In the format provided by the authors and unedited.

Single-shot quantitative phase gradient microscopy using a system of multifunctional metasurfaces

Hyounghan Kwon[1,2], **Ehsan Arbabi** [1,2], **Seyedeh Mahsa Kamali**[1,2], **MohammadSadegh Faraji-Dana**[1,2] **and Andrei Faraon** [1,2]*

[1]T. J. Watson Laboratory of Applied Physics and Kavli Nanoscience Institute, California Institute of Technology, Pasadena, CA, USA. [2]Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, USA. *e-mail: faraon@caltech.edu

Supplementary information: Single Shot Quantitative Phase Gradient Microscopy Using a System of Multifunctional Metasurfaces

Hyounghan Kwon,[1,2] Ehsan Arbabi,[1,2] Seyedeh Mahsa Kamali,[1,2]

MohammadSadegh Faraji-Dana,[1,2] and Andrei Faraon[1,2,*]

[1]*T. J. Watson Laboratory of Applied Physics and Kavli Nanoscience Institute,*

California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA

[2]*Department of Electrical Engineering, California Institute of Technology,*

1200 E. California Blvd., Pasadena, CA 91125, USA

[*] Corresponding author: A.F.: faraon@caltech.edu

SUPPLEMENTARY NOTE 1: WAVE PROPAGATION SIMULATION AND PHASE MAPS OF THE METASURFACES

In this section we present numerical analysis results of the QPGM. First, it is worth explaining why the QPGM consists of two metasurface layers, and a single layer is not capable of doing so. Then, we verify that the QPGM based on the two metasurface layers addresses the issues faced in the system using a single metasurface. We used wave-propagation simulations to analyze the systems in Fig. **S1**. In the simulations, the metasurfaces and the conventional lens are modeled as ideal phase plates. Moreover, the thickness of all fused silica substrates is fixed at 1 mm. For all metasurfaces, the distance between the optical axes for TE and TM polarizations, Δs, is fixed at 1.5 μm along the y-axis. Finally, the phase sample shown in Fig. 1**c** is used for the all simulations.

Figure **S1a** shows a schematic of a DIC system based on a single birefringent metasurface lens and a regular refractive lens, forming a 4-f imaging system. To implement the single metasurface layer system, we use the thin lens equation for the phase profiles of the metasurface lens (ML$_{1,a}$) and the conventional lens (Lens$_2$, a). More specifically, ML$_{1,a}$ has two different phase profiles for TE and TM polarizations, ($\phi_{\text{ML}_{1,a},\text{TE}}$ and $\phi_{\text{ML}_{1,a},\text{TM}}$), given by:

$$\phi_{\text{ML}_{1,a},\text{TE}} = -\frac{\pi}{\lambda f_1}(x^2 + (y + \frac{\Delta s}{2})^2) \tag{S1}$$

$$\phi_{\text{ML}_{1,a},\text{TM}} = -\frac{\pi}{\lambda f_1}(x^2 + (y - \frac{\Delta s}{2})^2), \tag{S2}$$

where x and y are Cartesian coordinates from the center of ML$_{1,a}$ and λ is the operating wavelength in vacuum. Moreover, the polarization-insensitive phase profile of Lens$_{2,a}$ is written as, $\phi_{\text{Lens}_{2,a}} = -\frac{\pi}{\lambda f_2}(x^2 + y^2)$. While one might expect that the configuration in Fig. **S1a** works similar to a conventional DIC microscope, an additional spurious intensity gradient shows up in the formed interference pattern, I_{single}, as seen in Fig. **S1b**. The reason is that despite the 4-f imaging system, the two slightly separated optical axes for the TE and TM polarizations cause the intensity gradient not present in the original target (see Supplementary Note 2 and Supplementary Fig. **S2** for experimental results of the single metasurface system and theoretical analysis about the degradation). Another clear issue with the system shown in in Fig. **S1a** which undermines its miniature size is that it would require a variable phase retarder to retrieve quantitative phase gradient information.

To resolve the spurious intensity gradient issue first, we replace the refractive lens with a second birefringent metasurface lens in Fig. **S1c**. For the system based on the two bifocal metasurface lenses shown in Fig. **S1c**, the metasurface lens 1, $ML_{1,c}$, is the same as $ML_{1,a}$ shown in Fig. **S1a**. A second birefringent metasurface lens, $ML_{2,c}$, is used in the system shown in Fig. **S1c** instead of $Lens_2$, a in Fig. **S1a**. The two phase profiles of $ML_{2,c}$ for TE and TM polarizations, ($\phi_{ML_{2,c},TE}$ and $\phi_{ML_{2,c},TM}$), are given by:

$$\phi_{ML_{2,c},TE} = -\frac{\pi}{\lambda f_2}(x^2 + (y + \frac{\Delta s}{2})^2) \tag{S3}$$

$$\phi_{ML_{2,c},TM} = -\frac{\pi}{\lambda f_2}(x^2 + (y - \frac{\Delta s}{2})^2) + \phi_o, \tag{S4}$$

where ϕ_o is the phase offset between the two orthogonal polarizations. ϕ_o is fixed at $\frac{3\pi}{4}$ for the simulations. As a result, the simulated interference intensity map shown in Fig. **S1d**, I_{double}, is a clear DIC image of the transparent object with no intensity artifacts (see Supplementary Note 2 for theoretical explanation about the system of the two birefringent metasurface lenses). In the experimental implementation, the values of ϕ_o are $\frac{3\pi}{4}$ and $\frac{\pi}{4}$ for the QPGM using two separate substrates and the double-sided QPGM, respectively.

In addition to birefringence, the capability of metasurfaces to simultaneously perform multiple independent functions allows us to eliminate the requirement of a variable retarder, since several images with different phase offsets can be captured simultaneously as shown schematically in Fig. **S1e**. Moreover, we employed the ray tracing method instead of the thin lens equation to mitigate geometric aberrations. In other words, the phase profiles are optimized for the metasurface layers 1 and 2 in Fig. **S1e**, $Layer_{1,e}$ and $Layer_{2,e}$, to minimize the off-axis aberrations and increase the field of view. Specifically, the phase profiles of the metasurfaces are defined by two terms. One is a sum of even-order polynomials of radial coordinates (mostly implementing the focusing), and the other is a linear phase gradient term associated with the three-directional splitting of light. The phase profiles of $Layer_{1,e}$ for TE and TM polarizations, $\phi_{Layer_{1,e},TE}$ and $\phi_{Layer_{1,e},TM}$, are given by:

$$\phi_{Layer_{1,e},TE} = \sum_{n=1}^{5} \frac{a_n}{R^{2n}}(x^2 + (y + \frac{\Delta s}{2})^2)^n - k_{grat,1}y \tag{S5}$$

$$\phi_{Layer_{1,e},TM} = \sum_{n=1}^{5} \frac{a_n}{R^{2n}}(x^2 + (y - \frac{\Delta s}{2})^2)^n - k_{grat,1}y \tag{S6}$$

$$\phi_{\text{Layer}_{1,e},\text{TE}} = \sum_{n=1}^{5} \frac{a_n}{R^{2n}}(x^2 + (y + \frac{\Delta s}{2})^2)^n - k_{grat,1}x \tag{S7}$$

$$\phi_{\text{Layer}_{1,e},\text{TM}} = \sum_{n=1}^{5} \frac{a_n}{R^{2n}}(x^2 + (y - \frac{\Delta s}{2})^2)^n - k_{grat,1}x \tag{S8}$$

$$\phi_{\text{Layer}_{1,e},\text{TE}} = \sum_{n=1}^{5} \frac{a_n}{R^{2n}}(x^2 + (y + \frac{\Delta s}{2})^2)^n + k_{grat,1}y \tag{S9}$$

$$\phi_{\text{Layer}_{1,e},\text{TM}} = \sum_{n=1}^{5} \frac{a_n}{R^{2n}}(x^2 + (y - \frac{\Delta s}{2})^2)^n + k_{grat,1}y, \tag{S10}$$

where a_n are the optimized coefficients of the even-order polynomials in the shifted radial coordinates, $k_{grat,1}$ is the linear phase gradient, and R denotes the radius of the metasurfaces. Detailed information about a_n, $k_{grat,1}$, and R is given in Table **S1**. Since a single set of rectangular nano-posts can only implement one pair of the birefringent phase maps, three different sets of rectangular nano-posts are designed to achieve the three pairs of phase maps in (Eqs. **S5** and **S6**), (Eqs. **S7** and **S8**), and (Eqs. **S9** and **S10**). Then, the three maps of the rectangular nano-posts are interleaved along the *x*-axis using the spatial multiplexing method [1–3]. With the phase profiles in Eqs. **S5**-**S10**, the Layer$_{1,e}$ in Fig. **S1e** plays the role of a lens and a three directional beam-splitter at the same time, at the cost of a drop in efficiency.

