OPTICA

Modeling off-axis diffraction with the least-sampling angular spectrum method

Haoyu Wei,[1,†] Xin Liu,[1,2,†] Xiang Hao,[2] Edmund Y. Lam,[1,3] and Yifan Peng[1,*]

[1]Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong SAR, China
[2]College of Optical Science and Engineering, Zhejiang University, Hangzhou, China
[3]elam@eee.hku.hk
[†]These authors contributed equally to this work.
*evanpeng@hku.hk

Received 23 March 2023; revised 22 May 2023; accepted 25 June 2023; published 19 July 2023

Accurately yet efficiently simulating off-axis diffraction is vital to design large-scale computational optics, but existing rigid sampling and modeling schemes fail to address this. Herein, we establish a universal least-sampling angular spectrum method that enables efficient off-axis diffraction modeling with high accuracy. Specifically, by employing the Fourier transform's shifting property to convert off-axis diffraction to quasi-on-axis, and by linking the angular spectrum to the transfer function, essential sampling requirements can be thoroughly optimized and adaptively determined across computation. Leveraging a flexible matrix-based Fourier transform, we demonstrate the off-axis point spread function of exemplary coded-aperture imaging systems. For the first time, to our knowledge, a significant speed boost of around 36× over the state of the art at 20° is demonstrated, and so is the viability of computing ultra-large angles such as 35° within seconds on a commercial computer. The applicability to high-frequency modulation is further investigated.

https://doi.org/10.1364/OPTICA.490223

Incorporating compact, computing-empowered diffractive optics into a wide range of imagers and sensory systems has emerged into a demanding research field, with its promising applications in holography [1], lens design [2,3], optical computing [4], and computational imaging [5–7]. However, most reported solutions are demonstrated with a limited field of view, due to the lack of an accurate yet efficient off-axis diffraction modeling scheme in the design space. That said, although fundamentals of diffraction have been well established [8,9] with many algorithms developed for numerical implementation [10,11], modeling light propagation at non-paraxial and large-scale fields is still problematic, requiring a significantly large number of samples. Even off-the-shelf optics design software such as Zemax [12] cannot faithfully support the simulation in a computing resource-efficient manner.

To simultaneously quest for computational accuracy and efficiency, one often adopts the angular spectrum method (ASM) based on fast Fourier transform (FFT) for off-axis diffraction modeling [9]. Unfortunately, conventional FFT-based ASM needs non-trivial zero-padding to avoid aliasing [13], suffering from a severe trade-off between computational efficiency and accuracy. Representative solutions to alleviate this sampling problem include the shifted band-limited ASM [14], generalized transfer function method [15], and shifted band-extended ASM (shift-BEASM) [16]. However, the samples are still rigidly determined subject to the respective requirements of the naive incident field and transfer function, which overlooks the intrinsic characteristics of the incident field and the link between the angular spectrum and the transfer function, resulting in excessive oversampling or harmful undersampling. More specifically, these works assume the sampling of the transfer function is independent of the incident field, which is incomplete since the correct computation of the final diffractive field requires a proper sampling of the product of the angular spectrum and the transfer function, rather than that from independent derivations.

To address the above issues, herein, we make the following contributions: (1) establishing an effective modeling scheme that enables accurate, fast, and flexible wave propagation simulation with a large incident angle, termed *least-sampling ASM (LS-ASM)*, which is achieved by identifying the optimal sampling strategy of ASM; (2) employing the shifting property of the Fourier transform to eliminate redundant samples induced by the off-axis configuration; (3) developing a universal approach to identifying both lower and upper sampling bounds that ensure correct results for all types of fields. The proposed method is implemented with a matrix-based discrete Fourier transform, enabling an arbitrary selection of sampling windows and intervals.

Unlike prior work, where the off-axis diffraction is considered with plane waves, we start with a more general and practical case. A spherical wave emitted from an off-axis point source at (x_0, y_0, z_0) impinges the aperture plane with the incident angle defined by $\sin\theta_x = x_0/r_0$ and $\sin\theta_y = y_0/r_0$ ($r_0 = \sqrt{x_0^2 + y_0^2 + z_0^2}$ is the distance from the point source to the origin of the aperture plane O), yielding a diffractive field on the observation plane [Fig. 1(a)]. Here we assume a circular aperture with a diameter of D for simplicity. Using ASM, the diffractive field at a distance z can be calculated by [9]

