

**Supplementary information**

---

**Evidence for unconventional  
superconductivity in twisted trilayer  
graphene**

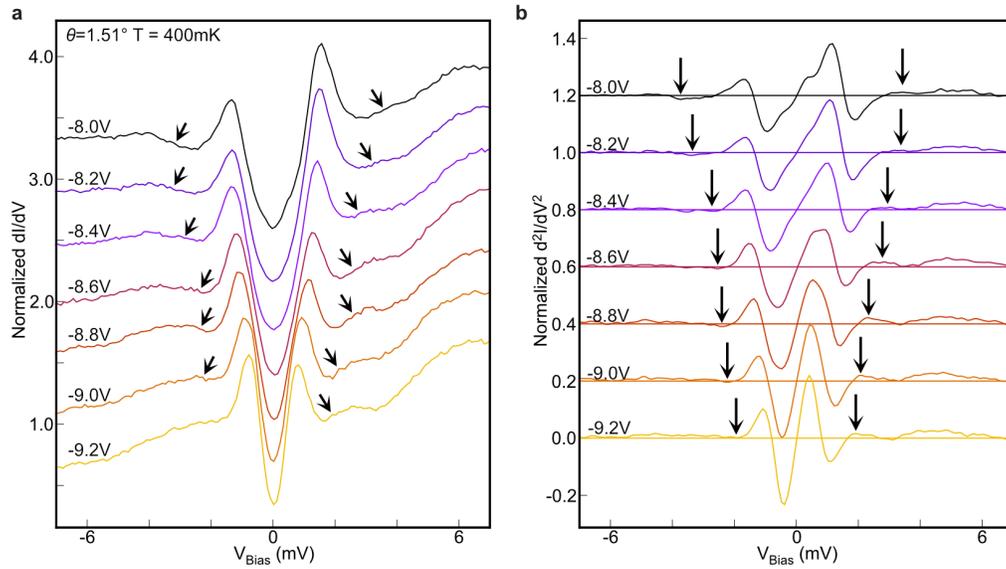
---

In the format provided by the  
authors and unedited

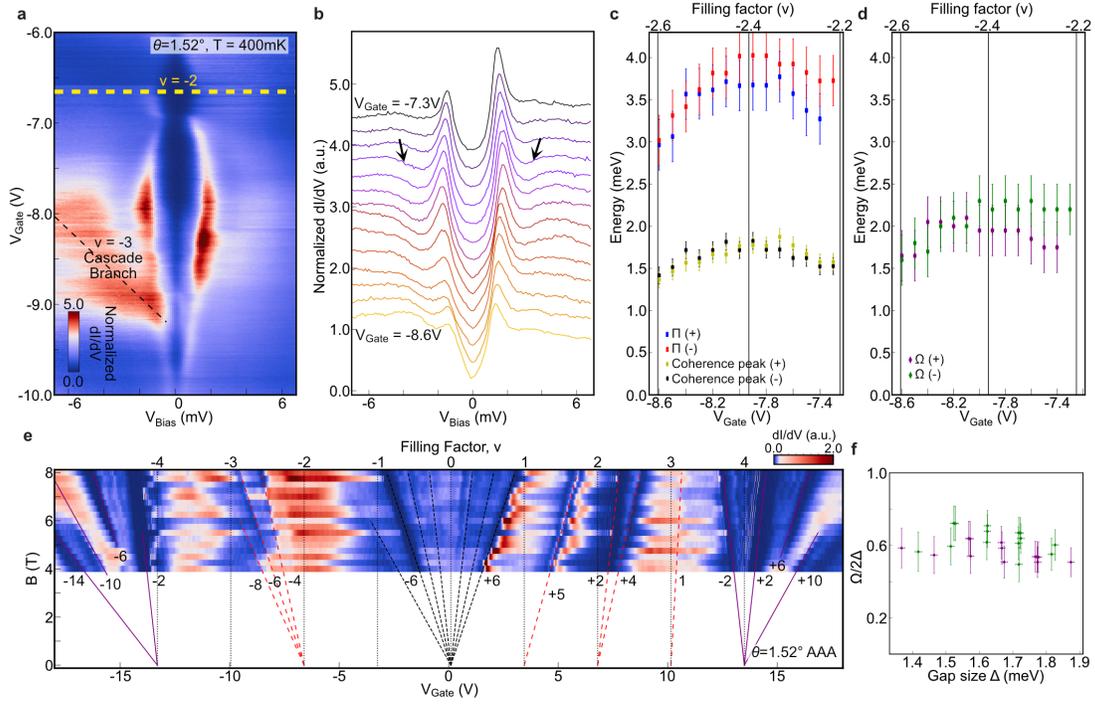
## Supplementary Information: Evidence for Unconventional Superconductivity in Twisted Trilayer Graphene

