

Intelligent infrared sensing enabled by tunable moiré quantum geometry

<https://doi.org/10.1038/s41586-022-04548-w>

Received: 22 September 2021

Accepted: 15 February 2022

Published online: 13 April 2022

 Check for updates

Chao Ma^{1,5}, Shaofan Yuan^{1,5}, Patrick Cheung^{2,5}, Kenji Watanabe³, Takashi Taniguchi⁴, Fan Zhang² & Fengnian Xia^{1,5}

Quantum geometric properties of Bloch wave functions in solids, that is, Berry curvature and the quantum metric, are known to significantly influence the ground- and excited-state behaviour of electrons^{1–5}. The bulk photovoltaic effect (BPVE), a nonlinear phenomenon depending on the polarization of excitation light, is largely governed by the quantum geometric properties in optical transitions^{6–10}. Infrared BPVE has yet to be observed in graphene or moiré systems, although exciting strongly correlated phenomena related to quantum geometry have been reported in this emergent platform^{11–14}. Here we report the observation of tunable mid-infrared BPVE at 5 μm and 7.7 μm in twisted double bilayer graphene (TDBG), arising from the moiré-induced strong symmetry breaking and quantum geometric contribution. The photoresponse depends substantially on the polarization state of the excitation light and is highly tunable by external electric fields. This wide tunability in quantum geometric properties enables us to use a convolutional neural network^{15,16} to achieve full-Stokes polarimetry together with wavelength detection simultaneously, using only one single TDBG device with a subwavelength footprint of merely $3 \times 3 \mu\text{m}^2$. Our work not only reveals the unique role of moiré engineered quantum geometry in tunable nonlinear light–matter interactions but also identifies a pathway for future intelligent sensing technologies in an extremely compact, on-chip manner.

Quantum geometric properties of Bloch wave functions, known as Berry curvature and quantum metric, are crucial in determining the ground-state behaviour of electrons, such as the electric polarization of crystals¹⁷, orbital magnetization^{18,19}, and quantum²⁰ and anomalous²¹ Hall effects. They also play critical roles in the recently discovered superconductivity^{11,13} and ferromagnetism^{11,14} in graphene moiré superlattices. In fact, the bulk photovoltaic effect (BPVE), a nonlinear direct-current photoresponse dependent on light polarization, is also governed by physical quantities associated with Berry curvature and quantum metric^{2–5}. In this work, we report the tunable infrared BPVE in twisted double bilayer graphene (TDBG). Using its tunability, we have generated photovoltage (V_{ph}) mappings that are unique for different polarization states and infrared wavelengths. Furthermore, using a trained convolutional neural network (CNN), we have demonstrated an intelligent TDBG photodetector with a footprint of $3 \times 3 \mu\text{m}^2$ capable of extracting the full Stokes parameters and the wavelength of an incident light simultaneously from the V_{ph} mapping. This device has a unique operational mechanism compared with previous demonstrations. For example, previous miniaturized polarimeters usually use meta-surfaces^{22–24} or anisotropic or chiral materials^{25–27}, in which light is selectively scattered or absorbed depending on its polarization state. However, anisotropic and chiral materials typically cannot provide full-Stokes detection; meta-surfaces require extra photodetectors

and usually have a narrow operational wavelength range owing to their resonant nature.

Transport properties of TDBG device

Figure 1a shows a schematic of our hexagonal boron nitride (hBN) encapsulated TDBG photodetector fabricated using a previously reported deterministic transfer technique (Methods). Independent control of the carrier concentration n and the vertical displacement field D in our TDBG device is realized by tuning both the top-gate (V_{TG}) and back-gate (V_{BG}) voltages (Methods). Here we use the moiré band structures of a 1.2° TDBG at different interlayer potential differences $\Delta V = 0, 50$ and 100 meV (Methods) to illustrate the band structure evolution under D (Fig. 1b). The superlattice-induced insulating states at full fillings of the lowest-energy moiré bands, that is, four electrons or holes per moiré unit cell (denoted by the light green bars) are suppressed with increasing D (central panel of Fig. 1b) and closed at large D (right panel of Fig. 1b). By contrast, a bandgap opens at the charge neutrality point (CNP) and increases with D (marked by the light red bar). Such evolution features have been reported in previous works^{28–32} and are generic for TDBG with a twist angle from -1.1° to 2° . The two-terminal resistance R measured at $T = 79$ K as a function of both V_{TG} and V_{BG} is shown in Fig. 1c. Prominent resistance peaks are observed

¹Department of Electrical Engineering, Yale University, New Haven, CT, USA. ²Department of Physics, The University of Texas at Dallas, Richardson, TX, USA. ³Research Center for Functional Materials, National Institute for Materials Science, Tsukuba, Japan. ⁴International Center for Materials Nanoarchitectonics, National Institute for Materials Science, Tsukuba, Japan. ⁵These authors contributed equally: Chao Ma, Shaofan Yuan, Patrick Cheung. ✉e-mail: shaofan.yuan@yale.edu; zhang@utdallas.edu; fengnian.xia@yale.edu

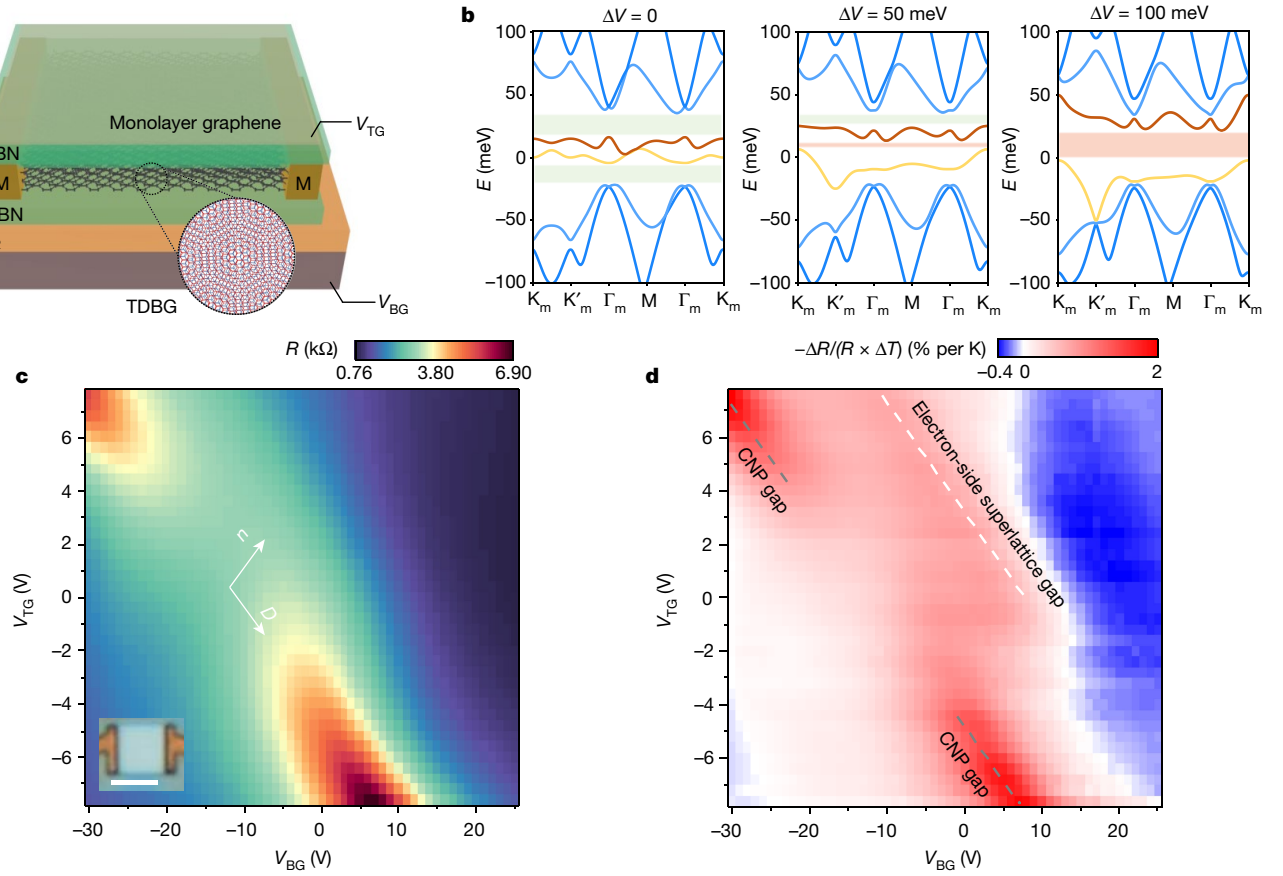


Fig. 1 | Transport properties of the TDBG photodetector. **a**, Schematic of the TDBG photodetector. A graphene monolayer is on top of the hBN encapsulated TDBG, functioning as the top-gate electrode. A degenerately doped silicon substrate is used as the back-gate electrode. Metal contacts (M) are made along the two edges of the stack. Inset: schematic of the TDBG moiré superlattice. **b**, Calculated band structures of 1.2° TDBG with interlayer potential difference ΔV of 0, 50 meV and 100 meV. Only one valley-spin is shown. Light green (red) bars denote the superlattice-induced (charge-neutrality) bandgaps. **c**, Resistance (R) mapping as a function of top (V_{TG}) and bottom (V_{BG}) biasing voltages at $T = 79$ K.

Resistance maxima can be identified at large displacement fields (D) at the CNP. Inset: optical image of the TDBG photodetector. Topmost hBN and monolayer graphene cover the entire image, although they are not visible owing to their high transparency. Scale bar, 3 μm . **d**, Temperature coefficient of resistance ($-\Delta R/(R \times \Delta T)$) mapping as a function of V_{TG} and V_{BG} . The resistance difference ΔR was calculated using two resistance mappings at $T = 79$ K and 84 K. Grey dashed lines denote bandgaps at the CNP under large D , whereas the white dashed line denotes the electron-side superlattice bandgap. The latter is closed under large D .

along the CNP at substantial D , consistent with our calculated band structure evolution and other transport studies^{28–32}. To determine the twist angle of TDBG, we performed temperature-dependent R measurements. The temperature coefficient of resistance (TCR), $-\Delta R/(\Delta T \times R)$, is plotted versus V_{TG} and V_{BG} in Fig. 1d, using the resistance mappings taken at $T = 79$ K and 84 K. Regions with positive TCR are expected to arise from insulating states in which thermal excitation of carriers reduces the resistance. TCR maxima occur at finite D along the CNP (grey dashed lines), corresponding to the field-induced bandgaps shown in Fig. 1b, c. In addition, local maxima of TCR also appear at $n = 3.2 \times 10^{12} \text{ cm}^{-2}$ (white dashed line), implying the presence of the superlattice gap at the electron side. The formation of the TDBG moiré superlattice is further confirmed by the temperature-dependent R measurements (Extended Data Fig. 1). We assign this carrier concentration to be the four-electron full filling and estimate the twist angle to be 1.2°, close to the angle targeted in the stacking.

Tunable BPVE in TDBG

The BPVE is a nonlinear optical phenomenon, consisting of shift (linear BPVE) and injection (circular BPVE) currents generated under linearly and circularly polarized excitations, respectively² (Methods). We first focus on the linear BPVE, which is related to the electric field $\mathbf{E}$ of

incident light through a third-rank nonlinear conductivity tensor σ by⁵ $V_i \propto J_i = \sum_{j,k} \sigma_{ijk} E_j E_k$ to the leading order. Here, V_i , J_i , σ_{ijk} and $E_{j,k}$ are elements of the photovoltage, photocurrent density, nonlinear conductivity and optical field with $i, j, k = \{x, y, z\}$ (Methods). The crystalline symmetry of a material sets strong constraints on its σ (Methods). Figure 2a shows a schematic of TDBG with varying atomic registries including ABBC, ABAC and ABAB. TDBG has three-fold rotation (C_{3z}) and two-fold rotation (C_{2x}) symmetries, and the latter is broken in our devices by D (see the detailed symmetry analysis in the Supplementary Information). Following our symmetry analysis (Methods), we deduce two non-trivial, independent elements, σ_{xxx} and σ_{yyy} , in the conductivity tensor, which give rise to the unique polarization dependence of the nonlinear photovoltage V_{ph} under normal incidence of linearly polarized light (Methods):

$$V_{ph} \propto E^2 (\sigma_{xxx} \cos(2\psi) - \sigma_{yyy} \sin(2\psi)),$$

where ψ is the orientation angle of the in-plane field $\mathbf{E}$ with respect to the voltage collection direction (x axis). The amplitude of the V_{ph} oscillation is proportional to $(\sigma_{xxx}^2 + \sigma_{yyy}^2)^{1/2}$, with a phase ψ_0 determined as $\psi_0 = \arctan(\sigma_{yyy}/\sigma_{xxx})$. Furthermore, from the expression of σ_{xxx} and σ_{yyy} (Methods) it is evident that the quantum geometric properties, that is, the interband Berry connections, play an essential role in

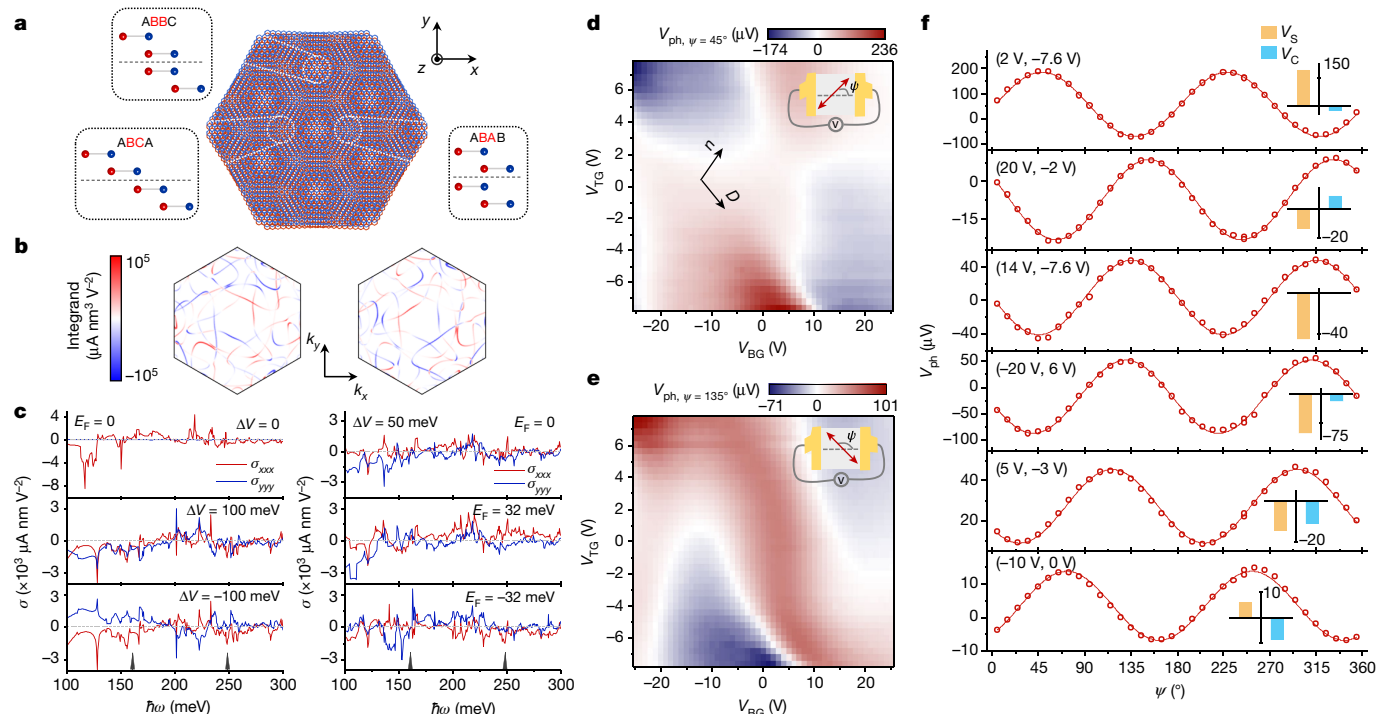


Fig. 2 | Tunable linear BPVE in TDBG. **a**, Schematic of the moiré pattern and atomic registries of TDBG. The red and blue spheres represent the two sublattices (A and B) in each graphene monolayer. The hexagonal moiré pattern is similar to that of twisted bilayer graphene (TBG) but with the AA, AB and BA atomic registries replaced by ABBC, ABBA and ABAB, respectively. **b**, Calculated integrands S_{xxx} (left panel) and S_{yyy} (right panel) used in the evaluation of the nonlinear conductivity elements σ_{xxx} and σ_{yyy} in the mBZ ($\sigma_{xxx(yyy)} = \int d^2k S_{xxx(yyy)}$). Only one valley-spin is shown, and the mBZ centre is the Γ point. The incident photon energy is set to be $\hbar\omega = 161$ meV, corresponding to the wavelength $\lambda = 7.7$ μm . The Fermi level (E_F) and interlayer potential difference (ΔV) are set to be $E_F = 0$ and $\Delta V = 100$ meV. **c**, Calculated nonlinear conductivity elements σ_{xxx} (red) and σ_{yyy} (blue) versus incident photon energy $\hbar\omega$ at $E_F = 0$, $\Delta V = 0, \pm 100$ meV

