The discovery of superconducting and insulating states in magic-angle twisted bilayer graphene (MATBG)\(^1\)\(^-\)\(^2\)\(^-\)\(^11\)\(^-\)\(^12\) has ignited considerable interest in understanding the nature of electronic interactions in this chemically pristine material. The transport properties of MATBG as a function of doping are similar to those of high-transition-temperature copper oxides and other unconventional superconductors\(^1\)\(^-\)\(^3\), which suggests that MATBG may be a highly interacting system. However, to our knowledge, there is no direct experimental evidence of strong many-body correlations in MATBG. Here we present high-resolution spectroscopic measurements, obtained using a scanning tunnelling microscope, that provide such evidence as a function of carrier density. MATBG displays unusual spectroscopic characteristics that can be attributed to electron–electron interactions over a wide range of doping levels, including those at which superconductivity emerges in this system. We show that our measurements cannot be explained with a mean-field approach for modelling electron–electron interactions in MATBG.

The breakdown of a mean-field approach when applied to other correlated superconductors, such as copper oxides, has long inspired the study of the highly correlated Hubbard model\(^7\). We show that a phenomenological extended-Hubbard-model cluster calculation, which is motivated by the nearly localized nature of the relevant electronic states of MATBG, produces spectroscopic features that are similar to those that we observed experimentally. Our findings demonstrate the critical role of many-body correlations in understanding the properties of MATBG.

The stacking of two graphene layers on top of each other results in a moiré superlattice with a periodicity that depends on the angle between the layers. Close to the ‘magic angle’ of approximately \(1^{\circ}\), a non-interacting continuum model of the band structure of this system predicts nearly flat low-energy valence and conduction bands\(^4\)\(^-\)\(^5\). Previously, studies of MATBG using a scanning transmission microscope (STM) have visualized the moiré superlattice, identified different regions of sublattice stacking (AA and AB/BA) of the graphene sheets, and have resolved two peaks in the tunnelling spectrum (\(dI/dV\)) associated with the large density of states (DOS) of the flat bands\(^6\)\(^-\)\(^10\). Transport studies show that the partial occupation of the two flat bands of MATBG results in a cascade of insulating phases and superconducting domes as a function of electron filling\(^12\)\(^-\)\(^11\)\(^-\)\(^12\). Considering the fourfold degeneracy (spin and valley) of the two flat bands, novel insulating phases occur at partial moiré band fillings of \(\nu = 0, \pm 1/4, \pm 1/2, \pm 3/4\), where \(\nu\) is the moiré band filling factor, which suggests that electron–electron interactions have a predominant role in the formation of the insulating phases. Interactions also result in the development of magnetism in MATBG\(^12\)\(^-\)\(^13\), although the breaking of the crystalline symmetry by the underlying hexagonal boron nitride substrate may be a requirement for such magnetic phases. The observation that resistivity increases linearly with temperature (\(T\)-linear resistivity) at high temperatures\(^14\) may also be indicative of interactions; however, the origin of this behaviour is still debated\(^15\). Beyond these observations and despite many theoretical studies\(^16\)\(^-\)\(^20\), there are many unresolved questions regarding the role of interactions in this system. Among these is the question of whether interactions in MATBG are not only strong when the system is insulating, but also when superconductivity emerges at other doping levels\(^1,2,11,12\). To construct the correct model of superconducting pairing, we require an accurate picture of how interactions influence the low-energy excitations of the system.

Here we perform high-resolution spectroscopy studies of MATBG using the STM to address these questions. We show that when the nearly flat valence and conduction bands are either filled or unoccupied, a non-interacting model—which includes the influence of strain and relaxation—captures the spectroscopic properties of MATBG. However, at partial band fillings, we demonstrate that the quasiparticle spectrum is strongly modified from that of a non-interacting model, over a wide range of energies that far exceed that of the bandwidth of the flat bands or their separation to remote bands. Without any theoretical modelling, these experimental observations illustrate that MATBG is a highly interacting problem, the physics of which cannot be captured with weak-coupling theoretical approaches. The strong correlations uncovered by our experiments are key to the properties of MATBG when superconductivity emerges in this system.

