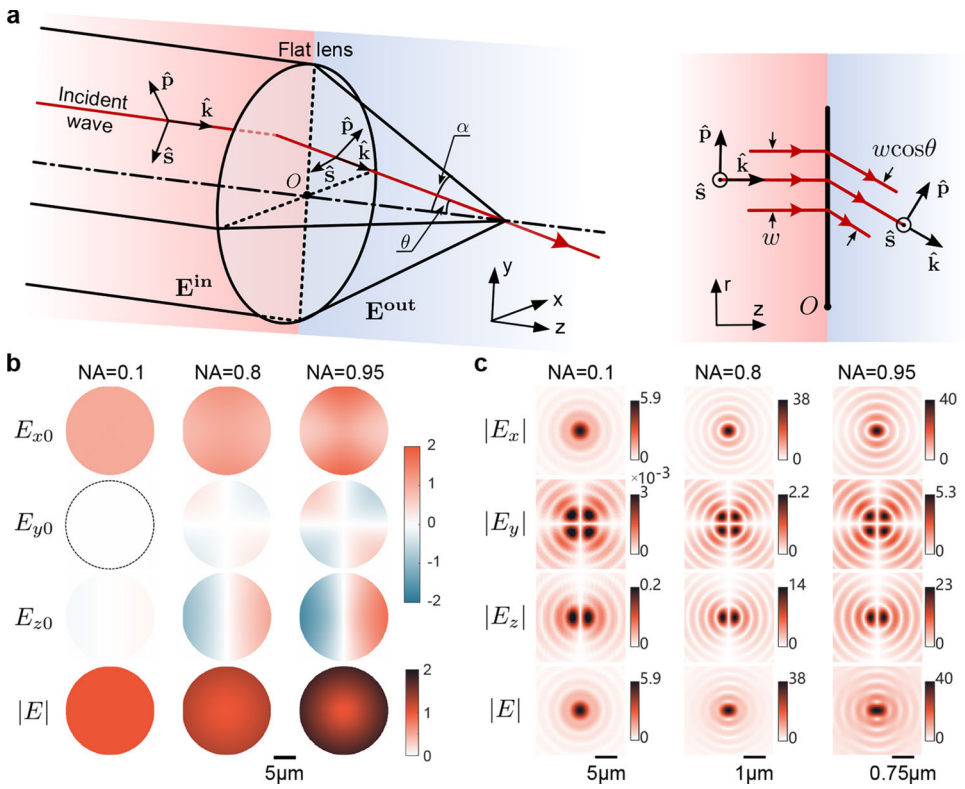

Figure 1. Vector diffraction theory of a planar metalens. a) Schematic ray diagram of light field tracking for a planar metalens. It focuses a vector light field with a maximum azimuthal angle α. The transverse area and wave-vector direction of light rays traversing the metalens has changed. b) Field magnitudes near the exit aperture of metalens for different NA calculated with vector diffraction theory. c) Focal spots in the image plane calculated by PWEM. The scale bar has been scaled for each NA for visualization.

field distribution and approximately pure spherical focus phase, with the negligible E_y and the relatively small E_z over the entire aperture. However, with the increment of NA, the electric intensity of metalens increases from the center region to the edge region. Simultaneously, owing to the depolarization effect,[34] the refracted electric field includes not only significant x components but also y and z components. More importantly, for the ultrahigh-NA metalens (NA = 0.95) with a large convergence angle, such depolarization effect becomes more apparent.

Figure 1c presents the diffracted focal spots in the image plane corresponding to the electric fields in Figure 1b, which are calculated by the PWEM.[35] As expected, the focal spots of low-NA metalens exhibit the circular Airy pattern by scalar diffraction theory due to the suppression of E_y and E_z components.[35] In contrast, increasingly significant ripples and asymmetry elliptical focal spots appear in the higher NA metalens. As the longitudinal z component of the electric field interferes in the focal region with comparable intensity as the x component, it results in the split of the bimodal focus; that is, the positions of the maximum intensity are away from the optical axis along the x-axis.

2.2. Efficiency Gap between Conventional Metalens and Vector Diffraction Theory

To thoroughly understand the origin of efficiency loss of conventional metalens, a series of typical metalens with different NAs is designed based on the unit-cell-based approach, as shown in **Figure 2**. With the assumption of gradually varying metalens, this approach decomposes the design space into locally periodic subwavelength-scale unit-cells. Figure 2b,c present the simulated electric field distributions for metalens with different NAs (see Experimental Section and Note S2, Supporting Information for more design details and results of other NAs). Specifically, Figure 2b shows that, with the increment of NA, the light intensity distributions become more aggregated in the longitudinal direction. As shown in Figure 2c, the light intensity distributions in the focal plane reveal that the low-NA metalens exhibit a circular spot, while the elliptical spot emerges in the high-NA metalens. Notably, these results are in accordance with the theoretical prediction presented in Figure 1c.

To clarify the efficiency gap from vector diffraction theory, the focusing characteristics of conventional forward-designed metalens are investigated in detail. We start from the full width at half-maximum (FWHM). Figure 2d shows the evolutions of FWHM with the increment of NA. Notably, due to the linear x-direction polarization incident, the FWHM of focal spots are totally different along x- and y-direction. For the high-NA metalens, the corresponding FWHM along y-direction is much smaller than that along x-direction, leading to the super-diffraction focusing phenomenon along y-direction. Essentially, owing to the destructive interferences between the E_y and E_z components, such phenomenon only depends on the E_x component distribution of the aperture field. The sub-diffractive focusing in the y-direction