(left panel) and $\Delta V = 50$ meV, $E_F = 0, \pm 32$ meV (right panel). Contributions from both valleys and both spins are included. Grey dashed lines indicate the zero value. Black triangles denote the photon energies corresponding to the wavelengths $\lambda = 5$ μm and 7.7 μm . **d, e**, Linear bulk photovoltaic voltage (V_{ph}) as a function of V_{TG} and V_{BG} at $T = 79$ K under linearly polarized light ($\lambda = 7.7$ μm) with an orientation angle $\psi = 45^\circ$ (**d**) and $\psi = 135^\circ$ (**e**) with respect to the voltage collection direction. Insets: schematics of the measurement scheme with different light polarizations. **f**, Polarization dependence of V_{ph} at a set of fixed gate voltage biases (V_{BG}, V_{TG}). The data are fitted by using $V_{ph} = V_c \cos(2\psi) + V_s \sin(2\psi) + V_{const}$. The extracted V_c and V_s are plotted in the inset for each oscillation curve. A strong tuning effect of V_c and V_s by gate voltage biases is evident.

determining σ and hence the BPVE. Figure 2b shows the calculated distributions of the integrands S_{xxx} and S_{yyy} , related to σ by $\sigma_{xxx(yyy)} = \int d^2k S_{xxx(yyy)}$, in the moiré Brillouin zone (mBZ) with the Fermi level $E_F = 0$, the interlayer potential difference $\Delta V = 100$ meV and the incident photon energy $\hbar\omega = 161$ meV (7.7 μm). By integrating S_{xxx} and S_{yyy} over the mBZ, we obtain the dependence of σ_{xxx} and σ_{yyy} on $\hbar\omega$ at different E_F and ΔV (Fig. 2c). Notably, both E_F and ΔV can effectively tune the magnitude and the polarity of σ_{xxx} and σ_{yyy} , and hence the resulting linear BPVE. In addition, σ_{xxx} and σ_{yyy} are strong and oscillating in a broad spectral range (Fig. 2c). Unlike any pristine material (Supplementary Information), the large interlayer moiré potential in TDBG leads to abundant states with substantial quantum geometric properties from multiple minibands available for the optical transitions over a wide energy range. Therefore, TDBG is the simplest system for achieving tunable BPVE enabled by moiré quantum geometry (Methods).

In experiments, we measure V_{ph} in our TDBG device under normal light incidence at 7.7 and 5 μm at different V_{TG} and V_{BG} , without applying an external in-plane bias (Methods). Half-wave plates at corresponding wavelengths are used to rotate the polarization of the incident linearly polarized light, and the device is held at $T = 79$ K for all measurements except for the temperature-dependent ones. We first fix ψ and measure V_{ph} as a function of V_{TG} and V_{BG} at 7.7 μm , as shown in Fig. 2d ($\psi = 45^\circ$) and Fig. 2e ($\psi = 135^\circ$). In V_{ph} mappings, tuning V_{TG} and V_{BG} changes both the magnitude and polarity of V_{ph} . Maximum and minimum V_{ph} occur around CNP at large D , and the polarity flips by reversing the direction

of D . Away from the extrema, V_{ph} shows complex dependences on V_{TG} and V_{BG} for both $\psi = 45^\circ$ and 135° . To better reveal the polarization dependence of V_{ph} , we measure V_{ph} as a function of ψ at multiple sets of (V_{BG}, V_{TG}), as shown in Fig. 2f. Each set of measured V_{ph} (red circles) can be well fitted by $V_{ph} = V_c \cos(2\psi) + V_s \sin(2\psi) + V_{const}$. Here V_{const} represents the offset voltage and V_c and V_s are fitting parameters of the polarization-dependent terms, which are plotted in the insets of Fig. 2f. The magnitudes and polarities of both V_c and V_s are largely tunable by gate voltages, leading to the tunable amplitude and phase of V_{ph} , $(V_c^2 + V_s^2)^{1/2}$ and $\psi_0' = \arctan(V_c/V_s)$, respectively. The maximum extrinsic responsivity reaches 3.7 V W^{-1} at 7.7 μm , which is defined as the peak amplitude divided by the power on the device. The converted short-circuit photocurrent response is 0.74 mA W^{-1} . This photoreponse is surprisingly strong in the infrared range^{8,9} considering its second-order nature. Our experimental observation is consistent with our theoretical analysis that two independent nonlinear conductivity elements σ_{xxx} and σ_{yyy} are tunable by E_F and D (Fig. 2c), giving rise to the unique polarization dependence of V_{ph} . Moreover, V_{ph} versus ψ taken under linearly polarized light at 5 μm is plotted in Extended Data Fig. 2. A similar tuning effect on V_{ph} by V_{BG} and V_{TG} is observed, and the magnitudes and polarities of V_c and V_s are different from those under the excitation of 7.7 μm light. Power, spatial and temperature dependences of V_{ph} are presented in Extended Data Figs. 3–5. In addition, a similarly strong, electrically tunable linear BPVE has been observed in another TDBG device (Extended Data Fig. 6). We expect that the linear BPVE is

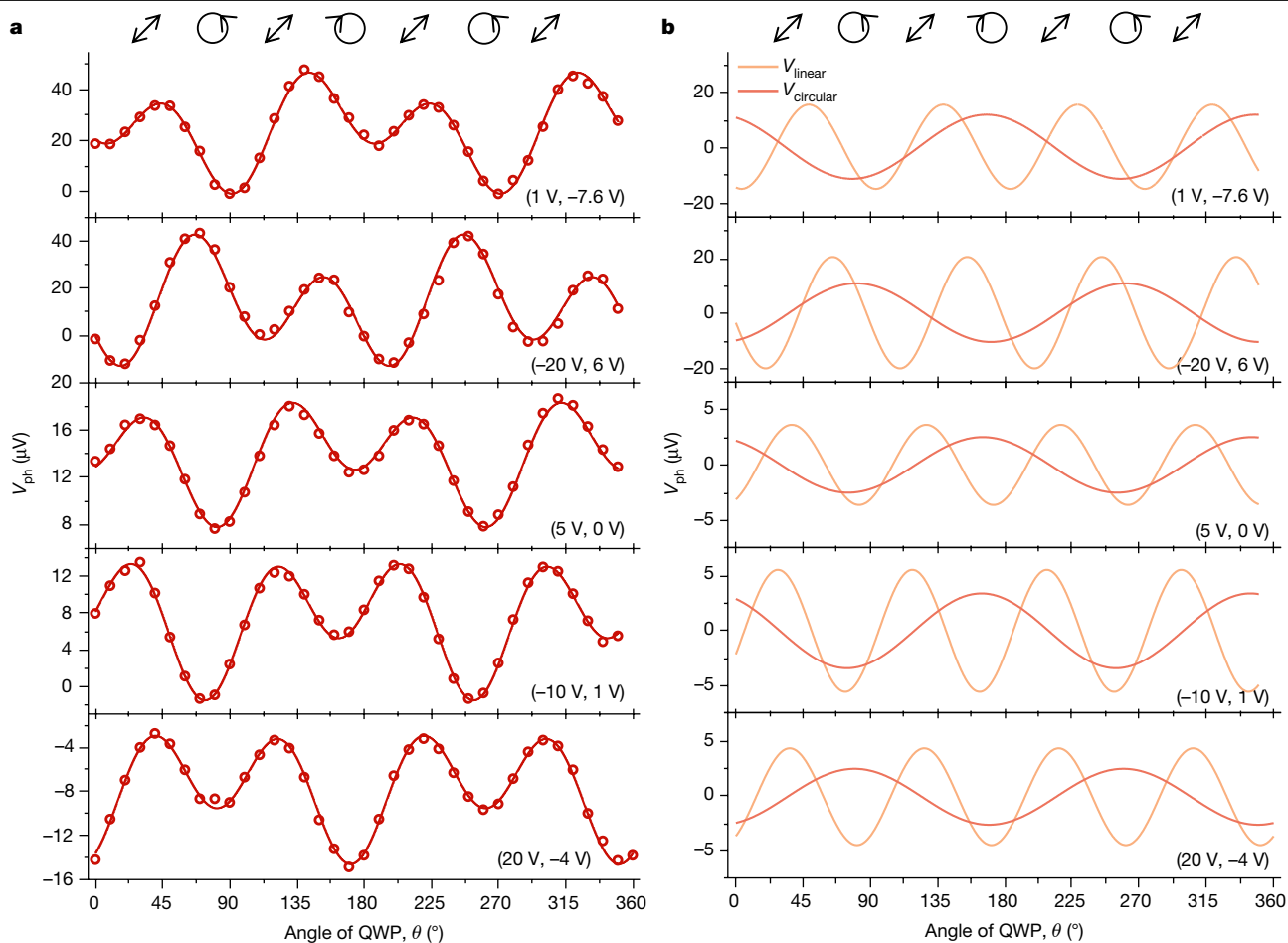


Fig. 3 | Tunable circular BPVE in TDBG. a, Photovoltage (V_{ph}) as a function of the angle of the quarter-wave plate (QWP) (θ) at different gate voltage biases (V_{BG} , V_{TG}), measured at $T = 79$ K and $\lambda = 5$ μ m. The incident light first passes through a half-wave plate (direction of fast axis fixed) and then a quarter-wave plate (direction of fast axis changing on rotation). The polarization states of the incident light are denoted at the top of the figure. The offset of the linear

polarization state from $\theta = 0$ is caused by the half-wave plate, of which the fast axis is approximately 60° away from the z axis in laboratory coordinates. Two components, V_{linear} and $V_{circular}$ with the periodicity of 90° and 180° can be identified by fittings. **b**, Constructed V_{linear} and $V_{circular}$ from fittings, from which a doubled periodicity of $V_{circular}$ compared to V_{linear} is observed. The amplitudes and phases of both components are tunable by gate voltage biases (V_{BG} , V_{TG}).

prominent in TDBG with a wide range of twist angles, as shown by our theoretical calculations (Extended Data Fig. 7).

Furthermore, we investigate the nonlinear photoresponse when the light is modulated between linear and circular polarizations by first passing it through a half-wave plate (fast axis fixed) and then a quarter-wave plate before focusing it onto the TDBG device. The values of V_{ph} measured at 5 μ m versus the angle of the quarter-wave plate θ (the direction of the fast axis with respect to the z axis in laboratory coordinates) at different biasing conditions are shown by the hollow red circles in Fig. 3a. The corresponding polarization states for different θ are denoted at the top. A two-periodicity waveform can be used for fitting the dependence of V_{ph} on θ , $V_{ph} = V_{circular} + V_{linear} + V_{const} = V_1 \cos(2\theta + \theta_0) + V_2 \sin(4\theta + \theta'_0) + V_{const}$, as shown by the solid red lines in Fig. 3a. The component V_{linear} with a 90° periodicity in θ stems from the linear BPVE discussed above (the light resumes linear polarization every 90° when rotating a quarter-wave plate), whereas the other component $V_{circular}$ arises from the circular BPVE that is closely related to the interband Berry curvature dipoles^{6,7,33,34} (Methods). We plot these two fitted components versus θ separately at different gate voltages in Fig. 3b. Similar to the linear BPVE (Fig. 2f), the amplitude of $V_{circular}$ can be tuned whereas the phase of $V_{circular}$ can be switched by 180° by the gate voltages, which implies the electrical tunability of the interband Berry curvature dipoles in TDBG. In principle,

the C_{3z} symmetry of TDBG results in zero in-plane circular BPVE, because the sum of the interband Berry curvature dipoles over the mBZ vanishes. However, possible strain³⁵ or interaction-induced nematic phase³⁶ may weakly break the C_{3z} symmetry, generating the observed circular BPVE (Methods). Further work is needed to clarify its origin. Nevertheless, the simultaneous observations of both linear and circular BPVE in TDBG and their tunability by gate voltages, to our knowledge, are unprecedented (Methods), leading to distinct patterns in the V_{ph} mappings that are both polarization and wavelength dependent.

Intelligent infrared sensing

Next, we discuss how the tunable BPVE can be used for sensing of the polarization, power and wavelength. Figure 4a is a representative V_{ph} mapping with a dimension of 20×26 pixels when the device is excited by elliptically polarized light at 5 μ m, which can be represented by a polarization ellipse with the ellipticity angle $\chi = 36.5^\circ$ and the orientation angle $\psi = 110^\circ$, as shown in the inset. As revealed above, the V_{ph} mapping shows the unique gate dependence. Figure 4b shows V_{ph} as a function of V_{BG} (extracted from corresponding V_{ph} mappings) under the excitation of 5 μ m light with different polarization states at $V_{TG} = 5.2$ V. For a different polarization, the dependence of V_{ph} on V_{BG} is distinct. Likewise, the dependence of V_{ph} on V_{TG} also critically depends

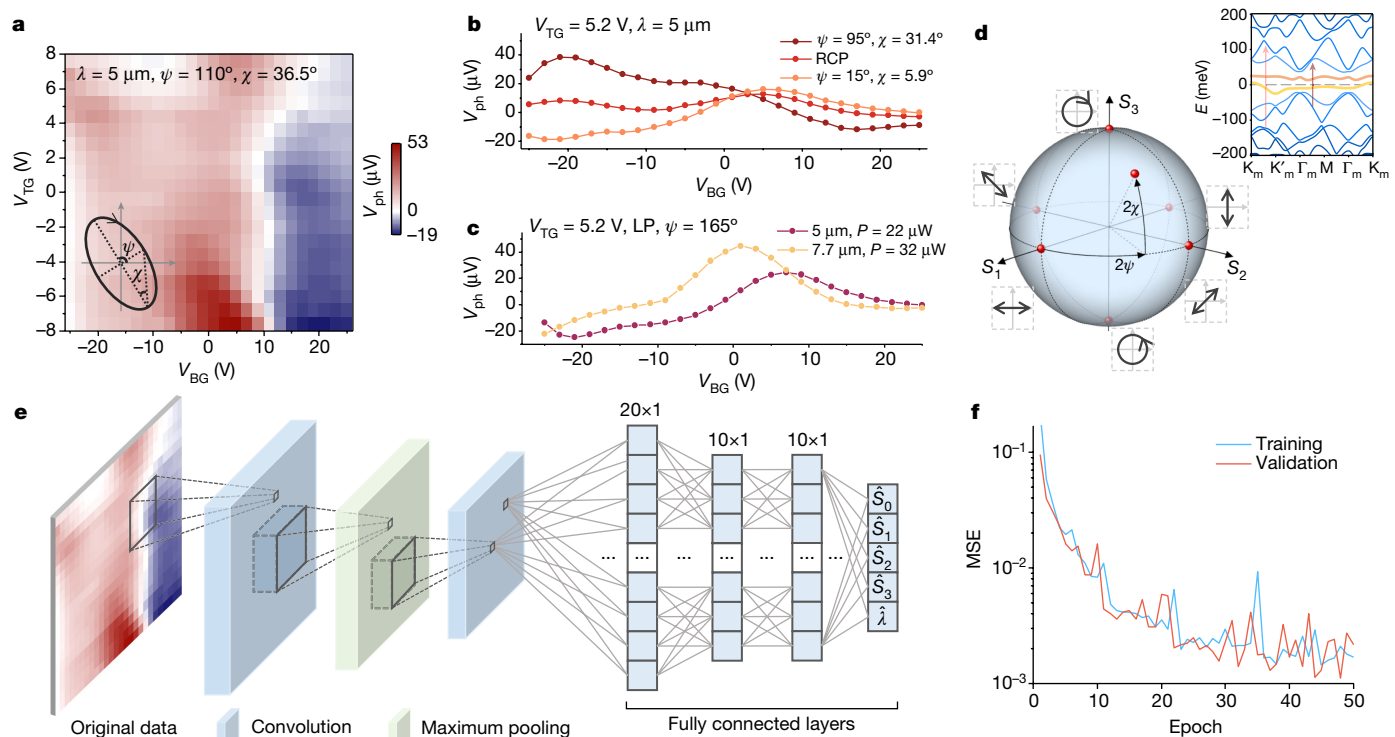


Fig. 4 | Principles of the neural network infrared polarimetry and wavelength detection. **a**, Photovoltage mapping excited by elliptically polarized light at 5 μm , with $\chi = 36.5^\circ$ and $\psi = 110^\circ$. χ and ψ are the ellipticity and orientation angles of the polarization ellipse in the inset, respectively. The photovoltage mapping consists of 20×26 pixels, corresponding to 20 values of the top-gate voltage (V_{TG}) and 26 values of the back-gate voltage (V_{BG}). **b**, Photovoltage as a function of V_{BG} for three representative polarization states at $\lambda = 5 \mu\text{m}$. RCP, right circularly polarized light. **c**, Photovoltage as a function of V_{BG} excited by 5 μm and 7.7 μm linearly polarized (LP) light with $\psi = 165^\circ$ and $V_{\text{TG}} = 5.2 \text{ V}$. **d**, Polarization states and wavelengths read out from photovoltage mappings due to the unique dependences shown in **b** and **c**. Polarization states of the incident light can be described by Stokes parameters and drawn on the Poincaré sphere (bottom left). Six representative polarization states are shown on the Poincaré sphere. The