We examine the properties of exposed MATBG as a function of electron density in back-gated devices (Fig. 1a, b; see Methods for details of fabrication\(^21\)\(^-\)\(^22\)) using a home-built ultrahigh-vacuum STM operating at 1.4 K. Consistent with previous studies\(^6\)\(^-\)\(^10\), the STM topographies of our devices reveal a moiré superlattice in which the bright (dark) regions correspond to the AA (AB/BA) stacking region, where high (low) local density of states (LDOS) is expected (Fig. 1c). From the observed periodicity of the moiré lattice, we confirm that the twist angle in this region of the sample is close to the magic-angle value of approximately \(1^{\circ}\). A more detailed examination of topographies also reveals the presence of strain and lattice relaxation, the information on which can be extracted from STM topographies (see Methods) and can be used to theoretically model the spectroscopic properties of MATBG when electron–electron interactions are not important. Figure 1d shows the \(dI/dV\) spectrum measured at the AA regions shown in Fig. 1c at a gate voltage \((V_g)\) of \(-4 \text{ V}\). The spectrum features two sharp peaks below the Fermi level and two weaker step-like features at other energies (arrows in Fig. 1d). As expected from the continuum model, and consistent with previous measurements\(^6\)\(^-\)\(^10\), the two sharp peaks are associated with the van Hove singularities of the occupied, nearly flat conduction and valence bands of MATBG. However, the original, non-interacting continuum model\(^4\)\(^-\)\(^5\) would predict these bands to have far shaper peaks than those that are observed in our experiments. Including the influence of strain\(^23\) and relaxation\(^24\)\(^-\)\(^25\) in the continuum model results in additional dispersion of the valence and conduction bands (Fig. 1e, f), which not only better captures the width and separation of the double peaks in the tunnelling spectra but also the presence of the step-like features at higher and lower energies (arrows in Fig. 1d, f). In our calculations, these step-like features are associated with the van Hove singularities of the bands.

Yonglong Xie\(^1\), Biao Lian\(^2\), Berthold Jäck\(^1\), Xiaomeng Liu\(^1\), Cheng-Li Chiu\(^1\), Kenji Watanabe\(^3\), Takashi Taniguchi\(^3\), B. Andrei Bernevig\(^1\) & Ali Yazdani\(^1\)*

---

1. Joseph Henry Laboratories and Department of Physics, Princeton University, Princeton, NJ, USA. 2. Princeton Center for Theoretical Science, Princeton University, Princeton, NJ, USA. 3. National Institute for Material Science, Tsukuba, Japan. **e-mail: yazdani@princeton.edu**
remote to the flat bands. We have repeated similar local modelling of the dl/dV spectra measured at AA sites using information extracted from topographies at other locations of our devices (see Extended Data Fig. 1). The description of the local spectra is satisfactory when disorder is weak, and when the double peaks associated with the flat bands are either below or above the chemical potential.

The breakdown of this single-particle description of the spectroscopic properties of MATBG when interactions are important becomes evident when we study the evolution of the quasiparticle spectra in our device as a function of electron density controlled by \( V_g \). Figure 2 shows dl/dV measurements on the AA region as a function of \( V_g \), which spans three different regions of occupation for the two flat bands: when the flat bands are both occupied (\( V_g > -5.5 \) V), when they are being depleted (\( -53.5 < V_g < -5.5 \) V), and after they have been depleted (\( V_g < -53.5 \) V). The rate of the shift of the flat-band peaks with \( V_g \) reflects the DOS at the Fermi level. Therefore, a change in slope of the lines in Fig. 2a signals a transition in band filling. The nearly vertical features signal the slow change of occupation of the flat bands with large DOS. In the range \(-58 < V_g < -53.5\) V, the change in slope might be related to the presence of an energy gap between the flat bands and the remote bands (estimated to be around 15 meV, which is roughly consistent with the calculated band structure in Fig. 1e). When the nearly flat bands are filled or fully depleted, the spectra—which are individually plotted in Fig. 2b, e—show relatively sharp double peaks at all gate voltages, and the widths of these peaks change weakly with their energy separation to the Fermi level (see Extended Data Fig. 2 and Methods). As described above, these spectra are consistent with those calculated from a non-interacting model that includes the effects of strain and relaxation. However, the most notable change in the quasiparticle spectra occurs when one of the flat bands begins to overlap with the Fermi level, as demonstrated by contrasting the data in Fig. 2c, d with those in Fig. 2b, e. In this region, as one of the flat bands is being depleted, not only does the peak associated with that band near the Fermi level develop substantial features and broaden, but the peak associated with the other valence (conduction) band below (above) the Fermi level is also substantially modified. Notably, the strong distortion of the shape of the quasiparticle spectra—which is caused by interactions during the partial filling of the flat bands—spans an energy range (30–50 meV) that is wider than both the separation of the flat bands to the remote bands and the apparent bandwidths of the flat bands, as measured when fully occupied or unoccupied. This observation demonstrates that the largest energy scale for determining the properties of MATBG at partial filling of the flat bands is set by electron–electron interactions. This signature of strong correlations occurs not just at commensurate fillings, but over all doping ranges for which transport studies have found superconductivity in this system below 1 K.

