Supplementary Information for On-Chip Multidimensional Control of Twisted Moiré Photonic Crystal for Adaptive Sensing and Imaging

Haoning Tang ⊥,[1,*] Beicheng Lou ⊥,[2] Fan Du,[1] Guangqi Gao,[1] Mingjie Zhang,[1] Xueqi Ni,[1] Evelyn Hu,[1] Amir Yacoby,[3] Yuan Cao,[3,4,†] Shanhui Fan,[2,‡] and Eric Mazur[1,§]

[1]*School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA*

[2]*Department of Applied Physics and Ginzton Laboratory, Stanford University, Stanford, CA 94305, USA*

[3]*Department of Physics, Faculty of Art and Sciences, Harvard University, Cambridge, MA 02138, USA*

[4]*Department of Electrical Engineering and Computer Science, University of California at Berkeley, Berkeley, CA 94720, USA*

CONTENTS

1. SAMPLE FABRICATION

The MEMS-optics chip is fabricated from silicon-on-insulator wafers, and is robust against heat, refrigeration, or magnetic fields. The fabrication procedure is outlined in Extended Data Fig.1. The main processing steps are

- Vias drilling and coating
- Photonic crystal (PhC) fabrication
- MEMS process
- Vapor releasing
- Top layer fabrication and bonding

A. Vias drilling and coating

In this fabrication, a 6-inch Silicon-on-insulator (SOI) wafer was used. The wafer contains a 500µm handle layer, a 2µm buried oxide (BOX) layer, and a 40µm Si device layer. Each wafer can accommodates 208 device dies. The Through-Silicon Via (TSV) provides electrical connectivity, grounding both the Silicon device and handle layers. As shown in Extended Data Fig. 1a, after a post-cleaning, a 50nm oxide layer is deposited on the wafer using low-pressure chemical vapor deposition (LPCVD). The TSV holes and global markers are patterned on the frontside using photolithography. Subsequent steps involves the etching of the 50nm oxide layer, silicon device layer, and BOX layer using reactive ion etching (RIE) and deep reactive ion etching (DRIE). A poly-silicon conductive layer is then added to the device frontside, covering the TSV holes. The TSV hole inversion is patterned using the photolithography, followed by a RIE process to remove excess conductive layer and the thin oxide layer.

B. Photonic crystal (PhC) fabrication

The PhC is integrated with MEMS, as shown in Extended Data Fig.1b, the PhC is an Al_2O_3(25 nm)-Si_3N_4(410 nm)-Al_2O_3(25 nm) layered structure. To fabricate the PhC layer, both sides of the wafer are first deposited by a 25 nm Al_2O_3 layer via atomic layer deposition (ALD), the front side of the wafer is then patterned into disks using photolithography and RIE, followed by a rapid thermal processing (RTP) annealing process. After cleaning the wafer, a 410nm Si_3N_4 layer is deposited on both sides using LPCVD and patterned into disks. The Ebeam-lithography and RIE etching is then used to pattern a 310nm thick PhC on the Si_3N_4 disk. Another 25 nm Al_2O_3 layer is deposited on both wafer sides, patterned into disks. The wafer is then annealed one more time via RTP.

C. MEMS process

The MEMS fabrication includes three primary steps: front side patterning, backside patterning, and Hydrofluoric acid (HF) release. First, an Al_2O_3 hard mask is deposited onto the frontside of the wafer, followed by MEMS actuator patterning using photolithography. This hard mask is etched using RIE, after which the silicon device layer is processed with DRIE (see Extended Data Fig.1c).

The purpose of backside processing is to suspend both the MEMS and the PhC. We first use RIE to remove the Al_2O_3, Si_3N_4, and SiO_2 layers. The photolithography is used for the backside patterning. The silicon handle layer is then etched via DRIE, followed by a BOX layer etching via RIE and another DRIE process for the silicon device layer (see Extended Data Fig.1d).

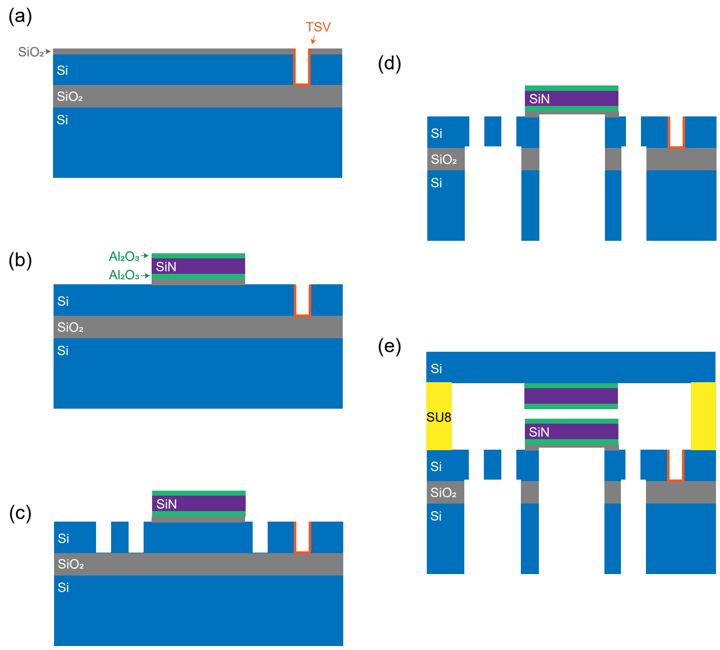
D. Vapor releasing

Before the HF releasing process, the wafer is thoroughly cleaned and dried in oxide plasma. The BOX layer is then etched via HF vapor etcher, the wafer is also segmented into small dies in this process.

E. Top layer fabrication and releasing

The top layer PhC is fabricated on a silicon wafer (see Extended Data Fig.1e). A 50 m oxide layer is first deposited to both wafer sides using LPCVD. This layer is subsequently patterned into small disks via photolithography and RIE. Following this, a 25 nm Al_2O_3 layer is deposited to both side of the wafer through ALD and then patterned into disks through photolithography and RIE. This Al_2O_3 layer is again annealed using RTP.

After a standard wafer cleaning process, a 410nm Si_3N_4 layer is deposited to both wafer sides using LPCVD and then patterned into disks. A PhC with a thickness of 310nm is etched onto the

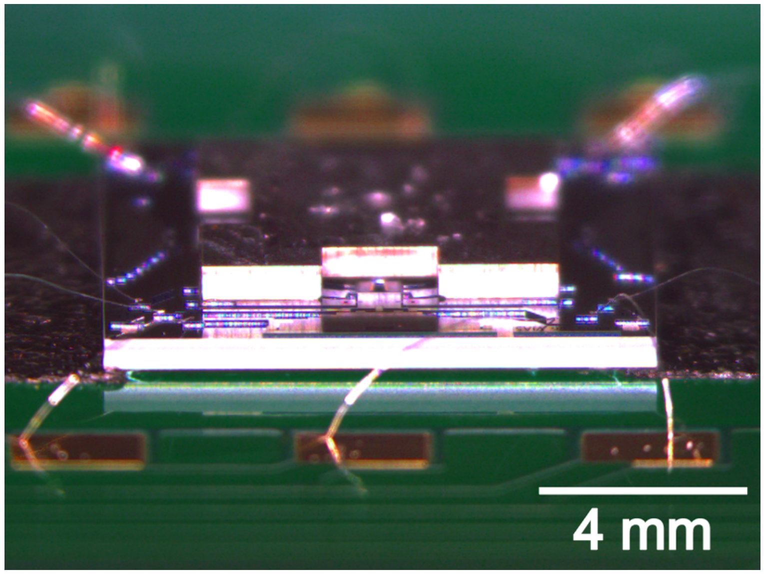


Extended Data Figure 1: **Illustration fabrication process.** (a) Vias drilling and coating. (b)Photonic crystal (PhC) fabrication. (c)MEMS process. (d)Vapor releasing. (e)Top layer fabrication and bonding. Blue: Silicon. Gray: Silicon oxide. Purple: Silicon Nitride. Orange: Conductive layer. Yellow: SU8.

Si_3N_4 disk using e-beam lithography and RIE. Another 25 nm Al_2O_3 layer is deposited on both wafer sides and patterned into disks. The wafer undergoes another RTP annealing process.

Then, a 3.8um thick SU8 adhesion layer is patterned onto the wafer through photolithograph. Areas excluding the PhC and SU8 are patterned and etched down by 20um through DRIE. The backside of the wafer is patterned and etched by DRIE. The silicon wafer is then segmented into individual dies in this step. Residual oxide layer is removed by immersing dies in Buffered HF.

The MEMS-optics layer is aligned and bonded to the top layer by adding heat and pressure. The final device is electrically connected to the printed circuit board through wire bonding (see Extended Data Fig.2).



Extended Data Figure 2: **The sideview of packaged MEMS-TMPhC device wire bonded to printed circuit broad**

2. MEMS WORKING PRINCIPLE

This section explains the functional dynamics of the cascade MEMS-optics device. MEMS actuators, widely utilized in both academia and industry, can be categorized into electrostatic, piezoelectric, and thermal types, each presenting distinct pros and cons. Specifically, electrostatic actuators stand out, delivering swift responses, zero backlash (hysteresis), and adaptability to a range of environments, including ultra-low temperatures and high magnetic fields. In comparison, piezoelectric actuators often show limited deformation (despite possible mechanical enhancement), noticeable hysteresis, and dimensional alterations due to environmental changes. Thus, electrostatic actuators are preferable for integrating with nanophotonics and addressing the rising demands in quantum-optics applications.

The basic operation principle of electrostatic actuators relies on a parallel plate capacitor. The energy of a parallel plate capacitor in vacuum is

$$E = \frac{1}{2}CV^2 = \frac{\varepsilon_0 A V^2}{2d}, \tag{1}$$

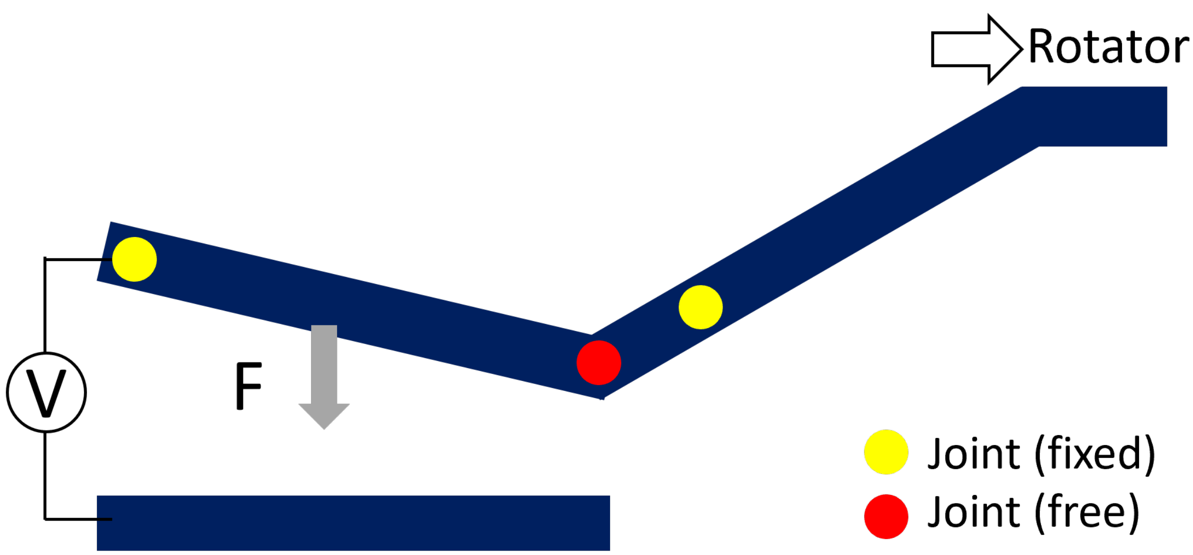
where C is capacitance, V is voltage on the capacitor, ε_0 is vacuum permittivity, A is the area of overlap and d is the distance between the plates. If a capacitor has a fixed plate and a movable plate that can displace perpendicularly (towards the other plate, which we call z direction) as well as parallely (which we call x direction), the electrostatic forces are

$$f_\perp = \frac{\mathrm{d}E}{\mathrm{d}z} = -\frac{\varepsilon_0 A V^2}{2d^2}, \tag{2}$$

$$f_\parallel = \frac{\mathrm{d}E}{\mathrm{d}x} = \frac{\varepsilon_0 l V^2}{2d}, \tag{3}$$

where l is the extent of the capacitor in y direction. All MEMS electrostatic actuators essentially use one of these two forces to balance the forces from a MEMS spring (simply modeled by the Hooke's law $f = k\Delta x$) to achieve a displacement $\Delta x = f/k$. Notice that the vertical actuation force often surpasses the parallel actuation force because the capacitor size generally exceeds the distance between the plates. However, a limiting factor is the pull-in instability inherent in these actuators, which typically restricts the maximum actuation to approximately 1/3 of the initial plate separation.