Layer$_{2,e}$ has three different birefringent metasurfaces which are identically displaced from the center of Layer$_{2,e}$. The distance from the center of the Layer$_{2,e}$ to the center of each lens, ΔD, is 660 μm. The three coordinates of the centers of the lenses measured from the center of the Layer$_{2,e}$ are (0,-ΔD) (-ΔD,0) and (0,ΔD). The six phase profiles of the three lenses for TE and TM polarizations are written as:

$$\phi_{\text{Layer}_{2,e},\text{TE}} = \sum_{n=1}^{5} \frac{b_n}{R^{2n}}(x^2 + (y + \Delta D + \frac{\Delta s}{2})^2)^n + k_{grat,2}y \tag{S11}$$

$$\phi_{\text{Layer}_{2,e},\text{TM}} = \sum_{n=1}^{5} \frac{b_n}{R^{2n}}(x^2 + (y + \Delta D - \frac{\Delta s}{2})^2)^n + k_{grat,2}y + \phi_o \tag{S12}$$

$$\phi_{\text{Layer}_{2,e},\text{TE}} = \sum_{n=1}^{5} \frac{b_n}{R^{2n}}((x + \Delta D)^2 + (y + \frac{\Delta s}{2})^2)^n + k_{grat,2}x \tag{S13}$$

4

$$\phi_{\text{Layer}_2,\text{e},\text{TM}} = \sum_{n=1}^{5} \frac{b_n}{R^{2n}} ((x + \Delta D)^2 + (y - \frac{\Delta s}{2})^2)^n + k_{grat,2}x + \phi_o + \frac{2\pi}{3} \tag{S14}$$

$$\phi_{\text{Layer}_2,\text{e},\text{TE}} = \sum_{n=1}^{5} \frac{b_n}{R^{2n}} (x^2 + (y - \Delta D + \frac{\Delta s}{2})^2)^n - k_{grat,2}y \tag{S15}$$

$$\phi_{\text{Layer}_2,\text{e},\text{TM}} = \sum_{n=1}^{5} \frac{b_n}{R^{2n}} (x^2 + (y - \Delta D - \frac{\Delta s}{2})^2)^n - k_{grat,2}y + \phi_o + \frac{4\pi}{3}, \tag{S16}$$

where b_n are the optimized coefficients of the even-order polynomials of the shifted radial coordinates and $k_{grat,2}$ is the linear phase gradient. The detailed information about b_n and $k_{grat,2}$ is given in Table S1. As a result, the two layers form the three different DIC images at the image plane. Specifically, combinations of (Eqs. S5, S6, S11, and S12), (Eqs. S7, S8, S13, and S14), and (Eqs. S9, S10, S15, and S16) result in the three phase-shifted DIC images in Fig. S1f, I_1, I_2 and I_3, respectively. To be specific, the desired phase offsets for the three-step phase shifting are achieved by the phase maps of the second metasurface layer in Eqs. S11-S16. Moreover, we should point out that I_1, I_2 and I_3 in Fig. S1f are comparable to the ideal results shown in Fig. 1d. Figure S1g shows the PGI calculated from the three DIC images in Fig. S1f by using Eq.2 in the main text, and is in good agreement with the ideal PGI shown in Fig. 1e.

As shown in Figs. 4a and 4b in the main text, the QPGM is also implemented using double-sided metasurfaces on a 1-mm thick fused silica wafer. The double-sided metasurface QPGM is designed through an identical process used for the design of the QPGM based on two metasurface layers on two separate substrates. The phase profiles of the double-sided metasurfaces are determined by Eqs. S5-S16 with the optimized phase profile parameters such as a_n, b_n, R, ΔD, Δs, $k_{grat,1}$, and $k_{grat,2}$ given in Table S2. The distances from the object plane to the metasurface layer 1 and from the image plane to the metasurface layer 2 are 317 μm and 397 μm, respectively.

5

SUPPLEMENTARY NOTE 2: EXPERIMENTAL AND THEORETICAL RESULTS WITH A SINGLE BIFOCAL METASURFACE LENS

We fabricated a single bifocal metasurface and measured its performance. The phase map of the device is determined by Eqs. **S1** and **S2**. For the fabricated single bifocal metasurface lens, the diameter, focal length, and separation of the focal points are 900 μm, 1.2 mm, and 2 μm, respectively. The two focal points are characterized through the optical setup shown in Fig. **S2a**. An 850-nm semiconductor laser (Thorlabs, L850P010) is coupled to a single mode fiber and a laser collimator for illumination. The linear polarizer in front of the laser collimator is aligned to 0° (for TE), 90° (for TM), and 45° for characterization of the bifocal metasurface lens. The measured intensity maps shown in Supplementary Fig. **S2b** clearly show the polarization dependent bifocal property. The intensity profiles on the black dashed lines in Fig. **S2b** are shown in Fig. **S2c**. In Fig. **S2c**, the FWHM of the two focal points and the distance between the two focal points are 1.21-1.35 μm and 2 μm, respectively. Moreover, the measured focusing efficiencies through a pin-hole with a diameter of about 4×FWHM at the focal plane are ~78% for both TE and TM polarizations.

We performed imaging experiments with the single bifocal metasurface lens using the optical setup schematically shown in Fig. **S2d**. We employed a variable retarder to capture three different DIC images. Moreover, oblique illumination from an LED was used to avoid saturation at the central pixels of the camera resulting from undiffracted light. To limit the bandwidth, we employed a band-pass filter with a center wavelength and bandwidth of 850 nm and 10 nm, respectively. The three captured DIC images, I_1, I_2, and I_3, are shown in Fig. **S2e**. The results clearly show that the spurious graded intensity patterns degrade the DIC images as expected from Fig. **S1b**. In Fig. **S2f**, the PGI is calculated from the three DIC images in Fig. **S2e** through the three-step phase shifting method [4]. The graded phase gradient in the background in Fig. **S2f** also indicates the imperfect imaging performance of the single bifocal metasurface lens.