To further relate our spectroscopic measurements to the transport properties of MATBG, in Fig. 3 we plot the tunnelling conductance at zero energy dl/dV(0)—which is a measure of the DOS at the Fermi level—as a function of \( V_g \). From the changes in dl/dV(0) and the measured energies of the van Hove singularity peaks in Fig. 2, we identify the \( V_g \) values that correspond to the point at which there is full local depletion of the conduction or valence flat bands (\( n = \pm n_0 \), where \( n_0 \) denote the carrier density of the system and that of a moiré band, respectively); at this point, the transport measurements reveal evidence for a band insulator. Further assignment of the occupation level within the flat bands with \( V_g \) is made complicated in our experiments by the presence of the STM tip (see Methods and Extended Data Fig. 3). Focusing on the gate region of \( V_g \) at which we are depleting the conduction flat band, a region where the tip-induced effects are minimal, we find some features in the tunnelling spectra that correlate with the transport studies. Most notably we find that, at half-filling of the conduction band (\( n = n_0/2 \))—where transport measurements reveal the strongest insulating behaviour—dl/dV(0) vanishes and a gap-like feature appears in the spectrum. Recent STM studies of MATBG have reported a similar gap feature, but the suppression of the DOS at the Fermi level was less than 50%. However, we caution that this gap at half-filling is much larger (about 20 times) than that observed in transport measurements, and may be related to a soft gap observed at other doping levels (Fig. 2c). Interactions, together with the localization of electrons either by disorder or large magnetic fields, are well known to induce soft Coulomb gaps in tunnelling spectroscopy. Close to the
At the charge neutrality point (CNP)—at which the two flat-band peaks are roughly symmetric about the Fermi level (green curve in Fig. 3b)—we observe a strong increase (around 20 meV) in the separation between these flat-band peaks. This enhanced separation is another indication of the importance of interactions in MATBG. Notably, the peaks associated with the fully occupied or unoccupied flat bands regain some of their sharpness near the CNP. The suppression of the DOS at the Fermi level near the CNP is also consistent with a recent report of insulating behaviour at zero doping. The apparent gap at $n = n_0/2$, and other finer features in the spectra at other rational fillings, show variability between devices; it may therefore be necessary to examine cleaner samples before these features can be fully established as intrinsic effects. Nevertheless, our key fundamental finding—the dramatic deviation from single-particle spectra at partial fillings—is reproducible (see Extended Data Fig. 4) and is unique to the magic-angle device (see Extended Data Fig. 5 for results from the non-magic-angle device). We can therefore use this information to discriminate between different models of interactions in MATBG.

Fig. 2 | Breakdown of non-interacting description. a, $dI/dV$ measured on an AA site as a function of sample bias and gate voltage ($V_{set} = 200$ mV, $I_{set} = 120$ pA, $V_{mod} = 1$ mV). b–e, Normalized $dI/dV$ spectra at different gate voltages extracted from a. The curves are shifted vertically by 0.5 each for clarity. c and d demonstrate the breakdown of the non-interacting description. The black dotted line indicates the Fermi level.

Fig. 3 | Evidence for correlated insulating state at half-filling. a, Conductance at the Fermi level as a function of gate voltage. Grey areas correspond to fully (un)occupied flat bands with sharp quasiparticles (QPs). The blue rectangles define the full filling of the two flat bands. The CNP is located inside the green region. The light yellow rectangle is the midpoint between the edges of the CNP and full filling. The black dotted line marks zero conductance. The asymmetry in gate responses between the conduction and valence flat bands is due to the tip band bending (see Methods). Individual spectra at different gate voltages from different shaded areas in a. The spectra are vertically shifted by 1.8 nS each for clarity. The black dotted line indicates the Fermi level.
Our first approach towards understanding the effects of electronic correlations on the spectroscopic properties of MATBG is to perform self-consistent Hartree–Fock calculations by adding the Coulomb interaction to the continuum model (see Methods). These calculations allow the possibility of interaction-induced spontaneous valley and/or spin polarization, without making any assumptions about the symmetry breaking in the ground state. The resulting spectra from these calculations (see Extended Data Fig. 6) show the pinning of the flat bands during filling (vertical feature in Fig. 2a) and the generic broadening of the bands, as well as the enhanced separation between the van Hove singularities near the CNP due to the exchange interactions. However, these calculated spectra fail to reproduce the abrupt distortion of the experimental quasiparticle spectra during the partial filling of the flat bands. During the partial filling of the conduction flat band, our calculations show no discernible distortion of the valence flat band. More generally, if a mean-field order parameter is added to the non-interacting band width $6t$, while maintaining $V_0 = U/6$, $V_1 = U/7.9$.

Next, we consider whether the features that we observe could be a consequence of the symmetry breaking in the ground state of MATBG. It is instructive to contrast our findings with recent STM studies of valley-polarized quantum Hall states, in which the occupation of the flat Landau level bands can be adjusted $29,30$. These experiments show that interactions produce symmetry-broken valley-polarized ground states, which induce spectral splitting at the Fermi level. However, in stark contrast to our observation, the line shape of the filled or unoccupied Landau levels in those experiments remain unaffected even when symmetry is broken in the system. We therefore conclude that in MATBG, although interactions can break the valley or spin symmetry at fractional fillings of the flat bands $12,20$, such changes should only alter the states of the partially filled band in which such symmetries are broken, and would not generically alter the fully occupied or unoccupied flat bands. We attribute this behaviour, which has been previously observed in quantum Hall ferromagnets $29,30$, to the fact that cyclotron energy is much larger than that of the interactions. By contrast, in MATBG, the lack of capacity of the weakly interacting model (mean-field) to explain the features of the data forces us to consider a physical picture in which the effects of interactions are dominant.

