PHOTONICS Research

Arbitrary large-gradient wavefront shaping: from local phase modulation to nonlocal diffraction engineering

Xipu Dong,[1] Jierong Cheng,[1,2,*] Yiwu Yuan,[1] Zhenyu Xing,[1] Fei Fan,[1] Xianghui Wang,[1] and Shengjiang Chang[1,3]

[1]Institute of Modern Optics, Nankai University, Tianjin 300350, China
[2]Tianjin Key Laboratory of Micro-scale Optical Information Science and Technology, Tianjin 300350, China
[3]Tianjin Key Laboratory of Optoelectronic Sensor and Sensing Network Technology, Tianjin 300350, China
*Corresponding author: chengjr@nankai.edu.cn

Received 19 July 2021; revised 27 January 2022; accepted 1 February 2022; posted 2 February 2022 (Doc. ID 438059); published 11 March 2022

The powerful wavefront manipulation capability of metasurfaces originates from their subwavelength or deep subwavelength elements with designable optical responses, especially phase responses. However, they usually suffer from performance degradation as the spatial phase gradient is large. To solve this issue, we propose an accurate and efficient nonlocal diffraction engineering mechanism to tailor an arbitrary large-gradient wavefront utilizing superwavelength-scale elements. The fast-varying phase profile is cut into segments according to 2π zones rather than subwavelength discretization. Each phase segment is accurately implemented by precisely tailoring the diffraction pattern of the element, where diffraction angles, efficiencies, and phases are controlled simultaneously. As proof of the concept, high numerical aperture cylindrical metalenses are designed using this method and experimentally validated at the terahertz band. The cylindrical metalens is further extended to a full-space metalens, which enables high-quality subwavelength imaging with resolved details of 0.65λ. The proposed mechanism offers an efficient way to capture the fast-varying wavefront using relatively coarse geometries with new physical insights.

https://doi.org/10.1364/PRJ.438059

1. INTRODUCTION

In the past decades, remarkable progress in metasurfaces has revolutionized wavefront shaping in a planar, miniaturized, and multifunctional way, which is enabled not by the phase accumulation during the propagation path but by the phase discontinuity at the interfacial boundary [1–3]. The abrupt phase discontinuity is encoded in elaborately designed subwavelength elements through various types of interaction with the incident light, such as electric and magnetic resonances [4,5], waveguide effect [6,7], and Pancharatnam–Berry phase [8]. The implementation of a continuous phase profile is through subwavelength discretization, and each element is only responsible for local phase modulation over its occupied subwavelength area. We name this type of wavefront shaping as local phase modulation (LPM), which is widely used in metasurface-based beam focusing [9–11], polarization control [12], holographic imaging [13,14], and so on.

Despite its remarkable success in wavefront control, LPM poses some limitations to metasurfaces. First, the phase response of each element is approximated by that of the element in locally periodic environment. This assumption fundamentally limits the application scenarios of LPM to slowly varying wavefronts. It further limits the material of metasurfaces to be high-index dielectrics or metals [15] as strong field localization is needed to minimize the coupling of adjacent elements in the aperiodic array. More importantly, LPM overlooks the impedance mismatch between the input and output beams, which inherently leads to increasingly strong parasitic scattering into undesired directions with the increase of the phase gradient [16]. Large-gradient phase profile is highly required in implementing tight focusing, beam multiplexing, or more complex wavefront shaping. In these cases, LPM-based metasurfaces become inefficient [6,17,18].

To circumvent these problems, metagratings [19] with superwavelength periods are proposed for near-perfect large-angle deflection by gathering energy into the first diffraction order (blazed diffraction) [20–25]. As constitutive elements of metasurfaces, metagratings with different periods and internal geometries are then combined for broad-angle beam scattering [26] and for metalenses with numerical aperture (NA) as high as 0.99 [27,28]. Each metagrating efficiently reroutes the beam following the ray path in geometric optics through efficient blazed diffraction.

However, the diffraction phase of metagratings is not studied yet. In fact, precise control of the diffraction phase is as important as maximization of the diffraction efficiency, as the incoming beam should be piecewise diffracted and constructively superimposed to build the desired wavefront. Furthermore, an arbitrary nonlinear-shaped large-gradient wavefront manipulation may not require piecewise blazed diffraction but other diffraction patterns. One should flexibly control the energy distribution among different orders and simultaneously control the diffraction phases so as to accurately reproduce a piece of arbitrary-shaped phase profile.

In this work, we propose a general mechanism for arbitrary large-gradient wavefront shaping by controlling a full set of diffraction parameters (direction, efficiency, and phase of each order) of superwavelength elements. The desired phase profile is cut into segments according to 2π zones, and the correlation between the phase segment and the desired diffraction parameters is revealed. The latter is then implemented by bilayer one-dimensional grating structures [29] with large design freedom. Different gratings are stitched into a metasurface to shape the wavefront in a desired way. As diffraction indicates the existence of lateral energy flow, we name this type of wavefront shaping mechanism as nonlocal diffraction modulation (nLDM).

As proof of the concept, we theoretically design and experimentally study high-NA cylindrical metalenses and 3D metalenses based on nLDM in the terahertz (THz) band. They show diffraction-limited focal spots, high efficiency, and subwavelength imaging resolution. As compared to LPM, nLDM serves as a better choice for large-gradient wavefront transformation. It brings additional merits of flexible choice of abundant dielectric materials and various processing methods, as strong field localization is not required. The method is general and can be extended for wavefront shaping at other frequencies.

2. NONLOCAL DIFFRACTION METHOD FOR LARGE-GRADIENT WAVEFRONT SHAPING

Figure 1(a) shows an arbitrary wavefront shaping over a metasurface. The tangential field before and after the metasurface can be represented by $E_i(x, z = 0_-)$ and $E_o(x, z = 0_+)$. The ideal metasurface sitting at $z = 0$ offers a transmission function $t(x)e^{j\varphi(x)}$ to enable $E_o(x, z = 0_+) = t(x)e^{j\varphi(x)} \times E_i(x, z = 0_-)$. $t(x)$ and $\varphi(x)$ describe the amplitude and phase responses of the metasurface, whose thickness is overlooked here. In order to ensure high efficiency and to avoid the use of gain and lossy material, the metasurface is elaborately designed to implement the phase response $\varphi(x)$, while keeping the amplitude as high as possible. Two design approaches of $\varphi(x)$, the tradiational LPM and the nLDM, are schematically compared in Figs. 1(b) and 1(c). LPM discritizes the continuous phase with uniform subwavelength steps. Each step with a constant phase is realized by an element with proper topology. Different elements are stitched into the metasurface in the bottom of Fig. 1(b). At the positions where the phase variation is dramatic, such discritization is not accurate enough to reproduce the original curve. The phase delay of the subwavelength element is calculated in a periodic array, where the impedances of incident and transmitted beams are the same. But when the nonperiodic metasurface is constructed, the phase gradient will reshape the wavefront and change the transmission wave impedance. This impedance mismatch in the periodic array and nonperiodic array leads to parasitic diffraction of the energy into undesired directions, which become increasely severe in implementing large-gradient phase curves.