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

### 1. ADDITIONAL DATA



**Supplementary Information Fig. 1. Peak-dip-hump analysis from  $d^2I/dV^2$  local minima/maxima. a,** Hole-side superconducting gap spectrum measured at various  $V_{\text{Gate}}$  ranging from -8.0 V to -9.2 V at  $\theta = 1.51^\circ$  region which is same dataset as Fig. 4a. **b,**  $d^2I/dV^2$  as a function of  $V_{\text{Bias}}$  obtained by taking the first derivative of the data in (a) and applying a Gaussian filter. The horizontal lines of the same color indicate the  $d^2I/dV^2 = 0$  for each  $V_{\text{Gate}}$ .



**Supplementary Information Fig. 2. Dip-hump structures observed at different magic-angle area a,** Gate spectroscopy measured at  $\theta = 1.51^\circ$ . **b**, Normalized point spectra at range of  $V_{\text{Gate}}$  from  $-8.6$  V to  $-7.3$  V. **c**, Extracted position of the dip-hump and a coherence peak versus  $V_{\text{Gate}}$  for  $V_{\text{Bias}} > 0$  (blue and yellow, respectively) and for  $V_{\text{Bias}} < 0$  (red and black, respectively). **d**, Energy of the bosonic mode versus  $V_{\text{Gate}}$ , obtained by subtracting the corresponding energies of the dip-hump feature and the coherence peak for  $V_{\text{Bias}} > 0$  (purple) and  $V_{\text{Bias}} < 0$  (green). **e**, LDOS Landau fan diagram measured at the same area as **a** on AAA region. Black lines indicate the gap between LLs emanating from CNP. Red dashed lines indicate gaps between LLs emanating from integer filling  $\nu \neq 0$  of the flat bands. **f**,  $\Omega/2\Delta$  versus  $\Delta$  obtained from **c,d**. In this particular area the dip-hump structure could be resolved mostly in U-shaped regime.

## 2. STRIPE SIMULATION

As mentioned in the main text, the stripes are believed to arise out of the restructuring of the moiré lattice. The flat-band scenario of interest arises when the top and bottom monolayers are AA stacked—all carbon atoms vertically aligned—while the middle layer is rotated by a twist angle  $\sim 1.5^\circ$ . While this situation seems understandably difficult to achieve during fabrication, it was shown in Ref. S1 that the desired AA stacking of the top and bottom is the energetically preferred configuration, and we therefore expect the system to relax into this configuration across large regions of the sample. This expectation is borne out by the observation of flat bands as well as the presence of a single moiré lattice constant as seen in STM.

There are two primary features in Fig. 1b that we wish to reproduce. The first, and most prominent, is the stripe, which can be obtained as follows. We let  $\mathbf{a}_1 = a(0, 1)$  and  $\mathbf{a}_2 = a(-\sqrt{3}/2, -1/2)$  denote the Bravais primitive vectors of the graphene monolayer, where  $a \approx 0.246$  nm is the graphene lattice constant, and let  $R(\phi) = e^{-i\phi\sigma^y}$  be a matrix that rotates by angle  $\phi$ . The bottom and middle lattices are simulated by plotting points at  $\Lambda_{\text{bot}} = \{R(-\theta/2)(n_1\mathbf{a}_1 + n_2\mathbf{a}_2), n_{1,2} \in \mathbb{Z}\}$  and  $\Lambda_{\text{mid}} = \{R(\theta/2)(n_1\mathbf{a}_1 + n_2\mathbf{a}_2), n_{1,2} \in \mathbb{Z}\}$ . The stripe is then entirely determined by strain in the top layer, where the points plotted are instead  $\Lambda_{\text{top}} = \{R(-\theta/2)(n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + f_w(-\frac{1}{2}n_1 + n_2)\mathbf{v}), n_{1,2} \in \mathbb{Z}\}$  where  $f_w(x) = \frac{1}{\pi}\arctan(2x/w) + \frac{1}{2}$  and  $\mathbf{v} = m_1\mathbf{a}_1 + m_2\mathbf{a}_2$ ,  $m_{1,2} \in \mathbb{Z}$  is a Bravais lattice vector. The function  $f_w(x)$  is essentially a smoothed step function in that it interpolates between 0 and 1:  $\lim_{x \rightarrow -\infty} f_w(x) = 0$  and  $\lim_{x \rightarrow +\infty} f_w(x) = 1$ . The size of the intermediate regime and hence the stripe width is determined by the parameter  $w > 0$ , with  $\lim_{w \rightarrow 0} f_w(x) = \Theta(x)$ , the Heaviside function. In our definition of  $\Lambda_{\text{top}}$ , we chose to have  $f_w$  as a function  $-\frac{1}{2}n_1 + n_2$  since it results in a stripe along the  $(-1/2, \sqrt{3}/2)$  direction and thus well represents the stripes shown in Fig. 1b. Putting these pieces together, one can see that in both regions where  $|-\frac{1}{2}n_1 + n_2|$  is large, the lattice points of  $\Lambda_{\text{bot}}$  and  $\Lambda_{\text{top}}$  should be directly above one another. In the region  $-w/2 \lesssim -\frac{1}{2}n_1 + n_2 \lesssim w/2$ , the registry of the top and bottom layers changes from AA to AB and then back to AA.