The salient features of our data can be captured within a phenomenological model in which the effects of Coulomb interactions of the nearly localized states of the moiré flat bands can be studied without relying on a mean-field approximation. To motivate our phenomenological model, we note that the maximally localized Wannier orbitals that correspond to the flat bands of the MATBG are shaped as three lobes symmetrically distributed around the AB/BA moiré sites, with their wavefunctions strongly localized at the three nearby AA sites $16-18$. Theoretical studies have emphasized that the shape of these Wannier functions suggests that it is necessary to include the nearest-neighbour Coulomb interactions, and extended Hubbard models have been constructed on the basis of this idea. Given that the charge density is peaked on the AA sites (Fig. 1c), the simplest model to consider is that of a triangular lattice with two orbitals per site at energies $\pm \epsilon$ (no spin/valley flavour). We include a hopping between nearest sites ($t$) and both the on-site ($U$) and the nearest-neighbour ($V_0$, $V_1$) Coulomb repulsions (Fig. 4a; see Methods). To compare the results of such a model with those of our experiments, we carry out an exact diagonalization calculation for small clusters (see Methods and Extended Data Fig. 7). The resulting local spectral weight from this toy model as a function of electron filling (Fig. 4b) exhibits the same behaviour as that observed experimentally in Fig. 2b–e. When the two orbitals in the model—corresponding to the two flat bands—are fully occupied or empty, the spectra show sharp peaks (broadened by a bandwidth $6t \approx 5$ meV), whereas at partial fillings of either orbitals (flat bands) we see substantial broadening of the two peaks. The broadening (or splitting) of the flat-band peaks in our model at partial fillings is tuned by the strength of Coulomb interactions: a value of $U/6t \approx 6$ produces spectra with a broadening (around 20 meV) that is comparable to our experiments. We note that the discrete broadening (splitting) is due to our use of singly degenerate orbitals to make the computation tractable, and we anticipate that restoring spin and valley flavours would further smooth the broadened spectra with the added presence of more peaks. We also observe an interaction-enhanced energy separation between the peaks when the chemical potential is in between the energies of the two flat bands. This agrees well with the experimental results observed near the CNP. The parameters used to calculate the spectra in Fig. 4b, the on-site $U$ (about 30 meV) and the nearest-neighbour values of $V$ (about 5 meV) are consistent with recent estimates for the Coulomb interactions in magic-angle moiré superlattices with an underlying hexagonal boron nitride substrate $16$. Similar calculations on a honeycomb lattice (see Extended Data Fig. 8) result in the same conclusions, indicating that they are independent of the choice of lattice structure.

Our toy model succeeds in capturing the broadening of MATBG quasiparticle spectra because it incorporates various correlated charge configurations at partial fillings, which considerably affect the electron addition and removal energies of the localized orbitals. The mean-field approach replaces these configurations with a simple average and ignores the fluctuations among them, therefore resulting in far less complex spectral features compared with those observed experimentally. More realistic extended Hubbard models with larger cluster calculations $31$ can be used to more precisely compute the electron addition and removal spectra for MATBG. Most broadly, our observation that $U$ is the largest energy scale in the problem indicates the inaccuracy of models that project the physics of electronic interactions onto the two lowest flat bands. Our results indicate that MATBG is truly a many-body problem—possibly not confined to the lowest two bands. This poses a formidable theoretical challenge for finding the
true ground-state for realistic models, because the size of the Hilbert space—which includes valley and spin flavours—is far beyond the limit of computation with a classical computer. Regardless, our experiments and toy model calculations demonstrate that many-body correlations—at least as complex as those present in a Hubbard model—are required to describe the low-energy properties of this system. Our finding therefore establishes a more robust connection between MATBG and high-transition-temperature superconducting copper oxides, beyond the phenomenological resemblance of their transport phase diagrams. Such a connection renders MATBG a potential model system in which a mechanism for strongly correlated superconductivity could be more precisely established.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests are available at https://doi.org/10.1038/s41586-019-1422-x.

Received: 7 April 2019; Accepted: 18 June 2019;
Published online 31 July 2019.


