

Correlated insulating states in slow Dirac fermions on a honeycomb moiré superlattice

Dongyang Yang^{1,2,†}, Jing Liang^{1,2,†}, Haodong Hu^{1,2,†}, Nitin Kaushal^{1,2}, Chih-En Hsu^{3,4}, Kenji Watanabe⁵, Takashi Taniguchi⁶, Jerry I. Dadap^{1,2}, Zhenglu Li⁴, Marcel Franz^{1,2}, and Ziliang Ye^{1,2,*}

¹Department of Physics and Astronomy, The University of British Columbia, Vancouver, BC V6T 1Z1, Canada

²Quantum Matter Institute, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada

³Department of Physics, Tamkang University, New Taipei City 251301, Taiwan

⁴Mork Family Department of Chemical Engineering and Materials Science, University of Southern California, Los Angeles, California 90089, USA

⁵Research Center for Functional Materials, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan

⁶International Center for Materials Nanoarchitectonics, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan

[†]These authors contributed equally to this work

*Corresponding author: zlye@phas.ubc.ca

Supplementary Figures S1-S16 and Supplementary Sections S1-S7

Supplementary Figures

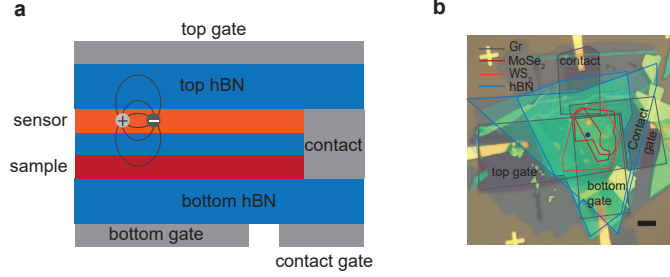


Figure S1. | Schematic and optical image of device D1. **a.** A schematic side view of Device D1. A contact gate covers the TMD-graphene contact region to minimize the nonlinear gating effect. **b.** The optical image of device D1. The twisted MoSe₂ homobilayer is outlined by the solid red curve. The grey curves outline the region of graphite gates. The blue and yellow dots denote the two measured positions P1 and P2 in D1. Scale bar: 5 μm .

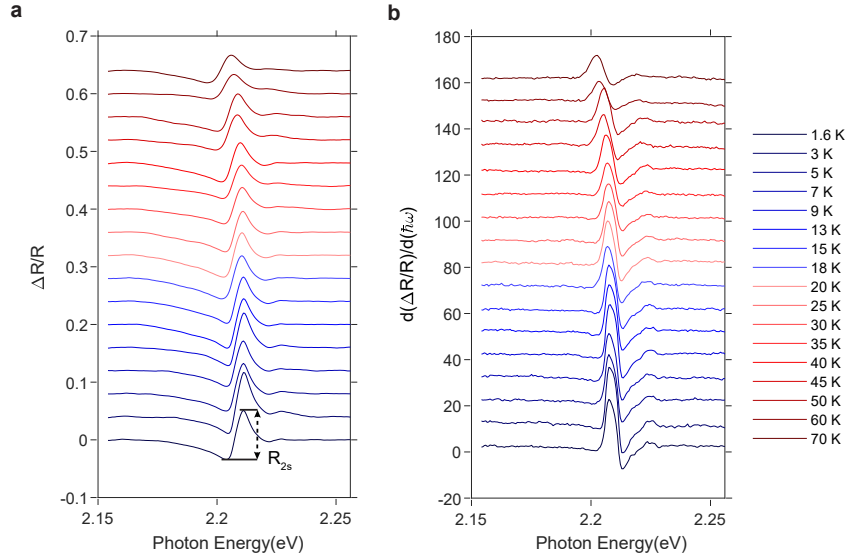


Figure S2. | Reflection contrast and its first energy derivative spectra of the 2s exciton. **a.** Reflection contrast spectra of the 2s exciton at different temperatures. The doping is fixed at the charge-neutral point of $\nu = 0$. We use the peak-to-peak amplitude at the 2s resonance to represent the oscillator strength, denoted as R_{2s} . **b.** The first derivative of the spectra in (a) with respect to photon energy.

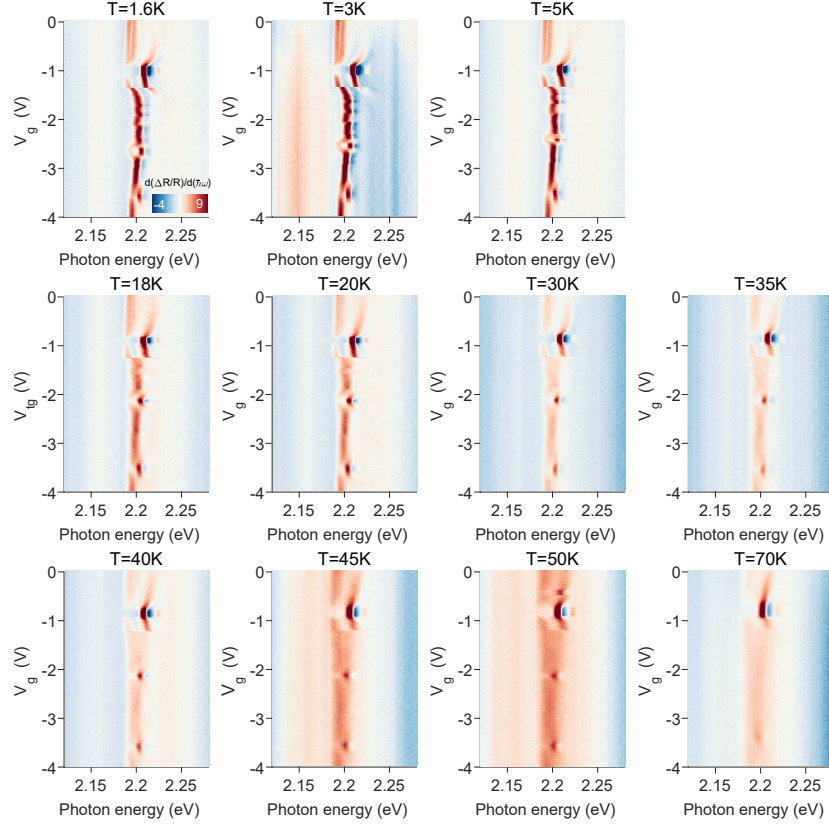


Figure S3. | Additional temperature dependence data for device D1. Contour plot of gate dependence data at $T=1.6, 3, 5, 18, 20, 30, 35, 40, 45, 50,$ and 70 K.