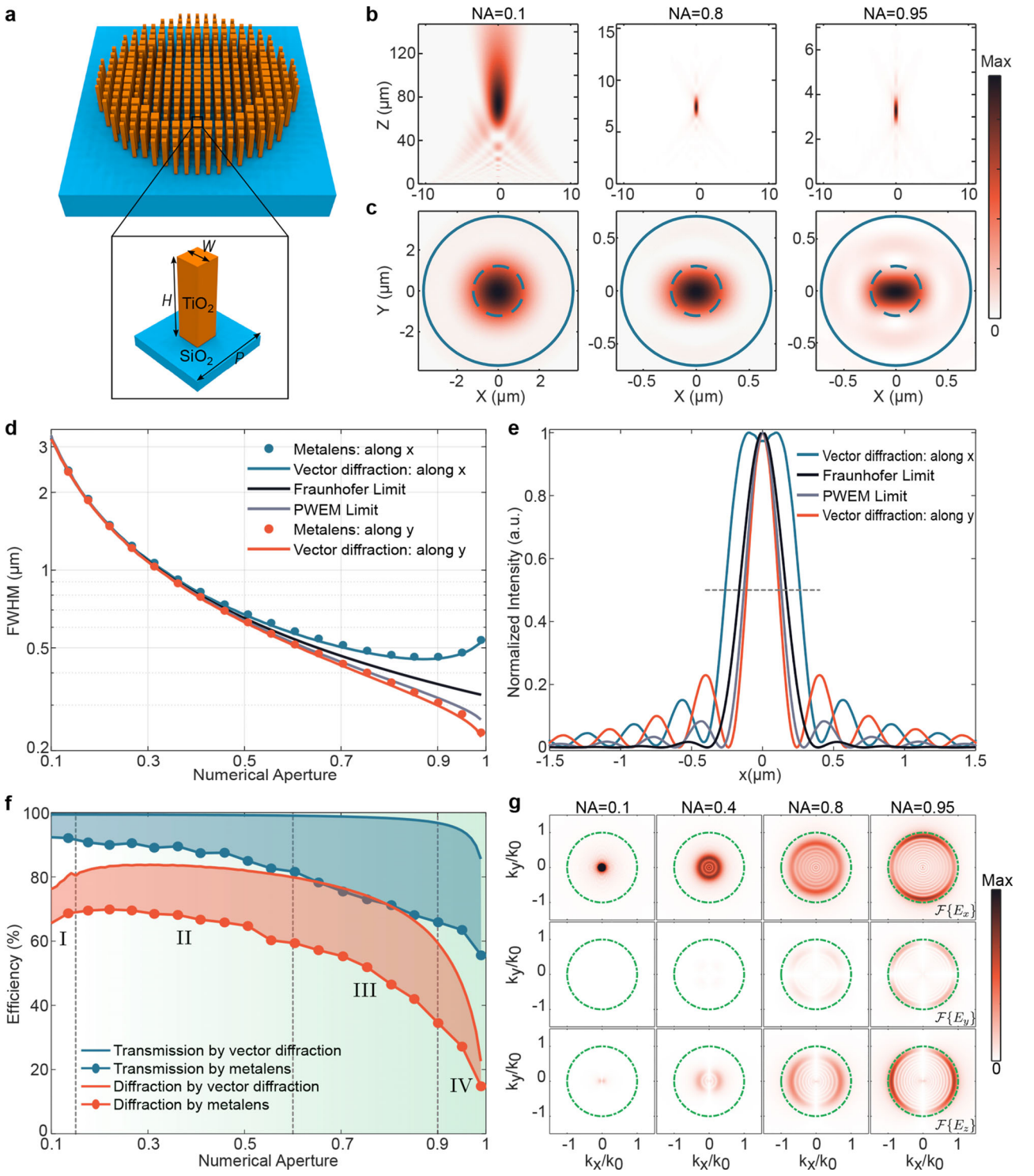

Figure 2. Comparison of conventional metalens and vector diffraction theory. a) Schematic of typical metalens operating in transmission mode. The inset shows a side view of the unit cell of metalens. For the design wavelength $\lambda_d = 633$ nm, the nanopillar height is $H = 600$ nm and the square lattice period is $P = 350$ nm. b,c) Normalized intensity profiles of the focal spot for different NAs at the *xoz* and *xoy* plane. The dotted line and solid line indicate the FWHM and focus area, respectively. d) FWHM comparison for conventional metalens and vector diffraction theory. The Fraunhofer scalar diffraction limit of a circular aperture (FWHM $= 0.514\lambda_d$/NA) and the FWHM of a circular aperture calculated by PWEM are shown. e) Intensity profiles of a lens with NA = 0.99 calculated from vector diffraction theory, Fraunhofer scalar diffraction limit, and PWEM limit. f) Transmission and diffraction efficiencies for conventional metalens and vector diffraction theory. The blue (red) region indicates the transmission (diffraction) efficiency gap between conventional metalens and vector diffraction theory. g) Transmission field distribution in the normalized *k*-space for different NAs. k_x and k_y are the wave-vector along the *x* and *y* directions, and the green dashed line represents the free space wave-vector k_0.

arises from the approximation of scalar diffraction limit and the inhomogeneous aperture distribution of the E_x. Moreover, Figure 2d clearly indicates that the FWHM $= 0.514\lambda/\mathrm{NA}$ obtained from the Fraunhofer diffraction pattern is no longer applicable in a large NA lens. Instead, the PWEM is a more accurate approach to calculate the field distribution in the focal plane. Figure 2e shows the difference of the Airy disc intensity distributions of the circular aperture calculated by the Fraunhofer diffraction theory and PWEM. Specifically, the central amplitude of the inhomogeneous exit aperture field is lower than the edge amplitude, resulting in a narrower focus point in focal plane in comparison with a uniform circular aperture. As a result, the corresponding energy of the main-lobe is greatly suppressed, leading to the amplified energy in the side-lobe.

We then investigate the focusing efficiency of metalens, which is calculated as the ratio between the optical intensity in an area with diameter of three times the FWHM at the focal plane and the total incident intensity, that is, $\eta = I^{\mathrm{focal}}_{2r<3\cdot\mathrm{FWHM}}/I^{\mathrm{in}}$. Note that different definitions of focusing efficiency have been proposed in the recent literatures,[19,31,36–42] and we select a more general one here.[43–45] Particularly, to distinguish the contributions of transmitted field and diffracted field, the focusing efficiency is divided into transmission efficiency and diffraction efficiency as

$$\eta = \eta_{\mathrm{t}} \cdot \eta_{\mathrm{d}} \tag{3}$$

where the transmission efficiency η_{t} represents the proportion of the energy transmitted through the planar metalens to the incident energy, and the diffraction efficiency η_{d} is the ratio of the focused light intensity to the transmitted light intensity.