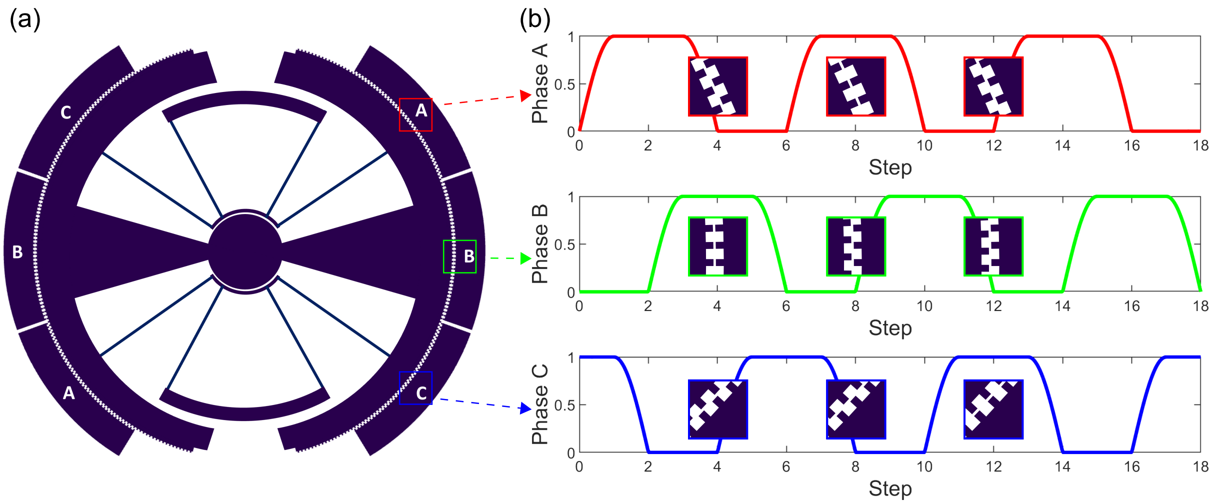
The mechanical part of our cascade MEMS-optics device consists of a rotary (R) actuator cascaded on the vertical (V) actuator. The V actuator uses the perpendicular force Eq. 2. As previously highlighted, such actuators can generate substantial forces but possess a restricted range. Consequently, a lever-based mechanism is utilized to magnify the motion (See extended data Fig. 3). The resulting actuator can provide up to 4.5 μm of motion when actuated to ~ 50 V.



Extended Data Figure 3: **Illustration of vertical actuator driving principle.** The vertical actuator is consist of a parallel capacitor that provides a driving force downwards, and a lever that magnifies the vertical translation for 4 times in the reversed direction.

In contrast, the R actuator derives its actuation force from Eq. 3. Given that this force is relatively smaller, it becomes advantageous to fragment the capacitor into numerous smaller capacitors until limitations from edge effects manifest. Following this design guideline, one typically arrives at either a 'comb' configuration or a 'teeth' arrangement. The latter proves to be more apt for rotary motion since actuator teeth can be designed around the rotary core, forming a stepper motor without compromising the balance between force and range. The R actuator therefore operates based on a three-phase micro-stepping motor mechanism, driven by energizing rotor and stator poles. The device employs a flexure mechanism that comprises six-beam pivot structures. Attached to the rotor, this six-beam pivot structure allows for a maximal angular rotation of 10° (±5°). The configuration includes three pairs of stators and a grounded rotor. Each stator phase pair is positioned diametrically opposite its counterpart to mitigate pull-in instability. The poles in phase-B align precisely with the rotor poles, while those in phases A and C misalign by 1/3 and -1/3 of the pitch, respectively. Each stator phase can be individually energized via a voltage waveform to attain the desired rotation (See Extended Data Fig.4 for the Illustration of stepper motor driving principle). Energizing the three phases induces an electrostatic force between the rotor and stator poles, prompting the rotor poles to re-align with the stator poles, neutralizing the generated electric field. Subsequently, the adjacent phase is activated, causing the rotor to once again align its poles with the relevant stator poles. Thus, at any given moment, only one-third of the rotor poles synchronize with the stator poles. Micro-stepping voltage can be employed for a finer rotary tuning. Micro-stepping enables a gradual voltage transition, ensuring the motor operates smoothly and achieves superior positioning resolution. The precision of the rotary angle control is 1/6 degree.

To cascade the R actuator on top of the V actuator, we designed an inter-stage electrical network with viases to connect the voltage necessary for the R stage. Proper design of this network ensures that there is minimal crosstalk between the R and V actuator.



Extended Data Figure 4: **Illustration of rotary three-phase stepper driving principle.** (a) Illustration of the rotary actuator with three-phase electrode. (b) Driving micro-stepping voltage waveform for driving rotary actuator

3. VAN DER WAALS FORCE BETWEEN TWO FLAT SURFACES

As for the bilayer photonic crystals in this paper, the Van der Waals force (per area) between two flat surfaces is given by[1]

$$F = -A/6\pi D^3 \tag{4}$$

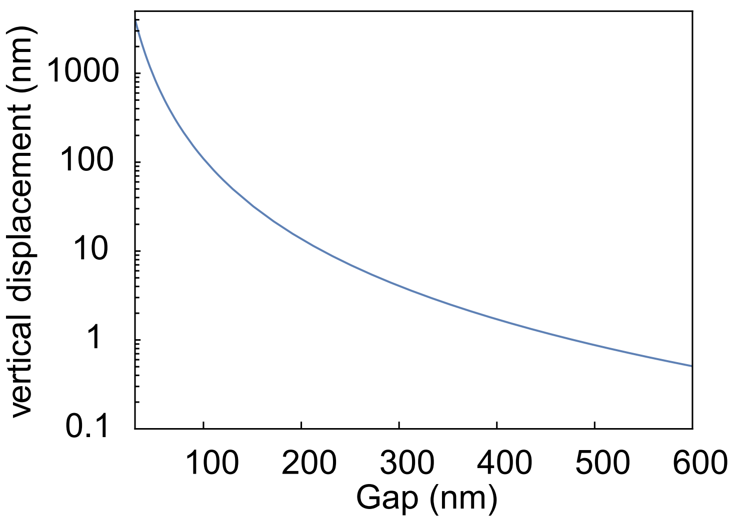
where D is the gap between them, and A is the Hamaker Constant. For two macroscopic medium 1 and 2 interacting across a medium 3, assuming all three media have the same absorption frequencies, we can derive an estimated formula for the nonretarded Hamaker constant as

$$\begin{aligned} A_{\text{total}} &= A_{v=0} + A_{v>0} \\ &\approx \frac{3}{4}kT\left(\frac{\varepsilon_1-\varepsilon_3}{\varepsilon_1+\varepsilon_3}\right)\left(\frac{\varepsilon_2-\varepsilon_3}{\varepsilon_2+\varepsilon_3}\right) \\ &+ \frac{3hv_e}{8\sqrt{2}}\frac{\left(n_1^2-n_3^2\right)\left(n_2^2-n_3^2\right)}{\left(n_1^2+n_3^2\right)^{1/2}\left(n_2^2+n_3^2\right)^{1/2}\left\{\left(n_1^2+n_3^2\right)^{1/2}+\left(n_2^2+n_3^2\right)^{1/2}\right\}}, \end{aligned} \tag{5}$$

where k is Boltzmann's constant, T is the temperature, h is Planck's constant, and $\varepsilon_1(n_1), \varepsilon_2(n_2)$, and $\varepsilon_3(n_3)$ are the permittivities (refractive indices) of medium 1, medium 2, and medium 3. In our case, medium 1 and 2 are the same, therefore this expression reduces to

$$A = \frac{3}{4}kT\left(\frac{\varepsilon_1-\varepsilon_3}{\varepsilon_1+\varepsilon_3}\right)^2 + \frac{3hv_e}{16\sqrt{2}}\frac{\left(n_1^2-n_3^2\right)^2}{\left(n_1^2+n_3^2\right)^{3/2}} \tag{6}$$

The effective attractive force per area is $F_{\text{eff}} = ff \times F$, where ff is the filling factor of dielectrics versus air in the photonic crystal. Consider a circular membrane photonic crystal chip with 502 nm radius and 1220 nm pitch. The maximum vertical displacement and effective area force versus the interlayer gap are shown in Fig. S5.



Extended Data Figure 5: **Calculated maximum vertical displacement versus interlayer gap.**

The minimum interlayer gap in our experiment is 300 nm, the Van der Waals attractive force is very small and will not cause the membranes to drop or stick together.

