Topological polaritons and photonic magic angles in twisted \( \alpha \)-MoO\(_3\) bilayers

Twisted two-dimensional bilayer materials exhibit many exotic electronic phenomena. Manipulating the ‘twist angle’ between the two layers enables fine control of the electronic band structure, resulting in magic-angle flat-band superconductivity\(^1\,^2\), the formation of moiré excitons\(^3\,^4\,^5\) and interlayer magnetism\(^6\). However, there are limited demonstrations of such concepts for photons. Here we show how analogous principles, combined with extreme anisotropy, enable control and manipulation of the photonic dispersion of phonon polaritons in van der Waals bilayers. We experimentally observe tunable topological transitions from open (hyperbolic) to closed (elliptical) dispersion contours in bilayers of \( \alpha \)-phase molybdenum trioxide (\( \alpha \)-MoO\(_3\)), arising when the rotation between the layers is at a photonic magic twist angle. These transitions are induced by polariton hybridization and are controlled by a topological quantity. At the transitions the bilayer dispersion flattens, exhibiting low-loss tunable polariton canalization and diffractionless propagation with a resolution of less than \( \lambda_0/40 \), where \( \lambda_0 \) is the free-space wavelength. Our findings extend twistronics\(^10\) and moiré physics to nanophotonics and polaritonics, with potential applications in nanoimaging, nanoscale light propagation, energy transfer and quantum physics.

Among the electronic phenomena observed in twisted stacks of two-dimensional (2D) materials are superconductivity at magic angles that induce a flat Fermi surface in bilayer graphene\(^1\,^2\), topological excitons resulting from twist-angle-dependent interlayer hopping in bilayer transition-metal dichalcogenides\(^3\,^4\,^5\) and interlayer magnetism\(^6\). These unusual electronic features emerge as a result of hybridization and the formation of moiré superlattices, and are controlled by the twist angle between the two layers, with potential in the developing field of twistronics\(^6\). Analogous concepts have recently been explored in phononics using atomically thin photonic crystals in bilayer graphene\(^11\,^12\) and in twisted hexagonal boron nitride\(^13\).

Recently, we theoretically proposed that a twisted stack of hyperbolic metasurfaces, each formed by densely packed graphene nanoribbons supporting hyperbolic plasmons\(^14\), might enable an unusual control of the plasmon dispersion—analogous to moiré physics but for photons\(^15\). However, the extreme anisotropy of these artificial hyperbolic metastructures is fundamentally limited by their granularity, inducing strong nonlocality\(^16\,^17\) that may hinder the practical verification of these concepts. In-plane hyperbolicity arises in metasurfaces when the imaginary part of the effective surface impedance along two orthogonal transverse directions have different signs\(^18\,^19\), leading to sub-diffractional surface-wave propagation and enhanced local density of states\(^14\). Notably, phonon polaritons (PhPs)—quasiparticles that arise from collective oscillations between photons and lattice vibrations—can be naturally endowed with a hyperbolic response in polar van der Waals nanomaterials\(^20\,^21\). Although most of these naturally occurring hyperbolic polaritons propagate out-of-plane, it has recently been shown that in-plane hyperbolic PhPs arise in \( \alpha \)-MoO\(_3\) flakes\(^22\,^23\). Within their Reststrahlen band, from 818 cm\(^{-1}\) to 974 cm\(^{-1}\), the real part of the permittivity of \( \alpha \)-MoO\(_3\) is negative along the \([100]\) direction but positive along the \([001]\) direction. Such extreme anisotropy enables in-plane low-loss hyperbolic PhP propagation, as has been experimentally observed recently\(^24\,^25\).

In this work we experimentally demonstrate precise control of the polariton dispersion through tailored interlayer coupling in twisted bilayered (tBL) \( \alpha \)-MoO\(_3\) flakes. Such control is possible because of the emergence of topological PhPs. Their topological character emerges in two ways: first, the dispersion contours of the polaritons change markedly in nature—from hyperbolic (open) to elliptical (closed)—as a function of the twist angle, yielding a photonic topological transition\(^26\) that is governed by the coupling of in-plane hyperbolic PhPs that are individually supported by each layer. In addition, this transition is determined by a topological quantity—that is, the number of anti-crossing points (denoted as \( N_{ACP} \)) of the dispersion lines of each isolated layer.

---

1. Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore. 2. Photonics Initiative, Advanced Science Research Center, City University of New York, New York, NY, USA. 3. Department of Materials Science and Engineering, and ARC Centre of Excellence in Future Low-Energy Electronics Technologies (FLEET), Monash University, Clayton, Victoria, Australia. 4. Melbourne Centre for Nanofabrication, Victorian Node of the Australian National Fabrication Facility, Clayton, Victoria, Australia. 5. Institute of Materials Research and Engineering, A*STAR (Agency for Science, Technology and Research), Singapore, Singapore. 6. Faculty of Materials Science and Chemistry, China University of Geosciences, Wuhan, People’s Republic of China. 7. Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX, USA. 8. Physics Program, Graduate Center, City University of New York, New York, NY, USA. 9. These authors contributed equally: Guangwei Hu, Qingdong Ou. 10. E-mail: qiaoliang.bao@gmail.com, chengwei.qiu@nus.edu.sg, aalu@gc.cuny.edu
Fig. 1 | Rotation-induced topological transition of phonon polaritons.