The degradation exists both in the simulation results shown in Fig. **S1b** and the measurement in Figs. **S2e** and **S2f**. The degradation can be explained using Fourier optics. If we consider the optical system consisting of one metasurface bi-focal lens and a normal thin lens as in Fig. **S1a** such that the z-axis passes through the center of the second lens (i.e. the optical axes of the first lens for TE and TM polarizations are Δs/2 distant from the z-axis), we can write down the relations between the field at the object plane (U) and the fields at the Fourier plane ($U_{F,TE}$ and $U_{F,TM}$) through the Fresnel diffraction. For simplicity, let's assume that the focal lengths of the two lenses

in Fig. **S1a** are identical. In this case, the two optical fields formed by the first metasurface lens, $U_{F,TE}$ and $U_{F,TM}$ are no longer simple Fourier transforms of U, but would instead be written as:

$$U_{F,TE}(x,y) = \iint \iint U(x'',y'')exp[i\frac{\pi}{\lambda f}((x'-x'')^2 + (y'-y'')^2)]dx''dy''$$

$$\times exp[-i\frac{\pi}{\lambda f}(x'^2 + (y' + \frac{\Delta s}{2})^2))])exp[i\frac{\pi}{\lambda f}((x-x')^2 + (y-y')^2)]dx'dy'$$

$$= A_1 \iint U(x'',y'')exp[i\frac{\pi}{\lambda f}(x^2 + x''^2 + y^2 + y''^2)]$$

$$\times \iint exp[i\frac{\pi}{\lambda f}(x'^2 + y'^2 - 2(x + x'')x' - 2(y + y'' + \frac{\Delta s}{2})y']dx'dy'dx''dy''$$

$$= A_2 exp(-i\frac{2\pi}{\lambda f}\frac{\Delta s}{2}y) \iint U(x'',y'')exp[-i\frac{2\pi}{\lambda f}(xx'' + (y + \frac{\Delta s}{2})y'')]dx''dy''$$

$$= A_2 exp(-i\frac{2\pi}{\lambda f}\frac{\Delta s}{2}y)\widetilde{U}(\frac{x}{\lambda f}, \frac{y + \frac{\Delta s}{2}}{\lambda f})$$

$$U_{F,TM}(x,y) = A_2 exp(i\frac{2\pi}{\lambda f}\frac{\Delta s}{2}y)\widetilde{U}(\frac{x}{\lambda f}, \frac{y - \frac{\Delta s}{2}}{\lambda f}).$$

where λ, f, and Δs are the wavelength, focal length of the two lenses, and the separation between the optical axes of the metasurface lens, respectively. Also, $\widetilde{U}$ represents the 2D Fourier transform of U. A_1 and A_2 are complex constants. Then, the fields at the image plane ($U_{i,TE}$ and $U_{i,TM}$) can be written through the 2D Fourier transform of $U_{F,TE}$ and $U_{F,TM}$;

$$U_{i,TE}(x,y) = \iint U_{F,TE}(x',y')exp[-i\frac{2\pi}{\lambda f}(xx' + yy')]dx'dy' = B_1 exp(+i\frac{2\pi}{\lambda f}\frac{\Delta s}{2}y)U(-x, -(y + \frac{\Delta s}{2}))$$

$$U_{i,TM}(x,y) = \iint U_{F,TM}(x',y')exp[-i\frac{2\pi}{\lambda f}(xx' + yy')]dx'dy' = B_1 exp(-i\frac{2\pi}{\lambda f}\frac{\Delta s}{2}y)U(-x, -(y - \frac{\Delta s}{2})),$$

where B_1 is a complex constant. We should point out that the fields at the image plane are not only shifted laterally by the separation Δs, but also accompanied with a different phase gradient for TE and TM polarizations. The different phase gradients for TE and TM polarizations cause the spurious degradation shown in Figs. **S1b**, **S2e**, and **S2f**. Considering that Δs is close to the diffraction-limit scale, the simulation result in Fig. **S1b** and the measurement in Figs. **S2e** and **S2f** clearly reveal that the system is very sensitive to even very small values of Δs. Furthermore, the inevitable axial misalignment between the double axes of the metasurface lens and the single axis of the normal lens results in the phase gradient difference at the image plane. In other words, lens-based DIC microscopy typically necessitates at least three optical elements which are two lenses and a birefringent crystal such as a Wollaston prism between the two lenses. Although one might suppose that the different linear phase gradients can be employed instead of the lateral shifts along the y-axis in Eqs. **S1** and **S2** to remove the spurious degradation, it is worth pointing out that both schemes are mathematically equivalent with a different complex constant and result in a similar degradation.

In contrast, the two metasurface layers can avoid the unwanted degradation because the two lenses have independent optical axes for both polarizations. For the system in Fig. **S1c**, the fields at the image plane can be written as the Fourier transform with respect to the optical axis of each

polarization. As a result, the fields at the image plane $U_{i,TE}$ and $U_{i,TM}$ can be written as;

$$U_{i,TE}(x, y) = U(-Mx, -M(y + \tfrac{\Delta s}{2}) - \tfrac{\Delta s}{2})$$
$$U_{i,TM}(x, y) = U(-Mx, -M(y - \tfrac{\Delta s}{2}) + \tfrac{\Delta s}{2}),$$

where M is the magnification of the system and is mainly determined by the phase maps of the two bi-axial lenses. These equations clearly show that the two birefringent metasurface lenses are able to capture the conventional DIC images with two polarizers. More interestingly, the simulation results based on the wave propagation in Fig. **S13** reveal that the proposed bi-axial system becomes very robust against other kinds of misalignments in the fabrication or optical alignment procedures once the separations of the optical axes for the both lenses are identical. As explained in the main manuscript, it is worth noting that the metasurfaces uniquely admit arbitrary phase maps in subwavelength scale for both polarizations. Thus, the separation is one of the most accurately controllable variables in the design process of the metasurfaces. This is one of the reasons why our system performs adequately in the experiments. Furthermore, we remark that bi-focal lenses having an accurate separation close to the diffraction-limit are difficult to implement using any conventional birefringent optics.

SUPPLEMENTARY NOTE 3: THREE-STEP PHASE SHIFTING METHOD FOR UNIDIRECTIONAL QUANTITATIVE PHASE GRADIENT IMAGING

Phase-shifting is a widely known technique for phase retrieval in common interferometers. Moreover, it has been also used for quantitative phase gradient imaging by modifying the classical DIC microscope [5]. The technique utilizes multiple phase-shifted interference patterns to retrieve the phase information. To be specific, a captured differential interference contrast (DIC) image at the object plane is effectively written as;

$$I_j = |U(x,y) - U(x, y - \Delta y)e^{i\phi_j}|^2 = B(x,y) - C(x,y)cos(\theta(x,y) - \phi_j) \tag{S17}$$

, where $U(x,y) = A(x,y)e^{i\phi(x,y)}$, $B(x,y) = |A(x,y)|^2 + |A(x, -\Delta y)|^2$, $C(x,y) = 2|A(x,y)A(x, -\Delta y)|$, $\theta(x,y) = \phi(x,y) - \phi(x, y - \Delta y)$, and ϕ_j is a phase offset. Since $B(x,y)$, $C(x,y)$, and $\theta(x,y)$ are unknown, it can be seen that a minimum of three independent measurements are required for unambiguous retrieval of all unknowns. This is why the second metasurface layer in the current work consists of three different polarization sensitive off-axis lenses. Moreover, it is worth noting here that extensive investigations have previously been done to develop two-step phase shifting algorithms with additional assumption, some form of a priori knowledge, or complicated computations [6–8]. In the three-step phase shifting techniques, ϕ_1, ϕ_2, and ϕ_3 are usually set to 0, $\frac{2\pi}{3}$, and $\frac{4\pi}{3}$. Using the Eq. S17, I_1, I_2, and I_3 are written as;

$$I_1 = B(x,y) - C(x,y)cos(\theta(x,y)),$$
$$I_2 = B(x,y) - C(x,y)cos(\theta(x,y) - \tfrac{2\pi}{3}), \text{ and}$$
$$I_3 = B(x,y) - C(x,y)cos(\theta(x,y) - \tfrac{4\pi}{3}).$$

With further calculations, one can see that $\theta(x,y)$ can be expressed in terms of I_1, I_2, and I_3;

$$\theta(x,y) = arctan(\sqrt{3}\tfrac{I_2 - I_3}{2I_1 - I_2 - I_3})$$

Considering that $\theta(x,y) = \phi(x,y) - \phi(x, y - \Delta y) \approx \nabla_y \phi \times \Delta y$ and Δy is small compared to the sample feature sizes, $\nabla_y \phi$ can be calculated as;

$$\nabla_y \phi \approx \tfrac{1}{\Delta y} arctan(\sqrt{3}\tfrac{I_2 - I_3}{2I_1 - I_2 - I_3})$$

In Eq. 2 in the main manuscript, a calibration term, $\nabla_y \phi_{cali}$, is added to remove any kind of unwanted background coming from imperfect experimental conditions. Specifically, $\nabla_y \phi_{cali}$ is the background signal measured without any sample. The calibration process also allows for arbitrarily choosing three phase offsets with the same difference of $\frac{2\pi}{3}$. For example, ϕ_1 is set to $\frac{3\pi}{4}$ and $\frac{\pi}{4}$

for the device using the two separate metasurface layers and the device based on the double-sided metasurfaces, respectively.