4. COMPARISON OF MEMS FLAT-OPTICS DEVICES

TABLE I: Comparison of different MEMS flat-optics device

Optics	Actuator type	Optical Material	DoFs	Cascaded (Multilayers)
Gratings[2]	Electrostatic	SiN	1(Vertical)	No
Metasurface alvarez lens[3]	Electrostatic	SiN	1(translational)	No
Metalens[4]	Elastomer	Si	1(translational)	No
Metalens[5]	Electrostatic	Au	1(tilt)	No
Metasurfaces[6]	Piezoelectric	Au	1(Vertical)	No
Metasurfaces[7]	Electrostatic	Si	1(switch)	No
Metasurfaces[8]	Electrostatic	Si	1(Vertical)	No
Metasurfaces[9]	Piezoelectric	Au	1(Vertical)	No
Metasurfaces[10]	Electrostatic	Al	1(translational)	No
Metamaterial[11]	Electrostatic	SiN	1(translational)	No
Metamaterial[12]	Electrostatic	SiN	1(tilt)	No
Metamaterial[13]	Thermal	Au, SiN	1(bend)	No
Metamaterial[14]	Electrostatic	Al	1(bend)	No
Metalens[15]	Electrostatic	SiN	1(Vertical)	Yes
Moiré optics (this work)	Electrostatic	SiN	3(rotational, translational, tilt)	Yes

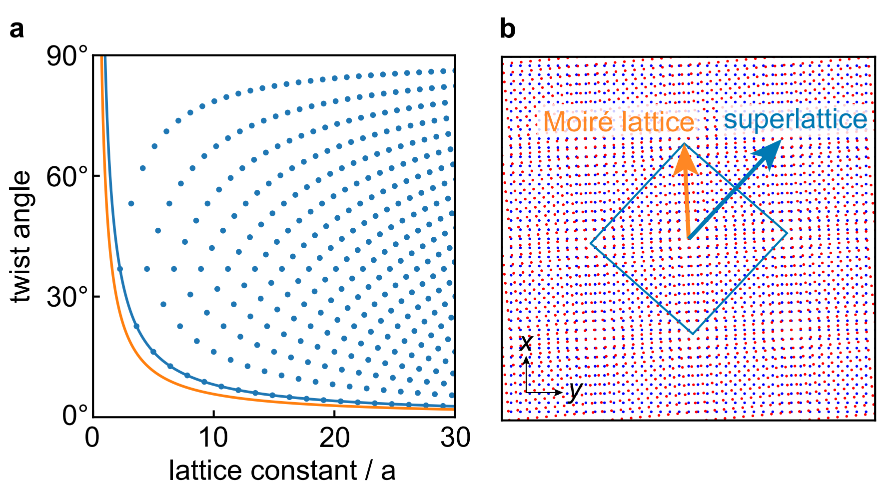
5. MOIRÉ SCATTERING

The moiré lattice is formed by two layers of identical 2D lattices with unit vectors $(\mathbf{a}_1, \mathbf{a}_2)$ and $(\mathbf{a}'_1, \mathbf{a}'_2)$ twisted against each other. As for the square lattice where $|\mathbf{a}_1| = |\mathbf{a}_2| = a$. At the commensurate angle there are integer pairs (m, n) and (m', n') satisfy

$$m^2 + n^2 = m'^2 + n'^2. \tag{7}$$

where m and n $(m < n)$ are two different positive integers, which generate Pythagorean set (A, B, C) by $A = n^2 - m^2, B = 2mn, C = m^2 + n^2$. The geometric meaning of (m, n) is shown in Extended Data Fig. 6a. The lattice constant of superlattice is $\sqrt{C}a$. The twist angle θ can also be read out as $\theta = \arctan\frac{n}{m} - \arctan\frac{m}{n}$. The smallest superlattice within the neighborhood of each twist angle comes from the generators satisfying $n = m + 1$. This series of generators forms the blue curve in Extended Data Fig. 6a.

When two identical square lattices are twisted, a expansive repetition pattern called the moiré lattice can be seen. Extended Data Fig. 6a shows the cell of a superlattice with $(m, n) = (10, 11)$ by the blue square. The first-order moiré wavevectors are $\mathbf{G}_{\mathrm{m},(\pm 1,0)}$ and $\mathbf{G}_{\mathrm{m},(0,\pm 1)}$, the first-order moiré lattice constant satisfies $a_{\mathrm{m}} = a/(2\sin(\theta/2))$, which is indicated by the orange arrow in Extended Data Fig. 6a. The first-order moiré lattice exhibits quasi-crystalline properties.



Extended Data Figure 6: **The superlattice and moiré lattice of twisted bilayer square lattices.** **a**, twist angle versus the lattice constant of the superlattice (in blue dots) and the moiré lattice (in orange). **b**, the smallest superlattice and the first order moiré lattice at $\theta = 5.453°$.

The moiré scattering process can be understood by analyzing the wavevectors from each layer. A PhC slab is a dielectric structure that is finite in the z-direction and periodic in $x - y$ plane.

This periodicity can be depicted through a lattice in either real space or reciprocal space. A square lattice in the reciprocal space is a set of points $\mathcal{G}_1 \equiv \{(i\hat{\mathbf{x}} + j\hat{\mathbf{y}})2\pi/a \mid i, j \in \mathbb{Z}\}$, where $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ are the unit vectors in the xy plane, and a is its period in real space. When two identical square lattices are in close proximity ($h < \Lambda$), the interlayer coupling strengthens, causing the electromagnetic waves in each layer couple with each others. Additionally, when the lattices are twisted relative to each other, they generate moiré lattices in real space. The reciprocal lattice of a moiré PhC features a broad spectrum of wavevectors, resulting from scattering effects generated by the periodic structures in each layer. The lattice of the twisted layer with a twist angle θ in the reciprocal space would be $\mathcal{G}_2 \equiv \{(i\hat{\mathbf{x}}' + j\hat{\mathbf{y}}')\, 2\pi/\theta \mid i, j \in \mathbb{Z}\}$, where $\hat{\mathbf{x}}'$, $\hat{\mathbf{y}}'$ are the unit vectors in the twisted xy plane, namely $\hat{\mathbf{x}}' = \cos(\theta)\hat{\mathbf{x}} + \sin(\theta)\hat{\mathbf{y}}$, $\hat{\mathbf{y}}' = \cos(\theta)\hat{\mathbf{y}} - \sin(\theta)\hat{\mathbf{x}}$. The reciprocal lattice for the moiré is the sum of the reciprocal lattices for each layer, i.e.

$$\begin{aligned} \mathcal{G}_{\mathrm{TB}} &\equiv \left\{\left(i\hat{\mathbf{x}} + j\hat{\mathbf{y}} + k\hat{\mathbf{x}}' + l\hat{\mathbf{y}}'\right) 2\pi/a \mid i, j, k, l \in \mathbb{Z}\right\} \\ &\equiv \left\{\mathbf{G}_{(i,j,k,l)} \mid i, j, k, l \in \mathbb{Z}\right\}, \end{aligned} \tag{8}$$

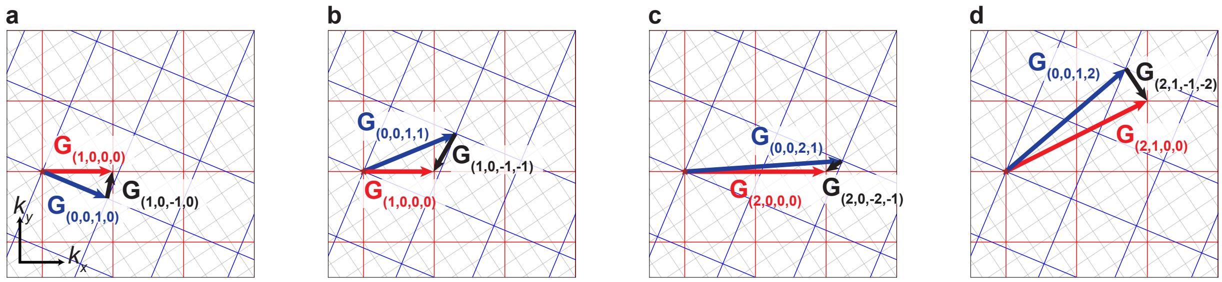
There are also additional subset of wavevectors needs to be considered for light scatters in the moiré structure. i.e.

$$\begin{aligned} \mathcal{G}_{\mathrm{m}} &\equiv \left\{\left(i\hat{\mathbf{x}} + j\hat{\mathbf{y}} - i\hat{\mathbf{x}}' - j\hat{\mathbf{y}}'\right) 2\pi/a \mid i, j \in \mathbb{Z}\right\} \\ &= \left\{\mathbf{G}_{(i,j,-i,-j)} \mid i, j \in \mathbb{Z}\right\} \\ &\equiv \left\{\mathbf{G}_{\mathrm{m},(i,j)} \mid i, j \in \mathbb{Z}\right\}, \end{aligned} \tag{9}$$

where $\mathbf{G}_{\mathrm{m},(i,j)}$ are moiré wavevectors. The properties of the moiré that are strongly dependent on the twist angle arise from a scattering process in which incident light with in-plane wavevector $\mathbf{k}_{\mathrm{inc}}$ is scattered by a moiré wavevector $\mathbf{G}_{\mathrm{m},(i,j)}$. Such scattering process gives rise to resonances when $\mathbf{k} = \mathbf{k}_{\mathrm{inc}} + \mathbf{G}_{\mathrm{m},(i,j)}$ matches with the wavevector of the resonant modes of the system.

Moiré wavevectors involved are of the form $\mathbf{G}_{(i,j,k,l)} = (i\hat{\mathbf{x}} + j\hat{\mathbf{y}} + k\hat{\mathbf{x}}' + l\hat{\mathbf{y}}')2\pi/a$, where $i, j, k, l \in \mathbb{Z}$, $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are the unit vectors for the square lattice of the first layer, $\hat{\mathbf{x}}'$ and $\hat{\mathbf{y}}'$ are the unit vectors for the square lattice of the second layer. The scattering process has two steps. The first step is the incident light is first scattered by the first photonic crystal slab, obtaining a change $\mathbf{G}_{(i,j,0,0)}$ in wavevector, then scattered by the second slab, obtaining another wavevector shift of $\mathbf{G}_{(0,0,k,l)}$. $\mathbf{G}_{(i,j,0,0)}$ and $\mathbf{G}_{(0,0,k,l)}$ denote the reciprocal lattice vectors of the first and second layer, respectively. While the moiré scattering process generates a vast array of vectors, only a select few are observable experimentally, constrained by factors such as scattering intensity and numerical aperture (NA).

As (i, j, k, l) increase, the higher-order moiré scattering becomes weaker. Extended Data Fig. 7 shows the first four leading orders of moiré scattering vectors.

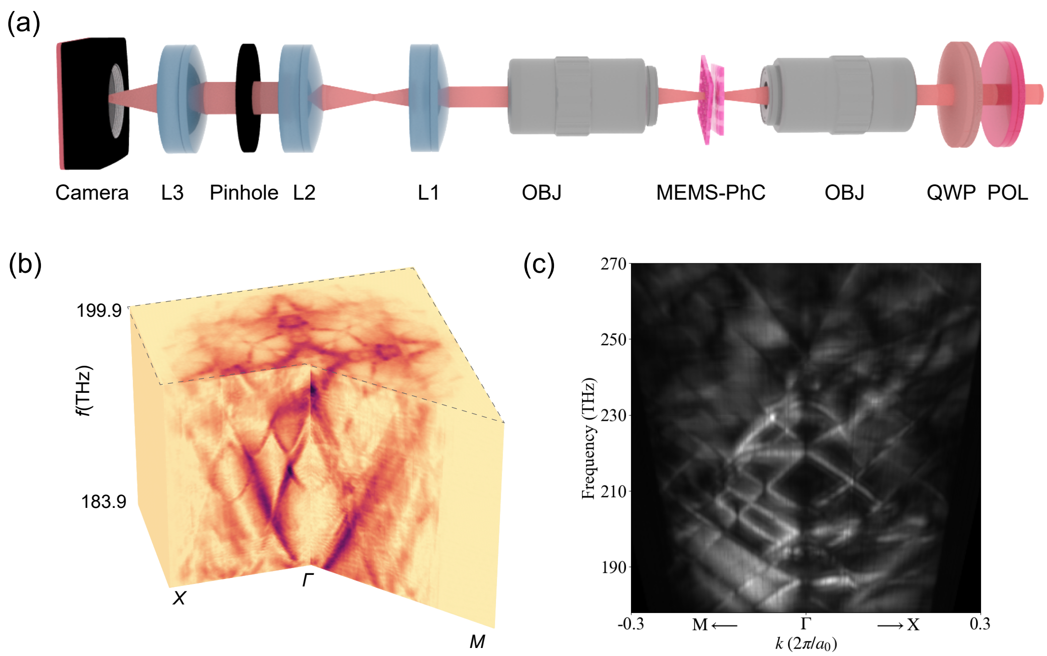


Extended Data Figure 7: **Leading orders of moiré scattering** Red (blue) grids indicate the reciprocal lattice of the first (second) photonic crystal layers. Gray grids represent the reciprocal superlattice of moiré photonic crystal. Red (blue) arrows indicate the wavevector that incident light obtained from the first (second) layer. Black arrows indicate the moiré scattering vectors. **a**, $\mathbf{G}_{(1,0,-1,0)} = \mathbf{G}_{(1,0,0,0)} - \mathbf{G}_{(0,0,1,0)}$. **b**, $\mathbf{G}_{(1,0,-1,-1)} = \mathbf{G}_{(1,0,0,0)} - \mathbf{G}_{(0,0,1,1)}$. **c**, $\mathbf{G}_{(2,0,-2,-1)} = \mathbf{G}_{(2,0,0,0)} - \mathbf{G}_{(0,0,2,1)}$. **d**, $\mathbf{G}_{(2,1,-1,-2)} = \mathbf{G}_{(2,1,0,0)} - \mathbf{G}_{(0,0,1,2)}$.