5 μm and 7.7 μm incident light excite different optical transitions in TDBG (top right), and they can be distinguished by the trained CNN. Two representative transitions for 5 μm (light red arrow) and 7.7 μm (dark red arrow) are labelled on the calculated band structure. **e**, Schematic of the CNN used for the demonstration of the polarimeter and wavelength detection. The input layer is the measured 20×26 -pixel photovoltage mapping, and the output is a five-element vector, $(\hat{S}_0, \hat{S}_1, \hat{S}_2, \hat{S}_3, \hat{\lambda})$. The hidden layers consist of the first convolutional layer, the maximum-pooling layer, the second convolutional layer and three subsequent fully connected layers. Stokes parameters can be directly calculated based on $(\hat{S}_0, \hat{S}_1, \hat{S}_2, \hat{S}_3)$. The wavelength labels $\hat{\lambda} = \pm 1$ correspond to 5 μm and 7.7 μm , respectively. **f**, MSE as a function of training epoch for the training and validation sets. MSE is markedly reduced in the training process.

on the excitation light polarization. Moreover, the initial and final states in an optical transition are closely related to the incident photon energy, which leads to the wavelength dependence of V_{ph} . Figure 4c shows a marked difference in V_{ph} between the 5 μm and 7.7 μm incident light of the same polarization state. Lastly, the power of light can also be deduced from V_{ph} as both the polarization-dependent and -independent components are linearly proportional to the power (Extended Data Fig. 3). Therefore, by evaluating V_{ph} and its gate dependence of the incident light, it is possible to read out its polarization, wavelength and power. The power and polarization state can be described by the well-known Stokes parameters, $S_0 = I$, $S_1 = I p \cos(2\psi) \cos(2\chi)$, $S_2 = I p \sin(2\psi) \cos(2\chi)$ and $S_3 = I p \sin(2\chi)$, where I is the incidence power, p is the degree of polarization, and ψ and χ are the orientation and ellipticity angles (Fig. 4a inset), respectively. Figure 4d shows the Poincaré sphere with six representative polarization states to illustrate this Stokes parameterization. We focus on the 5 μm and 7.7 μm excitations, as illustrated by the possible transitions shown in the inset of Fig. 4d. Decoding the Stokes parameters and wavelength in Fig. 4d from a measured V_{ph} mapping in Fig. 4a is not trivial (Methods). Here the polarization state, power and wavelength of unknown incident light are all encoded in its two-dimensional (2D) V_{ph} mapping implicitly. We resort to a CNN to decipher the physical properties of incident light, as a CNN is most suitable for the interpretation of 2D images^{15,16}. The

structure of the CNN is shown in Fig. 4e. The input layer is a mapping matrix consisting of 20 by 26 V_{ph} values measured at different combinations of V_{TG} and V_{BG} . The hidden layers consist of the first convolutional layer, a maximum-pooling layer, the second convolutional layer and then three fully connected layers^{16,37,38}. The output layer provides a five-component vector (predicted values) including four scaled Stokes parameters $(\hat{S}_0, \hat{S}_1, \hat{S}_2, \hat{S}_3)$ and a wavelength label $\hat{\lambda}$ set to be -1 and 1 for 5 μm and 7.7 μm , respectively (Methods). The training and validation datasets consist of 9,100 V_{ph} mappings generated from 91 measured original sets and their corresponding output vectors (Methods). In the training process, a loss function, which is the mean squared error (MSE) between the measured and predicted values of the output vector plus l2-norm regularization terms¹⁵, is minimized using the Adam optimizer³⁹. As shown in Fig. 4f, the MSEs decrease significantly to around 2×10^{-3} for both sets as the number of training epochs increases. Detailed configurations of the CNN and the training processes are presented in the Methods.

After the CNN is properly trained, the performance of this intelligent light sensor is tested. Representative results, including 12 test and 35 training data, are summarized in Fig. 5 (see Extended Data Fig. 8 for a complete plot from the total 12 test and 91 training data sets). Figure 5a compares the polarization states (three Stokes parameters S_1 – S_3) obtained from the CNN outputs (red spheres for training data and

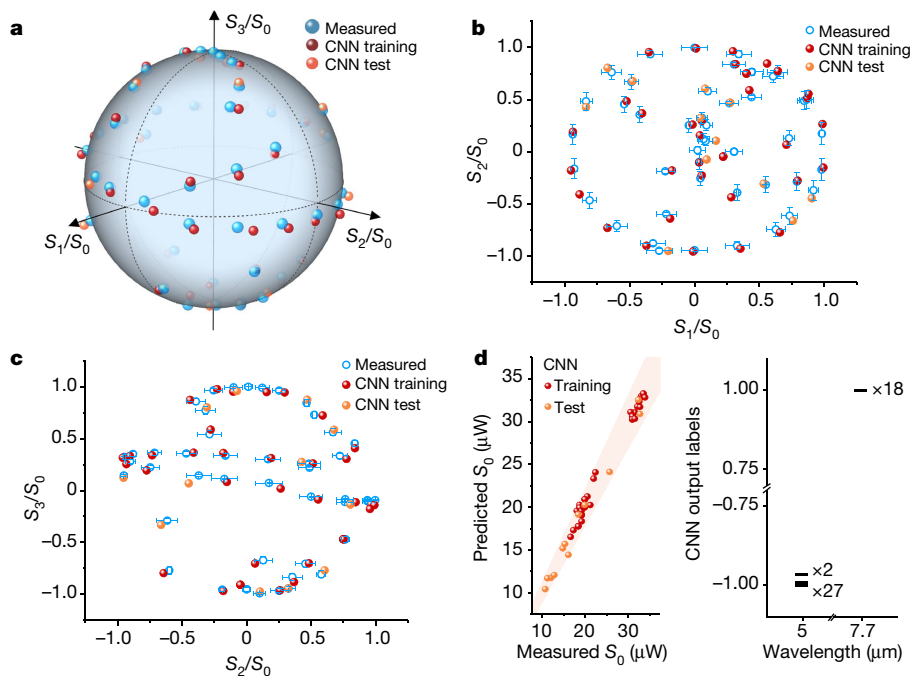


Fig. 5 | Intelligent light sensing by leveraging a CNN. **a**, Polarization states (47 in total, including 35 training and 12 test data) from the output of the CNN training (red spheres) and test (orange spheres) data sets with their corresponding measured values (blue spheres) plotted on the Poincaré sphere. There are 91 original training data in this work. The number of training data was expanded to 9,100 by applying linear extrapolation and adding random noise for CNN training (Methods). All the Stokes parameters are normalized by the power S_0 for clarity. Some of the data points from the CNN do not perfectly reside on the Poincaré sphere, owing to the small deviation of the polarization degree p from 1 in CNN outputs despite having $p = 1$ for all measured data used

orange spheres for test data) with those directly measured values (blue spheres). Evidently, the CNN predicts consistent results for all input polarization states. For better clarity and error analysis, Fig. 5b, c project the results in Fig. 5a onto the S_1 - S_2 and S_2 - S_3 planes, and error bars are included in the measured ψ and χ values (Methods). The left panel of Fig. 5d plots the predicted S_0 values by CNN as a function of the directly measured power. The error in the measured S_0 is estimated to be 10% from multiple measurement results read from the power meter, denoted by the light red area. The right panel of Fig. 5d presents the correctly predicted wavelength labels, in the vicinities of -1 and 1 for $5 \mu\text{m}$ and $7.7 \mu\text{m}$, respectively. To conclude, our implemented CNN provides an accurate prediction of the light polarization, power and wavelength.

Outlook

The utilization of a CNN as a decoder, together with the tunable BPVE of TDBG as an encoder, enables the simultaneous decipherment of light polarization, power and wavelength using a single on-chip device with a subwavelength footprint. Although in this work we focus on the mid-infrared spectral range, we expect that such a nonlinear photoreponse shaped by the quantum geometric properties in TDBG persists down to the terahertz regime (Supplementary Information), and that the demonstrated intelligent graphene sensor can be operational from the mid-infrared to the terahertz range. A similar nonlinear photoreponse may also be prominent at visible or near-infrared wavelengths in hetero- or homo-bilayers of transition metal dichalcogenides because of the analogous symmetry reduction and geometric properties^{11,40}. Our demonstration thus opens up a potential avenue for on-chip intelligent light sensing in a broad spectral range.

in training. **b, c**, The results in **a** are projected onto the S_1 - S_2 (**b**) and S_2 - S_3 (**c**) planes for clarity and error analysis. The error bars are calculated by considering the errors in measuring the orientation angle ψ and the ellipticity angle χ (Methods). **d**, Left panel: predicted power (S_0 , red spheres for the training set and orange spheres for the test set) by the CNN at different incidence powers measured by a commercial infrared detector. The light red area corresponds to a 10% error in power measurements. Right panel: the predicted wavelength labels for the 47 measurements. The $5 \mu\text{m}$ and $7.7 \mu\text{m}$ incident light are expected to have labels -1 and 1 , respectively.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-022-04548-w>.

- Xiao, D., Chang, M.-C. & Niu, Q. Berry phase effects on electronic properties. *Rev. Mod. Phys.* **82**, 1959–2007 (2010).
- Ma, Q., Grushin, A. G. & Burch, K. S. Topology and geometry under the nonlinear electromagnetic spotlight. *Nat. Mater.* **20**, 1601–1614 (2021).
- Cook, A. M., Fregoso, B. M., De Juan, F., Coh, S. & Moore, J. E. Design principles for shift current photovoltaics. *Nat. Commun.* **8**, 14176 (2017).
- Morimoto, T. & Nagaosa, N. Topological nature of nonlinear optical effects in solids. *Sci. Adv.* **2**, e1501524 (2016).
- Sipe, J. & Shkrebtii, A. Second-order optical response in semiconductors. *Phys. Rev. B* **61**, 5337–5352 (2000).
- Ma, Q. et al. Direct optical detection of Weyl fermion chirality in a topological semimetal. *Nat. Phys.* **13**, 842–847 (2017).
- Xu, S.-Y. et al. Electrically switchable Berry curvature dipole in the monolayer topological insulator WTe_2 . *Nat. Phys.* **14**, 900–906 (2018).
- Osterhoudt, G. B. et al. Colossal mid-infrared bulk photovoltaic effect in a type-I Weyl semimetal. *Nat. Mater.* **18**, 471–475 (2019).
- Ma, J. et al. Nonlinear photoresponse of type-II Weyl semimetals. *Nat. Mater.* **18**, 476–481 (2019).
- Akamatsu, T. et al. A van der Waals interface that creates in-plane polarization and a spontaneous photovoltaic effect. *Science* **372**, 68–72 (2021).
- Lau, C. N., Bockrath, M. W., Mak, K. F. & Zhang, F. Reproducibility in the fabrication and physics of moiré materials. *Nature* **602**, 41–50 (2022).
- Cao, Y. et al. Correlated insulator behaviour at half-filling in magic-angle graphene superlattices. *Nature* **556**, 80–84 (2018).
- Cao, Y. et al. Unconventional superconductivity in magic-angle graphene superlattices. *Nature* **556**, 43–50 (2018).

14. Sharpe, A. L. et al. Emergent ferromagnetism near three-quarters filling in twisted bilayer graphene. *Science* **365**, 605–608 (2019).
15. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
16. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2324 (1998).
17. King-Smith, R. & Vanderbilt, D. Theory of polarization of crystalline solids. *Phys. Rev. B* **47**, 1651–1654 (1993).
18. Thonhauser, T., Ceresoli, D., Vanderbilt, D. & Resta, R. Orbital magnetization in periodic insulators. *Phys. Rev. Lett.* **95**, 137205 (2005).
19. Xiao, D., Shi, J. & Niu, Q. Berry phase correction to electron density of states in solids. *Phys. Rev. Lett.* **95**, 137204 (2005).
20. Thouless, D. J., Kohmoto, M., Nightingale, M. P. & den Nijs, M. Quantized Hall conductance in a two-dimensional periodic potential. *Phys. Rev. Lett.* **49**, 405–408 (1982).
21. Nagaosa, N., Sinova, J., Onoda, S., MacDonald, A. H. & Ong, N. P. Anomalous Hall effect. *Rev. Mod. Phys.* **82**, 1539–1592 (2010).
22. Wei, J., Xu, C., Dong, B., Qiu, C.-W. & Lee, C. Mid-infrared semimetal polarization detectors with configurable polarity transition. *Nat. Photonics* **15**, 614–621 (2021).
23. Maguid, E. et al. Photonic spin-controlled multifunctional shared-aperture antenna array. *Science* **352**, 1202–1206 (2016).
24. Rubin, N. A. et al. Matrix Fourier optics enables a compact full-Stokes polarization camera. *Science* **365**, eaax1839 (2019).
25. Tong, L. et al. Stable mid-infrared polarization imaging based on quasi-2D tellurium at room temperature. *Nat. Commun.* **11**, 2308 (2020).
26. Chen, C. et al. Circularly polarized light detection using chiral hybrid perovskite. *Nat. Commun.* **10**, 1927 (2019).
27. Ishii, A. & Miyasaka, T. Direct detection of circular polarized light in helical 1D perovskite-based photodiode. *Sci. Adv.* **6**, eabd3274 (2020).
28. Shen, C. et al. Correlated states in twisted double bilayer graphene. *Nat. Phys.* **16**, 520–525 (2020).
29. Liu, X. et al. Tunable spin-polarized correlated states in twisted double bilayer graphene. *Nature* **583**, 221–225 (2020).
30. He, M. et al. Symmetry breaking in twisted double bilayer graphene. *Nat. Phys.* **17**, 26–30 (2021).
31. Cao, Y. et al. Tunable correlated states and spin-polarized phases in twisted bilayer–bilayer graphene. *Nature* **583**, 215–220 (2020).
32. Burg, G. W. et al. Correlated insulating states in twisted double bilayer graphene. *Phys. Rev. Lett.* **123**, 197702 (2019).
33. de Juan, F., Grushin, A. G., Morimoto, T. & Moore, J. E. Quantized circular photogalvanic effect in Weyl semimetals. *Nat. Commun.* **8**, 15995 (2017).
34. Sodemann, I. & Fu, L. Quantum nonlinear Hall effect induced by Berry curvature dipole in time-reversal invariant materials. *Phys. Rev. Lett.* **115**, 216806 (2015).
35. Arora, A., Kong, J. F. & Song, J. C. Strain-induced large injection current in twisted bilayer graphene. *Phys. Rev. B* **104**, L241404 (2021).
36. Rubio-Verdú, C. et al. Moiré nematic phase in twisted double bilayer graphene. *Nat. Phys.* **18**, 196–202 (2022).
37. Abadi, M. et al. TensorFlow: a system for large-scale machine learning. In *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16)* 265–283 (USENIX Association, 2016).
38. Chollet, F. et al. Keras. GitHub <https://github.com/fchollet/keras> (2015).
39. Kingma, D. P. & Ba, J. Adam: a method for stochastic optimization. Preprint at <https://arxiv.org/abs/1412.6980> (2014).
40. Wilson, N. P., Yao, W., Shan, J. & Xu, X. Excitons and emergent quantum phenomena in stacked 2D semiconductors. *Nature* **599**, 383–392 (2021).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

Methods

Device fabrication processes

The hBN encapsulated TDBG was fabricated using a ‘tear-and-stack’ dry-transfer technique^{41,42}. Half of an AB-stack graphene bilayer was first teared by a hBN flake, then rotated by around 1.3° and used to pick up the remaining part. The entire heterostructure hBN/TDBG/hBN was placed on a SiO₂/Si substrate and then etched into the desired shapes. Electrical contacts were made along the edges with Cr/Au (3 nm/47 nm). Another hBN flake and monolayer graphene were transferred onto the fabricated device, to serve as an additional dielectric layer for gating and the top-gate electrode.