$$u_z(x, y) = \mathcal{F}^{-1}[U(f_X, f_Y)\, H_z(f_X, f_Y)]\,, \tag{1}$$

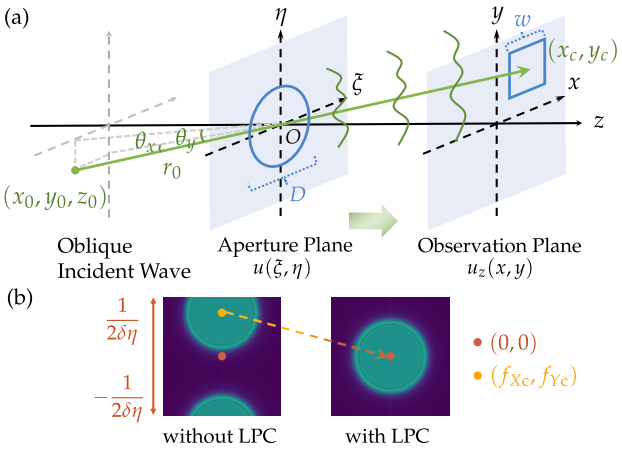

Fig. 1. (a) Diagram of off-axis diffraction. (b) The LPC shifts the angular spectrum of the input field to the frequency center. $\delta\eta$ is the sampling interval along η axis in the spatial domain.

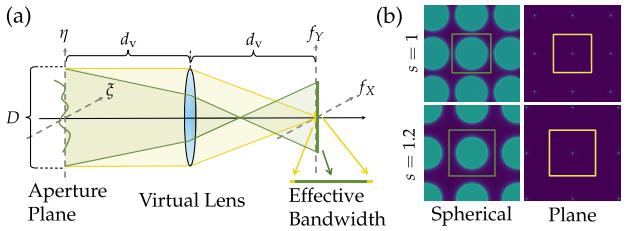

Fig. 2. (a) Fourier transform interpreted as a virtual thin lens model. The green and yellow areas denote spherical and plane waves, respectively. (b) Influence of oversampling factor. The inner squares indicate the sampling region of a spherical wave (left) and a plane wave (right) with oversampling factor s of one (top) and 1.2 (bottom), respectively. A larger region is visualized than the sampling rate to illustrate the aliasing artifact.

where u_z is the optical field on the observation plane (x, y). $U(f_X, f_Y) = \mathcal{F}[u(\xi, \eta)]$ is the angular spectrum, in the (spatial) frequency domain (f_X, f_Y) of the field right after the aperture u defined in the spatial domain (ξ, η), which we call the input field. Here $\mathcal{F}$ represents the Fourier transform operation, and $H_z(f_X, f_Y) = \exp[ikz\sqrt{1-(\lambda f_X)^2-(\lambda f_Y)^2}]$ is the transfer function of the angular spectrum, where $k = 2\pi/\lambda$ is the wavenumber associated with the wavelength λ.

In the scope of this paper, we mainly discuss the sampling of the input field's phase with the temporal assumption that the amplitude is a gradually varying function of (ξ, η) and the phase is smooth. Without loss of generality, here we employ an additional phase modulation to the incident field, which is denoted by ϕ_{mod}. Then the phase of the input field can be expressed as

$$\phi_u = \phi_{\text{inc}} + \phi_{\text{mod}}, \tag{2}$$

where $\phi_{\text{inc}}(\xi, \eta) = k\sqrt{(\xi - x_0)^2 + (\eta - y_0)^2 + z_0^2}$ is the phase of the incident wave.

Analyzing the effective angular spectrum of the incident wave reveals that it has asymmetric coordinates relative to the origin, primarily caused by a linear phase component. This asymmetry causes an increase in the sampling rate as the effective angular spectrum is located in the high-frequency area, according to the Nyquist sampling theorem [9]. To tackle this problem, we apply a linear phase compensation (LPC) to the input field, such that the center of the effective angular spectrum can be shifted to the origin, as determined through an exploration of the shifting property of the Fourier transform, i.e.,

$$\phi_{\text{on-axis}} = \phi_u + \phi_{\text{comp}}, \tag{3}$$

where $\phi_{\text{comp}}(\xi, \eta) = -2\pi(f_{Xc}\xi + f_{Yc}\eta)$, and the center of the effective angular spectrum is (f_{Xc}, f_{Yc}). In this way, the off-axis wave can be converted to an equivalent quasi-on-axis one, whose Fourier transform is still the effective angular spectrum of the input field but requires many fewer samples. Notably, the quasi-on-axis wave refers to any (compensated) one with the angular spectrum centering around zero frequency, such as most on-axis waves and some off-axis ones possessing intrinsic frequency offsets, the angular spectra of which are shifted back to zero (Supplement 1, Section 1A). Herein, we refer to "quasi-on-axis" as "on-axis" for short.