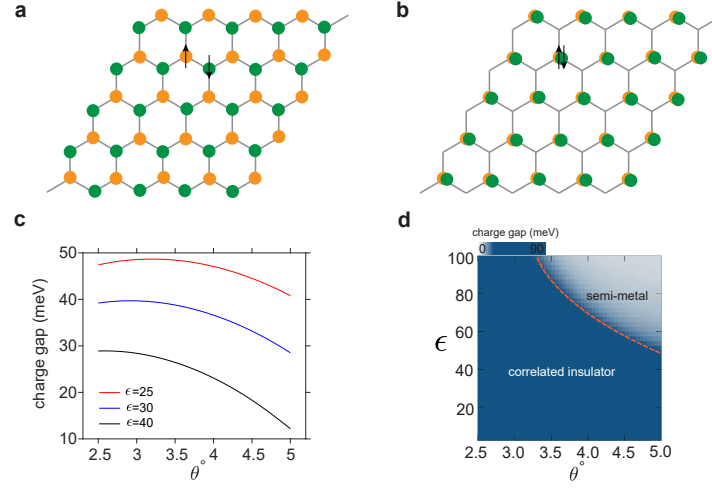


Figure S4. | Hartree-Fock calculations of the charge gap at $\nu = -2$
a,b. Schematics of two possible charge configurations for $\nu = -2$ state. When the ratio of U_0/t is large enough, the ground state can be a spin-staggered antiferromagnetic insulator (**a**). Due to competition between U_0 and U_1 , another possible charge distribution of the ground state is a sublattice-polarized charge density wave state (**b**). **c.** Calculated gap size of $\nu = -2$ state as a function of twist angle for three different effective dielectric permittivity. **d.** Charge gap of $\nu = -2$ state as a function of the effective dielectric permittivity (ϵ) and twist angle (θ), calculated using unrestricted Hartree-Fock method. The dashed orange denotes the boundary between the correlated insulator and semimetal phases.

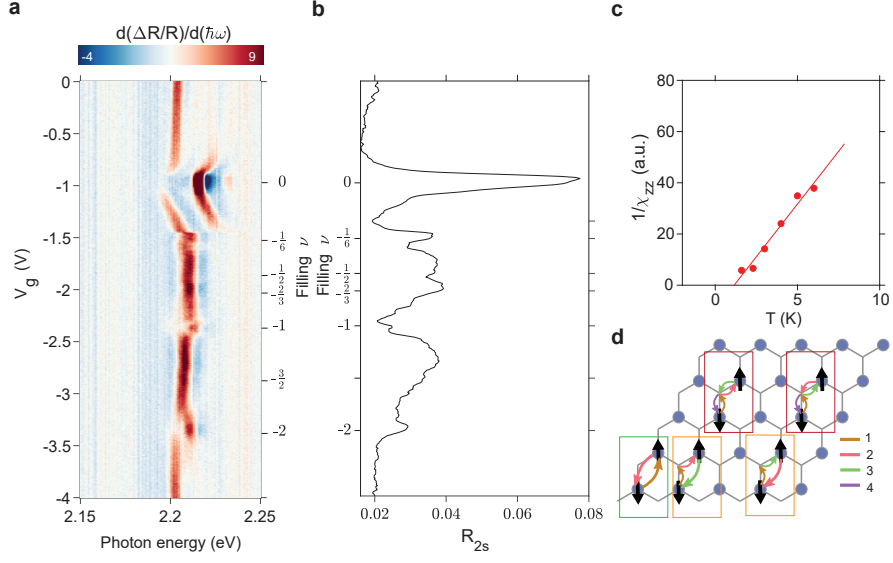


Figure S5. | Additional data for another position P2 of device D1. **a.** First energy derivative of the reflection contrast spectra as a function of top gate voltage. The experimental temperature is 7 K. **b.** Filling number versus oscillator strength of the 2s exciton at position P2. **c.** Curie-Weiss fit for the magnetic susceptibility near zero magnetic field as a function of temperature at filling $\nu = -1$ and position P2. **d.** A schematic of the hopping processes contributing to the leading terms in the spin exchange. Black arrows represent the local moments. Green, yellow, and red boxes outline the second-, third-, and fourth-order processes, respectively. The color of the arrows corresponds to the step number in the hopping process.

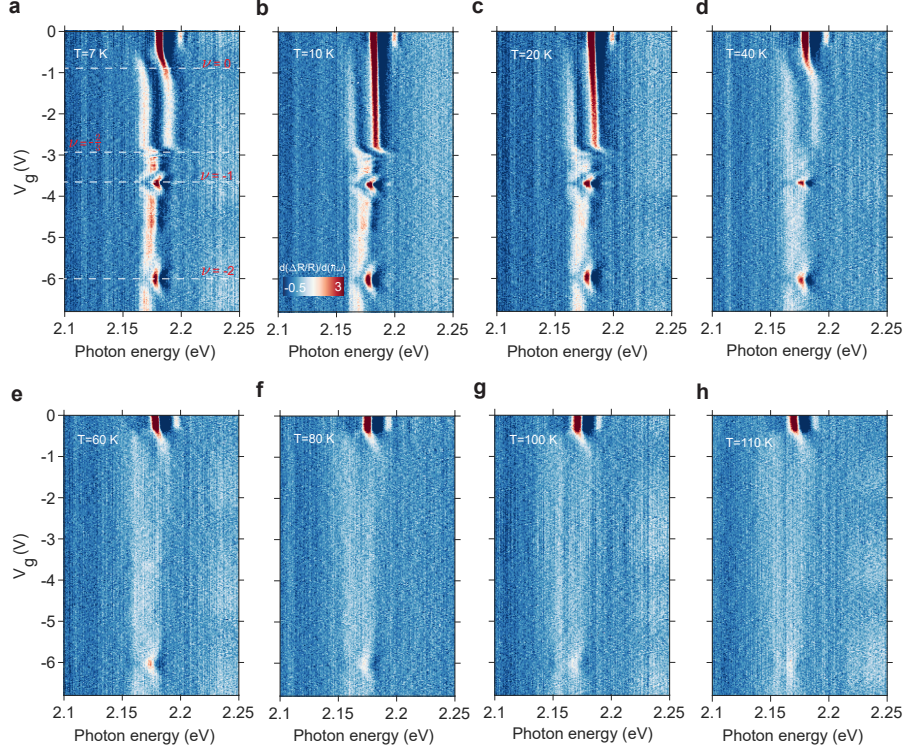


Figure S6. | Temperature dependence of the insulating states in device D2. a-h Contour plot of gate-dependent reflection contrast spectra at $T=7, 10, 20, 40, 60, 80, 100$, and 110 K. The white dashed lines in (a) denote the filling of $\nu = 0, -\frac{2}{3}, -1$, and -2 states. The calibrated twist angle for D2 is 2.9° .

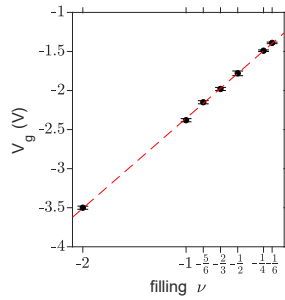


Figure S7. | Assignment of the filling number of the insulating states. The plot shows the gate voltage V_g as a function of the filling factor ν . The black dots represent experimental data, extracted from the peak oscillator strength of the $2s$ exciton, with error bars indicating the measurement uncertainty. The red dashed line is a linear fit to all observed states, yielding a slope of approximately 1.15 V per filling number. The error bar indicates the uncertainty of the R_{2s} peak position.