Electrical and optical measurements and error estimations

A dual-gate two-probe scheme was used for photovoltage measurements. The two gate voltages V_{TG} and V_{BG} , applied to the top monolayer graphene and degenerately doped Si substrate, respectively, can independently control the carrier concentration n and displacement field D in TDBG through the following relations⁴³: $n = -(C_{TG}(V_{TG} - V_{TG,0}) + C_{BG}(V_{BG} - V_{BG,0}))/e$ and $D = (-C_{TG}(V_{TG} - V_{TG,0}) + C_{BG}(V_{BG} - V_{BG,0}))/2\epsilon_0$, where C_{TG} (C_{BG}), $V_{TG,0}$ ($V_{BG,0}$), e and ϵ_0 are the top (back) gate capacitance, top (back) offset voltage, elementary charge and vacuum permittivity, respectively. The device was placed in a cryostat filled with argon gas for variable temperature measurements down to 79 K. Infrared light (at 5 or 7.7 μm) from a quantum cascade laser first passed through a half-wave plate and/or a quarter-wave plate at the corresponding wavelength, then was chopped by an optical chopper at 967 Hz and finally was focused on the sample by an infrared microscope. The photovoltage was collected without an external in-plane bias using a lock-in amplifier with reference to the chopping frequency. States of the light polarization described by χ and ψ in the polarization ellipse were measured by a commercial polarizer and a mercury cadmium telluride (MCT) infrared detector. Specifically, the polarizer was used to detect the direction of the major axis of the polarization ellipse. The angle between it and the photovoltage collecting direction defined the orientation angle ψ . By rotating the polarizer and recording the maximum and minimum intensity values from the readings of the MCT detector, denoted as $I_{\max}$ and $I_{\min}$, respectively, we obtained the absolute value of $|\chi| = \arctan(I_{\min}/I_{\max})^{1/2}$. The sign of χ was deduced by examining the angle α between the polarization direction of the incident linearly polarized light and the fast axis of the quarter-wave plate. For $0 > \alpha > 90^\circ$, $\chi > 0$ and for $90^\circ < \alpha < 180^\circ$, $\chi < 0$. The incidence power S_0 was obtained by measuring the total power under a microscope and the beam profile. The measurement error was estimated to be $\delta S_0/S_0 = 10\%$. The other three Stokes parameters S_1 – S_3 were obtained using the measured ψ and χ values and the relations $S_1/S_0 = p \cos(2\psi) \cos(2\chi)$, $S_2/S_0 = p \sin(2\psi) \cos(2\chi)$ and $S_3/S_0 = p \sin(2\chi)$. In this work $p = 1$, as polarized laser light was used. We further took the finite difference to estimate the measurement errors in these quantities. The error in the orientation angle was $\delta\psi = 3^\circ$ from the uncertainty in measuring the direction of the major axis of the polarization ellipse. The error in the ellipticity angle χ was given by $\delta\chi = (S_3/S_0) \cdot \sqrt{2} \delta I/I$, in which $\delta I/I = \delta I_{\max}/I_{\max} = \delta I_{\min}/I_{\min} \approx 10\%$ from the fluctuation of the intensity values read out by the MCT detector. The error bars of the Stokes parameters S_1/S_0 , S_2/S_0 and S_3/S_0 in Fig. 5b, c and Extended Data Fig. 8 were then calculated by taking the finite difference of their expressions presented above and substituting for $\delta\psi$ and $\delta\chi$. Two TDBG devices were measured. All the data presented in this work were taken from TDBG D1 except for that in Extended Data Fig. 6, which was taken from TDBG D2.

CNN structure and training processes

The CNN used in this work was developed based on Tensorflow library³⁷ and Keras interface³⁸. In the CNN (Fig. 4e), the dot product of a learnable filter and its receptive field (the squares in the previous

layers) produces a new element in the convolutional layers, and performing the convolution gives us the whole layer¹⁶. The maximum-pooling layer selects maximum values in the corresponding field as its elements, which retain the major features of the data and reduce their dimensions¹⁶. The weight and bias parameters in the convolutional layers and fully connected layers are optimized in the training process through back propagation. Specifically, the first convolutional layer consisted of 32 kernels (filters), with a size of 5 by 5. The maximum-pooling layer had a pool size of 2 by 2. The second convolutional layer had 16 kernels, with a size of 3 by 3. The rectified linear activation function (ReLU) was used for all hidden layers, and the hyperbolic tangent function was used for the output layer¹⁵. The l2-norm regularizations were applied to the convolutional layers and fully connected layers, and both regularization coefficients were chosen to be 10^{-4} based on the performance in cross validation. The CNN outputs $\hat{S}_i$ ($i = 0, 1, 2, 3$) given by the hyperbolic tangent activation function³⁷ were in the range of -1 to 1 . We used $\hat{S}_0 = S_0/20 - 1$ and $\hat{S}_i = S_i/S_0$, $i = 1, 2, 3$, to map them to -1 to 1 . The value of λ was set to be -1 and 1 for 5 μm and 7.7 μm incidence light, respectively. We measured 91 original V_{ph} mappings for training and validation, and their corresponding output vectors were calculated based on the equations above. To expand the training and validation sets, we applied linear extrapolation and added random noises to generate a total of 9,100 V_{ph} mappings based on the original 91 measured mappings, and they were randomly split into the training set (90%) and the validation set (10%). The extrapolation was appropriate as the overall V_{ph} scaled linearly with the excitation power as shown in Extended Data Fig. 3. These techniques were widely used in machine learning when a larger training set was desired⁴⁴. Acquiring more training data was expected to further improve the performance. In the training process, a loss function was minimized using the Adam optimizer³⁹. The loss function in our case is the MSE between the measured and predicted values of the output vector plus l2-norm regularization terms¹⁵. The directly measured polarization state, wavelength and power (as discussed in the last section) are used as the ground truth in the training. After each epoch, the MSEs are calculated separately for the training and validation datasets. In Fig. 5d, we used $S_0 = (\hat{S}_0 + 1) \times 20$ to recover the power S_0 from $\hat{S}_0$, which was generated from the CNN.

Moiré band structure modelling

The following Hamiltonian for the AB–AB TDBG was used for calculating its band structures⁴⁵ and nonlinear conductivity tensors:

$$H = \begin{pmatrix} h^- + \frac{\Delta}{2} & t^- & 0 & 0 \\ t^+ & h^- + \frac{\Delta}{6} & T(\mathbf{r}) & 0 \\ 0 & T^\dagger(\mathbf{r}) & h^+ - \frac{\Delta}{6} & t^+ \\ 0 & 0 & t^{++} & h^+ - \frac{\Delta}{2} \end{pmatrix},$$

where $h^\pm$ are the low-energy effective Hamiltonians of monolayer graphene near the K point, $t^\pm$ are the interlayer couplings within each AB bilayer graphene, $T(\mathbf{r})$ is the interlayer coupling between the two twisted AB bilayer graphene, Δ is the interlayer potential difference, and the dagger denotes the Hermitian conjugate. In this model,

$h^\pm = \hbar v_0 \begin{pmatrix} 0 & \pi^{\pm\dagger} \\ \pi^\pm & 0 \end{pmatrix}$ and $t^\pm = \begin{pmatrix} -v_4 \pi^{\pm\dagger} & -v_3 \pi^\pm \\ t_1 & -v_4 \pi^{\pm\dagger} \end{pmatrix}$ with $v_i = \frac{\sqrt{3}}{2\hbar} |t_i| a$, a is the graphene lattice constant, $\pi^\pm = \mathbf{k} e^{i(\theta_{\mathbf{k}} \mp \frac{\theta}{2})}$, $\theta_{\mathbf{k}}$ is the momentum orientation angle and θ is the TDBG twist angle. The twisted bilayer part follows the standard Bistritzer–MacDonald continuum model⁴⁶, with the interlayer AA and AB couplings respectively expressed as $w_{AA} = (-0.1835t_1^2 + 1.036t_1 - 0.06736)/3$ and $w_{AB} = t_1/3$. In our

Article

calculations, $\theta = 1.2^\circ$, $t_0 = -3.100$ eV, $t_1 = 0.360$ eV, $t_3 = 0.283$ eV and $t_4 = 0.138$ eV were used⁴⁵. Moreover, a total of 81 moiré reciprocal lattice points were used, resulting in a 648×648 Hamiltonian matrix at each mBZ momentum per spin-valley.

BPVE

With summation notation, the rectified nonlinear current density $\mathbf{J}$ is related to the electric field $\mathbf{E}$ through a third-rank conductivity tensor as follows:

$$J_i = \chi_{ijk}(0; -\omega, \omega) E_j(-\omega) E_k(\omega) = \chi_{ijk}(0; -\omega, \omega) E_j(\omega)^* E_k(\omega),$$

where $i, j, k = \{x, y, z\}$, ω is the frequency of the incident light, and the asterisk denotes the complex conjugation. χ_{ijk} vanishes for inversion symmetric systems. χ_{ijk} can be decomposed into its real ($\Re$) and imaginary ($\Im$) parts, which are symmetric and anti-symmetric, respectively, in the last two indexes. The current density can then be written as⁴⁷

$$J_i = \sigma_{ijk}(E_j(\omega) E_k(\omega)^* + E_j(\omega)^* E_k(\omega))/2 + i\gamma_{il}(\mathbf{E}^*(\omega) \times \mathbf{E}(\omega))_l,$$

where $\sigma_{ijk} = \Re(\chi_{ijk})$, and $\gamma_{il} = \Im(\chi_{ijk})/2$ is a second-rank pseudo-tensor composed of the anti-symmetric part of χ_{ijk} . The two separated terms are responsible for the linear and circular BPVE, respectively.

We first focus on the intrinsic shift current contribution to the linear BPVE. For linearly polarized light with its electric field polarized in the $\hat{j}$ direction, the conductivity tensor can be expressed as^{3,5,48}

$$\sigma_{ij}(0; \omega, -\omega) = \frac{\pi e^3}{\hbar^2} \int \frac{d^2 k}{(2\pi)^2} \sum_{n,m} f_{nm} \Im(r_{mn}^j r_{nm;i}^j) \delta(\omega_{nm} - \omega),$$

where f_{nm} is the difference in Fermi occupation, $r_{nm}^j = i \langle n | \partial_{k_j} | m \rangle$ is the non-abelian Berry connection, $r_{nm;i}^j = \partial_{k_i} r_{nm}^j - i(r_{nn}^j - r_{mm}^j) r_{nm}^j$ is the generalized derivative and $\hbar\omega_{nm} = \hbar\omega_n - \hbar\omega_m$ is the energy difference between the n th and m th bands. $\hbar\omega$ is the energy of the irradiation photons and the integration is carried out in the first mBZ.

The pseudo-tensor for the intrinsic injection current rate can be calculated by^{3,5}

$$\beta_{ij}(\omega) = \frac{\pi e^3}{\hbar^2} \epsilon_{jlk} \int \frac{d^2 k}{(2\pi)^2} \sum_{n,m} f_{nm} \Delta_{nm}^i r_{nm}^l r_{mn}^k \delta(\omega_{nm} - \omega).$$

Under uniaxial heterostrain⁴⁹ (which breaks C_{3z} symmetry and induces circular BPVE), the graphene Dirac cones are shifted such that $\mathbf{k}_{l,\xi} \rightarrow \mathbf{k}_{l,\xi} + [\mathbf{1} + \epsilon_l^T][\mathbf{q} - \mathbf{D}_{l,\xi}]$, where $\mathbf{D}_{l,\xi} = [\mathbf{1} - \epsilon_l^T]\mathbf{K}_{l,\xi} - \xi \mathbf{A}_l$ gives the positions of the Dirac points, $\mathbf{A}_l = \frac{\sqrt{3}\zeta}{2a}(\epsilon_l^{xx} - \epsilon_l^{yy}, -2\epsilon_l^{xy})$ is the induced gauge field, ξ and l denote the valley and layer degrees of freedom, respectively, and

$$\epsilon_l = \mathcal{F}_l \begin{pmatrix} -\cos^2\phi + v\sin^2\phi & (1+v)\cos\phi\sin\phi \\ (1+v)\cos\phi\sin\phi & v\cos^2\phi - \sin^2\phi \end{pmatrix}.$$

We assume that the bottom (b) two and top (t) two layers of the TDBG are strained in the opposite directions, that is, $\mathcal{F}_{b,t} = \mp 1/2$, but with the same strength ϵ . In the simulations, the graphene hopping modulus factor $\zeta = 3.14$, the graphene Poisson ratio $v = 0.165$, $\epsilon = 0\%$, 0.1% and 0.6% , and the strain direction $\phi = 30^\circ$ were used.

Nonlinear conductivity calculations

In evaluating the conductivity tensor $\sigma_{ij}(0; \omega, -\omega)$, a 'sum rule' introduced previously^{3,5} was used such that the differentiations act on the Hamiltonian matrix instead of the Dirac kets:

$$r_{nm}^i = \frac{v_{nm}^i}{i\omega_{nm}}, r_{nm;j}^i = \frac{i}{\omega_{nm}} \left[\frac{v_{nm}^i \Delta_{nm}^j + v_{nm}^j \Delta_{nm}^i}{\omega_{nm}} - w_{nm}^{ij} + \sum_{p \neq n, m} \left(\frac{v_{np}^i v_{pm}^j}{\omega_{pm}} - \frac{v_{np}^j v_{pm}^i}{\omega_{np}} \right) \right]$$

with $v_{nm}^i = \langle u_n | \partial_{k_i} H | u_m \rangle / \hbar$, $\Delta_{nm}^i = \partial_{k_i} \omega_{nm} = v_{nn}^i - v_{mm}^i$ and $w_{nm}^{ij} = \langle u_n | \partial_{k_i} \partial_{k_j} H | u_m \rangle / \hbar$. Moreover, the Dirac delta function was smeared by a Gaussian function g with a smearing parameter W as follows:

$$g(\mathbf{k}) = \frac{1}{\sqrt{2\pi}W} \exp\left(-\frac{(\omega - \omega_{mn}(\mathbf{k}))^2}{2W^2}\right).$$

Given that the moiré bands can be extremely flat, the adaptive smearing approach⁵⁰ was used. In this approach, the smearing parameter is state dependent and given by

$$W_{nm}(\mathbf{k}) = \left| \frac{\partial \omega_n}{\partial \mathbf{k}} - \frac{\partial \omega_m}{\partial \mathbf{k}} \right| \Delta k,$$

where Δk is the distance between the neighbouring sampling points. The mBZ was discretized into 346×346 sampling points, which are sufficiently dense to ensure convergence of the calculations. Contributions to σ were ignored if $|\omega - \omega_{mn}(\mathbf{k})| > 6W(\mathbf{k})$.