The Nyquist sampling barrier of the input field is now lifted. In contrast, without LPC, the naive ϕ_u requires additional samples due to the band asymmetry to avoid aliasing. Figure 1(b) illustrates that sampling an off-axis spherical wave following the reduced requirement results in an aliased angular spectrum, while with LPC, we correctly calculate it.

Although the LPC reduces the sampling requirement of off-axis waves, they still need to be properly sampled according to the requirement of on-axis waves. Notably, while phase gradient analysis can well determine the on-axis wave's bandwidth, the truncation by the finite-size aperture broadens the angular spectrum to the infinite. Here we seek a straightforward solution based on the optical insight that a positive thin lens's back focal plane is the Fourier transform of the field at the front focal plane [9].

We apply a *virtual lens* to focus light in a limited region on the back focal plane, acting as a Fourier transform operator and identifying the effective bandwidth of the input field, i.e., the on-axis wave [Fig. 2(a)]. The angular spectrum of the input field can be viewed as the ideal one convolved with the Fraunhofer diffraction field of the aperture without modulation. Therefore, the effective bandwidth is the sum of the phase gradient term and the size of the Airy disk. Given the focal length of virtual lens d_v, the effective width of the Airy disk is specified as $W = \frac{41.2\lambda d_v}{D}$, which covers around 99% energy [8].

The virtual lens can also determine the effective bandwidth of a spherical wave under the paraxial assumption, but it may be inaccurate in cases where the assumption is invalid or when the input field has a complicated phase. In such cases, we adopt the phase gradient analysis, which determines bandwidth from maximum gradients (Supplement 1, Sections 1B and 1C). Finally, the sampling rate of the input field can be well determined as

$$\tau_{\text{spat}}^{(p)} = s f_{\text{b}}^{(p)} = s\frac{W}{\lambda d_v} = s\left(\frac{1}{\pi}\left|\frac{\partial}{\partial p}\phi_{\text{on-axis}}\right|_{\max} + \frac{41.2}{D}\right), \tag{4}$$

where $p \in \{\xi, \eta\}$, $f_{\text{b}}^{(p)}$ is the effective bandwidth, and $s \geq 1$ is a heuristic oversampling factor. As mentioned, the input field is not band-limited, so aliasing may still occur at high frequencies even if involving most energy by sampling with Eq. (4). Thus, the oversampling factor s is incorporated to alleviate this problem. The influence of s is shown in Fig. 2(b). Given $s = 1$, we observe noticeable aliasing at the high frequency as expected. With an increased s, e.g., 1.2, this problem can be greatly alleviated.

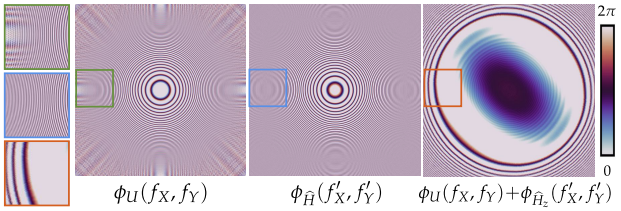

Fig. 3. Effect of combined sampling in frequency domain. From left to right, the first column is the zoomed view from the right ones. Second to fourth columns are the phase of the angular spectrum of an off-axis converging spherical wave with LPC, the shifted transfer function, and the superposition, respectively.

After establishing an optimal sampling for the input field in the spatial domain, we examine the counterpart in the frequency domain. We demonstrate that manipulating the observation window (OW) can substantially decrease the sampling requirement in this domain. Specifically, we virtually shift the OW from (x_c, y_c) to $(x_c - \Delta x, y_c - \Delta y)$ and incorporate the shift as an additional linear phase in the transfer function, expressed as

$$\hat{H}_z(f'_X, f'_Y) = H_z(f'_X, f'_Y) \exp\left[ik\lambda\left(f'_X \Delta x + f'_Y \Delta y\right)\right], \quad (5)$$

where $f'_X = f_X + f_{Xc}$, and $f'_Y = f_Y + f_{Yc}$. We derive the optimal values of Δx and Δy that guarantee the minimum sampling rate given any possible choice of the OW, and prove that our approach greatly reduces the sampling requirements in all scenarios (Supplement 1, Section 2A).