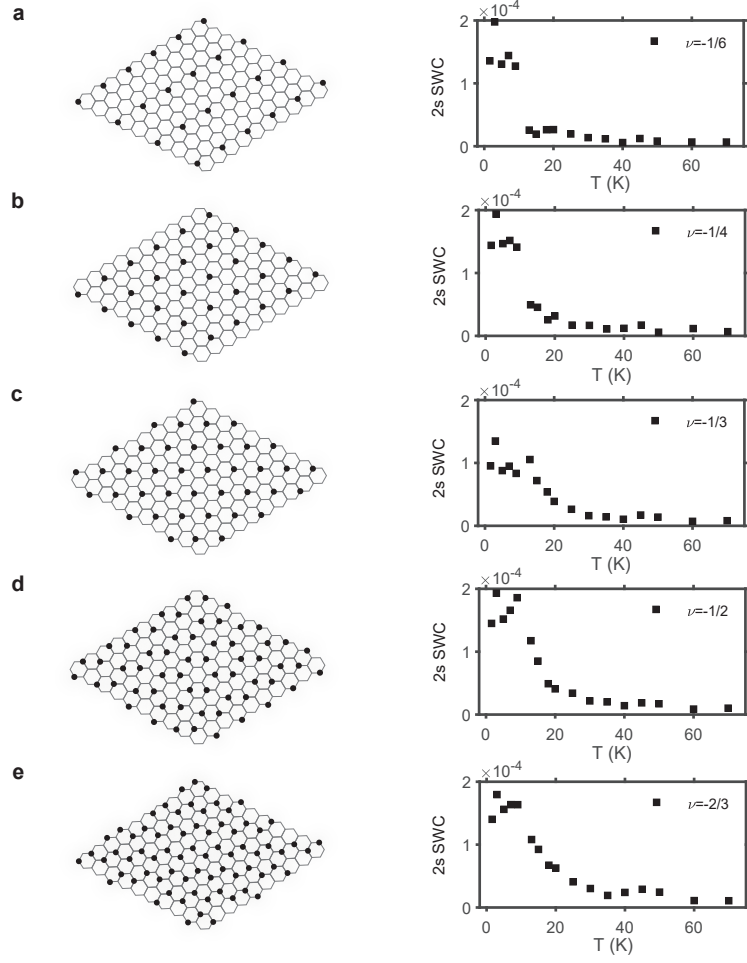


Figure S8. | Generalized Wigner crystals.a-e. Left: Proposed charge configuration for the Wigner crystal state at fractional fillings. Right: Temperature dependence of 2s spectral weight contrast (2s SWC).

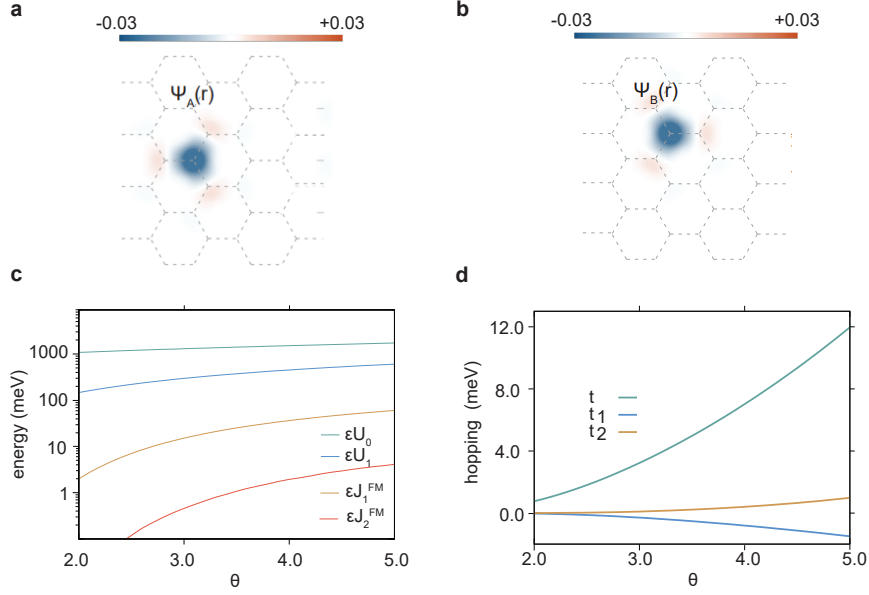


Figure S9. | Hartree-Fock calculation parameters **a.** and **b.** Wannier orbitals attained using the first two moiré bands in Figure 1d for A and B sublattices. **c.** various interaction parameters as a function of the twist angle, namely onsite Coulomb repulsion ϵU_0 , nearest neighbor density-density repulsion ϵU_1 , nearest neighbor (J_1^{FM}) and next nearest neighbor direct (J_2^{FM}) ferromagnetic exchanges. **d.** Hopping parameters, (nearest t , second nearest t_1 and third nearest neighbor t_2) as a function of the twist angle

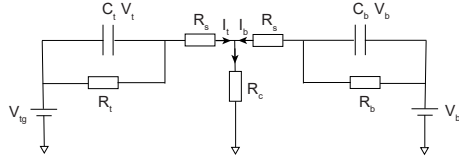


Figure S10. | Model of Nonlinear gating effect The equivalent circuit of the dual-gated device. $R_{t/b}$, R_s , and R_c are the tunneling, sample, and contact resistance, respectively. The leak currents are denoted as I_t and I_b . The voltage drop on the top and bottom ideal capacitors are labeled as V_t and V_b .

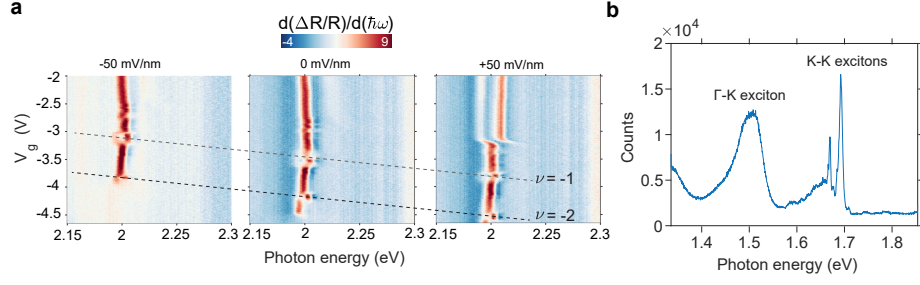


Figure S11. | Additional spectral data for D1. **a.** With the contact gates turned off, the voltages for integer fillings disperse significantly with the electric field. The three panels from left to right are the first energy derivative of the reflection contrast measured at different electric fields of -50, 0, and +50 mV/nm, respectively. **b.** Photoluminescence spectrum of t-MoSe₂. The $\Gamma - K$ indirect exciton emission is centered around 1.5 eV, while the other two peaks at high energy are from the $K - K$ bright excitons.