The procedure detailed above yields a stripe, but does not account for a second feature of Fig. 1b: the moiré lattices on either side of the stripe are offset by about half of a moiré unit cell in the vertical ( $\hat{y}$ ) direction, which corresponds to a displacement of  $\mathbf{D}_{\text{shift}} = \frac{\sqrt{3}}{2}L_M\hat{y}$ ,  $L_M = a/(2\sin(\theta/2))$ . This offset at the moiré lattice scale is a result of a shift of the top and bottom lattices relative to the middle lattice occurring at the level of the microscopic scale of monolayer graphene. In particular, displacing the top and bottom layers by  $\mathbf{v}_{\text{shift}} \approx \theta\hat{z} \times \mathbf{D}_{\text{shift}} \approx -\frac{\sqrt{3}}{2}a\hat{x}$  moves the moiré lattice by  $\mathbf{D}_{\text{shift}}$ . Such a shift is readily implemented numerically by replacing the lattices  $\Lambda_{\text{bot}}$  and  $\Lambda_{\text{top}}$  with  $\Lambda'_{\text{bot}} = \{R(-\theta/2)(n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + f_w(-\frac{1}{2}n_1 + n_2)\mathbf{v}_{\text{shift}}), n_{1,2} \in \mathbb{Z}\}$  and  $\Lambda'_{\text{top}} = \{R(-\theta/2)(n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + f_w(-\frac{1}{2}n_1 + n_2)(\mathbf{v} + \mathbf{v}_{\text{shift}})), n_{1,2} \in \mathbb{Z}\}$ . The middle layer is defined through  $\Lambda_{\text{mid}}$  as in the previous paragraph. For ease of visualization,  $\Lambda'_{\text{top}}$  and  $\Lambda'_{\text{bot}}$  are plotted in black while  $\Lambda_{\text{mid}}$  is plotted in red.

We emphasize that the primary purpose of this calculation is to reproduce the stripe in the simplest possible manner. A more complete study requires understanding the energetics, which would not only be needed to predict that width of the stripe (here, simply an input parameter), but which would also result in lattice relaxation within a unit cell.

### 3. CONTINUUM MODEL AND INTERACTION-DRIVEN BAND STRUCTURE RENORMALIZATION

#### A. Continuum model

In this section, we summarize the continuum model<sup>S1,S2</sup> used to capture the low-energy theory of twisted trilayer graphene. In particular, we consider the case where the top and bottom layers are directly atop one another (AA stacked) and twisted by  $-\theta/2$ , while the middle layer is twisted by  $+\theta/2$ . The electronic structure of MATTG is obtained by an extension<sup>S2</sup> of the continuum model developed originally for twisted bilayer graphene (TBG)<sup>S3</sup>. As in that case, there are two independent sectors in the non-interacting limit distinguished by the valley  $K$  and  $K'$ . Without loss of generality, we therefore focus on valley  $K$  in this section; the model relevant to valley  $K'$  may be obtained in a straightforward manner through time reversal. We let  $\psi_t$ ,  $\psi_m$ , and  $\psi_b$  denote the spinors one obtains by expanding the dispersion of monolayer graphene about valley  $K$  for the top, middle and bottom layers, respectively. In terms of the microscopic operators of the graphene monolayers, that means  $\psi_\ell(\mathbf{k}) = f_\ell(\mathbf{k} + \mathbf{K}_\ell)$ ,  $\ell = t, m, b$ . Importantly, as a result of the twist, the  $K$  points of the different layers are not the same. The model is composed of a ‘diagonal’ Dirac piece and an ‘off-diagonal’ tunneling piece accounting for the moiré interlayer coupling:  $H_{\text{cont}} = H_D + H_{\text{tun}}$ . The Dirac term is broken up into three components, one for each layer, with  $H_D = H_t + H_m + H_b$  where

$$H_\ell = \int_{\mathbf{k}} \psi_\ell^\dagger(\mathbf{k}) h_{\theta_\ell}(\mathbf{k}) \psi_\ell(\mathbf{k}), \quad h_{\theta_\ell}(\mathbf{k}) = -v_0 e^{i\theta_\ell \sigma^z / 2} (k_x \sigma^x + k_y \sigma^y) e^{-i\theta_\ell / 2}. \quad (1)$$