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

© The Author(s), under exclusive licence to Springer Nature Limited 2019
METHODS

Sample preparation. The device was made using a modified ‘tear and stack’
technique.21,22. We used a polypropylene carbonate (PCP) film and polydimethyl-
siloxane (PDMS) stack on a glass slide to first pick up a 30–40 nm-thick hexagonal
boron nitride (hBN) flake. Then we used the van der Waals force between hBN
and monolayer graphene to tear and pick up half of the graphene flake. The remaining
graphene flake on the Si substrate was rotated by 1.3° and picked up. In order to
expose the graphene surface, the resulting stack with PCP was transferred onto a
second PDMS stamp. After dissolving the PCP film in aceton, the inverted stack
was placed in between pre-patterned Au contacts deposited on a SiO₂ wafer with
a Si back gate (Fig. 1a, b). Before inserting into the scanning transmission micro-
scope, the sample was annealed for 10 h in ultrahigh vacuum at 250°C. In these
devices, we adjusted the electron density in situ in the microscope by adjusting the
back-gate voltage. The 1.72° device (Extended Data Fig. 5) was fabricated by first
picking up a graphite flake followed by the preparation of an hBN/twisted bilayer
graphene stack. The resulting stack with the PCP film was dropped off directly on
a pre-patterned SiO₂ wafer. The PCP film was removed by radiative heating
in a high-vacuum chamber. Finally, a second graphite flake was placed to ensure
the connectivity between the twisted bilayer graphene and the pre-patterned Au
contact.

STM measurements. All measurements were performed on a home-built ultrahigh-
vacuum scanning transmission microscope operating at T = 1.4 K, which can store
two samples simultaneously. The dI/dV spectra were acquired using a standard
lock-in technique with a modulation voltage V_mod (root mean square, f = 4 kHz)
and a time constant of 5 ms while keeping the feedback opened using an appropriate
bias 1 V. We stabilized the tip–sample distance using a uniaxial strain model as described
previously.23 We interpret the value of λ that depends on the details of avail-
able scattering channels near the Fermi surface, as expected from the Fermi liquid
theory, until the VHS energy is about 20 meV below the Fermi level. For VHS ener-
gies lower than ~20 meV, the peak width appears to reach a plateau: 10 meV for
the conduction flat band and 13.5 meV for the valence flat band. We speculate that the
deviation from Fermi liquid theory possibly arises because of the greater distance
to the Fermi level and the detailed Fermi surface when the Fermi level is inside the
first conduction remote band(s). For fully emptied bands (Extended Data Fig. 2b),
the peak width first decreases until the VHS energy reaches 17 meV. We interpret
this behaviour as the suppression of electron–electron scattering channels due to
the presence of an energy gap separating the valence flat band and the first valence
remote band(s). The VHS energy from which the peak width starts to increase is
approximately 17 meV, which is consistent with the gap value found in Fig. 2a. For
VHS energies higher than ~17 meV, the peak width can be fit with an E² function
in which the value of λ is almost 30 times smaller, possibly due to a different Fermi
surface when the Fermi level is crossing the first valence remote band(s).

Gate voltage asymmetry between the two flat bands. In Fig. 3a, we find that
the scale of the gate voltage for the positive side of the band (depletion of conduction
flat band) is consistent with that estimated from capacitive coupling to a silicon back
gate for a twist angle of 1.01°. The scale of the gate voltage for the depletion of the
valence flat band—that is, negative side of the CNP—is, however, much smaller. This asymmetry is due to the tip-induced band bending caused by
the difference in the work function between the STM tip and the sample. This
band bending changes the effective doping of the region underneath the tip rela-
tive to regions further away. In this situation, the region in which measurements
of the DOS are performed (under the tip) is doped not only by the gate but also
by the regions surrounding it—both of which respond to varying gate voltage.
As a result, the density under the tip is not simply a linear function of V_g as it
would be in a transport experiment. To understand this effect of tip gating, we have
performed self-consistent finite element electrostatic simulation using the
band structure shown in Fig. 1e without considering interactions. The detailed
geometry of the simulation is shown in Extended Data Fig. 3a. The results of the
calculations in Extended Data Fig. 3b confirm that the density in the region under
the STM tip (blue curve) is not simply a linear function of V_g. This density changes
more quickly when the valence flat band overlaps with the Fermi level because of
the influence of regions in the sample away from the tip. As a result, the gate
voltages required to deplete the conductance and valence flat bands are different,
as measured by the gate dependence of dI/dV(0) in both the model (Extended Data
Fig. 3c) and measurements in Fig. 3a. The simulation in Extended Data Fig. 3c shows
a 2:1 ratio for the difference in depletion of the conductance and the valence
bands. This ratio for the data in Fig. 3a is 3.8:1; however, the exact results in the
model depend on microscopic details such as the exact tip shape, the difference
in work function and the tip–sample distance. Extended Data Fig. 3d, e shows the
measurement on the same AA site with the same microscopic STM tip (as Fig. 3a)
but with different set point conditions, which adjust the tip height. We find that the
gate range for the valence flat band is sensitive to the tip height—consistent with
our model—the tip–sample of conduction to valence flat bands is reduced from
that of Fig. 3a. The sensitivity of the dI/dV(0) to fractional filling of valence
and conduction bands (that is, gaps), and the gate voltages at which they occur,
also change with the height of the tip. Extended Data Fig. 3d, e also shows that the
regions of zero conductance near the CNP depends on the tip height. Using
different tip conditioning, we also find an example in which the gate voltage for
depleting the lowest conduction and valence bands is nearly symmetric (Extended
Data Fig. 5). As noted in the manuscript, further experiments are required to
associate features in dI/dV(0) with intrinsic behaviour of MATBG and to cor-
relate them with transport studies precisely. Last, in the finite element
electrostatic simulation, we find that the tip band bending effect has negligible
influence on the line shape of the flat bands.