Similar to the spatial domain, the frequency-domain sampling rate must be analyzed collectively, taking into account the angular spectrum U of the input field and the shifted transfer function $\hat{H}_z$. While the angular spectrum is generally not expressible in a closed form, it can be approximately analyzed via the virtual lens and an upper-bound approach, which includes amplitude sampling as it cannot be assumed to be slowly varying. We also observe that $\hat{H}_z$ can be regarded as a converging spherical wave with a negative quadratic phase. Depending on the characteristics of the angular spectrum, the quadratic phase may counteract or accumulate to the corresponding one of U when considered together, resulting in fewer or more samples than those of sampling U and $\hat{H}_z$ separately, as illustrated in Fig. 3. The maximum gradient of $\phi_U + \phi_{\hat{H}_z}$, which is at the boundary, is much smaller than that of ϕ_U and $\phi_{\hat{H}_z}$. In addition, a broadening factor is included to account for the frequency cut-off, similar to the effect induced by the finite-sized aperture in the spatial domain. The detailed derivations of combined sampling are in Supplement 1, Section 2B.

Determined by not only the combined sampling of U and $\hat{H}_z$ but also the OW size w, the sampling rate in the frequency domain with $f \in \{f_X, f_Y\}$ can be derived as

$$\tau^{(f)}_{\text{freq}} = s \max\left\{\min\left\{2D + \tau^{(f)}_{\gamma\hat{H}_z}, D + \tau^{(f)}_{\hat{H}_z}\right\} + \frac{41.2}{\tau^{(\rho)}_{\text{spat}}}, 2|q'_c| + w\right\}, \quad (6)$$

where $\tau^{(f)}_{\gamma\hat{H}_z} = \frac{1}{\pi}\left|\frac{\partial}{\partial f}\phi_{\gamma\hat{H}_z}\right|_{\max}$ and $\tau^{(f)}_{\hat{H}_z} = \frac{1}{\pi}\left|\frac{\partial}{\partial f}\phi_{\hat{H}_z}\right|_{\max}$ denote the minimum sampling rates decided by phases of $U\hat{H}_z$ and $\hat{H}_z$, respectively, $\gamma = e^{i\phi_U}$, and $q'_c \in \{x_c - \Delta x, y_c - \Delta y\}$ is the OW center after virtual shift. We note that the first term in the max operator can also be used to forecast the OW size, if not specified beforehand, to cover most energy of the diffractive field. Thereby, the LS requirements of ASM are adaptively determined by Eqs. (4) and (6).

Adaptive sampling involves arbitrary samplings across the two domains, which cannot be handled by conventional FFT. To improve computational flexibility, algorithms such as chirp-Z transform (CZT) [17], Non-uniform FFT (NUFFT) [16], or matrix triple product (MTP) [18–20] can be used. However, CZT requires three conventional FFTs, and NUFFT often suffers from numerical precision issues. As such, we adopt MTP due to its numerical stability and intrinsic linear algebraic operation, which can be highly optimized with modern parallel computing techniques. To provide a detailed explanation of LS-ASM's procedure, we summarize the overall computational flow in Supplement 1, Section 3.

We validate the proposed LS-ASM by comparing it with the state-of-the-art shift-BEASM [16] to compute the off-axis complex point spread function (PSF) for a thin lens camera at 500 nm wavelength with a focal length of 35 mm and f-number of 16. The PSF is formed by focusing an off-axis spherical wave that originates from a distance of $z_0 = 1.7$ m away from the aperture. The OW size is set to $D/8$ with 512×512 pixels. The oversampling factors in both domains are set as 1.5. All numerical simulations are run on a workstation with an AMD Ryzen Threadripper Pro 3955WX CPU, using `Python 3.7`, `PyTorch 1.10` for the matrix product in LS-ASM, and the `finufft` library [21] for NUFFT implementation in shift-BEASM.

Figure 4 presents visual examples at a 3° incident angle and quantitative results from 0° to 20°. LS-ASM follows the sampling strategy mentioned before, while shift-BEASM is implemented in two different cases to highlight the superiority of our approach. The results from the first Rayleigh–Sommerfeld (RS) integral are used as the ground truth [9]. We assess the accuracy of output fields using the signal-to-noise ratio (SNR) [16] and root-mean-square error (RMSE). For visualization purposes, a common linear phase term $k(x \sin\theta_x + y \sin\theta_y)$ is removed from the results, and the remainder is denoted as $\hat{\phi}_{u_z}(x, y)$.

