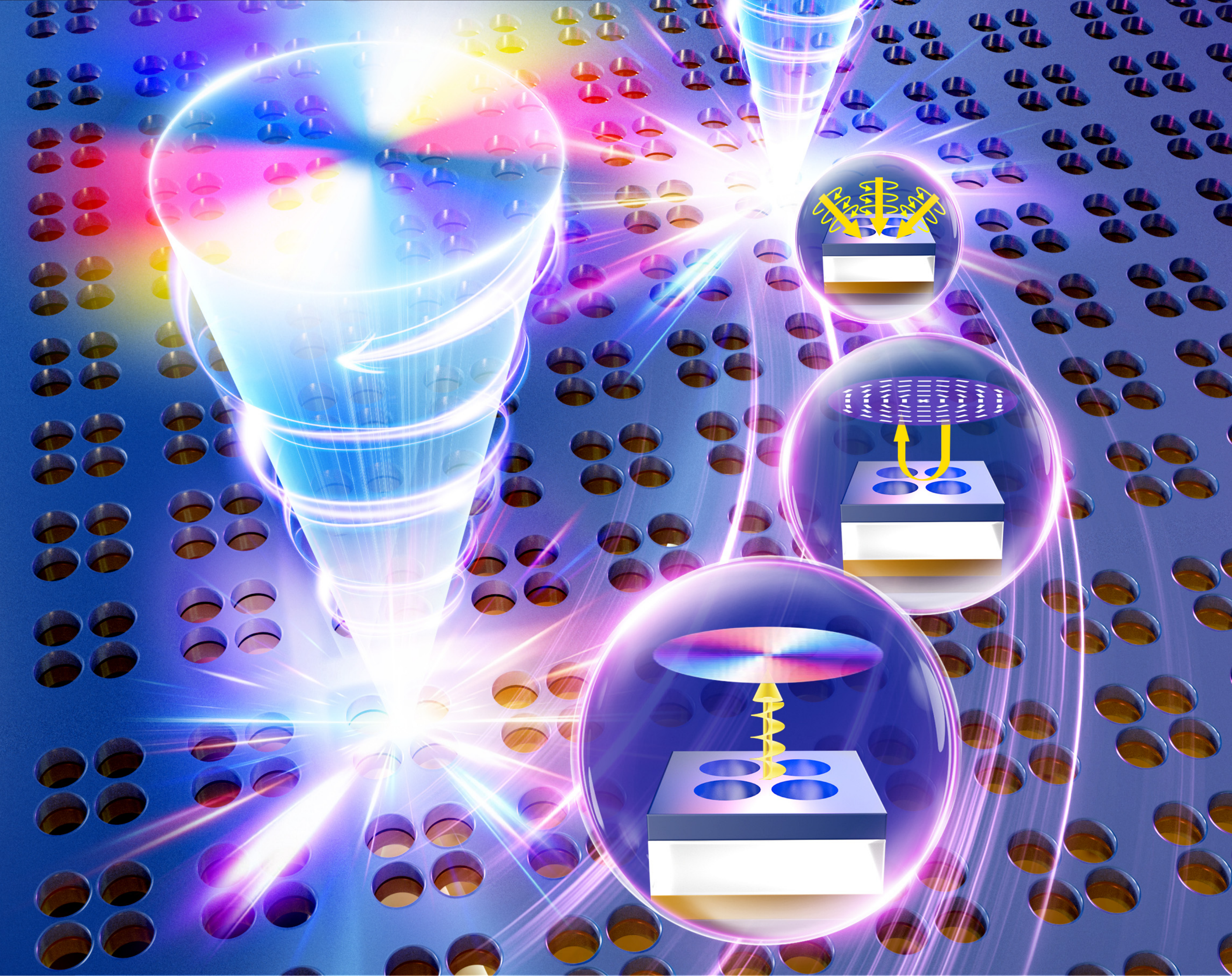
OEA | Opto-Electronic Advances

Cavity-assisted nonlocal metasurfaces for momentum-space broadband-operational optical vortice generation with maximum efficiency approaching 80%

Keren Wang†, Kaili Sun†, Jing Du, Peijuan Dai, Hao Zhou, Lujun Huang*, Zhanghua Han* and Wei Wang*

https://doi.org/10.29026/oea.2026.250296



OEA | Opto-Electronic Advances

CN 51-1781/TN ISSN (Print) 2096-4579 ISSN (Online) 2097-3993

Cavity-assisted nonlocal metasurfaces for momentum-space broadband-operational optical vortice generation with maximum efficiency approaching 80%

Keren Wang, Kaili Sun, Jing Du, Peijuan Dai, Hao Zhou, Lujun Huang*, Zhanghua Han* and Wei Wang*

Citation: Wang KR, Sun KL, Du J et al. Cavity-assisted nonlocal metasurfaces for momentum-space broadband-operational optical vortice generation with maximum efficiency approaching 80%. *Opto-Electron Adv* **9**, 250296 (2026).

https://doi.org/10.29026/oea.2026.250296

Received: 4 November 2025; Accepted: 23 January 2026; Published online: 2 March 2026

OEA | Opto-Electronic Advances

Article

July 2026, Vol. 9, No. 7

Open Access

Cavity-assisted nonlocal metasurfaces for momentum-space broadband-operational optical vortice generation with maximum efficiency approaching 80%

Keren Wang[1†], Kaili Sun[2†], Jing Du[1], Peijuan Dai[1], Hao Zhou[3], Lujun Huang[4*], Zhanghua Han[2*] and Wei Wang[1*]

Abstract: Nonlocal metasurfaces generate optical vortices via momentum-space topological singularities, eliminating the need for complex meta-atom phasing or precise beam alignment. Their utility has been constrained by narrow bandwidth and low efficiency stemming from the steep dispersion of high-*Q* resonances. Here, we introduce a cavity-assisted reflective nonlocal metasurface that overcomes these limits by hybridizing bound states in the continuum (BICs) with degeneracy points (DPs). A Fabry-Pérot cavity supplies a single, addressable control for BIC-DP coupling, enabling deterministic tuning of dispersion, radiative *Q*, and polarization to realize quasi-flat bands with strong scattering. Simulations predict near-unity on-resonance conversion and >90% overall efficiency—3–4× higher and >15× broader than conventional designs. Experiments confirm operation from 1480 to 1600 nm with ~80% peak efficiency and 91.7% OAM purity, while suppressing edge effects and markedly reducing sensitivity to alignment, beam profile, and numerical aperture. Crucially, the device enables broadband, efficient conversion of zero-order Bessel beams into high-quality, OAM-carrying perfect vortex beams—performance not achieved by prior metasurfaces. These results establish a practical, scalable route to broadband, high-efficiency vortex generation for high-dimensional optical communications, advanced imaging, and quantum photonics.

Keywords: nonlocal metasurfaces; orbital angular momentum; bound states in the continuum; band engineering; broadband operation; vortex beam generation

DOI: 10.29026/oea.2026.250296 | CSTR: 32247.14.oea.2026.250296

Citation: Wang KR, Sun KL, Du J et al. Cavity-assisted nonlocal metasurfaces for momentum-space broadband-operational optical vortice generation with maximum efficiency approaching 80%. *Opto-Electron Adv* **9**, 250296 (2026).

1 Introduction

Vortex beams, defined by helical phase fronts and orbital angular momentum (OAM), are essential in many cutting-edge applications, including optical communications[1], quantum information processing[2,3], spin-orbit interactions of light[4], and super-resolution imaging[5,6]. They allow researchers and scientists to encode information in OAM[7], greatly expanding the capacity for high-dimensional data transmission and quantum entanglement[8].

Conventional approaches of generating vortex beam leverage phase modulation in real space[8–15], including dynamic phase and geometric phase[11]. Typical examples include spiral phase plates[12] and Q-plates[11,16]. For the former case, it achieves spiral phase accumulation by varying the thickness or refractive index of the materials along the radial direction, and usually operates in the low-frequency range. For Q-plates operating in the visible wavelength range, a cross-polarized circularly polarized transmitted light gains the geometric Pancharatnam-Berry (PB) phase with a gradient along the azimuthal direction. Although straightforward, these designs are bulky and highly sensitive to wavelength, limiting their versatility in many applications. Recently, a more compact platform, the so-called metasurfaces, has

Received: 4 November 2025

Accepted: 23 January 2026

Published online: 2 March 2026

[1]College of Physics, Sichuan University, Chengdu 610064, China; [2]Shandong Provincial Key Laboratory of Optics and Photonic Devices, Center of Light Manipulation and Applications, School of Physics and Electronics, Shandong Normal University, Jinan 250358, China; [3]College of Electronics and Information Engineering, Sichuan University, Chengdu 610064, China; [4]State Key Laboratory of Precision Spectroscopy, School of Physics, East China Normal University, Shanghai 200241, China.