In contrast, nLDM divides the continuous phase curve into several segments, which can be described as

$$\varphi(x) = \begin{cases} \varphi_1(x) & x_1 < x \le x_2 \\ \varphi_2(x) & x_2 < x \le x_3 \\ \cdots & \cdots \\ \varphi_i(x) & x_i < x \le x_{i+1} \\ \cdots & \cdots \end{cases}. \tag{1}$$

There is no discritization error as compared to Fig. 1(b). The starting point and ending point (x_i and x_{i+1}) of each segment can be chosen in different ways. As the phase profile is usually folded in 2π, a segment can be simply chosen as a range from 0 to 2π, as shown in Fig. 1(c). The length of the segments

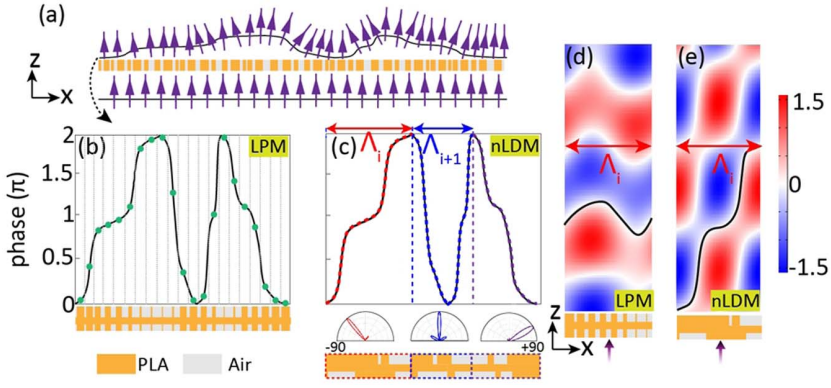

Fig. 1. (a) Schematic of wavefront shaping through a metasurface. (b) Implementation of the desired continuous phase profile (solid curve) using the LPM method. The discretized phases are shown by the dots. Each of them is implemented by a subwavelength element shown below the phase curve. (c) Implementation of the desired continuous phase profile (solid curve) using the nLDM method. The desired phase profile is divided into several segments (shown by different colors) according to the 2π zones. Each phase segment is translated into a diffraction pattern and then implemented by a superwavelength-scale metagrating. (d), (e) The field distribution after the metasurface in the first 2π zone designed by the LPM method (d) and by the nLDM method (e).

(Λ_i) varies in space and it will be slightly larger than the wavelengths in the large-gradient case. Please note that the choice of the segment can be flexible. If one segment is in the subwavelength scale, it can be merged into the latter one, as shown by the blue dashed line in Fig. 1(c). The starting and ending points can be arbitrary values not limited to 0 and 2π. It is not straightforward to find a proper metagrating geometry to reproduce the exact phase segment. By considering the metagrating in an infinitely periodic configuration with $\varphi_i(x) = \varphi_i(x + \Lambda_i)$, the far field diffraction corresponds to the coefficients of the Fourier series expansion of the transmission function $e^{j\varphi_i(x)}$ as [30]

$$e^{j\varphi_i(x)} = \sum_m t_m^d e^{j2m\pi x/\Lambda_i}, \tag{2}$$

where t_m^d is the desired complex diffraction coefficient in the m-th order, which not only determines the diffraction efficiency as $\eta_m = |t_m^d|^2$, but also defines the diffraction phase $\psi_m = \arg(t_m^d)$. Using Eq. (2), a piece of continuous phase profile in real space is mapped to M diffraction efficiencies and M diffraction phases, which can be used to find the target metagrating geometries. If the metagratings with desired diffraction behavior can be found, the phase profile can be implemented with high accuracy.

In order to meet so many diffraction requirements, we combine an optimization algorithm with electromagnetic simulation to find the metagrating quickly and automatically. We use the gradient descent optimization function fmincon in MATLAB and rigorously coupled wave analysis (RCWA) to approach the ideal diffraction by minimizing the following cost function:

$$\text{cost} = \sum_m |t_m^d - t_m^p(\boldsymbol{g})|, \tag{3}$$

where t_m^d and t_m^p are the desired diffraction coefficients and the practical ones realized by the metagrating. $\boldsymbol{g} = (g_1, g_2, g_3, \ldots, g_N)$ represents the geometric parameters (widths of ridges and gaps) to be optimized. Since the optimization is not for the whole metasurface but its constitutive metagratings, it usually takes tens of seconds to find the optimized geometry. Equation (2) ensures that the ideal total diffraction efficiency is 100%, meaning that the beam should be totally transmitted and distributed among different orders. So high transmission is inherently considered during optimization, and the impedance mismatch issue is avoided. Coherent superposition of the segmented diffraction fields is realized by taking the diffraction phases into account.

Lastly, the metagrating should have enough degree of freedom so as to achieve different types of diffraction patterns. Here it is composed of bilayer one-dimensional ridges [Fig. 1(c)] made of polymer polylactic acid (PLA), whose refractive index is 1.57 at the target frequency 0.14 THz. Our previous study has proven the success of this type of metagrating for efficient large-angle diffraction [24,31], beam splitting [29], and asymmetric diffraction [32]. To ensure a planar profile of the metasurface, the thickness of each layer is not tuned. Throughout the manuscript, the ridges in the incident side are 2.0 mm thick, the slab is 2.1 mm thick, and the ridges in the transmission side are 3.1 mm thick if not mentioned.

As an example, we implement the red phase segment in Fig. 1(c) using LPM and nLDM at 0.14 THz. The length of this segment is 1.55λ. The LPM design in Fig. 1(d) contains eight elements with fixed distance of 0.42 mm and spatially varied width. In the nLDM design, this phase profile is mapped to the diffraction coefficients of ($0.88e^{j0.5\pi}$, 0, $0.35e^{j0.8\pi}$) in the (–1, 0, 1) orders, which is achieved by the metagrating in Fig. 1(e). It contains only two ridges in each layer and has much larger feature size as compared to the LPM one. Simulations of the two designs are carried out in Lumerical FDTD Solutions. The field distributions in Figs. 1(d) and 1(e) show that the nLDM design can accurately reproduce the desired wavefront, while the LPM design fails.

