nature portfolio

https://doi.org/10.1038/s41567-021-01359-0

Supplementary information

Interaction-driven band flattening and correlated phases in twisted bilayer graphene

In the format provided by the authors and unedited

Supplementary Information:

Interaction-driven Band Flattening and Correlated Phases in Twisted Bilayer Graphene

Youngjoon Choi, Hyunjin Kim, Cyprian Lewandowski, Yang Peng, Alex Thomson, Robert Polski, Yiran Zhang, Kenji Watanabe, Takashi Taniguchi, Jason Alicea, Stevan Nadj-Perge

1. DEVICE FABRICATION

The device was fabricated using the Polydimethylsiloxane (PDMS)-assisted stack-and-flip technique[S1,S2]. A ∼30nm hBN, monolayer WSe_2 (HQ Graphene), and monolayer graphene are exfoliated on SiO_2 and identified optically. We used a poly(bisphenol A carbonate) (PC)/PDMS stamp to pick up hBN, WSe_2, and used the stack to tear and twist graphene layers[S3], all at ∼90°C. We rotated the two graphene pieces by 1.0° and the final stack attached to PC is then transferred onto a clean PDMS, with TBG side facing the PDMS. PC is disolved in N-Methyl-2-pyrrolidone (NMP) and the stack is transferred on a chip with a graphite gate and gold electrodes. In the final step, TBG is connected to the electrodes by another graphite flake. This fabrication method resulted in high sample quality that gives relatively large (400 nm × 400 nm) contamination-free areas[S2].

2. STM MEASUREMENTS

The STM measurements were performed in a Unisoku USM 1300J STM/AFM system. We used a Pt/Ir tip, first prepared on a Ag(111) crystal by observing quasiparticle interference and a good spectrum that shows Ag(111) surface state. Additional tip cleaning is performed on the chip with TBG using the gold electrode surface to pulse the tip[S4] until a clean tip is restored. All reported features are observed with >10 different microtips. Unless specified otherwise, parameters for dI/dV spectroscopy measurements were $V_{\text{Bias}} = 100$ mV and $I = 1$ nA, and the lock-in parameters were modulation voltage $V_{\text{mod}} = 0.5 - 1$ mV and frequency $f = 973$ Hz. The piezo scanner is calibrated on a Pb(110) crystal by matching the lattice constant and verified by measuring the distance between carbon atoms. The twist-angle uncertainty is approximately ±0.01°, and determined by measuring distances between AA sites in topography at 2 K. Filling factor assignment has been performed by taking Landau fan diagrams as discussed previously[S2]. At higher temperatures, where piezo calibration is known to change, the twist angle is deduced similarly as in transport measurements by determining V_{Gate} values at which full filling of the flat bands ($\nu = \pm 4$) is reached.

3. EFFECT OF WSE_2 SUBSTRATE

In our work, we used WSe_2 as a substrate, contrary to more widely used MATBG samples that employ an hBN substrate. This method allows us to find large areas (400 nm×400 nm) that have small twist angle disorder which we believe is a mechanical effect coming from WSe_2. Our previous study[S3] also pointed out the effect of spin-orbit interaction in this system and extracted the energy scale to be of order 1 meV. This scale is similar to the Landau-level broadening observed in our STM spectroscopy at 2K, and it is thus more challenging to observe SOC-related features with STM measurements employed here.

The interaction energy scale and the corresponding filling-dependent band flattening is about 20 meV, an order of magnitude larger than SOC. We anticipate that the interaction-driven band flattening physics described in this manuscript does not change in the absence or presence of SOC. Previous observations of Chern insulating states[S5] and the cascade of symmetry-breaking transitions[S6] in samples that do not use WSe_2 substrate support our argument. We also made a second device where TBG is directly placed on hBN (device #2) to have direct comparison with the data taken with a WSe_2 substrate at similar twist angle (Supplementary Fig. S4). Supplementary Fig. S4b shows band flattening effects that are clearly visible at 1.3°. Comparing to the data taken on WSe_2 (Supplementary Fig. S4c) we observed no qualitative difference when TBG is placed on hBN and WSe_2. We leave a more detailed comparison between TBG devices on hBN and WSe_2 for future studies.

4. EFFECT OF HETEROSTRAIN ON THE VHS SEPARATION

Gate spectroscopy taken at different spatial points were compared to check the effect of heterostrain on the VHS separation around the magic angle (Supplementary Fig. S11). We took the VHS separation at full filling to exclude additional increase in the separation that happens at partial flat-band filling due to interactions. Supplementary Fig. S11f shows that the VHS separation increases by 3.1 ± 0.5 meV per increase in 0.1% heterostrain in the twist angle range 1.01°∼1.06°. A similar value has been predicted from theory[S7] and confirmed through experiments. Because our sample has heterostrain of $0.15 - 0.3\%$, we expect an enhanced VHS separation of around $4 - 8$ meV. This enhancement could potentially affect the onset of the cascade, but the role of the Hartree-driven band flattening remains. Robustness of the band-flattening physics stems from the fact that the nature of the inhomogenous charge distribution coming from different momentum points persists in a strained system as well. Such inhomogeneity is corroborated through gate spectroscopy taken at two different points having the same twist angle but different heterostrain; both points show similar real-space charge profiles in γ and κ Landau levels (Supplementary Fig. S12). The interaction-driven band flattening mechanism is thus valid within the range of heterostrain where the measurement is taken as supported by further theory simulations (see Supplementary Fig. S14 and further discussion in Sec. 8).

5. BAND FLATTENING AND MAGNETIC FIELD

Since the observed band flattening is detected by Landau Level spectroscopy, one may ask whether this is only a high field phenomena or holds also for B = 0T. Theory of Hartree corrections strongly suggests that flattening occurs regardless of the field[S8–S10]. Nevertheless, this question remains relevant near the magic angle where finite $B > 0$ can possibly change the system's ground state.

This high-field regime is mostly relevant when the field is large enough that the magnetic length is comparable to the moiré lattice scale. The effect of correlations can be accounted for in part by tracking the behaviour of the band-flattening mechanism as the magnetic field is lowered. We are able to track the evolution of the γLLs as a function of twist angle down to B = 3T (see Supplementary Fig. S11d).

Even at this much lower field (significantly lower than B = 7T), the observation are qualitatively the same: we observe similar band flattening, and correlated gaps develop

only for angles where the γLLs are fully merged with the VHS (at the corresponding filling factors). In the regime under consideration, Hofstadter physics does not dominate, and magnetic fields cannot significantly perturb the electrostatic interactions responsible for band flattening. The first point is illustrated by noting that the magnetic length, $l_B = \hbar/(eB) \approx 26\text{nm}/[B(\text{Tesla})]^{1/2}$, is larger than the moiré length scale, $l_m \approx 10 - 13\text{nm}$, for magnetic fields less than $B < 4\text{T}$. The second statement follows by noting that, at these scales, the magnetic field may be treated semiclassically (as long as there is no symmetry-breaking transitions), allowing us to determine the LL spectrum directly from quantized orbits of the B=0T band structure using the Bohr-Sommerfeld quantization rules (with some modifications due the Dirac physics). More precisely, the magnetic field in this case can be included as a perturbation to the $B = 0\text{T}$ Hamiltonian that mixes the eigenstates within a certain energy window roughly determined by the LL energy separation, which is at most 2-3 meV in this case (at B=3T).

Since we clearly resolve γLLs and κLLs in the majority of the twist-angle range close to the magic angle, there cannot be substantial mixing between eigenstates originating from different pockets in the mini-Brillouin zone (MBZ). This picture holds, to some degree, even in case of the symmetry-breaking cascade where gaps are present and γLLs are obscured. Even in this case, κLLs remain visible around charge neutrality, indicating that there is not large hybridization between different parts of the MBZ. As emphasized, the Hartree-induced band-flattening is a consequence of the different spatial profiles of the eigenstates at the γ and κ point of the MBZ; the absence of magnetic field-induced mixing between states in different part of the MBZ therefore suggests that the electrostatic interaction has a similar effect at $B = 3\text{T}$ as it does at $B = 0\text{T}$. Also, as shown before (in Ref. S2) κLLs can be observed down to even much smaller fields of $B = 0.5\text{T}$, further suggesting that there is no significant mixing of κLLs with other parts of the MBZ.

Finally, we note that development of the LLs for all angles happens gradually and continuously, whereas large modifications of the band structure with magnetic field would likely manifest themselves as some abrupt jumps with field. The continuous evolution holds even in the case where Chern-insulating phases are present. In this case we see gaps opening close to the Fermi energy, but tens of meV away from it evolution of the features (LLs and VHS) is smooth upon field change. This further corroborates that the band flattening is valid for a range of magnetic fields (at least for $B < 4\text{T}$).

6. MODELLING: MEAN-FIELD HAMILTONIAN

This section reviews the continuum model used to capture the non-interacting band structure of TBG and explains how the Hartree and Fock corrections are calculated within that continuum-model framework.

We describe the non-interacting TBG band structure using the Bistritzer-MacDonald model[S11], adopting the parametrization and notation of Ref. S12. For completeness we summarize here the key elements of the modelling and underlying assumptions. In our analysis we study a small range of twist angles $0.8^\circ < \theta < 1.32^\circ$ in vicinity of the magic angle regime. To construct the non-interacting Hamiltonian H_0, we start by taking two AA-stacked graphene layers and rotating layers 1 and 2 by $-\theta/2$ and $\theta/2$, respectively. For the range of twist angle of interest, the moiré real-space lattice constant $L_M = a/2\sin(\theta/2)$ ranges from ~ 18 nm to ~ 11 nm. This distance is two orders of magnitudes greater than graphene's lattice constant $a = 0.246$ nm—justifying the use of the continuum approach.

In momentum space, the layer rotation translates into two graphene Brillioun zones rotated by an angle θ relative to each other. Both Brillouin zones (BZs) are centered at the same Γ point but the K (and K') points of the two layers are separated by a small momentum $4\pi/(3L_M)$. As the moiré periodicity L_M is much greater than the lattice constant a, we can ignore mixing between the two valleys K and K'—labeled $\xi = -1, 1$ below—of the original graphene layers. The non-interacting Hamiltonian H_0 therefore becomes block diagonal in the valley index. The blocks $H_0^{(\xi)}$ describing each of the two valleys take the form

$$H_0^{(\xi)} = \begin{pmatrix} H_1 & U^\dagger \\ U & H_2 \end{pmatrix} \tag{1}$$

in the basis (A_1, B_1, A_2, B_2), where A, B denotes sublattice and $1, 2$ subscripts are layer indices. The matrices H_l correspond to intralayer Hamiltonians for layers $l = 1, 2$ and, due to the length-scale separation between L_M and a, can be approximated by performing a standard $k \cdot p$ expansion around the points K and K'. This $k \cdot p$ procedure yields 2×2 Dirac Hamiltonians centered at $\boldsymbol{K}_\xi^{(l)}$ points (defined below):

$$H_l = -\hbar v \left[R\left(s_l \theta/2\right) \left(\boldsymbol{k} - \boldsymbol{K}_\xi^{(l)}\right) \right] \cdot \left(\xi \sigma_x, \sigma_y\right), \tag{2}$$

where $\boldsymbol{k}$ is a momentum in the BZ of the original graphene layers, $R\left(\phi\right)$ is the 2×2 2D counterclockwise rotation matrix that accounts for rotation of the BZs for the original graphene layers, and $s_{l=1} = +1$ while $s_{l=2} = -1$. For all twist angles we use $\hbar v/a = 2.1354$ eV as the energy scale for the Hamiltonians H_l. The vectors $\boldsymbol{K}_1^{(l)}$, $\boldsymbol{K}_{-1}^{(l)}$, which denote the Dirac points K and K' of the two layers, are given by

$$\boldsymbol{K}_\xi^{(1)} = -\xi \frac{4\pi}{3a} R\left(-\theta/2\right) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{K}_\xi^{(2)} = -\xi \frac{4\pi}{3a} R\left(\theta/2\right) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{3}$$

respectively. In the analysis below the moiré superlattice BZ is defined using two reciprocal lattice vectors

$$\boldsymbol{G}_1^M = -\frac{2\pi}{\sqrt{3} L_M} \begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}, \quad \boldsymbol{G}_2^M = \frac{4\pi}{\sqrt{3} L_M} \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{4}$$

We denote the reciprocal lattice vector length as $G_M = |\boldsymbol{G}_1^M| = |\boldsymbol{G}_2^M| = 4\pi/\sqrt{3} L_M$. Interlayer hopping is encoded by the matrix U given by

$$U = \begin{pmatrix} u & u' \\ u' & u \end{pmatrix} + \begin{pmatrix} u & u' e^{-i2\pi\xi/3} \\ u' e^{i2\pi\xi/3} & u \end{pmatrix} e^{i\xi \boldsymbol{G}_1^M \cdot \boldsymbol{r}} + \begin{pmatrix} u & u' e^{i2\pi\xi/3} \\ u' e^{-i2\pi\xi/3} & u \end{pmatrix} e^{i\xi \left(\boldsymbol{G}_1^M + \boldsymbol{G}_2^M\right) \cdot \boldsymbol{r}}. \tag{5}$$

We treat the interlayer couplings u and u' as fitting parameters for the band structure according to the procedure explained below. To determine the energy bands and the eigenstates we take the Bloch wavefunction ansatz for valley ξ as

$$\Psi_{\xi, n, \boldsymbol{k}}^X(\boldsymbol{r}) = \sum_{\boldsymbol{G}} C_{\xi, n, \boldsymbol{k}}^X(\boldsymbol{G}) e^{i(\boldsymbol{k} + \boldsymbol{G}) \cdot \boldsymbol{r}}, \tag{6}$$

where $X = A_1,\ B_1,\ A_2,\ B_2$ labels the spinor components, n is a band index, and $\boldsymbol{k}$ is the Bloch wave vector in the BZ of the original graphene layers. The $\boldsymbol{G}$ sum runs over all possible integer combinations of the reciprocal lattice vectors $\boldsymbol{G} = m_1 \boldsymbol{G}_1^M + m_2 \boldsymbol{G}_2^M$ with integer m_1 and m_2.