I. REFERENCES

---

[1] Maguid, E. *et al.* Photonic spin-controlled multifunctional shared-aperture antenna array. *Science* **352**, 1202-1206 (2016).

[2] Arbabi, E., Arbabi, A., Kamali, S. M., Horie, Y. & Faraon, A. Multiwavelength polarization-insensitive lenses based on dielectric metasurfaces with meta-molecules. *Optica* **3**, 628–633 (2016).

[3] Lin, D. *et al.* Photonic multitasking interleaved Si nanoantenna phased array. *Nano Lett.* **16**, 7671–7676 (2016).

[4] Huang, P. S. & Zhang, S. Fast three-step phase-shifting algorithm. *Appl. Opt.* **45**, 5086–5091 (2006).

[5] Cogswell, C. J., Smith, N. I., Larkin, K. G. & Hariharan, P. Quantitative DIC microscopy using a geometric phase shifter. *Proc. SPIE* **2984**, 72–81 (1997).

[6] Meng, X. F. *et al.* Two-step phase-shifting interferometry and its application in image encryption. *Opt. Lett.* **31**, 1414–1416 (2006).

[7] Liu, J.-P. & Poon, T.-C. Two-step-only quadrature phase-shifting digital holography. *Opt. Lett.* **34**, 250–252 (2009).

[8] Zhang, Y, Tian, X & Liang, R. Random two-step phase shifting interferometry based on Lissajous ellipse fitting and least squares technologies. *Opt. Express* **26**, 15059–15071 (2018).

Metasurface	R (μm)	ΔD (μm)	Δs (μm)	a_1	a_2	a_3	a_4	a_5	$k_{grat,1}$ (rad/μm)
Layer1	300	660	1.5	-4.70×10^2	2.05×10^1	-3.88×10^0	7.69×10^{-1}	-7.30×10^{-2}	2.23

Metasurface	R (μm)	ΔD (μm)	Δs (μm)	b_1	b_2	b_3	b_4	b_5	$k_{grat,2}$ (rad/μm)
Layer2	300	660	1.5	-2.37×10^2	3.58×10^{-1}	1.07×10^1	-7.58×10^0	1.92×10^1	2.22

Table S1 Phase profile parameters for the two separate metasurface layers

Metasurface	R (μm)	ΔD (μm)	Δs (μm)	a_1	a_2	a_3	a_4	a_5	$k_{grat,1}$ (rad/μm)
Layer1	100	210	1.5	-1.29×10^2	7.96×10^0	-1.17×10^1	6.98×10^0	-1.34×10^0	2.21

Metasurface	R (μm)	ΔD (μm)	Δs (μm)	b_1	b_2	b_3	b_4	b_5	$k_{grat,2}$ (rad/μm)
Layer2	100	210	1.5	-7.88×10^1	-5.19×10^0	4.53×10^0	-1.43×10^0	1.61×10^{-1}	2.19

Table S2 Phase profile parameters for the double-sided metasurface QPGM
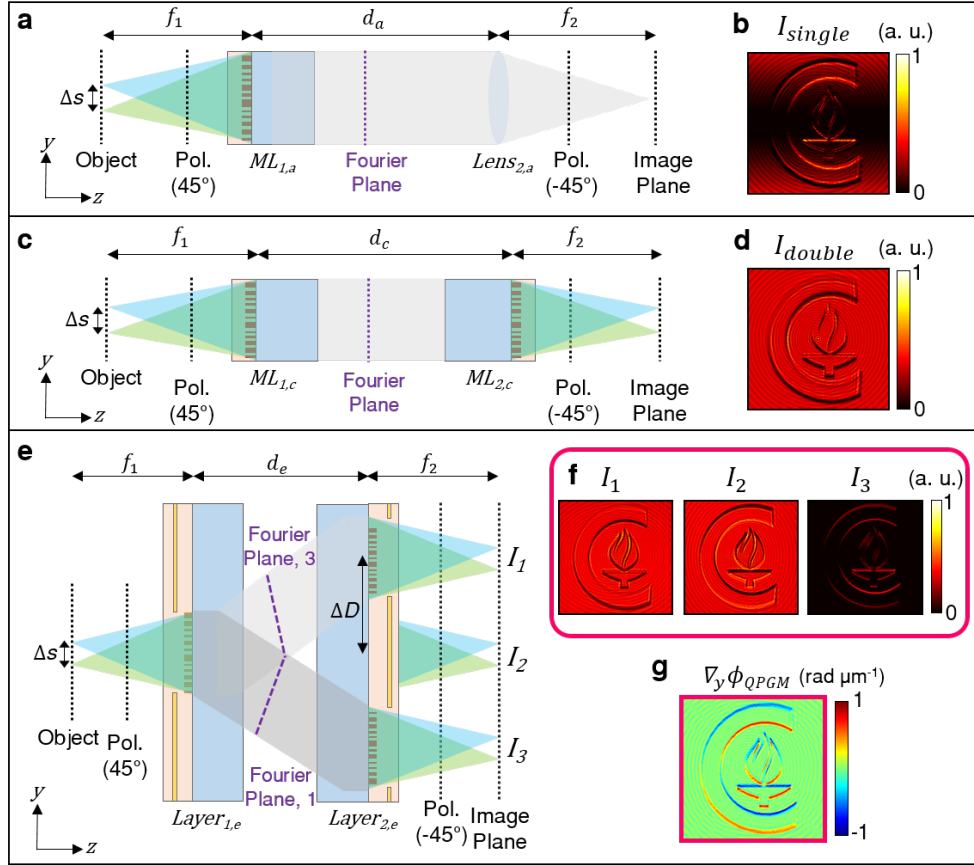
Figure S1 System-level design and numerical analysis of the QPGM. a Schematic of an optical system consisting of a single metasurface and a conventional thin lens. The metasurface works as a bifocal lens with two focal points for TE and TM polarizations separated along the *y* axis. **b** Simulated intensity map at the image plane formed by the system shown in **a**. **c** Schematic of an optical system consisting of two birefringent metasurfaces. Each metasurface acts as a bifocal lens with two focal points for TE and TM polarizations separated along the *y* axis. **d** Simulated intensity map formed in the image plane by the system shown in **c**. **e** Schematic of the optical system composed of one multi-functional metasurface and three bifocal lenses. As mentioned in Fig. **1b**, the first metasurface collimates light from two focal points for the TE and TM polarizations that are separated in the *y* direction and splits it in three directions towards the three metasurfaces on the second layer. Similar to the system in **c**, the first metasurface forms a separate imaging system with each of the metasurface lenses on layer 2. **f** Three simulated DIC images at the image plane using the system shown in **e**. **g** The PGI calculated from the three DIC images in **f**. Pol.: linear polarizer; ML: metasurface lens; d_a: 2.20 mm; d_c: 2.81 mm; d_e: 2.70 mm ; f_1: 687 μm; and f_2: 1.51 mm.
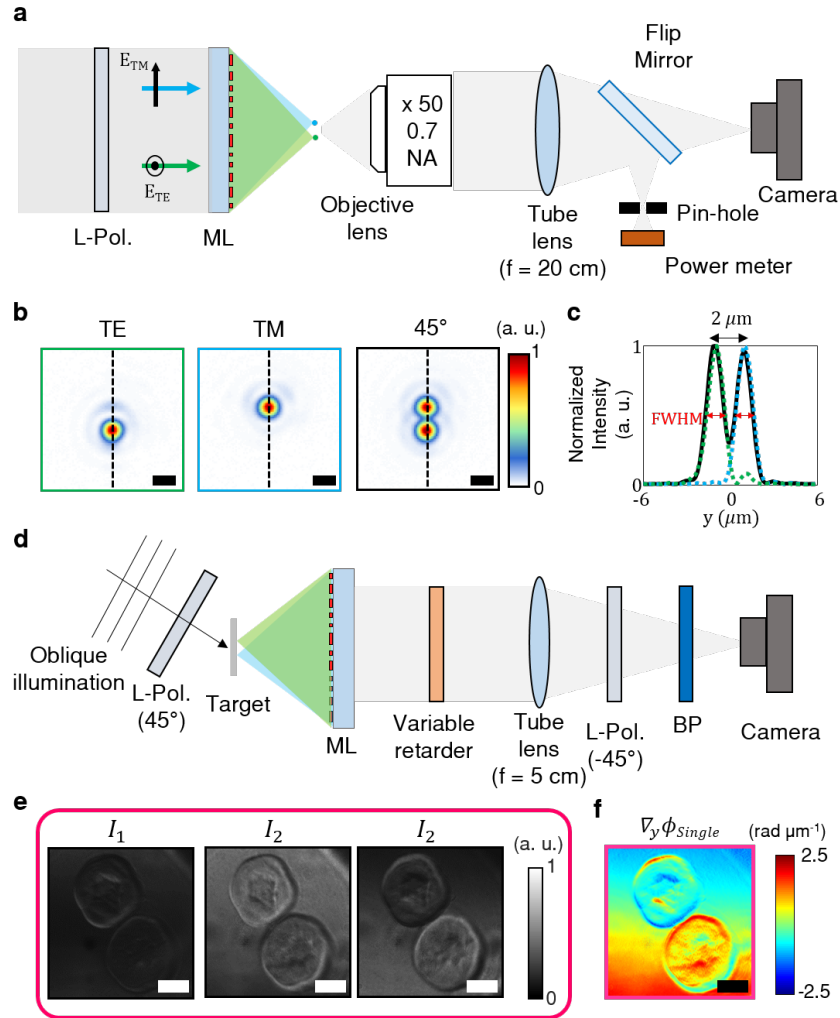
14