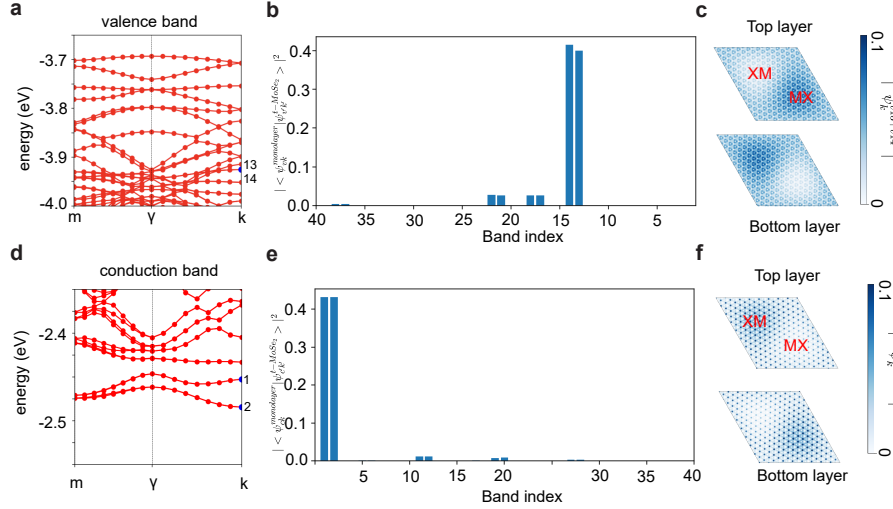


Figure S12. | DFT calculation results. **a,d** Valence and conduction band structures of a 3.89° twisted MoSe₂. The folded valence band maximum in the K valley contributes mostly to band 13 and 14 (degenerated at k point), while the folded K-valley conduction band minimum forms band 1 and 2. **b,e** The projection calculation of $|\langle \psi_{v(c)k}^{monolayer} | \psi_{v'(c')k'}^{t-MoSe_2} \rangle|^2$. **c,f** The wavefunction of the folded K-valley valence and conduction band extremes is mostly distributed in the XM/MX site in the top/bottom layer, respectively.

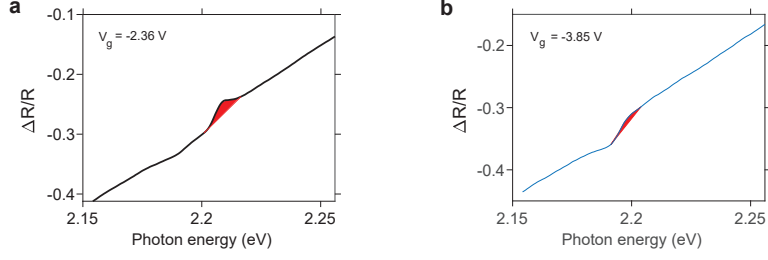


Figure S13. | Definition of the 2s spectral weight contrast. **a.** The 2s exciton reflection contrast at $T=7$ K in device D1 (Black). The doping is set to $V_{tg} = -2.36$ V ($\nu = -1$). We assume a linear background near the resonance region. **b.** The 2s exciton reflection contrast in a conductive region ($V_{tg} = -3.85$ V). The red area difference between **(a)** and **(b)** underneath the two curves is defined as the 2s spectral weight contrast. A 14-meV integration window covers the 2s exciton resonance at all temperatures in both **(a)** and **(b)**.

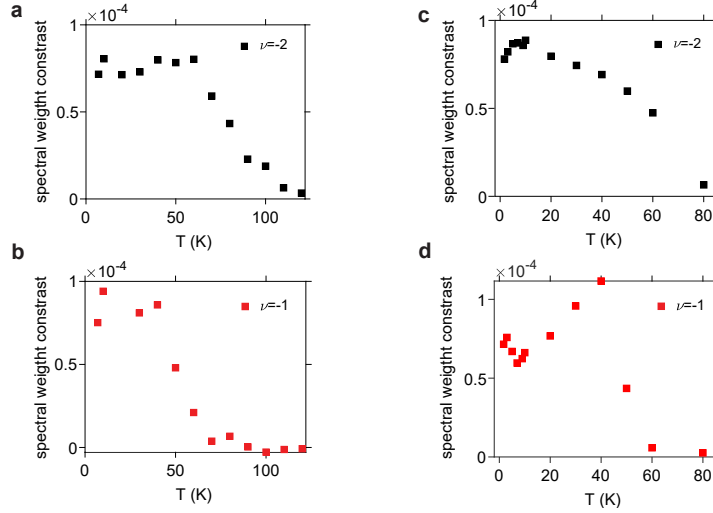


Figure S14. | Melting of $\nu = -2$ and -1 states in device D2 and D3. **a and b.** The temperature dependence of the 2s spectral weight contrast when the filling is set at $\nu = -2$ (black) and $\nu = -1$ (red) in device D2. **c and d.** Same plots for device D3. The peak of the 2s spectral weight contrast for $\nu = -1$ is around 25 K in this device.

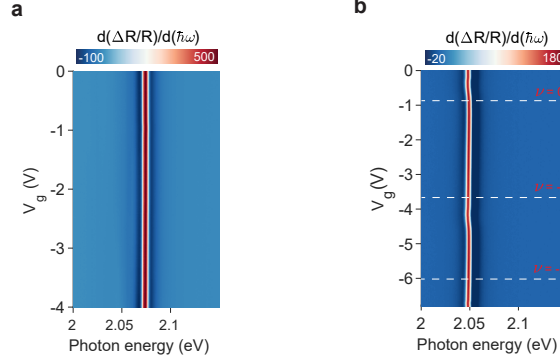


Figure S15. | 1s exciton of WS₂ sensor in device D1 and D2 **a,b** The spectra of 1s exciton in the WS₂ sensor layer as a function of doping V_g in the device D1 (**a**) and device D2 (**b**). The doping density is controlled by the effective gate voltage $V_g = \frac{V_{tg} + (C_b/C_t)V_{bg}}{2}$, where C_b and C_t are the geometrical capacitance of the top and bottom hBN.

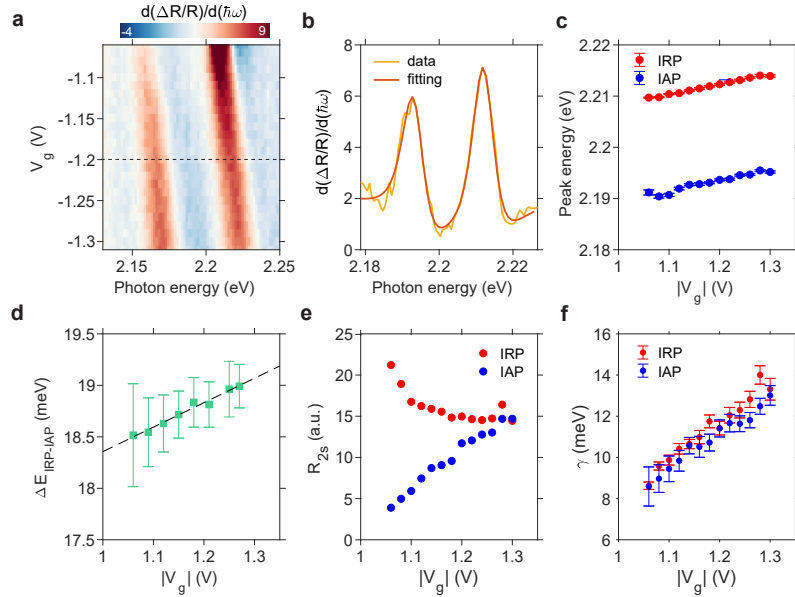


Figure S16. | Interlayer exciton-polaron **a** Enlarged contour plot of 2s interlayer exciton-polarons (*IAP* and *IRP*) in the dilute doping region. **b**. A linecut of fitting results at $V_g = -1.2$ V. **c**. Extracted peak energies of *IAP* (blue) and *IRP* (red). **d**. Energy difference between *IRP* and *IAP* as a function of gate voltage. **e**. The Oscillator strength (R_{2s}) transfers from *IRP* to *IAP* as the doping increases. **f**. Extracted linewidth (γ) increases with doping for both *IRP* and *IAP*. The error bar arises from the uncertainty of fitting.