†These authors contributed equally to this work.

*Correspondence: LJ Huang, ljhuang@phy.ecnu.edu.cn; ZH Han, zhan@sdnu.edu.cn; W Wang, w.wang@scu.edu.cn

been developed to generate vortex beams in the visible and near-infrared wavelength range[15,17,18]. Metasurfaces, which rely on dynamic[10,13] or geometric phases[14,19], provide greater flexibility and wavelength independence[20–22], but are costly to fabricate and demand strict optical alignment to geometric centers and suffer from efficiency bottlenecks imposed by fabrication imperfections, especially on a short-wavelength scale. In addition, metasurfaces made of discrete meta-atoms can only modulate phase discontinuously[10,13,20], limiting the quality of vortex beams.

More recently, momentum space polarization vortices centered at the bound states in the continuum (BICs) have been harnessed to generate vortex beams in a nonlocal metasurface[23]. This approach encodes vortex PB phase into cross-polarized light, opening a new avenue for light manipulation[24–27]. Except for BICs, degeneracy points (DPs) carrying polarization vortex have also been used to generate vortex beam in momentum space[28,29]. Unlike phase modulation in real space[10,30], momentum-space modulation eliminates the need for beam alignment with a geometric center and offers smoother phase modulation. However, the pursuit of a vortex platform concurrently delivering broadband operation, near-unity efficiency, and diffraction-limited purity —without positional sensitivity —has persisted as a critical challenge in photonics[31]. Specifically, the on-resonance conversion efficiency—defined at a specific wavevector—is limited by high-quality factor (*Q*-factor) resonances near the BICs. More importantly, the overall efficiency remains low, as the steep dispersion in the photonic crystal slabs (PCS) restricts the working bandwidth of the wavevector within the iso-frequency contour in momentum space[32]. These restrictions are common in both BIC- and DP-based approaches, resulting in efficiency being highly dependent on the working wavelengths and wavevector, i.e., numerical apertures (NAs), which limits the tunability and flexibility of versatile optical platforms. To illustrate the limitations of existing designs, Table 1 summarizes representative nonlocal metasurface implementations, highlighting their narrow operational bandwidths, low overall efficiencies, and stringent NA constraints. This comparison motivates the need for a new platform capable of delivering high-efficiency, broadband operation without sacrificing beam quality or positional robustness.

In this study, we demonstrate broadband operational vortex beam generation with high efficiency through leveraging BICs and DPs synergistically in a nonlocal metasurface. We show that the dispersion, *Q*-factor, and topological states of polarization (SOPs) properties together in momentum space can be tailored through engineering the coupling between BICs and DPs. This approach effectively decouples efficiency limitations from the working wavelength and NAs seen in previous methods, enabling simultaneously high efficiency and wide tunability on wavelength. Numerical simulations show that the on-resonance efficiency and overall efficiency can reach as high as 96% and 94%, respectively, with OAM purity as high as 99.5%. Besides, such a metasurface can generate vortex beam with the operation wavelength tunable from 1500 nm to 1600 nm. Following the theoretical design, we fabricated the relevant metasurfaces and successfully demonstrated the generation of vortex beam with overall efficiencies approaching 80%, highlighting a significant improvement over previous studies. Furthermore, our measurements confirm that the overall efficiency remains reasonably high across the 1480–1600 nm under various NAs while simultaneously showing an exceptional robustness to edge effect in periodic structures. This enables wavelength-tunable vortex generation with exceptional flexibility. Finally, by directly illuminating the structure with a zero-order Bessel beam, we experimentally generated efficient second-order Bessel beams (also known as perfect vortex beam[33,34]), showcasing the versatility and flexibility of our design across diverse optical platforms.

2 Results

2.1 Principle

Vortex beams can be generated in momentum space via nonlocal metasurfaces by exploiting the polarization vortices carried by guided resonances (GRs) in photonic crystal slabs (PCS). Such momentum-space modulation provides smooth, symmetry-protected phase profiles[13] and high tolerance to fabrication imperfections[14]. A detailed derivation of the origin of the vortex phase generation in momentum space, induced by the polarization vortex around the BIC/DP, is provided in Supplementary Section 1.

However, existing schemes based on single BICs or DPs suffer from intrinsic efficiency bottlenecks, as illustrated in Fig. 1(a–c). To quantitatively analyze this, we note that the vortex conversion efficiency is governed by the cross-polarization anisotropy $\Delta s_{\text{res}} = s_{\text{res-}x} - s_{\text{res-}y}$, scaling as $\eta \propto |\Delta s_{\text{res}}|^2$.

Table 1 | Comparison of our design with previous non-local metasurface.

Ref.	Design	Wavelength band	Sim. on-res eff.	Exp. overall eff.	Operational bandwidth	Operational NA
ref.[23]	PCS	Visible	-	-	Narrow band	Limited
ref.[35]	Metal-backed PCS	Visible	86%	-	Narrow band	Limited
ref.[28]	Metal-backed PCS	Visible	<80%	13%	Narrow band	*NA*=0.07 (limited)
ref.[32]	Metallic PCS	mm-wave	41%	19.6%	Narrow band	*NA*=0.4 (limited)
This work	Metal-backed BZF PCS + FP cavity	Near-Infrared	94%	77%	Broadband	*NA*=0.06–0.42 (flexible)

The strength of the individual resonance terms s_{res} is governed by temporal coupled mode theory (TCMT)[36,37]:

$$s_{res} = -\frac{d_s^2 + d_p^2}{\gamma_{rad} + \gamma_{abs} - i(\omega - \omega_0)}, \quad (1)$$

where γ_{rad} and γ_{abs} represent radiative and absorption leakage rates. In a reflection-type PCS, energy conservation dictates $|d_s|^2 + |d_p|^2 = 2\gamma_{rad}$[31,35], so on resonance ($\omega = \omega_0$) the cross-conversion efficiency simplifies to $\eta = \left|\frac{1}{2}s_{res}\right|^2 \sim (Q_{abs}/(Q_{rad}+Q_{abs}))^2$, where Q_{rad} and Q_{abs} are the radiative and absorptive Q-factors, defined by $Q = \omega/(2\gamma)$ for a complex eigenfrequency $\tilde{\omega} = \omega - i\gamma$. This formulation reveals two distinct mechanisms for the failure of conventional designs near the Γ point (Fig. 1(b)): (i) For BIC-based designs, the radiative Q-factor diverges ($Q_{rad} \to \infty$) at the singularity. According to Eq. (1), the resonance amplitude itself vanishes ($s_{res} \to 0$), inherently driving $\Delta s_{res} \to 0$. (ii) For DP-based designs, although Q_{rad} is finite (yielding strong s_{res}), the symmetry-enforced degeneracy results in orthogonal polarization branches with equal amplitude (SOP cancellation). This implies $s_{res\text{-}x} \approx s_{res\text{-}y}$, leading to $\Delta s_{res} \to 0$ despite the strong resonance. A rigorous TCMT- and scattering-matrix-based analysis of this dual-mode system is provided in Supplementary Section 1.

Furthermore, aside from the central void, the overall efficiency is severely limited by steep dispersion, which restricts the momentum-match condition to a narrow iso-frequency contour (Fig. 1(c)). Most incident energy falls outside this narrow ring, rendering the device inefficient for broadband or high-NA inputs.

To overcome these barriers, we introduce a deterministic band-engineering strategy that simultaneously tailors dispersion, radiative Q-factor, and polarization topology. As shown in Fig. 1(d–f), by hybridizing BICs with multiple DPs, we engineer GRs with quasi-flat dispersion and exceptionally