We now proceed to include Coulomb interactions into the Hamiltonian. Our approach follows that of Refs. S8,S13,S14. The full Hamiltonian is given by

$$\hat{H} = \hat{H}_0 + \hat{H}_C \,, \tag{7}$$

where $\hat{H}_0$ corresponds to the non-interacting part and $\hat{H}_C$ accounts for interactions. They are given by

$$\hat{H}_0 = \sum_{\xi,\sigma} \int_\Omega d^2\boldsymbol{r} \Psi^\dagger_{\xi,\sigma}(\boldsymbol{r}) H_0^{(\xi)} \Psi_{\xi,\sigma}(\boldsymbol{r}) \,, \quad \hat{H}_C = \frac{1}{2} \int_\Omega d^2\boldsymbol{r} d^2\boldsymbol{r}' \delta\rho(\boldsymbol{r}) V_C(\boldsymbol{r}-\boldsymbol{r}') \delta\rho(\boldsymbol{r}') \,. \tag{8}$$

Here $\Psi_{\xi,\sigma}(\boldsymbol{r})$ is a fermion spinor operator in the basis (A_1, B_1, A_2, B_2) of the non-interacting Hamiltonian of Eq. (1); the summation in $\hat{H}_0$ ranges over valleys $\xi = \pm 1$ and spins $\sigma = \uparrow, \downarrow$; real space integration is over a moiré unit cell Ω; $\delta\rho(\boldsymbol{r}) = \sum_{\xi,\sigma} \Psi^\dagger_{\xi,\sigma}(\boldsymbol{r}) \Psi_{\xi,\sigma}(\boldsymbol{r}) - \rho_{CN}(\boldsymbol{r})$ tracks the density relative to that of non-interacting TBG at charge neutrality, $\rho_{CN}(\boldsymbol{r})$; and finally, $V_C(\boldsymbol{r})$ is the Coulomb potential, whose Fourier transform we take as $V_C(\boldsymbol{q}) = 2\pi e^2/\epsilon q$. We treat ϵ, which encodes dielectric screening by the substrate, as a free parameter.

We treat interactions at a mean-field level and approximate the Hamiltonian in Eq. (7) as

$$\hat{H} \approx \hat{H}_0 + \hat{H}_H + \hat{H}_F - E_{int}, \tag{9}$$

where $\hat{H}_0$ is the non-interacting Hamiltonian as before, $\hat{H}_H$ and $\hat{H}_F$ respectively encode Hartree and Fock terms. Here, E_{int} corresponds to the constant energy shift equal to half of the expectation of $< \hat{H}_H + \hat{H}_F >$ with respect to the ground state, which corrects for double-counting in the mean-field theory (see section 11 on Cascade modelling for further discussion and Ref.S14 for similar procedure applied to a TBG Hamiltonian). We now proceed to analyze $\hat{H}_H$ and $\hat{H}_F$ individually, detailing their precise form and clarifying their role in altering the band structure.

The Hartree term is given by a contraction of operators at the same position in $\hat{H}_C$ of Eq. (8). Explicitly, we have

$$\hat{H}_H = \sum_{\xi,\sigma} \int_\Omega d^2\boldsymbol{r} \Psi^\dagger_{\xi,\sigma}(\boldsymbol{r}) V_H(\boldsymbol{r}) \Psi_{\xi,\sigma}(\boldsymbol{r}) \,, \quad V_H(\boldsymbol{r}) = \int_\Omega d^2\boldsymbol{r}' V_C(\boldsymbol{r}-\boldsymbol{r}') \left\langle \delta\rho(\boldsymbol{r}') \right\rangle_{occ.} \,, \tag{10}$$

where the Hartree potential $V_H(\boldsymbol{r})$ contains a summation over all spins and valley implicitly through the definition of $\delta\rho(\boldsymbol{r})$. The symbol $\langle \dots \rangle_{occ.}$ denotes a summation over occupied states measured from the CNP ($\nu = 0$)S8; by construction the Hartree potential therefore vanishes at charge neutrality. For convenience of implementation we rewrite Eq. (10) in the crystal momentum basis $\boldsymbol{k}$ of bare graphene:

$$\begin{aligned} \left\langle \boldsymbol{k}+\boldsymbol{G}, s, X \left| \hat{H}_H \right| \boldsymbol{k}+\boldsymbol{G}', s', X' \right\rangle =& \delta_{s,s'} \delta_{X,X'} V_C(\boldsymbol{G}-\boldsymbol{G}') \times \\ & \times \sum_{s'',\boldsymbol{k}',\boldsymbol{G}'',X''}{}' \left(C^{X''}_{s''\boldsymbol{k}'}(\boldsymbol{G}'') \right)^* C^{X''}_{s''\boldsymbol{k}'}(\boldsymbol{G}''-\boldsymbol{G}+\boldsymbol{G}') \,, \end{aligned} \tag{11}$$

where for clarity we introduced a composite index $s = (\xi, \sigma)$ tracking the valley/spin degree of freedom. In the above expression the $\sum'$ denotes summation over occupied states for a given filling measured with respect to the CNP as explained previously. As argued in

Refs. S8,S13,S14 since the Hartree potential is controlled primarily by the contribution of the first-star of reciprocal lattice vectors, it is sufficient to consider combinations of $\boldsymbol{G}$, $\boldsymbol{G}'$ satisfying $\boldsymbol{G}-\boldsymbol{G}'=m\boldsymbol{G}_1^M+n\boldsymbol{G}_1^M$ with $(m,n)=\{(\pm1,0),(0,\pm1),(\pm1,\pm1)\}$. Note that as usual the $\boldsymbol{G}=\boldsymbol{G}'=0$ contribution is excluded since it is cancelled by the positive ionic background.

Before detailing the self-consistent procedure, we discuss first the physics behind the Hartree renormalization of the band structure[S8–S10]. When electrons are added to a charge-neutral system starting from the κ points of the mini-Brillouin zone, in real space these additional states generate a charge density buildup at AA sites. In contrast only near large fillings—when the flat electron bands are fully filled, i.e., near the γ point of the mini-Brillouin zone—do electrons contribute to charge density buildup away from the AA sites. This non-uniform spatial charge distribution implies a large potential energy build-up and a corresponding significant Hartree potential. The role of Hartree, which reflects the triangular symmetry of the effective TBG lattice, is to redistribute charge density further away from the AA sites. In momentum space this redistribution translates to an effective potential that lowers the energy of states located near the γ points; see Fig. S16.

As argued in the main text and demonstrated in Fig. S16, the closer the twist angle comes to the magic-angle (at which the non-interacting bandwidth is minimized), the more pronounced is the effect of Hartree. Consequently, for each filling ν there is a critical twist angle $\theta_{\text{crit}}(\nu)$ at which the band becomes flattened near the γ point. Such Hartree-induced band renormalization shifts the VHS in the density of states away from the location dictated by the parameters of the non-interacting band structure and enhances the density of states at the Fermi level as shown in Fig. S13. Experimentally, this effect is seen in the LL sweeps of Figs. 1 and 2 as the filling at which the top-most LL from the γ point joins with the signal arising from the VHS. Below the critical angle for a given filling, the band structure will continue becoming more and more inverted at the γ point, ultimately producing enormous inversions[S10,S13] giving rise to experimentally inconsistent band structures. Furthermore, as we will see in the following section, these large band structure inversions result in huge kinetic energies that would prohibit the appearance of cascade transitions.

To avoid the appearance of these experimentially inconsistent effects of Hartree on the band structure, we argue that it is necessary to include Fock corrections (See section Sec.7 for discussion of other mechanisms). Within the mean-field formalism, the Fock term $\hat{H}_F$ in Eq. (9) becomes

$$\hat{H}_F=\sum_{X,X',s}\int_\Omega d^2\boldsymbol{r}d^2\boldsymbol{r}'\left(\Psi_s^X(\boldsymbol{r})\right)^\dagger V_{F;s}^{XX'}(\boldsymbol{r},\boldsymbol{r}')\Psi_s^{X'}(\boldsymbol{r}')\,, \tag{12}$$

corresponding to a Wick contraction between creation and annihilation operators at $\boldsymbol{r}$ and $\boldsymbol{r}'$, respectively. Here the index $s=(\xi,\sigma)$ once again tracks the valley/spin degree of freedom. The function $V_{F;s}^{XX'}(\boldsymbol{r},\boldsymbol{r}')$ denoting the non-local Fock potential is given by:

$$V_{F;s}^{XX'}(\boldsymbol{r},\boldsymbol{r}')=-V_C(\boldsymbol{r}-\boldsymbol{r}')\left\langle\left(\Psi_s^{X'}(\boldsymbol{r}')\right)^\dagger\Psi_s^X(\boldsymbol{r})\right\rangle_{occ.}\,. \tag{13}$$

Note how the Fock potential, unlike the Hartree one, does not contain a summation over valley/spin degree of freedom. Furthermore, due to the block-diagonal nature of the non-interacting Hamiltonian, Eq. (1), neither the Fock potential nor the mean-field Fock term, Eq. (12), contains any inter-flavour terms. We also do not allow such inter-flavour terms to

be generated spontaneously, in line with Ref. S15,S16 but unlike the analysis of Ref. S17. This assumption is justified as we are focusing on qualitative changes to the band structure (e.g., dispersion renormalization and cascade onset) as opposed to resolving the precise nature of correlated insulators. Finally the notation $\langle \dots \rangle_{occ.}$ corresponds to a summation over occupied states, but as different conventions are present in the literature we relegate discussion of this detail to the following section that describes the modelling process. As above, for convenience of notation we rewrite Eq. (13) in the crystal-momentum basis $\boldsymbol{k}$ of bare graphene,

$$\left\langle \boldsymbol{k}+\boldsymbol{G}, s, X \left| \hat{H}_F \right| \boldsymbol{k}+\boldsymbol{G}', s', X' \right\rangle = -\delta_{s,s'} \sum_{\boldsymbol{k}',\boldsymbol{G}''}'' V_C(\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G}'-\boldsymbol{G}'')\times \\ \times \left(C^{X'}_{s\boldsymbol{k}'}(\boldsymbol{G}'') \right)^* C^{X}_{s\boldsymbol{k}'}(\boldsymbol{G}''+\boldsymbol{G}-\boldsymbol{G}'), \quad (14)$$

where the $\sum''$ denotes summation over occupied states in a manner to be prescribed later. We stress how, due to the non-local nature of the Fock potential, the self-energy explicitly depends on the crystal momentum $\boldsymbol{k}$. This dependence imposes significant numerical complexity for self-consistent calculations since [unlike the Hartree form, Eq. (11)] here the self-energy due to the Fock term has to be determined separately for each momentum $\boldsymbol{k}$.

We conclude this section by clarifying the qualitative role that the Fock term plays in altering the band structure. As argued in the main text and demonstrated in Fig. S16, when Fock corrections are included in the calculation of the self-consistent band structure, band inversions stemming from the Hartree potential are counteracted as has been noted recently[S16]. Two key mechanisms are at work: (i) an overall increase in bandwidth due to Fock corrections and (ii) an "opposite" contribution of the Fock term, Eq. (14), to the self-energy as compared to the Hartree term, Eq. (11). We stress that the qualitative behavior of the Hartree and Fock terms, e.g., their contribution for flattening the band structure in opposite directions, is a general feature not dependent on the particular model used.

7. FOCK CORRECTIONS IN TBG

In this section, we motivate the need for including Fock corrections in capturing qualitative trends seen in the experiment. We focus on four observations that are most easily explained with the help of a Fock term, but we also comment on other mechanisms that could provide a plausible explanation.

First, a fundamental assumption of Hartree-only theory is that corrections are absent or nearly absent at charge neutrality, as recognized already in earlier theory works from continuum model[S8,S13,S14], effective tight-binding[S9] and ab-initio[S10] calculations. The absence of Hartree corrections is exact in a particle-hole (PH) symmetric system. And since the majority of the moiré unit cell in TBG is either AA or AB/BA stacked, each of which is PH symmetric, it is reasonable to expect an approximate PH symmetry in TBG. The near-PH symmetry of the continuum model (in its simplest versions) supports this simplification. As a result, the bandwidth—which in STM is measured by the van Hove to van Hove separation—should be given charge neutrality by that of the non-interacting band structure. Experiments, by contrast, resolve far broader bandwidths. The closer the twist angle comes to the magic angle regime, the more pronounced this effect is. This behavior persists at

finite fillings, whilst Hartree-only theory would predict a narrowing of the van Hove to van Hove separation at finite filling.

Secondly, the $B = 0$ STM results[S1,S2,S6,S18,S19] demonstrate a significant broadening of each van Hove singularity around the charge neutrality point. These widths become substantially narrowed close to full filling—a feature that is hard to reconcile with Hartree-only theory.

Third, due to the local-probe nature of STM, it is possible to image LLs at different locations in the moiré unit cell. As a result, as argued in the main text, one can disentangle whether a given LL originates from κ points of the mini Brillouin zone (more signal at the AA sites) or the γ point of the mini Brillouin zone (more signal at the AB sites). This allows us to identify whether a LL comes from a Dirac point or the γ-pocket and how these LLs are located energy-wise with respect to the van Hove singularity. We find that a LL from the γ-pocket never drops significantly below the van Hove singularity (in fact as soon as a LL from the γ-pocket at a given filling merges with the van Hove singularity a cascade behavior is seen to occur, c.f. Fig. 3 and the surrounding discussion in the main text); the same is true for LLs originating from the Dirac points. Since Hartree theory seems to predict an ever-increasing band inversion[S8,S10,S13,S14], where the γ pocket falls below the Dirac points, such a band structure appears incompatible with the measured LL spectrum.

Finally, when Hartree only corrections are included, the resulting band inversions considerably increase the kinetic energy of filled electronic states (See the main text and SI, section 11). In such situations, our calculations of Fig. 3 and Fig. S16, suggest that cascade of electronic transitions becomes unfavorable, contrary to experimental observations.

The first three points of this list are independent of our modeling and rely only on the experimental data and the existing body of theory work. From this reasoning, we therefore conclude that another mechanism needs to be at play to reconcile these findings—a most natural one being Fock corrections (see. Fig. S15a to demonstrate how these claims are met within our treatment). It is, however, possible that other mechanisms such as the interplay of strain with interactions (see Fig. S15b-c and Ref. S20 for a discussion of a possible quantum phase transition), an effective Hartree+U theory[S21], or self-consistent screening (RPA-like self-energy corrections), could also explain some of the experimental observations.