Supplementary Sections S1-S7

S1. Continuum model for t-MoSe₂ homobilayer

In this section, we will discuss the procedure to solve the continuum model in detail. We start from a continuum model for twisted MoSe₂ homobilayer, following the ref.[1,4], as shown below:

$$H_\tau(\mathbf{r}) = \Psi_\tau^\dagger(\mathbf{r}) \left(-\frac{\hbar^2 \hat{\mathbf{k}}^2}{2m^*} + \Delta(\mathbf{r}) \right) \Psi_\tau(\mathbf{r}) \quad (1)$$

where $\Psi_\tau^\dagger(\mathbf{r})$ is the field operator for the spin τ . The $\Delta(\mathbf{r}) = \sum_s \sum_{j=1}^6 V_s e^{i(\mathbf{g}_j^s \cdot \mathbf{r} + \phi_s)}$ is the moiré potential as shown in Figure 1b. We fixed the parameters of the continuum model corresponding to the twisted MoSe₂ bilayer system, i.e. $\{V_1, V_2, V_3\} = \{36.8, 8.4, 10.2\}$ meV, $\phi_s = \pi$, $m^* = 1.17m_e$, and monolayer lattice constant $a_0 = 0.33$ nm.

We use the plane wave expansion $\Psi_\tau(\mathbf{r}) = \frac{1}{\sqrt{A}} \sum_{\mathbf{k}} f_{\mathbf{k}\tau}^\dagger e^{i\mathbf{k} \cdot \mathbf{r}}$, with inverse transformation $f_{\mathbf{k}\tau}^\dagger = \frac{1}{\sqrt{A}} \int \Psi_\tau(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}} d^2\mathbf{r}$, where $f_{\mathbf{k}\tau}^\dagger$ is the hole creation operator with momentum $\mathbf{k}$ and spin τ . A is the area of the sample. The noninteracting Hamiltonian, i.e. $H_\tau = \int H_\tau(\mathbf{r}) d^2\mathbf{r}$, after performing the plane wave expansion, can be written as $H_\tau = \sum_{\mathbf{k} \in mbz} H_\tau(\mathbf{k})$, where

$$H_\tau(\mathbf{k}) = - \sum_{\mathbf{G}} [\epsilon(\mathbf{k} + \mathbf{G}) f_{\mathbf{k}+\mathbf{G}\tau}^\dagger f_{\mathbf{k}+\mathbf{G}\tau} + \sum_{\substack{j=1,3,5 \\ s}} V_s (e^{i\phi_s} f_{\mathbf{k}+\mathbf{g}_j+\mathbf{G}\tau}^\dagger f_{\mathbf{k}+\mathbf{G}\tau} + \text{h.c.})] \quad (2)$$

where, $\epsilon_\tau(\mathbf{k}) = -\frac{\hbar^2 k^2}{2m^*}$. The sum in $\mathbf{G}$ spans the grid of moiré reciprocal lattice vectors around the momentum $\mathbf{k}$ which lies in the moiré Brillouin zone (mbz). We noticed that finite grids of $\mathbf{G}$ of size $\tilde{L} \times \tilde{L}$ quickly give convergence in the low-energy eigenspectrum of $H_\tau(\mathbf{k})$, we use the 18×18 grid of $\mathbf{G}$. The single particle Hamiltonian is diagonal in moiré Bloch states $c_{\mathbf{k}n\tau}^\dagger = \sum_{\mathbf{G}} u_{\mathbf{G}\tau}^n(\mathbf{k}) f_{\mathbf{k}+\mathbf{G}\tau}^\dagger$ i.e. $H_\tau(\mathbf{k}) = \sum_n \tilde{\epsilon}_{n\tau}(\mathbf{k}) c_{\mathbf{k}n\tau}^\dagger c_{\mathbf{k}n\tau}$, where n is moiré band index. The total number of moiré bands we get in the continuum model is $\tilde{L}^2$, out of which mainly the few lowest energy moiré bands (in hole representation, as used in this work) are generally relevant.

S2. Unrestricted Hartree-Fock calculation

We start with the interaction Hamiltonian written for the holes, as shown below

$$H_{int} = \frac{1}{2A} \sum_{\substack{\mathbf{k}_1 \mathbf{k}_2 \mathbf{q} \\ \tau \tau'}} V(\mathbf{q}) f_{(\mathbf{k}_1 - \mathbf{q})\tau}^\dagger f_{(\mathbf{k}_2 + \mathbf{q})\tau'}^\dagger f_{\mathbf{k}_2\tau'} f_{\mathbf{k}_1\tau} \quad (3)$$

where, $V(\mathbf{q}) = \frac{2\pi e^2 \tanh(|\mathbf{q}|d)}{\epsilon|\mathbf{q}|}$, ϵ is the effective dielectric constant used to compare the experiment and theory, and $d \approx 10$ nm is distance between the gate and the sample. The area of the sample is $A = (\sqrt{3}/2)l_1 l_2 a_M^2$, with $l_{1(2)}$ moiré unit cells along the Bravais lattice vector $\mathbf{a}_{1(2)}$. In equation (3), the sum over all possible momentum's namely $\sum_{\mathbf{k}}$ can be replaced by $\sum_{\mathbf{k}+\mathbf{G}}$ where now $\mathbf{k} \in mbz$

(moiré Brillouin zone). Moreover, using $f_{\mathbf{k}+\mathbf{G}l\tau} = \sum_n c_{\mathbf{k}n\tau} u_{\mathbf{G}l\tau}^n(\mathbf{k})$, we write the interaction term in the moiré band basis as following

$$H_{int} = \frac{1}{2A} \sum_{\substack{\mathbf{k}_1 \mathbf{k}_2 \mathbf{q} \tau \tau' \\ n_1 n_2 n_3 n_4}} \tilde{V}_{n_1 n_2 n_3 n_4}^{\tau \tau'}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}) c_{(\mathbf{k}_1 - \mathbf{q})n_1 \tau}^\dagger c_{(\mathbf{k}_2 + \mathbf{q})n_2 \tau'}^\dagger c_{\mathbf{k}_2 n_3 \tau'} c_{\mathbf{k}_1 n_4 \tau} \quad (4)$$

where,

$$\tilde{V}_{n_1 n_2 n_3 n_4}^{\tau \tau'}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}) = \sum_{\mathbf{G}} V(\mathbf{q} + \mathbf{G}) \lambda_{n_1 n_2 \tau}(\mathbf{k}_1 - \mathbf{q}, \mathbf{k}_1 + \mathbf{G}) \lambda_{n_3 n_2 \tau'}^*(\mathbf{k}_2, \mathbf{k}_2 + \mathbf{q} + \mathbf{G}) \quad (5)$$