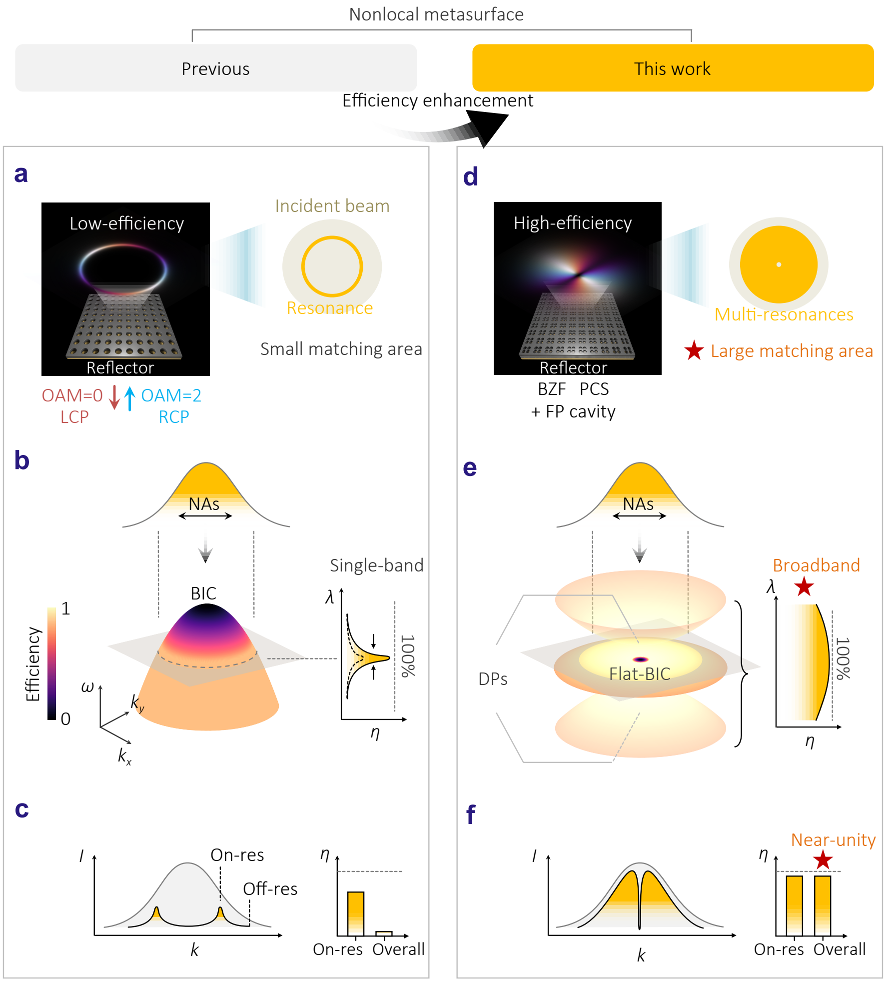


Fig. 1 | Overview of our theoretical principle. (a–c) For conventional nonlocal metasurfaces based on a single BIC approach, the presence of a narrow efficiency ring in (k_x, k_y) space (a) and steep modal dispersion (b) severely limits overall vortex beam conversion efficiency (c). (d–f) In contrast, our reflection-type BZF nonlocal metasurface (d) employs tailored band engineering to construct quasi-flat dispersion harnessing both BIC and DP modes (e), leading to broadband high-efficiency operation (f).

strong scattering (low Q-factor). This ensures the momentum matching condition is satisfied over a broad range, leading to near-unity overall efficiency and robust operation. Specifically, we strategically introduce mode coupling to establish a link between BICs and multiple DPs, as illustrated in Fig. 2.

2.2 Numerical implementation

Our implementation, illustrated in Fig. 2(a), employs a reflection-type PCS vortex generator: an air-hole PCS in silicon atop a SiO_2 spacer backed by a gold mirror. The high-index PCS confines the GR modes, while the SiO_2 spacer decouples the fields from the Au substrate, significantly reducing Ohmic losses. Furthermore, this configuration forms a FP (Fabry-Pérot) cavity between the top PCS and the gold substrate (Fig. 2(b)), where the spacer thickness t_0 allows independent tuning of the FP resonance wavelength while suppressing Au absorption.

Periodic perturbations are introduced into the PCS lattice, preserving C_4 symmetry (Fig. 2(a), middle panel), to enable Brillouin zone folding (BZF)[38–41], generating GR modes with controllable wavelength and Q factor (Fig. 2(a), bottom panel). In this hybrid system, the BIC arises from a high-Q GR band of the PCS, while DPs are contributed by FP cavity modes with two low-Q bands. All modes carry a topological charge of +1 but exhibit different SOPs (AP vs. RP).

To systematically achieve efficient vortex generation with multidimensional tunability, we start with a simplified system including DP_1 and BIC singularities. The optical bands that relate with DP_1, BIC, DP_2 are labeled from bottom to top as FP_1^A, FP_1^R, GR^A, FP_2^R, and FP_2^A, where the superscripts A/R denote the azimuthal and radial SOPs, respectively (see left panel in Fig. 2(b)). For clarity, we can distinguish these modes based on their degeneracy at the Γ point: the GR band hosts a BIC and thus is non-degenerate at Γ, whereas the FP bands host DPs and exhibit twofold degeneracy at Γ.

Initially, as illustrated in Fig. 2(b), neither BIC nor DP efficiently generates vortex beams —the former is due to excessive Q-factor restricting bandwidth, and the latter is due to SOPs cancellation between FP_1^A and FP_1^R bands (left panel). Furthermore, the steep dispersion of the initial structure permits only a narrow tolerance for wavelength and wavevector.

Subsequently, we tune the wavelength of the first DP (DP_1) to facilitate the interaction between FP_1^A and GR^A (avoided crossing in Fig. 2(c)). This tuning establishes strong coupling between GR^A and FP_1^A[41], resulting in an avoided crossing that flattens the dispersion of GR^A around the BIC. Simultaneously, the off-Γ Q-factor is reduced to FP levels, dramatically enhancing the conversion efficiency while maintaining the strongly polarization-dependent response. Notably, the low Q-factor signifies a short

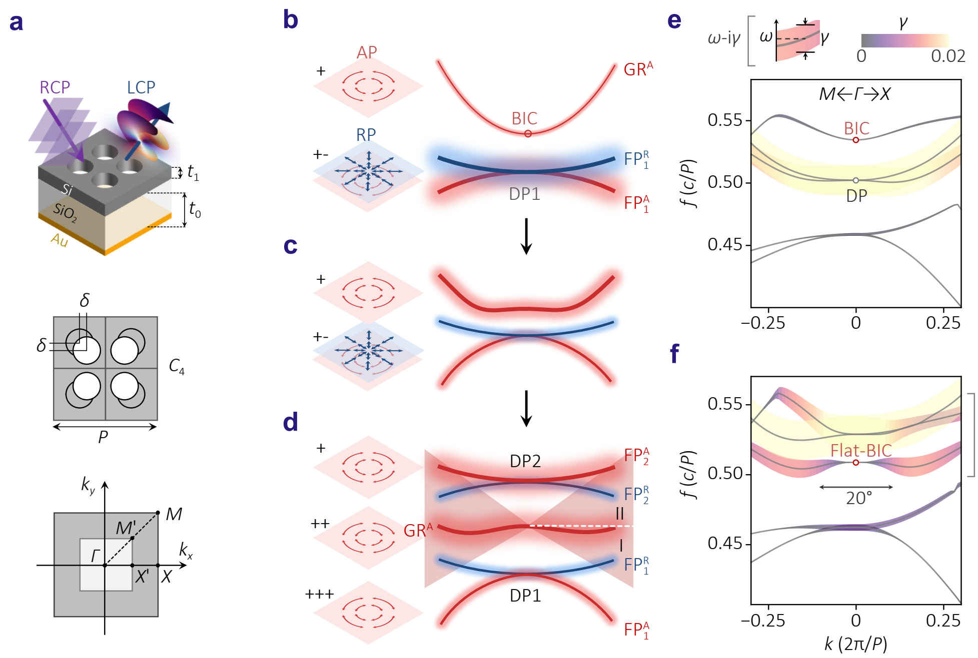


Fig. 2 | Overview of our implementation. (a) Top panel: BZF PCS structure converting an incident right-handed circularly polarized plane wave into a left-handed circularly polarized vortex beam. Middle panel: Periodic perturbation maintaining C_4 symmetry. Bottom panel Brillouin zone folding in momentum space. (b–d) Band-structure engineering diagrams; glowing regions represent mode bandwidth (Q factor), red/blue colors denote AP/RP SOPs respectively. Broadband high-efficiency regimes I and II are formed in (d) through synergistic effects of BIC- and DP-related modes. (e, f) Simulated band structures for conventional and optimal designs, with line thickness and color indicating mode bandwidth.

resonance lifetime, which enhances the system's robustness against edge effects in periodic structures, as the resonances cannot travel far before radiating. This behavior is later confirmed by our experimental results. The engineered flat dispersion (Fig. 2(c)) broadens the NA tuning range and establishes a flat, triangular high-efficiency region (indicated by the red background) around the BIC frequency. However, this single mode approach can only generate vortex beams within a narrowband wavelength range.

To further expand the efficiency regime, we introduce a secondary DP (DP_2), as shown in Fig. 2(d). First, we adjust the wavelength of the GR^A-FP_1^A mode to mitigate the disruptive effects of FP_1^R in the lower-frequency Region I. Then, we promote an interaction between GR^A and FP_2^A. This interaction suppresses the Q-factor of FP_2^A and benefits radiation in the azimuthal SOP over the radial SOP (FP_2^R). Consequently, a higher-frequency Region II is created that supports efficient cross-polarized conversion. The avoided crossing between GR^A and FP_2^A further flattens the BIC band at larger wavevectors. As shown in Fig. 2(d), such a bidirectional band engineering, driven by the synergistic contributions of GR^A, FP_2^A and FP_1^A, dominates the radiation in azimuthal SOPs (left panel) and establishes a special hourglass-shaped high-efficiency regime I/II, and provides unprecedented wavelength/NA tunability for vortex beam generation.