Above,  $\ell = t, m, b$  identifies the layers,  $v_0 \sim 10^6$  m/s is the Fermi velocity of the Dirac cones of monolayer layer graphene, and  $\sigma^{x,y,z}$  are Pauli matrices acting on the A/B sublattice indices of the spinors  $\psi_\ell$ . The angle  $\theta_\ell$  indicates the angle by which each layer is rotated, with  $\theta_t = \theta_b = -\theta/2$  and  $\theta_m = +\theta/2$ . The magic angle for this model occurs for  $\theta \approx 1.5^\circ$ , which is related to the magic angle of TBG through a prefactor of  $\sqrt{2}$ :  $\theta = 1.5^\circ \approx \sqrt{2} \times 1.05^\circ$ . The origins of this relation trace back to a similarity transformation that maps the MATTG continuum model into one of a decoupled TBG-like band structure with an interlayer coupling (to be discussed) multiplied by  $\sqrt{2}$  and a graphene-like Dirac cone. We refer to Ref. S2 for an in-depth explanation of this relation.

We assume that tunneling only occurs between adjacent layers:

$$H_{\text{tun}} = \sum_{j=1,2,3} \int_{\mathbf{k}} \left( \psi_t^\dagger(\mathbf{k}) + \psi_b^\dagger(\mathbf{k}) \right) T_j \psi_m(\mathbf{k} + \mathbf{q}_j) + h.c., \quad (2)$$

where the momenta shift and the tunneling matrices are given by

$$\mathbf{q}_j = \frac{4\pi}{3L_M} R \left( \frac{2\pi}{3}(j-1) \right) \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad T_j = w_0 + w_1 \left( e^{-2\pi(j-1)i/3} \sigma^+ + e^{2\pi(j-1)i/3} \sigma^- \right) \quad (3)$$

with  $R(\phi) = e^{-i\phi\sigma^y}$  is a  $2 \times 2$  matrix acting on vector indices,  $L_M = a/[2 \sin(\theta/2)]$ , and  $\sigma^\pm = (\sigma^x \pm i\sigma^y)/2$ . The tunneling strength is determined by the parameters  $w_0$  and  $w_1$ ; in this paper we set  $(w_0, w_1) = (55, 105)$  meV. (Note that the conventions used in this section are rotated by  $90^\circ$  relative to those of section 2.)

This model possesses a number of symmetries. We have already alluded to time reversal, with which one may obtain the continuum model Hamiltonian corresponding to the valley  $K' = -K$ .

We therefore re-introduce a valley label, writing  $\psi_\ell \rightarrow \psi_{v,\ell}$  with  $v = K, K'$ . A number of spatial symmetries are also present in this model, but for our purposes it is sufficient to note that the model is invariant under rotations by  $60^\circ$ , under which the spinors transform as  $\psi_\ell(\mathbf{k}) \rightarrow \tau^x \sigma^x e^{2\pi i \tau^z \sigma^z / 3} \psi_\ell(R(2\pi/6)\mathbf{k})$ , where  $\tau^{x,y,z}$  are Pauli matrices acting on the (now suppressed) valley indices.

To diagonalize the continuum model, we recall that the spinor operators  $\psi_\ell$  are not all defined about a common momentum point. Hence the tunneling term in Eq. (2) does not involve a momentum exchange of  $\mathbf{q}_j$ , but rather  $K_t = K_b = K_m + \mathbf{q}_j$  and  $K'_t = K'_b = K_m - \mathbf{q}_j$ , which differ by a moiré reciprocal lattice vector. We therefore define operators  $\Psi_{v,\ell}$  about a common momentum point for each valley through  $\Psi_{v,t/b}(\mathbf{k}) = \psi_{v,t/b}(\mathbf{k})$  and  $\Psi_{K/K',m}(\mathbf{k}) = \psi_{K/K',m}(\mathbf{k} \pm \mathbf{q}_1)$ , where the  $+$  ( $-$ ) corresponds to  $K$  ( $K'$ ) (the choice  $\mathbf{q}_1$  is arbitrary— $\mathbf{q}_2$  and  $\mathbf{q}_3$  could be equally chosen). Grouping the valley, layer, sublattice, and spin labels into a single indice,  $\Psi_\alpha$ , we can express  $H_{\text{cont}}$  in matrix form as