In reciprocal space—with the isofrequency contours adopting the role of a Fermi surface for electrons, analogous to a Lifshitz transition in electronics\(^{16,37}\). The robust nature of these transitions—which arise when the integer number of anti-crossings \(N_{ACP}\) changes—is not affected by continuous small perturbations or disorder, enabling the experimental observation of the associated phenomena using real-space nanoimaging\(^{38}\). At the twist angle at which the topological transition arises the dispersion must necessarily flatten, as its topology changes from closed to open, yielding diffractionless and low-loss directional PhP canalization\(^{14,29}\). This phenomenon is reminiscent of the recently reported superconductivity at flat bands, which is controlled by twist-angle-dependent interlayer coupling in twisted graphene bilayers\(^{1,2}\).

The thicknesses of the top and bottom layers of our \(\alpha\)-MoO\(_3\) tBLs are denoted \(d_1\) and \(d_2\), respectively (Fig. 1a). A global Cartesian coordinate is defined such that the \(x\) and \(y\) axes are along the [100] and [001] crystal directions of the bottom \(\alpha\)-MoO\(_3\) flake; \(\Delta \theta (\in [−90°, 90°])\) is the twist angle of the [100] crystal direction of the top flake with respect to the \(x\) axis, and is defined as positive in the anticlockwise direction. A representative image of a tBL \(\alpha\)-MoO\(_3\) sample is shown in Fig. 1b. To study this system, we solve source-free Maxwell equations to obtain the modal dispersion, treating each \(\alpha\)-MoO\(_3\) flake as a 2D conductivity sheet (see Methods and Supplementary Information). In this model, we account for interlayer hybridization by making the reasonable assumption that the two surfaces are separated by an infinitesimally thin dielectric spacer (see Methods). For a single layer, hyperbolic PhPs propagate at the open angle \(\beta = \arctan\left(\frac{-\varepsilon_{1x}\varepsilon_{2y} - \varepsilon_{1y}\varepsilon_{2x}}{\varepsilon_{1x}\varepsilon_{2y}}\right)\), where \(\varepsilon_{1x}\) (\(\varepsilon_{1y}\)) is the permittivity component along the [100] ([001]) crystal direction (Extended Data Fig. 1). For tBL flakes, the two hyperbolic bands of the individual layers hybridize and strongly couple to each other at the points in reciprocal space at which they cross, leading to anti-crossing. Notably, the number of anti-crossing points \(N_{ACP}\) in reciprocal space directly determines the hyperbolic or elliptical nature of the iso-frequency contour of tBL \(\alpha\)-MoO\(_3\) (ref. 1). In particular, when \(N_{ACP} = 2\) the bilayer dispersion remains hyperbolic, whereas for \(N_{ACP} = 4\) it becomes elliptical. The topological integer \(N_{ACP}\) is solely determined by the twist angle \(\Delta \theta\), which rotates one dispersion band with respect to the other, and the open angle \(\beta\) of the hyperbolic branches of each isolated layer. Therefore, \(N_{ACP}\) provides control over the bilayer dispersion with inherent topological robustness, as low levels of disorder or small imperfections do not affect this quantity (see Supplementary Information).
sections 2, 3). As summarized in Fig. 1c, $\beta > 45^\circ$ in the frequency range from 860 cm$^{-1}$ to 940 cm$^{-1}$ (Extended Data Fig. 1). We expect hyperbolic dispersion ($N_{\alpha\omega}=2$) when $|\Delta \theta| < |180^\circ - 2\beta|$, and elliptical dispersion ($N_{\alpha\omega}=4$) when $|\Delta \theta| > |180^\circ - 2\beta|$ (Extended Data Fig. 2). Figure 1d illustrates an example at the fixed frequency $\omega = 925.9$ cm$^{-1}$, as numerically verified by full-wave simulations (Fig. 1e–i, Methods). The electric field pattern changes from hyperbolic to elliptical with an increase in the twist angle; this is reflected in the calculated Fourier spectrum and in the analytical dispersion (Fig. 1e–i), with the expected topological transition from open to closed dispersion arising at the critical transition angle ($180^\circ - 2\beta$). At this twist angle, the dispersion flattens and the field becomes highly directional (Fig. 1g, h), reminiscent of the flat Fermi surface that is responsible for magic-angle superconductivity in bilayer graphene$^{13}$. We therefore term this angle a topological transition magic angle. Notably, this form of dispersion engineering, and the open or closed nature of the resulting dispersion bands, does not require that the layers are identical to each other. Applying geometrical variations—for example changing the thickness of the layers (Extended Data Figs. 3–5, Supplementary Information section 3)—reveals the inherent robustness of these transitions associated with their topological nature. We do not imply here that the resulting polaritons are not affected by impurities or disorder in terms of their propagation length or scattering; however, their topological features—that is, the open or closed nature of their dispersion bands and the occurrence of a topological transition at the magic angle at which the integer $N_{\alpha\omega}$ changes—are inherently robust.