Figure S2 Phase gradient imaging with a single bifocal metasurface lens. a Schematic illustration of the optical setup used for capturing focuses of the single bifocal metasurface lens and measuring focusing efficiencies. The linear polarizer (L-Pol.) in front of the laser collimator is accordingly adjusted to confirm the input polarization states. **b** Measured focal points for different input polarization states. Left, center, and right images are the measured intensity maps at the focal plane with the linear polarizer aligned to 0° (TE), 90° (TM), and 45°, respectively. Scale bars: 2 μm. **c** Normalized intensity profiles at the focal plane on the three black dashed lines in **b**. Green and blue dashed lines are the intensity profiles for TE and TM polarized input light, respectively. The black line is the intensity profile with the 45° linear polarized light. **d** Schematic illustration of the optical setup capturing the three interference intensity patterns with a single bifocal metasurface and a variable retarder. An LED is used for the oblique illumination. BP: bandpass filter. **e** Measured interference intensity patterns, I_1, I_2, and I_3 that have phase offsets of 0, $\frac{2\pi}{3}$, and $\frac{4\pi}{3}$ between the TE and TM polarizations, respectively. **f** PGI calculated through the three-step phase shifting method from the three interference intensity maps in **e**. Scale bars: 30 μm.
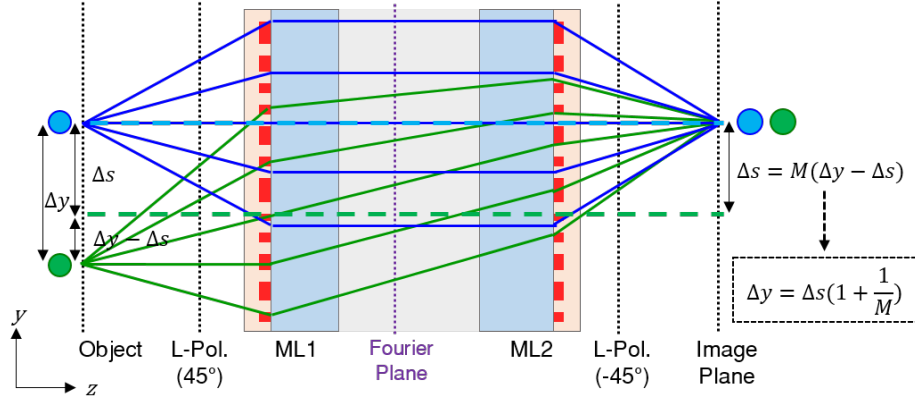
Figure S3 Schematic illustration showing the relation between Δs and Δy. Δs is the separation between the two optical axes for TE and TM polarizations. The green and blue dashed lines denote the optical axes for TE and TM polarizations, respectively. Δy is the effective shearing distance at the object plane. In other words, two points that are Δy apart along the y-axis in the object plane, are imaged to the same points at the image plane for the two different polarizations. The green and blue solid lines represent the rays coming from the green and blue points at the object plane, respectively. While the blue dot in the object plane is imaged along the blue dashed line representing the optical axis of TM polarization, the green dot is actually off-axis imaged at the position of the blue dot with respect to the green dashed line showing the optical axis of TE polarization. The black dashed box shows the relation between Δs and Δy given by $\Delta y = \Delta s(1 + \frac{1}{M})$, where M is the magnification of the optical system.
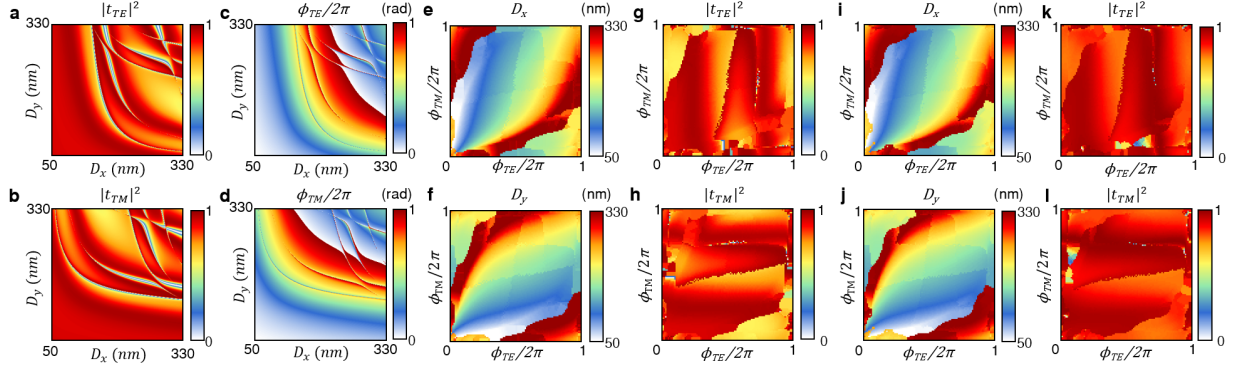
Figure S4 Simulation results of the nano-posts at the wavelength of 850 nm. a-d Simulated transmittance and transmitted phase of TE and TM polarized light for periodic arrays of meta-atoms as functions of D_x and D_y. The amorphous silicon layer is 664 nm thick, and the lattice constant is 380 nm. **a** and **b**: transmittance, **c** and **d**: transmitted phase **e-h** The optimized simulation results calculated from **a-d** for complete polarization and phase control. Calculated optimal D_x and D_y as functions of the required ϕ_{TE} and ϕ_{TM} are shown in **e** and **f**, respectively. Simulated transmittance of TE and TM polarized light as functions of ϕ_{TE} and ϕ_{TM} are shown in **g** and **h**, respectively. **i-l** The