Optimized parameters through full-wave simulations yield: Si thickness $t_1 = 157$ nm, SiO_2 thickness $t_0 = 545$ nm, hole radius $r = 133$ nm, hole offset $\delta = 52$ nm, and lattice constant $P = 2a = 1570$ nm. The system behavior under various geometric parameters is discussed in the Supplementary information of Section 7. Note that material absorption plays a critical role in conversion efficiency, so we assign an imaginary refractive -index component of 10^{-4}—a realistic value for chemical vapor deposition (CVD) -deposited films. Specifically, we set $n_{Si} = 3.58 + 10^{-4}i$ and $n_{SiO_2} = 1.45 + 10^{-4}i$. For comparison, a reference design—representative of the typical efficiency limitations reported in previous studies—is obtained by setting $\delta' = 0.7\delta$. The band structures of both the optimal and alternative designs are shown in Fig. 2(e, f), respectively, successfully validating our proposed band-engineering concept. Detailed Q-factor distributions are provided in Section 10 of the Supplementary information, while a more thorough performance comparison can be found in Section 3 of the Supplementary Materials.

The strong coupling between BIC and DP modes is evidenced by an avoided crossing in the parameter space (Supplementary Section 7), resulting in a system that exhibits dominant AP contributions from GR^A and FP_2^A. This enables broadband vortex generation, with the BIC-based GR^A governing efficient conversion below 1550 nm and the DP-based FP_2^A dominating above 1550 nm, as illustrated by the hourglass-shaped high-efficiency region in Fig. 3(d). Detailed SOP characteristics for each band are provided in the Supplementary information of Section 6. Building on the above mode -coupling analysis, we now examine the performance contrast between the optimized and reference structures.

For the reference design (Fig. 3(a–c)), a BIC appears near 1480 nm and a DP near 1500 nm, as shown in Fig. 2(e). The efficiency η shows only limited peaks around the BIC, consistent with the resonance linewidths of R_s. At 1500 nm, DP conversion is suppressed due to SOP cancellation. As shown in Fig. 3(b), a narrow efficiency ring emerges in momentum space, in agreement with dispersion-limited matching calculated in Supplementary Section 2.

In contrast, the optimized design (Fig. 3(d–f)) achieves significantly enhanced broadband efficiency. The +1 topological-charge BIC, relocated near 1550 nm, enables maximal radiative resonance utilization across 1550–1600 nm, reaching up to ~96% on-resonance efficiency at (1549 nm, 3°). Meanwhile, the DP at 1500 nm enhances conversion in the 1500–1550 nm range. Together, these coupled modes form an hourglass-shaped high-efficiency region from 1500 to 1600 nm (Fig. 3(d)), outperforming the conventional structure in Fig. 3(a). Importantly, the preserved C_4 symmetry ensures azimuthally uniform conversion profiles near center, while the widely flattened dispersion minimizes the influence of central efficiency voids (seen in iso-contours in colored panels). Compare Fig. 3(b, e), since the narrow conversion ring has been greatly expanded, the cross-shaped conversion efficiency near the edges can be effectively eliminated in practical applications by limiting the beam's focus angle (i.e., NA), to achieve optimal azimuthal uniformity and OAM purity without sacrificing efficiency. Figure 3(c, f) show the simulated overall efficiency under flat-top beam incidence, obtained by averaging the maps in Fig. 3(b, e). Compared to conventional nonlocal metasurfaces (Fig. 3(c)), our design achieves more than threefold higher efficiency with flat-top beams (Fig. 3(f)), fourfold higher efficiency with Gaussian beams, over fifteenfold broader bandwidth with Bessel beams, and far greater tolerance to varying NA. The overall efficiency reaches 94% at 1549 nm within a NA=0.21, with near-unity (up to 99.5%) OAM purity and smooth phase manipulation. Note that while the "flexible operational NA" mentioned in Table 1 implies a significantly broadened acceptance angle compared to previous studies, it is physically bounded by the flat-band region. Specifically, our design maintains high efficiency within an NA range of approximately 0.06 to 0.42. Within this window, the quasi-flat dispersion ensures robust momentum matching; beyond $NA \approx 0.42$, the efficiency gradually declines as the flat-band condition deteriorates. Nevertheless, this range covers most standard objectives and accommodates diverse incident beam profiles (e.g., various Gaussian waists or Bessel cone angles) without requiring strict alignment, thereby offering practical flexibility. Full comparisons for different beam types are provided in Supplementary

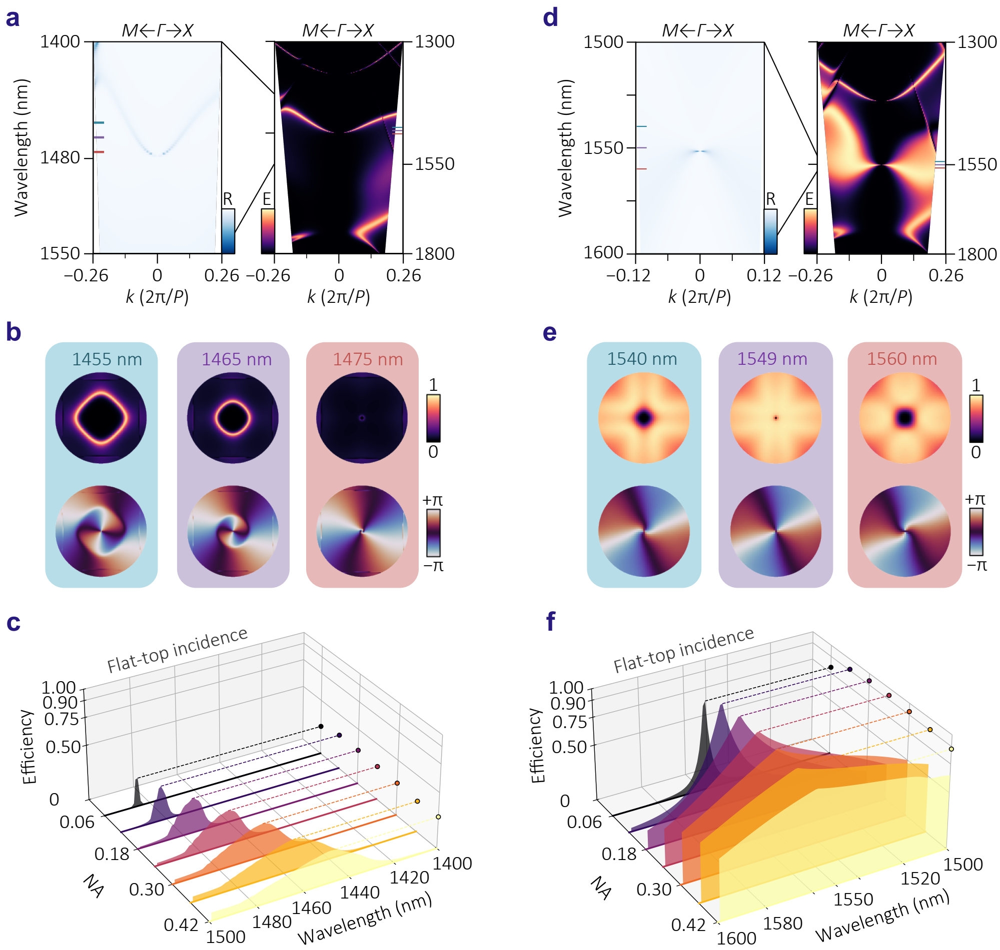


Fig. 3 | Numerical results of our design. (a) Detailed spectrum of conversion efficiency $\eta(\omega, \boldsymbol{k})$ for the comparison structure. Left: s-polarized reflection spectrum (R_s); right: cross-polarized conversion efficiency spectrum (η). (b) Colored panels: cross-sectional efficiency and phase at different wavelengths under NA=0.42. (c) Simulated overall efficiency of flat-top-beam illumination under different NA. (d–f) Same as (a–c) but for optimal structure.

Section 4, while OAM quality is evaluated via LG-mode decomposition in Section 9.

Overall, our design surpasses conventional nonlocal metasurfaces in both efficiency and tunability, offering robust momentum-space generation and strong resilience against edge effects, as further confirmed experimentally.

2.3 Experimental demonstration

To confirm the superior performance predicted by our simulations, we experimentally demonstrate the vortex beam generation using our optimized structural parameters. A Si/SiO_2 multilayer was grown through plasma-enhanced chemical vapor deposition (PECVD). PCS structures with four air holes in a super unit cell are fabricated with electron beam lithography followed by inductively coupled plasma reactive ion etching, as shown in Fig. 4(b). More details of the fabrication process can be found in Section *Methods*.