To validate our predictions, we prepared tBL samples by using mechanical exfoliation to obtain individual $\alpha$-MoO$_3$ flakes and then transferring one flake onto the other with deterministic alignment and stacking order (Fig. 2b–d, Methods). For experimental observations, we used infrared real-space nanoimaging based on scattering-type scanning near-field optical microscopy (s-SNOM, Fig. 2a). Because it is challenging to excite and image the PhPs in far-field owing to the large mismatch between free-space and polariton momenta, previous studies have used a resonant gold antenna or silver nanowire to excite the PhPs$^{25}$ and a metallic atomic force microscopy (AFM) tip assembled in s-SNOM to couple out and map them. However, using solution-based lithography to fabricate metallic nanostructures on the $\alpha$-MoO$_3$ bilayer could result in damage or contamination$^{30}$. We therefore used a focused ion beam to make a point defect in the form of a hole in the sample (Fig. 2b–d) and excited the PhPs using a metallic tip. The tip-launched PhPs reflect at the point defect and produce polariton interference, yielding near-field images (Extended Data Fig. 6).

As a benchmark, we first investigated PhPs in a single $\alpha$-MoO$_3$ layer—the top layer of the tBL $\alpha$-MoO$_3$ sample with $\Delta \theta = -77^\circ$. This revealed a clear hyperbolic wavefront (Fig. 2e), further confirmed by the hyperbolic Fourier spectrum (Fig. 2i), thus validating our point-defect-assisted excitation and imaging technique. Next, we analysed the response of tBL $\alpha$-MoO$_3$ at the same frequency. The measured wavefront and retrieved dispersion curves remain hyperbolic for the tBL $\alpha$-MoO$_3$ sample with $\Delta \theta = -44^\circ$ (Fig. 2f, j). Such hyperbolicity reduces to a nearly straight PhP wavefront for $\Delta \theta = 65^\circ$ (Fig. 2g, k), yielding a topological transition, and elliptical wavefronts are observed at $\Delta \theta = -77^\circ$ (Fig. 2h, l). The extent to which the dispersion in these twisted bilayers can be tuned is notable, considering that each of them individually supports only hyperbolic polaritons over this frequency range. These observations provide experimental evidence for our theoretically predicted twist-induced topological transition and extreme dispersion engineering controlled by $N_{\alpha\omega}$. Our experimental results were also numerically verified with excellent agreement (Extended Data Fig. 7). The experimentally measured dispersion fits analytical and numerical curves with a discrepancy in wavenumber of a factor of two; this is to be expected because we measured the PhPs reflected by the point defect (Extended Data Figs. 6, 7). All measurements in Fig. 2 measured near-field distribution ($s_2$) near a point defect (PD) denoted by white dashed circles. The image in e is measured near the point defect in the top layer of the sample with $\Delta \theta = -77^\circ$ (marked by dashed squares in d); f–h are measured in the tBL $\alpha$-MoO$_3$ flake in the areas marked by dashed squares in b–d, respectively. The diameter of the point defect is around 200 nm. a–l, Dispersion curves obtained by Fourier transform of the measured near-field signals ($s_2$) in e–h, respectively. For all measurements, $\omega = 903.8$ cm$^{-1}$, a.u., arbitrary units.
Fig. 3 | Tunable topological transition magic angels. a–h, Near-field images (a, c, e, g) and corresponding flattened dispersions obtained by Fourier transform (b, d, f, h) for tBL α-MoO₃ flakes. For a, b, \( \omega = 887.9 \text{ cm}^{-1}, \Delta \theta = -44^\circ \), \( d_1 = d_2 = 128 \text{ nm} \). For c, d, \( \omega = 893.3 \text{ cm}^{-1}, \Delta \theta = 50^\circ , d_1 = 120 \text{ nm}, d_2 = 117 \text{ nm} \). For e, f, \( \omega = 903.8 \text{ cm}^{-1}, \Delta \theta = -63^\circ, d_1 = d_2 = 117 \text{ nm} \). For g, h, \( \omega = 921.7 \text{ cm}^{-1}, \Delta \theta = 79^\circ \), \( d_1 = 223 \text{ nm}, d_2 = 213 \text{ nm} \). The point defects are shown by the white dashed circles in the near-field images and have a diameter of around 200 nm. All near-field images were measured near the topological transition magic angles, showing diffractionless polariton propagation and flattened dispersion.

The measured field distributions and the retrieved dispersion contours are shown in Fig. 3. Topological transitions at the magic angles can be clearly observed, offering a new avenue through which to tune photonic topological transitions at the desired frequency and topologically engineer the polariton dispersion. This differs from previous approaches based on sweeping the excitation frequency, for which the topological transition is typically associated with resonant material damping at the transition frequency.

Close to the magic angle, the isofrequency dispersion contour flattens and PhPs are highly collimated, directive and diffractionless. This is to be expected, because we are operating in the canalization regime, which is characterized by almost-fixed group velocity directions (\( \mathbf{V}_g = \nabla_\mathbf{k} \omega \)). Minor variations in group velocity may still arise, especially for modes with small wavevectors that are more sensitive to realistic imperfections and scattering. Canalization supported by such flat bandstructure has previously been explored for non-diffractive wave