As shown in Figure 2f, depending on the evolution pattern of transmission efficiency and diffraction efficiency, the complex mechanism of efficiency loss behaviors can be distinguished from four separate regions: I) Single Fresnel-zone region. When NA < 0.15, the diffraction efficiency increases with NA. Since the selected diameter is $D = 20$ μm, the metalens does not have a complete Fresnel zone simultaneously, which means that the phase change is less than 2π. Therefore, the focusing performance is mainly limited by the diffraction ability, making incomplete focal point with the focal position along z-axis deviating from the theoretical value. II) Scalar diffraction region. When $0.15 \leq \mathrm{NA} < 0.6$, the transmission efficiency is nearly 100% and the diffraction efficiency is about 80%, which is consistent with the results of Fraunhofer scalar diffraction of circular aperture.[35] In this region, the loss of focusing efficiency is attributed to the diffraction effect of light, that is, diffraction energy is not entirely concentrated on a small focal point in the focal plane with a small amount of energy leaking to form ripples. III) Vector diffraction region. The transmission efficiency remains near unity when the NA is ≈ 0.6–0.9. However, with the increase of the E_y and E_z components, scalar diffraction is no longer applicable. The field of E_y and E_z components destructively interfere at the focal point, resulting in decreased diffraction efficiency. IV) Evanescent wave region. When NA exceeds 0.9, the vector diffraction effect remains dominant and it further leads to a sharp decline in diffraction efficiency as the E_y and E_z components dramatically increase. Meanwhile, the corresponding transmission efficiency declines to nearly 80%, which could be explained from the transformation theory between propagating wave and evanescent wave.[46] As shown in Figure 2g, for an ultrahigh-NA metalens, the main deflection wave-vector is close to the unit circle k_0 and then the leakage spectrum partially exceeds the unit circle. Here, most of the evanescent wave may not contribute to the focusing spot when it is scattered and re-emitted into the free-space, resulting in a significant loss of transmitted energy. As expected, the spectral leakage due to finite aperture decreases as the size of planar metalens increases (see Note S3, Supporting Information for details).

Consequently, the presuppositions hidden in the unit-cell-based design approach lead to a huge efficiency gap between conventional metalens and vector diffraction theory, especially in the vector diffraction region III and the evanescent wave region IV. Specifically, the diffraction efficiency gap results from the predictable field deviation between the predesigned responses of periodic unit-cells and the actual responses of nonperiodic configuration in conventional metalens, as well as the discrete phase sampling design strategy. By contrast, the transmission efficiency gap is mainly due to the inevitable reflection caused by a spatial impedance mismatch, as it no longer satisfies the nonreflection condition assumed in the vector diffraction theory.

2.3. Topology-Optimized Freeform Ultrahigh-NA Metalens

To overcome the shortcomings of unit-cell-based design approach and improve the efficiency of conventional ultrahigh-NA metalens, we inversely designed the metalens with adjoint-based topology optimization. As shown in **Figure 3**a, in the adjoint simulation, the metalens is excited by backward-propagating adjoint sources that are determined by the desired electric field distribution. Specifically, the figure of merit (FoM) of topology optimization is defined as the norm of complex-valued projection integral that characterizes the minimum deviation of output fields from the desired field distribution

$$F(\mathbf{E}(\mathbf{x}_0)) = \int_{x_0\in\Omega} f(\mathbf{E}(\mathbf{x}_0))\mathrm{d}s = \int_{x_0\in\Omega} |\mathbf{E}(\mathbf{x}_0)\cdot\mathbf{E}_{\mathbf{d}}^{*}|^2\mathrm{d}s \tag{4}$$

where $\mathbf{E}(\mathbf{x}_0)$ represents the electric field excited by a normally incident electric field $\mathbf{E}_{\mathbf{i}}$ in the forward simulation, $\mathbf{E}_{\mathbf{d}}$ indicates the desired electric field distribution in the exit aperture region Ω, and $*$ represents the complex conjugate operation. Especially, $\mathbf{E}_{\mathbf{d}}$ is described by the Equation (2). As a result, the backward-propagating adjoint sources in the adjoint simulation are governed by the derivative with respect to the electric field

$$\frac{\partial f}{\partial \mathbf{E}} = \mathbf{E}_{\mathbf{d}}^{*}(\mathbf{E}^{*}\cdot\mathbf{E}_{\mathbf{d}}) \tag{5}$$

Based on the Lorentz reciprocity, the partial derivative of F relative to permittivity ε is given by $\partial F/\partial\varepsilon(\mathbf{x}') = 2\Re[\mathbf{E}(\mathbf{x}')\cdot\mathbf{E}_{\mathbf{adj}}(\mathbf{x}')]$ (see Note S4, Supporting Information for details), where $\mathbf{E}_{\mathbf{adj}}(\mathbf{x}')$ denotes the electric field in the adjoint simulation and $\mathbf{x}'$ represents the design domain. Then, the topology optimization starts from an initial design and is updated iteratively by the gradient ascent algorithm. In each iteration, the permittivity in every pixel position is adjusted to approach the local optimal value, leading to a freeform topology structure (see Experimental Section for details). Figure 3b presents a typical structure of freeform metalens