3. HIGH-NA CYLINDRICAL METALENSES

As proofs of the concept, we design two high-NA cylindrical metalenses with one and four focal spots, respectively, using the nLDM method as shown in Figs. 2(a) and 2(b). The diameter of both metalenses is 82 mm (37λ). The focal length of metalens 1 is $f = 15$ mm (7λ), leading to a designed NA of 0.94. In metalens 2, four foci along the optical axis are located

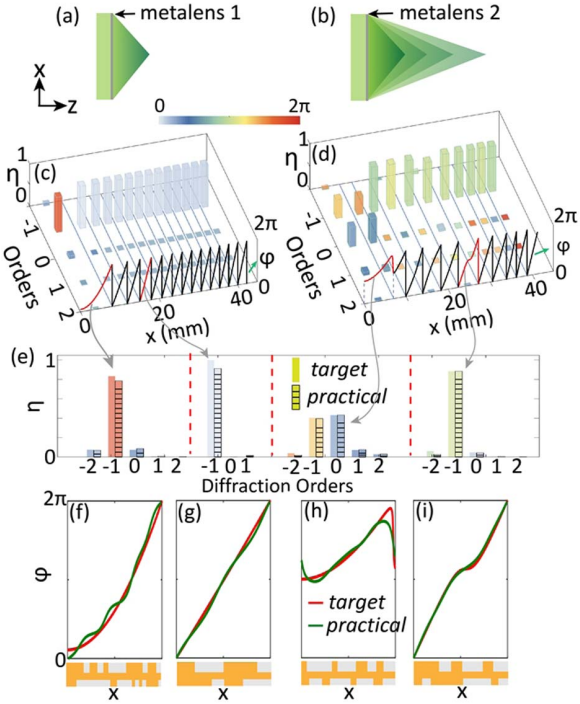

Fig. 2. Metalenses for high-NA focusing with (a) a single focal spot and (b) four focal spots. The desired phase profiles of metalenses 1 and 2 folded within 2π are shown by the black curves in (c) and (d), respectively. The diffraction pattern of each 2π zone is shown by the histogram in (c) and (d), with the height of the bar indicating the diffraction efficiency and the color of the bar indicating the diffraction phase in each order. (e) The target diffraction patterns and the practical diffraction patterns of the optimized metagrating elements corresponding to four phase segments marked by the red curves in (c) and (d). The efficiency and phase in each diffraction order are shown by the height and color of the bar. The phase distributions after the optimized metagratings and the target phase segments are compared in (f)–(i).

at $f_1 = 15$ mm, $f_2 = 25$ mm, $f_3 = 35$ mm, and $f_4 = 45$ mm. The ideal phase profile offered by the metalens can be calculated by backpropagation of point sources from the foci as

$$\varphi(x) = \arg\left(\sum_{n=1}^{n=N} \xi_n \times \frac{e^{jk_0 R_n}}{4\pi R_n}\right), \tag{4}$$

where k_0 is the wave vector in vacuum, $R_n = \sqrt{f_n^2 + x^2}$, and ξ_n is the weight to balance the intensity ratio among different focal spots. $\xi = 1$ for $N = 1$ and $\xi_{1-4} = (0.3, 0.6, 0.8, 1)$ for $N = 4$ are used here. Figures 2(c) and 2(d) show half of the ideal phase profiles along the x axis folded from 0 to 2π. They are divided into segments whenever the phase reaches 2π. Metalens 1 contains 14 segments in its half area and metalens 2 contains 11 segments. After Fourier transformation using Eq. (2), each phase segment is translated into a series of diffraction coefficients, whose amplitudes and phases are shown by the height and color of the histogram in Figs. 2(c) and 2(d). For metalens 1, all the segments require blazed diffraction with the same phase except for the first one. The largest blazed diffraction angle is 70°. The first phase segment is a quadratic shape, whose -1st order diffraction efficiency is 83% with the rest of the energy into the -2nd and 0th orders. The phase is also different from that of other metagratings. For metalens 2, the phase profile cannot be considered linear in several segments, leading to energy distribution in more than one diffraction order and variation of diffraction phases.

Next, using RCWA and the gradient descent optimization algorithm, we successfully find all the metagratings to reproduce the target diffraction patterns with x-polarized plane wave excitation (TM mode). Here we pick two segments in each metalens to show the optimization results, which are marked by the red phase segments in Figs. 2(c) and 2(d). The target and optimized diffraction efficiency (height) and phase (color) in each order is compared in Fig. 2(e), showing a good agreement. The phase profiles after the four optimized metagratings are plotted in Figs. 2(f)–2(i), which are consistent with the ideal phase segments quite well. The metagrating under each phase curve in Figs. 2(f)–2(i) shows that the relatively coarse geometry can reproduce the fast-varying phase profiles and even capture the deep subwavelength variation details as long as it exhibits the desired diffraction behavior. Details of the reproduction of phase profiles using the metagrating segments and the origin of the slight fitting error can be found in Appendix A.

The two metalenses are then 3D printed and experimentally tested. The metalenses are fabricated by a commercial 3D printer (Raise3D N2) using a fused-deposition modeling method. The diameter of the nozzle is 200 μm. The base plate and printing nozzle are set to 60°C and 215°C, respectively. The printing speed has a constant value of 60 mm/s. The metalens is piled up layer by layer along the y direction. Printing of each metalens takes 10 h approximately. The designed and 3D-printed metalenses are shown in Figs. 3(a) and 3(b), and the experiment setup in Fig. 3(c) is used to characterize the focusing performance. The source is a continuous-wave IMPATT diode at 0.14 THz. The beam is collimated by a Teflon convex lens with a focal length of 200 mm and a diameter of 4 inches.

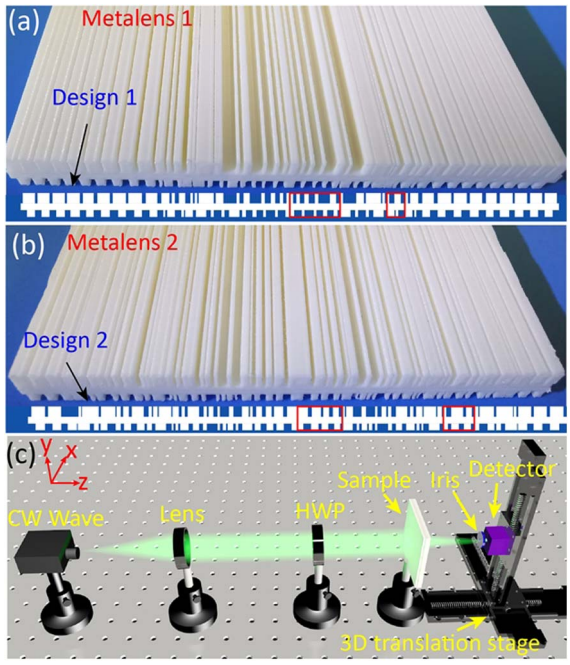

Fig. 3. (a) Photograph of 3D-printed metalens 1 and its designed cross section. (b) Photograph of 3D-printed metalens 2 and its designed cross section. The segments analyzed in Fig. 2 are marked by the red boxes. (c) Experimental setup for testing the metalenses.