Self-consistent Hartree–Fock calculations. We apply the self-consistent Hartree–
Fock method to the extended continuum model in the presence of Coulomb
interaction. The Coulomb interaction takes the form

\[ H_{\text{Coulomb}} = \frac{1}{2} \sum_{\alpha \beta, q, k, k'} \frac{V(q)}{A} \langle \psi_\alpha^{\dagger} \psi_\beta \rangle \langle \psi_\beta \psi_\alpha \rangle \delta_{kk'} \]

where V(q) = 2π²ε²/q² (for q = 0; for q = 0 one has V(0) = 0) is the Fourier trans-
form of Coulomb interaction ε²/er, A is the area of the sample. \( \psi_\alpha \) is the anhil-
iation operator of the Dirac electron of graphene at momentum k, and \( \alpha \) denotes
the sublattice, layer, spin and valley indices. The mean-field Hartree (\( \Sigma_H \)) and Fock
(\( \Sigma_F \)) terms are given by

\[ \Sigma_H = \sum_{\alpha \beta, q, k, k'} \frac{V(q)}{A} \langle \psi_\alpha^{\dagger} \psi_{\alpha-q} \psi_{\beta-k} \psi_{\beta-k+q} \rangle \]

\[ \Sigma_F = -\sum_{\alpha \beta, q, k, k'} \frac{V(q)}{A} \langle \psi_\alpha^{\dagger} \psi_{\alpha-q} \psi_{\beta-k} \psi_{\beta-k+q} \rangle + h.c. \]

where h.c. is the Hermitian conjugate.

In the extended continuum model, the momenta \( k + m g_1 \) and \( k + m g_2 \) of the
same spin and same valley are coupled via (multiple) momentum space hoppings,
where \( m_1 \) and \( m_2 \) are the reciprocal vectors of the moiré superlattice. We assume
that the translation symmetry is unbroken, so \( \psi_\alpha^{\dagger} \psi_{\beta-k} \) is non-zero only if

\[ k' = k + m_1 g_1 + m_2 g_2 \ (m_1, m_2 \in Z) \]

and \( \alpha, \beta \) belong to the same spin and valley. The mean-field Hartree–Fock Hamiltonian is then given by

\[ H = H_0 + \Sigma_H + \Sigma_F \]

where \( H_0 \) is the free Hamiltonian of the continuum model. For each fixed total
filling density, we diagonalize \( H \) and calculate \( \Sigma_H \) and \( \Sigma_F \) iteratively to find the
self-consistent solution. In the calculations, we allow the four different spin and
valley flavours to have different fillings, so that the system could spontaneously
develop a flavour polarization. We keep six momentum shells in the continuum
model (74 bands per spin per valley) for our self-consistent Hartree–Fock calcu-
lations. Our calculations (Extended Data Fig. 5b) show generic broadening of the
flat bands independent of the fillings, in contrast to the abrupt broadening seen in the experiments and possibly a gap near half-filling of the conduction flat band. In addition, we find a small spontaneous spin and/or valley polarization (when the Fermi level is in the flat bands) with maximum amplitude near half-filling of the flat bands.

**Exact diagonalization of an extended Hubbard model.** To capture the flat band broadening, we use exact diagonalization to study two single-flavour toy models of two nearly flat bands with on-site and neighboring-site interactions. The first model is defined on a triangular lattice (representing the lattice of AA stacking positions), with two orbitals per site corresponding to the two flat bands. The Hamiltonian is given by

\[ H = \sum_i [\epsilon(n_{i,1} - n_{i,2}) + U_0(n_{i,1}n_{i,2})] + \sum_{\langle ij \rangle} \eta^{ij} c_i^\dagger c_j + c_j^\dagger c_i + \text{h.c.} \]

\[ + \sum_{\langle ij \rangle} \epsilon_{ij} n_{i,1} n_{i,1} + n_{i,2} n_{i,2} + \epsilon_{ij} n_{i,1} n_{i,2} + n_{i,2} n_{i,1} \]