optimized simulation results of the nano-posts composed of 664-nm thick amorphous silicon and 60-nm thick Al_2O_3 on the fused silica substrate. The nano-posts are arranged on a square lattice with a 390-nm lattice constant and cladded by an $\sim$8-μm-thick SU-8 layer. Calculated optimal D_x and D_y as functions of the required ϕ_{TE} and ϕ_{TM} are shown in **i** and **j**, respectively. Simulated transmittance of TE and TM polarized light as functions of ϕ_{TE} and ϕ_{TM} are shown in **k** and **l**, respectively.
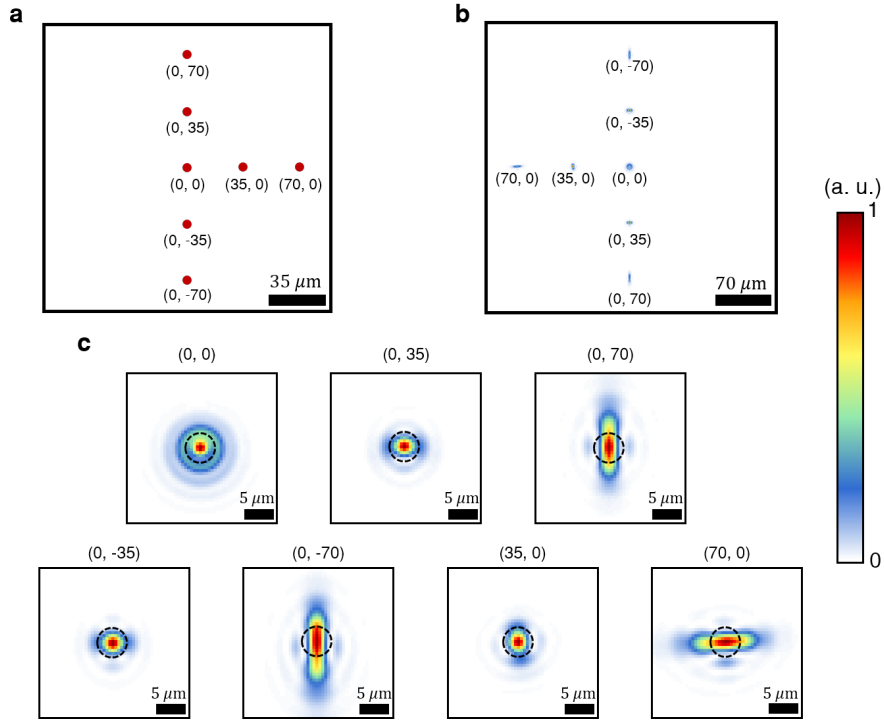
Figure S5 Simulated point spread functions of the QPGM. a Schematic showing the locations of point sources used at the object plane for characterizing the point spread functions (PSFs). Cartesian coordinates of each point are given under it. The coordinates are calculated from the center of the metasurface layer 1. **b** The normalized PSFs at the image plane for TE polarization of the optical system shown in Fig. **S1e**. The coordinates under the PSFs are the coordinates of the corresponding point sources. It is assumed that the PSFs are identical for TE and TM polarizations. **c** The magnified intensity maps of the normalized PSFs in **b**. The coordinates of the corresponding point sources are shown above the intensity maps. The dashed black circles are airy disks whose radii are 2.54 μm at the image plane. All coordinates in the figures are given in microns.
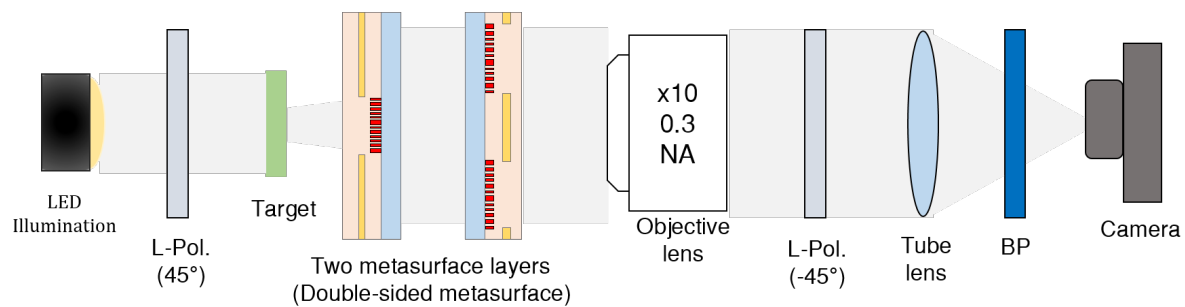
Figure S6 Schematic of the measurement setup. Schematics of the custom-built microscope setup used to measure the three DIC images captured by the QPGMs. L-Pol.: linear polarizer; BP: band-pass filter.
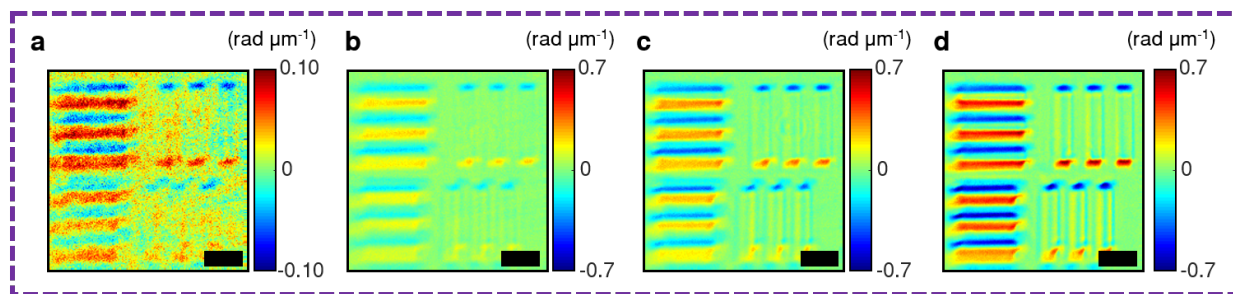
Figure S7 Magnified PGIs of the phase resolution targets having different thicknesses. The thicknesses of the targets are as follows; **a**: 54 nm; **b**: 159 nm; **c**: 261 nm; and **d**: 371 nm. Note that the color bar scale in **a** is different from the other panels. Scale bars: 15 μm.
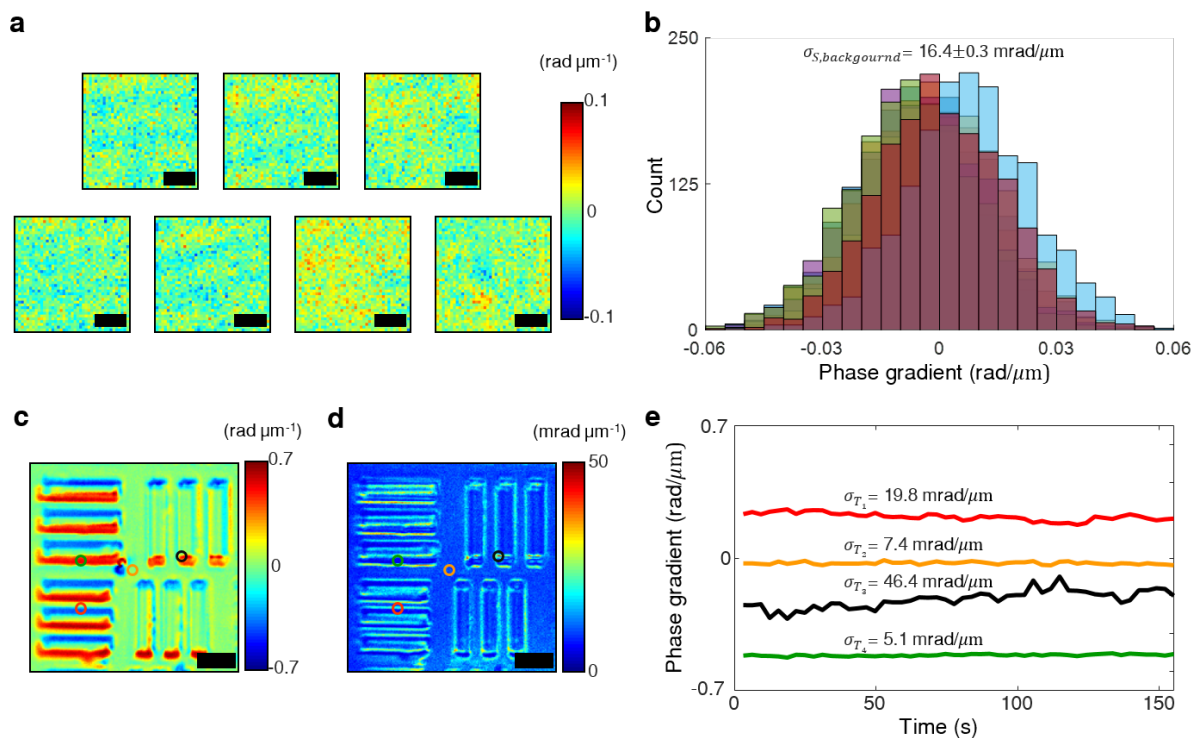
Figure S8 Measurement of spatial and temporal noise levels. a Background phase gradient images captured in an area with dimensions of 17.6×17.6 μm^2 in the central part of the seven different measurements with the seven different phase resolution targets used in Fig. 3d. Scale bars: 5 μm. **b** Phase gradient histograms of the phase gradient images in **a**. Each image in a has 1681 points. The spatial noise level is 16.4±0.3 mrad/μm calculated by measuring the standard deviation of the phase gradients in **b**. **c** and **d** The 371-nm thick resolution target used in Fig. S7d is measured 50 times during 155 seconds. **c** The first phase gradient image at t = 3.1s. **d** Map of the standard deviations over 50 frames. The average standard deviation over the map is 11.4 mrad/μm. **e** Plots of the temporal phase gradient signals at the different points marked with red, orange, black and green circles in **c** and **d**. The colors of the plots are matched with the colors of the circles in **c** and **d**. The corresponding standard variations are noted over the plots. Especially, the black line shows the maximum fluctuation having a standard deviation of 46.4 mrad/μm in the map.
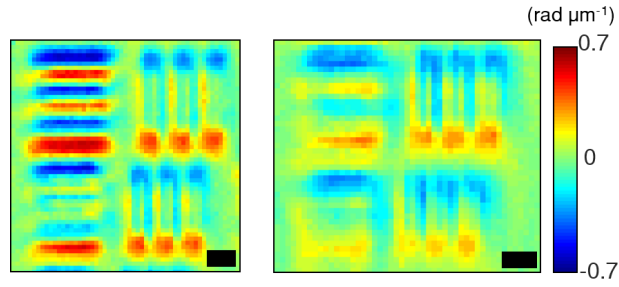
21