6. OPTICAL MEASUREMENT

The principle of this approach lies in utilizing optical lenses to execute the optical Fourier Transform, thereby capturing the momentum space information at the lens back focal-plane. In this measurement, a laser beam traverses a polarizer and a quarter-wave plate before reaching the back focal plane of an infinity-corrected objective lens. The light subsequently focuses on the sample plane, consistently exciting the eigenmodes of the moiré PhC at a specific wavelength. These eigenmodes are scattered by the reciprocal lattices wavevector $\mathcal{G}_{\mathrm{TB}}$ and moiré wavevector $\mathcal{G}_{\mathrm{m}}$. A subsequent confocal objective lens gathers the transmitted light, projecting it into momentum space on its back focal plane. After passing through a 4-f system, the iso-frequency contour undergoes magnification and is captured on a monochromatic CCD camera. By aggregating the iso-frequency contours across various wavelengths, we derive the bandstructure of the moiré PhC.

The band structure of the sample is measured using a free-space band structure measurement setup (see Extended Data Fig.8a). A TSL-550 Santec tunable laser, with a range between 1500 nm and 1630 nm and a frequency resolution of 0.003 nm, serves as the light source. This laser traverses the first polarizer (P1) and the quarter-wave plate (QWP) before entering the back focal plane of an infinity-corrected objective lens (O1). Light with a defined polarization state focuses on the sample plane, exciting the eigenmodes of the moiré PhC slabs at a corresponding wavelength. These eigenmodes scatter by the reciprocal lattices, resulting in resonances with finite linewidths. Additionally, fabrication disorder contributes to the broadening of these resonances. A subsequent confocal objective lens (O2) captures the transmitted light, projecting it into momentum space on its back focal plane. After transitioning through a 4-f system, this momentum space image undergoes a 1.33-fold magnification and is captured on a monochromatic CCD camera. The imaging system achieves a k-space resolution of 1.48×10^4 m^{-1}/pixel. The specific k-space imaging range and resolution, expressed in Brillouin zone units, vary depending on the numerical aperture of the objective lens, incident wavelength, and lattice constant. For instance, with an incident wavelength of 1550 nm and a single-layer PhC lattice constant of $a = 1220$ nm, the k-space imaging range spans $0.53b$, with a resolution of $0.003b$/pixel, where $b = 2\pi/a$ denotes the single-layer reciprocal lattice constant. For a 10° twist bilayer structure, the momentum space imaging range covers 5.6 b_m, with a resolution of 0.02 b_m/pixel, and $b_m = 2\pi/a_m = 4\pi\sin(\theta/2)/a$ represents the moiré reciprocal lattice constant. In the context of the transmission measurement for the optical sensing, a pinhole selects the momentum of the emitted light, and the light intensity is recorded by a photo-detector. By aggregating the iso-frequency contours across various wavelengths, we derive the bandstructure of the moiré PhC, as showcased in Extended Data Fig.8b. Using SuperK continuum laser(Laser ranges 1100 to 1700 nm with a frequency resolution of 6.4 to 19.8 nm), we can measure a wider range of band structure as Extended Data Fig.8c.



Extended Data Figure 8: **Banstructure measurement of moiré PhC.** (a) Schematic of the measurement setup. The red line represents the incident light, its direct transmission, and radiation-induced scattering from the bilayer lattice.POL - Polarizer; QWP - Quarter-Wave Plate; L - Lens; OBJ - Objective Lens. (b) Band structure measured by stacking isofrequency contours for each wavelength. The dashed line area on the top surface shows the isofrequency contour, and the cross-section reveals the band structure. (c)Bandstructure measured by SuperK laser

7. VIDEO OF MEMS, MOIRÉ PATTERN, BANDSTRUCTURES, ISO-FREQUENCY CONTOURS

See supplementary videos for

- **Realtime measurement for the dynamic tuning of the moiré pattern**
- **Bandstructures for different Vz (interlayer gap)**
- **Iso-frequency contour for different Vz (interlayer gap)**
- **Bandstructures for different twist angle**
- **Iso-frequency contour for different twist angle**
- **Iso-frequency contour when rotating quater-wave plate**
- **Iso-frequency contour when rotating polarizer**

8. RIGOROUS COUPLED-WAVE ANALYSIS (RCWA) SIMULATION, HAMILTONIAN, AND MEASUREMENT RESULT OF THE MOIRÉ PHC BANDSTRUCTURE

The measured bandstructure results matches prefectly with the rigorous coupled-wave analysis (RCWA) simulation and the Hamiltonian calculation, as shown in Extended Data Fig. 9. The detailed Hamiltonian calculation approach is explained in this section.

Single layer: The Hamiltonian can be written starting from a uniform dielectric slab, where the permittivity $\varepsilon(\mathbf{r})$ is translationally invariant in the xy plane, and waves with different in-plane wavevectors are decoupled. The periodic perturbation of the refractive index in single-layer photonic crystal slabs can be written as

$$\varepsilon(\mathbf{r}) = \sum_{\mathbf{G}_{(i,j)}} \tilde{\varepsilon}_{(i,j)}(z) e^{i\mathbf{G}_{(i,j)}\cdot\mathbf{r}_{\parallel}} \tag{10}$$

where $\mathbf{G}_{(i,j)} = i\frac{2\pi}{a}\hat{\mathbf{x}} + j\frac{2\pi}{a}\hat{\mathbf{y}}$ is the C_{4v} reciprocal lattice vector. The electromagnetic waves are modulated and scattered, facilitated by the lattice, with the following governing equation:

$$\iint \nabla\times\nabla\times\tilde{\mathbf{E}}\left(\mathbf{k}_{\parallel}, z\right) e^{i\mathbf{k}_{\parallel}\cdot\mathbf{r}_{\parallel}}\ \mathrm{d}k_x\mathrm{d}k_y = \sum_{\mathbf{G}_{(i,j)}} \iint \left(\frac{\omega}{c}\right)^2 \tilde{\varepsilon}_{(i,j)}(z)\tilde{\mathbf{E}}\left(\mathbf{k}_{\parallel} - \mathbf{G}_{(i,j)}, z\right) e^{i\mathbf{k}_{\parallel}\cdot\mathbf{r}_{\parallel}}\ \mathrm{d}k_x\mathrm{d}k_y, \tag{11}$$

where $\mathbf{k}_{\parallel}$ represents the wavevector of a specific mode. We then set the unperturbed uniform slab modes as bases and treat the lattice-facilitated scattering as coupling terms to construct the Hamiltonian. Considering the mode resonance at $\mathbf{k}_{\mathrm{inc}}$. We denote bases as

$$|\mathbf{G}_{(i,j)}\rangle \equiv \phi(z) e^{-i\left(\mathbf{G}_{(i,j)}+\mathbf{k}_{\mathrm{inc}}\right)\cdot\mathbf{r}_{\parallel}}, \tag{12}$$

where $\phi(z)$ is the profile of the bases in the z direction. A minimal set of bases to account for the scattering phenomenon can be chosen as

$$\text{Bases } = \{|\mathbf{G}_{(0,0)}\rangle, |\mathbf{G}_{(\pm1,0)}\rangle, |\mathbf{G}_{(0,\pm1)}\rangle\}. \tag{13}$$

The Hamiltonian is a 5-by-5 matrix given by

$$\left\langle \mathbf{G}_{(i,j)} | H | \mathbf{G}_{(i',j')} \right\rangle = \begin{cases} \epsilon_{(i,j)}, & \text{if } |(i,j)-(i',j')| = (0,0) \\ U_{10}, & \text{if } |(i,j)-(i',j')| = (1,0),(0,1) \\ U_{11}, & \text{if } |(i,j)-(i',j')| = (1,1) \\ U_{20}, & \text{if } |(i,j)-(i',j')| = (2,0),(0,2) \end{cases} \tag{14}$$

where $\epsilon_{(i,j)}$ is the energy for the plane wave bases. The TE/TM fundamental waveguide dispersion for an effective uniform slab is a good approximation for $\epsilon_{(i,j)}$.

AA Bilayer: For bilayer photonic crystal slabs, the bases can be localized in either layer. Namely,

$$|\mathbf{G}_{(i,j)}\rangle_h \equiv \phi_h(z) e^{-i\left(\mathbf{G}_{(i,j)}+\mathbf{k}_{\text{inc}}\right)\cdot\mathbf{r}_\parallel} \tag{15}$$

where $\phi_h(z)$ $(h = 1, 2)$ is the z-profile of the bases in the first/second layer. A minimal set of bases to account for the scattering phenomenon can be chosen as

$$\text{Bases } = \left\{ |\mathbf{G}_{(0,0)}\rangle_1, |\mathbf{G}_{(\pm1,0)}\rangle_1, |\mathbf{G}_{(0,\pm1)}\rangle_1, |\mathbf{G}_{(0,0)}\rangle_2, |\mathbf{G}_{(\pm1,0)}\rangle_2, |\mathbf{G}_{(0,\pm1)}\rangle_2 \right\}. \tag{16}$$

The Hamiltonian H is a 10-by-10 matrix given by

$$_h\left\langle \mathbf{G}_{(i,j)} | H | \mathbf{G}_{(i',j')} \right\rangle_{h'} = \begin{cases} \epsilon_{(i,j)}, & \text{if } |(i,j)-(i',j')| = (0,0) & \text{and } h = h' \\ U_{10}, & \text{if } |(i,j)-(i',j')| = (1,0),(0,1) & \text{and } h = h' \\ U_{11}, & \text{if } |(i,j)-(i',j')| = (1,1) & \text{and } h = h' \\ U_{20}, & \text{if } |(i,j)-(i',j')| = (2,0),(0,2) & \text{and } h = h' \\ V, & \text{if } |(i,j)-(i',j')| = (0,0) & \text{and } h \neq h' \end{cases} \tag{17}$$