and the form factors $\lambda_{n_1 n_2 \tau}(\mathbf{k}_1, \mathbf{k}_2) = \sum_{\mathbf{G}} (u_{\mathbf{G}\tau}^{n_1}(\mathbf{k}_1))^* u_{\mathbf{G}\tau}^{n_2}(\mathbf{k}_2)$. In equation (3), the total number of terms is $4(l_1 l_2)^3 \tilde{L}^6$, whereas if we project the equation (4) to only few N_b relevant moiré bands then the total number of terms in equation (4) is $4N_b^4(l_1 l_2)^3$. One may look at the total single particle fermionic states required for both equations, for equation (3) $2l_1 l_2 \tilde{L}^2$ are required whereas for the equation (4) $2N_b l_1 l_2$ states are required, which suggests a huge gain if we use the interaction Hamiltonian projected to few moiré bands of interest. In our calculations, we project the interactions only to the first two moiré bands by restricting the sum over band indices in the equation (4).

Then we perform the mean field decomposition of equation (4) to get the following Hartree and Fock terms

$$H_{Hartree} = \frac{1}{2A} \sum_{\substack{\mathbf{k}_1 \mathbf{k}_2 \mathbf{q} \tau \tau' \\ n_1 n_2 n_3 n_4}} (\tilde{V}_{n_1 n_2 n_3 n_4}^{\tau \tau'}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}) + \tilde{V}_{n_2 n_1 n_4 n_3}^{\tau' \tau}(\mathbf{k}_2, \mathbf{k}_1, -\mathbf{q})) \langle c_{(\mathbf{k}_1 - \mathbf{q})n_1 \tau}^\dagger c_{\mathbf{k}_1 n_4 \tau} \rangle c_{(\mathbf{k}_2 + \mathbf{q})n_2 \tau'}^\dagger c_{\mathbf{k}_2 n_3 \tau'} \quad (6)$$

$$H_{Fock} = -\frac{1}{2A} \sum_{\substack{\mathbf{k}_1 \mathbf{k}_2 \mathbf{q} \tau \tau' \\ n_1 n_2 n_3 n_4}} (\tilde{V}_{n_1 n_2 n_3 n_4}^{\tau \tau'}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}) + \tilde{V}_{n_2 n_1 n_4 n_3}^{\tau' \tau}(\mathbf{k}_2, \mathbf{k}_1, -\mathbf{q})) \langle c_{(\mathbf{k}_1 - \mathbf{q})n_1 \tau}^\dagger c_{\mathbf{k}_2 n_3 \tau'} \rangle c_{(\mathbf{k}_2 + \mathbf{q})n_2 \tau'}^\dagger c_{\mathbf{k}_1 n_4 \tau} \quad (7)$$

The total mean field Hamiltonian can be written as follows,

$$\begin{aligned} H_{HF} = & H_{Hartree} + H_{Fock} - \frac{1}{2} (\langle H_{Hartree} \rangle + \langle H_{Fock} \rangle) \\ & - \frac{1}{2A} \sum_{\substack{\mathbf{k}_2 \mathbf{q} \tau \\ n_1 n_2 n_3}} \tilde{V}_{n_1 n_2 n_3 n_2}^{\tau \tau}(\mathbf{k}_2 + \mathbf{q}, \mathbf{k}_2, \mathbf{q}) c_{\mathbf{k}_2 n_1 \tau}^\dagger c_{\mathbf{k}_2 n_3 \tau} \\ & + \frac{1}{2A} \sum_{\substack{\mathbf{k}_1 \mathbf{k}_2 \tau \tau' \\ n_1 n_2 n_4}} \tilde{V}_{n_1 n_2 n_2 n_4}^{\tau \tau'}(\mathbf{k}_1, \mathbf{k}_2, 0) c_{\mathbf{k}_1 n_1 \tau}^\dagger c_{\mathbf{k}_1 n_4 \tau} \end{aligned} \quad (8)$$

The last two quadratic fermionic terms come from the anti-commutation of Fermions while reordering fermionic operators for the Hartree and Fock terms. Equation H_{HF} is solved for the given filling of holes maintaining self-consistency in the order parameters, namely the single-particle density matrix elements $\langle c_{\mathbf{k}_1 n_1 \tau}^\dagger c_{\mathbf{k}_2 n_2 \tau'} \rangle$. We solve the above mean-field Hamiltonian using unrestricted

Hartree-Fock, starting from random initial conditions for the order parameters and then gaining the self-consistency for the order parameters in an iterative way.

After gaining convergence, we also calculate the charge gap Δ_c to deduce whether the system is a metal or insulator, as shown in Figure S4(d). In addition, we calculate the real space resolved density of doped holes $\langle n(\mathbf{r}) \rangle = \sum_{\tau} \langle \Psi_{\tau}(\mathbf{r}) \Psi_{\tau}^{\dagger}(\mathbf{r}) \rangle$ to visualize the Wigner Crystals using the equation (9). The theoretically proposed charge configuration of the generalized Wigner crystals using the above algorithm and their critical temperatures observed in the experiments are presented in Figure S8.

$$\langle \Psi_{\tau}(\mathbf{r}) \Psi_{\tau'}^{\dagger}(\mathbf{r}) \rangle = \frac{1}{A} \sum_{\substack{\mathbf{G}_1 \mathbf{G}_2 \mathbf{k}_1 \mathbf{k}_2 \\ n_1 n_2}} e^{i\mathbf{r} \cdot (\mathbf{k}_1 - \mathbf{k}_2)} e^{i\mathbf{r} \cdot (\mathbf{G}_1 - \mathbf{G}_2)} (u_{\mathbf{G}_1 \tau}^{n_1}(\mathbf{k}_1))^* u_{\mathbf{G}_2 \tau}^{n_2}(\mathbf{k}_2) \langle c_{\mathbf{k}_1 n_1 \tau}^{\dagger} c_{\mathbf{k}_2 n_2 \tau'} \rangle \quad (9)$$

S3. Spin Model for $\nu = -1$ Wigner crystal

Our Hartree-Fock analysis establishes the presence of an emergent triangular Wigner crystal for filling $\nu = -1$. Furthermore, for $\theta = 3.6^\circ$, we find very robust ferromagnetism up to large $\epsilon \approx 45$ until the Wigner crystal starts melting. This is in contrast to the $\nu = -2$ Mott insulator state, where we find a clear transition of ferromagnetic insulator to antiferromagnetic insulator near $\epsilon \approx 16$. The Wigner crystal nature of $\nu = -1$ state is also consistent with the weak angle dependence of its melting temperature (Figure 3d).