Fig. 4 | Polariton canalization near the topological transition magic angle. a, Near-field images (s) of the sample. Seven edges (abbreviated as ‘E’) and one point defect were fabricated: edge 1 and edge 2 are parallel along the [100] crystal axis of the bottom layer; edge 3 and edge 4 are parallel along the [001] crystal axis of the top layer. Edge 5 is along the [001] crystal axis of the bottom layer. The red arrows correspond to the directions of the calculated group velocity. The inset shows a scanning electron microscopy image of the sample (\( \Delta \theta = -63^\circ, d_1 = d_2 = 117 \text{ nm} \)). Edges with different orientations and one point defect (inside the white square) were fabricated in a high-quality large-area sample. b–e, Expanded view of the measured signals near edges 1, 2, 3 and 4, respectively. f, Calculated dispersion of the sample at \( \omega = 903.8 \text{ cm}^{-1} \). Black dashed lines represent the polariton wavevectors and red solid lines denote the group velocity directions, which are parallel to each other. g, Line plot of measured PhPs along the purple line in a, near the point defect. The red dots are the measured signal and the purple solid line is the Gaussian fitting. h, Line plot of the signal along the blue trajectory in a near edge 6. Red dots are the measured signals and the blue solid line is the fitting.
propagation, which presents opportunities for nanoimaging, radiative energy transfer and enhanced local density of states\textsuperscript{29}. However, its practical realization—based on metamaterials\textsuperscript{29,31} and photonic crystals\textsuperscript{32}—has been hindered by loss, bandwidth restrictions and lack of tunability. Our tunable flatband topological transitions can address these issues and provide polariton canalization over a wide range of frequencies, extending within a region of low-loss propagation. As further experimental evidence, we fabricated multiple edges of the sample for $\Delta \theta \approx 63^\circ$ (see scanning electron microscope image in the inset of Fig. 4a), and then compared polariton propagation in isolated layers (edge 1 and edge 3) and in bilayers (edges 2, 4, 5, 6 and 7). The measured point-defect-scattered PPHs in the bilayer are indeed highly collimated along the reference group velocity direction (Fig. 4a).

The PPHs near the edge all exhibit nearly parallel group velocity, independent of their edge orientations. As a comparison, edge 1 on the bottom-layer $\alpha$-MoO$_3$, and edge 2 on the $\mathrm{TiB}_2\alpha$-MoO$_3$ flakes are parallel. However, edge 1 shows a hyperbolic expanding wavefront (Fig. 4b), whereas edge 2 shows a diffractionless collimated beam (Fig. 4c).

Similar phenomena can be observed by comparing edge 3 (Fig. 4d) and edge 4 (Fig. 4e). Moreover, by carefully examining the PPHs at the end of edges 5, 6 and 7 in Fig. 4a, we find that PPHs consistently propagate along the reference group velocity direction. These features confirm our analytical and numerical results that predict a flat band at $\omega = 903.8$ cm$^{-1}$, corresponding to an almost constant group constant velocity in the proximity of $\theta = 90.3^\circ$. We therefore conclude that the canalization associated with the topological transition, which presents opportunities for robust, diffractionless, low-loss polariton propagation over uniaxial metasurfaces.

To further demonstrate the diffractionless nature of PPHs at the transition angle, we drew a line plot of the signal measured near the point defect that is perpendicular to the reference group velocity direction, as indicated by the purple lines in Fig. 4a. The results and the fit by a Gaussian function (Fig. 4g) give a full-width-at-half-maximum (FWHM) of around $280 \pm 20$ nm (approximately $\lambda_0/40$, where $\lambda_0$ is the free-space wavelength), showing that PPHs are indeed deeply confined. Such a super-resolved channelization PPH mode can be exploited for hyperlensing\textsuperscript{29,31,32}. To characterize the channelization effect we calculated the channelization angle, defined as the relative change of FWHM as the mode propagates (see Extended Data Fig. 9, Supplementary Information section 4). We found this angle to be $6.6^\circ$, which indicates strong polariton canalization despite the presence of realistic loss. To investigate the propagation loss at the observed flattened band, we analyse in Fig. 4f the trajectory of the measured PPH signals along the blue line (parallel to the group velocity direction). After fitting, the measured exponential decay length was approximately $2.77 \mu$m—several times the PPH wavelength—with an inverse damping ratio of around $17$; by contrast, the calculated PPH channelization for the isolated layer—with extreme resonant anisotropy—has a decay length\textsuperscript{33} of approximately $0.35 \mu$m (Extended Data Fig. 10). We therefore conclude that the channelization associated with the topological transition, controlled by the twist angle, can be tuned over a broad range of frequencies for which low-loss collimated polariton propagation is possible. Our results reveal opportunities for robust, diffractionless, low-loss polariton canalization at topological transition magic angles.

We have demonstrated dispersion engineering and polariton topological transitions in TiB$_2\alpha$-MoO$_3$ flakes. The dispersion contours vary from hyperbolic to elliptical over a broad frequency range, and we directly linked the topological nature of the transitions with $N_{\text{ACP}}$ in reciprocal space, as a result of twist-controlled hybridization of the dispersion of each layer. A flat polariton band necessarily emerges at a topological transition, supporting highly collimated diffractionless and low-loss polariton propagation and a tunable PPH channelization mode. Our studies provide an avenue through which to translate recent advances in moiré physics, twistronics and topological engineering in low-dimensional electronic materials to photonics and polaritons. We also reveal a route for twist-angle-controlled dispersion engineering and nanoscale light–matter interactions in natural van der Waals materials and metamaterials\textsuperscript{14,35}, with potential applications in imaging, sensing, radiative energy control and quantum nanophotonics\textsuperscript{16}.
### Article

#### Methods

**Theoretical modelling**

Modelling ultra-thin low-dimensional nanomaterials as a surface conductivity sheet to study their polaritonic properties has been widely adopted.\(^6\) We treat each layer as an anisotropic surface with conductivity \(\sigma = -i \omega \varepsilon_0 \mu_0 \delta^2 \) for large permittivity. \(\delta\) is the thickness of the \(\alpha\)-MoO\(_3\) flake, \(\omega\) is the frequency and \(\varepsilon_0\) is the vacuum permittivity.\(^4\) The relative permittivities along the [100] and [001] directions are denoted as \(\varepsilon_{1\omega}\) and \(\varepsilon_{2\omega}\), respectively. Because each surface is assumed to be of zero thickness, we artificially include in our model a dielectric spacer with relative permittivity \(\varepsilon_d\) and thickness \(d\). Nevertheless, we can take the limit \(d \to 0\) for a realistic scenario. The substrate and superstrate have permittivities of \(\varepsilon_{\text{sub}}\) and \(\varepsilon_{\text{sup}}\), respectively.