Twisted bilayer: As for twisted bilayer photonic crystal slabs, two layers have two different sets of reciprocal lattices, $\mathcal{G}_1$ and $\mathcal{G}_2$. Namely,

$$\begin{aligned} \mathcal{G}_1 &\equiv \{(i\hat{\mathbf{x}} + j\hat{\mathbf{y}})2\pi/a \mid i,j \in \mathbb{Z}\} \equiv \{\mathbf{G}_{1,(i,j)} \mid i,j \in \mathbb{Z}\} \\ \mathcal{G}_2 &\equiv \{(k\hat{\mathbf{x}}' + l\hat{\mathbf{y}}')2\pi/a \mid k,l \in \mathbb{Z}\} \equiv \{\mathbf{G}_{2,(k,l)} \mid k,l \in \mathbb{Z}\} \end{aligned}, \tag{18}$$

where $\hat{\mathbf{x}}' = \cos(\alpha)\hat{\mathbf{x}} + \sin(\alpha)\hat{\mathbf{y}}$, $\hat{\mathbf{y}}' = \cos(\alpha)\hat{\mathbf{y}} - \sin(\alpha)\hat{\mathbf{x}}$. The reciprocal lattice for the moiré photonic crystal will be the sum of the reciprocal lattices for each layer, i.e.

$$\mathcal{G}_{\mathrm{TB}} \equiv \left\{ \left(i\hat{\mathbf{x}} + j\hat{\mathbf{y}} + k\hat{\mathbf{x}}' + l\hat{\mathbf{y}}' \right) 2\pi/a \mid i,j,k,l \in \mathbb{Z} \right\} \equiv \left\{ \mathbf{G}_{(i,j,k,l)} \mid i,j,k,l \in \mathbb{Z} \right\} \tag{19}$$

So we denote the bases as

$$|\mathbf{G}_{(i,j,k,l)}\rangle_h = \phi_h(z) e^{-i\left(\mathbf{G}_{(i,j,k,l)} + \mathbf{k}_{\mathrm{inc}}\right)\cdot \mathbf{r}_\parallel}. \tag{20}$$

To show the lowest order moiré scattering $\left\{ \mathbf{G}_{(\pm 1,0,\mp 1,0)}, \mathbf{G}_{(0,\pm 1,0,\mp 1)} \right\}$, we construct bases starting from the following three sets

$$\begin{aligned} \mathcal{G}_1' &\equiv \left\{ \mathbf{G}_{(0,0,0,0)}, \mathbf{G}_{(\pm 1,0,0,0)}, \mathbf{G}_{(0,\pm 1,0,0)} \right\} \\ \mathcal{G}_2' &\equiv \left\{ \mathbf{G}_{(0,0,0,0)}, \mathbf{G}_{(0,0,\pm 1,0)}, \mathbf{G}_{(0,0,0,\pm 1)} \right\} \quad , \\ \mathcal{G}_{\mathrm{m}}' &\equiv \left\{ \mathbf{G}_{(0,0,0,0)}, \mathbf{G}_{(\pm 1,0,\mp 1,0)}, \mathbf{G}_{(0,\pm 1,0,\mp 1)} \right\} \end{aligned} \tag{21}$$

and the bases are set as

$$\begin{aligned} \mathcal{G}_{\mathrm{bases}} &= \{ \mathbf{G}_1' + \mathbf{G}_{\mathrm{m}}' \mid \mathbf{G}_1' \in \mathcal{G}_1', \mathbf{G}_m' \in \mathcal{G}_{\mathrm{m}}' \} \cup \{ \mathbf{G}_2' + \mathbf{G}_{\mathrm{m}}' \mid \mathbf{G}_2' \in \mathcal{G}_2', \mathbf{G}_m' \in \mathcal{G}_{\mathrm{m}}' \} \\ \text{Bases } &= \{ |\mathbf{G}_{\mathrm{bases}}\rangle_h \mid \mathbf{G}_{\mathrm{bases}} \in \mathcal{G}_{\mathrm{bases}}; h = 1,2 \} \end{aligned} \tag{22}$$

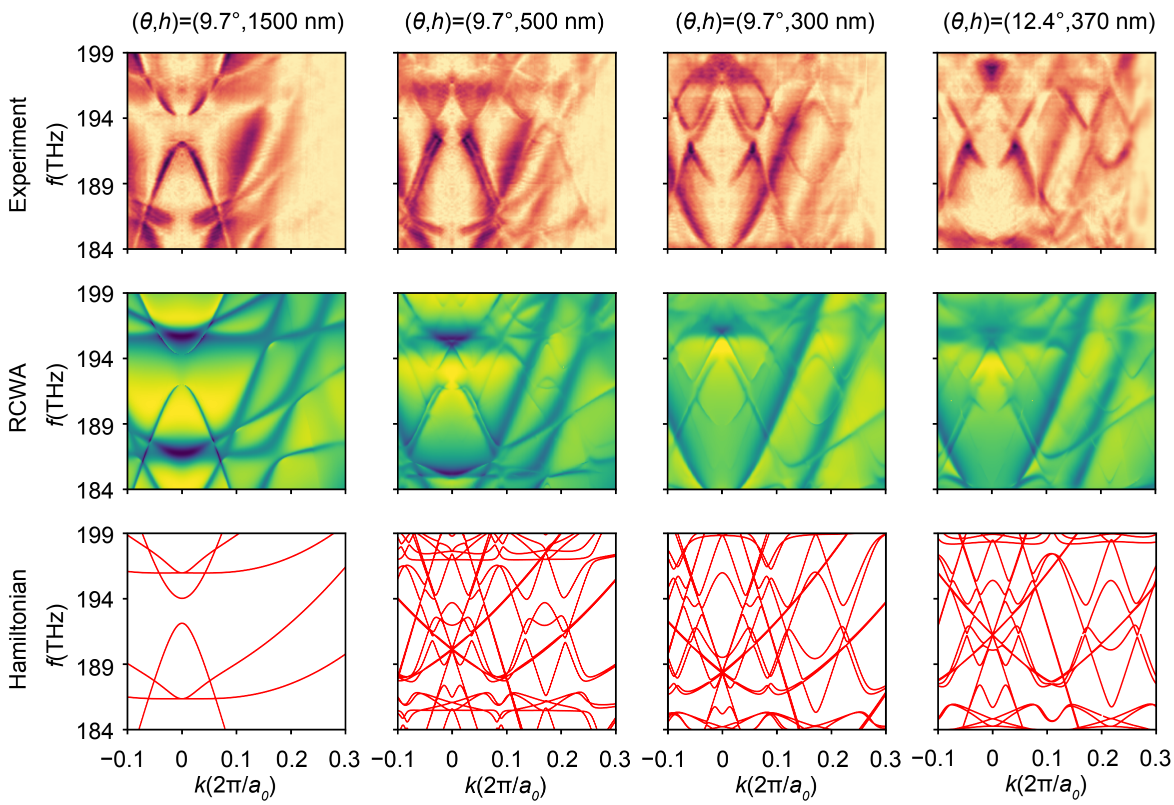
The Hamiltonian H is

$$\begin{aligned} &{}_h\left\langle \mathbf{G}_{(i,j,k,l)} | H | \mathbf{G}_{(i',j',k',l')} \right\rangle_{h'} \\ &= \begin{cases} \epsilon_{(i,j,k,l)}, & \text{if } |(i,j,k,l) - (i',j',k',l')| = (0,0,0,0) & \text{and } h = h' \\ U_{10}, & \text{if } |(i,j,k,l) - (i',j',k',l')| = (1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1) & \text{and } h = h' \\ U_{11}, & \text{if } |(i,j,k,l) - (i',j',k',l')| = (1,1,0,0),(0,0,1,1) & \text{and } h = h' \\ U_{20}, & \text{if } |(i,j,k,l) - (i',j',k',l')| = (2,0,0,0),(0,2,0,0),(0,0,2,0),(0,0,0,2) & \text{and } h = h' \\ V, & \text{if } |(i,j,k,l) - (i',j',k',l')| = (0,0,0,0) & \text{and } h \neq h' \end{cases} \end{aligned} \tag{23}$$

To calculate the coupling strength of both input and output light. We define the parameter

$$\begin{aligned} \tilde{\kappa} = & \sum_{\lambda \in \left\{ \mathbf{G}_{(i,j,0,0)_1} \right\}} \frac{x_\lambda^2 U_\lambda}{\sum_{\nu \in \left\{ \mathbf{G}_{(i,j,k,l)_1} \right\}} x_\nu^2} \\ & \times \sum_{\lambda \in \left\{ \mathbf{G}_{(0,0,k,l)_2} \right\}} \frac{x_\lambda^2 U_\lambda}{\sum_{\nu \in \left\{ \mathbf{G}_{(i,j,k,l)_2} \right\}} x_\nu^2} \end{aligned} \tag{24}$$

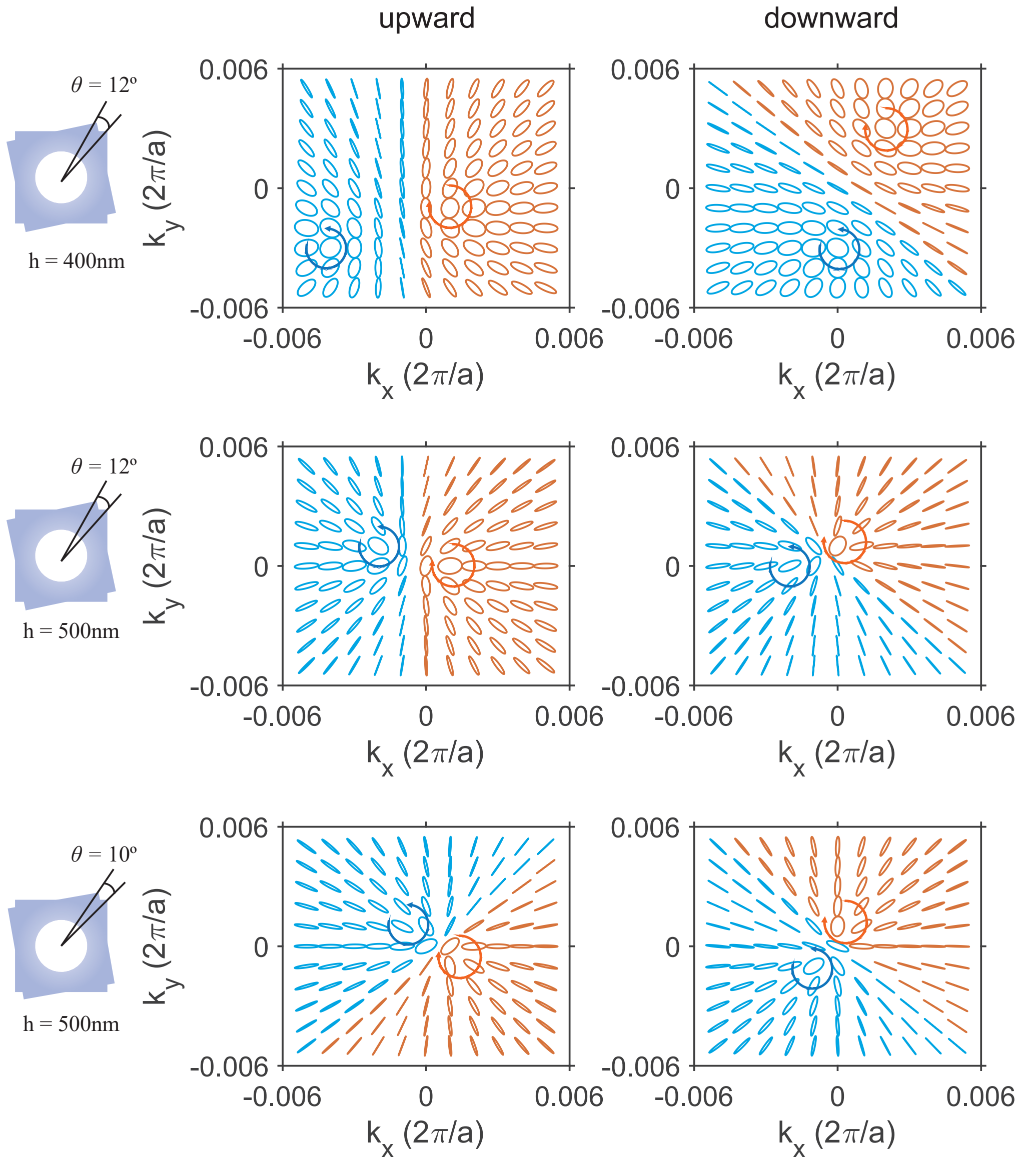
as an indicator for the total coupling strength of an eigenmode, where λ and ν are the bases in each layer, x_λ (x_ν) is the weight of bases λ (ν) in the eigenmode, U_λ is the coupling strength between the bases λ and the continuum.