Symmetry analysis

In our work, only in-plane photovoltages were collected, and the light was at normal incidence to the material, so we find that $\mathbf{J}$ can be expressed as

$$\begin{pmatrix} J_x \\ J_y \end{pmatrix} = \begin{pmatrix} \sigma_{xxx} & \sigma_{xxy} & \sigma_{xyx} & \sigma_{xyy} \\ \sigma_{yxx} & \sigma_{yyx} & \sigma_{yxy} & \sigma_{yyy} \end{pmatrix} \begin{pmatrix} E_x^2 \\ E_x E_y \\ E_y E_x \\ E_y^2 \end{pmatrix}.$$

In the experimental setup, the current or the electric field directions may not be aligned with the principal axes of the crystal coordinate system (that is, the x and y axes used above), owing to the difficulty in knowing the crystallographic axes of TDBG, but with an angle φ measured from the x axis. In this device coordinate system, denoted by the prime symbol, the current and electric field are related to those in the crystal coordinate system by the following rotations:

$$\begin{pmatrix} J'_x \\ J'_y \end{pmatrix} = R \begin{pmatrix} J_x \\ J_y \end{pmatrix}, \begin{pmatrix} E'_x \\ E'_y \end{pmatrix} = R \begin{pmatrix} E_x \\ E_y \end{pmatrix}, \begin{pmatrix} E'_x E'_y \\ E'_y E'_x \end{pmatrix} = R' \begin{pmatrix} E_x^2 \\ E_x E_y \\ E_y E_x \\ E_y^2 \end{pmatrix},$$

where the rotations R and R' are given by

$$R = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}, R' = \begin{pmatrix} \cos^2\varphi & \cos\varphi\sin\varphi & \cos\varphi\sin\varphi & \sin^2\varphi \\ -\cos\varphi\sin\varphi & \cos^2\varphi & -\sin^2\varphi & \cos\varphi\sin\varphi \\ -\cos\varphi\sin\varphi & -\sin^2\varphi & \cos^2\varphi & \cos\varphi\sin\varphi \\ \sin^2\varphi & -\cos\varphi\sin\varphi & -\cos\varphi\sin\varphi & \cos^2\varphi \end{pmatrix}.$$

In the presence of a C_{3z} symmetry in AB-AB TDBG, there is a symmetry constraint $\sigma = R^{-1} \sigma R'$ with $\varphi = 2\pi/3$ that leads to

$$\sigma = \begin{pmatrix} \sigma_{xxx} & -\sigma_{yyy} & -\sigma_{yyy} & -\sigma_{xxx} \\ -\sigma_{yyy} & -\sigma_{xxx} & -\sigma_{xxx} & \sigma_{yyy} \end{pmatrix}.$$

There are only two independent elements for σ . For an electric field of magnitude E that makes an angle ψ above the x axis, the photocurrent density J exhibits a 2ψ dependence with respect to the optical field polarization:

$$J_x = E^2 (\sigma_{xxx} \cos 2\psi - \sigma_{yyy} \sin 2\psi), J_y = E^2 (-\sigma_{xxx} \sin 2\psi - \sigma_{yyy} \cos 2\psi).$$

It follows that the polarization for the maximum or minimum photovoltage ($V \propto J$) depends on the ratio between σ_{xxx} and σ_{yyy} ; it changes with the top and bottom gate voltages in our experiments. In the numerical calculations of σ_{xxx} and σ_{yyy} (Fig. 2c), the crystal coordinate system is used. In the device coordinate system, if the x' axis (the photocurrent and photovoltage collection direction) makes an angle φ with the x axis, $J_{x'} = J_x \cos \varphi + J_y \sin \varphi$ and $J_{y'} = -J_x \sin \varphi + J_y \cos \varphi$, equivalent to a change of variable $2\psi \rightarrow 2\psi + \varphi$ in the current density equations above.

With an extra mirror symmetry $\mathcal{M}_x$ (the symmetry group C_{3z} is enlarged to C_{3v}), any element in the conductivity tensor that contains an odd number of the index x should vanish, resulting in

$$\sigma = \begin{pmatrix} 0 & -\sigma_{yyy} & -\sigma_{yyy} & 0 \\ -\sigma_{yyy} & 0 & 0 & \sigma_{yyy} \end{pmatrix}.$$

In this case, there is only one independent non-zero element, and the photocurrent is given by

$$J_x = -E^2 \sigma_{yyy} \sin 2\psi, J_y = -E^2 \sigma_{yyy} \cos 2\psi.$$

In contrast to the C_{3z} case, the polarization for the maximum or minimum photovoltage does not change with σ_{yyy} or the top and bottom gate voltages in experiments. The analysis for the mirror symmetry $\mathcal{M}_y$ is similar. Note that both mirror symmetries are broken in AB-AB TDBG.

In AB-AB TDBG, there is a C_{2x} symmetry, in addition to the C_{3z} symmetry. As a result, any element in the conductivity tensor with an odd number of the index y should vanish, leaving only one independent non-zero element as shown in Fig. 2c (top left panel):

$$\sigma = \begin{pmatrix} \sigma_{xxx} & 0 & 0 & -\sigma_{xxx} \\ 0 & -\sigma_{xxx} & -\sigma_{xxx} & 0 \end{pmatrix}.$$

Without any external bias to break the C_{2x} symmetry, the polarization for the maximum or minimum photovoltage does not change with σ_{xxx} . However, the top and bottom gate voltages in our experiments break the C_{2x} symmetry. The analysis for the C_{2y} symmetry is similar.

A C_{2z} symmetry is present in twisted bilayer graphene (TBG) but not in TDBG or other graphene moiré systems. As a consequence of the C_{2z} symmetry, any element in the conductivity tensor with an odd number of the index x or y should vanish. Therefore, $\sigma = 0$ in TBG.

For circularly polarized light at normal incidence to a 2D material, $\mathbf{E}(\omega) \times \mathbf{E}^*(\omega)$ is in the z direction. As a result, the in-plane injection current cannot be generated in TDBG, given its intrinsic C_{3z} symmetry. However, strain effects³⁵ and an interaction-induced nematic phase³⁶ may weakly break the C_{3z} symmetry and account for our observation of the non-vanishing circular BPVE in TDBG as shown in Fig. 3. The exact origin of the observed circular BPVE deserves further research.

Detailed discussions on BPVE under different symmetries and associated calculations can be found in the Supplementary Information.

Uniqueness of TDBG in BPVE

The BPVE in TDBG is unique in the following three ways. First, the moiré-engineered C_{3z} symmetry enables the aforementioned σ_{xxx} and

σ_{yyy} to be non-trivial and independent. This gives rise to linear BPVE with electrically tunable amplitude and phase in the V_{ph} oscillation. Notably, an additional symmetry such as C_{2x} (in TDBG with $\Delta V = 0$), C_{2z} (present in TBG but absent in all other graphene moiré systems) or vertical mirror (in non-moiré graphene systems) dictates that σ_{xxx} or/and σ_{yyy} vanish, whereas the inversion symmetry dictates all conductivity elements to be zero (Supplementary Information). Second, both the amplitude and phase of the BPVE response in TDBG are strongly tunable by external electric fields because ΔV can largely modify the band structure (Fig. 1b) and wave functions of TDBG, and thus its quantum geometric properties; varying E_f in TDBG can select moiré minibands with different quantum geometric properties for the generation of BPVE. Third, the widely tunable BPVE in TDBG is strong in a broad spectral range. Unlike any pristine material (Supplementary Information), the large interlayer moiré potential and the moiré-induced Brillouin zone folding in TDBG lead to abundant states from multiple minibands being available for the interband optical transitions over a wide energy range. This is manifested in the strongly oscillating behaviour of the calculated σ_{xxx} and σ_{yyy} with $\hbar\omega$ (Fig. 2c), as even a small change in $\hbar\omega$ can alter the optical transitions substantially. To summarize, the demonstrated BPVE is universal to graphene moiré systems with no C_{2z} symmetry, and TDBG is the simplest.

Essential role of the CNN

To read out the information of an unknown incident light from a V_{ph} mapping is challenging, because light polarizations and power are represented by the continuous Stokes parameters, and because wavelength is also continuous (although in this work we used only two wavelengths). Direct quantitative modelling of the nonlinear photoresponse to match the experimentally measured V_{ph} is not feasible, because of the inevitable complexity in TDBG devices, such as lattice relaxation and extrinsic scattering. Inferring the information of the incident light based on analytical fittings (Fig. 2f and Fig. 3a) is not practical because it requires numerous polarization-dependent measurements at multiple different gate biases, wavelengths and power levels for a reliable assessment. Moreover, to extract Stokes parameters and wavelength by directly comparing the V_{ph} mapping under a testing light with those already measured under the excitation of light with known information is highly unreliable, as the number of possible polarization states is, in principle, infinite and the incidence wavelength and power can also vary. By contrast, machine learning methods are well known for their capability of completing specific tasks without being specifically programmed⁵¹. Machine learning techniques have been used for many applications including imaging recognition¹⁵, ultrafast optics⁵² and analysis of the angular momentum of light⁵³. Among all machine learning methods, the CNN is most suitable for the interpretation of 2D images, such as object recognition and classification^{15,16}, and indeed the polarization state, power and wavelength of each incident light are all encoded uniquely in its 2D V_{ph} mapping. Introducing convolutional layers to the neural network greatly reduces the number of training parameters and retains the essential features in the input data¹⁵. After the training of the CNN with a reasonably large number of experimentally measured V_{ph} mappings under the incident light with known polarizations, powers and wavelengths, the CNN can predict these continuous variables based on a new V_{ph} mapping under the excitation of an unknown light.

Temperature dependence of the resistance measured in TDBG D1

To confirm the presence of the superlattice-induced insulating state and thus the formation of the TDBG moiré superlattice, we show the temperature-dependent resistance (R) at selected carrier concentrations (n) and displacement fields (D) in TDBG D1. Six sets of gate biases (marked by coloured dots) are selected in Extended Data Fig. 1a (same as Fig. 1c with additional coloured dots), and their $\ln(R)$ as functions of $100/T$ are plotted in Extended Data Fig. 1b. For the resistances

measured at charge neutrality ($n = 0$) and at $D = 0.54 \text{ V nm}^{-1}$ (red dot) and -0.54 V nm^{-1} (orange dot), we observe insulating behaviours, that is, R increases with a decrease in temperature, implying that there exist bandgap openings at charge neutrality under D fields. Similarly, at $n = 3.2 \times 10^{12} \text{ cm}^{-2}$ (the white dashed line assigned to the electron-side superlattice gap) and small displacement fields, $D = 0$ (light orange dot) and 0.15 V nm^{-1} (yellow dot), we also observe insulating behaviours, indicating the presence of a superlattice-induced gap. Moreover, at the same concentration $n = 3.2 \times 10^{12} \text{ cm}^{-2}$ but a large field, $D = 0.6 \text{ V nm}^{-1}$ (the light blue dot in Extended Data Fig. 1a), the resistance increases slightly with elevated temperature (light blue circles in Extended Data Fig. 1b). This observation implies that the superlattice gap is closed and the TDBG is metallic. In addition, introducing more electrons also turns the TDBG into a metallic state again (see, for example, the dark blue dot in Extended Data Fig. 1a corresponding to $n = 4.8 \times 10^{12} \text{ cm}^{-2}$ and $D = 0.2 \text{ V nm}^{-1}$). These observations are consistent with our theoretical results and other previously reported experiments^{28–32}, supporting our conclusion that the electron-side superlattice-induced gap exists at $n = 3.2 \times 10^{12} \text{ cm}^{-2}$ and the TDBG moiré superlattice is successfully formed. We note that the hole-side full filling $n = -n_s$ is not evident in the TCR mapping, and that similar features of a less insulating hole-side superlattice gap were also reported previously^{28–32}.

Linear BPVE under 5 μm light excitation in TDBG D1

Extended Data Figure 2 shows V_{ph} as a function of ψ measured under 5 μm linearly polarized light excitation at different sets of gate voltages ($V_{\text{BG}}, V_{\text{TG}}$). Similar to our observations at 7.7 μm , the photoresponse is highly dependent on the polarization, which can be fitted by using the formula $V_{\text{ph}} = V_c \cos(2\psi) + V_s \sin(2\psi) + V_{\text{const}}$. Gate voltages ($V_{\text{BG}}, V_{\text{TG}}$) significantly modify the polarization dependent components V_c and V_s , which are shown in the insets of Extended Data Fig. 2. Another notable feature is that these components are also wavelength dependent. For instance, the curve shown in the third panel of Fig. 2f (under 7.7 μm excitation) has $V_c \approx 0$ and $V_s \approx -45 \mu\text{V}$, whereas the curve shown in the fifth panel of Extended Data Fig. 2 (under 5 μm excitation) has $V_c \approx -13 \mu\text{V}$ and $V_s \approx -19 \mu\text{V}$, although they were measured under very similar gate voltages. As discussed in the main text, the amplitude and the phase of the V_{ph} oscillation curve are $(V_c^2 + V_s^2)^{1/2}$ and $\psi'_0 = \arctan(V_c/V_s)$ respectively, which lead to drastically different photoresponses for the two incident wavelengths. This wavelength dependent V_{ph} enables the wavelength detection in our tunable TDBG photodetector using the CNN.

Power dependence of the photovoltage V_{ph}

Extended Data Figure 3a shows the V_{ph} as a function of the orientation angle ψ under the 7.7 μm linearly polarized light at different excitation powers. The results from two sets of ($V_{\text{BG}}, V_{\text{TG}}$), (0 V, -7.6 V) and (14 V, -7.6 V), are shown. As shown in Extended Data Fig. 3b, the amplitude of the V_{ph} oscillation, $(V_c^2 + V_s^2)^{1/2}$, linearly depends on the incident power for all four sets of gate voltage combinations in this experiment. The linear-in-power dependence confirms the second-order nature of the polarization-dependent V_{ph} , which is consistent with previous observations of BPVE^{6–10}. Moreover, we examine the power dependence of the overall V_{ph} , including both the polarization-dependent $(V_c^2 + V_s^2)^{1/2}$ and the polarization-independent V_{const} components, as shown in Extended Data Fig. 3c, d. A linear relation between the overall V_{ph} and the incident power is evident for different polarizations and gate voltages. This feature implies that the polarization-independent component V_{const} is also linearly dependent on the power, consistent with our argument that it mainly comes from the photothermal effect, which depends linearly on the incident power (as discussed in the next section). In addition, this linear dependence of V_{ph} on power provides the basis for the expansion of the training and validation data sets as discussed above.

Spatial dependence of the photovoltage V_{ph}

In our V_{ph} measurements, a knife-edge technique is used to determine the size of the laser spot and we assume a Gaussian profile for the intensity distribution. The standard deviation σ of the laser intensity profile is approximately 7 μm and 11 μm for 5 μm and 7.7 μm incident light, respectively. Such beam spots are larger than our TDBG photodetector ($3 \times 3 \mu\text{m}^2$). Although a large beam spot can reduce the contribution from the photo-thermal effect due to the enhanced light intensity uniformity within the device area, the photo-thermal voltage arising from non-uniform illumination can still play a role, as discussed in previous works^{2,6–9}. As discussed below, the photo-thermal effect produces a polarization-independent background in the measured photoresponse V_{ph} , which is minimized when aligning the light spot with the centre of the device.

In general, the photo-thermal voltage V_{Th} can be modelled as^{54–56} $V_{\text{Th}} = (S_2 - S_1)\Delta T$. Here $S_{1,2}$ is the Seebeck coefficient of material 1 or 2, and ΔT is the temperature difference induced by light illumination. In our TDBG photodetector, the photothermal-effect-induced photovoltage mainly occurs at the two contact regions, given by $V_{\text{Thermal}, \pm} = \pm (S_{\text{TDBG}} - S_{\text{metal}})\Delta T$, because of the difference in the Seebeck coefficients of the TDBG and metal electrodes. Ideally, when the beam spot is placed at the centre of the TDBG, the overall voltage induced by the photothermal effect at the two metal–TDBG interfaces, $V_{\text{Thermal}} = V_{\text{Thermal}, +} + V_{\text{Thermal}, -}$, vanishes in a symmetric device because of the cancellation of the photo-thermal voltages at the two interfaces. However, a deviation of the beam spot from the device centre, together with unintentional asymmetry in the contact regions, can lead to a non-vanishing V_{Thermal} in the measured V_{ph} . As the photo-thermal voltage is proportional to the local laser power, the dependence of V_{Thermal} on the beam spot position x with respect to the centre of the device along the voltage collecting path can be modelled by^{7,8} $V_{\text{Thermal}} = A_1 (e^{\frac{(x-x_0)^2}{2w^2}} - e^{\frac{(x+x_0)^2}{2w^2}})$. Here $x_0 = 1.5 \mu\text{m}$ is half of the length of the device, x is the displacement of the beam spot with respect to the centre of the device (shown in the inset of Extended Data Fig. 4a), A_1 is proportional to the absorbed power of the incident light and $w \approx 7 \mu\text{m}$ is the fitting parameter representing the standard deviation of the Gaussian beam profile along the x axis. This model accounts for the contributions from the two interfaces and is thus an odd function of x . By contrast, at a fixed polarization, the BPVE photovoltage V_{BPVE} is proportional to the light power and can be modelled by^{7,8} $V_{\text{BPVE}} = A_2 e^{-\frac{x^2}{2w^2}}$, which is an even function of x . In Extended Data Fig. 4a, we plot the measured V_{ph} as a function of the displacement x of the laser beam with respect to the device centre. The simulated V_{Thermal} (blue line) and V_{BPVE} (black line) are also shown in Extended Data Fig. 4a, and their summation (red line) matches the experimental results well. This observation indicates that the photo-thermal effect can indeed contribute to V_{ph} in the measured mappings, because of the imperfect beam alignment with respect to the device centre and possible device asymmetry. However, it will not affect the device performance owing to its insensitivity to polarization.