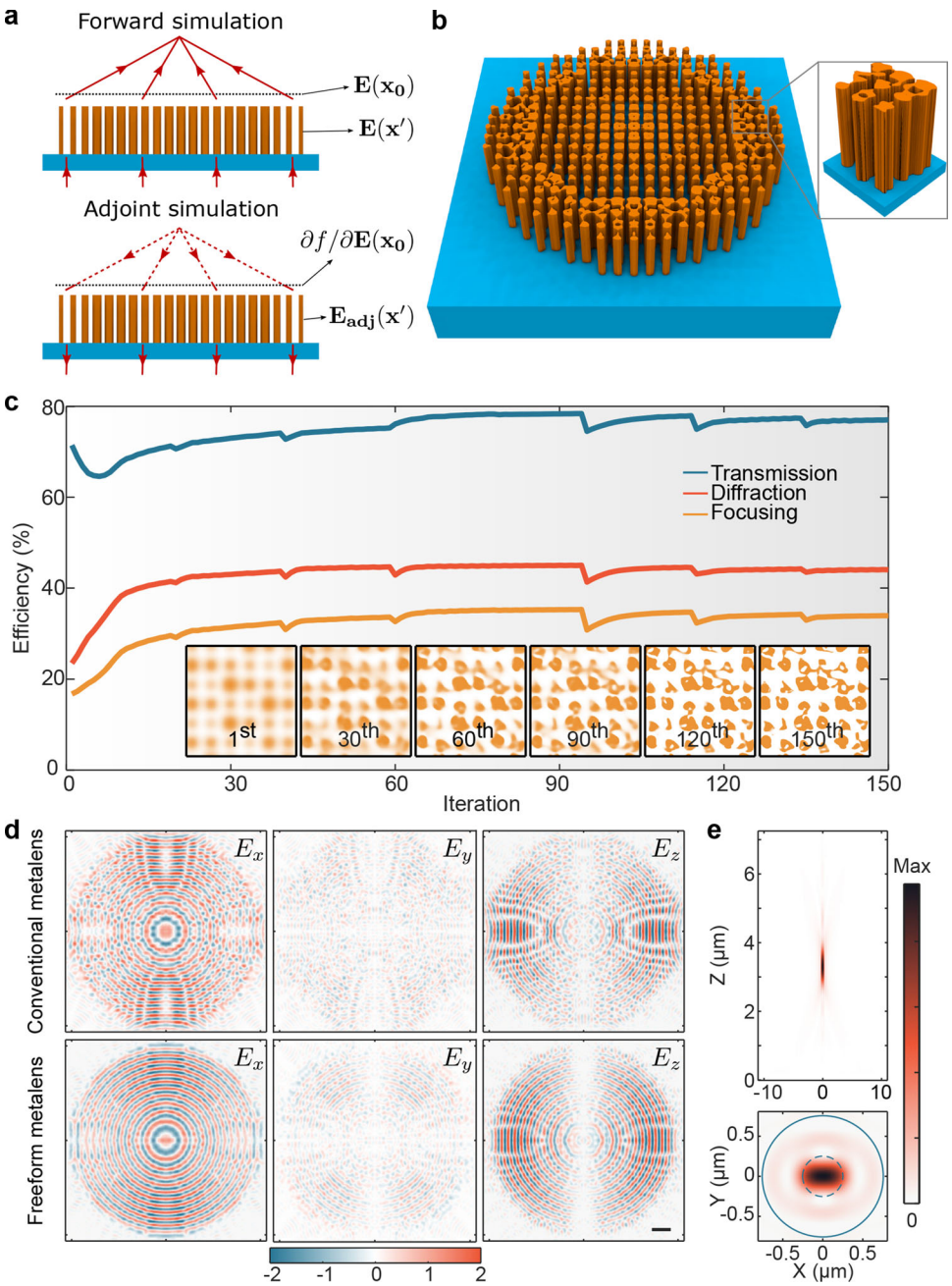

Figure 3. Topology optimization of ultrahigh-NA metalens. a) Schematic of forward simulation and adjoint simulation technique. The dashed line indicates the exit aperture region used for FoM calculation. b) A typical structure of topology-optimized freeform metalens. c) Evolution of different efficiencies during the optimization process for NA = 0.95 metalens. Inset shows the density geometry of metalens at different iterations during the optimization process. d) Comparison of electric field distributions between conventional metalens (up) and topology-optimized freeform metalens (down) for NA = 0.95. Scale bar: 2 μm. e) Simulated normalized intensity profile of the focal spot at the *xoz* plane and *xoy* plane of freeform metalens with NAs = 0.95. The dotted line indicates the FWHM, and the solid line indicates its focus area.

designed by topology optimization, which is optimized based on the initial structure as shown in Figure 2a.

To demonstrate the topology optimization method, we designed a series of metalens with different NAs ranging from 0.1 to 0.99, whose intensity profiles at *xoz* and *xoy* plane are displayed in Figure S5, Supporting Information. For a conventional metalens with NA = 0.95, the focusing efficiency is only 17.2%. However, as shown in Figure 3c, the corresponding focusing efficiency is greatly improved to 34.2% after 150 iterations of topology optimization, which is close to the theoretical limit of 44%. Moreover, Figure 3d shows the electric field distribution of different components at the exit aperture for conventional metalens and topology-optimized freeform metalens. For the conventional metalens, the in-plane electric field distribution is discrete