A half-wave plate is used to convert the polarization from the y to the x direction, so that the electric field is perpendicular to the ridges of cylindrical metalenses. To measure the field distribution after the metalens, a Schottky diode detector with an iris (diameter of 1 mm) is mounted on a three-dimensional translation stage and scanned with the step of 0.25 mm in the x, y, and z directions. For terahertz imaging, the imaging sample is mounted on the translation stage for scanning along the x and y directions in the focal plane. The detector is fixed at the optical axis after the sample for detection. A lock-in amplifier is connected to the source and the detector for modulation and lock-in detection.

For better comparison, simulations of the two metalenses are carried out in Lumerical FDTD Solutions with the Gaussian beam source polarized along the x direction. The beam waist is 30 mm, which is consistent with that in the experiment. Perfectly matched layer boundary conditions are used in the x and z directions, and the periodic boundary condition is used in the y direction. The exit interface of the metalens is set as $z = 0$. The simulated and measured focusing performances of the metalens 1 are shown in Figs. 4(a) and 4(b). The focal length is exactly 15.0 mm in simulation and 14.75 mm in measurement. The simulated FWHM of the x component is 1.15 mm (0.54λ), equal to the diffraction-limited value. The FWHM becomes 1.55 mm (0.72λ) if the total intensity is considered. This is contributed by the strong longitudinal component away from the center due to the vector property of the field [33,34]. In experiment, the detector is responsive to the intensity of the total electric field, so the measured FWHM well agrees with the simulated one. Here we

define the focusing efficiency as the ratio between the focused power and the input power, where the focused power is the transmitted power within the radius of 3 times of the FWHM in the focal plane [6]. Simulated focusing efficiency is 52% when FWHM = 1.55 mm is considered. The measured focusing efficiency is 29.0%. Further study reveals that the discrepancy between the simulated and measured efficiency is likely caused by the fabrication error. The actual dimensions of the metalens show random deviation from the designed ones within ±0.2 mm. When the dimension error is considered for the metalens ridges and air gaps, the simulated focusing efficiency drops to 39% and the focal length remains fixed. So there is still room for improvement of the focusing efficiency in experiment by increasing the fabrication accuracy. Considering the simulated focusing efficiency above 40%, the focusing bandwidth covers 0.12–0.15 THz. The focal length is in linear proportion with the frequency, which is a signature of diffractive lenses.

In addition, the simulated focusing efficiency of 52% of the metalens 1 seems not a significant improvement if compared with some state-of-the-art metalenses in the optical region with similar NA designed using the LPM method [6,35,36]. However, they are all made of high-index dielectric on low-index substrate with very large aspect ratio, which is challenging at the terahertz band [10]. If we redesign the metalens 1 using the LPM method, the focusing efficiency is only 25.2%. So the 52% focusing efficiency does indicate the effectiveness of nLDM in boosting the performance of high-NA metalenses. In addition, the focusing efficiency can be improved to 59% and 71% in silicon-on-silicon metalenses and silicon-on-SiO_2 metalenses based on the nLDM method, as observed in Appendix B. Therefore, for any high-NA metalens with fixed material and fixed thickness, the nonlocal diffraction design in segments will always be more efficient than the local phase modulation of subwavelength elements. Another advantage is that the nLDM makes it possible to use very low-index dielectrics, such as plastic, glass, or photoresists, to build high-NA metalenses.

The simulated and measured field distributions through metalens 2 are shown in Figs. 4(d) and 4(e), which show excellent coincidence. Figure 4(f) plots the intensity along the optical axis, where four peaks can be identified. However, since the second and third peaks are very close, they cannot be distinguished through the measurement. Still the general trend and the relative intensity of the foci agree well with the simulation.

4. 3D TIGHT FOCUSING AND IMAGING

To further generalize the functionality from 2D focusing into full-space 3D focusing, the cylindrical metalens can simply be rotated around its center forming the concentric ring metalens. To ensure good quality of 3D tight focusing in the concentric ring metalens with linearly polarized excitation, it is necessary to have polarization-insensitive 2D metagratings, meaning the same diffraction patterns for both TE and TM polarization states. Since the metalens 1 is only optimized for TM polarization, we redesigned all the 14 segments in Fig. 2(c) by minimizing the cost function in both polarization states as $\text{cost} = \sum[|t^d_m - t^p_{m\_\text{TM}}(\boldsymbol{g})| + |t^d_m - t^p_{m\_\text{TE}}(\boldsymbol{g})|]$. As shown in Fig. 5(b), the optimized diffraction efficiencies in the dominate order (−1st order) for TE and TM polarizations are close to each other in each segment, and the diffraction phases of both polarizations are close to the target value π. The thicknesses of