8. MODELLING: CHOICE OF PARAMETERS

Here we complete the analysis of Sec. 6 by detailing the procedure used to fit the parameters of the continuum model, specific approximations taken to model Hartree and Fock potentials, as well as the self-consistent procedure. Our main objective is to provide an experimentally inspired modelling of the qualitative behavior of band structure and corresponding Bloch wavefunctions as a function of twist angle and filling, building off of existing work in the literature that analyzes Hartree-Fock contributions (e.g., Refs. S13,S14,S16,S17,S22,S23). In no way does the procedure detailed here suggest that a more comprehensive treatment of Fock corrections is unnecessary. Most importantly, we do not address here the detailed nature of polarized states and the presence or absence of Fock-induced gaps at integer fillings.

The non-interacting continuum model as reviewed in the previous section has a range of different parameters, notably the moiré interlayer couplings u, u', bare graphene velocity v, and twist angle θ. Furthermore, the non-interacting Hamiltonian can be further augmented to include non-local moiré tunneling parameterized by another series of coefficients. Al-

though dependence of most of these parameters with twist angle has been studied through ab-initio methods[S24,S25], making it in principle possible to include within the mean-field analysis, here we choose a simpler approach intended to highlight the important interaction-driven qualitative changes to the band structure. To that end we assume a simple form of the moiré interlayer coupling of Eq. (5) characterized by two parameters u and u' corresponding to same-sublattice and opposite-sublattice interlayer tunneling. Their ratio $\eta = u/u'$ is denoted in the literature as a relaxation parameter. To fix these two parameters u and u' we focus on the largest experimentally measured angle of $\theta = 1.32°$ where the role of interactions is least important. We fix $u' = 90$ meV and $u = 0.4u'$ to correctly reproduce the LL spectrum at $\nu = 0$ as seen in Fig. 1. In the following analysis we maintain these same u and u' values for all twist angles considered. This approximation misses the subtle role relaxation physics plays on increasing the ratio η as the twist angle is brought closer to the magic angle[S24]. Qualitatively, reducing the twist angle still decreases the non-interacting bandwidth; however, the magic angle (defined as the twist angle θ_M at which the non-interacting bandwidth is minimized) occurs at $\theta_M \approx 0.99°$. This value differs modestly from the typical magic-angle values in the literature of $\theta_M \approx 1.1°$.

We now proceed to parametrize the strength of the dielectric screening ϵ in the Coulomb potential $V_C(q) = 2\pi e^2/\epsilon q$. Nominally a value set by the substrate should be used, in this case corresponding to a mean value of the dielectric constants of the materials surrounding TBG, i.e., $\epsilon \approx 5$. This value, however, massively overestimates the role of Hartree and Fock processes, leading to band structures with large γ point inversions (even at $\theta = 1.32°$) that are not observed experimentally. Such behavior was also seen in existing theory literature, prompting authors[S8,S13,S16] to use a range of values ranging from 5 to 66. In the same spirit we choose $\epsilon = 15$ to quantitatively capture the following three experimental characteristics at $\theta = 1.32°$: (i) the energy spacing between LLs arising from the γ point band structure, (ii) the energy spacing between the highest energy LL from the flat band and the lowest energy LL from the dispersive bands, and (iii) the critical angle $\theta_{\text{crit}}(\nu = 4)$ at which γLL_0 joins the VHS. These criteria are met by $\epsilon = 15$ which we then keep constant for all values of θ studied. On theory grounds $\epsilon > 5$ is expected due to screening originating from the dispersive bands (typically on the order of 10) which when combined with that of the substrate (typically for hBN 5) gives the $\epsilon \approx 15$.[S10,S17]. With this parametrisation we find that including Hartree-only adequately captures experimental observations as described in the main text, suggesting that the Fock term plays a subdominant role at least at this large twist angle as indeed seen by no increase in the van Hove-van Hove separation at the charge neutrality point.

Next we explain the procedure behind modelling the Fock potential. As noted in the previous section, the Fock potential is non-local and thus carries a high computational cost, making a parameter sweep like that in Fig. S16 prohibitive. Motivated by the physical intuition that the role of Fock is to oppose the Hartree potential, further substantiated by recent works[S16], we approximate Eq. (14) as

$$\left\langle \boldsymbol{k}+\boldsymbol{G}, s, X \left| \hat{H}_F \right| \boldsymbol{k}+\boldsymbol{G}', s', X' \right\rangle \approx -\delta_{s,s'} g(\theta) V_C(\boldsymbol{G}-\boldsymbol{G}') \times \\ \times \sum_{\boldsymbol{k}',\boldsymbol{G}''}'' \left(C^{X'}_{s\boldsymbol{k}'}(\boldsymbol{G}'') \right)^* C^{X}_{s\boldsymbol{k}'}(\boldsymbol{G}''+\boldsymbol{G}-\boldsymbol{G}') \,. \tag{15}$$

In particular, we replaced the non-local dependence with a constant, possibly twist-angle dependent prefactor $g(\theta)$. The characteristic energy scale of the Fock interaction is set by the Hartree potential. Note that despite similarity to a local approximation, e.g., $V_C(\boldsymbol{r}) \propto \delta(\boldsymbol{r})$,

the form of Eq. (15) carries the additional dependence on reciprocal momenta $\boldsymbol{G}, \boldsymbol{G}'$. In analogy to the Hartree potential, we limit the reciprocal momenta included in the self-energy to a difference $\boldsymbol{G} - \boldsymbol{G}'$ residing in the first star of reciprocal lattice vectors. Unsurprisingly, the Fock potential so constructed acts in opposition to the Hartree potential due to the overall minus sign above. This approximation captures the qualitative behavior of Fock in preventing large Hartree-driven band inversions, capturing also also bandwidth broadening (see Fig. S15) albeit not as pronounced as reported by other authors, e.g., Refs.[S13,S14,S16]. Note also that the above form would not yield a complete cancellation of the contribution from the Hartree potential, even for $g(\theta) = 4$, as both Hartree and Fock potentials exhibit different dependence on spinor indices.

We considered a few forms for the prefactor of the Fock interaction $g(\theta)$. A twist-angle-independent prefactor does not manage to capture the qualitative band structure behavior seen in experiments: A constant value that is too large prevents unphysical band inversion near the magic angle, yet also prevents band flattening for $\nu = 3, 4$ seen in the experiment at intermediate angles that could be recovered with a Hartree-only analysis. A too-small constant value allows for band flattening seen in experiments, but does not prevent unphysical band inversions near the magic angle. To remedy these issues we choose $g(\theta) = 0$ for $\theta \geq 1.14°$ and $\theta \leq 0.84°$, whereas for $0.84° \leq \theta \leq 1.14°$ range we assume a triangular profile with a maximum of $g(\theta = 0.99°) = 1.875$. We verified that the qualitative trends seen and explained in the main text are not fine-tuned and indeed stem from band-flattening physics.

A final ingredient in the description of the Fock term is the specification of what bands are included in the summation $\sum''$ in Eqs. (14) and (15). Several summation schemes are present in the literature with the key difference being whether only flat bands[S13,S14] or also dispersive bands[S17,S23] are included. Qualitatively, at $\nu = 4$ the Fock term arises predominantly from the flat bands; on the other hand, at $\nu = -4$ the dominant contribution arises from the dispersive valence bands. The contribution from the dispersive valence bands is also opposite in sign to that of the flat bands as seen in Ref. S16. We verified that this behavior holds within our approximation of Eq. (15), but due to the local form of the Fock potential contribution from dispersive bands at $\nu = -4$, one finds gap opening near the κ, κ' points beyond what is experimentally plausible. To qualitatively capture the above sign trend in the Fock term while mitigating the preceding gap issue, we include in the summation $\sum''$ the flat bands both at $\nu = 4$ and $\nu = -4$, but change the sign of the contribution at $\nu = -4$. For other fillings, motivated by Ref. S16, we interpolate our solution as explained next.

As the contribution of the Hartree and Fock terms to the self-energy depends linearly (in the absence of cascade and gap opening at the CNP) on filling[S10,S13,S16], for ease of computation we linearly interpolate between the solutions at $\nu = 4$ and $\nu = -4$. We verified that the results do not qualitatively change if a self-consistent approach is used at every filling. To that end, the mean-field interacting Hartree-Fock Hamiltonians for filling ν are taken to be

$$\hat{H}_F(\nu) = \frac{1}{2}\left[\left(\hat{H}_F(\nu = 4) + \hat{H}_F(\nu = -4)\right) + \frac{\nu}{4}\left(\hat{H}_F(\nu = 4) - \hat{H}_F(\nu = -4)\right)\right] \tag{16}$$

$$\hat{H}_H(\nu) = \frac{1}{2}\left[\left(\hat{H}_H(\nu = 4) + \hat{H}_H(\nu = -4)\right) + \frac{\nu}{4}\left(\hat{H}_H(\nu = 4) - \hat{H}_H(\nu = -4)\right)\right]. \tag{17}$$

For the simulations presented in this paper, we evaluate the solutions at $\nu = 4$ and $\nu = -4$ self-consistently until convergence is reached. A grid of 1681 k-points is used for the analysis with the convergence reached after a few iterations.

9. MODELLING: LANDAU-LEVEL SPECTRUM AND THE ROLE OF DISPLACEMENT FIELD

To simulate the Landau-level spectrum shown in Fig. 1f and Fig. S3c, we take a two step approach: (i) Hartree potential is computed at $B = 0$ as detailed in Sections 6-8; (ii) such Hartree potential term is then treated as constant and used in computing the LL spectrum. In what follows we describe details of the numerical procedure used, analyze the role of displacement field, and justify this two-step procedure.

9.1. Non-interacting continuum model with magnetic field

The noninteracting continuum model of TBG with magnetic field was introduced in Refs. S26,S27, and was also discussed in the supplemental material of our previous work[S2]. We consider a perpendicular magnetic field $\boldsymbol{B} = B\hat{\boldsymbol{z}}$, which gives rise to a vector potential $\boldsymbol{A} = B(-y, 0)$ in the Landau gauge. Via minimal coupling, we substitute $\hat{\boldsymbol{p}} \to \hat{\boldsymbol{p}} + e\boldsymbol{A}$. One can further introduce the Landau level basis $\{|L, \sigma, n, k\rangle\}$, where the integer n is the Landau level index (or harmonic oscillator quantum number), k is the momenta along x direction that counts the Landau level degeneracy, together with the layer and spin indices L and σ. In this basis, the intralayer Hamiltonian can be written in terms of the creation and annhilation operators of the Landau levels, and the inter layer tunneling terms further couples these Landau levels states at different n and k. In defining the non-interacting Hamiltonian in the magnetic field we use the same bandstructure parameters as explained in section 8.

9.2. Adding the Hartree potential

As mentioned in the beginning of this section we use a two step process to compute the LL spectrum with the band-flattening effects included. First we compute the self-consistent bandstructure at $B = 0$ and then compute the corresponding LL spectrum. The validity of this approximation is bolstered by the experimental data wherein the charge density at AA sites is higher than at AB sites even at finite magnetic fields. Therefore the Hartree potential at finite magnetic fields still reflects triangular symmetry of the $B = 0$ effective lattice. We stress also that since experimentally there was no significant enhancement of van-Hove to van-Hove separation at charge neutrality point (an indication of exchange effects) we did not include Fock correctionst in calculating the LL spectrum of Fig. 1.

With this in mind, the Hartree potential can be written as (see section 6)

$$V_H(\boldsymbol{r}) = V_0 \delta\rho_G \sum_G e^{i\boldsymbol{G}\cdot\boldsymbol{r}}\,, \tag{18}$$

where $V_0 = V_C(\boldsymbol{G} - \boldsymbol{G}')$ is the interaction strength [see Eq. (11)], $\delta\rho_G$ is a self-consistent parameter corresponding to the first-harmonic of the charge density $\delta\rho(\boldsymbol{r})$ measured from the CNP, and the sum is over the first star of reciprocal lattice vectors $\boldsymbol{G}$ as explained previously. For every filling, we take $\delta\rho_G$ computed in the system at zero magnetic field, i.e., we take the self-consistent band structure as obtained by following the procedure of section 6. As we are focusing here on a large twist angle, $\theta = 1.32^{\circ}$, where Hartree-term alone

gives a good agreement with experiment, we do not include a Fock term in the LL-spectrum analysis. In the Landau level basis, the term $\exp(i\boldsymbol{G}\cdot\boldsymbol{r})$ has the following matrix elements

$$\langle L',\sigma',n';k'|\exp(-i\boldsymbol{G}\cdot\boldsymbol{r})|L,\sigma,n;k\rangle = \delta_{L,L'}\delta_{\sigma,\sigma'}\delta_{k',k-G_x}e^{-iG_ykl_b^2}e^{-iG_yk_\theta\frac{\sqrt{3}l_b^2}{2}}e^{iG_xG_y\frac{l_b^2}{2}}F_{n'n}(z) \tag{19}$$

with $z=(G_x+iG_y)l_b/\sqrt{2}$, and

$$F_{n',n}(z)=\begin{cases}\sqrt{\frac{n!}{n'!}}(-\bar{z})^{n'-n}\exp(-|z|^2/2)L_n^{(n'-n)}(|z|^2) & n'\geq n\\ \sqrt{\frac{n'!}{n!}}z^{n-n'}\exp(-|z|^2/2)L_{n'}^{(n-n')}(|z|^2) & n'<n.\end{cases} \tag{20}$$

In the above Eq. (19) G_x, G_y denote the x,y components of the reciprocal lattice vector $\boldsymbol{G}$ and $l_b=\sqrt{\hbar/eB}$.

9.3. Role of displacement field

When the backgate voltage $V_{\rm bg}$ is applied, a displacement field between the two graphene layers can be generated. This can be modeled by introducing a potential difference D, which is linear in $V_{\rm bg}$, between the two layers. Namely, we add a term $\pm D/2$, which is diagonal in spin, valley and Landau level indices, to the two layers. Note that a displacement field alone will shift the conduction and valence bands in the same direction, see Fig. S3c, and thus cannot produce the features observed in Fig. 1f.