Extended Data Figure 9: **Experiment measurement (first row), Rigourous Coupled Wave Analysis (RCWA) simulation (second row), and Hamiltonian analysis (third row) of the moiré PhC bandstructure, for configurations** $(\theta, h) = (9.7°, 1500\ \mathrm{nm}), (\theta, h) = (9.7°, 500\ \mathrm{nm}), (\theta, h) = (9.7°, 300\ \mathrm{nm})\ (\theta, h) = (12.4°, 370\ \mathrm{nm})$.

9. POLARIZATION DEPENDENCY AND MOMENTUM SPACE POLARIZATION FIELD

The moiré photonic crystal offers tunable chirality[16] and tunable polarization singularities[17–19], situated around the center of momentum space. In this study, it is demonstrated that our moiré photonic crystal presents a polarized far field with a winding angle of π (C-point). Additionally, by introducing a twist angle to break the σ_z-symmetry, one can achieve both an asymmetry in the upward and downward directions and a tunable chirality. This is achieved through dynamic adjustments of the twist angle and interlayer gap, as illustrated in Extended Data Fig. 10. For example, near the vicinity of Γ-point, the photonic crystal exhibits elliptical polarization states characterized by ellipticity and azimuth angle. When varying gap h and twist angle θ, the far-field polarization response in the momentum space of each configurations are different.The chirality, represented by left-hand polarization (LCP) states (blue ellipse) and right-hand polarization (RCP) states (orange ellipse) can be largely affect by h and θ. Besides the chirality, the position of C-point is also affected by h and θ, as represented by the circular winding center in Extended Data Fig. 10. As a result of breaking the σ_z-symmetry, the original bound-state-in-the continuum V-point breaks into two C-point, and the position of two C-point in momentum space undergoes a shift. In this research, the tunable polarization state in moiré photonic crystal holds potential for expanding applications in structured light, quantum optics, and twistronics for photons.



Extended Data Figure 10: **Simulated far-field polarization distribution in the momentum space for different moiré photonic crystal configurations** (a) Far-field polarization distribution in the momentum space when $\theta = 12$ and $h = 400nm$ (top row), $\theta = 12$ and $h = 500nm$ (middle row), and $\theta = 10$ and $h = 500nm$ (bottom row). The polarization fields of emitting upward and downward are different. The blue ellipse represent left-hand polarization (LCP) states and the orange ellipse represent right-hand polarization (RCP) states. The winding centers of the LCP and RCP are the polarization singularity C-point. The rotational center of two PhC slabs are perfectly aligned.

10. CALIBRATION TO THE WAVELENGTH DEPENDENT QUARTER-WAVE PLATE

The quarter-wave plate retardance is highly depended on the wavelength. It is therefore necessary to calibrate the quarter-wave plate during the broadband measurement. We used two polarizers, one quarter-wave plate and one powermeter to calibrate the retardance of quarter-wave plate. Direct the light source through polarizer 1 and then through polarizer 2, aiming it at the power meter. Adjust the angle of polarizer 1 until the light is completely blocked. Insert a quarter-wave plate between polarizer 1 and polarizer 2, and adjust the angle of polarizer 2 until the light is completely blocked. Rotate the quarter-wave plate by 45°. At this wavelength, the combination of polarizer 1 and the quarter-wave plate will produce light that is as close to circularly polarized as possible. Rotate polarizer 2 and measure the intensity of the major and minor axes of this elliptically polarized light. Calculate the ellipticity of this light, and then determine the retardance of the quarter-wave plate.

11. SMART SENSING ALGORITHM

Extended Data Fig. 11a shows a few examples of device configurations and their corresponding observation functions, sampled over a discrete set of bases. It is critical that different signals correspond to distinct observation functions, otherwise, they would be indiscernible. In realistic setting, the adjustment of device configuration has a finite accuracy in each DoF, and therefore, a large number of DoF is desirable for high sensing resolution. Several approaches can be used for signal reconstruction. For example, one commonly adopted formulation of compressed sensing is to reconstruct the signal in a space defined by a set of bases, where the signal is expected to have relative sparsity. The sequence of measurement under different device configurations needs to be determined for a more informative measurement. Extended Data Fig.11b illustrates two ways for this purpose. One way is to sample the configuration space in a predetermined sequence. A uniform sampling across the configuration space, based on the desired number of measurements, could be sufficiently informative. This non-adaptive approach is favorable when a large quantity of measurements is affordable. Alternatively, one can iteratively select the next measurement configuration based on updated signal knowledge. One common algorithm for adaptive selection of measurement bases is to maximize the information gain based on updated prior knowledge[20], potentially reducing the number of measurements compared to the non-adaptive approach.

A. Signal reconstruction from measurement

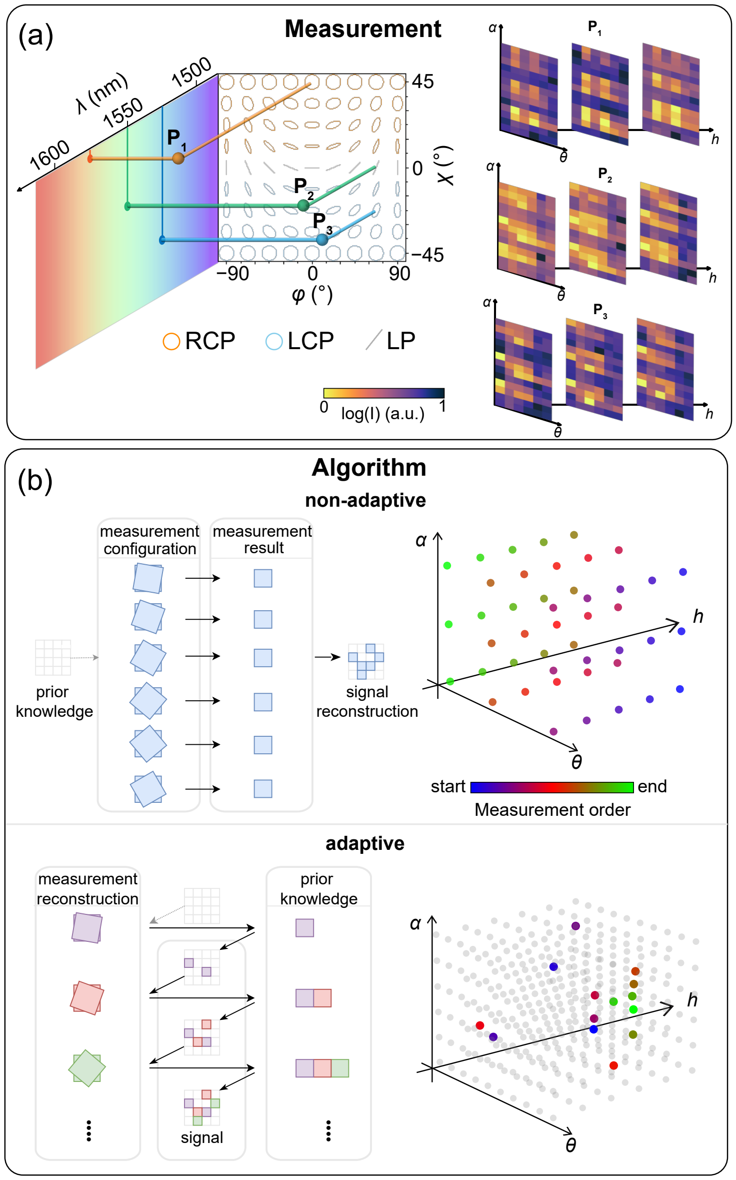
In typical compressed sensing framework, with $x \in \mathbb{R}^n$ being the signal to be recovered and $y \in \mathbb{R}^m$ being the measurement results, the signal reconstruction process can be mathematically written as:

$$\begin{aligned} &\arg\min_x \|x\|_0 \\ &\text{subject to } Ax = y \end{aligned} \tag{25}$$

where $A \in \mathbb{R}^{m \times n}$ is the sensing matrix that maps signals in to measurement results.

With some relaxation for noise, Eq25 becomes:

$$\begin{aligned} &\arg\min_x \|x\|_0 \\ &\text{subject to } \|Ax - y\| < \epsilon \end{aligned} \tag{26}$$



Extended Data Figure 11: **Smart sensing principle**(a) Right: Expected intensity measurements for three signals, as a function of twist angle (α), interlayer gap distance (h), and detection angle (represented by in-plane wave vector (θ). Left: The physical properties of the signals, defined by wavelength (λ) and polarization, where polarization is represented by azimuth (ϕ) and ellipticity (χ). (b) Right: Schematic of non-adaptive (top) and adaptive sensing (bottom) approach. The non-adaptive sensing has all measurements involved synchronized, the adaptive sensing pro forms measurement and optimization one-by-one. Left: The configurations and measurement order of two sensing approaches.

Since $L0$-norm optimization is not tractable, a common practice is to relax the problem into

$L1$-norm optimization:

$$\begin{aligned} &\arg\min_x \|x\|_1 \\ &\text{subject to } \|Ax - y\| < \epsilon \end{aligned} \tag{27}$$

It is often desirable to relax the problem further into a convenient unconstrained optimization:

$$\arg\min_x \|Ax - y\|_2^2 + \lambda_1 \|x\|_1 \tag{28}$$

where λ_1 is a hyperparameter governing how much we penalize large $L1$ norm.

The variations of mathematical formulations have been discussed extensively in the literature, and are not the focus of our discussion here. We use Eq28 in our demonstrations, but our framework is compatible with other methods for signal reconstruction, too.

B. Adaptive choice of measurement basis

In our system, the sensing device can be reconfigured in real time, which allows adaptive selection of bases, in addition to the benefits that these bases allow compressed sensing.

Each measurement result is subject to noise, of the following form:

$$y_i = a_i^T x + w_i \tag{29}$$

where $w_i \sim \mathcal{N}(0, \sigma_w^2)$ is the measurement noise. The formulation can be easily modified if other forms of noise is desired.

The problem with adaptive choice of measurement basis is to select $a_i \in \mathcal{A}$ in a smart way so that the signal reconstruction is of higher fidelity. This could either result in a higher accuracy of the recovered signal, or an early termination that ends up using fewer measurements.