In case 1, where the samples in shift-BEASM are set to achieve the same level of SNR as LS-ASM, LS-ASM consistently performs at a remarkably short runtime even at large angles (e.g., 20°), outperforming shift-BEASM by over 35×. This is because the effective angular spectrum of shift-BEASM is decentered, requiring more samples to avoid aliasing, whereas LS-ASM always places the effective angular spectrum at the origin in the frequency domain. In addition, LS-ASM optimally specifies the sampling rate in the frequency domain, whereas shift-BEASM does not. It is worth noting that RS is in fact biased towards shift-BEASM as the unshifted transfer function is equivalent to the Fourier transform of the first RS impulse response [9]. Nevertheless, LS-ASM still exhibits 7× fewer spatial samples and 3× fewer frequency samples at 20°, resulting in less memory usage and enables larger-batch parallel computing, which further accelerates the computation.

In case 2, where the samples in shift-BEASM are set as the same rate as LS-ASM, the effective angular spectrum of the input field is partially cropped due to insufficient spatial samples, resulting in an incorrect result even at 5°. In contrast, LS-ASM enables efficient computations at much larger angles.

We further verify our method by showcasing the results at an ultra-large angle of 35°, the computation of which requires only 1.2 s. LS-ASM uses 17× and 3× fewer samples in each domain than shift-BEASM. Another example of a complex input field is

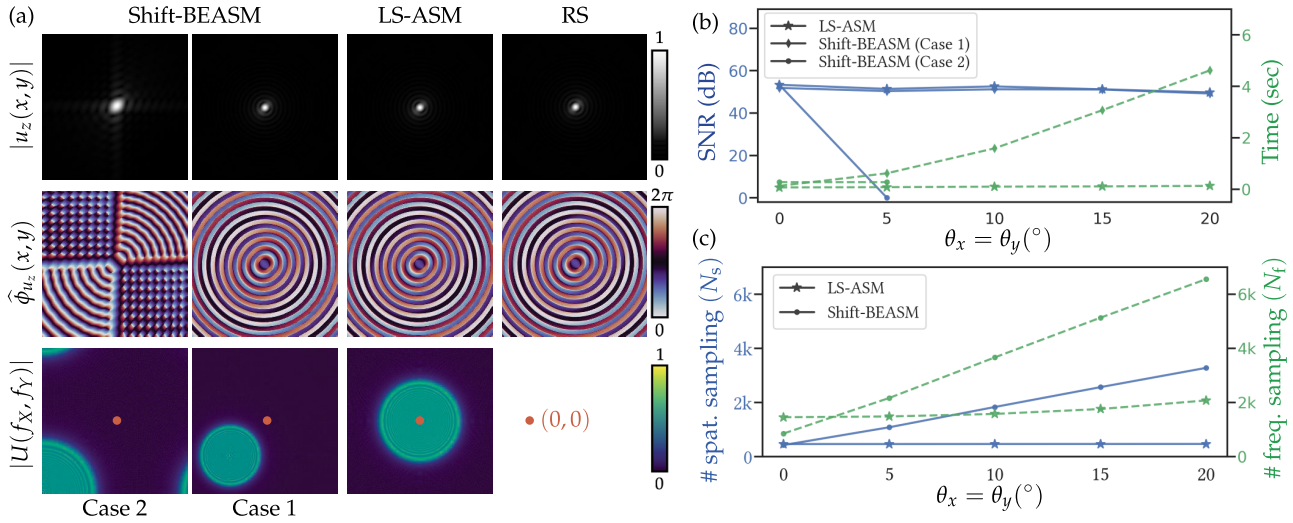

Fig. 4. (a) Visual results of the amplitude (first row) and phase (second row) of the complex PSF $u_z(x, y)$, and the amplitude for the angular spectrum $U(f_X, f_Y)$ (third row) of the input field at 3°. The amplitude of the PSF and the angular spectrum is normalized by respective maxima. In the third row, the red dots denote the center of the angular spectrum. (b) SNR and runtime w.r.t. incident angles. (c) Sampling number w.r.t. incident angles. More results with different incident angles and metrics can be found in Supplement 1, Section 4.

presented with a cubic phase plate placed against the thin lens at 10°. The cubic phase is a special case that possesses an intrinsic frequency offset but without any linear phase terms. Our method can accurately handle such cases. In addition, a comparison with Zemax and an ablation study on the oversampling factor s is also presented, from which we observe a balance between accuracy and runtime when $s = 1.5$ (Supplement 1, Section 4).