A home-built reflection-type Fourier-optics imaging system with dual operation modes (imaging and interferometry) was developed for optical characterization, as schematically illustrated in Fig. 4(a). The system is composed of three functional modules: (i) beam generation, (ii) vortex beam conversion, and (iii) converted beam collection/imaging. In imaging mode, left-handed circularly polarized light from a laser source (generated by a linear polarizer and quarter-wave plate (JCOPTIX, China)) was focused onto the PCS through an objective lens with NA=0.42, which is consistent with our simulation NA range. The cross-polarized reflected light underwent Fourier transformation through the same optical path before being captured by an InGaAs CMOS camera (JCOPTIX, China) in the beam collection/imaging module.

For phase characterization, a self-referenced interferometric configuration was implemented. The front-surface reflection from the linear polarizer P served as the reference beam. Although it is linearly polarized, it interferes effectively with the circularly polarized vortex beam (due to non-orthogonality), creating off-axis fringes suitable for Fourier-transform phase retrieval[42]. By adjusting the yaw angle of

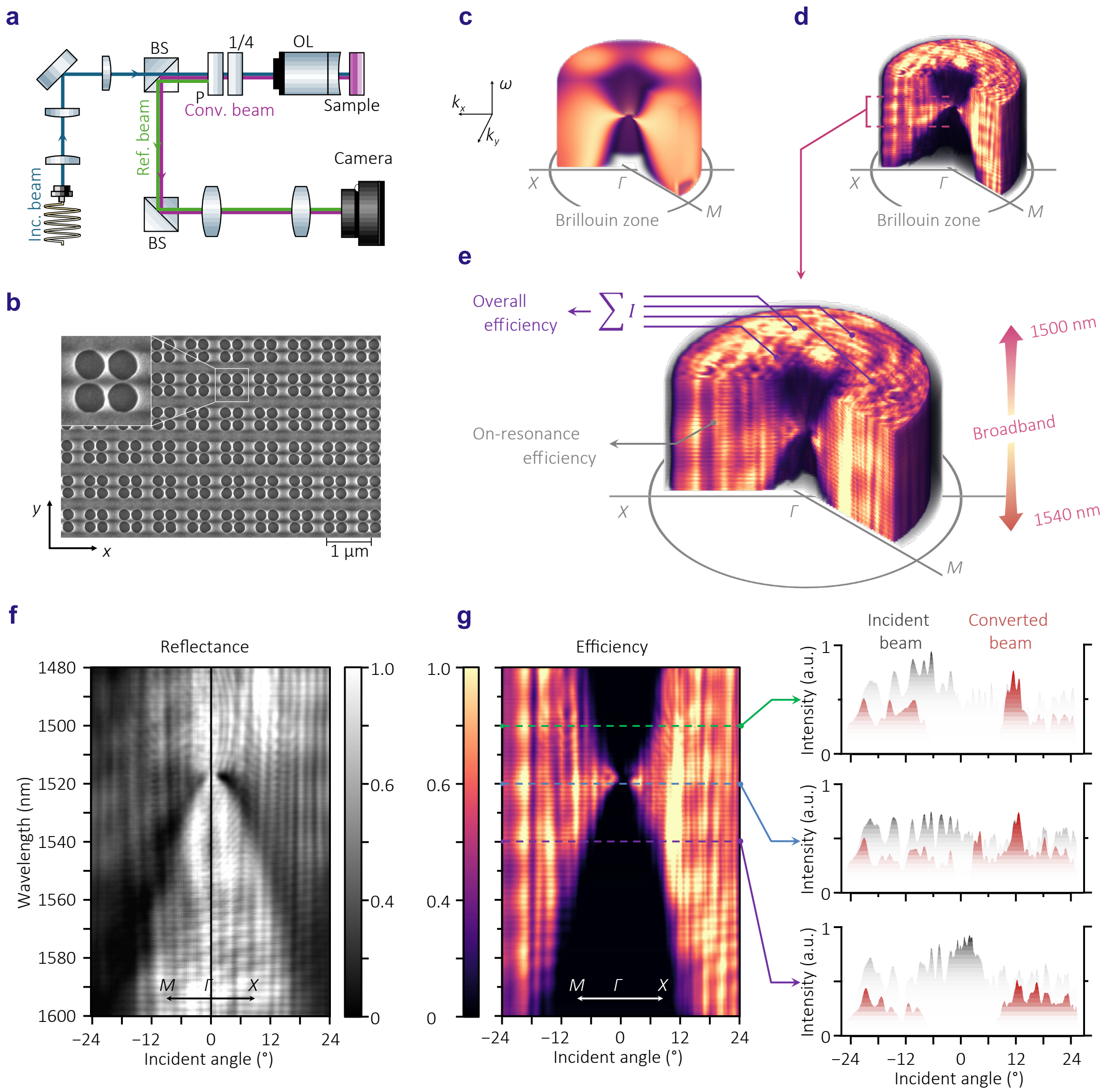


Fig. 4 | (a) Schematic of the momentum-space imaging optical setup employed in the experiment. (b) Top-view SEM micrograph of the fabricated metasurface. (c–e) Three-dimensional full ω-$\boldsymbol{k}$ dispersion of the cross-polarization conversion efficiency: (c) simulated results; (d, e) experimental measurements, with (e) highlighting broadband and high overall conversion efficiency across the 1500–1540 nm range. (f) Measured s-polarized reflection spectrum R_s. (g) Cross-polarized conversion efficiency spectrum. Colored dashed lines in (g) mark the wavelengths—1500 nm (green), 1520 nm (blue) and 1540 nm (purple)—for which the right subpanels show the absolute intensities of the incident (black) and converted (red) beams.

polarizer P, three interference modes were achieved: reference-free (for conversion efficiency measurement), tilted-reference (for oblique phase imaging), and collinear-reference (for coaxial phase retrieval). Wavelength-dependent spectral features were obtained by scanning the laser wavelength, revealing the full ω-$\boldsymbol{k}$ dispersion relation (Fig. 4(c–e), also see Supplementary Video 1 for details). The broadband high-efficiency, and the difference between overall efficiency and the on-resonance efficiency are demonstrated in Fig. 4(e). The s-polarized reflection spectra (obtained by rotating the quarter-wave plate) and cross-polarized conversion efficiency profiles are shown in Fig. 4(f, g), respectively, demonstrating a broadband "hourglass" region of high efficiency consistent with theoretical predictions. The BIC is located around 1520 nm, slightly deviating from the numerical simulations primarily due to discrepancies between the actual refractive index of the structure and the refractive index used in the simulations. This mismatch leads to a globally slight blueshift in the spectrum, but does not significantly affect the structural performance (see Supplementary information of Section 7 for details). The right subpanels of Fig. 4(g) show the absolute intensities of the incident non-vortex beam (black) and the converted vortex beam (red) around 1500, 1520, and 1540 nm. In particular, the mid-subpanel of 1520 nm demonstrates remarkable robustness of the converted beam intensity against variations in incident angle, which underpins the device's high overall efficiency. Note that, since we sample only along two one-dimensional cuts—along the *Γ*-*X* and *Γ*-*M* directions —through the imaged two-dimensional momentum-space intensity distribution, diffraction-speckle fluctuations are observed.

Next, we confirm the robustness of our design by verifying the decentralization of momentum-space vortex beam generation and examining its dependence on wavelength and NAs. The decentralization of vortex generation was first verified by probing regions from the structure's edge ($< 10P$) to its center (Fig. 5(a)). Considering the 3P-sized beam spot in NA=0.42, consistent vortex beam generation across all positions confirmed the position-independent operation and robustness against edge effects. As we have predicted, the disappearance of edge effects in the PCS is due to the low Q-factor of the emission mode we utilize, causing the resonance to radiate spontaneously before it can propagate to the edge.

The wavelength- and NA-dependent efficiency was then systematically investigated (Fig. 5(b)). By illuminating the sample with momentum-space flat-top beam, the overall efficiency was calculated by integrating the total beam intensity from the acquired images and comparing it to a reference beam reflected by an Au mirror. More details are presented in Supplementary information of Section 4. The efficiency shows an exceptional robustness against wavelengths under NA=0.42, and the efficiency and quality can be further enhanced by reducing the NA. At NA=0.24, the structure shows improved overall conversion efficiency over 60% with a ±25 nm bandwidth, peaking at 77% at the central wavelength of 1520 nm, as shown in Fig. 5(b). The reduction in experimental efficiency compared to numerical predictions can be attributed to three main factors. First, the roughness of the fabricated sample induces scattering losses, reducing the reflectance. Second, fabrication errors cause deviations from the optimal geometry. Third, interference speckle in the imaged beam introduces errors in the intensity integration. By focusing the beam to minimize the influence of interference speckles, efficiencies exceeding 80% under NA=0.42 were achieved. Further details on this efficiency measurement are provided in the Supplementary Section 4. Overall, the efficiency was preserved across a wide range of NAs and wavelengths, demonstrating the tunability and flexibility of our platform for diverse optical configurations.