A common approach is to use Bayesian analysis, iteratively deriving a posterior distribution for the signal and deciding the next measurement basis based on mutual information maximization[?] . The posterior can be calculated as:

$$P(x|\{y_j, a_j\}_{j<i}) = \frac{P(\{y_j, a_j\}_{j<i}|x)P(x)}{\sum_{x' \in \mathcal{X}} P(\{y_j, a_j\}_{j<i}|x')P(x')} \tag{30}$$

where $\mathcal{X}$ is a discrete set of possible signals, and the measurement basis can be chosen as:

$$a_i = \arg\max_{a_i \in \mathcal{A}} \mathbb{I}[x; a_i^T x + w_i | \{y_j, a_j\}_{j<i}] \tag{31}$$

where $\mathbb{I}[\cdot;\cdot]$ denotes the mutual information between two random variables.

The adaptive sensing process starts with an initial choice of measurement basis a_0, which results in measurement result y_0, and iterates Eq30 and 31 until either the number of measurements reaches a certain limit, or the mutual information is smaller than a certain threshold. Note that the posterior estimation here makes use of the prior distribution of the signal $P(x)$. In principle, the prior knowledge that the signal is sparse could be taken into account in $P(x)$, but in practice, explicitly writing out the sparsity prior may be infeasible. In some cases, as will be demonstrated, one can simply replace the posterior calculation with a compressed sensing routine, and treat the recovered mixture of signals as a posterior distribution, as will be demonstrated in the following section.

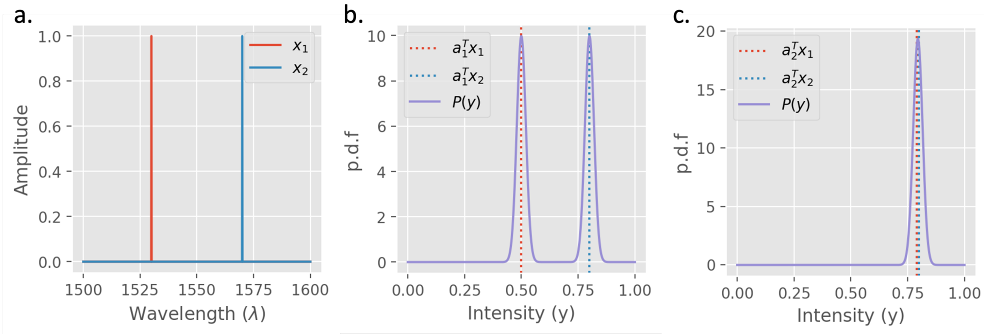
C. A detailed example with monochromatic signal

For simplicity, we demonstrate the algorithm on the problem of spectrum reconstruction for a monochromatic signal. Under realistic resolution, the spectrum is a delta function $x^* = \delta_{\lambda,\lambda^*}$ with the peak at $\lambda^* \in \Lambda$, where $\Lambda = \{\lambda_1, ..., \lambda_n\}$ is a discrete set of possible frequencies.

The measurement process is an intensity measurement of the form $y_i = a_i^T x + w_i$, where a_i is the transmission coefficients over the frequency bases and $w_i \sim \mathcal{N}(0, \sigma_w^s)$ is the measurement noise.

Consider the case of $n = 2$. Suppose we have two possible wavelengths for the signal: $\lambda_1 = 1530nm$ and $\lambda_2 = 1570nm$. The corresponding signals are $x_1 = \delta_{\lambda,\lambda_1}$ and $x_2 = \delta_{\lambda,\lambda_2}$, as illustrated in Extended Data Fig.12a. Before we perform any measurement, if we do not have any prior knowledge on which signal is more likely, one common choice is to start with prior distribution of $P(x)$ being a uniform distribution over $\mathcal{X} = \{x_1, x_2\}$.

To see how different choices of measurement bases could have different mutual information with the prior knowledge, which informs our adaptive selection, we consider two measurement bases a_1 and a_2, with their responses to the signals shown in Extended Data Fig.12b and c. Namely, intensity measurement using a_1 renders a reading of around 0.5 for signal x_1 and a reading of around 0.8 for signal x_2, whereas intensity measurement using a_2 renders a reading of around 0.79 for signal x_1 and a reading of around 0.8 for signal x_2. Note that there is a little uncertainty in the measurement, which is necessary to motivate the adaptation of measurement bases.



Extended Data Figure 12: Illustration of two signals (a) and their measurement results under different choices of measurement bases (b,c). The measurement results that distinguish the two signals (b) is more informative than the ones that do not distinguish the two signals (c).

The mutual information can be calculated as:

$$\mathbb{I}[x; y] = \mathbb{H}[y] - \mathbb{H}[y|x] \tag{32}$$

where

$$\mathbb{H}[y] = -\sum_{y' \in \mathcal{Y}} P_y(y') \log P_y(y') \tag{33}$$

$$\mathbb{H}[y|x] = -\sum_{x' \in \mathcal{X}} P_x(x') \big[\sum_{y' \in \mathcal{Y}} P_y(y'|x') \log P_y(y'|x') \big] \tag{34}$$

Here $\mathbb{H}[y]$ is the marginal entropy for the measurement result y, which is a random variable due to noise in measurement. Similarly, $\mathbb{H}[x]$ is the marginal entropy for the signal x, which can also be treated as a random variable due to uncertainty in our prior knowledge. The log is in base 2 by convention, so that the entropy values are in the unit of bits. In our example, we know $\mathbb{H}[x] = 1$, since there is only 1 bit of information in the knowledge of which signal it is.

$\mathbb{H}[y|x]$ is the conditional entropy, which describes the amount of information in the measurement results given knowledge of the true signal. $\mathbb{H}[y|x]$ only depends on the level of measurement noise σ_w and does not depend on the choice of measurement bases. When there is no noise, $\mathbb{H}[y|x] = 0$, since the measurement does not contain any new information given x. When there is large noise, $\mathbb{H}[y|x]$ becomes large as the measurement does contain information, which may not be useful for us though. Using the distribution of noise w, we have:

$$\mathbb{H}[y|x] = \int P_w(w') \log P_w(w') dw' \tag{35}$$

For simplicity, let $h \equiv \mathbb{H}[y|x]$.

In the case of *distinguishable* measurement results (Extended Data Fig.12b),

$$\mathbb{H}[y] \approx \Big(\int \frac{P_w(w')}{2} \log \frac{P_w(w')}{2} dw'\Big) + \Big(\int \frac{P_w(w')}{2} \log \frac{P_w(w')}{2} dw'\Big) \tag{36}$$

$$= h + 1 \tag{37}$$

The mutual information between the measurement results and the signal distribution is therefore:

$$\mathbb{H}[y] - \mathbb{H}[y|x] = 1 \tag{38}$$

Namely, there is 1 bit of useful information in the measurement results that can help us tell the signals apart.

In the case of *non-distinguishable* measurement results (Extended Data Fig.12c),

$$\mathbb{H}[y] \approx \int \Big(\frac{P_w(w')}{2} + \frac{P_w(w')}{2}\Big) \log\Big(\frac{P_w(w')}{2} + \frac{P_w(w')}{2}\Big) dw' \tag{39}$$

$$= h \tag{40}$$

The mutual information between the measurement results and the signal distribution is therefore:

$$\mathbb{H}[y] - \mathbb{H}[y|x] = 0 \tag{41}$$

Namely, the measurement results do not contain any useful information that can help us tell the signals apart.

While this example is for $n = 2$, the formulation is general and robust for large n. The following sections explain some details on generalization to more dimensions and some expedient measures for more efficient computations.

D. A few words on generalization

While the previous example assumes a monochramatic signal, the method can be easily generalized to more complicated prior distribution with the help of a dictionary of possible signals[?] , namely:

$$x = \Phi c \tag{42}$$

where Φ is a projection matrix, preferably an orthonormal one.

The signal dimension can be higher than 1. For example, it may be desirable to reconstruct frequency, polarization and phase together. In that case, one can extend the formulation discussed in the previous sections by treating x as three random variables, e.g. X_1, X_2, X_3. The measurement result is still only one variable y, and the mutual information with the signal can still be used as the metric for selection of measurement bases.

Note that in many traditional methods, the signal distribution in each dimension has to be reconstructed separately. For example, one could vary the angle of polarizers to find out the distributiion over polarizations, and then scan through various frequency filters or Fabry-Perot lineshapes to find out the distribution over frequencies. Such methods are highly inefficient when there is correlation across the dimensions. In our framework, the prior distribution of the signal is a joint distribution over all the dimensions, and can easily accommodate correlations across the dimensions. Namely, computational sensing in our framework allows more flexibility in encoding prior knowledge when

$$P(X_1, X_2, ..., X_n) \neq P(X_1) \times P(X_2) \times ... \times P(X_n) \tag{43}$$

E. A few words on heuristic variations

In practice, many parameters are not precisely known. For example, the noise distribution $sigma_w$ may not be well characterized. The noise process may also take more complicated forms. Therefore, it often makes sense to replace some time-consuming steps in the Bayesian method by some heuristic variations.

In the case of monochromatic signal reconstruction, or more generally, when we know the signal is a delta distribution over a discrete set, we can recover the signal from merely a comparison of dot products with a pre-calculated dictionary of vectors. Namely, over a number of measurements

$\{a_1, a_2, ..., a_m\}$, the measurement results can be collected into a vector $y \in \mathbb{R}^m_{\geq 0}$. For each possible signal $x \in \{x_1, x_2, ..., x_n\}$, we have a dictionary of the expected measurement result vectors $y \in \{y_1, y_2, ..., y_n\}$, with y_i being the measurement result vector for x_i. Here each y is of dimension m and is normalized to have total $L2$ norm of 1. The signal recovery can be implemented as finding the signal corresponding to :

$$\arg\max_i y_i^T y \tag{44}$$

Note that maximizing dot product as in Eq.44 corresponds to maximum likelihood estimation when the noise process is an additive white noise, as assumed in the previous sections. Namely, if the true signal is x_i, the likelihood of observing y is:

$$P(y|x = x_i) \propto \exp\left(-\frac{\|y - y_i\|_2^2}{2\sigma_w^2}\right) \tag{45}$$

To maximize the likelihood of observing y is equivalently maximizing the logarithm of Eq.45, which is to maximize:

$$-\|y - y_i\|_2^2 = -\|y\|_2^2 - \|y_i\|_2^2 + 2y_i^T y \tag{46}$$

Since $\|y_i\|_2^2 = 1$ from normalization, we are effectively maximizing $y_i^T y$, as is achieved from Eq.44.

While Eq.44 is a optimal under additive white noise assumption, it reduces to a mere convenient approximation to the maximum likelihood estimation when the noise process deviates from the additive white noise assumption. For example, in experiment, we often observes multiplicative white noise, of the form:

$$y_i = (1 + w_i)a_i^T x \tag{47}$$

where $w_i \sim \mathcal{N}(0, \sigma_w^2)$ is the measurement noise.

In the case of Eq.47, the conditional probability of observing y under the expected measurement result vector y_i^* is:

$$P(y|x = x_i) \propto \exp\left(-\frac{\sum_{j=1}^m ([y]_j - [y_i]_j)^2/[y_i]_j^2}{2\sigma_w^2}\right) \tag{48}$$

where $[\cdot]_j$ denotes the j-th entry of the vector in the bracket. To maximize the likelihood of observing y is therefore equivalent to:

$$\arg\min_i \sum_{j=1}^{m}([y]_j - [y_i]_j)^2/[y_i]_j^2 \tag{49}$$

The variations above can be used in the place of posterior calculation and signal reconstruction. On the other hand, for the adaptive selection process, the mutual information calculation can be replaced with simpler heuristic variations, too.