The dispersion can be found by solving source-free Maxwell equations for the model associated with the in-plane wavevectors \((k_x, k_y)\). The electromagnetic field in the superstrate, substrate and dielectric spacer can be written as linear independent superpositions of tangential electric (TE) and tangential magnetic (TM) modes. We can then write the boundary conditions for the TE and TM fields in the presence of a surface conductivity sheet and obtain an eigenvalue problem with dispersion given by the \(4 \times 4\) matrix\(^{11}\):

\[
\begin{pmatrix}
\frac{k_{z,\text{sup}}}{k_0} + \frac{k_{z,d}}{k_0} + \frac{\eta_{\text{sup}}}{k_0}
\\
\frac{\eta_{\text{sup}}}{k_0}
\\
\frac{k_{z,d}}{k_0} + \frac{\eta_{\text{sup}}}{k_0}
\\
\eta_{\text{sup}}
\end{pmatrix}
\begin{pmatrix}
\frac{k_{z,\text{sup}}}{k_0}
\\
\frac{\eta_{\text{sup}}}{k_0}
\\
\frac{k_{z,d}}{k_0} + \frac{\eta_{\text{sup}}}{k_0}
\\
\eta_{\text{sup}}
\end{pmatrix}
= 0
\]

Here, \([\sigma_{\text{ph},0}]_{\text{ph},0}\) is the projection of the conductivity tensors in the auxiliary coordinate frame defined by \((\rho, \tilde{z})\), with \(\rho = k_x / k_{z,\text{sup}} + k_y / k_{z,d}, \tilde{z} = k_{\text{sup}} / k_{z,\text{sup}} + k_{\text{sub}} / k_{z,d}\) and \(v = \{\text{sup, sub, } d\}\) denote superstrate (positive sign), substrate (negative signs) and dielectric spacer, respectively. More details and discussion can be found in Supplementary Information section I.

We found that the parameter \(\varepsilon_{\text{d}}\) has a weak influence on the final dispersion, and the thickness \(d\) is important to determine the coupling strength between the two surfaces. To calculate the dispersion in tBL \(\alpha\)-MoO\(_3\) flakes, we simply assume \(\varepsilon_{\text{d}} = 1\) and \(d = 1\) nm, which gives an excellent fit of the dispersion obtained both theoretically (Fig. 1e–i) and experimentally (Extended Data Fig. 7). Besides, the thickness of the dielectric spacer does not have a strong influence on the obtained dispersion (see more discussion in Extended Data Fig. 2 and Supplementary Information section I).

**Numerical simulations**

The finite-difference time-domain method was used in our full-wave simulations and is based on commercially available software (Lumerical FDTD, 2017b, http://www.lumerical.com/tcad-products/fdtd/). To launch highly confined PhP modes, we use a dipole polarized along the \(z\) direction, and the distance between the dipole and the uppermost surface of the sample is 200 nm. The boundary condition is a perfectly matched layer. We monitor the real part of \(E_z\) at 20 nm on top of the uppermost surface of the sample, which is then used as the input of a fast Fourier transform to extract the iso-frequency dispersion contours. For the simulation result in Fig. 1, the sample is placed in free space. For the simulation of realistic samples, we used silicon dioxide as our substrate. The simulation of the field distribution of the sample shown in Fig. 2 is provided in Extended Data Fig. 7, and shows an excellent agreement with the experimental results. The permittivity of \(\alpha\)-MoO\(_3\) layer is obtained by fitting the experimental results in refs.\(^{24,25}\) with a Lorentzian model, as shown in Extended Data Fig. 1. The permittivity of the silicon dioxide substrate is obtained from ref.\(^{41}\).

**Sample preparation**

Bulk crystals of \(\alpha\)-MoO\(_3\) were prepared through chemical vapour deposition, in which the MoO\(_3\) powder (Alfa Aesar) was evaporated at 780 °C in a quartz tube and \(\alpha\)-MoO\(_3\) bulk crystals were then re-deposited at around 350 °C downstream of the tube. Mixed carrier gases of argon (around 200 standard cubic centimetres per minute, sccm) and oxygen (around 50 sccm) were introduced to reduce the possibility of oxygen defects during the redeposition process. We mechanically exfoliated the as-grown bulk crystals onto a polydimethylsiloxane sheet and then inspected the quality of the target \(\alpha\)-MoO\(_3\) flakes via an optical microscope and s-SNOM. Large, homogeneous and ultrathin flakes were then selected and transferred onto Si/SiO\(_2\) (300 nm) substrates.