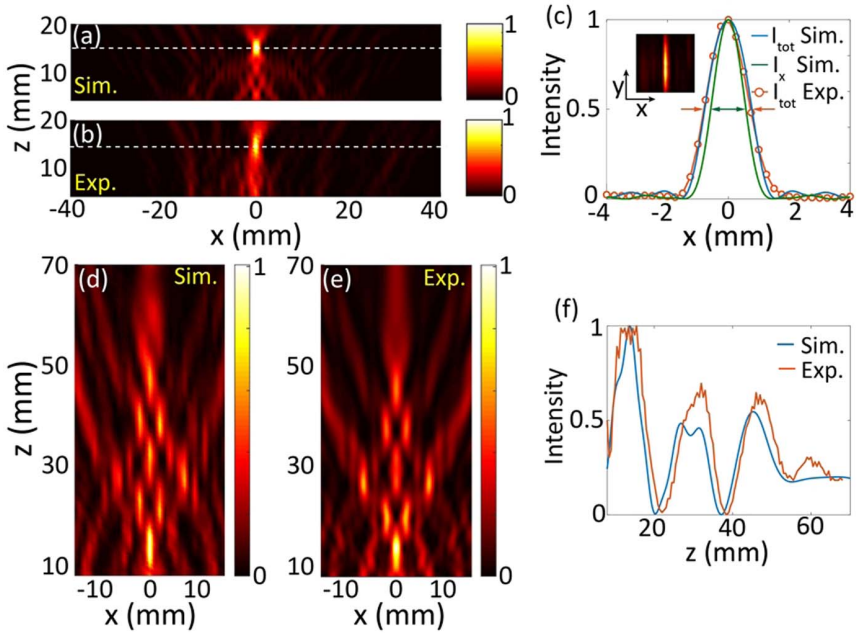

Fig. 4. (a) Simulated and (b) measured total electric field intensity distribution of metalens 1 in the *xz* plane. (c) Normalized intensity distribution along *x* in the focal plane. The simulated transverse field intensity and total intensity distribution are different due to the contribution of the longitudinal field component. The inset is the measured field distribution in the focal plane, showing a cylindrical focusing effect. (d) Simulated and (e) measured intensity distribution of metalens 2. (f) Intensity distribution along the optical axis.

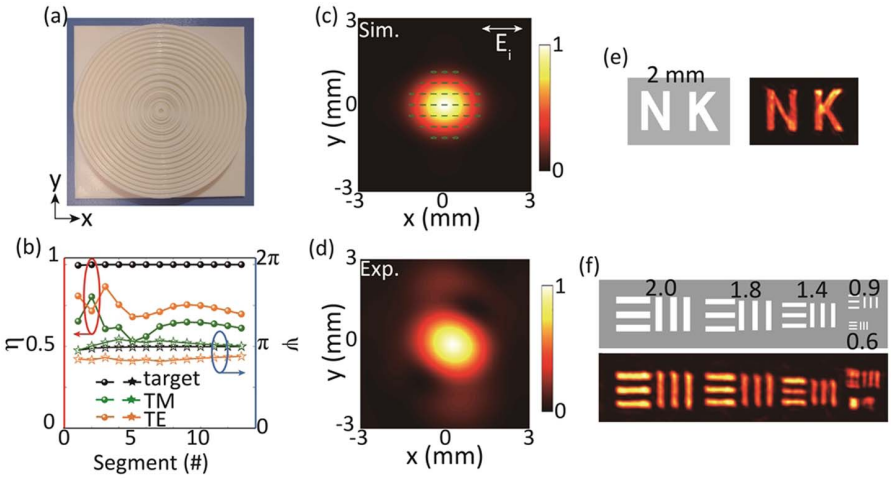

Fig. 5. (a) Photograph of 3D-printed polarization-insensitive metalens 3 with NA = 0.94. (b) −1st order diffraction efficiencies and phases of constituent segments in the metalens. The dotted lines are diffraction efficiencies η and star lines are diffraction phases ψ. Black ones are the target from Eq. (2), while green and orange ones are the practically optimized ones for TM and TE polarization, respectively. (c) Simulated and (d) measured intensity distribution in the focal plane. Polarization of the focus is superimposed in (c). (e), (f) Samples and measured images with different linewidths.

the bottom ridges, the slab, and the top ridges are 1.6 mm, 3.5 mm, and 1.6 mm, respectively. These polarization-insensitive metagratings are stitched and rotated into the metalens 3 in Fig. 5(a). This metalens is different from traditional Fresnel lenses realized by varying the thickness [37], as a full set of diffraction efficiencies and phases in each 2π zone are finely tuned. The focal spot is simulated and measured in Figs. 5(c) and 5(d). The polarization distribution of the focus is superimposed on the pseudo color of Fig. 5(c), showing linear polarization the same as the excited one.

To characterize the imaging performance of the metalens, we mount the imaging sample on a translation stage and scan it in the focal plane of the 3D metalens. A Schottky detector with a pin hole is fixed behind the sample to record the field distribution. The sample in Fig. 5(e) is a metallic sheet with etched "NK" letters whose linewidth is 2 mm. The detected intensity distribution can well resolve the letters. Next, we change the imaging sample into a resolution chart in Fig. 5(f). From left to right, the width of the bar is 2.0 mm, 1.8 mm, 1.4 mm, 0.9 mm, and 0.6 mm. The detected image shows a resolution of 1.4 mm, which is 0.65λ. This agrees with the diffraction limit for NA = 0.94 considering the vectorial property.

5. DISCUSSION AND CONCLUSION

We notice that each segment is designed and optimized in a periodic environment, while the whole device is aperiodic. This will inevitably bring some error. If we take the segmented metalens 1 as an initial solution and further numerically optimize the entire structure, the focusing efficiency is slightly boosted from 52% to 60%, indicating the rationality of the periodic assumption. This assumption now is in the superwavelength scale, therefore being better satisfied than the subwavelength periodic assumption in the LPM. For the case when the phase segment is drastically different from the neighboring ones, one can merge it to the neighboring one to form a larger segment. It will make the wavefront shaping more accurate. Nevertheless, the periodic assumption makes it possible to translate a continuous phase segment into a few discrete diffraction coefficients, which significantly reduce the information dimension, and simplify the optimization.

We know the LPM-based metasurface can be readily generated regardless of the device size once some basic subwavelength elements are found. In the nLDM, the optimization time will increase linearly with the scale of the metasurface, since each segment is optimized separately. However, the average optimization time for each segment is around 7 s. It is not a problem to deal with large-scale devices. In contrast, global optimization of the full metalens, though being a good solution to push the NA and to improve the efficiency [38–41], needs extensive computational resources and severely limits the scale of the devices. The proposed nLDM strategy is much more time-efficient in design and provides rich physical insights into the wavefront transformation.

Finally, the nonlocal diffraction engineering method is general enough to be applied at other frequency bands. When working at higher frequencies, it is possible to etch both sides of the glass plate or to build bi-layer photoresist structures similar to the one demonstrated here. Instead, one can simply choose high-index dielectric monolayer structures for nonlocal diffraction engineering. In Appendix B, we also give a monolayer metalens made of silicon with improved efficiency when designed using the nonlocal diffraction engineering, which can be scaled to optical frequencies and fabricated by conventional lithography techniques.

To summarize, we demonstrate that arbitrary large-gradient wavefront shaping can be accurately and effectively achieved from the nonlocal diffraction engineering of the superwavelength-scale metagratings. Different from the local phase

modulation where each phase point corresponds to a subwavelength element, the proposed method here establishes correlation between the phase segments and the diffraction patterns, and further finds the proper metagrating geometry to fulfill the diffraction properties including diffraction angles, diffraction efficiencies, and diffraction phases. High-NA cylindrical metalenses composed of a series of metagratings are designed and experimentally tested, showing the effectiveness and flexibility of this method for arbitrary 2D wavefront shaping. The design is further extended to 3D metalenses by utilizing polarization-insensitive metagratings, whose application in high-resolution imaging is well characterized. These findings indicate that the nonlocal diffraction engineering offers a new perspective on beam transformation over metasurface devices, well suited for large-gradient cases.