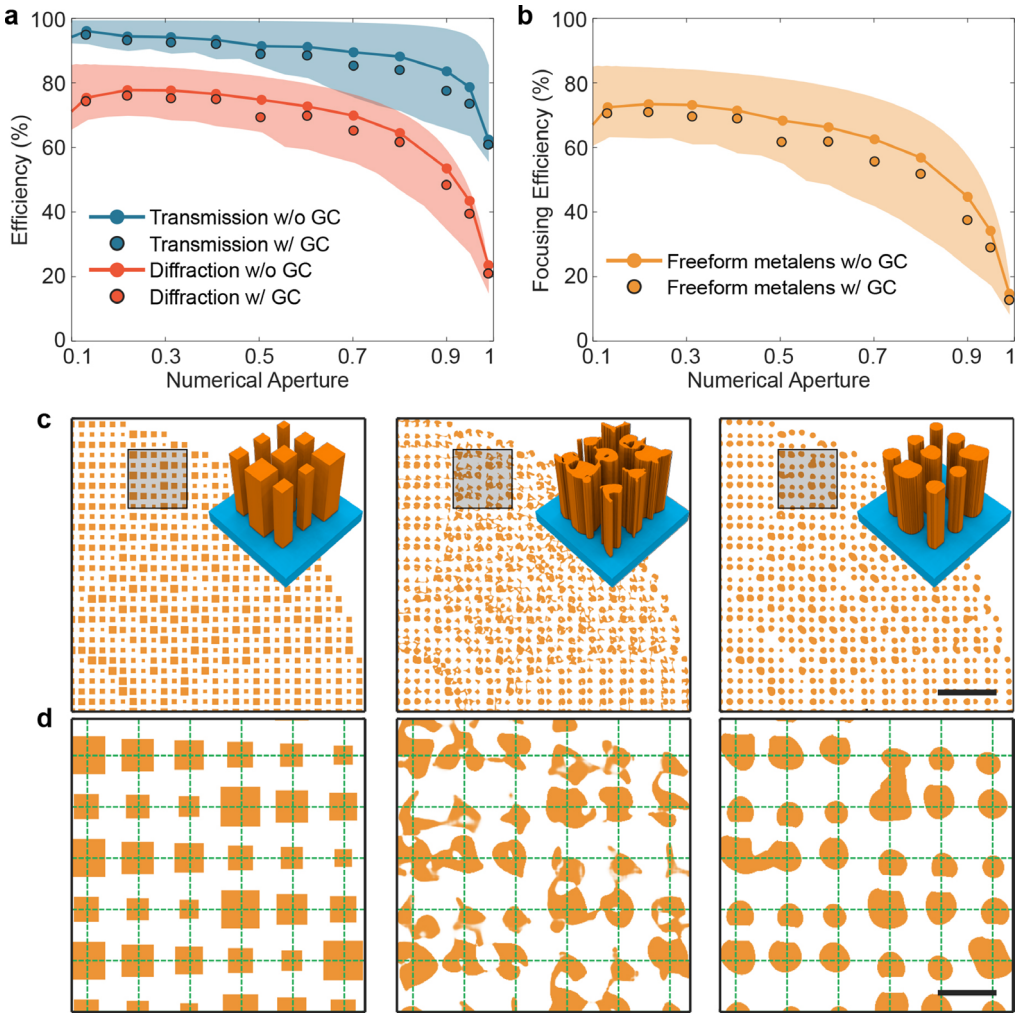

Figure 4. Topology-optimized freeform metalens with and without geometric constraints. a) Transmission and diffraction efficiencies, and b) focusing efficiency of topology-optimized freeform metalens. The shaded areas represent the efficiency gap between conventional metalens and vector diffraction theory, as shown in Figure 2f. c) Representative structures of conventional metalens, freeform metalens, and freeform metalens with GC, in which NA = 0.95. Scale bar: 2 μm. d) Enlarged views of the gray box in (c). The green dotted line indicates the grid of sampling points at the center of the periodic unit-cell. Scale bar: 0.4 μm.

and inhomogeneous accompanying with the frequently appeared abrupt shifts, leading to additional random intensity fluctuations. In contrast, owing to the smaller field distribution error, the in-plane electric field distribution of topology-optimized freeform metalens is more continuous and smooth, and the abrupt shifts are greatly suppressed. As shown in Figure 3e, unlike the significant changes in the in-plane near-field distribution, the farfield distribution of topology-optimized freeform metalens appears similar with that of the conventional metalens, although the focusing efficiency is doubled.

Although the focusing efficiency is greatly improved, the freeform structure has distinct porous structures and disconnected patterns with feature sizes even down to 5 nm, which would be unrealistic for practical fabrication. Therefore, blurring filter and geometric constraint (GC) method are further adopted to eliminate the undesired patterns to meet the fabrication requirements,[47,48] while keeping the focusing efficiency not substantially collapsed (see Experimental Section for details).

Figure 4a presents the transmission and diffraction efficiency of topology-optimized freeform metalens with and without GC method. For the topology-optimized freeform metalens without GC method, the transmission efficiency is significantly improved especially in the vector diffraction region III and evanescent wave region IV, where the reflections of inevitable impedance mismatch makes the optimized efficiency only up to – ≈ 80–90%. As predicted by the vector diffraction theory, in the evanescent wave region IV where NA > 0.9, the transmission efficiency gradually decreases due to the evanescent wave loss. In addition, the diffraction efficiency in each region has been significantly improved, as the topology optimization affords certain compensations for the field distribution distortions caused by the assumptions of local periodicity and normal response, as well as dramatically inhibiting the electric field discontinuities due to discrete phase sampling. For the topology-optimized freeform metalens with GC, both transmission and diffraction efficiencies are slightly reduced due to the topology structure limited by geomet-

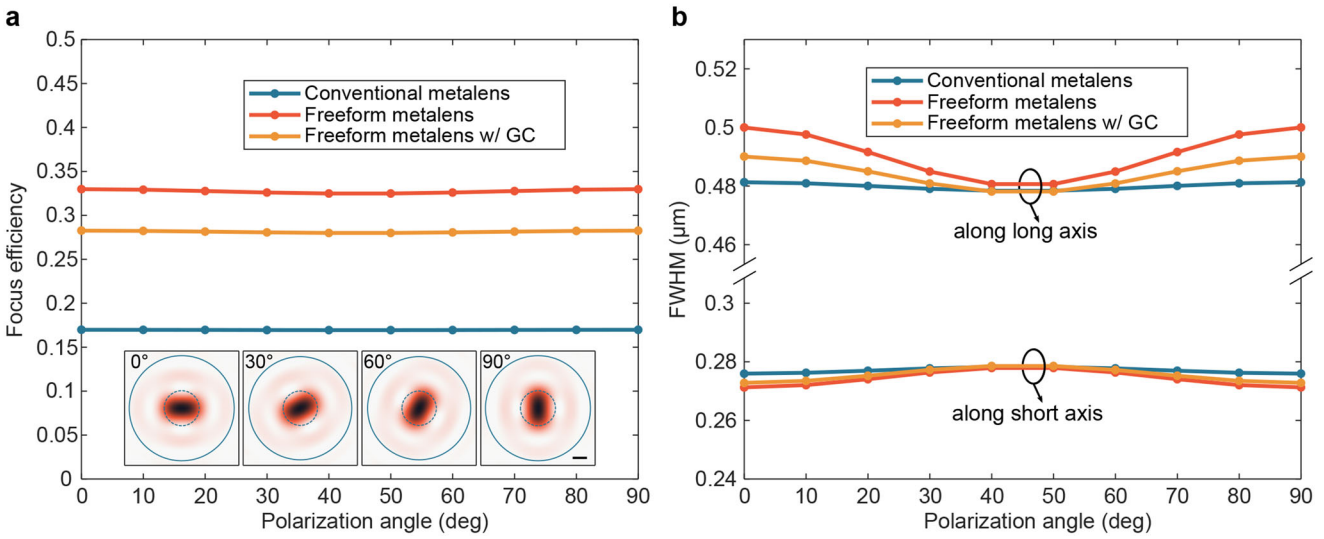

Figure 5. a) Focusing efficiency and b) FWHM of the metalens with NA = 0.95 at different incident polarization angles. The inset is normalized intensity distribution in the *xoy* plane for the freeform metalens with GC, where the dotted line indicates the FWHM and the solid line indicates its focus area. Scale bar: 200 nm.

ric constraint, especially in the high-NA region where more freedom is demanded for rapidly changing structures.