where \( \eta^{ij} \) represents neighbouring sites in a triangular lattice. We perform our exact diagonalization for both a six-site cluster and a seven-site cluster and extract out the spectral weight at various fillings. The chemical potential \( \mu = \frac{d E_g}{d N} \) is calculated from the ground-state energy \( E_g(N) \) at various total electron numbers \( N \). For the six-site exact diagonalization, we assume a 2\( \times \)2 triangular lattice (see Extended Data Fig. 6a) with periodic boundary conditions in both directions. For the seven-site exact diagonalization (see Fig. 4a), periodic boundary conditions are impossible; instead, we assume that \( \{ij\} \) runs over all pairs of sites, so that each site still has six neighbours (as is true for a triangular lattice), and all the sites are equivalent. Owing to the delocalized nature of the remote bands, the interaction between the remote band and the flat bands is insignificant. When both flat bands are empty or fully occupied, and the Fermi level is in the remote band, the spectral weight only shifts in energy with respect to the electron filling. The slope is determined by the density of states of the remote band. The second model is defined on a honeycomb lattice (representing the lattice of AB and BA stacking positions, see Extended Data Fig. 7a), where each site has one orbital, so that there are two bands in total. The Hamiltonian is given by

\[ H = \sum_i \epsilon_i n_i + \sum_{\langle ij \rangle} [\eta^{ij} c_i^\dagger c_j + \text{h.c.}] + \sum_{\langle ij \rangle} \epsilon_{ij} n_i n_j \]

where \( i \) runs over all the honeycomb lattice sites, \( \eta^{ij} = +1 \) and \(-1\) for sublattices AB and BA, respectively, and \( \epsilon_i \) is a staggered potential between the two sublattices. \( \langle ij \rangle \) and \( \langle \langle ij \rangle \rangle \) denote the nearest and next-nearest neighbours, respectively. \( t \) is the nearest-neighbour hopping, whereas \( V_0 \) and \( V_1 \) are the interactions between nearest neighbour and next-nearest neighbour, respectively. Because each site has only one orbital, there is no on-site interaction. Our exact diagonalization calculation for this model is performed for six unit cells, which form a 2\( \times \)3 honeycomb lattice with periodic boundary conditions. Independent of the choice of the lattice, our exact diagonalization calculations show a generic broadening (splitting) of the band broadening stemming from the interplay between strong correlations and flat bands, as seen in the experiment. Adding spin and valley degeneracy would create more possible charge configurations for a given filling, thus would further smoothen the broadened spectra with the added presence of more peaks, as can already be seen in the honeycomb calculations (see Extended Data Fig. 7b).

**Data availability**
The data that support the findings of this study are available from the corresponding author on reasonable request.

**Acknowledgements** We acknowledge discussions with N. P. Ong, S. Wu, D. Wong, P. Sambo, L. Glazman, P. Phillips and A. MacDonald. We also acknowledge collaborations with E. Tutuc, K. Kim and Y. Wang during the initial stages of this project. This work has been primarily supported by the Gordon and Betty Moore Foundation as part of the EPiQS initiative (GBMF4530) and DOE-BES grant DE-FG02-07ER464419. Other support for the experimental work was provided by NSF-MRSEC programmes through the Princeton Center for Complex Materials DMR-142054, NSF-DMR-1608848, ExxonMobil through the Andlinger Center for Energy and the Environment at Princeton, and the Princeton Catalysis Initiative. B.J. acknowledges funding from the Alexander von Humboldt foundation through a Feodor Lynen postdoctoral fellowship. K.W. and T.T. acknowledge support from the Elemental Strategy Initiative conducted by the MEXT, Japan, A3 Foresight by JSPS and the CREST (JPMJCR15F3). JST. B.L. acknowledges support from the Princeton Center for Theoretical Science at Princeton University. B.A.B. acknowledges support from the Department of Energy DE-SC0016239, Simons Investigator Award, the Packard Foundation, the Schmidt Fund for Innovative Research, NSF EAGER grant DMR-1643312, and NSF-MRSEC DMR-1420541.

**Author contributions** Y.X., B.J., C.-L.C., X.L. and A.Y. designed, performed and analysed the STM experiments. X.L., Y.X. and B.J. fabricated the sample. X.L. carried out the finite element electrostatic simulation. B.L. and B.A.B. performed the theoretical calculations. K.W. and T.T. provided hBN crystals. All authors contributed to the writing of the manuscript.