In addition, we also performed measurements of the spatial dependence of V_{ph} in the direction perpendicular to the photovoltage collection pathway (y axis). As shown in Extended Data Fig. 4b, red squares are measured V_{ph} at different laser spot positions along the y axis ($x = 0$ and $y \neq 0$). By contrast with the spatial dependence along the x axis (Extended Data Fig. 4a), the measured V_{ph} shows a symmetric dependence on the displacement y and reaches its maximum at $y = 0$, which can be fitted by the black curve $V = A' e^{-\frac{y^2}{2w'^2}}$, where V is the photovoltage, A' is the peak amplitude and $w' \approx 6 \mu\text{m}$. No obvious odd-function component in the response is observed. The deviation from the centre along the y axis reduces the overall incident power on the device, and both the BPVE and photo-thermal voltages are proportional to the incident power on the device. This observation is very different from

moving the light spot along the x axis, in which the asymmetric illumination at two metal–graphene interfaces is introduced and the photo-thermal voltage can flip the sign. Overall, the spatial dependence of the V_{ph} in our work is consistent with previous works^{7,8}.

Furthermore, we note that the photo-thermal effect in TDBG is expected to be independent of polarization, as the absorption of TDBG is isotropic. Thus, the polarization-independent components in the measured V_{ph} can be mainly attributed to $V_{Thermal}$ in our TDBG devices. This property is very different from that of previously reported Weyl semimetal systems^{2,8,9}, in which the absorptions can be significantly different in different directions, giving rise to a polarization-dependent $V_{Thermal}$.

Temperature dependence of the photovoltage V_{ph}

Extended Data Figure 5a shows the measured V_{ph} versus the orientation angles ψ at different temperatures and gate voltages under the linearly polarized, 7.7 μm light excitation. The overall V_{ph} shows a strong temperature dependence and decreases to nearly zero above 250 K. By plotting $(V_c^2 + V_s^2)^{1/2}$ as a function of the temperature (Extended Data Fig. 5b), we observe an approximately ten times smaller oscillation amplitude from 79 K to 250 K at all sets of gate voltage combinations. Such a reduction in the BPVE signal can be attributed to the shortening of the relaxation time of hot carriers excited by photons because of stronger scatterings at higher temperatures^{57,58}. At a steady state, the nonlinear photoresponse is proportional to the relaxation time and decreases with increasing temperature, as reported by previous works on BPVE in Weyl semimetals^{6,9}. In addition, thermal smearing of resonant optical transitions could also reduce the nonlinear photoresponse at higher temperatures.

Linear BPVE in TDBG D2

Here we show the experimental data from another TDBG with a twist angle of around 1.1° (TDBG D2). Extended Data Figure 6a–d shows zero in-plane bias photovoltage V_{ph} mappings as a function of top- (V_{TG}) and back- (V_{BG}) gate biases under the excitation of 5 μm linearly polarized light at orientation angles of $\psi = 40^\circ, 130^\circ, 0^\circ$ and 90° , all measured at $T = 79$ K. In the V_{ph} mappings taken at different light polarizations, there is a strong tuning effect of V_{ph} both in the magnitude and polarity by the gate biases, similar to that in the device (D1) as shown in the main text (Fig. 2d, e). Each of the V_{ph} mappings taken at a certain polarization shows a distinct pattern that encodes the information of the polarization state. The V_{ph} dependence on the orientation angle ψ at different sets of gate biases is shown in Extended Data Fig. 6e. Measured V_{ph} (red circles) can be well fitted by $V_{ph} = V_c \cos(2\psi) + V_s \sin(2\psi) + V_{const}$. The polarization-dependent components V_c and V_s are extracted and plotted in the insets of Extended Data Fig. 6e. A strong tuning effect of V_c and V_s by gate biases is evident, and it is consistent with our theoretical prediction that two independent, non-vanishing, tunable nonlinear conductivity elements exist in TDBG. The amplitude of V_{ph} , that is, $(V_c^2 + V_s^2)^{1/2}$, as a function of the power on the device is plotted in Extended Data Fig. 6f, in which the linear dependence confirms the second-order nature of the photoreponse. In addition, as shown in Extended Data Fig. 6g, the amplitude of V_{ph} strongly depends on the temperature. In brief, all the major observations on the linear BPVE are consistent with those measured with D1 as presented in the main text.

Twist-angle-dependence of the linear BPVE

Here we discuss the twist angle dependence of the linear BPVE. We have calculated the nonlinear conductivity elements, σ_{xxx} and σ_{yyy} , in TDBG with twist angles of 1.2°, 1.4°, 1.6° and 1.8°. Extended Data Fig. 7a–d shows σ_{xxx} and σ_{yyy} as functions of $\hbar\omega$ when E_F is placed in the CNP gap and the electron-side superlattice gap, all at $\Delta V = 50$ meV. Because the band structure of TDBG depends on the twist angle, the nonlinear conductivity elements determined by the quantum geometric properties

and the available interband transitions are tunable by the twist angle. In Extended Data Fig. 7e, f, we plot σ_{xxx} and σ_{yyy} at $\hbar\omega = 161$ meV and 248 meV (corresponding to the wavelengths 7.7 μm and 5 μm used in the experiments) against the twist angle. The conductivity elements are substantial for all four twist angles. In brief, the twist angle is another degree of freedom with which to tune the BPVE in TDBG.

Complete plots of the polarization outputs

In Extended Data Fig. 8 we plot the polarization outputs S_1 – S_3 with their corresponding measured values of the total 103 CNN data sets (91 training and 12 test) used in this work. In Fig. 5 of the main text, we select 35 sets from these training data sets and plot them together with the 12 test data sets for better clarity.

Device performance metrics

Here we discuss the responsivity, response time, noise equivalent power and polarization resolution of the TDBG photodetectors. In our work, the BPVE can be largely tuned by electrical biases, and thus the extrinsic responsivity R_{ex} , defined as the amplitude of the V_{ph} oscillation, $(V_c^2 + V_s^2)^{1/2}$, divided by the power on the device, is tunable by (V_{BG}, V_{TG}) . The maximum R_{ex} in our measurement reaches approximately 3.7 V W^{-1} at 7.7 μm in TDBG D1, 3.1 V W^{-1} at 5 μm in TDBG D1 and approximately 3.3 V W^{-1} at 5 μm in TDBG D2. These values are considerable, given that our observed BPVE is a second-order photoresponse, and even comparable to the regular first-order photoresponse in monolayer graphene with a similar structure⁵⁹. According to our theoretical calculations (Supplementary Fig. 2b), the BPVE should remain substantial in a wide range of the spectrum and peaks in the terahertz regime. (Strong shift current response at terahertz frequencies has also been theoretically predicted in TBG with broken C_{2z} symmetry^{60,61}.) This observation shows the feasibility of constructing a high-sensitivity, broadband intelligent sensor based on TDBG.

In Extended Data Fig. 9a, we show the photovoltage V_{ph} as a function of the intensity modulation frequency, measured at $(V_{BG}, V_{TG}) = (2 \text{ V}, -7.6 \text{ V})$ and linearly polarized light of 7.7 μm and $\psi = 45^\circ$. We find that V_{ph} remains at around 170 μV and does not show any decrease up to 10 kHz (the highest available chopping frequency with our measurement instrument).

Next, we present the noise characteristic of our TDBG device. Extended Data Figure 9b shows the voltage noise spectral density measured at $T = 79$ K and $(V_{BG}, V_{TG}) = (2 \text{ V}, -7.6 \text{ V})$ without light illumination in TDBG D1. We observe that the $1/f$ noise dominates in the low-frequency range; however, the voltage noise quickly levels off at around 11 $\text{nV Hz}^{-1/2}$ beyond 200 Hz. The thermal noise, which can be estimated by $\sqrt{4k_B TR}$, where k_B is the Boltzmann constant, gives a value of around 5 $\text{nV Hz}^{-1/2}$ at $T = 79$ K and $R = 5,333 \Omega$ at $(V_{BG}, V_{TG}) = (2 \text{ V}, -7.6 \text{ V})$. The difference between the measured value and the actual noise floor is attributed to the noise from the instrumentation used in the measurements. Furthermore, we use the thermal noise of approximately 5 $\text{nV Hz}^{-1/2}$ and the responsivity of approximately 3.7 V W^{-1} (at 7.7 μm , as discussed above) of our TDBG device to compute the noise equivalent power, resulting in a value of approximately 1.4 $\text{nW Hz}^{-1/2}$. This noise equivalent power value is smaller than that of a regular graphene photodetector based on the linear photoresponse⁵⁹.

Finally, we discuss the polarization detection resolution in our intelligent light sensing scheme. The polarization state of an incident light can be represented by the Stokes parameters $(S_1/S_0, S_2/S_0, S_3/S_0)$, whereas the CNN predicted values are $(\hat{S}_1, \hat{S}_2, \hat{S}_3)$. Thus, any two unknown polarization states whose Stokes parameters differences are smaller than $(\Delta S_1, \Delta S_2, \Delta S_3) = (S_1/S_0 - \hat{S}_1, S_2/S_0 - \hat{S}_2, S_3/S_0 - \hat{S}_3)$ cannot be accurately distinguished by our scheme. We notice that $(\Delta S_1, \Delta S_2, \Delta S_3)$ depend on the position of the polarization state on the Poincaré sphere, that is, the actual values of $(S_1/S_0, S_2/S_0, S_3/S_0)$. The complete information about $(\Delta S_1, \Delta S_2, \Delta S_3)$ for all measured states in our work can be read out from the projection S_1 – S_2 and S_2 – S_3 planes as shown in Fig. 5b, c and Extended Data Fig. 8.

Data availability

The data that support the plots within this paper are available from the corresponding authors upon request. Source data are provided with this paper.

Code availability

The code and CNN used in this paper are available from the corresponding authors upon request.

41. Wang, L. et al. One-dimensional electrical contact to a two-dimensional material. *Science* **342**, 614–617 (2013).
42. Kim, K. et al. van der Waals heterostructures with high accuracy rotational alignment. *Nano Lett.* **16**, 1989–1995 (2016).
43. Zhang, Y. et al. Direct observation of a widely tunable bandgap in bilayer graphene. *Nature* **459**, 820–823 (2009).
44. Shorten, C. & Khoshgoftaar, T. M. A survey on image data augmentation for deep learning. *J. Big Data* **6**, 60 (2019).
45. Chebrolu, N. R., Chittari, B. L. & Jung, J. Flat bands in twisted double bilayer graphene. *Phys. Rev. B* **99**, 235417 (2019).
46. Bistritzer, R. & MacDonald, A. H. Moiré bands in twisted double-layer graphene. *Proc. Natl Acad. Sci. USA* **108**, 12233 (2011).
47. Hosur, P. Circular photogalvanic effect on topological insulator surfaces: Berry-curvature-dependent response. *Phys. Rev. B* **83**, 035309 (2011).
48. von Baltz, R. & Kraut, W. Theory of the bulk photovoltaic effect in pure crystals. *Phys. Rev. B* **23**, 5590–5596 (1981).
49. Bi, Z., Yuan, N. F. Q. & Fu, L. Designing flat bands by strain. *Phys. Rev. B* **100**, 035448 (2019).
50. Yates, J. R., Wang, X., Vanderbilt, D. & Souza, I. Spectral and Fermi surface properties from Wannier interpolation. *Phys. Rev. B* **75**, 195121 (2007).
51. Samuel, A. L. Some studies in machine learning using the game of checkers. *IBM J. Res. Dev.* **3**, 210–229 (1959).
52. Genty, G. et al. Machine learning and applications in ultrafast photonics. *Nat. Photon.* **15**, 91–101 (2021).
53. Giordani, T. et al. Machine learning-based classification of vector vortex beams. *Phys. Rev. Lett.* **124**, 160401 (2020).
54. Gabor, N. M. et al. Hot carrier-assisted intrinsic photoresponse in graphene. *Science* **334**, 648–652 (2011).
55. Xu, X., Gabor, N. M., Alden, J. S., Van Der Zande, A. M. & McEuen, P. L. Photo-thermoelectric effect at a graphene interface junction. *Nano Lett.* **10**, 562–566 (2010).
56. Song, J. C. W., Rudner, M. S., Marcus, C. M. & Levitov, L. S. Hot carrier transport and photocurrent response in graphene. *Nano Lett.* **11**, 4688–4692 (2011).
57. Frenzel, A. J., Lui, C. H., Shin, Y. C., Kong, J. & Gedik, N. Semiconducting-to-metallic photoconductivity crossover and temperature-dependent Drude weight in graphene. *Phys. Rev. Lett.* **113**, 056602 (2014).
58. Ni, G. X. et al. Fundamental limits to graphene plasmonics. *Nature* **557**, 530–533 (2018).
59. Yuan, S. et al. Room temperature graphene mid-infrared bolometer with a broad operational wavelength range. *ACS Photon.* **7**, 1206–1215 (2020).
60. Kaplan, D., Holder, T. & Yan, B. Momentum shift current at terahertz frequencies in twisted bilayer graphene. Preprint at <https://arxiv.org/abs/2101.07539> (2021).
61. Chaudhary, S., Lewandowski, C. & Refael, G. Shift-current response as a probe of quantum geometry and electron-electron interactions in twisted bilayer graphene. Preprint at <https://arxiv.org/abs/2107.09090> (2021).

Acknowledgements F.X., S.Y. and C.M. acknowledge support by the National Science Foundation EFRI NewLaw programme under grant number 1741693 and the Yale Raymond John Wean Foundation. F.Z. and P.C. acknowledge the Texas Advanced Computing Center (TACC) for providing resources that have contributed to the research results reported in this work. F.Z. and P.C. acknowledge support by the Army Research Office under grant number W911NF-18-1-0416 and by the National Science Foundation under grant numbers DMR-1945351 through the CAREER program, DMR-1921581 through the DMREF program and DMR-2105139 through the CMP program. Growth of hBN crystals by K.W. and T.T. was supported by the Elemental Strategy Initiative conducted by the Japan Ministry of Education, Culture, Sports, Science and Technology (MEXT; grant number JPMXP0112101001), the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (JSPS KAKENHI; grant number JP20H00354) and the Japan Science and Technology Agency Core Research for Evolutional Science and Technology (CREST; grant number JPMJCR15F3).

Author contributions C.M., S.Y., F.Z. and F.X. conceived the project. C.M. and S.Y. fabricated the devices, performed the measurements and developed the program for applying CNN. P.C. and F.Z. performed the theoretical analysis and calculations. K.W. and T.T. synthesized the hBN crystals. F.Z. and F.X. supervised the project. C.M., S.Y., P.C., F.Z. and F.X. analysed the results and wrote the manuscript.

Competing interests C.M., S.Y., P.C., F.Z. and F.X. are evaluating the feasibility of a patent application based on the concepts and results in this work with their intellectual property offices.