As shown in Figure 4b, the focusing capability of freeform metalens has been greatly improved in comparison with conventional metalens and is approaching to the theoretical limit, especially in the ultrahigh-NA scenario. Notably, even with an additional GC control, the focusing efficiency of freeform metalens has no serious degradation. For example, it still has a relatively high focusing efficiency of 29.1% at NA = 0.95. Furthermore, Figure 4c,d present the structure variations from the conventional discrete metalens to the continuous freeform metalens with and without GC (see Figures S6 and S7, Supporting Information for the other NAs). In the continuous freeform metalens, both the distinct periodic boundary and the lattice center are vanished, thus implementing spatially continuous modulation of incident electromagnetic waves. More importantly, the porous structures have been completely eliminated in the freeform metalens with GC control and the minimum feature size is greatly optimized from 5 to 100 nm (corresponding to a 6:1 aspect ratio), making it more easy-to-fabricate than the initial freeform metalens. Note that the designed freeform metalens could be fabricated with atomic layer deposition and electron beam lithography technique (see Note S5, Supporting Information for details).

Due to the high degree of structure symmetry (along the *x*-, *y*-, and $y = x$ axes), both the FWHM and the focusing efficiency of the designed metalens are theoretically polarization-insensitive. As shown in **Figure 5**a, the focusing efficiencies almost keep constant as the incident polarization state rotates from *x*-axis to *y*-axis. Importantly, the corresponding elliptical focusing spot rotates with the incident polarization to keep the ellipse's long axis in line with the polarization direction. Moreover, Figure 5b indicates that the FWHMs along the long and short axes of the ellipse also exhibit excellent polarization insensitivity with variation range of less than 20 nm. Particularly, for the 45° polarization incidence, different types of metalens have very similar values of FWHMs.

3. Conclusion

To summarize, we comprehensively demonstrated that, due to the presuppositions hidden in unit-cell-based design approach, conventional metalens has a huge efficiency gap from the theoretical limitation obtained by vector diffraction theory and PWEM. Importantly, we also revealed that diffraction capability of vector field and evanescent wave attenuation play different roles on the efficiency decline of high-NA metalens. On the one hand, the light field cannot be perfectly focused due to the vector diffraction theory; on the other hand, the evanescent wave caused by the finite aperture effect of ultrahigh-NA metalens makes a part of transmitted energy to be scattered into the free-space, leading to the rapid descent of transmission efficiency as well as focusing efficiency. Based on topology optimization, we demonstrated a freeform ultrahigh-NA metalens with efficiency approached to the theoretical limit. By using GC control, the undesired porous structures are successfully eliminated and the corresponding minimum feature size is greatly improved while keeping high efficiency, leading to the high-efficiency ultrahigh-NA freeform metalens appropriated to practical fabrication requirements.

Table 1 summarizes the specific performance comparison with polarization-insensitive transmissive metalens in previous literatures where clear definitions of focusing efficiency are presented. It is shown that, for different NA regions, the focusing efficiencies of our freeform metalens are significantly improved compared with conventional metalens, even under GC control. Particularly, the efficiency improvement is more remarkable for the near-unity NA. Moreover, the ratio of focusing efficiency to theoretical limit efficiency is adopted as an auxiliary reference index to characterize the efficiency gap between the designed metalens and the theoretical limit. As presented in Table 1, such efficiency gaps of our freeform metalens are narrower than conventional metalens. For comparison, the focusing efficiencies of our freeform metalens with a definite of $\eta = I^{\text{focal}}_{2r<6\cdot\text{FWHM}}/I^{\text{in}}$ are further presented in Table S1, Supporting Information.

Table 1. Performance comparisons of polarization-insensitive transmissive metalens.

NA	Ref.	λ [nm]	D [μm]	Def. ($2r <$)	$\eta_{\text{theo.}}$ [%]	η [%]	$\frac{\eta}{\eta_{\text{theo.}}}$ [%]
0.11	[42]	625	112	6 FWHM	93.6	45 (exp.)	48
0.13	Our work	633	20		93.0	84.9[a] →82.5[b]	89
0.56	[45]	640	10.22	3 FWHM	80.1	47 (sim.)	59
0.60	Our work	633	20		78.9	66.3[a] →61.8[b]	78
0.99	[19]	715	600	≈10.4 FWHM	38.4	10 (exp.)	26
	Our work	633	20		41.3	29.4[a] →25.6[b]	62

[a] Focusing efficiency of freeform metalens without GC control; [b] focusing efficiency of freeform metalens with GC control.

The adjoint-based topology optimization technique offers a simple yet effective approach for overcoming the unit-cell-based design limitations. By introducing a multi-objective optimization framework, such adjoint-based topology optimization technique can be easily extended to other important functionality designs, such as broadband achromatic metalens, wide-angle metalens, and complex multifunctional single-layer metalens.[21] To solve the design problem of large-scale metalens, we suggest that the combination of topology optimization and fast electromagnetic solution methods, including adaptive step size, domain-decomposition approximations, and axisymmetric restrictions,[28,31,36] allows for the optimization of metalens with dimensions as large as a thousand wavelengths. Moreover, as topology optimization is a gradient-based optimization method, the presented freeform metalens are locally optimal rather than globally optimal, which means that the focusing efficiency could be further improved. In fact, deep neural network could serve as a kind of robust global optimizers design by combining topology optimization,[49,50] offering a promising direction for large-scale meta-devices design.

We hope that the results presented here can serve as an important step toward understanding the complex roles of vector light effect on the efficiency loss issue in high-NA metalens, which may also be helpful in designing other high-performance subwavelength meta-structures close to vector diffraction-based theoretical limit. Finally, an exciting next step is to investigating the promising applications of topology optimization in the field of vector meta-optics, such as near-field surface-plasmon polariton lithography.