**Competing interests** The authors declare no competing interests.

**Additional information**

**Correspondence and requests for materials** should be addressed to A.Y.

**Peer review information** *Nature* thanks Roman Pisarev, Joaquín Fernández Rossier and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at http://www.nature.com/reprints.
Extended Data Fig. 1 | Non-interacting spectroscopic properties on another region. a, STM topography showing the moiré superlattice with \( \theta = 1.14^\circ \). b, Scanning tunnelling spectra measured on an AA site for slight electron doping (\( V_g = -10 \text{ V}, V_{set} = -200 \text{ mV}, I_{set} = 500 \text{ pA}, V_{mod} = 0.5 \text{ mV} \)). The blue and green arrows mark the step-like features. c, Band structure calculated using the continuum model including the effects of strain and relaxation. \( \Lambda_m \) is a non-high-symmetry direction along which the Dirac points (protected by \( C_2/T \) symmetry) are located. The black dotted line indicates the Fermi level. The blue (green) dashed line corresponds to the VHS of the first conduction (valence) remote band. d, Corresponding \( \sqrt{\text{LDOS}} \) (offset by \(-28 \text{ meV}\)) on an AA site. The blue and green arrows mark the VHS of the first conduction and valence remote bands.
Extended Data Fig. 2 | Lifetime broadening. a, b. Peak width of the VHS as a function of the VHS energy when both flat bands are filled (a) or emptied (b). The green (a) and yellow (b) curves are the fit using $p(E) = p_0 + \lambda E^2$ with $p_0 = 5.25$ meV, $\lambda = 0.021$ meV$^{-2}$ for a and with $p_0 = 5.1$ meV, $\lambda = 0.00072$ meV$^{-2}$ for b. The black dotted line marks $E = 17$ meV.
Extended Data Fig. 3 | Tip band bending effect. a, Schematic (not to scale) of the electrostatic simulation for the device geometry in our experiment. b, Density under and away from the tip as a function of gate voltage, calculated using the geometry in a with height $z = 4\,\text{Å}$, radius $r = 0.6\,\text{nm}$, a work function difference of $0.25\,\text{V}$ and the band structure from Fig. 1e. c, Density and DOS at the Fermi level as a function of gate voltage under the tip with the same parameters as in b. d, e, Zero-bias conductance as a function of gate voltage with different set point conditions ($V_{\text{set}} = +200\,\text{mV}, I_{\text{set}} = 120\,\text{pA}$ for d and $V_{\text{set}} = -200\,\text{mV}, I_{\text{set}} = 500\,\text{pA}$ for e) showing different apparent gate efficiencies.
Extended Data Fig. 4 | Additional example of the breakdown of the non-interacting description. a, \( \frac{dI}{dV} \) spectra measured on an AA site as a function of sample bias (energy) and gate voltage on the region with \( \theta = 1.14^\circ \) (\( V_{\text{set}} = -200 \text{ mV}, I_{\text{set}} = 500 \text{ pA}, V_{\text{mod}} = 0.5 \text{ mV} \)). b, Normalized \( \frac{dI}{dV} \) spectra at different gate voltages extracted from a. The curves are shifted vertically by 0.5 nS each for clarity. The black dotted line indicates the Fermi level.
Extended Data Fig. 5 | Density-dependent spectroscopy on a non-magic-angle device. a, dI/dV measured on an AA site as a function of sample bias and gate voltage on the region with θ = 1.72° (V_{set} = −100 mV, I_{set} = 180 pA, V_{mod} = 2 mV). b–e, Normalized dI/dV spectra at different gate voltages extracted from a. The curves are shifted vertically by 0.4 each for clarity. The black dotted line indicates the Fermi level. We note that the electron and hole dopings are symmetric, which can be achieved by different tip conditioning (see Methods for discussion).
Extended Data Fig. 6 | Mean-field calculations. a, b, Normalized $\text{d}I/\text{d}V$ spectra obtained from a non-interacting model (a) and a Hartree–Fock calculation (b) for different filling factors. The curves are each shifted vertically by 0.3 each for clarity. c, Individual flavour filling as a function of total filling factor, indicating the presence of spontaneous spin or valley polarization near half-filling of the flat bands from the Hartree–Fock calculation.
Extended Data Fig. 7 | Extended Hubbard model with six sites.

**a**, Schematic of a six-site cluster two-orbital Hubbard model with on-site energy $\pm \epsilon$, hopping $t$, on-site Coulomb interaction $U$, near-neighbour interactions $V_0$ (same orbital) and $V_1$ (different orbitals). **b**, Local spectral weight computed from the exact diagonalization of a six-site cluster Hubbard model for different filling factors with $\epsilon = 9$, $t = 0.75$, $U = 30$, $V_0 = 5$, $V_1 = 3.8$ (see Methods). The curves beyond $\pm n_0$ are obtained by assuming a constant DOS at the Fermi level from the remote bands. The curves are shifted vertically by 0.25 each for clarity. **c**, Band broadening as a function of the ratio of the on-site Coulomb interaction $U$ to the non-interacting band width $6t$, while maintaining $V_0 = U/6$, $V_1 = U/7.9$. 
Extended Data Fig. 8 | Extended Hubbard model on a honeycomb lattice. 

**a**, Schematics of a 6-unit-cell lattice with one orbital per site, hopping $t$, nearest site interaction $V_0$ (different sublattices) and next-nearest site interaction $V_1$ (same lattices). 

**b**, Local spectral weight computed from the exact diagonalization of a 6-unit-cell lattice Hubbard model for different filling factors with $\epsilon = 8.5$, $t = 0.75$, $V_0 = 18.2$, $V_1 = 4.4$ (see Methods). The curves beyond $\pm n_0$ are obtained by assuming a constant DOS at the Fermi level from the remote bands. The curves are shifted vertically by 0.25 each for clarity. 

**c**, Band broadening as a function of the ratio of the nearest site interaction $V_0$ to the non-interacting band width $3t$, while maintaining $V_0 = 4.2V_1$. 