It is noteworthy that the light intensity distribution of beams commonly used in imaging and computational applications is uniform in momentum space (i.e., momentum-space flat-top beams). Consequently, the calculated efficiency corresponds to an unweighted average over momentum space, as illustrated in Fig. 3(e, f). In contrast, another widely employed beam type —Gaussian beams —exhibits a

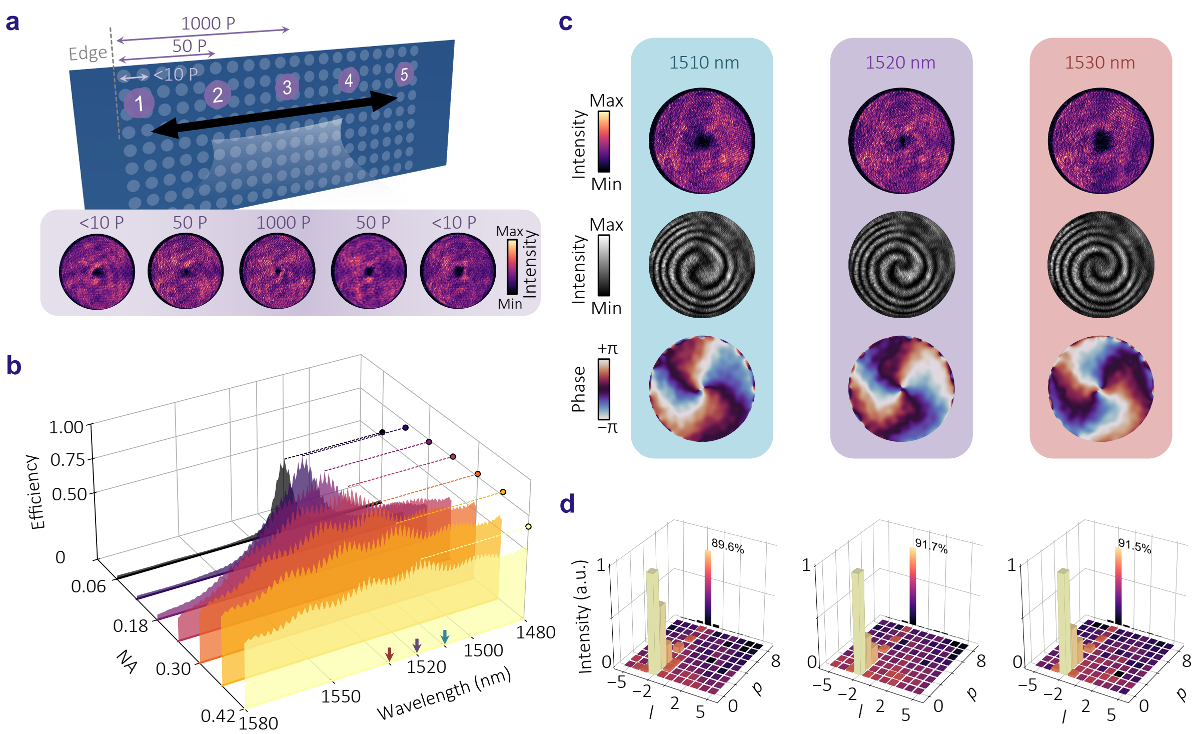


Fig. 5 | (a) Measured cross-polarization converted beam profile at different illumination positions. (b) Measured overall averaged efficiency under different wavelength and NA. (c) Measured generated beam profile. From top to bottom: intensity, interference pattern and extracted phase. (d) LG mode decomposition of the converted vortex beam under NA=0.21 at 1510, 1520, and 1530 nm. Inset: OAM (l) distribution, where the purity of l=−2 can reach as high as 91.7% at 1520 nm.

generation efficiency that deviates from that of flat-top beams by a Gaussian weighting factor[28]. Given our structure's robust performance across various wavevectors, the overall efficiency and quality remain comparable between these beam types. One can find more details in Supplementary information of Section 4.

Finally, the vortex phase and the quality of the generated beam in momentum-space were analyzed through interferometric imaging (Fig. 5(c)). The reference beam for interference, generated by front-surface reflections from polarizer P, was combined with the converted vortex beam for phase analysis via Fourier methods[42] (see Supplementary Section 9 for details). Three critical aspects were demonstrated: Firstly, the robustness of high-quality vortex generation across wavelengths was reconfirmed, with measured intensity profiles (first row) showing excellent agreement with simulated conversion efficiencies. Secondly, collinear interference with a reference plane wave (second row) exhibited two distinct helical arms in the fringe pattern, unambiguously verifying the topological charge of OAM = 2. Thirdly, reconstructed phase distributions (third row) revealed smooth 0–4π azimuthal variations, confirming high-fidelity vortex phase generation. To quantify beam quality, Laguerre-Gaussian (LG) mode decomposition was performed[43,44]. The decomposition component is illustrated in Fig. 5(d), showing >90% OAM purity under NA=0.21 (see Supplementary Section 9 for details). Here, the OAM purity did not reach theoretical prediction (>99%) due to experimental measurement errors and the azimuthal non-uniformity near the edge of high NA components. Notably, reducing the NA can improve both beam quality and efficiency, as detailed in LG mode decomposition provided in the Supplementary Section 9. These results collectively validate the high quality and adaptability of our vortex generation mechanism.

In addition to vortex beams generated under flat-top illumination, we investigated the performance of Bessel-beam incidence due to their distinctive momentum-space characteristics[8]. Although Bessel beams exhibit complex intensity profiles in real space, their momentum-space representations reduce to simple annular distributions, as shown in Fig. 6(a). This property makes them particularly suitable for momentum-space light-field modulation.

As illustrated in Fig. 6(b), zero-order Bessel beams were generated using axicon lenses, and subsequently converted into high-quality higher-order Bessel vortex beams by illuminating the same sample at NA=0.42. The reflected light, analyzed with the circular polarizer in Fig. 4(a), revealed a second-order Bessel vortex beam carrying OAM of $\ell = +2$. Conversion efficiency was quantified by comparing the intensity of the converted beam to that of a reference beam reflected by a gold mirror (labeled as "ref" in Fig. 6(d)). Although direct phase measurements were not performed for this proof-of-concept demonstration, the pure cross-polarization conversion via the PB phase mechanism, combined with the high phase quality verified in the flat-top experiments, ensures that the generated beams carry the expected OAM topological charge. The annular profile in momentum space enabled selective utilization of the regions with the highest efficiency, thereby improving performance metrics such as generation efficiency, wavelength tunability, and NA tolerance.

As shown in Fig. 6(c), the efficiency remains highly robust across wavelengths (1480–1600 nm), even under high NA conditions (NA = 0.42, where optimal NA is around 0.2 by simulation), in good agreement with theoretical predictions. Moreover, the annular intensity profile observed in Fig. 6(d)—characteristic of a "perfect vortex beam"—confirms

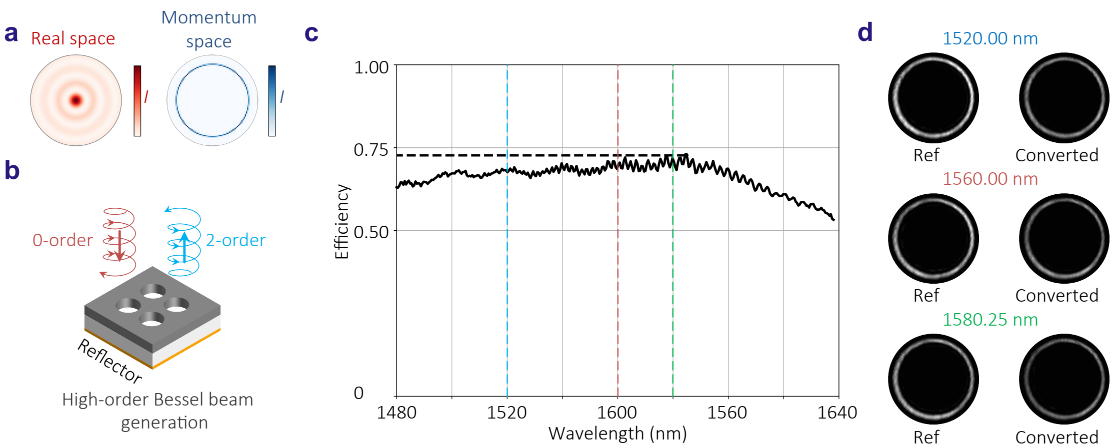


Fig. 6 | (a) Beam profile in real space (left) and momentum space (right) for zero-order Bessel beam. (b) Generation of high-order Bessel beam by direct illuminating the circular polarized zero-order Bessel beam at NA=0.42. (c) Measured conversion efficiency at various wavelengths. (d) Beam profiles at three different wavelengths. The leftmost column ("ref") shows the incident zero-order Bessel beam reflected by a gold mirror reference, while the right columns show the converted high-order Bessel vortex beams. The annular intensity distributions confirm high-quality "perfect vortex" generation.

the structure's ability to reliably impart vortex phase across diverse beam configurations without altering energy distributions. Detailed experimental setups are provided in Supplementary Section 8.