9.4. Calculation of density of states with finite magnetic field and Hartree potential

To calculate the density of states as a function of $V_{\rm bias}$ as well as the filling factor ν Fig. 1f, we need use the fact that $\nu\propto V_{\rm bg}\propto D$. Since $\delta\rho_G$ is a function of ν, which means it is also a function D. Numerically, at given B, D, $\delta\rho_G$, we can calculate the density of states at energy E of the system

$$\rho(E;D,B,\delta\rho_G)=\frac{\gamma}{\pi}\sum_j\frac{1}{(E-E_j)^2+\gamma^2}, \tag{21}$$

where the summation is over all eigen energies E_j of the system and γ is a phenomenological LL broadening. We set it as 1 meV in agreement with the experimental broadening of 1$\sim$2 meV.

The electron density (also linear in D) n measured from charge neutrality up to the chemical potential μ can be obtained by integrating the density of states, which gives

$$n(\mu;D,B,\delta\rho_G)=\sum_j\left[\arctan\left(\frac{\mu-E_j}{\gamma}-\arctan\left(\frac{E_{\rm CN}-E_j}{\gamma}\right)\right)\right], \tag{22}$$

with $E_{\rm CN}$ is the eigen energy corresponding to charge neutrality.

One can use the data computed in the previous step to obtain the chemical potential $\mu(n,D,B,\rho_G)$ as a function of n, D, B, and $\delta\rho_G$. Furthermore, since n is linear in D, and $\delta\rho_G(D)$ is also a function of D, we have the chemical potential $\mu(D,B)$ is only a function of D and B.

Then, the density of states as a function of $V_{\text{bias}} = E - \mu(D, B)$ can be obtained from the calculated density of states, by replacing $E \to \mu(D, B) + V_{\text{bias}}$, namely $\rho(\mu(D, B) + V_{\text{bias}}; D, B, \delta\rho_G(D))$. Here, the only parameter that need to be fixed is the coefficient c_n in $n = c_n D$. Finally, one can replace the D axis with the axis of filling factor ν. The last two steps can be done manually in order to match the experimental plot.

10. DETERMINING THE ONSET ANGLE OF CORRELATED GAPS

In Fig. 3 of the main text the onset angle are determined as the maximal angles where the local minima in the LDOS at the Fermi energy develops near integer-filling factors. We note that there are as we go from higher to lower angles the LDOS at Fermi energy drops both from the development of the symmetry breaking cascade, which can be seen in our data as an overall suppression of the LDOS as well as through the development of the pronounced dips pinned to integer filling factors. We interpret these dips as originating from the development of correlated gaps (see linecuts in Supplementary Fig. S7b-e and Supplementary Fig. S8c-f where the corresponding onsets can be clearly resolved).

We can also define the onset angle in the spectroscopy at $\nu = \pm 2$ as the maximum angle below which correlated gaps develop Supplementary Fig. S9. Spectroscopic measurements allow for potentially distinguishing these gaps that are expected to be pinned to Fermi energy from other LDOS suppressions (near E_F) that are of different origin. To illustrate further this point, we focus on the spectrum at $\theta = 1.21^\circ$ at $\nu = 2$ (Supplementary Fig. S9b, third trace from the top). This trace indeed shows gap-like feature in the gate spectroscopy, Supplementary Fig. S9c,d show that this "gap"-like minima are not pinned to Fermi energy. Instead, it follows the evolution of the VHS peaks and crosses from positive to negative bias voltages as filling factor is increased. For this reason, we attribute these gaps-like features to the non-interacting band structure (similar features are present for even larger angles Supplementary Fig. S10). In addition to this type of trivial peaks, we also observe weak (soft) LDOS suppression in the whole filling factor range, which we attribute to the Coulomb gap[S28]. Note that we purposefully focused here on the data at $\theta = 1.21^\circ$ (or $\theta = 1.26^\circ$ in Supplementary Fig. S10 where presumably physics related to the magic angle (cascade of symmetry breaking transitions, formation of correlated insulating states) does not occur. This distinction (pinning to E_F) allows us to identify onset of correlated gaps more precisely.

Focusing now on the regime around the magic angle, here we can still observe peaks in the spectrum crossing the Fermi energy (one such peak is shown in Supplementary Fig. S9e). These peaks likely occur due to emerging symmetry breaking cascade of electronic transitions near the magic angle TBG and may originate from the flavor resolved VHS. Details of these transition are beyond the scope of this work; however, as Supplementary Fig. S9e shows that even at $\theta = 1.12^\circ$ an LDOS peak crosses the Fermi energy near $\nu = +2$ which in turn can produce unexpected dip or peak-like structure in a single line traces (shown in Supplementary Fig. S9b at $\theta = 1.14^\circ; 1.12^\circ$). For $\nu = +2$, only below $\theta = 1.09^\circ$, we observed a pronounced gapped feature pinned to E_F regardless of presence (or absence) of other fine features and below this angle, the gap is always present down to the lowest accessible angle ($\theta = 0.98^\circ$).

While this analysis gives onset angle $\theta = 1.09^\circ$ for $\nu = +2$, for $\nu = -2$ similar onset occurs around $\theta \approx 1.12^\circ - 1.14^\circ$. This illustrates presence of of electron-hole asymmetry (Supplementary Fig. S9e-j) in the system. For example, the gap at $\nu = -2$ is clearly visible

at $\theta = 1.12^\circ$ (and for any other angle below it) where $\nu = +2$ gap is yet to emerge. This correlated gap on the hole side ($\nu = -2$) becomes the most pronounced at around $\theta = 1.1^\circ$ after which becomes shallower. In contrast, the gap at $\nu = 2$ starts emerging around $\theta = 1.09^\circ$ and becomes deeper and more developed down to $\theta = 0.98^\circ$. We note that, the observed electron-hole asymmetry in observed here is inline with previous observations in TBG-WSe_2 stuctures[S2,S3].

11. CASCADE IN FLAT BANDS

To study the appearance of a cascade in TBG as a function of twist angle and filling, we modify the mean-field Hamiltonian prescription of the previous section to allow for unequal filling of different flavours. Specifically, we focus on two candidate ground states at integer fillings ν: (i) all flavours are filled equally ("unpolarized" state) and (ii) electrons (for $\nu > 0$) or holes (for $\nu < 0$) added to the CNP completely fill $|\nu|$ flavors, leaving the other $4 - |\nu|$ flavours at charge neutrality ("polarized" state). Of course other candidate fillings are possible as we comment on below.

To determine whether a cascade occurs we compare energies for the unpolarized and the polarized solutions. For the unpolarized solution we evaluate the energy

$$E_{unpol.}(\nu) = \sum_{\boldsymbol{k},s}{}' E_s(\boldsymbol{k}, \nu)\,, \tag{23}$$

where $E_s(\boldsymbol{k}, \nu)$ is the energy of a state with momentum $\boldsymbol{k}$ from flavor s obtained by the diagonalization of the mean-field Eq. (9) with filling ν. The summation $\sum'$ ranges over occupied states from the CNP to ν. Note that $E_s(\boldsymbol{k}, \nu)$ implicitly includes the constant energy shift E_{int} that avoids over-counting in the mean-field energy. The energy for the polarized solution is similarly given by

$$E_{pol.}(\nu) = \sum_{\boldsymbol{k},s}{}''' E_s(\boldsymbol{k}, \nu)\,, \tag{24}$$

but here the summation $\sum'''$ ranges only over the relevant number of flavors (ν). We stress that for the polarized solution at integer fillings, $|\nu|$ flavors are completely filled in either the conduction or valence band. Therefore the corresponding Hartree and Fock terms for the same ν in the polarized and unpolarized case are distinctly different as they involve different momentum eigenstates that are included in the summation. For example, at $\nu = 1$ the majority of the states in the unpolarized solution originate from the vicinity of the κ points of the mini-BZ for four flavours, whilst in the polarized solution they come from all momentum eigenstates in the mini-BZ for the one occupied flavor in the conduction band. At this level of analysis there is no difference which specific flavour is the completely filled one, unlike those of more sophisticated models[S14,S17].

Since the role of Hartree is to flatten the bands and reduce the potential energy by smoothing the charge distribution in the unit cell, we expect cascading to be promoted by these interaction corrections. To quantify this scenario we consider the following two schemes—which we refer to as "0^{th} order" cascade—for estimating the onset of cascade without the benefit of band flattening. Scheme (I), shown in Fig. S16i, estimates the total

energy with an expectation of Eq. (9) taken over occupied states as computed from the non-interacting model. In other words it is the total energy at the first step of the self-consistent procedure, where mean-field potentials are computed using the non-interacting model but have not yet been used to modify the electronic spectrum. In scheme (II), shown in Fig. S16j, motivated by the model of Ref. S29, we consider the following approximation for the unpolarized and polarized energies,

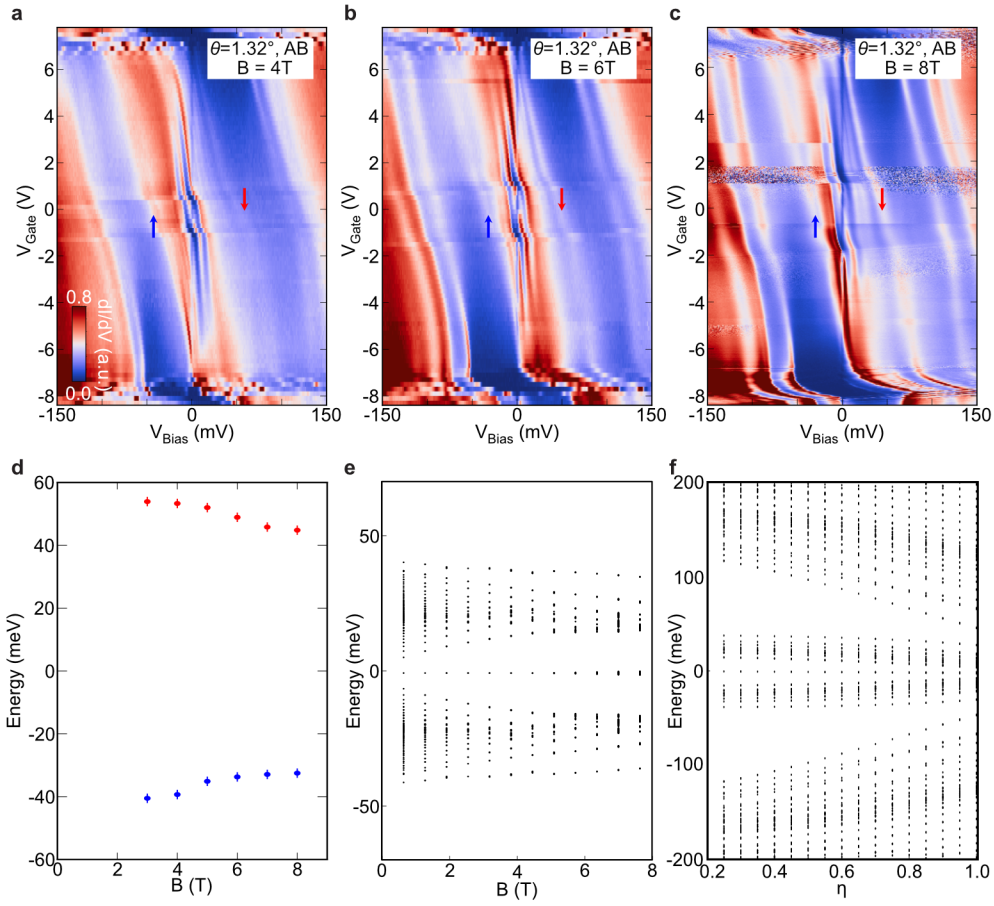
$$E^0_{unpol.} = \sum_{\boldsymbol{k},s}{}' E_s(\boldsymbol{k})^0 + \frac{1}{2}\sum_{s\neq s'}\int_\Omega d^2\boldsymbol{r}\int_\Omega d^2\boldsymbol{r}' V_C(\boldsymbol{r}-\boldsymbol{r}')\langle\delta\rho(\boldsymbol{r})\rangle_{occ.}\langle\delta\rho(\boldsymbol{r}')\rangle_{occ.} \tag{25}$$

$$E^0_{pol.} = \sum_{\boldsymbol{k},s}{}''' E_s(\boldsymbol{k})^0 + \frac{1}{2}\sum_{s\neq s'}\int_\Omega d^2\boldsymbol{r}\int_\Omega d^2\boldsymbol{r}' V_C(\boldsymbol{r}-\boldsymbol{r}')\langle\delta\rho(\boldsymbol{r})\rangle_{occ.}\langle\delta\rho(\boldsymbol{r}')\rangle_{occ.}\,, \tag{26}$$

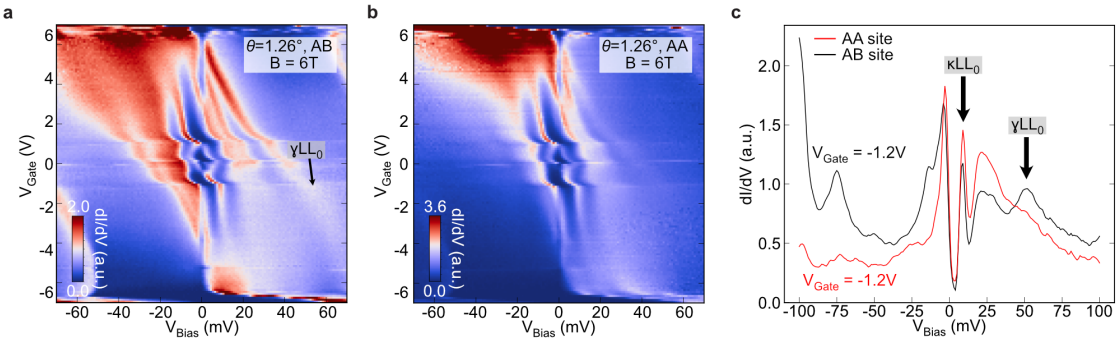
where $V_C(\boldsymbol{r})$ is the Coulomb potential as before, $E_s(\boldsymbol{k})^0$ is the non-interacting energy of a state with momentum $\boldsymbol{k}$ and flavor s obtained by diagonalizing the H_0 Hamiltonian, and $\delta\rho(\boldsymbol{r})$ is the charge density deviation from the CNP of the non-interacting TBG charge density. The summations $\sum'$ and $\sum'''$ are taken according to the same prescription outlined above. Most crucially, as argued in Ref. S29, we exclude the same-flavor interaction in the potential term, and the $\langle\ldots\rangle_{occ.}$ average is taken by obtaining the charge density deviation $\delta\rho(\boldsymbol{r})$ at a corresponding filling. Most crucially we stress that our second scheme (II) is not exactly the model of Ref. S29 where a local Coulomb potential is used unlike the long-range interaction we use. Irrespective of which scheme is used we note that depending on whether the unpolarized or polarized solution is considered, different states and flavors are included in the ground state average and thus both the kinetic energy and the potential energies in both schemes in the polarized solution will be different from the unpolarized solution.