The tBL \(\alpha\)-MoO\(_3\) sample was prepared through an all-dry viscoelastic stamping method without introducing any chemical contamination. The bottom \(\alpha\)-MoO\(_3\) flake was first transferred onto a Si/SiO\(_2\) substrate, followed by a deterministic transfer of the top \(\alpha\)-MoO\(_3\) under an optical microscope. The top \(\alpha\)-MoO\(_3\) flake was then stacked above the bottom \(\alpha\)-MoO\(_3\) flake such that the crystal axes were rotationally aligned by using a rotation stage.

The point defect in our \(\alpha\)-MoO\(_3\) samples was used to assist the observation of the polaritons. Specifically, the single-layer or tBL \(\alpha\)-MoO\(_3\) samples were milled through with a hole diameter of 200 nm using high-resolution focused ion beam (FIB) lithography in a FEI Helios 600 Nanolab dual-beam FIB–SEM system. Similarly, line defects were fabricated with a length of 4 μm and a width of 200 nm. Note that the milling current was set as low as 28 pA to minimize the implantation of Ga ions, and the milling depth was controlled until reaching the substrate. After FIB etching, the samples were thermally annealed at 300 °C for 2 h to eliminate the influence of the trace Ga-ion implantation and recover the high-profile polariton propagation. Compared with conventional metallic antennas that are used to launch polariton wavefronts at the surface of a van der Waals crystal, the FIB etching technique developed here provides a highly efficient alternative means to observe and study polariton wavefront and propagation, and has advantages such as ultra-high resolution and high-speed dry etching with deterministic positioning and orientation.

**s-SNOM measurements**

We used an s-SNOM (NeaSpec) to perform the infrared nanoimaging, in which a metallic AFM tip (Arrow-NCpt, NanoWorld) was used as both an...
infrared antenna and a near-field probe. With an oscillating frequency of around 300 kHz and an amplitude of around 70 nm, the metallic tip focuses the incident field into a nanoscale hotspot at the apex that interacts with the α-MoO3 sample, when illuminated by p-polarized infrared light of tunable frequencies from a CO2 laser. The twisted bilayer α-MoO3 samples were then raster-scanned below the oscillating tip; the tip-scattered field was recorded by a pseudo-heterodyne interferometer and demodulated at the third harmonic (for example, $s_3$, the near-field amplitude) of the tip resonant frequency. The tip-launched polaritons are reflected and scattered at the sample edges and artificial point/line defects, with interference fringes in the near-field images showing a distance of half the polariton wavelength.

**Data availability**

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

37. Correas-Serrano, D., Aliu, A. & Sebastian Gomez-Diaz, J. Plasmon canalization and tunneling over anisotropic metasurfaces. Phys. Rev. B 96, 075436 (2017).

38. Tilley, D. R. Surface Polaritons: Electromagnetic Waves at Surfaces and Interfaces (Elsevier, 1982).

39. Tretyakov, S. Analytical Modeling in Applied Electromagnetics (Artech House, 2003).

40. Zheng, Z. et al. Highly confined and tunable hyperbolic phonon polaritons in van der Waals semiconducting transition metal oxides. Adv. Mater. 30, e1705318 (2018).

41. Palik, E. D., (ed.) Handbook of Optical Constants of Solids Vol. 3 (Academic, 1998).

**Acknowledgements**

We thank Z. Q. Xu for discussions and help with the sample fabrication. This work was performed at the CUNY Advanced Science Research Center (ASRC), National University of Singapore (NUS) and the Melbourne Centre for Nanofabrication (MCN) in the Victorian Node of the Australian National Fabrication Facility (ANFF). G.H., A.K., Y.M. and A.A. acknowledge support from the Air Force Office of Scientific Research (MURI grant no. FA9550-18-1-0379), the Department of Defense under the Vannevar Bush Fellowship program, the Office of Naval Research (grant no. N00014-19-1-2011), the Simons Foundation and the National Science Foundation. Q.O. acknowledges support from ARC Centre of Excellence in Future Low Energy Electronics Technologies (FLEET). J.W. acknowledges the A*STAR Career Development Award Funding from the Science and Engineering Research Council (grant no. A1820g00087). Z.D. acknowledges the National Natural Science Foundation of China (grant no. 51601131). Q.B. acknowledges support from the Australian Research Council (ARC, FT150100450, CE170100039 and IH150100009). C.-W.Q. acknowledges financial support from A*STAR Pharos Program (grant number 15270 00014, with project number R-263-000-891-305).

**Author contributions**

G.H. and A.A. conceived the idea. G.H., A.A., Y.M., A.K. and C.-W.Q. developed the theory. G.H. performed the simulations and advised in the experimental designs. Q.O. and Q.B. led the experiments. Q.O. designed the structures, and performed the optical measurements. G.S. and Q.O. fabricated the defects and edges on the sample. Y.W. contributed to material synthesis. G.H., Q.O., J.W., Z.D., Q.Z., Q.B., C.-W.Q. and A.A. analysed the data and all authors discussed the results. G.H., A.A., Q.O., Q.B. and C.-W.Q. wrote the manuscript, with input and comments from all authors. A.A., Q.B. and C.-W.Q. supervised the project.

**Competing interests**

The authors declare no competing interests.

**Additional information**

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2359-9.