3 Discussion

The central physics of this work is that broadband, high-efficiency momentum-space vortex generation requires the synergistic engineering of the radiative Q-factor and dispersion. In our cavity-assisted nonlocal metasurface, coupling a symmetry-protected BIC to FP cavity modes simultaneously reduces the radiative Q-factor and flattens the dispersion band, yet these two effects play distinct roles. The reduced Q-factor enables strong cross-polarized conversion and a broad spectral response, while the quasi-flat dispersion maximizes momentum-space accessibility by allowing a wide range of in-plane wavevectors to satisfy the resonance condition at a fixed wavelength. A flat band alone is insufficient: a perfectly flat but high-Q band would still suffer from weak radiation ($Q_{\mathrm{rad}} \gg Q_{\mathrm{abs}}$), resulting in low efficiency and a narrow operational bandwidth. A reduced Q-factor alone is also insufficient: a dispersive but low-Q BIC band can enhance the on-resonance conversion efficiency and spectral linewidth, yet strong dispersion still confines efficient conversion to a ring-shaped region in momentum space, limiting the overall efficiency integrated over numerical aperture. Only their combination enables broadband, high-efficiency, and large-NA vortex generation.

We further emphasize that our cavity-coupling strategy provides a distinct approach to modulating the Q-factor around BICs in momentum space. It is well known that a broad family of perturbative methodologies, such as asymmetry perturbations[45] or Brillouin-zone-folding approaches[46–48], can transform ideal BICs into observable quasi-BICs. However, in these schemes the accessible radiative Q-factors typically remain in the high-Q regime ($Q \gtrsim 10^2$), which may limit conversion efficiency in the presence of material absorption. By contrast, FP-cavity coupling enables access to an ultra-low-Q regime ($Q \sim 10$), providing strong radiative coupling that is essential for overcoming absorption losses and achieving near-unity conversion efficiency. This capability also offers a promising platform for exploring rich non-Hermitian physics in the vicinity of BICs.

We also note that the cavity-assisted platform can, in principle, support merging BICs via Friedrich-Wintgen interference[49]. Such states are characterized by ultra-high Q-factors and strongly suppressed radiation over a large range of wavevectors, which are well suited for applications such as lasing or sensing but are incompatible with efficient vortex generation.

From an application perspective, this work resolves a long-standing limitation of nonlocal metasurfaces. While local metasurfaces manipulate phase in real space and can achieve high efficiency, their operation is inherently position-dependent and requires precise alignment. Nonlocal metasurfaces, by encoding phase in momentum space, offer alignment-free and position-independent functionality but have historically suffered from low overall efficiency due to dispersion mismatch. By overcoming this efficiency bottleneck through cavity-assisted band engineering, our work establishes nonlocal metasurfaces as a practical and efficient complement to local designs, particularly for broadband vortex generation.

Looking forward, we anticipate that nonlocal vortex generation will attract broader interest and find increasingly diverse applications. The proposed approach is, in principle, extendable to the visible spectral range and may be further developed for generating nondiffracting beams or vortex beam arrays[34]. We also note that the recently proposed generalized geometric phase, which provides a unified framework for phase manipulation, has laid the theoretical foundation for the evolution of meta-optics from version 1.0 to 2.0[50]. This emerging concept may introduce additional degrees of freedom for nonlocal wavefront engineering and further enrich the design space for momentum-space vortex generation.

4 Conclusions

In summary, we propose an innovative approach to achieve broadband operational vortex beam generation with high efficiency and high purity with a nonlocal metasurface. By establishing a hybrid coupling mechanism that links BIC with two DPs, we show that such a metasurface allows for precise, deterministic control over dispersion, Q-factor, and polarization state of the guided resonances. Simulation results show that such vortex beam generation can operate across a broad wavelength range (1500–1600 nm) while preserving near-unity conversion efficiency. We demonstrated such vortex beam generation experimentally by fabricating and characterizing the designed metasurfaces. Measurement results show the overall efficiency can approach 80% under both Gaussian and Bessel beams' illumination. We also confirmed enhanced generation quality and tunability across the 1480–1600 nm range under varied NA. Besides, we experimentally observed OAM purities of up to 91.7% and exceptionally robust performance against edge effects in finite-size arrays. This approach surpasses the intrinsic trade-offs of previous techniques, establishing a practical pathway for broadband, high-efficiency momentum-space vortex beam generation in real-world applications.

5 Methods

5.1 Theoretical analysis

Please see the Supplementary Section 1, 2 for the derivations.

5.2 Device fabrication

To fabricate the desired structure, an 8 nm thick titanium adhesion, a 70 nm thick Au film and a subsequent 5 nm thick titanium layer were successively sequentially deposited onto a Si substrate using electron beam evaporation (EBE; DETECH, DE500C) at a deposition rate of 0.6 Å/s (1 Å = 10^{-10} m) under a chamber pressure of 9×10^{-7} Torr (1 Torr = 133.3 Pa). A 545 nm SiO_2 layer followed by a 157 nm Si layer were deposited using plasma-enhanced chemical vapor deposition (PECVD). Next, the PMMA resist (950 A4 from KayakuAM) was spin-coated onto the sample at 3000 rpm for 60 s, and baked at 180 °C for 90 s, producing a uniform 220 nm thick resist film. The reverse designed structure patterns were then exposed onto the PMMA layer using a 50 kV electron beam lithography (EBL) system (Raith Voyager), and a subsequent development of the exposed resist was performed to obtain the desired PMMA pattern. A post development bake was carried out on a hot plate at 100 °C for 90 s. The final Si structures were obtained via inductively coupled plasma-enhanced reactive ion etching (ICP-RIE, Oxford Instruments, Plasma Pro 80 Cobra). Finally, the sample was immersed in acetone for 5 minutes to completely remove the residual PMMA, followed by ethanol rinsing to ensure a clean and contaminant-free surface. The morphology of the metasurface sample was characterized by a scanning electron microscope (Sigma 500, Zeiss).

5.3 Numerical simulation

The eigenfrequencies of the structure were computed using the Electromagnetic Waves, Frequency Domain (EWFD) module in COMSOL Multiphysics.

The s-polarized reflectance spectrum R_s was obtained through full-wave simulations conducted with the Rigorous Coupled-Wave Analysis (RCWA) module in ANSYS Lumerical FDTD. Both solvers used the same complex refractive indices to capture the influence of material absorption on conversion efficiency: $n_{Si} = 3.58 + 10^{-4}i$, $n_{SiO_2} = 1.45 + 10^{-4}i$. The chosen imaginary part $(k = 10^{-4})$ reflects realistic losses in silicon and silicon dioxide films grown by CVD and is critical for accurate prediction of cross-polarization conversion efficiencies. The experimentally observed shift of the BIC resonance from approximately 1550 nm to 1520 nm is attributed to the slightly higher refractive index n_{Si} used in the simulations compared to the experimental material, as confirmed by numerical comparison detailed in the Supplementary information of Section 7.