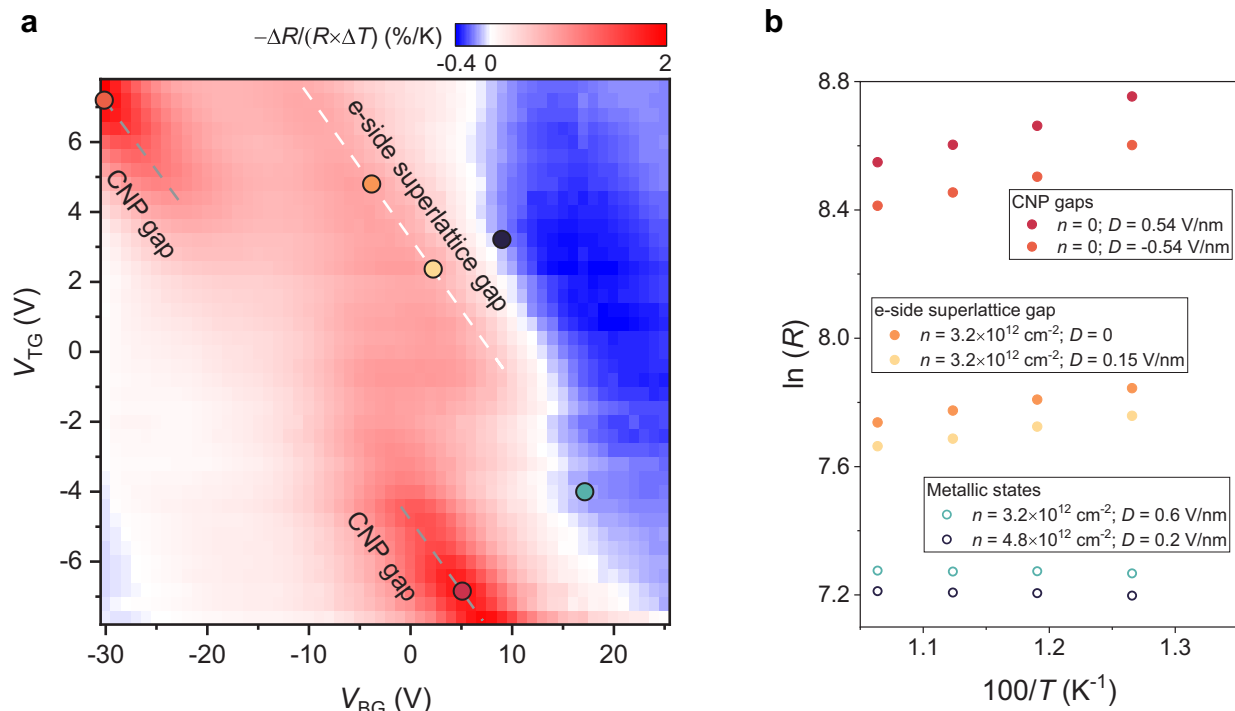
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-022-04548-w>.

Correspondence and requests for materials should be addressed to Shaofan Yuan, Fan Zhang or Fengnian Xia.

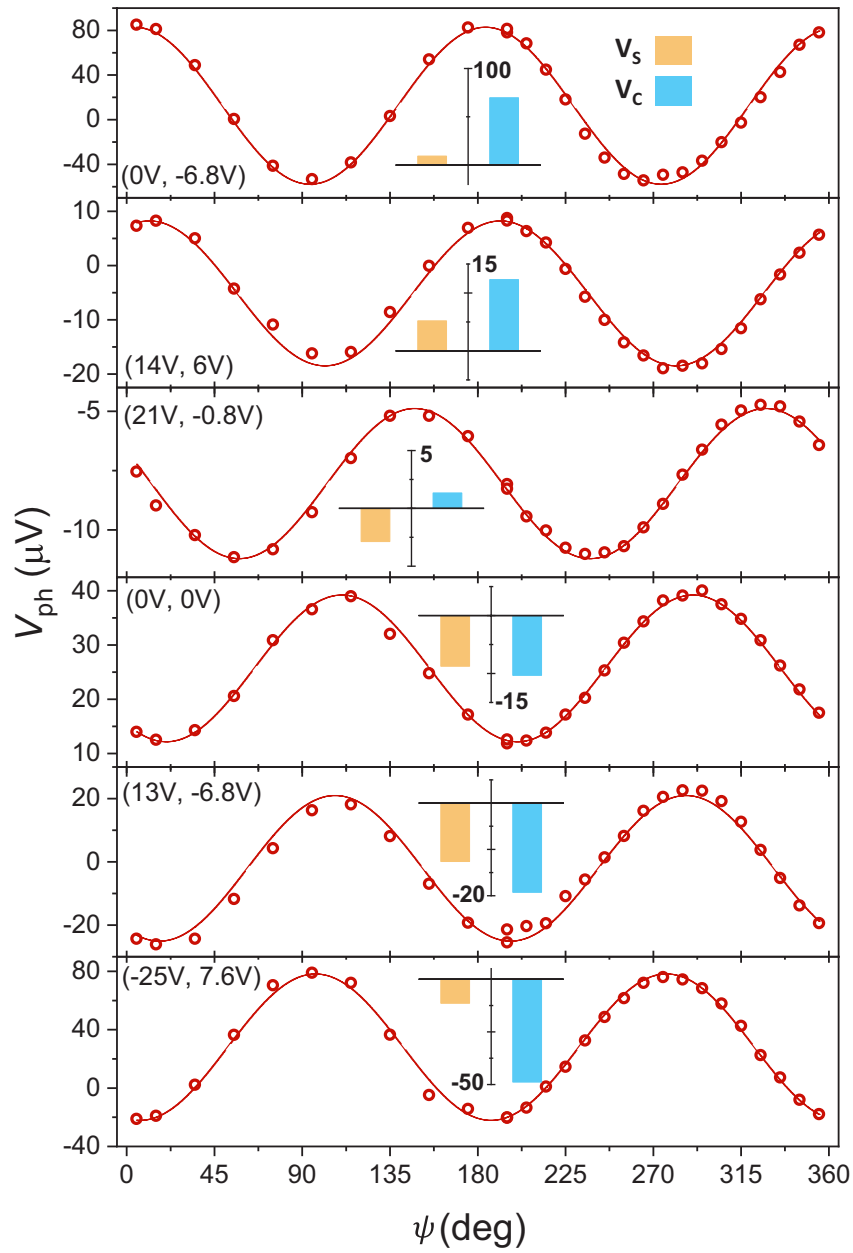
Peer review information *Nature* thanks Chengkuo Lee and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

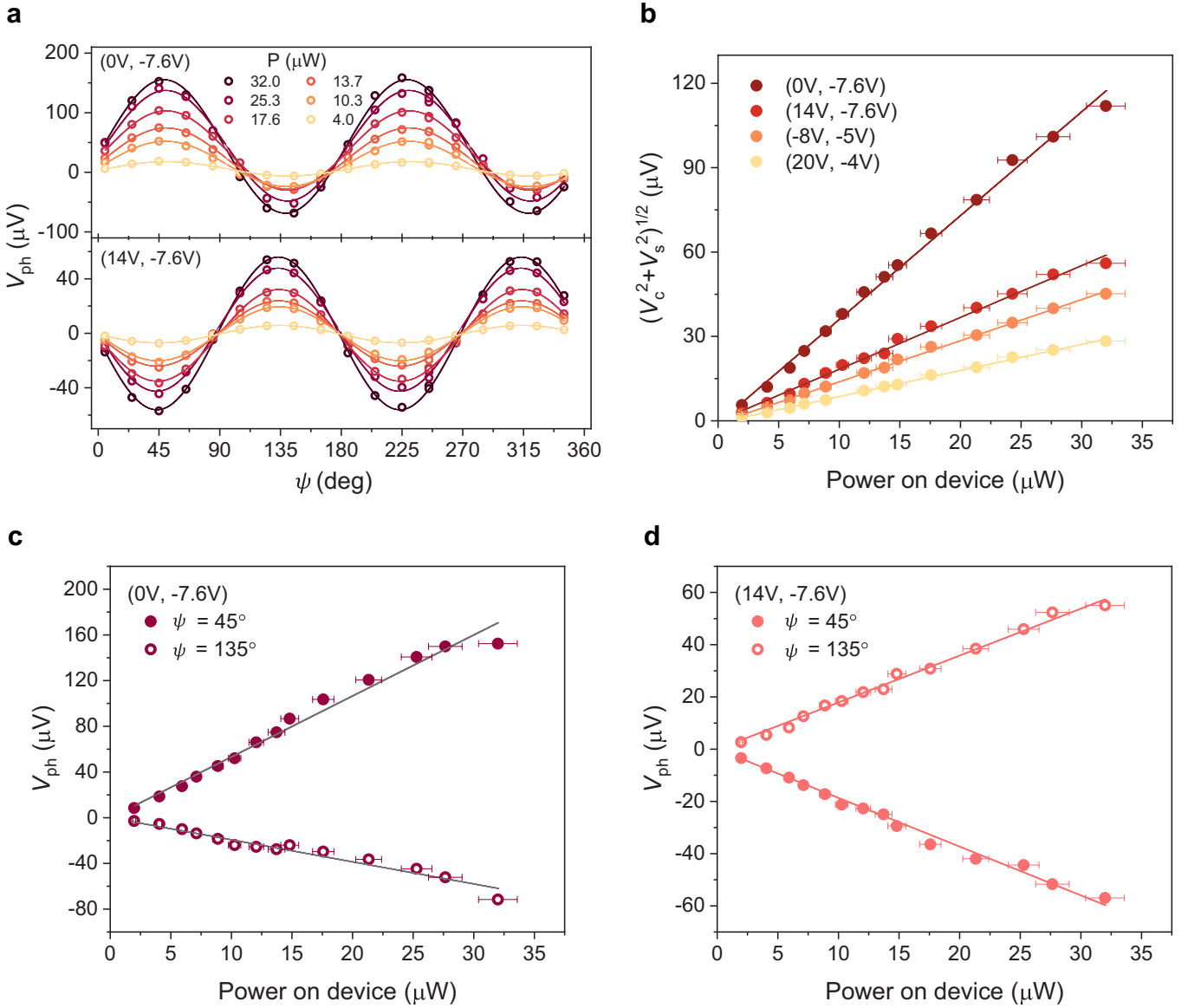


Extended Data Fig. 1 | Temperature dependence of the resistance of TDBG
D1. a, Temperature coefficient of resistance mapping (the same as Fig. 1d) with the selected gate biases measured in **b** highlighted. **b**, Logarithmic values of the

resistance R as a function of $100/T$, measured at different sets of n and D highlighted in **a**.

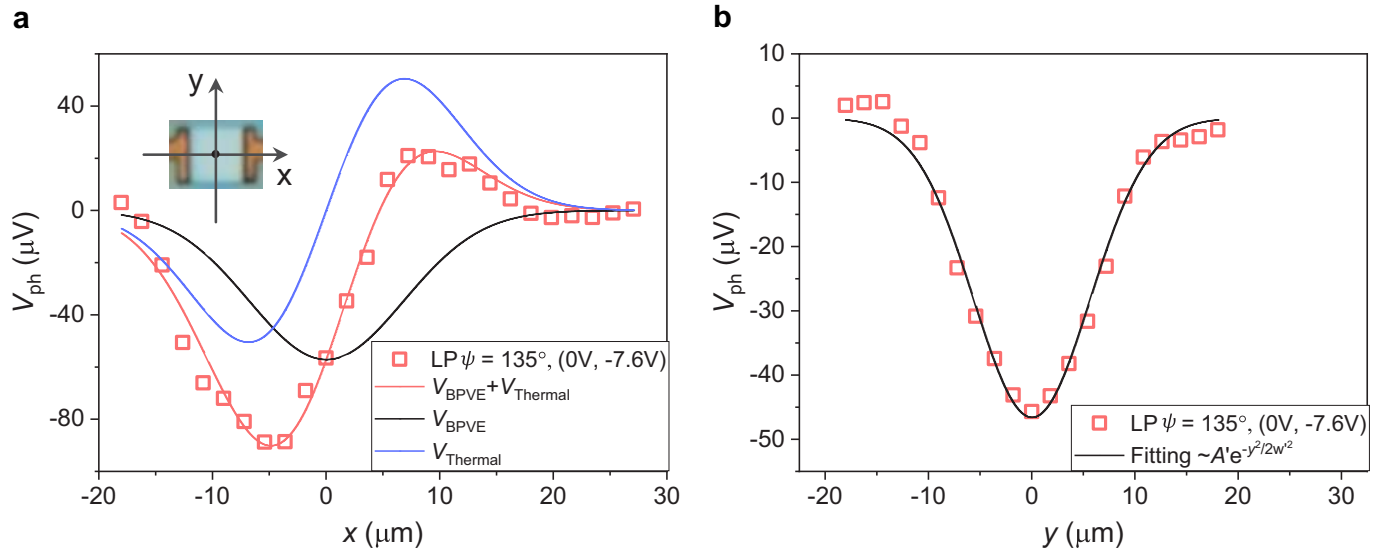


Extended Data Fig. 2 | Polarization dependence of V_{ph} versus ψ under 5 μm linearly polarized light in TDBG D1. Measured photovoltage V_{ph} as a function of the orientation angle ψ under 5 μm linearly polarized light at different sets of gate voltages. Inset: extracted amplitudes of V_c and V_s .



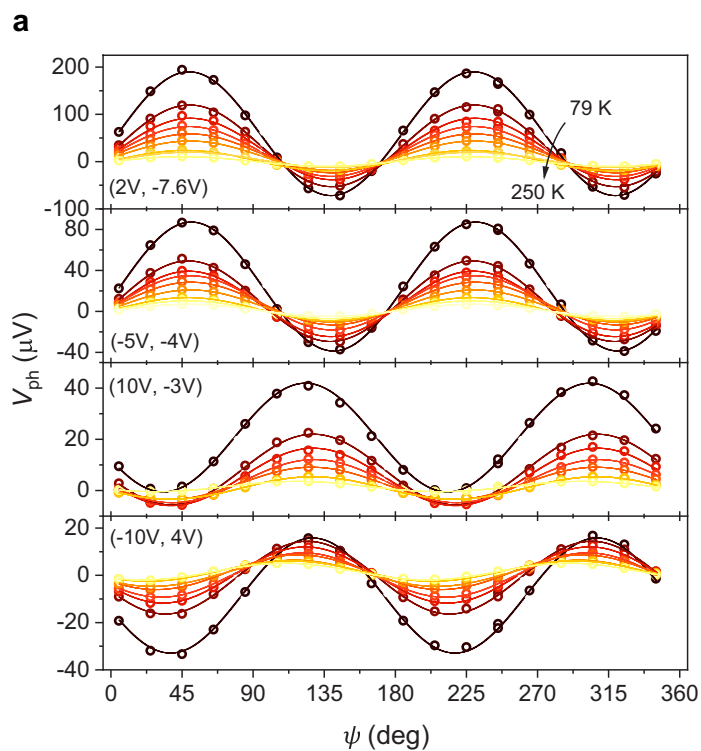
Extended Data Fig. 3 | Power dependence of the photovoltage V_{ph} . **a**, Selected V_{ph} oscillations as a function of ψ , under different excitation powers. **b**, Amplitudes of the V_{ph} oscillation curves, $(V_c^2 + V_s^2)^{1/2}$, versus power on device at different sets of (V_{BG}, V_{TG}) combinations. The polarization-dependent components scale linearly with the power, confirming the second-order nature

of the photoresponse. **c** and **d**, The overall photoresponse V_{ph} , including both the polarization-dependent and -independent components, exhibiting a linear dependence on the power. The curves are measured at two selected polarization angles $\psi = 45^\circ$ and 135° with two sets of gate voltages (0V, -7.6V) and (14V, -7.6V).