**Correspondence and requests for materials** should be addressed to Q.B., C.-W.Q. or A.A.

**Peer review information** Nature thanks Zhe Fei, Vincenzo Giannini and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at http://www.nature.com/reprints.
Extended Data Fig. 1 | Properties of a single α-MoO$_3$ layer. a, Permittivity of α-MoO$_3$. $\varepsilon_{11}$ and $\varepsilon_{22}$ are two permittivity tensor components along the [100] and [001] crystal directions, respectively. The permittivity is obtained by fitting the obtained data (refs. 24,25,40) with the Lorentzian model. b, In-plane open angle ($\beta$) of a single layer hyperbolic α-MoO$_3$ flake. c, Isofrequency dispersion of a single α-MoO$_3$ layer (with a thickness of 100 nm) in free space at 925.9 cm$^{-1}$. The red line is the theoretical dispersion, the dashed black lines are the asymptotic line of the hyperbolic dispersion and the green arrow shows the group velocity direction at large wavevector. The open angle therefore illustrates the behaviour of the asymptotic dispersion line as well as the group velocity directions.
Extended Data Fig. 2 | Rotation-induced dispersion engineering of tBL α-MoO₃.

a, b, The crossing of two dispersion lines of the top and bottom layer upon anticlockwise (a) and clockwise (b) rotation. The black line is the dispersion of the bottom layer; the yellow, purple, green, blue and red lines correspond to the dispersion of the top layer at twist angles of ±30°, ±45°, ±70°, ±80° and ±90°, respectively.

c–f, The dispersion of a tBL α-MoO₃ sample \(d₁ = d₂ = 100 \text{ nm}, \epsilon_d = 1\) at twist angles of ±30° (c), ±70° (d), ±80° (e) and ±90° (f). The plotted dispersion lines correspond to the sample of the bottom layer without rotation (solid black lines), the dispersion of the top layer (solid purple lines), the analytical dispersion for tBL α-MoO₃ with \(d_s = 1 \text{ nm}\) (solid red lines), \(d_s = 0.1 \text{ nm}\) (dashed blue lines), \(d_s = 10 \text{ nm}\) (dashed green lines) and \(d_s = 0 \text{ nm}\) (dashed cyan lines).

g–j, The dispersion of a tBL α-MoO₃ sample \(d₁ = d₂ = 100 \text{ nm}, d_s = 1 \text{ nm}\). The plotted dispersion lines correspond to the sample of the bottom layer without rotation (solid black lines), the dispersion of the top layer (solid purple lines) and the analytical dispersion for tBL α-MoO₃ with \(\epsilon_d = 1\) (solid red lines), \(\epsilon_d = 2\) (dashed green lines), \(\epsilon_d = 3\) (dashed blue lines), \(\epsilon_d = 4\) (dashed cyan lines), \(\epsilon_d = 5\) (dashed dark yellow lines) and \(\epsilon_d = 10\) (dashed orange lines).
Extended Data Fig. 3 | Numerical field distributions and dispersions for a single layer α-MoO₃ at ω = 925.93 cm⁻¹. a–j The numerical field distribution (Re(E); top), the numerically obtained dispersions (FFT(Re(E)); contours, bottom) and the analytical dispersions (red dashed lines, bottom), at thicknesses (H) of 40 nm (a), 50 nm (b), 60 nm (c), 80 nm (d), 100 nm (e), 200 nm (f), 300 nm (g), 400 nm (h), 500 nm (i) and 600 nm (j).
Extended Data Fig. 4 | Numerical field distributions and dispersions of twisted bilayer α-MoO₃ at a fixed thickness of $d_2 = 200$ nm. a-e. The numerical field distribution ($\text{Re}(E_z)$; top), the numerically obtained dispersions ($\text{FFT(Re}(E_z))$; contours, bottom) and the analytical dispersions (red dashed lines, bottom) for top-layer thicknesses ($d_1$) of 40 nm (a), 100 nm (b), 200 nm (c), 300 nm (d) and 400 nm (e) at $\omega = 925.93$ cm⁻¹. The dashed light magenta lines, dashed cyan lines and solid red lines are the analytical dispersions of the individual bottom layer, individual top layers and twisted bilayers, respectively. The twist angle is 30° in all cases.
Extended Data Fig. 5 | Numerical result for a top layer of fixed thickness of \( d_1 = 200 \) nm. a–e, The numerical field distribution (\( \text{Re}(E_z) \); top), the numerically obtained dispersions (FFT(\( \text{Re}(E_z) \)); contours, bottom) and the analytical dispersions (red dashed lines, bottom) for bottom-layer thicknesses \( (d_2) \) of 40 nm (a), 100 nm (b), 200 nm (c), 300 nm (d) and 400 nm (e) at \( \omega = 925.93 \) cm\(^{-1} \). The dashed light magenta lines, dashed cyan lines and solid red lines are the analytical dispersions of individual bottom layers, individual top layers and twisted bilayers, respectively. The twist angle is 30° in all cases.
Extended Data Fig. 6 | Mechanism of defect-enabled observation of PhPs.