References

1. Paterson C. Atmospheric turbulence and orbital angular momentum of single photons for optical communication. *Phys Rev Lett* **94**, 153901 (2005).
2. Fickler R, Campbell G, Buchler B et al. Quantum entanglement of angular momentum states with quantum numbers up to 10, 010. *Proc Natl Acad Sci USA* **113**, 13642–13647 (2016).
3. Ndagano B, Perez-Garcia B, Roux FS et al. Characterizing quantum channels with non-separable states of classical light. *Nat Phys* **13**, 397–402 (2017).
4. Guo YH, Zhang SC, Pu MB et al. Spin-decoupled metasurface for simultaneous detection of spin and orbital angular momenta via momentum transformation. *Light Sci Appl* **10**, 63 (2021).
5. Fürhapter S, Jesacher A, Bernet S et al. Spiral phase contrast imaging in microscopy. *Opt Express* **13**, 689–694 (2005).
6. Tamburini F, Anzolin G, Umbriaco G et al. Overcoming the Rayleigh criterion limit with optical vortices. *Phys Rev Lett* **97**, 163903 (2006).
7. Zhang Z, Li GY, Liu YL et al. Robust measurement of orbital angular momentum of a partially coherent vortex beam under amplitude and phase perturbations. *Opto-Electron Sci* **3**, 240001 (2024).
8. Lian YD, Qi X, Wang YH et al. OAM beam generation in space and its applications: a review. *Opt Lasers Eng* **151**, 106923 (2022).
9. Shen YJ, Wang XJ, Xie ZW et al. Optical vortices 30 years on: OAM manipulation from topological charge to multiple singularities. *Light Sci Appl* **8**, 90 (2019).
10. Yu NF, Genevet P, Kats MA et al. Light propagation with phase discontinuities: generalized laws of reflection and refraction. *Science* **334**, 333–337 (2011).
11. Marrucci L, Manzo C, Paparo D. Optical spin-to-orbital angular momentum conversion in inhomogeneous anisotropic media. *Phys Rev Lett* **96**, 163905 (2006).
12. Kotlyar VV, Almazov AA, Khonina SN et al. Generation of phase singularity through diffracting a plane or Gaussian beam by a spiral phase plate. *J Opt Soc Am A* **22**, 849–861 (2005).
13. Chen QM, Qu GY, Yin J et al. Highly efficient vortex generation at the nanoscale. *Nat Nanotechnol* **19**, 1000–1006 (2024).
14. Mehmood MQ, Mei ST, Hussain S et al. Visible-frequency metasurface for structuring and spatially multiplexing optical vortices. *Adv Mater* **28**, 2533–2539 (2016).
15. Yu NF, Capasso F. Flat optics with designer metasurfaces. *Nat Mater* **13**, 139–150 (2014).
16. Berry MV. The adiabatic phase and Pancharatnam's phase for polarized light. *J Mod Opt* **34**, 1401–1407 (1987).
17. Zhou Y, Zheng HY, Kravchenko II et al. Flat optics for image differentiation. *Nat Photonics* **14**, 316–323 (2020).
18. Pu MB, Li X, Ma XL et al. Catenary optics for achromatic generation of perfect optical angular momentum. *Sci Adv* **1**, e1500396 (2015).
19. Guo YH, Pu MB, Zhao ZY et al. Merging geometric phase and plasmon retardation phase in continuously shaped metasurfaces for arbitrary orbital angular momentum generation. *ACS Photonics* **3**, 2022–2029 (2016).
20. Zheng YQ, Liao Y, Fan YL et al. Unevenly distributed pixel-based camouflage metasurface hiding multiwavelength holograms in color printing. *Adv Photonics* **7**, 016003 (2025).
21. He GL, Zheng YQ, Zhou CD et al. Multiplexed manipulation of orbital angular momentum and wavelength in metasurfaces based on arbitrary complex-amplitude control. *Light Sci Appl* **13**, 98 (2024).
22. Zhou YJ, Liu T, Dai CH et al. Functionality multiplexing in high-efficiency metasurfaces based on coherent wave interferences. *Opto-Electron Adv* **7**, 240086 (2024).
23. Wang B, Liu WZ, Zhao MX et al. Generating optical vortex beams by momentum-space polarization vortices centred at bound states in the continuum. *Nat Photonics* **14**, 623–628 (2020).
24. Zhen B, Hsu CW, Lu L et al. Topological nature of optical bound states in the continuum. *Phys Rev Lett* **113**, 257401 (2014).
25. Zhang YW, Chen A, Liu WZ et al. Observation of polarization vortices in momentum space. *Phys Rev Lett* **120**, 186103 (2018).
26. Chen Y, Wang MJ, Si JH et al. Observation of chiral emission enabled by collective guided resonances. *Nat Nanotechnol* **20**,

1205–1212 (2025).
27. Qin HY, Su ZP, Zhang Z et al. Disorder-assisted real-momentum topological photonic crystal. *Nature* **639**, 602–608 (2025).
28. Zhang WJ, Chu J, Deng RH et al. Optical vortices generation via a self-assembly photonic crystal slab. *Adv Opt Mater* **12**, 2400088 (2024).
29. De Tommasi E, Romano S, Mocella V et al. Half-integer topological charge polarization of quasi-dirac bound states in the continuum. *Adv Opt Mater* **11**, 2300475 (2023).
30. Zhang TC, Dong KC, Li JC et al. Twisted moiré photonic crystal enabled optical vortex generation through bound states in the continuum. *Nat Commun* **14**, 6014 (2023).
31. Liu WZ, Shi L, Zi J et al. Ways to achieve efficient non-local vortex beam generation. *Nanophotonics* **10**, 4297–4304 (2021).
32. Xu H, Chen JG, Wang B et al. Non-local metasurface generates highly efficient transmission vortex by intrinsic singularity and generalized Kerker effect. *PhotoniX* **6**, 8 (2025).
33. Liu XJ, Gu MN, Tian Y et al. Research progress on generating perfect vortex beams based on metasurfaces. *Opto-Electron Sci* **4**, 250007 (2025).
34. Zhou CD, Zheng YQ, He GL et al. Generating arrays of perfect vector vortex beam with six on-demand degrees of freedom in metasurfaces. *Sci China Phys, Mech Astron* **68**, 254211 (2025).
35. Li TY, Wang JJ, Zhang WJ et al. High-efficiency nonlocal reflection-type vortex beam generation based on bound states in the continuum. *Natl Sci Rev* **10**, nwac234 (2023).
36. Suh W, Wang Z, Fan SH. Temporal coupled-mode theory and the presence of non-orthogonal modes in lossless multimode cavities. *IEEE J Quantum Electron* **40**, 1511–1518 (2004).
37. Huang LJ, Xu L, Powell DA et al. Resonant leaky modes in all-dielectric metasystems: fundamentals and applications. *Phys Rep* **1008**, 1–66 (2023).
38. Sun KL, Wei H, Chen WJ et al. Infinite-Q guided modes radiate in the continuum. *Phys Rev B* **107**, 115415 (2023).
39. Wang KR, Sun KL, Ding Q et al. High-Q resonance engineering in momentum space for highly coherent and rainbow-free thermal emission. *Nano Lett* **25**, 3613–3619 (2025).
40. Sun KL, Levy U, Han ZH. Exploiting zone-folding induced quasi-bound modes to achieve highly coherent thermal emissions. *Nano Lett* **24**, 764–769 (2024).
41. Hu P, Wang JJ, Jiang Q et al. Global phase diagram of bound states in the continuum. *Optica* **9**, 1353 (2022).
42. Carlon Zambon N, St-Jean P, Milićević M et al. Optically controlling the emission chirality of microlasers. *Nat Photonics* **13**, 283–288 (2019).
43. Borghi R, Gori F, Santarsiero M. Optimization of Laguerre-Gauss truncated series. *Opt Commun* **125**, 197–203 (1996).
44. Xiao Y, Tang XH, Wan CH et al. Laguerre-Gaussian mode expansion for arbitrary optical fields using a subspace projection method. *Opt Lett* **44**, 1615–1618 (2019).
45. Koshelev K, Lepeshov S, Liu MK et al. Asymmetric metasurfaces with high-Q resonances governed by bound states in the continuum. *Phys Rev Lett* **121**, 193903 (2018).
46. Wang WH, Srivastava YK, Tan TC et al. Brillouin zone folding driven bound states in the continuum. *Nat Commun* **14**, 2811 (2023).
47. You SJ, Zhou MM, Xu L et al. Quasi-bound states in the continuum with a stable resonance wavelength in dimer dielectric metasurfaces. *Nanophotonics* **12**, 2051–2060 (2023).
48. Zhong HZ, Huang LJ, Li SL et al. Toroidal dipole bound states in the continuum in asymmetric dimer metasurfaces. *Appl Phys Rev* **11**, 031404 (2024).
49. Tu QA, Zhou HX, Ji CL et al. Dynamic control of polarization singularities in momentum space by a hybrid photonic system. *Laser Photonics Rev e01685* (2025).
50. Xie X, Pu MB, Jin JJ et al. Generalized Pancharatnam-berry phase in rotationally symmetric meta-atoms. *Phys Rev Lett* **126**, 183902 (2021).

Acknowledgements

This work was supported by the Natural Science Foundation of Sichuan Province (Grant No. 2024NSFSC0460) and the National Natural Science Foundation of China (Grant No. 12274269, 12474377)

Author contributions

Keren Wang and Kaili Sun contributed equally to this work.

Competing interests

The authors declare no competing financial interests.

Data availability

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

Supplementary information

Supplementary information for this paper is available at
https://doi.org/10.29026/oea.2026.250296

Published by Editorial Office of *Opto-Electronic Advance*, Institute of Optics and Electronics, Chinese Academy of Sciences.