4. Experimental Section

Numerical Simulations: As shown in Figure 2a, the metalens consists of TiO_2 meta-atoms ($n_{TiO_2} = 2.58$) located on a fused silica substrate ($n_{SiO2} = 1.45$), where the height of TiO_2 circular pillar is 600 nm and the square lattice constant is 350 nm. First, a rigorous coupled-wave analysis solver RETICOLO[51] was used for the parameter sweeps of large-scale unit-cell library during the forward design of conventional metalens. Then, both the electromagnetic properties analysis and the field calculation of topology optimization were performed by a commercial electromagnetic software (Lumerical FDTD solutions) based on finite-difference time-domain method, and the results were converted from time to frequency domain. In the simulations, the cutoff of convergence accuracy was set as 10^{-6}. The metalens were illuminated by normally incident *x*-polarized plane waves and the incident field outside the metalens aperture was blocked by a perfect electric conductor layer. The field in a small volume (21 μm × 21 μm × 1.35 μm) around the metalens was simulated and terminated on all sides by perfect matched layer boundary condition. Owing to the symmetry of structure and field, symmetric boundary conditions were used along both *x* and *y* axes in order to reduce the required computational resources. Moreover, the electrical field on the focal plane of farfield was obtained by plane wave expansion method in Fourier angular spectrum domain.

Topology Optimization: As shown in Figure 3a, the forward simulation and adjoint simulation were excited by a *x*-polarized plane wave and the $\partial f/\partial \mathbf{E}$ adjoint source, respectively. The optimization space $\rho(\mathbf{x}')$ was represented as the density of pixel point that varied between 0 and 1 in the each iteration of topology optimization process, eventually converging to 0 (air) or 1 (dielectric) at each pixel point. The pixel size was chosen to be 5 nm (about $\lambda/125$) to provide sufficient control while avoiding the overly fine features. The topology optimization started with a simple initial design $\rho^0(\mathbf{x}')$ (e.g., conventional metalens with partial fuzziness for deinformation) and updated iteratively by gradient ascent algorithm until all parameters converge to local optimal values or a maximum number of iterations is reached. The optimization simulation was carried out on a workstation with Intel Gold 6256 CPU and each simulation used 24 cores. Particularly, each of iteration of topology optimization took ≈160 s, including both forward and adjoint simulation.

Geometric Constraints: As a simple yet effective way to control variations in spatial field, the density filter was adopted for the geometric constrains of topology optimization. Specifically, a filtered field $\tilde{\rho}(\mathbf{x}')$ can be calculated from a design field $\rho(\mathbf{x}')$ as ref. [48]

$$\widetilde{\rho}_i(\mathbf{x}') = \frac{\sum_{j\in\mathbb{N}_i}\omega(\mathbf{x}_j')\nu_j\rho_j(\mathbf{x}')}{\sum_{j\in\mathbb{N}_i}\omega(\mathbf{x}_j')\nu_j}, w(\mathbf{x}_j') = R - \left|\mathbf{x}_i' - \mathbf{x}_j'\right| \tag{6}$$

where $\mathbb{N}_i$ indicates the neighborhood set of element *i* lying within the filter radius, *R* is the spatial filtering radius, and ν_j is the volume of the element *j*. Then, the projected field $\bar{\rho}_\eta(\mathbf{x}')$ can be characterized by an approximate Heaviside function, where η is the threshold value that implicitly determines the minimum length scale of the projected field. Therefore, when the three design fields satisfied the following sufficient condition ref. [48]

$$\begin{cases} \tilde{\rho}(\mathbf{x}') \geq \eta_e, \forall \mathbf{x}' \in \{\mathbf{x}' | \bar{\rho}_\eta = 1 \text{ and } \nabla\tilde{\rho} = 0\} \\ \tilde{\rho}(\mathbf{x}') \leqslant \eta_d, \forall \mathbf{x}' \in \{\mathbf{x}' | \bar{\rho}_\eta = 0 \text{ and } \nabla\tilde{\rho} = 0\} \end{cases} \tag{7}$$

all the projected fields within the threshold of $\eta \in (\eta_d, \eta_e)$ shared a consistent topology, which meant that a minimum length scale was expected on the final field in the air and dielectric space.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

D.S. and M.X. contributed equally to this work. This work was supported by the National Natural Science Foundation of China (NSFC) (Grants No. 62105338 and U20A20217), the Frontier Research Fund of Institute of Optics and Electronics, China Academy of Sciences (Grant No. C21K007), and the Sichuan Science and Technology Program (Grant No. 2021ZYCD001).

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

all-dielectric metalens, geometric constraints, topology optimization, ultrahigh numerical apertures, vector diffraction

Received: April 19, 2022
Revised: June 25, 2022
Published online: August 11, 2022