The metallic AFM tip in the s-SNOM setup could scatter the incident field and launch the highly confined PhPs (referred to as tip-launched PhPs and shown as red/blue solid lines) via near-field light-matter interactions. For case A, the AFM tip is close to the point defect in the sense that the tip-launched PhPs could mostly reach and then be scattered by the point defect (referred to as reflected PhPs and shown as red and blue dashed lines). Thus the wave could be reflected, and strong signals of total PhPs would be expected. For case B, the AFM tip is far from the point defect in the sense that the tip-launched PhPs cannot reach the point defect or the reflected signals cannot be collected by the AFM tip before it is mostly decayed. Therefore, no signal or only weak signals are expected from the reflected PhPs. Note that the optical path should be twice the distance between the tip and point defect, leading to the measured field $e^{-ikr/2\piPD}$. Therefore, when performing the spatial Fourier transform, the obtained spectrum should be of $2k_{\text{PD}}$, that is, the value of the dispersion lines should be multiplied by 2.
Extended Data Fig. 7 | Comparison of experimental and numerical results of the field distribution and the dispersion at $\omega = 903.8$ cm$^{-1}$.

a, Experimentally measured field distribution (left) and dispersion (right) of a single $\alpha$-MoO$_3$ layer of thickness 125 nm (the same as in Fig. 2e, i). 
b, Numerical field distribution (left) launched by the dipole and the obtained dispersion (right) of a single $\alpha$-MoO$_3$ layer of thickness 125 nm. 
c, Experimentally measured field distribution (left) and the dispersion (right) of the tBL $\alpha$-MoO$_3$ flake with $\Delta \theta = -44^\circ$, $d_1 = d_2 = 128$ nm (the same as in Fig. 2f, j). 
d, Numerical field distribution (left) launched by the dipole and the obtained dispersion (right) of the tBL $\alpha$-MoO$_3$ flake with $\Delta \theta = -44^\circ$, $d_1 = d_2 = 128$ nm.

e, Experimentally measured field distribution (left) launched by the dipole and the obtained dispersion (right) of the tBL $\alpha$-MoO$_3$ flake with $\Delta \theta = 65^\circ$, $d_1 = 120$ nm and $d_2 = 235$ nm (the same as in Fig. 2g, k). 
f, Numerical field distribution (left) launched by the dipole and the obtained dispersion (right) of the tBL $\alpha$-MoO$_3$ flake with $\Delta \theta = 65^\circ$, $d_1 = 120$ nm and $d_2 = 235$ nm.

g, Experimentally measured field distribution (left) and dispersion (right) of the tBL $\alpha$-MoO$_3$ flake with $\Delta \theta = -77^\circ$, $d_1 = 125$ nm, and $d_2 = 210$ nm (the same as in Fig. 2h, l). 
h, Numerical field distribution (left) launched by the dipole and the obtained dispersion (right) of the tBL $\alpha$-MoO$_3$ flake with $\Delta \theta = -77^\circ$, $d_1 = 125$ nm and $d_2 = 210$ nm. In the dispersion plot, the red lines correspond to the analytical dispersions ($k_x$, $k_y$) and the white lines correspond to the analytical dispersion with a factor 2—that is, $(2k_x, 2k_y)$. The experimentally obtained dispersion fits better for the white-coloured dispersion $(2k_x, 2k_y)$ owing to doubled optical path via the collection of field by the tip, whereas the numerical one fits better for the dashed red dispersion line $(k_x, k_y)$ as it is propagating wave.
Extended Data Fig. 8 | Observation of the topological transition at \( \omega = 925.9 \text{ cm}^{-1} \).

**a–i.** Optical images (**a, d, g**), experimentally measured field distributions (**b, e, h**), and the obtained dispersions (**c, f, i**) of tBL\( \alpha \)-MoO\(_3\) flakes. For **a–c, \( \Delta \theta = -33^\circ\), \( d_1 = 224 \text{ nm} \) and \( d_2 = 280 \text{ nm} \). For **d–f, \( \Delta \theta = 65^\circ\), \( d_1 = 120 \text{ nm} \) and \( d_2 = 235 \text{ nm} \). For **g–i, \( \Delta \theta = -77^\circ\), \( d_1 = 125 \text{ nm} \) and \( d_2 = 210 \text{ nm} \). These results show that for small twist angles (\( |\Delta \theta| = 33^\circ, 65^\circ\); that is, <72\(^\circ\)), the obtained dispersions are hyperbolic, and for large twist angles (\( |\Delta \theta| = 77^\circ > 72^\circ\)), the dispersion is elliptical.
Extended Data Fig. 9 | The canalization angle at $\omega = 903.8 \text{ cm}^{-1}$. a, The near-field images near the point defect in Fig. 4a. b, The cut-line plots of the PhP signal at different propagation distances along the group velocity $v_g$ from the point defect. c, The extracted FWHM from different line profiles at different propagation lengths, along the group velocity $v_g$. The reference of zero distance is the black line.
Extended Data Fig. 10 | Field canalization and extreme anisotropy at $\omega = 903.8 \text{ cm}^{-1}$.

a, The field distribution ($|E_z|$) of a single $\alpha$-MoO$_3$ flake with a thickness of 100 nm. The permittivity is artificially set as $\varepsilon_{11} = -100 + 100i$, $\varepsilon_{22} = 1.1$ and $\varepsilon_{zz} = 8.9$. The energy mostly propagates along the $x$ axis, which is caused by a flattened band at the extreme anisotropy. 

b, The line plot of the field along the dashed blacked line in a. The red dots are the numerical value and the blue line is the fitting. This plot gives the exponential decay $k_x = 0.35 \mu\text{m}$. 

c, The decay length with respect to $\varepsilon_i$ (the imaginary part of $\varepsilon_{11}$), as $\varepsilon_{11} = -100 + \varepsilon_i \times i$, while the other permittivity components do not change.