APPENDIX A: PHASE PROFILE REPRODUCTION BASED ON THE nLDM

As aforementioned, metalens 1 contains 14 metagrating segments, and metalens 2 contains 11 metagrating segments. In order to show the accuracy of the optimized metagratings, the phase profile after each metagrating is shown piecewise on top of the ideal phase profile in Figs. 6(a) and 6(c). In addition, the target and practical diffraction behaviors in the dominate order (−1st order) of all the metagrating segments are compared in Fig. 6(b) for metalens 1 and Fig. 6(d) for metalens 2. The practical diffraction efficiencies η_{-1} are slightly smaller than the ideal ones in most of the segments, but all the diffraction phases are accurately implemented.

Despite the relatively good consistency between the ideal phase and the practical one, it is still beneficial to analyze the origin of the fitting errors, which seem stronger in the nonlinear-shaped phase segments. The nLDM contains two steps. The first step is to find the diffraction coefficients corresponding to the phase segment, by Eq. (2) through Fourier series expansion. It means the desired wavefront can be expanded into a series of plane waves with the diffraction coefficients as the complex amplitudes and the diffraction angles as the propagation directions. The second step is to optimize the metagrating geometry so that its diffraction coefficient in each order best fits the desired one. Both steps may bring some errors. In the first step, only a finite number of plane waves corresponding to the propagation diffraction orders are considered. The drop of evanescent orders may introduce error depending on the shape of the phase profile. In the second step, the optimized metagrating may not show the desired diffraction pattern limited by the degree of freedom.

For generality, we choose the first segment (a quadratic shape) in Fig. 6(a), the fourth segment (a linear shape) in Fig. 6(a), and the first segment (an arbitrary shape) in Fig. 6(c) to characterize the origin of the fitting error. We expand the ideal phase segment by plane waves corresponding to the propagation diffraction orders. The superposition of the plane waves gives a theoretical phase profile as the purple lines in Fig. 7. It is observed that the quadratic- and linear-shaped phases can be reproduced well in Figs. 7(a) and 7(b), while drop of the evanescent waves does bring some error to the arbitrary-shaped phase in Fig. 7(c). Adding more evanescent plane waves gradually pulls the curve to the ideal case, as shown by the purple dashed line in Fig. 7(c).

The practical phases realized by the optimized metagratings are shown by the green curves. The optimization brings some error to all the three phase segments, and it is more pronounced in the nonlinear shaped ones in Figs. 7(a) and 7(d). This is because the linear-shaped phase segment corresponds to a single-order plane wave. The amplitude deviation does not affect the phase profile as long as the diffraction in other orders is suppressed. The nonlinear-shaped phase is implemented by the superposition of several plane waves. So the deviation in amplitude or phase of any diffraction order may affect the final

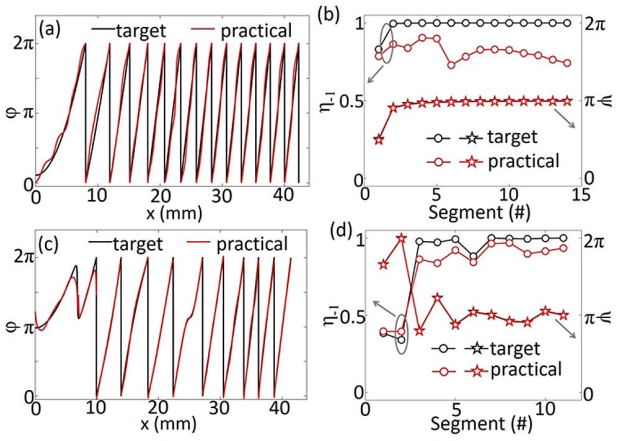

Fig. 6. (a) Target phase profile (black) of metalens 1 and the reproduced phase segments (red) by the optimized metagratings. (b) Summarized diffraction efficiencies and phases of the −1st order for all the segments in metalens 1. (c) Target phase profile (black) of metalens 2 and the reproduced phase segments (red) by the optimized metagratings. (d) Summarized diffraction efficiencies and phases of the −1st order for all the segments in metalens 2.

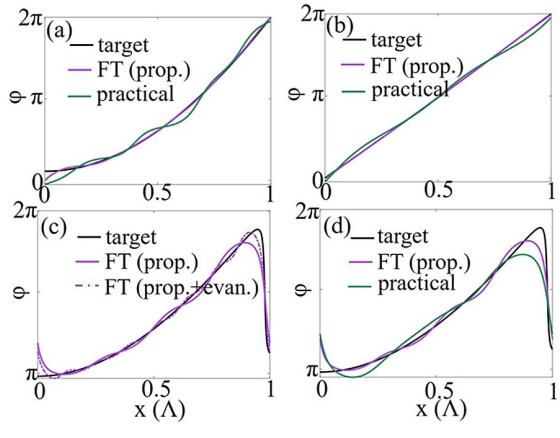

Fig. 7. Detailed comparison of the target and reproduced phase segments of different shapes. (a) The first segment in Fig. 6(a). (b) The fourth segment in Fig. 6(a). (c), (d) The first segment in Fig. 6(c). The black lines are the target phase curves. The purple lines are the results of plane wave superposition in all the propagation orders. The purple dash line in (c) is the result of plane wave superposition in all the propagation orders and additional 8 evanescent orders. The green lines are the practical phases realized by the optimized metagratings.

shape of the wavefront. In addition, the analysis of the arbitrary-shaped phase segment in Figs. 7(c) and 7(d) shows that the starting and ending points of the phase segment can be arbitrary and are not necessary to be 0 or 2π.

APPENDIX B: COMPARISON OF LPM- AND nLDM-BASED METALENSES MADE OF DIFFERENT DIELECTRICS

Here we compare the focusing performance of the same metalens designed using the LPM and nLDM methods. In order to show that the nLDM is better than LPM no matter of the metalens material, we carry out this comparison in low-index metalenses made of PLA and in high-index metalenses made of silicon, respectively.