For fillings other than half filling, Hartree-Fock studies generally tend to overestimate ferromagnetic states[1, 2, 3]. In this section, we derived the effective low-energy spin 1/2 model $H = J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j$ by performing the t/U_0 perturbation theory up to the fourth order. The most dominant exchange processes contributing to the spin correlation J are shown in Figure S5(d). They are all included in the following equation,

$$J \approx J_2^{\text{FM}} + \frac{4t_1^2}{U_0} + \frac{t^2 t_1}{U_1^2} + \frac{8t^2 t_1}{U U_1} + \frac{t^4}{U_0 U_1^2} + \frac{2t^4}{(U_0 + 2U_1)U_1^2} \quad (10)$$

We also include the second nearest neighbor (in terms of the underlying Honeycomb lattice) direct ferromagnetic exchange J_2^{FM} because it is in direct competition with antiferromagnetic superexchange terms coming from second and fourth-order terms (Figure S5(d)). t and t_1 are the nearest and second nearest hopping energies on a honeycomb lattice, while U_0 and U_1 denote the on-site and nearest neighbor Coulomb repulsion. It should be noted that the negative sign of t_1 (see Figure S9(c)) also favors ferromagnetism via the third-order terms present in the equation (10). In deriving the above exchange term using perturbation theory, for simplicity, we have only used density-density repulsion terms up to the nearest neighbor and also ignored the much weaker longer-range hopping. We believe that including the long $1/r$ tail of density-density repulsion and other weaker interaction terms should not change the qualitative picture attained by this analysis.

To estimate J , we calculated the hopping and interaction parameters using the Wannier functions $\Psi_{A(B)}(\mathbf{r})$ (shown in Figure S9(a) and Figure S9(b))[6]. The dependence of these real-space model parameters on the twist angle is shown

in Figure S9(c) and Figure S9(d). The inset of Figure 5d shows the evolution of the spin exchange as a function of ϵ , for a twist angle of 3.6° . The divergence for very small ϵ is expected because all Coulomb interaction terms J_1^{FM}, U_0 and U_1 diverge as ϵ^{-1} . We found that for the twist angle $\theta = 3.6^\circ$, the system favors ferromagnetism up to a large ϵ of nearly 45. The above study explains our experimental finding of the weak ferromagnetic spin-spin correlation for $\nu = -1$ and suggests that direct exchange and second-nearest-neighbor hopping overcome the antiferromagnetic super-exchange interaction.

S4. Circuit model for the nonlinear gating effect

In this section, we build a circuit model to explain the dispersive feature in the $E - V_g$ phase diagram (Figure 2c), as shown in Figure S10. We denote $V_g = \frac{C_t V_{tg} + C_b V_{bg}}{2C_t}$ as half of the effective gate voltage that characterizes the doping density under ideal conditions (with no contact resistance or leak current through gates). $C_t = \frac{\epsilon_h B N \epsilon_0}{d_t}$ and $C_b = \frac{\epsilon_h B N \epsilon_0}{d_b}$ are the geometric capacitance of the top and bottom gates. The tunneling resistances of two dielectric gates (R_t and R_b) decrease when the gate voltages (V_{tg} and V_{bg}) are close to the breakdown limit. Due to the finite contact resistance between t-MoSe₂ and the electrode, the voltage drop on the contact becomes non-negligible. We also denote the leak currents in the top and bottom gates as I_t and I_b , and the voltage drops on the two ideal capacitors as V_t and V_b , respectively.

During the experiment, we use V_g and ΔV to control the doping and the field independently. To be specific, the voltage physically applied on the top and bottom gates is $V_{tg} = V_g + \Delta V$ and $V_{bg} = \frac{d_b}{d_t}(V_g - \Delta V)$. $E = \frac{1}{2}(\frac{V_{tg}}{d_t + d_s + d_0} - \frac{V_{bg}}{d_b}) \approx \frac{\Delta V}{d_t}$ is the electric field applied on the bilayer ($d_t \gg d_0, d_s$). For a fixed filling number ν , the total charge density accumulated in the two capacitors remains constant.

$$C_t V_t + C_b V_b = e \nu n_M \quad (11)$$

Due to the leak current and contact resistance, the voltage drops on the ideal capacitors are smaller than the physically applied voltages:

$$\begin{aligned} V_t &= V_{tg} - I_t R_s - (I_t + I_b) R_c \\ V_b &= V_{bg} - I_b R_s - (I_t + I_b) R_c \end{aligned} \quad (12)$$

I_t and I_b can be determined by the Kirchhoff law.

$$\begin{aligned} V_{tg} &= I_t R_1 + I_b R_c \\ V_{bg} &= I_t R_c + I_b R_1 \end{aligned} \quad (13)$$

We assume a symmetric where $R_t \approx R_b$ and we define $R_1 = R_t + R_s + R_c$

$$\begin{aligned} I_t &= \frac{R_1 V_{tg} - R_c V_{bg}}{R_1^2 - R_c^2} \\ I_b &= \frac{R_1 V_{bg} - R_c V_{tg}}{R_1^2 - R_c^2} \end{aligned} \quad (14)$$

Thus we arrive at an explicit relationship between V_g and E .

$$\begin{aligned}
C_t V_t + C_b V_b &= C_t (V_{tg} - I_t R_s - (I_t + I_b) R_c) + C_b (V_{bg} - I_b R_s + (I_t + I_b) R_c) \\
&= (2 - \frac{(R_1 - R_c) R_c + 2 R_1 R_s}{R_1^2 - R_c^2} - \frac{(R_1 - R_c) R_c - R_1 R_s}{R_1^2 - R_c^2} \frac{C_t^2 + C_b^2}{C_t C_b}) C_t V_g \\
&\quad - \frac{(R_1 - R_c) R_c - R_1 R_s}{R_1^2 - R_c^2} \frac{C_b^2 - C_t^2}{C_b C_t} \epsilon_0 \epsilon_{hBN} E = e \nu n_M
\end{aligned} \tag{15}$$

Since the gate tunneling resistance is much higher than the contact and sample resistance ($R_1 \gg R_c, R_s$) especially when the contact gate is turned on, we can drop the R_c^2 term and further simplify the equation (15).

$$\begin{aligned}
V_g &\approx \frac{1}{(2 - \alpha) C_t} (e \nu n_M + \frac{\epsilon_0 \epsilon_{hBN} (R_s - R_c) (C_t^2 - C_b^2)}{R_1 (C_t C_b)} E) \\
\alpha &= \frac{R_c + 2 R_s}{R_1} + \frac{(R_c - R_s) (C_t^2 + C_b^2)}{R_1 C_t C_b}
\end{aligned} \tag{16}$$

As the equation above shows, when $R_1 \gg R_c, R_s$, V_g does not disperse with E . But when the doping of the sample is close to $1.0 \times 10^{13} \text{cm}^{-2}$, R_1 is reduced to the same level as R_c and R_s . As a result, the voltage level that maintains the same doping level in the sample starts to change with the electric field — an effect known as the nonlinear gating effect in the literature.

Regarding our device D1, C_t is larger than C_b (Table 1), and the sample resistance should be larger than the contact resistance ($R_s > R_c$) among insulating states happening at the high doping level, such as $\nu = -2$. As a result, the $\nu = -2$ state is expected to disperse positively with E , which is consistent with the measurement results. The slope in the low doping region ($\nu < -\frac{1}{2}$) becomes negative, likely because the relationship between R_s and R_c is reversed. The curvy feature in the dispersion should arise from the nonlinear change of the contact resistance.