Having explained how a 0^{th} order cascade can be estimated and defined the Hartree/Hartree-Fock self-consistent cascade, we proceed to analyze their differences in greater detail. Supplementary Fig. S16 presents band structures and energies obtained within these schemes at different fillings. As hinted in the above paragraph, band flattening coming from the Hartree term decreases the kinetic and potential energy. Kinetic energy decreases because the band structure is flattened, whereas the potential energy decreases because the charge distribution is smoothed out within the unit cell. This reduction allows for a cascade to occur more easily; cf. the $\nu = 3$ curves from the 0^{th} orders and Hartree-only cascade in panels i,j,k of Supplementary Fig. S16. However, as the twist angle comes closer to the magic angle, the Hartree-only correction predicts large band inversion at the γ point (panels c,d). This inversion increases the bandwidth and either prevents a cascade altogether in a Hartree-only description for $\nu = 2, 1$ or makes the system polarize and then unpolarize as for $\nu = 3$ in the panel k. As discussed earlier the Fock term cures this unphysical feature of Hartree; cf. the band structure in panels g,h. In the comparison of polarized/unpolarized energies in panel l—which includes Hartree and Fock—we indeed see that the system undergoes a cascade at $\nu = 1, 2, 3$ (contrary to Hartree only), and due to the band structure flattening the cascade occurs further away from the magic angle, as compared to the 0^{th} order cascade (where it either does not occur all together in i or occurs only near the magic angle j). As argued in the main text this is the key role played by band-flattening: it redefines the twist-angle regime where correlated behavior is observed, as also pointed out recently in Ref. S21. We also stress that we consider here only two potential ground states, either completely unpolarized or completely polarized. In reality more complex ground states could arise, including states with intermediate polarization (e.g., at $\nu = 3$ two flavors

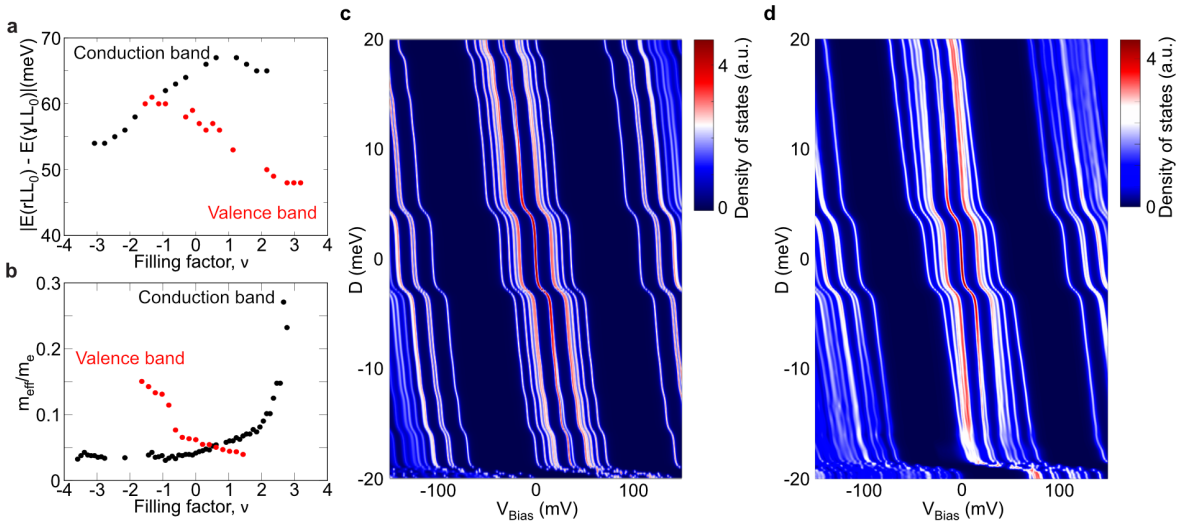
in the conduction band could be occupied while the other two are half-filled). Nevertheless, the analysis of Fig. 3h and Supplementary Fig. S16k predicts a hierarchy of twist angles at which cascade happens: upon approaching the magic angle from above, first the system cascades only at $\nu = 3$, then also at $\nu = 2$, and finally also at $\nu = 1$. This hierarchy matches the experimental observation from Fig. 3a on the electron side, but not on the hole side. Furthermore, experimentally the insulating state at $\nu = \pm 2$ is typically more robust and more frequently seen in experiments—which could possibly be explained by partially polarized states descending from the $\nu = 2$ polarized state becoming more favorable under certain conditions than the fully polarized $\nu = 3$ state we considered. These experimental observations require further theory investigation that is beyond the scope of the simple analysis presented here.



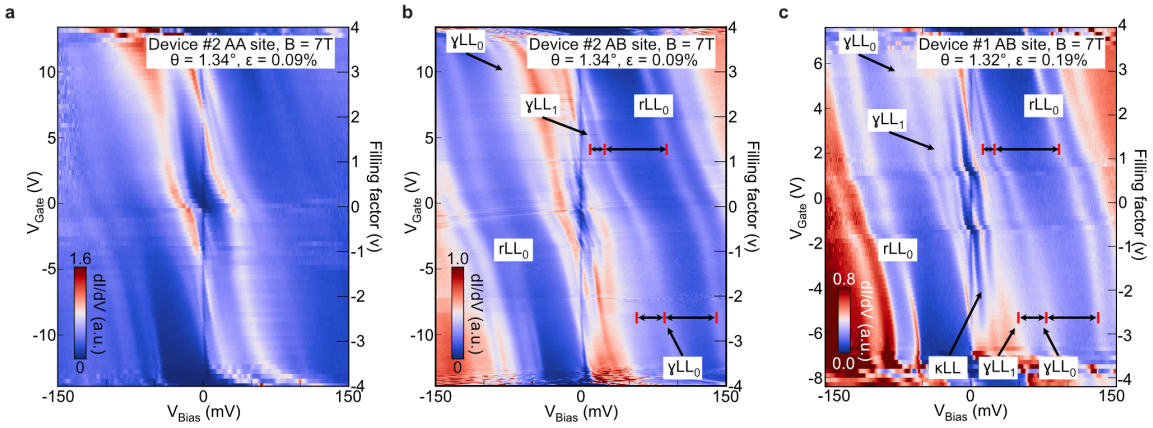
Supplementary Figure S1. **Magnetic field spectrum of flat band LLs at $\theta = 1.32°$. a-c**, Point spectroscopy on AB site as a function of V_{Gate} (T = 2 K) at twist angle 1.32° with different perpendicular magnetic fields B = 4 T (**a**), B = 6 T (**b**), and B = 8 T (**c**). Arrows mark the positions of the zeroth γLLs. As the field increases, γLLs move closer towards the center of the band indicating that the dispersion near the γ pocket is hole-like (electron-like) for the conduction (valence) band. **d**, Peak position of the zeroth γLL of conduction (red) and valence (blue) flat band as a function of magnetic field. **e**, Continuum model calculation of the TBG Landau level spectrum at a twist angle 1.32° showing qualitatively similar behaviour of the zeroth γLLs as seen in the experiment ($\eta = u/u' = 0.4$). **f**, Continuum model calculation of the TBG Landau level spectrum as a function of parameter η (see sections 8 and 9) at a fixed magnetic field B = 4 T. Range of $\eta = 0.4 - 0.6$ gives good quantitative agreement with experimental results.



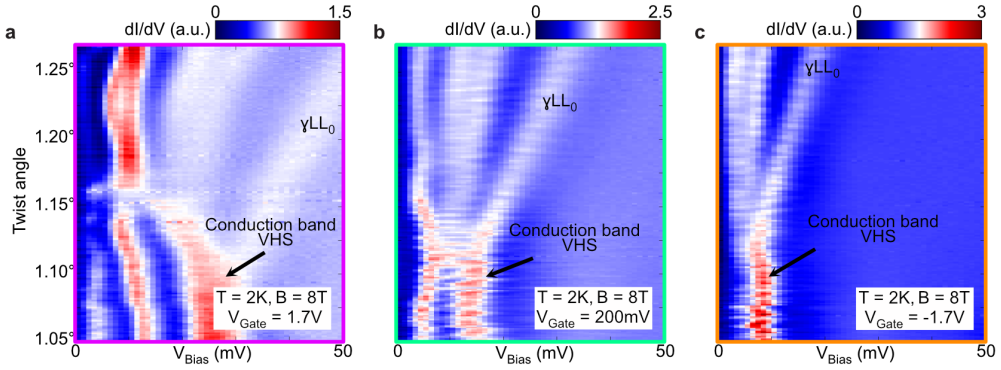
Supplementary Figure S2. **Spatial dependence of relative LDOS between γLL and κLL. a, b**, Point spectroscopy as a function of V_{Gate} at twist angle 1.26° with B = 6 T (T = 2 K) on AB site (**a**), and AA site (**b**). γLL_0 is clearly seen on the AB site while almost absent on the AA site. **c**, dI/dV taken at AA site (red) and AB site (black) at $V_{Gate} = -1.2$ V. While κLL_0 is prominent on both sites, γLL_0 is relatively weak on the AA site.



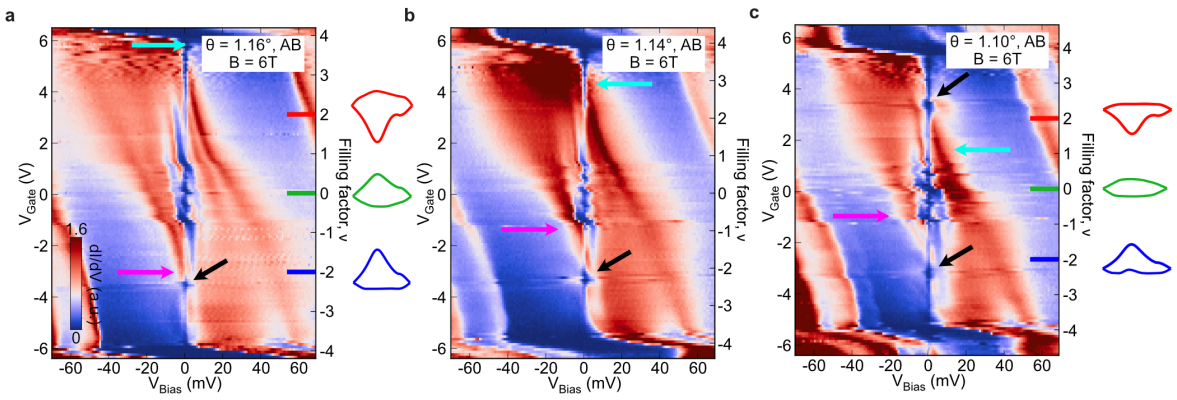
Supplementary Figure S3. **Filling-dependence of LL separation to remote bands, effective mass, and effect of displacement field at $\theta = 1.32^\circ$ and B = 7 T. a,** Energy difference between zeroth γLL and zeroth remote band LL as a function of filling factor for the conduction and valence flat bands obtained from Fig. 1c. The difference increases with increasing the filling factor for conduction flat band in contrast to the decreasing trend in the valence flat band. **b**, Effective mass in the γ pocket of the conduction (black) and valence (red) flat bands from the energy separation of zeroth and first γLL in Fig. 1c, highlighting the flattening of the band. **c**, Theory calculation of density of states as a function of doping and V_{Bias} at 1.32° and B = 7 T, with perpendicular displacement field effect included but without Hartree correction. We assume the perpendicular displacement field is linearly proportional to doping. In this case distance between γLLs does not change with doping (or displacement field). Also, large displacement field is expected to change separation between γLLs in conduction and valence bands symmetrically, which is opposite to the asymmetric trends shown in Fig. 1g and Supplementary Fig. S3a. **d**, Theory calculation as (**c**), when both perpendicular displacement field and Hartree correction are considered. See SI, section 9 for definitions of V_{Bias} and D in the model calculations.



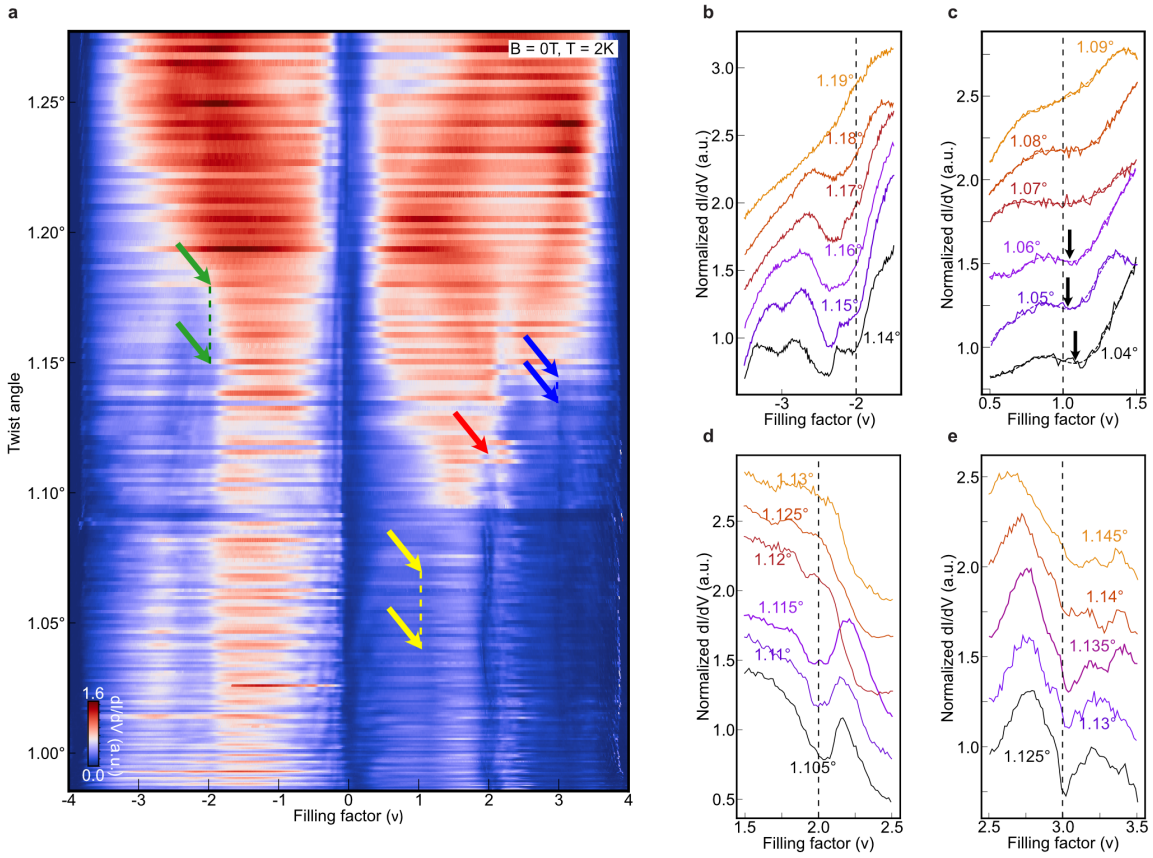
Supplementary Figure S4. **Comparison of gate spectroscopy taken at two different devices without and with WSe_2 substrate a,b**, Gate spectroscopy taken at the second device where TBG is put directly on an hBN substrate (without WSe_2) on AA site(**a**) and AB site(**b**). As in the first device that has been discussed in the main text the LLs originating from the γ point are located near the AB sites, accounting for the differences in spectral weight between (**a**) and (**b**). (**c**), For comparison we show gate spectroscopy from Fig. 1c taken in the first device. The band flattening physics is almost identical in the two devices.