The performance of the nLDM-based metalens made of PLA has been studied. Here we show the design and response of the LPM-based metalens made of PLA. Figure 8(a) shows the transmission amplitude and phase of the subwavelength PLA grating elements under different duty cycles W/Λ. The period Λ of the element is 0.65λ. The transmission amplitude is near unity and the phase response covers 2π. As shown by the stars in Fig. 8(a), eight elements with fixed phase difference of 0.25π are selected to reproduce the phase profile of the metalens with the same size and NA as that of metalens 1. The geometries of the nLDM-based metalens 1 and the LPM-based metalens are shown in Fig. 8(c), and the intensities in the focal plane of the two metalenses are shown in Fig. 8(b). The FWHMs of the focal spots in the two metalenses are both 0.72λ. The focusing efficiency improves from 25.2% in the LPM-based one to 52% in the nLDM-based one.

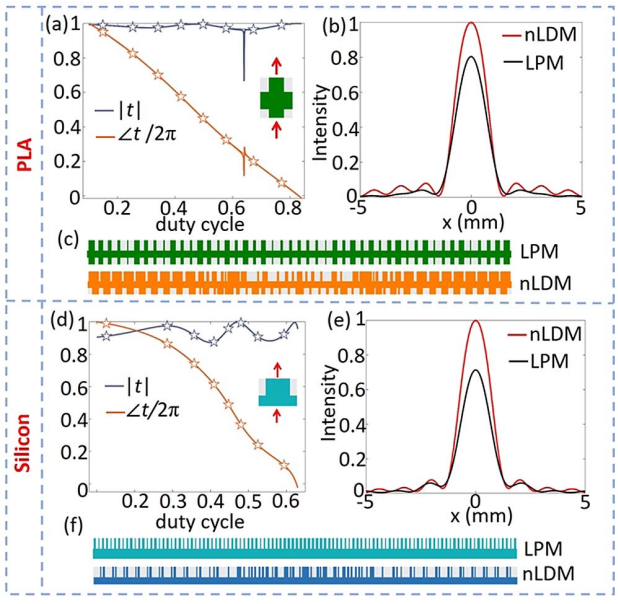

Fig. 8. (a) Transmission and phase response of subwavelength PLA elements with different duty cycles. The stars show eight elements with fixed phase difference of 0.25π to reproduce the phase profile of the metalens with NA of 0.94. The inset is one element. (b) Simulated normalized intensity distributions on the focal plane. The black line is for the LPM-based metalens and the red line is for the nLDM-based metalens 1 in the main text. (c) Geometries of the nLDM-based metalens 1 and the LPM-based metalens. (d)–(f) The same information as (a)–(c) for metalenses made of silicon.

Similarly, two metalenses made of silicon (refractive index of 3.5) are designed by the LPM and nLDM methods, respectively. The thicknesses of the silicon grating ridges and the silicon substrate are chosen as 0.7λ and 0.45λ. For the LPM design, the subwavelength grating has the period of 0.37λ. Eight elements are chosen to cover the 2π phase delay from Fig. 8(d) and are arranged properly for desired spatial phase modulation. For the nLDM design, 14 metagratings are optimized in sequence, and each metagrating is targeted to achieve the same set of diffraction parameters as metalens 1 does. The geometries of the silicon metalenses designed using the two methods are shown in Fig. 8(f). Their intensity distributions along the focal plane are calculated by FDTD in Fig. 8(e). The simulated focusing efficiency is 32.6% for the LPM-based one and 59% for the nLDM-based one. If the substrate is replaced as SiO_2, the focusing efficiency can be further boosted to 71%. So no matter what dielectric material the metalens is made of, the nLDM mechanism is more effective in large-gradient phase modulation, and can further boost the efficiency of the metadevices.

Funding. National Key Research and Development Program of China (2017YFA0701000); National Natural Science Foundation of China (62175118, 61831012, 61805123).

Disclosures. The authors declare no conflicts of interest.

Data Availability. The data used to support the findings of this study are available from the corresponding author upon request.

REFERENCES

1. H. Chen, A. J. Taylor, and N. Yu, "A review of metasurfaces: physics and applications," Rep. Prog. Phys. **79**, 076401 (2016).
2. S. Chang, X. Guo, and X. Ni, "Optical metasurfaces: progress and applications," Annu. Rev. Mater. Res. **48**, 279–302 (2018).
3. F. Ding, A. Pors, and S. I. Bozhevolnyi, "Gradient metasurfaces: a review of fundamentals and applications," Rep. Prog. Phys. **81**, 026401 (2018).
4. M. Decker, I. Staude, M. Falkner, J. Dominguez, D. N. Neshev, I. Brener, T. Pertsch, and Y. S. Kivshar, "High-efficiency dielectric Huygens' surfaces," Adv. Opt. Mater. **3**, 813–820 (2015).
5. J. Cheng, D. Ansari-Oghol-Beig, and H. Mosallaei, "Wave manipulation with designer dielectric metasurfaces," Opt. Lett. **39**, 6285–6288 (2014).
6. A. Arbabi, E. Arbabi, S. M. Kamali, Y. Horie, S. Han, and A. Faraon, "Miniature optical planar camera based on a wide-angle metasurface doublet corrected for monochromatic aberrations," Nat. Commun. **7**, 13682 (2016).
7. Z. Fan, Z. Shao, M. Xie, X. Pang, W. Ruan, F. Zhao, Y. Chen, S. Yu, and J. Dong, "Silicon nitride metalenses for close-to-one numerical aperture and wide-angle visible imaging," Phys. Rev. Appl. **10**, 014005 (2018).
8. M. Khorasaninejad, W. Chen, R. Devlin, J. Oh, A. Y. Zhu, and F. Capasso, "Metalenses at visible wavelengths: diffraction limited focusing and subwavelength resolution imaging," Science **352**, 1190–1194 (2016).
9. M. Khorasaninejad and F. Capasso, "Metalenses: versatile multifunctional photonic components," Science **358**, eaam8100 (2017).