Figure S9 Investigation of the lateral resolution. PGIs of the smallest resolution groups that the QPGM based on the two separate metasurface layers was able to resolve. The measured lateral resolutions along the *x* and *y* directions are 2.76 μm and 3.48 μm, respectively. The thickness of the resolution target is 374 nm. Left: elements 2 and 3 in group 8; Right: elements 4 and 5 in group 8. Scale bars: 2 μm.
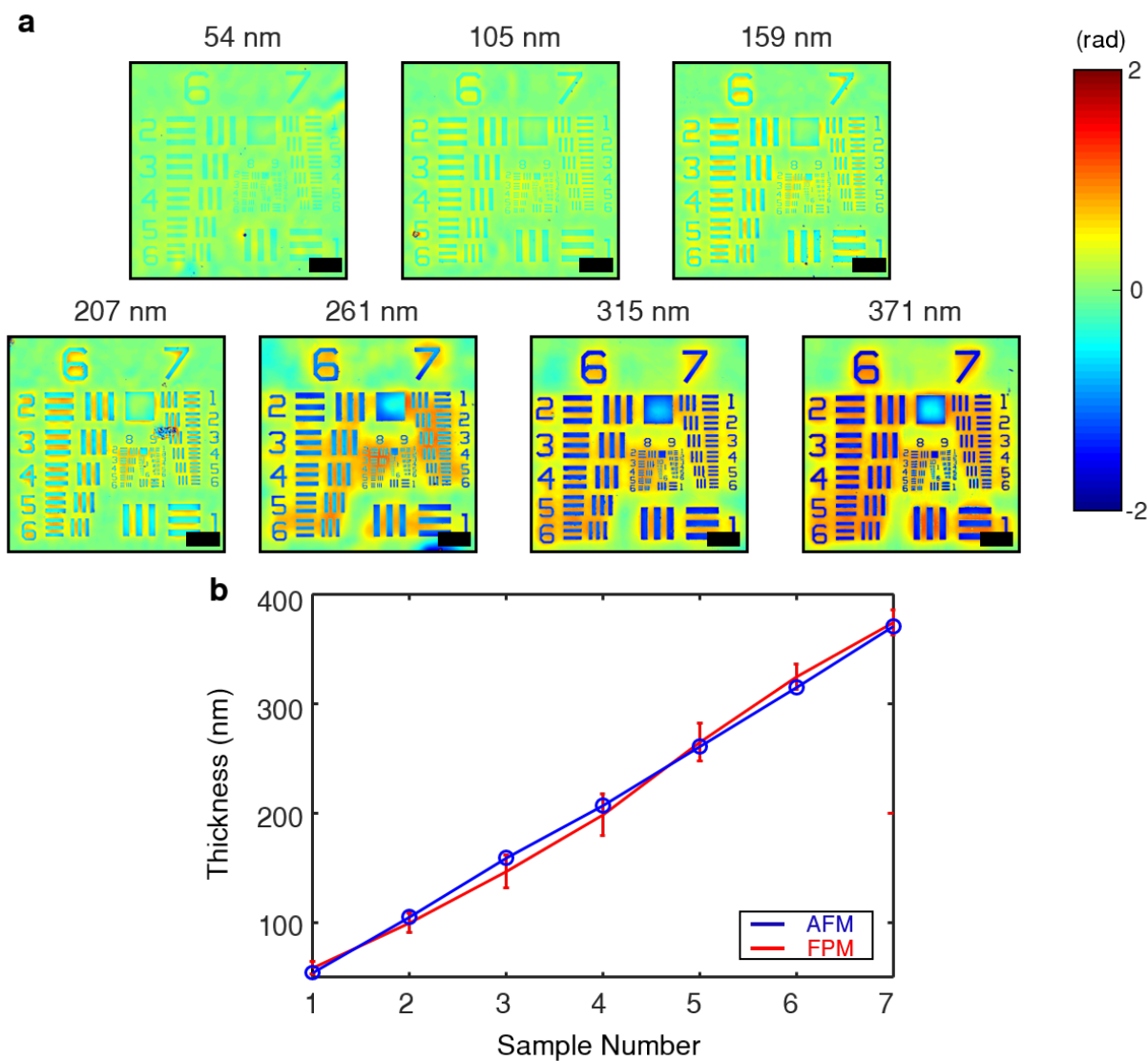
Figure S10 Phase maps of the USAF 1951 phase resolution targets measured by Fourier ptychography. a Measured phase maps of the seven USAF 1951 phase resolution targets through the Fourier ptychography method (FPM). An array of green LEDs with an operation wavelength of 522 nm was utilized for Fourier ptychography. Scale bars: 40 μm. **b** Thicknesses of seven different parts of the phase target calculated from the phase maps captured by Fourier ptychography (red), and those measured by AFM (blue). The plotted estimated height values through Fourier ptychography are averaged over 100 arbitrarily chosen points. Error bars represent standard deviations of the estimated values.
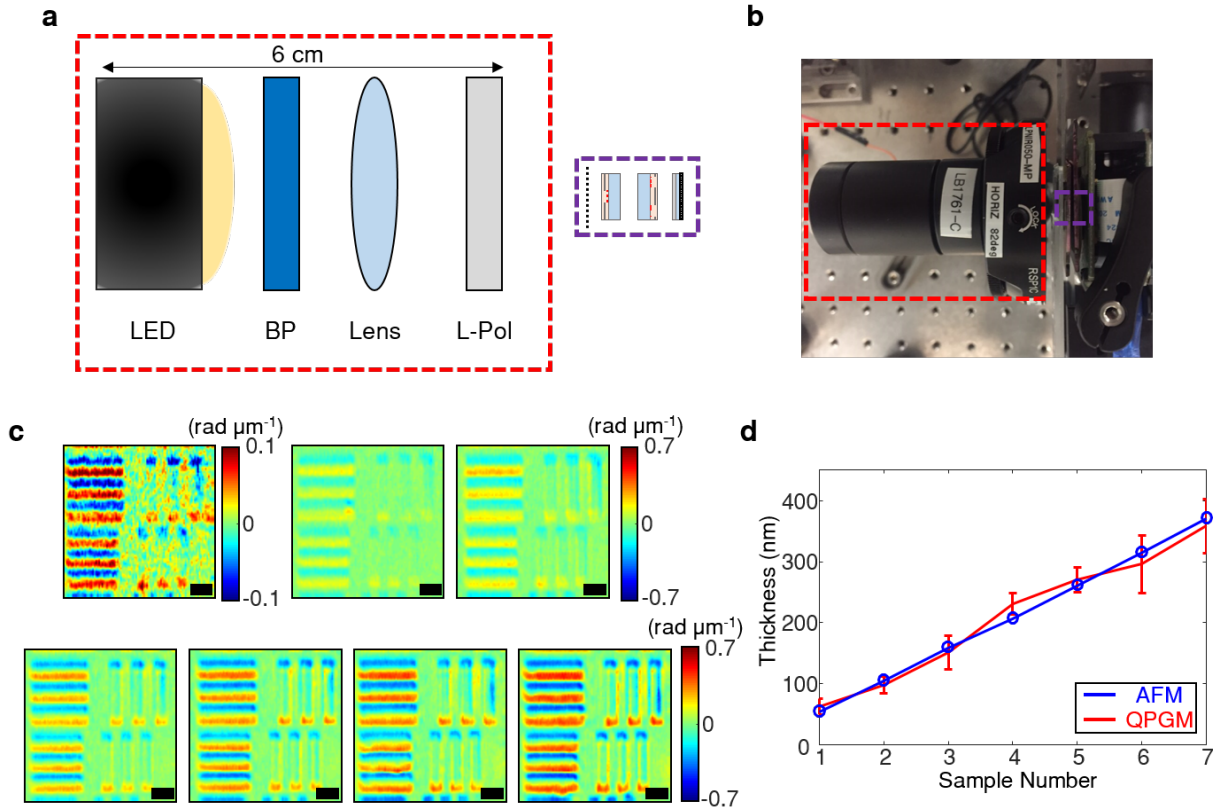
Figure S11 Compact QPGM setup and measurement results using a CMOS image sensor. a Schematics of the custom-built miniaturized microscope setup using a CMOS image sensor. The red and purple dashed boxes illustrate the illumination and detection parts of the setup, respectively. The magnified schematic of the detection part is shown in Fig. **3f** in the main manuscript. L-Pol.: linear polarizer; BP: band-pass filter. **b** Optical image of the setup in **a**. The illumination and detection parts are represented by the red and purple dashed boxes, respectively. **c** Phase gradient images captured for the same parts of the phase target used in Figs. 3c and S7. The thicknesses of the three targets from top left to top right and the four targets from bottom left to bottom right are as follows: Top: 54 nm; 105nm; and 159 nm. Bottom: 207 nm; 261 nm; 314 nm; and 371 nm. Note that the color bar scale for the thinnest sample is different from the other panels. Scale bars: 15 μm. **d** Thicknesses of the phase targets calculated from the PGIs in **c** and those measured by AFM. The plotted estimated thickness values through QPGM are averaged over 100 arbitrarily chosen points at the sample edges. Error bars represent standard deviations of the estimated values.
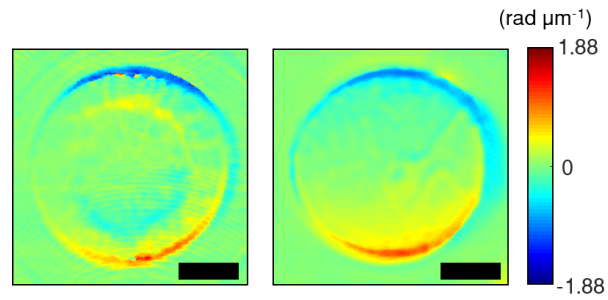
24