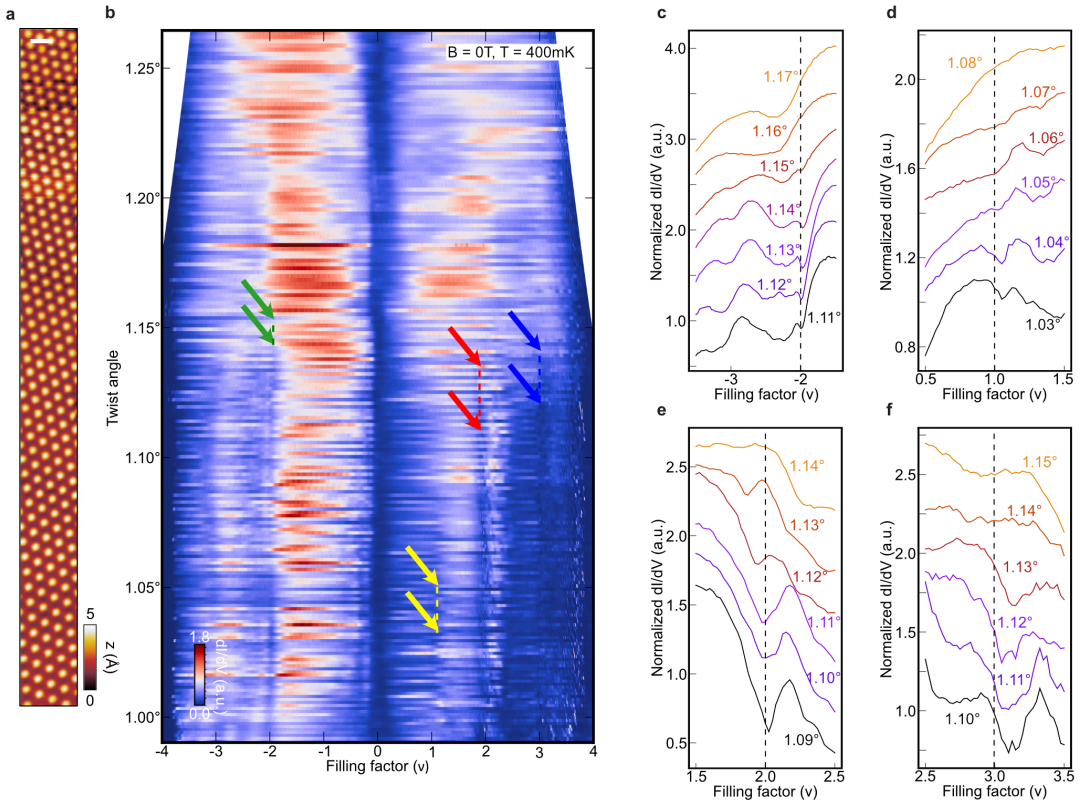
Supplementary Figure S5. **Mapping the evolution of conduction band LLs with twist angle at B = 8 T. a-c**, Spectroscopic dI/dV map as in Fig. 2c-e, but focusing on the conduction-band γLL_0 at different dopings ($V_{Gate} = 1.7$ V (**a**), electron doping; $V_{Gate} = 0.2$ V (**b**), near the CNP; $V_{Gate} = -1.7$ V (**c**), hole doping). In the electron doped regime, γLL_0 merges with the conduction band VHS at significantly lower twist angle compared to the hole doped regime, showing the effects of interaction-driven band flattening.



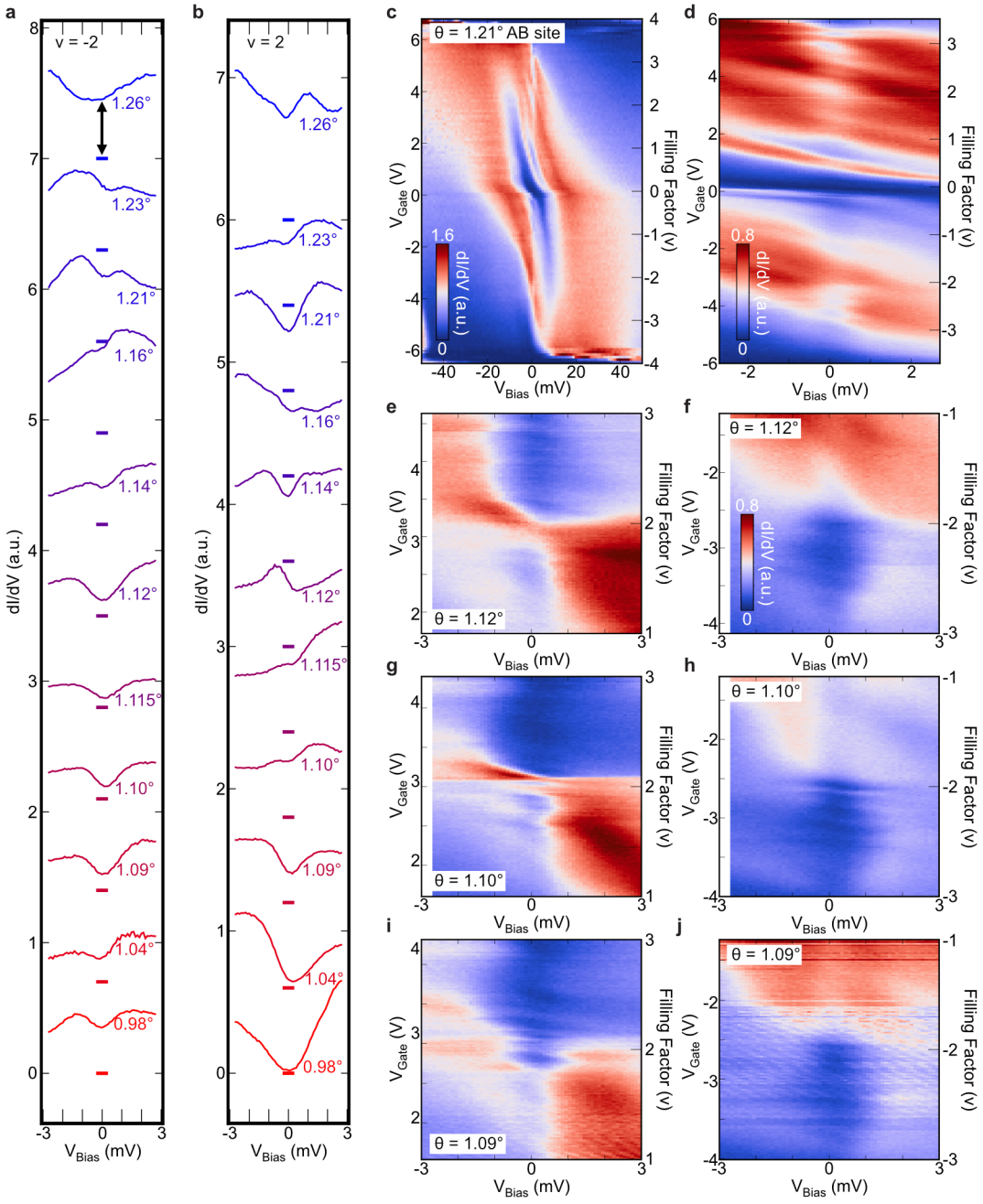
Supplementary Figure S6. **Examples of correlated Chern insulating phases at different magnetic field (B = 6 T) and different twist angles. a-c**, Point spectroscopy as a function of V_{Gate} at $\theta = 1.16°$ (**a**), $\theta = 1.14°$ (**b**), and $\theta = 1.10°$ (**c**) (T = 2 K). Horizontal arrows mark where the conduction-band (cyan) and the valence-band (magenta) γLL merge with the VHSs. Black arrows indicate the appearance of Chern insulating phases. We note that the correlated Chern phases appear after γLL merges with the VHS. Schematic band structure drawn on the right side of (**a**, **c**) illustrates effects of band flattening.



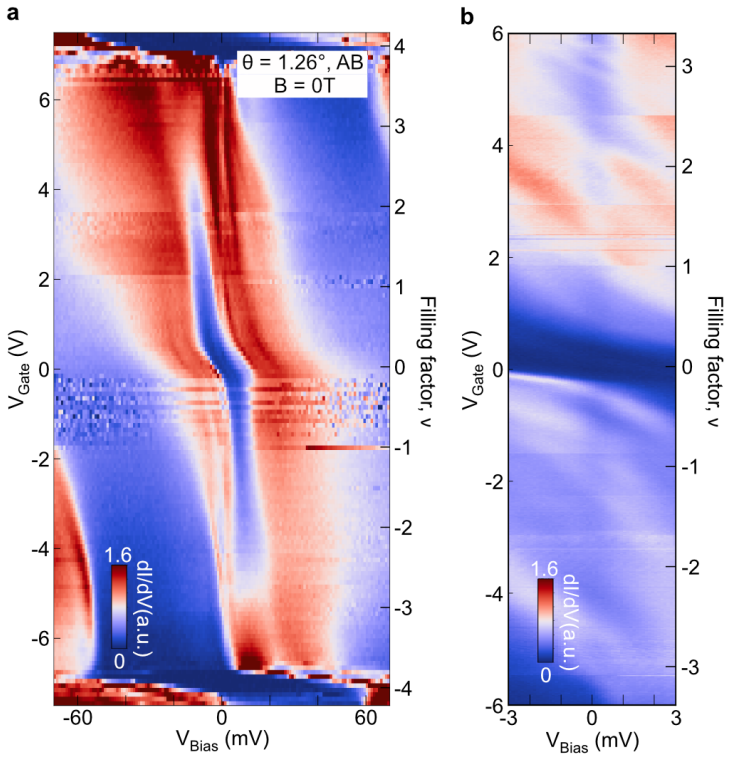
Supplementary Figure S7. **Angle dependence of LDOS suppression near the Fermi energy. a,** Same as Fig. 3a for reference. Colored arrows connected with dashed vertical lines indicate twist angle range identified as the onset of integer filling correlated states: 1.15°- 1.18° ($\nu = -2$), 1.04°- 1.07° ($\nu = 1$), 1.115° ($\nu = 2$), 1.135°- 1.145° ($\nu = 3$). **b,c,d,e,** Linecuts from (**a**), around $\nu = -2$(**b**), 1(**c**), 2(**d**), 3(**e**) showing dips that develop below the onset angle. Each spectrum is normalized by dividing by mean value and offset for clarity. Vertical black dashed lines is the position of $\nu = -2, 1, 2, 3$. Vertical black arrows denote the position of the dip at $\nu = 1$ by taking local minimum.



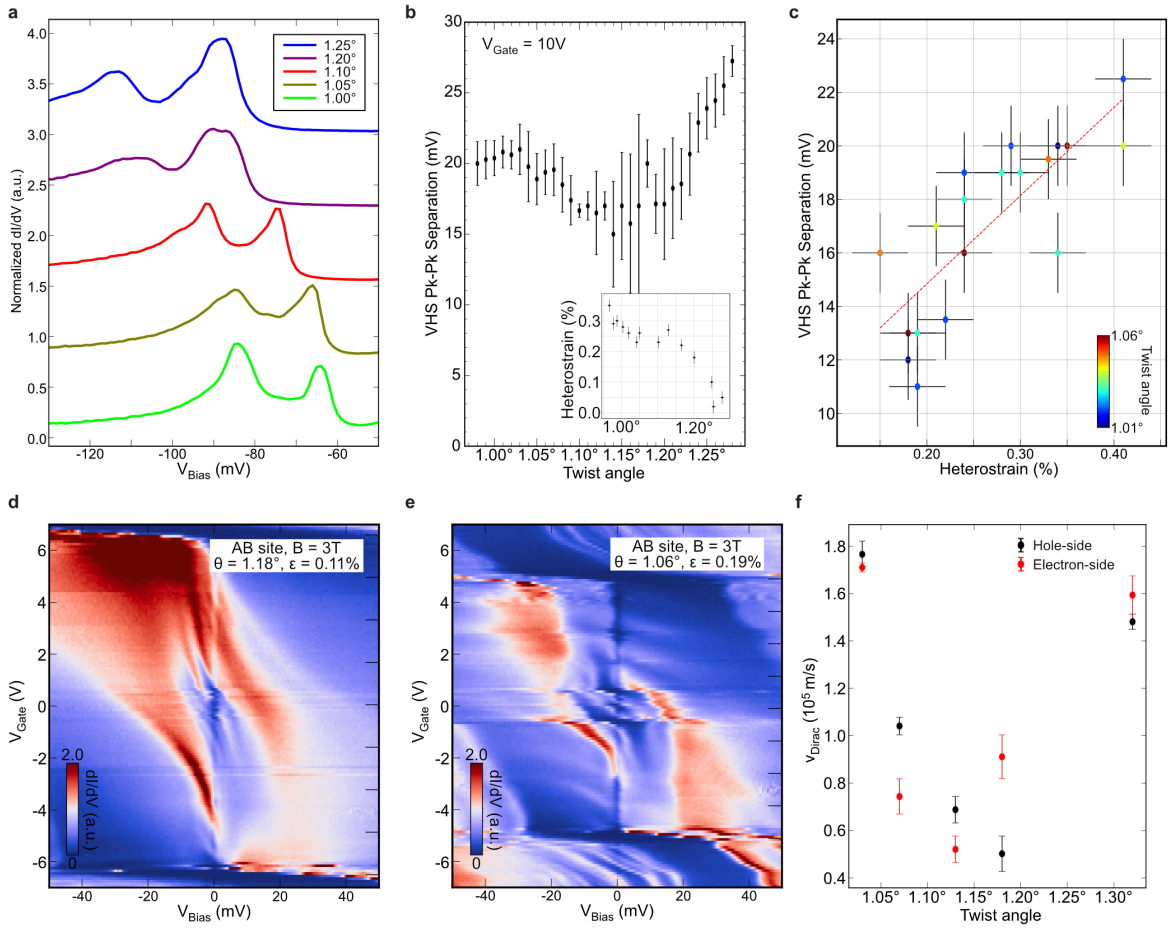
Supplementary Figure S8. **Twist angle vs Filling factor map measured at different area.** **a**, 647nm × 52nm area where twist angle is slowly varying from 1.26° to 0.99° (scale bar corresponds to 20nm, set point parameters: $V_{Bias} = 100$mV, I = 20pA) in a different area. **b**, Angle and filling factor dependence of dI/dV near E_F ($V_{Bias} = 0.3$mV) at B = 0T and T = 400mK measured at the area of (**a**). **c-f**, dI/dV as a function of filling factor and twist angle by taking average of linecuts of (**a**) over small twist angle window of ±0.05° to mitigate the fluctuation in the spectra coming from moiré sites, around $\nu = -2$(**c**), 1(**d**), 2(**e**), 3(**f**). Each spectrum is normalized by dividing into mean value for clarity. The vertical black dashed lines indicate the position of $\nu = -2, 1, 2, 3$. This data set is taken several months apart from the data in Fig. 3 (Supplementary Fig. S7).