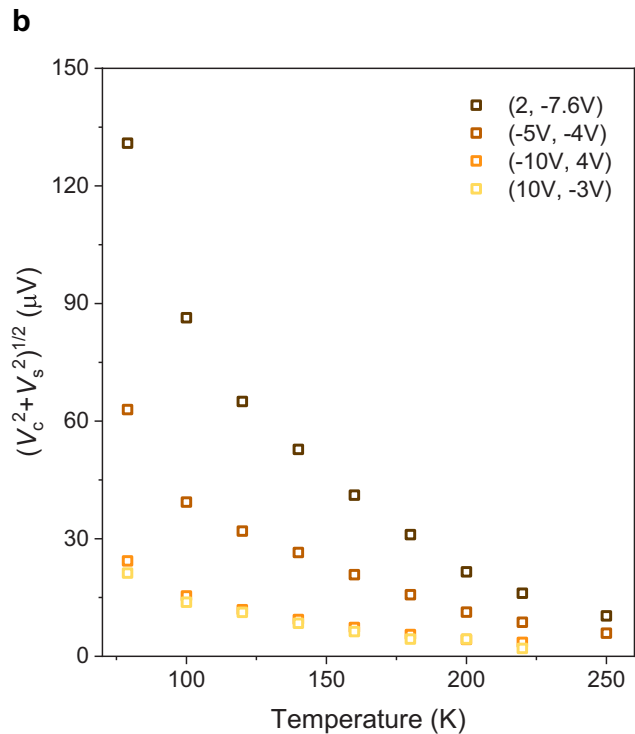


Extended Data Fig. 4 | Spatial dependence of the photovoltage V_{ph} . **a**, Spatial dependence of V_{ph} as a function of the beam spot position x under linear polarized, $7.7 \mu m$ light excitation with $\psi = 135^\circ$ and $(V_{BG}, V_{TG}) = (0 V, -7.6 V)$. Blue, black, and red curves represent the photo-thermal effect, linear BPVE, and their summation, respectively. Inset: Optical image of the TDBG photodetector

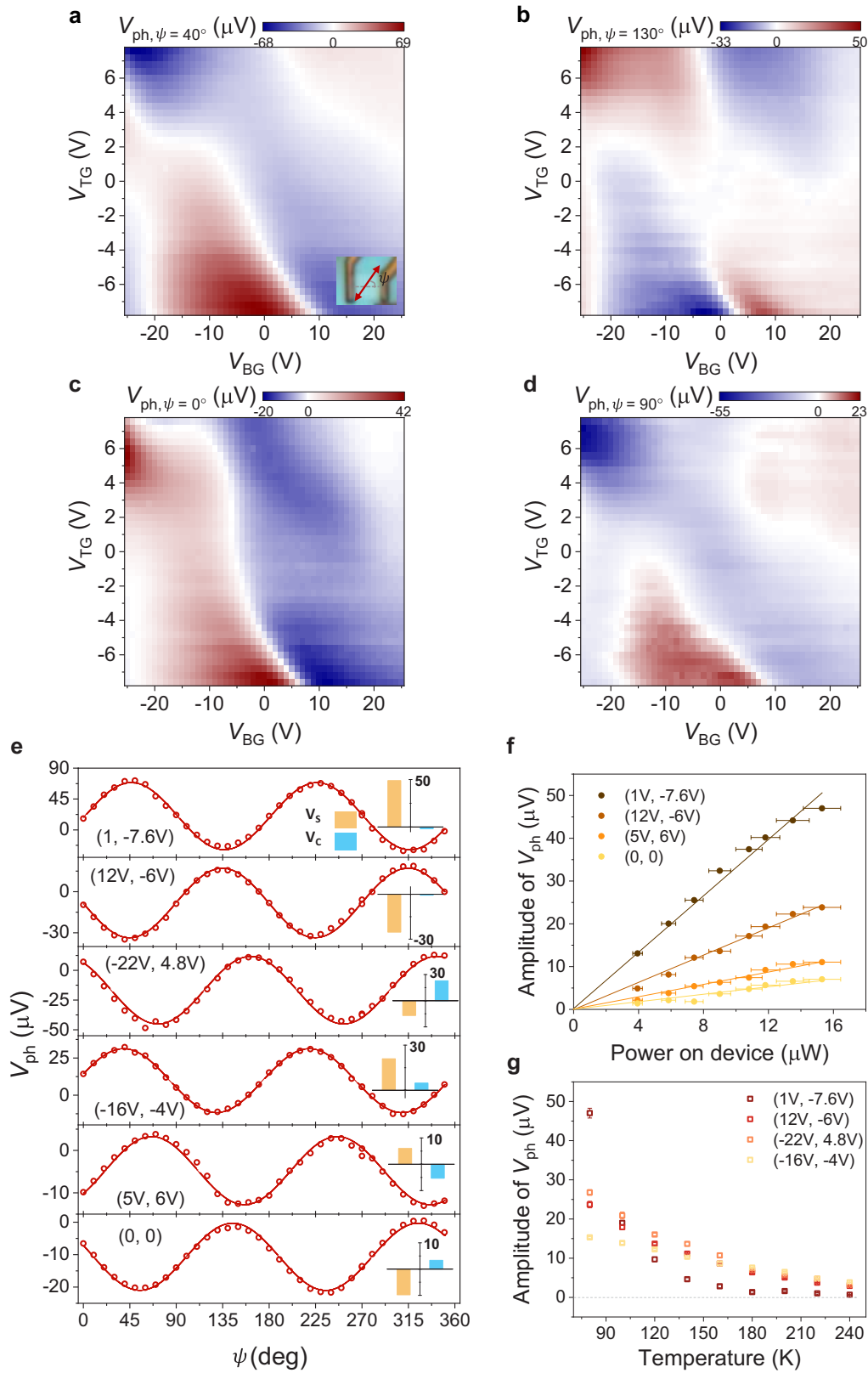
with the x - and y - axes. The device center is defined to be the origin ($x, y = 0$). **b**, Spatial dependence of the V_{ph} along the y -axis (perpendicular to the V_{ph} collection path) measured at $(V_{BG}, V_{TG}) = (0 V, -7.6 V)$ under linearly polarized light with $\psi = 135^\circ$ in TDBG D1, which can be fitted by a Gaussian peak function $A'e^{-y^2/2w^2}$ (black curve).



Extended Data Fig. 5 | Temperature dependence of the photovoltage V_{ph} .
a, Temperature dependence of V_{ph} oscillation curves at different sets of gate voltage combinations under the excitation of $7.7\ \mu\text{m}$ linearly polarized light.

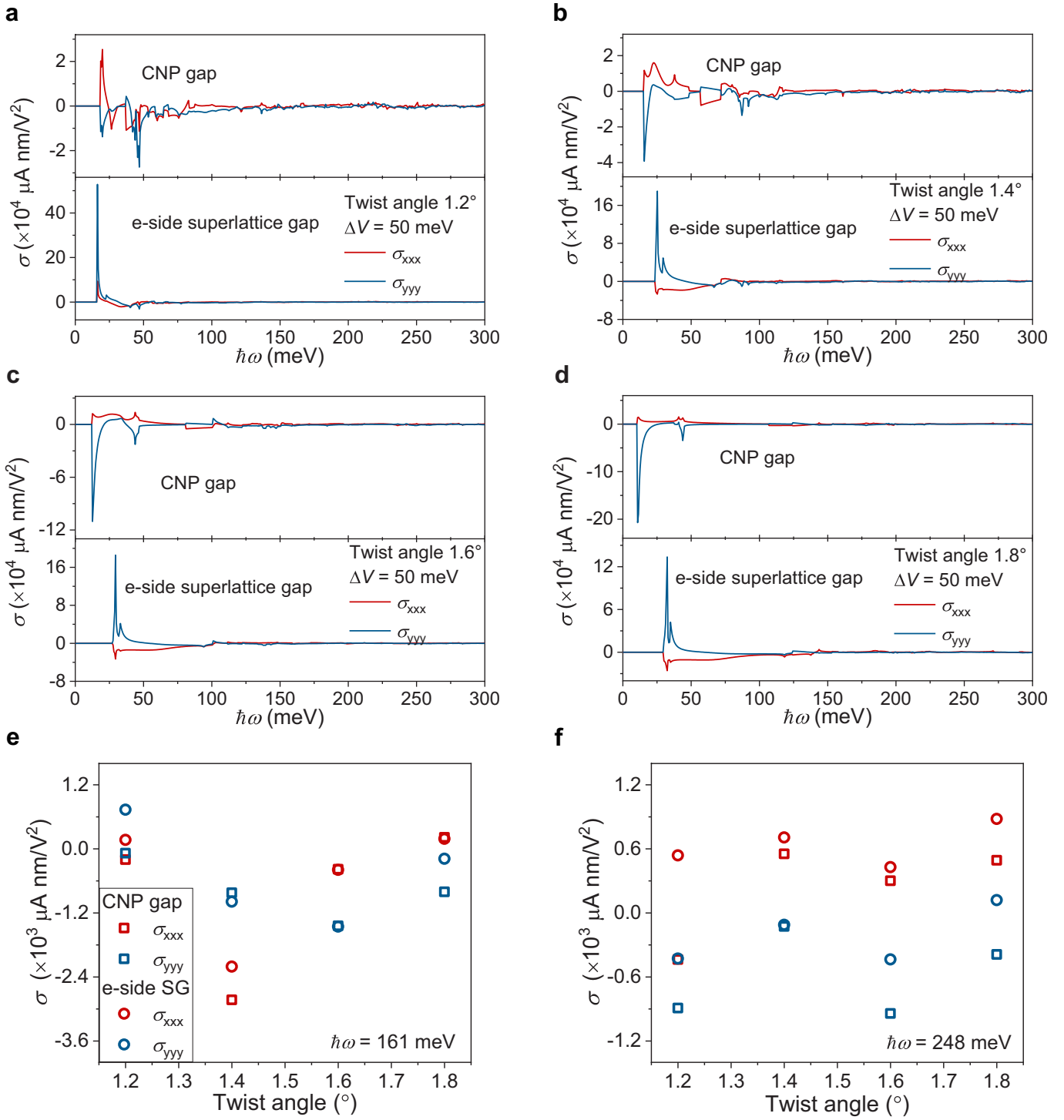


b, Extracted oscillation amplitudes, $(V_c^2 + V_s^2)^{1/2}$, versus temperature at different sets of gate voltage combinations.



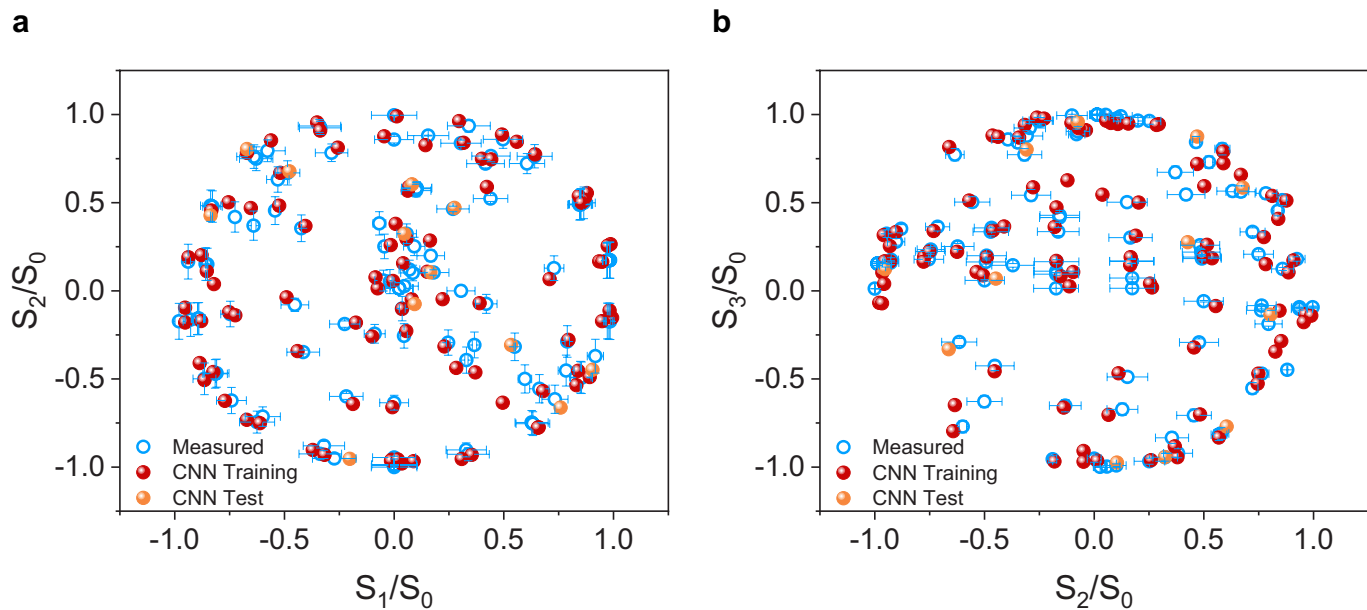
Extended Data Fig. 6 | Linear BPVE in TDBG D2. a-d, Photovoltage V_{ph} mappings as functions of top- (V_{TG}) and back- (V_{BG}) gate biases under the excitation of $5 \mu m$ linearly polarized light with orientation angles $\psi = 40^\circ, 130^\circ, 0^\circ$ and 90° , measured at $T = 79$ K. Inset: optical image of D2 and schematic of the orientation angle ψ of the incident light. D2 has a dimension of $2.2 \mu m \times$

$2 \mu m$. **e,** Polarization dependence of V_{ph} at different gate biases. Insets: extracted polarization-dependent components V_s and V_c . **f,** Linear dependence of the V_{ph} amplitudes on the power at selected gate biases. **g,** Temperature dependence of the amplitudes of V_{ph} .



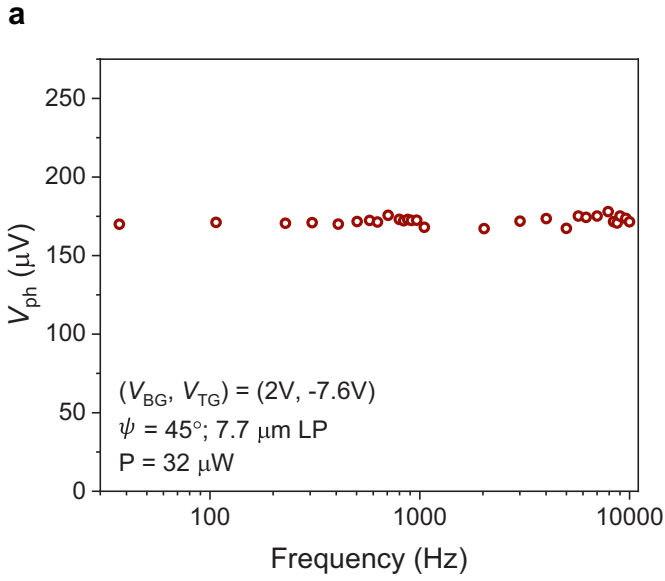
Extended Data Fig. 7 | Twist angle-dependence of the linear bulk photovoltaic effect in TDBG. a-d, Nonlinear conductivity elements σ_{xxx} (red) and σ_{yyy} (blue) as functions of the incident photon energy $\hbar\omega$ when the Fermi level is placed in the CNP gap (top) and the electron-side superlattice gap

(bottom) in TDBG with $\Delta V = 50$ meV and the twist angle of 1.2° , 1.4° , 1.6° and 1.8° . Extracted σ_{xxx} and σ_{yyy} at **e**, $\hbar\omega = 161$ meV and **f**, 248 meV, corresponding to the wavelengths of $7.7 \mu\text{m}$ and $5 \mu\text{m}$ used in our experiments, respectively.

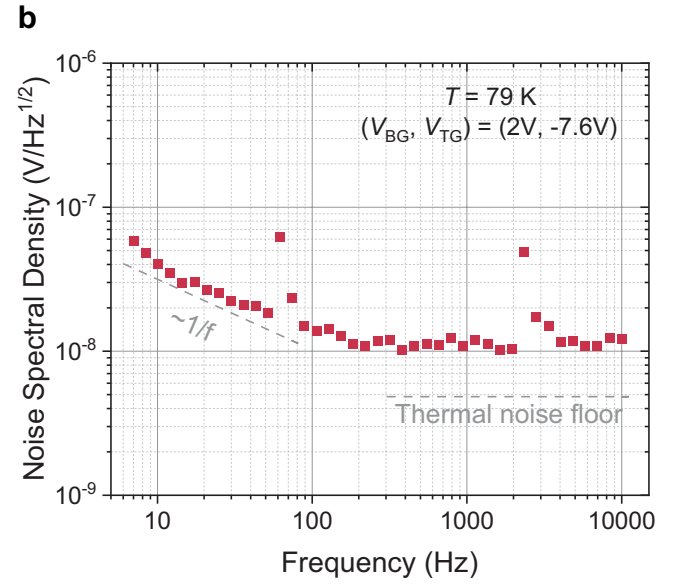


Extended Data Fig. 8 | Complete plots of the polarization outputs from the total 91 training and 12 test data sets. Complete polarization outputs from the total 103 CNN data sets projected to **a**, S_1 - S_2 and **b**, S_2 - S_3 planes. Blue circles

are measured values with error bars determined by the errors in measuring the orientation angle ψ and the ellipticity angle χ (Methods). Red (Orange) spheres are corresponding outputs from CNN training (test) data sets.



Extended Data Fig. 9 | Speed and noise characterizations of the TDBG as a photodetector. a, Photovoltage V_{ph} as a function of the chopping frequency measured under the excitation of $7.7 \mu m$ linearly polarized light with $\psi = 45^\circ$. **b,** Voltage noise spectral density measured with zero in-plane electric bias



without light illumination. Noise density spikes at multiples of 60 Hz are caused by the power line “hums” in the ground loop. Both measurements were taken at $T = 79 K$ and $(V_{BG}, V_{TG}) = (2 V, -7.6 V)$ in TDBG D1.