Figure S12 PGIs of the sea urchin samples captured by the doublet QPGM. PGIs of two sea urchin samples measured using the double-sided metasurface QPGM shown in Fig. **4b** of the main text. Scale bars: 25 μm.
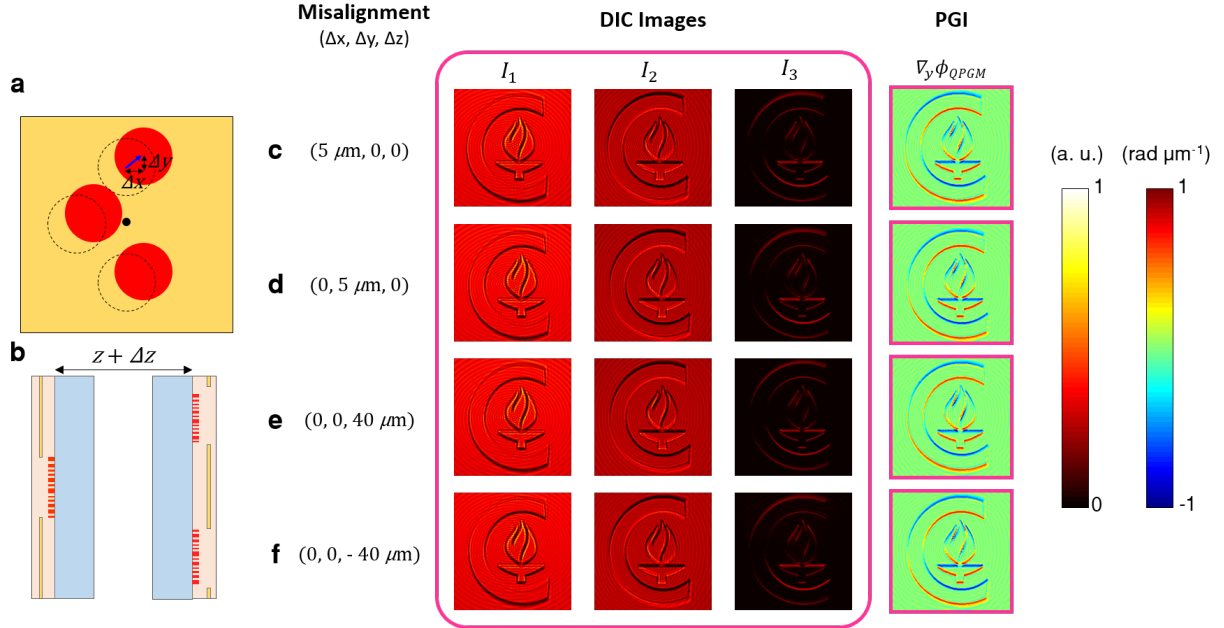
Figure S13 Numerical investigation of the effects of misalignment between the two metasurface layers. a, Schematic illustration of the the top view of the metasurface layer 2. The lateral misalignments along x and y directions, Δx and Δy, are shown. **b** Schematic illustration of the the side view of the QPGM. The axial misalignment along the z axis, Δz, is shown. We assumed that the thicknesses of the two substrates is identically changed by $\frac{\Delta z}{2}$. **c-f** Simulated DIC images and PGIs for four different types of misalignment. In simulation, the QPGM is based on the system shown in Fig. **S1e** with the phase profile parameters given in Table **S1**. Left: misalignment vectors, $(\Delta x, \Delta y, \Delta z)$. Center: three different DIC images at the image plane corresponding to misalignment vectors shown on the left. Right: PGIs calculated from the DIC images at the center. For calibration, we subtract the PGI calculated in the absence of the sample with each misalignment from the PGI calculated from the three DIC images at the center.