10. H. Chen, Z. Wu, Z. Li, Z. Luo, X. Jiang, Z. Wen, L. Zhu, X. Zhou, X. H. Li, Z. Shang, Z. Zhang, K. Zhang, G. Liang, S. Jiang, L. Du, and G. Chen, "Sub-wavelength tight-focusing of terahertz waves by polarization-independent high numerical-aperture dielectric metalens," Opt. Express **26**, 29817–29825 (2018).
11. H. Zhang, X. Zhang, Q. Xu, C. Tian, Q. Wang, Y. Xu, Y. Li, J. Gu, Z. Tian, C. Ouyang, X. Zhang, C. Hu, J. Han, and W. Zhang, "High-efficiency dielectric metasurfaces for polarization-dependent terahertz wavefront manipulation," Adv. Opt. Mater. **6**, 1700773 (2018).
12. X. Cai, R. Tang, H. Zhou, Q. Li, S. Ma, D. Wang, T. Liu, X. Ling, W. Tan, Q. He, S. Xiao, and L. Zhou, "Dynamically controlling terahertz wavefronts with cascaded metasurfaces," Adv. Photonics **3**, 036003 (2021).
13. G. Y. Lee, J. Sung, and B. Lee, "Recent advances in metasurface hologram technologies," ETRI J. **41**, 10–22 (2019).
14. G. Zheng, H. Mühlenbernd, M. Kenney, G. Li, T. Zentgraf, and S. Zhang, "Metasurface holograms reaching 80% efficiency," Nat. Nanotechnol. **10**, 308–312 (2015).
15. J. Yang and J. A. Fan, "Analysis of material selection on dielectric metasurface performance," Opt. Express **25**, 23899–23909 (2017).
16. V. S. Asadchy, M. Albooyeh, S. N. Tcvetkova, A. Díaz-Rubio, Y. Ra'di, and S. A. Tretyakov, "Perfect control of reflection and refraction using spatially dispersive metasurfaces," Phys. Rev. B **94**, 075142 (2016).
17. H. Liang, A. Martins, B.-H. V. Borges, J. Zhou, E. R. Martins, J. Li, and T. F. Krauss, "High performance metalenses: numerical aperture, aberrations, chromaticity, and trade-offs," Optica **6**, 1461–1470 (2019).
18. V. S. Asadchy, A. Wickberg, A. Díaz-Rubio, and M. Wegener, "Eliminating scattering loss in anomalously reflecting optical metasurfaces," ACS Photonics **4**, 1264–1270 (2017).
19. Y. Ra'di, D. L. Sounas, and A. Alu, "Metagratings: beyond the limits of graded metasurfaces for wave front control," Phys. Rev. Lett. **119**, 067404 (2017).
20. A. Díaz-Rubio, V. S. Asadchy, A. Elsakka, and S. A. Tretyakov, "From the generalized reflection law to the realization of perfect anomalous reflectors," Sci. Adv. **3**, e1602714 (2017).
21. A. M. H. Wong and G. V. Eleftheriades, "Perfect anomalous reflection with a bipartite Huygens' metasurface," Phys. Rev. X **8**, 011036 (2018).
22. D. Sell, J. Yang, S. Doshay, R. Yang, and J. A. Fan, "Large-angle, multifunctional metagratings based on freeform multimode geometries," Nano Lett. **17**, 3752–3757 (2017).
23. Z. Fan, M. R. Shcherbakov, M. Allen, J. Allen, B. Wenner, and G. Shvets, "Perfect diffraction with multiresonant bianisotropic metagratings," ACS Photonics **5**, 4303–4311 (2018).
24. X. Dong, J. Cheng, F. Fan, and S. Chang, "Low-index second-order metagratings for large-angle anomalous reflection," Opt. Lett. **44**, 939–942 (2019).
25. Z. Zhang, M. Kang, X. Zhang, X. Feng, Y. Xu, X. Chen, H. Zhang, Q. Xu, Z. Tian, W. Zhang, A. Krasnok, J. Han, and A. Alù, "Coherent perfect diffraction in metagratings," Adv. Mater. **32**, 2002341 (2020).
26. V. Neder, Y. Ra'Di, A. Alù, and A. Polman, "Combined metagratings for efficient broad-angle scattering metasurface," ACS Photonics **6**, 1010–1017 (2019).
27. 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, and A. I. Kuznetsov, "A metalens with a near-unity numerical aperture," Nano Lett. **18**, 2124–2132 (2018).
28. M. Kang, Y. Ra'di, D. Farfan, and A. Alù, "Efficient focusing with large numerical aperture using a hybrid metalens," Phys. Rev. Appl. **13**, 044016 (2020).
29. X. Dong, J. Cheng, F. Fan, X. Wang, and S. Chang, "Efficient wideband large-angle refraction and splitting of a terahertz beam by low-index 3D-printed bilayer metagratings," Phys. Rev. Appl. **14**, 014064 (2020).
30. S. Larouche and D. R. Smith, "Reconciliation of generalized refraction with diffraction theory," Opt. Lett. **37**, 2391–2393 (2012).
31. X. Dong, J. Cheng, F. Fan, Z. Zhang, Y. Liu, X. Wang, and S. Chang, "Extremely large-angle beam deflection based on low-index sparse dielectric metagratings," J. Phys. D **53**, 245101 (2020).
32. X. Dong, J. Cheng, Y. Yuan, F. Fan, X. Wang, and S. Chang, "An efficient bi-functional metagrating via asymmetric diffraction of terahertz beams," IEEE Photonics Technol. Lett. **33**, 441–444 (2021).
33. F. Zhang, H. Yu, J. Fang, M. Zhang, S. Chen, J. Wang, A. He, and J. Chen, "Efficient generation and tight focusing of radially polarized beam from linearly polarized beam with all-dielectric metasurface," Opt. Express **24**, 6656–6664 (2016).
34. Z. Wang, Q. Li, and F. Yan, "A high numerical aperture terahertz all-silicon metalens with sub-diffraction focus and long depth of focus," J. Phys. D **54**, 085103 (2020).
35. H. Liang, Q. Lin, X. Xie, Q. Sun, Y. Wang, L. Zhou, L. Liu, X. Yu, J. Zhou, T. F. Krauss, and J. Li, "Ultrahigh numerical aperture metalens at visible wavelengths," Nano Lett. **18**, 4460–4466 (2018).
36. S. J. Byrnes, A. Lenef, F. Aieta, and F. Capasso, "Designing large, high-efficiency, high-numerical-aperture, transmissive meta-lenses for visible light," Opt. Express **24**, 5110–5124 (2016).
37. E. Noponen, J. Turunen, and A. Vasara, "Electromagnetic theory and design of diffractive-lens arrays," J. Opt. Soc. Am. A **10**, 434–443 (1993).
38. S. Banerji and B. Sensale-Rodriguez, "A computational design framework for efficient, fabrication error-tolerant, planar THz diffractive optical elements," Sci. Rep. **9**, 5801 (2019).
39. M. Mansouree, A. McClung, S. Samudrala, and A. Arbabi, "Large-scale parametrized metasurface design using adjoint optimization," ACS Photonics **8**, 455–463 (2021).
40. Z. Lin, B. Groever, F. Capasso, A. W. Rodriguez, and M. Lončar, "Topology-optimized multilayered metaoptics," Phys. Rev. Appl. **9**, 044030 (2018).
41. H. Chung and O. D. Miller, "High-NA achromatic metalenses by inverse design," Opt. Express **28**, 6945–6965 (2020).