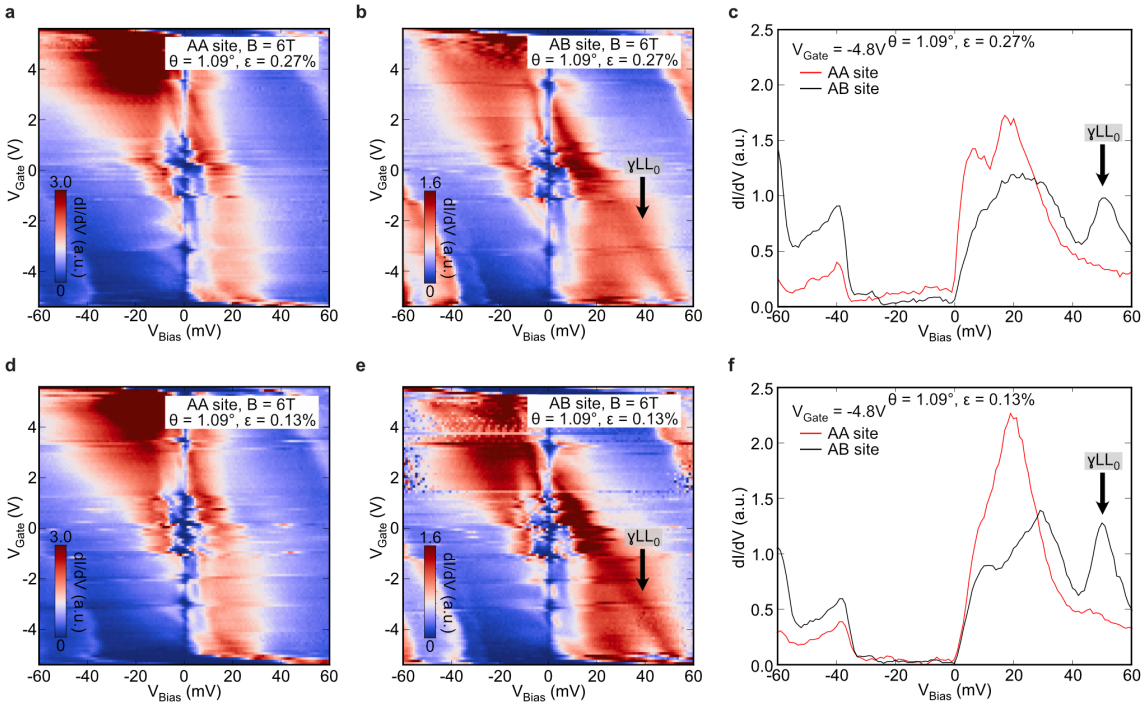
Supplementary Figure S9. **Spectroscopy around $\nu = \pm 2$ taken in the same area as Fig. 2a and in the local twist angle range $\theta = 1.26° - 0.98°$. a, b**, Linecuts at $\nu = -2$ (**a**), 2(**b**). Each linecut is averaged over filling factor ± 0.05 centered around $\nu = \pm 2$. Horizontal line of the same color mark corresponding zero dI/dV value. Offset added for clarity. (**c, d**), V_{Gate}-dependent point spectroscopy for $\theta = 1.21°$. (**e-j**), High resolution point spectroscopy in a small V_{Bias} (V_{Gate}) range around $\nu = 2$(**e, g, i**) and $\nu = -2$(**f, h, j**) for $\theta = 1.12°$ (**e, f**; $\theta = 1.10°$ **g,h**; and $\theta = 1.09°$ **i,j**. Data measured on AB sites at 400 mK.



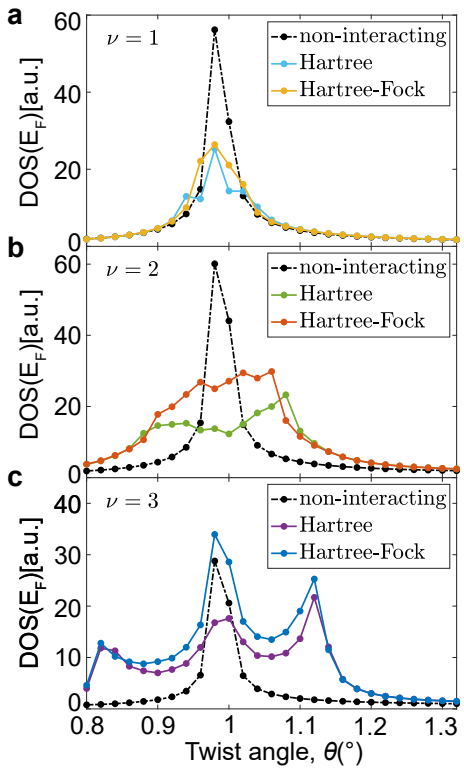
Supplementary Figure S10. **Trivial LDOS suppressions away from the magic angle.** **a**, Point spectroscopy as a function of V_{Gate} for $\theta = 1.26^\circ$ (T = 0.4 K). Signatures of correlations (symmetry-broken cascade of phase transitions or VHS separation enhancement around the CNP) are not observed. **b**, High resolution point spectroscopy in a smaller V_{Bias} range for the same point as (**a**). The spectrum here illustrates that care must be exercised in order to distinguish LDOS suppression that can occur at the Fermi energy due to strong correlations versus incidental band structure features. Here, multiple local minima in the LDOS develop around the Fermi energy; however, these move parallel to each other and with the VHS upon doping. Also, these LDOS features do not disappear even when they are away from the Fermi energy, suggesting that their likely origin is related to band structure details that are beyond the scope of this work. Another trivial reduction of LDOS close to the Fermi energy can arise from Coulomb gaps[S28,S30] that create shallow dips at all filling factors. Gaps such as those shown in Fig. 3c,d at $\nu = \pm 2$, by contrast are significantly deeper and are pinned to the Fermi energy (Supplementary Fig. S9).



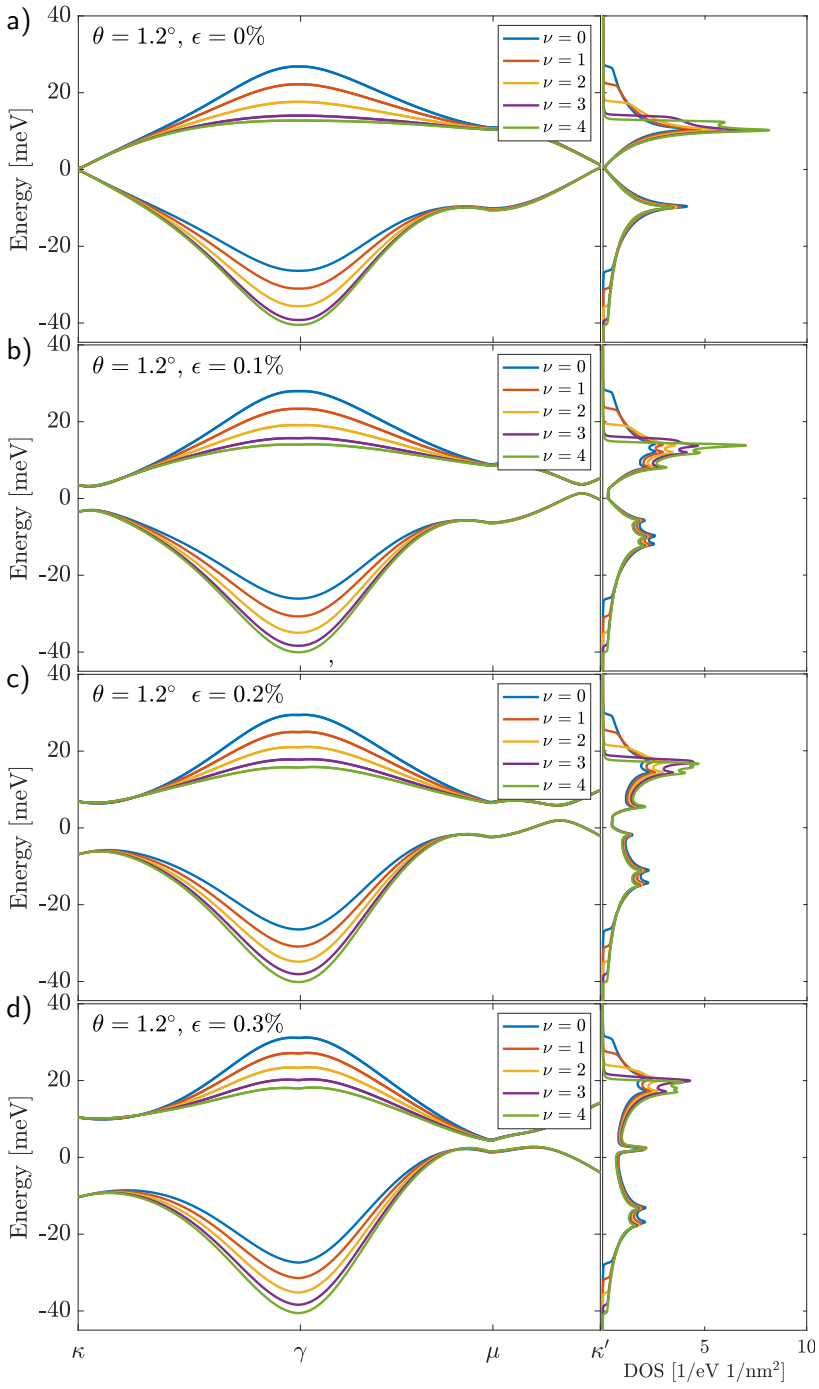
Supplementary Figure S11. **VHS separation measured at $\nu > +4$ where interaction effects are minimal and Dirac velocity as a function of twist angle.** **a**, dI/dV spectroscopy showing two VHSs from TBG flat bands at $V_{Gate} = 10$ V for twist angles from 1.25° to 1.00° (T = 2 K). At this value of V_{Gate}, the flat bands are fully occupied ($\nu > +4$). In this regime the Fermi level is in the conduction remote band and it is expected that the interaction effects, which tend to increase the VHS separation when the Fermi level sits in between the two VHSs[S1], are minimized. **b**, Extracted conduction band and valence band VHS separation as a function of twist angle ($\nu > +4$). **c**,VHS Peak-to-Peak Separation measured at full fillings of the flat bands taken at various areas of the sample with different level of heterostrain but similar values of twist angles near the magic angle. The precise local twist angle is plotted as a color of each data point. Red dashed line is a linear regression of the dataset. The slope and error is $3.1meV \pm 0.5meV$ per 0.1% heterostrain. **d,e**, Point spectroscopy examples as a function of V_{Gate} at B = 3 T for extracting Dirac velocity. **f** Extracted Dirac velocity by measuring the energy separation between the zeroth κLL and the next closest level inside the band, assuming linear dispersion. Note that the Dirac velocity at 1.18°, where no strong correlation effects are seen, is much smaller than the Dirac velocity at 1.04°, where cascade of phase transitions is observed and correlated Chern insulators are clearly resolved.