While our scheme has demonstrated flexibility and efficiency with assumptions of slowly varying amplitude and smooth phase, it applies also to non-smooth, i.e., suddenly varying, phase and amplitude. This can be achieved by applying Eq. (4) to local sub-apertures where the field is smooth. The largest incident angle that our method can handle is limited by the validity of the scalar diffraction theory and available computational resources. In addition, although a long propagation distance may increase the sampling burden, it can be alleviated by the semi-analytical Fourier transform method, which enables analytical computation of quadratic phase [22]. Relevant analysis and demonstrations are provided in Supplement 1, Section 5.

In conclusion, we have explored a flexible and efficient numerical wave propagation modeling scheme that is tailored to represent off-axis diffractive propagation. Compared to prior works, our approach provides the LS strategy for optical fields and is applicable to all kinds. We envision a wide range of applications where efficient and accurate modeling of wave propagation is essential, particularly where a large off-axis incident angle is expected.

Funding. Hong Kong University Grants Committee (ECS 27212822, GRF 17201822, GRF 17208023); National Natural Science Foundation of China.

Acknowledgment. We thank Wolfgang Heidrich for discussion.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper may be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

REFERENCES

1. Y. Peng, S. Choi, J. Kim, and G. Wetzstein, Sci. Adv. **7**, eabg5040 (2021).
2. L. Yang, I. Badar, C. Hellmann, and F. Wyrowski, Opt. Express **29**, 3621 (2021).
3. S. Banerji, M. Meem, A. Majumder, B. Sensale-Rodriguez, and R. Menon, Optica **7**, 214 (2020).
4. X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan, Science **361**, 1004 (2018).
5. V. Sitzmann, S. Diamond, Y. Peng, X. Dun, S. Boyd, W. Heidrich, F. Heide, and G. Wetzstein, ACM Trans. Graph. **37**, 1 (2018).
6. Y. Peng, Q. Sun, X. Dun, G. Wetzstein, W. Heidrich, and F. Heide, ACM Trans. Graph. **38**, 219 (2019).
7. V. Boominathan, J. T. Robinson, L. Waller, and A. Veeraraghavan, Optica **9**, 1 (2022).
8. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th ed. (Cambridge University, 1999).
9. J. W. Goodman, *Introduction to Fourier Optics*, 4th ed. (W.H. Freeman and Company, 2017).
10. W. Zhang, H. Zhang, C. J. R. Sheppard, and G. Jin, J. Opt. Soc. Am. A **37**, 1748 (2020).
11. X. Liu, L. Li, X. Liu, X. Hao, and Y. Peng, Opt. Express **30**, 36973 (2022).
12. J. M. Geary, *Introduction to Lens Design: With Practical ZEMAX Examples* (Willmann-Bell Richmond, 2002).
13. J.-P. Liu, J. Opt. Soc. Am. A **29**, 1956 (2012).
14. K. Matsushima, Opt. Express **18**, 18453 (2010).
15. C.-S. Guo, Y.-Y. Xie, and B. Sha, Opt. Lett. **39**, 2338 (2014).
16. W. Zhang, H. Zhang, K. Matsushima, and G. Jin, Opt. Express **29**, 10089 (2021).
17. Y. Hu, Z. Wang, X. Wang, S. Ji, C. Zhang, J. Li, W. Zhu, D. Wu, and J. Chu, Light Sci. Appl. **9**, 119 (2020).
18. A. S. Jurling, M. D. Bergkoetter, and J. R. Fienup, J. Opt. Soc. Am. A **35**, 1784 (2018).
19. W. Zhao, C. Wei, C. Yuan, C. Chang, J. Ma, and R. Zhu, Opt. Lett. **45**, 5937 (2020).
20. X. Liu, Y. Hu, S. Tu, C. Kuang, X. Liu, and X. Hao, Opt. Lasers Eng. **162**, 107405 (2023).
21. A. H. Barnett, J. Magland, and L. af Klinteberg, SIAM J. Sci. Comput. **41**, C479 (2019).
22. Z. Wang, S. Zhang, O. Baladron-Zorita, C. Hellmann, and F. Wyrowski, Opt. Express **27**, 15335 (2019).