[1] N. I. Zheludev, Y. S. Kivshar, *Nat. Mater.* **2012**, *11*, 917.
[2] A. V. Kildishev, A. Boltasseva, V. M. Shalaev, *Science* **2013**, *339*, 1232009.
[3] N. Yu, F. Capasso, *Nat. Mater.* **2014**, *13*, 139.
[4] P. Genevet, F. Capasso, F. Aieta, M. Khorasaninejad, R. Devlin, *Optica* **2017**, *4*, 139.
[5] E. Khaidarov, H. Hao, R. Paniagua-Domínguez, Y. F. Yu, Y. H. Fu, V. Valuckas, S. L. K. Yap, Y. T. Toh, J. S. K. Ng, A. I. Kuznetsov, *Nano Lett.* **2017**, *17*, 6267.
[6] A. Arbabi, E. Arbabi, Y. Horie, S. M. Kamali, A. Faraon, *Nat. Photonics* **2017**, *11*, 415.
[7] F. Aieta, P. Genevet, M. A. Kats, N. Yu, R. Blanchard, Z. Gaburro, F. Capasso, *Nano Lett.* **2012**, *12*, 4932.
[8] S. Banerji, M. Meem, A. Majumder, F. G. Vasquez, B. Sensale-Rodriguez, R. Menon, *Optica* **2019**, *6*, 805.
[9] Y. Wang, Q. Fan, T. Xu, *Opto-Electron. Adv.* **2021**, *4*, 01200008.
[10] Q. Fan, W. Xu, X. Hu, W. Zhu, T. Yue, C. Zhang, F. Yan, L. Chen, H. J. Lezec, Y. Lu, A. Agrawal, T. Xu, *Nat. Commun.* **2022**, *13*, 2130.
[11] M. Liu, Q. Fan, L. Yu, T. Xu, *Opt. Express* **2019**, *27*, 10738.
[12] H. Gao, X. Fan, W. Xiong, M. Hong, *Opto-Electron. Adv.* **2021**, *4*, 1210030.
[13] K. E. Chong, I. Staude, A. James, J. Dominguez, S. Liu, S. Campione, G. S. Subramania, T. S. Luk, M. Decker, D. N. Neshev, I. Brener, Y. S. Kivshar, *Nano Lett.* **2015**, *15*, 5369.
[14] J. B. Mueller, N. A. Rubin, R. C. Devlin, B. Groever, F. Capasso, *Phys. Rev. Lett.* **2017**, *118*, 113901.
[15] M. Khorasaninejad, F. Capasso, *Science* **2017**, *358*, eaam8100.
[16] M. Khorasaninejad, W. T. Chen, R. C. Devlin, J. Oh, A. Y. Zhu, F. Capasso, *Science* **2016**, *352*, 1190.
[17] M. Faraji-Dana, E. Arbabi, A. Arbabi, S. M. Kamali, H. Kwon, A. Faraon, *Nat. Commun.* **2018**, *9*, 4196.
[18] Q. Fan, M. Liu, C. Yang, L. Yu, F. Yan, T. Xu, *Appl. Phys. Lett.* **2018**, *113*, 201104.
[19] R. Paniagua-Dominguez, Y. F. Yu, E. Khaidarov, S. Choi, V. Leong, R. M. Bakker, X. Liang, Y. H. Fu, V. Valuckas, L. A. Krivitsky, A. I. Kuznetsov, *Nano Lett.* **2018**, *18*, 2124.
[20] A. Arbabi, E. Arbabi, M. Mansouree, S. Han, S. M. Kamali, Y. Horie, A. Faraon, *Sci. Rep.* **2020**, *10*, 7124.
[21] H. Chung, O. D. Miller, *Opt. Express* **2020**, *28*, 6945.
[22] M. Torfeh, A. Arbabi, *ACS Photonics* **2020**, *7*, 941.
[23] M. Mansouree, A. McClung, S. Samudrala, A. Arbabi, *ACS Photonics* **2021**, *8*, 455.
[24] M. Miyata, M. Nakajima, T. Hashimoto, *J. Appl. Phys.* **2019**, *125*, 103106.
[25] B. Richards, E. Wolf, *Proc. R. Soc. Lond., Ser. A* **1959**, *253*, 358.
[26] A. McClung, M. Mansouree, S. Samudrala, A. Arbabi, in *Conference on Lasers and Electro-Optics*, IEEE, Piscataway, NJ **2020**, pp. 1–2.
[27] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, A. W. Rodriguez, *Nat. Photonics* **2018**, *12*, 659.
[28] T. Phan, D. Sell, E. W. Wang, S. Doshay, K. Edee, J. Yang, J. A. Fan, *Light Sci. Appl.* **2019**, *8*, 48.
[29] M. Mansouree, H. Kwon, E. Arbabi, A. McClung, A. Faraon, A. Arbabi, *Optica* **2020**, *7*, 77.
[30] D. Sell, J. Yang, S. Doshay, R. Yang, J. A. Fan, *Nano Lett.* **2017**, *17*, 3752.
[31] Z. Lin, V. Liu, R. Pestourie, S. G. Johnson, *Opt. Express* **2019**, *27*, 15765.
[32] M. Xu, M. Pu, D. Sang, Y. Zheng, X. Li, X. Ma, Y. Guo, R. Zhang, X. Luo, *Opt. Express* **2021**, *29*, 10181.
[33] M. Xu, Q. He, M. Pu, F. Zhang, L. Li, D. Sang, Y. Guo, R. Zhang, X. Li, X. Ma, X. Luo, *Adv. Mater.* **2022**, 2108709.
[34] Z.-P. Zhuang, R. Chen, Z.-B. Fan, X.-N. Pang, J.-W. Dong, *Nanophotonics* **2019**, *8*, 1279.
[35] M. Born, E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, Elsevier, New York **2013**.
[36] S. J. Byrnes, A. Lenef, F. Aieta, F. Capasso, *Opt. Express* **2016**, *24*, 5110.
[37] M. Khorasaninejad, A. Y. Zhu, C. Roques-Carmes, W. T. Chen, J. Oh, I. Mishra, R. C. Devlin, F. Capasso, *Nano Lett.* **2016**, *16*, 7229.
[38] M. Kang, Y. Ra'di, D. Farfan, A. Alù, *Phys. Rev. Appl.* **2020**, *13*, 044016.
[39] C. Zhang, S. Divitt, Q. Fan, W. Zhu, A. Agrawal, Y. Lu, T. Xu, H. J. Lezec, *Light Sci. Appl.* **2020**, *9*, 55.
[40] A. Ndao, L. Hsu, J. Ha, J.-H. Park, C. Chang-Hasnain, B. Kanté, *Nat. Commun.* **2020**, *11*, 3205.
[41] A. Arbabi, Y. Horie, A. J. Ball, M. Bagheri, A. Faraon, *Nat. Commun.* **2015**, *6*, 7069.
[42] A. Zhan, S. Colburn, R. Trivedi, T. K. Fryett, C. M. Dodson, A. Majumdar, *ACS Photonics* **2016**, *3*, 209.
[43] S. Banerji, M. Meem, A. Majumder, F. G. Vasquez, B. Sensale-Rodriguez, R. Menon, *Optica* **2019**, *6*, 805.
[44] F. Balli, M. Sultan, S. K. Lami, J. T. Hastings, *Nat. Commun.* **2020**, *11*, 3892.
[45] M. M. Elsawy, A. Gourdin, M. Binois, R. Duvigneau, D. Felbacq, S. Khadir, P. Genevet, S. Lanteri, *ACS Photonics* **2021**, *8*, 2498.
[46] M. Pu, X. Ma, X. Li, Y. Guo, X. Luo, *J. Mater. Chem. C* **2017**, *5*, 4361.
[47] R. E. Christiansen, O. Sigmund, *J. Opt. Soc. Am. B* **2021**, *38*, 496.
[48] M. Zhou, B. S. Lazarov, F. Wang, O. Sigmund, *Comput. Methods Appl. Mech. Eng.* **2015**, *293*, 266.
[49] J. Jiang, M. Chen, J. A. Fan, *Nat. Rev. Mater.* **2021**, *6*, 679.
[50] T. Ma, M. Tobah, H. Wang, L. J. Guo, *Opto-Electron. Sci.* **2022**, *1*, 210012.
[51] P. Lalanne, J.-P. Hugonin, *Reticolo Software for Grating Analysis*, Cornell University, Ithaca, NY **2005**.