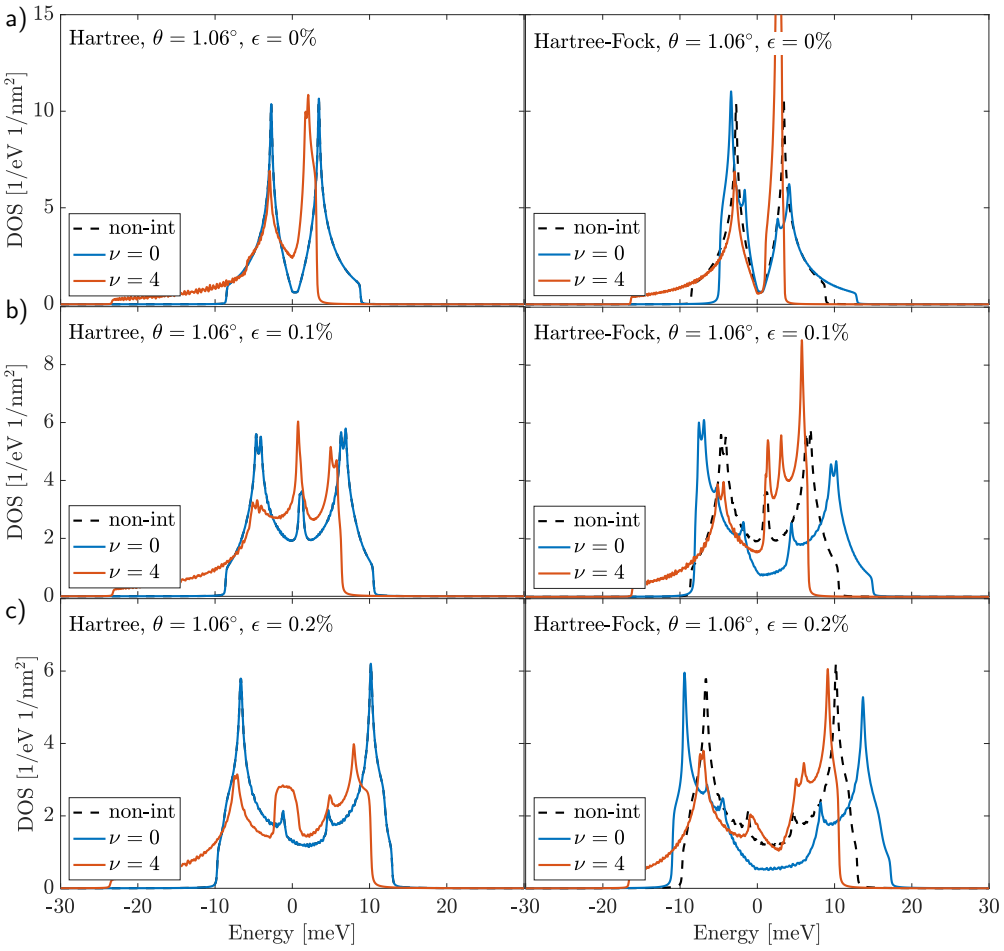
Supplementary Figure S12. **Finite perpendicular magnetic field spectroscopy taken at two points with different heterostrain. a,b**, Gate spectroscopy on an AA and AB site in area with 0.27% heterostrain and $\theta = 1.09°$. **c**, Linecut taken at $V_{Gate} = -4.8$ V (hole-doped) showing that the γLL is more pronounced on the AB site. **d,e**, Gate spectroscopy on an AA and AB site of 0.13% heterostrain area at the same twist angle. **f**, Linecut taken at $V_{Gate} = -4.8$ V. Despite different levels of strain, the flattening of the bands is similar



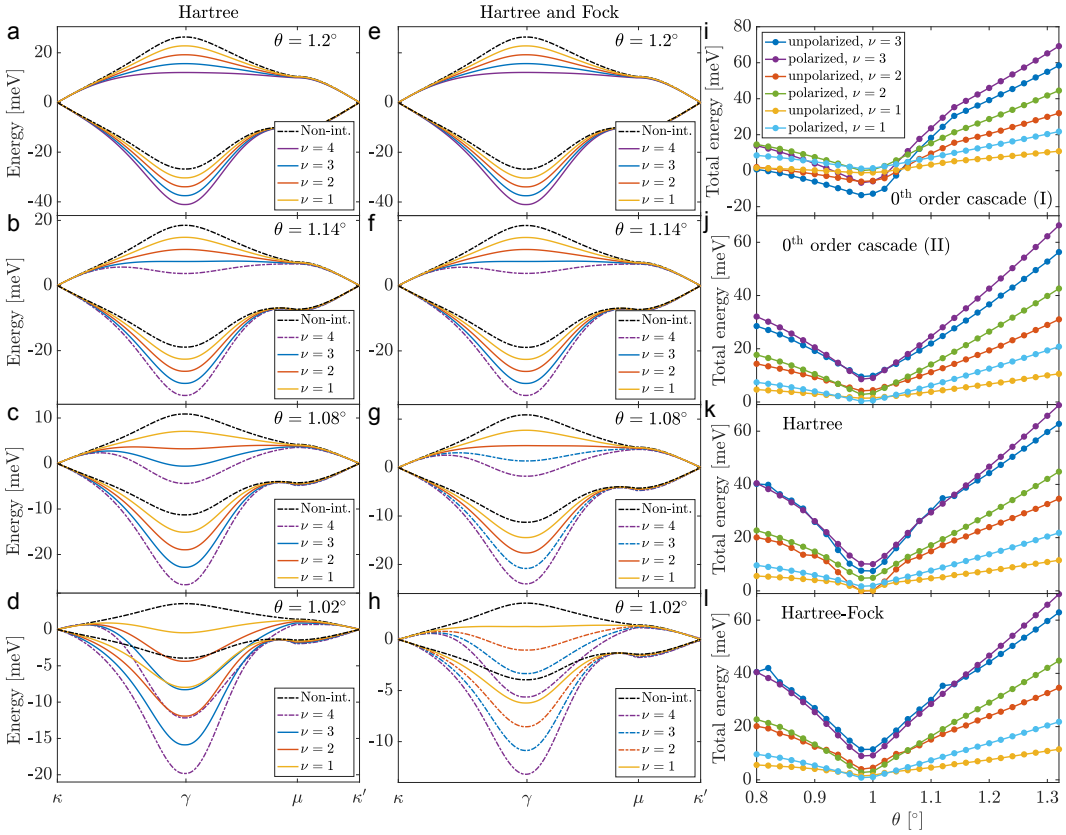
Supplementary Figure S13. **Fermi level density of states for different fillings and twist angles. a-c**, Density of states for the TBG band structure calculated using three different models: non-interacting model; Hartree-only modified band structure; and Hartree-Fock modified band structure. Note how the density of states increases as twist angle comes closer to the vicinity of the magic-angle $\theta_M \approx 1^\circ$ (expected from the bandwidth reduction). Note also how the onset of band-flattening is indicated by peaks in the band structures with interactions, particularly in (**c**). Angle range where the DOS exceeds a certain threshold is much broader with interactions.



Supplementary Figure S14. **Interaction-driven band-flattening in the presence of strain** We implemented uniaxial heterostrain in the direction $\phi = 30°$ following the Ref. S7. The high symmetry points in panels (a-d) correspond to the unstrained symmetry points. We see that even with modest strain (0.3%) the filling-dependent band flattening remains robust. Side panels show the corresponding density of states illustrating increasing van Hove-to van Hove separation with strain as expected together with the Hartree-driven band flattening. Results are effectively for a Hartree-only calculation as at this twist angle we expect small exchange corrections.



Supplementary Figure S15. **Side-side by comparison of Hartree and Hartree-Fock modeling in the presence of strain**; Density of states obtained from the continuum model calculations that incorporate Hartree and Hartree-Fock corrections. We also include a uniaxial heterostrain[S7] (indicated with ϵ) and focus on two fillings: charge neutrality point and full filling for simplicity. In the unstrained case (**a**) we find that our treatment of Fock (unlike Hartree-only treatment) correctly predicts bandwidth enhancement (specifically increase in van Hove to van Hove separation as argued in section 7) and splitting (broadening) of van Hove peaks at the charge neutrality point. When strain is added to the system (**b**, **c**) bandstructure overall broadens and multiple peaks in DOS develop. In (**a**, **b**, **c**) we solve self-energy corrections self-consistently at each filling and do not employ the interpolation procedure of Eq.(17) to capture Fock-induced features at $\nu = 0$.



Supplementary Figure S16. **Modeling of TBG band structure and the cascade. a-h**, Band structure with Hartree (**a-d**) and Hartree-Fock (**e-h**) corrections for a valley $\xi = 1$: black dashed lines indicate the non-interacting model; colored lines correspond to the band structure with interactions corrections with the dashed lines indicating the band structure if cascade were not to occur. Note how Hartree-only solution gives rise to significantly larger γ point inversions (**c-d**) compared to the case when the Fock term is included (**g-h**) (see also Sec. 7 for other mechanisms that could counteract Hartree-driven band inversions). For clarity (also in the main text) when plotting we centered all band structures near the κ, κ' points. **i-l**, Total energy of the system in an unpolarized and polarized ground state for different cascade schemes: 0^{th} order cascade, scheme (I) (**i**) and scheme (II) (**j**) relying only on the non-interacting band structure (see section 11); cascade incorporating Hartree (**k**) and Hartree-Fock (**l**) corrections to the band structure. When energy of the polarized solution falls below that of the unpolarized solution a cascade is energetically favorable. For the 0^{th} order cascade we see no cascade for type (I) (**i**) and cascade in a narrow range of angles for type (II) (**j**). Note how the Hartree-only cascade, due to the large kinetic energy stemming from γ point inversions (**c-d**), is energetically unfavorable compared to the Hartree-Fock solution and thus does not predict a sequence of transitions seen in (**j**).

[S1] Choi, Y. *et al.* Electronic correlations in twisted bilayer graphene near the magic angle. *Nature Physics* **15**, 1174–1180 (2019).

[S2] Choi, Y. *et al.* Correlation-driven topological phases in magic-angle twisted bilayer graphene. *Nature* **589**, 536–541 (2021).

[S3] Arora, H. S. *et al.* Superconductivity in metallic twisted bilayer graphene stabilized by WSe 2. *Nature* **583**, 379–384 (2020).

[S4] Walkup, D. *et al.* Tuning single-electron charging and interactions between compressible Landau level islands in graphene. *Physical Review B* **101**, 035428 (2020).

[S5] Nuckolls, K. P. *et al.* Strongly correlated Chern insulators in magic-angle twisted bilayer graphene. *Nature* **588**, 610–615 (2020).

[S6] Wong, D. *et al.* Cascade of electronic transitions in magic-angle twisted bilayer graphene. *Nature* **582**, 198–202 (2020).

[S7] Bi, Z., Yuan, N. F. Q. & Fu, L. Designing flat bands by strain. *Physical Review B* **100**, 035448 (2019).

[S8] Guinea, F. & Walet, N. R. Electrostatic effects, band distortions, and superconductivity in twisted graphene bilayers. *Proceedings of the National Academy of Sciences* **115**, 13174–13179 (2018).

[S9] Rademaker, L., Abanin, D. A. & Mellado, P. Charge smoothening and band flattening due to Hartree corrections in twisted bilayer graphene. *Physical Review B* **100**, 205114 (2019).

[S10] Goodwin, Z. A. H., Vitale, V., Liang, X., Mostofi, A. A. & Lischner, J. Hartree theory calculations of quasiparticle properties in twisted bilayer graphene. *Electronic Structure* **2**, 034001 (2020).

[S11] Bistritzer, R. & MacDonald, A. H. Moiré bands in twisted double-layer graphene. *Proceedings of the National Academy of Sciences* **108**, 12233–12237 (2011).

[S12] Koshino, M. *et al.* Maximally Localized Wannier Orbitals and the Extended Hubbard Model for Twisted Bilayer Graphene. *Physical Review X* **8**, 031087 (2018).

[S13] Cea, T., Walet, N. R. & Guinea, F. Electronic band structure and pinning of Fermi energy to Van Hove singularities in twisted bilayer graphene: A self-consistent approach. *Physical Review B* **100**, 205113 (2019).

[S14] Cea, T. & Guinea, F. Band structure and insulating states driven by Coulomb interaction in twisted bilayer graphene. *Physical Review B* **102**, 045107 (2020).

[S15] Guinea, F. & Walet, N. R. Continuum models for twisted bilayer graphene: Effect of lattice deformation and hopping parameters. *Physical Review B* **99**, 205134 (2019).

[S16] Xie, M. & MacDonald, A. H. Weak-field Hall Resistivity and Spin/Valley Flavor Symmetry Breaking in MAtBG. *arXiv:2010.07928 [cond-mat]* (2020). 2010.07928.

[S17] Bultinck, N. *et al.* Ground State and Hidden Symmetry of Magic-Angle Graphene at Even Integer Filling. *Physical Review X* **10**, 031034 (2020). 1911.02045.

[S18] Kerelsky, A. *et al.* Maximized electron interactions at the magic angle in twisted bilayer graphene. *Nature* **572**, 95–100 (2019).

[S19] Xie, Y. *et al.* Spectroscopic signatures of many-body correlations in magic-angle twisted bilayer graphene. *Nature* **572**, 101–105 (2019).

[S20] Parker, D. E., Soejima, T., Hauschild, J., Zaletel, M. P. & Bultinck, N. Strain-Induced Quantum Phase Transitions in Magic-Angle Graphene. *Physical Review Letters* **127**, 027601

(2021).

[S21] Klebl, L., Goodwin, Z. A. H., Mostofi, A. A., Kennes, D. M. & Lischner, J. Importance of long-ranged electron-electron interactions for the magnetic phase diagram of twisted bilayer graphene. *Physical Review B* **103**, 195127 (2021).

[S22] Khalaf, E., Chatterjee, S., Bultinck, N., Zaletel, M. P. & Vishwanath, A. Charged skyrmions and topological origin of superconductivity in magic-angle graphene. *Science Advances* **7**, eabf5299 (2021). 2004.00638.

[S23] Xie, M. & MacDonald, A. H. Nature of the Correlated Insulator States in Twisted Bilayer Graphene. *Physical Review Letters* **124**, 097601 (2020).

[S24] Carr, S., Fang, S., Zhu, Z. & Kaxiras, E. Exact continuum model for low-energy electronic states of twisted bilayer graphene. *Physical Review Research* **1**, 013001 (2019).

[S25] Fang, S., Carr, S., Zhu, Z., Massatt, D. & Kaxiras, E. Angle-Dependent {\it Ab initio} Low-Energy Hamiltonians for a Relaxed Twisted Bilayer Graphene Heterostructure. *1908.00058[cond.mat]* (2019).

[S26] Bistritzer, R. & MacDonald, A. H. Moir\'e butterflies in twisted bilayer graphene. *Physical Review B* **84**, 035440 (2011).

[S27] Hejazi, K., Liu, C. & Balents, L. Landau levels in twisted bilayer graphene and semiclassical orbits. *Physical Review B* **100**, 035115 (2019).

[S28] Efros, A. L. Coulomb gap in disordered systems. *Journal of Physics C: Solid State Physics* **9**, 2021–2030 (1976).

[S29] Zondiner, U. *et al.* Cascade of phase transitions and Dirac revivals in magic-angle graphene. *Nature* **582**, 203–208 (2020).

[S30] Jung, S. *et al.* Evolution of microscopic localization in graphene in a magnetic field from scattering resonances to quantum dots. *Nature Physics* **7**, 245–251 (2011).