As a control experiment, when the contact gate is turned off, the contact resistance becomes much larger, and the nonlinear gating effect becomes significantly stronger, as shown in Figure S11(a).

S5. DFT calculation of t-MoSe₂ homobilayer

We have performed a Density Functional Theory (DFT) calculation of the t-MoSe₂ homobilayer to confirm the spatial distribution of the two moiré triions observed (X_1^+ and X_2^+). The simulated moiré superlattice is constructed using two monolayers of MoSe₂ with a twist angle of 3.89° and a lattice constant of 0.33 nm in the primitive cell. Structure relaxation is carried out using force-field calculations implemented in the LAMMPS package[7]. Specifically, the Stillinger-Weber force-field[8, 9] is used to describe the intralayer interactions, while the Kolmogorov-Creps force-field[10] accounts for the interlayer coupling. Electronic structure calculations based on the density functional theory of twisted MoSe₂ (t-MoSe₂) are performed using the SIESTA package[11] with the Perdew-Burke-Ernzerhof functional[12]. The wavefunctions are expanded using a double-zeta plus polarization basis.

Since the optical absorption is dominated by moiré states that are originally folded from the $\pm K$ point of the pristine unit cell, we also calculate the wavefunction of a monolayer MoSe₂ whose band edges are located at the $\pm K$ point of the pristine Brillouin zone. After getting all the wavefunctions, we perform a projection calculation, $|\langle \psi_{vk}^{monolayer} | \psi_{v'k'}^{t-MoSe_2} \rangle|^2$ and $|\langle \psi_{ck}^{monolayer} | \psi_{c'k'}^{t-MoSe_2} \rangle|^2$, to identify the moiré valence and conduction bands folded from the pristine K valley, respectively. The calculation results are summarized in Figure S12. We find that the highest moiré valence band folded from the K valley is ~ 218 meV below the moiré valence band folded from the Γ valley, and the lowest moiré conduction band folded from the K valley is also the moiré conduction band minimum. The states in these moiré bands are consistently localized at the MX and XM stacking regions, suggesting that the bright moiré exciton is located at these sites after folding, and there should therefore exist two types of bright trions, depending on the exciton-carrier distance, consistent with our experimental observations. In addition, the DFT results are also consistent with our PL measurement results (Figure S11(b)), where the indirect Γ -K excitons are about 200 meV lower than the direct K-K excitons.

S6. Interlayer exciton-polaron

To model the reflectance spectra near the attractive and repulsive interlayer Fermi polaron resonances (*IAP* and *IRP*) shown in Figure S16(a), we employ a multi-Lorentzian line shape to fit the experimental data. Specifically, we fit the data using the derivative form:

$$\frac{d}{d\hbar\omega} \frac{\Delta R}{R} = \frac{d}{d\hbar\omega} \left(\frac{A \cos(\alpha) \gamma^2 / 2}{(\hbar\omega - \hbar\omega_0)^2 + \gamma^2 / 4} + \frac{A \sin(\alpha) \gamma (\hbar\omega - \hbar\omega_0)}{(\hbar\omega - \hbar\omega_0)^2 + \gamma^2 / 4} \right) + const$$

Here, A is the oscillator strength, $\hbar\omega_0$ is the resonance energy, and γ is the full linewidth at half maximum. And α is a phase factor that depends on the photon energy and the gate voltage. A representative line cut and its corresponding fit are shown in Figure S16(b). The extracted parameters, such as peak energies, oscillator strengths, and linewidths, are summarized in Figures S16(c)-(f).

As shown in Figure S16(c), both *IAP* and *IRP* resonances exhibit a blueshift with increasing doping. The energy separation between *IRP* and *IAP* increases linearly with gate voltage [Figure S16(d)], from which we estimate the *IAP* binding energy to be approximately 18.4(8) meV. The oscillator strength gradually transfers from *IRP* to the lower-energy branch *IAP* as doping increases, consistent with the expected behavior of Fermi polarons. Furthermore, within this doping window, the linewidths of both *IAP* and *IRP* increase with carrier density, as shown in Figure S16(f).

S7. Device parameters of D1 to D3

In this section, we list the key parameters of the D1-D3 devices, such as the target twist angle, the temperature when the twisted bilayer is fabricated, and the thickness of the hBN gate.

Table 1: Parameters of D1-D3

Device	D1	D2	D3
target angle	3.66	3.50	4.0
calibrated angle	3.8	2.9	4.0
Contact gate	✓	✓	×
Pick-up temperature	60	100	60
top hBN (nm)	6.6	26.5	9.0
bottom hBN (nm)	9.5	30.2	9.6
spacer (nm)	1.0	0.6	1.2

References

- [1] Nitin Kaushal, Nicolás Morales-Durán, Allan H MacDonald, and Elbio Dagotto. Magnetic ground states of honeycomb lattice wigner crystals. *Communications Physics*, 5(1):289, 2022.
- [2] Patrik Fazekas. *Lecture notes on electron correlation and magnetism*, volume 5. World scientific, 1999.
- [3] Thoralf Hanisch, Burkhard Kleine, Afra Ritzl, and Erwin Müller-Hartmann. Ferromagnetism in the Hubbard model: instability of the Nagaoka state on the triangular, honeycomb and kagome lattices. *Annalen der Physik*, 507(4):303–328, 1995.
- [4] Mattia Angeli and Allan H MacDonald. Γ valley transition metal dichalcogenide moiré bands. *Proceedings of the National Academy of Sciences*, 118(10):e2021826118, 2021.
- [5] Yang Zhang, Noah FQ Yuan, and Liang Fu. Moiré quantum chemistry: charge transfer in transition metal dichalcogenide superlattices. *Physical Review B*, 102(20):201115, 2020.
- [6] Nitin Kaushal and Elbio Dagotto. Moiré kanamori-hubbard model for transition metal dichalcogenide homobilayers. *Physical Review B*, 107(20):L201118, 2023.
- [7] Aidan P Thompson, H Metin Aktulga, Richard Berger, Dan S Bolintineanu, W Michael Brown, Paul S Crozier, Pieter J In’t Veld, Axel Kohlmeyer, Stan G Moore, Trung Dac Nguyen, et al. Lammmps-a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer physics communications*, 271:108171, 2022.
- [8] Jin-Wu Jiang. Parametrization of stillinger–weber potential based on valence force field model: application to single-layer MoS_2 and black phosphorus. *Nanotechnology*, 26(31):315706, 2015.
- [9] Frank H Stillinger and Thomas A Weber. Computer simulation of local order in condensed phases of silicon. *Physical review B*, 31(8):5262, 1985.
- [10] Mit H Naik, Indrajit Maity, Prabal K Maiti, and Manish Jain. Kolmogorov–crespi potential for multilayer transition-metal dichalcogenides: capturing

structural transformations in moiré superlattices. *The Journal of Physical Chemistry C*, 123(15):9770–9778, 2019.

- [11] Alberto García, Nick Papior, Arsalan Akhtar, Emilio Artacho, Volker Blum, Emanuele Bosoni, Pedro Brandimarte, Mads Brandbyge, Jorge I Cerdá, Fabiano Corsetti, et al. Siesta: Recent developments and applications. *The Journal of chemical physics*, 152(20), 2020.
- [12] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.