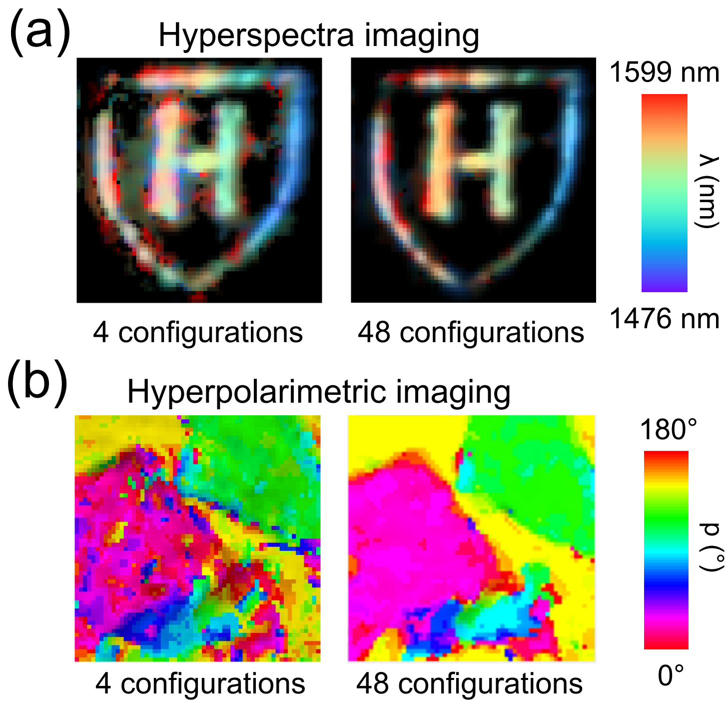
Instead of maximize mutual information between potential measurement result and the prior distribution, one can maximize the variance of the expected measurement result over the prior distribution. Namely, for each of measurement bases $a_i \in \{a_1, a_2, ..., a_m\}$, calculate $\mathrm{Var}[y_i]$ over the prior distribution and pick the measurement basis corresponding to:

$$\arg\max_i \mathrm{Var}[y_i] \tag{50}$$

which is an approximation to maximizing $\mathbb{H}[y]$ and therefore maximizing $\mathbb{I}[x; y]$.

Another alternative that allows more hyperparameter tuning is to use an objective function that considers the tradeoff between maximizing expected response to the signal according to the prior distribution, and minimizing overlap to the already measured bases. The first consideration is for the sake of signal-noise-ratio, whereas the second consideration is for the sake of gaining more new information. The relative weights between these two considerations are hyperparameters that can be tuned given the specific circumstances.

One can go further can encode the whole adaptation process to a neural network. For example, with a transformer architecture, the input can be a sequence of previous measurements, each in its own encoding, and the model predicts which measurement basis as a next one can give the most informative result. The training labels can either be calculated from the empirical accuracy for various sequences of measurements. There could also be a pre-training curriculum where one uses the methods discussed in the previous sections as a teacher to bring the model onto the right track. Such a neural-network model may be desirable especially when the tuning knobs are high-dimensional and potentially continuous variables. For example, in the case of twisted multi-layer photonic crystal, the possible set of measurement bases $\mathcal{A}$ is huge as it contains all possible combinations of twist angles for each layer. A neural network can efficiently parametrize the response functions for these different measurement bases. In an end-to-end fashion, the neural network that parametrizes the response function and the neural network that maximizes mutual information can be achieved in one model with further model compression and regularization. It can also accommodate more sophisticated objective function taking physical constraints into account. For example, there could be a different cost function associated with each measurement basis some twist angles further away are more time-consuming to reach, or potentially less accurate given the current configuration (see Fig.13).



Extended Data Figure 13: **Comparison of using different configurations in hyperspectral and hyperpolarimetric imaging** (a) Comparison of using 4 and 48 configurations in the hyperspectral imaging. (b) Comparison of using 4 and 48 configurations in the hyperpolarimetric imaging.

12. COMPARISON OF DIFFERENT MINIATURIZED SENSORS

TABLE II: Comparison of different miniaturized sensors (selected)

Method	Operation enviroment	Spectral resolution	Spectral range (nm)	Polarimetric range	Device reproduction	Smart sensing	Type (reference)
Photonic-crystal slabs[21]	Room temperature	2 nm	550 - 750	NA	Yes	No	Conventional
Dielectric metasurfaces[22]	Room temperature	16.5 nm	5650 - 7300	NA	Yes	No	Conventional
Folded metasurfaces[23]	Room temperature	1.2 nm	760 - 860	NA	Yes	No	Conventional
Mach-Zehnder interferometer[24]	Room temperature	0.2 nm	1550 - 1570	NA	Yes	No	Conventional
Colloidal quantum dots[25]	Room temperature	3.2 nm	400 - 700	NA	Yes	No	Computational
Multiple silicon nanowires[26]	Room temperature	6 nm	450 - 800	NA	Yes	No	Computational
Single nanowire[27]	Room temperature	10 nm	500 - 630	NA	Yes	No	Computational
Black phosphorus[28]	Cryogenic	90 nm	2000 - 9000	NA	Yes	No	Computational
Superconducting nanowire[29]	Cryogenic	6 nm	660 - 1900	NA	No	No	Computational
ReS2/Au/WSe2 vdW heterostructure[30]	Cryogenic	20 nm	1150 - 1470	NA	No	No	Computational
MoS2/WSe2 vdW heterojunction[31]	Room temperature	3 nm	405 - 845	NA	No	No	Computational
Metasurface[32]	Room temperature	0.23 nm	450–750	NA	Yes	No	Computational
Twisted double-layer graphene[33]	Cryogenic	NA	NA	full-stoke	No	No	Computational, CNN
Dielectric metasurface[34]	Room temperature	NA	3000 - 4000	full-stoke	Yes	No	Computational
Moiré optics	Room temperature	0.1 nm	1500 - 1630 (In principle 1100 - 1700)	full-stoke	Yes	Yes	Computational, adaptive

---

* hat431@g.harvard.edu

† caoyuan@berkeley.edu

‡ shanhui@stanford.edu

§ mazur@seas.harvard.edu

[1] Israelachvili, J. N. *Intermolecular and Surface Forces* (Academic Press, 2010).

[2] Wang, Y., Stellinga, D., Klemm, A. B., Reardon, C. P. & Krauss, T. F. Tunable Optical Filters Based on Silicon Nitride High Contrast Gratings. *IEEE Journal of Selected Topics in Quantum Electronics* **21**, 108–113 (2015).

[3] Han, Z., Colburn, S., Majumdar, A. & Böhringer, K. F. MEMS-actuated metasurface Alvarez lens. *Microsystems & Nanoengineering* **6**, 1–11 (2020).

[4] Roy, T. *et al.* Dynamic metasurface lens based on MEMS technology. *APL Photonics* **3**, 021302 (2018).

[5] Dynamic piezoelectric MEMS-based optical metasurfaces | Science Advances.

[6] She, A., Zhang, S., Shian, S., Clarke, D. R. & Capasso, F. Adaptive metalenses with simultaneous electrical control of focal length, astigmatism, and shift. *Science Advances* **4**, eaap9957 (2018).

[7] MEMS cantilever–controlled plasmonic colors for sustainable optical displays | Science Advances.

[8] Holsteen, A. L., Cihan, A. F. & Brongersma, M. L. Temporal color mixing and dynamic beam shaping with silicon metasurfaces. *Science* **365**, 257–260 (2019).

[9] MEMS Tunable Metasurfaces Based on Gap Plasmon or Fabry–Pérot Resonances | Nano Letters.

[10] Zhu, W. M. *et al.* Switchable Magnetic Metamaterials Using Micromachining Processes. *Advanced Materials* **23**, 1792–1796 (2011).

[11] Ou, J.-Y., Plum, E., Zhang, J. & Zheludev, N. I. An electromechanically reconfigurable plasmonic metamaterial operating in the near-infrared. *Nature Nanotechnology* **8**, 252–255 (2013).

[12] Tao, H. *et al.* Reconfigurable Terahertz Metamaterials. *Physical Review Letters* **103**, 147401 (2009).

[13] Ou, J. Y., Plum, E., Jiang, L. & Zheludev, N. I. Reconfigurable Photonic Metamaterials. *Nano Letters* **11**, 2142–2144 (2011).

[14] Zhao, X. *et al.* Electromechanically tunable metasurface transmission waveplate at terahertz frequencies. *Optica* **5**, 303–310 (2018).

[15] Arbabi, E. *et al.* MEMS-tunable dielectric metasurface lens. *Nat. Commun.* **9**, 812 (2018).

[16] Chen, M. K. *et al.* Chiral-magic angle of nanoimprint meta-device. *Nanophotonics* **12**, 2479–2490 (2023).

[17] Qin, H. *et al.* Arbitrarily polarized bound states in the continuum with twisted photonic crystal slabs. *Light Sci. Appl.* **12**, 66 (2023).

[18] Overvig, A., Yu, N. & Alù, A. Chiral quasi-bound states in the continuum. *Phys. Rev. Lett.* **126** (2021).

[19] Huang, L., Zhang, W. & Zhang, X. Moir\'e Quasibound States in the Continuum. *Physical Review Letters* **128**, 253901 (2022).

[20] Braun, G., Pokutta, S. & Xie, Y. Info-greedy sequential adaptive compressed sensing. In *2014 52nd*

Annual Allerton Conference on Communication, Control, and Computing (Allerton), 858–865 (2014).

[21] Wang, Z. *et al.* Single-shot on-chip spectral sensors based on photonic crystal slabs. *Nature Communications* **10**, 1020 (2019).

[22] Imaging-based molecular barcoding with pixelated dielectric metasurfaces | Science.

[23] Faraji-Dana, M. *et al.* Compact folded metasurface spectrometer. *Nature Communications* **9**, 4196 (2018).

[24] Kita, D. M. *et al.* High-performance and scalable on-chip digital Fourier transform spectroscopy. *Nature Communications* **9**, 4405 (2018).

[25] Bao, J. & Bawendi, M. G. A colloidal quantum dot spectrometer. *Nature* **523**, 67–70 (2015).

[26] Detector-Only Spectrometer Based on Structurally Colored Silicon Nanowires and a Reconstruction Algorithm | Nano Letters.

[27] Single-nanowire spectrometers | Science.

[28] Yuan, S., Naveh, D., Watanabe, K., Taniguchi, T. & Xia, F. A wavelength-scale black phosphorus spectrometer. *Nature Photonics* **15**, 601–607 (2021).

[29] Kong, L. *et al.* Single-Detector Spectrometer Using a Superconducting Nanowire. *Nano Letters* **21**, 9625–9632 (2021).

[30] Deng, W. *et al.* Electrically tunable two-dimensional heterojunctions for miniaturized near-infrared spectrometers. *Nature Communications* **13**, 4627 (2022).

[31] Miniaturized spectrometers with a tunable van der Waals junction | Science.

[32] Xiong, J. *et al.* Dynamic brain spectrum acquired by a real-time ultraspectral imaging chip with reconfigurable metasurfaces. *Optica* **9**, 461–468 (2022).

[33] Ma, C. *et al.* Intelligent infrared sensing enabled by tunable moiré quantum geometry. *Nature* **604**, 266–272 (2022).

[34] Wang, J. *et al.* Continuous-Spectrum–Polarization Recombinant Optical Encryption with a Dielectric Metasurface. *Advanced Materials* **n/a**, 2304161.