$$H_{\text{cont}} = \sum_{\mathbf{G}, \mathbf{G}'} \int_{\mathbf{k} \in \text{mBZ}} \Psi_\alpha^\dagger(\mathbf{k} + \mathbf{G}) h_{\alpha, \mathbf{G}; \alpha', \mathbf{G}'}^{(\text{cont})}(\mathbf{k}) \Psi_{\alpha'}(\mathbf{k} + \mathbf{G}'); \quad (4)$$

$\mathbf{G}$  and  $\mathbf{G}'$  are moiré reciprocal lattice vectors defined via  $\mathbf{G} = n_1 \mathbf{G}_1 + n_2 \mathbf{G}_2$ ,  $n_{1,2} \in \mathbb{Z}$  where  $\mathbf{G}_1 = \mathbf{q}_2 - \mathbf{q}_1$  and  $\mathbf{G}_2 = \mathbf{q}_3 - \mathbf{q}_1$ . The integration over  $\mathbf{k}$  includes only those momenta within the moiré Brillouin zone (mBZ).

## B. Interaction-driven band structure renormalization

The presence of flat bands in MATTG necessitates the consideration of interaction-driven band structure corrections. As demonstrated experimentally in our previous work on twisted graphene bilayers<sup>S4</sup>, filling-dependent interaction effects, specifically Hartree and Fock corrections, drastically alter the electron dispersion. Here we incorporate only a Hartree mechanism<sup>S5–S8</sup> in the analysis. In TBG we found<sup>S4</sup> that the main role of the Fock correction, provided that one does not consider the nature of the correlated states and the cascade, is to broaden the band structure at the charge neutrality point ( $\nu = 0$ ) and to counteract band inversions at the zone center promoted by Hartree effects. For comparison with the experiment presented in Fig. 2, where we focus only on  $\nu = \pm 4$ , we can thus ignore Fock corrections as a first approximation. Similar Hartree-driven band structure renormalizations were considered recently in the literature<sup>S9,S10</sup>, and our analysis together with the experimental results are consistent with their conclusions.

We introduce Coulomb interaction into the system through

$$H_C = \frac{1}{2} \int d^2\mathbf{r} d^2\mathbf{r}' \delta\rho(\mathbf{r}) V(\mathbf{r} - \mathbf{r}') \delta\rho(\mathbf{r}'). \quad (5)$$

Here,  $V(\mathbf{r}) = e^2/(4\pi\epsilon|\mathbf{r}|)$  is the Coulomb potential and  $\delta\rho(\mathbf{r}) = \Psi^\dagger(\mathbf{r})\Psi(\mathbf{r}) - \rho_{\text{CN}}(\mathbf{r})$ , where  $\rho_{\text{CN}}(\mathbf{r}) = \langle \Psi^\dagger(\mathbf{r})\Psi(\mathbf{r}) \rangle_{\text{CN}}$  is the expectation value of the density at the charge neutrality point. The use of  $\delta\rho(\mathbf{r})$  instead of  $\rho(\mathbf{r})$  in the interaction is motivated by the expectation that the input parameters of the model  $H_{\text{cont}}$  already include the effect of interactions at the charge neutrality point. Although numerically expedient, this assumption is not strictly correct since the input parameters in actuality refer to three independent graphene monolayers. Nevertheless, for the purpose of making qualitative comparisons with Fig. 2, we do not expect this distinction to be important. The dielectric constant  $\epsilon$  in the definition of  $V(\mathbf{r})$  is used as a fitting parameter; see section 3 C for details.

We study the effect of the interacting continuum model of MATTG through a self-consistent Hartree mean-field calculation. Instead of solving the many-body problem, we obtain the quadratic Hamiltonian that best approximates the full model when only the symmetric contributions of  $H_C$  are included, i.e., the Fock term is neglected as explained above. Thus instead of  $H_{\text{cont}} + H_C$ , we study the Hamiltonian

$$H_{\text{MF}}^{(\nu)} = H_{\text{cont}} + H_{\text{H}}^{(\nu)} - \frac{1}{2} \langle H_{\text{H}}^{(\nu)} \rangle_{\nu}, \quad (6)$$

where  $H_{\text{H}}^{(\nu)}$  is the Hartree term at filling  $\nu$ ,

$$H_{\text{H}}^{(\nu)} = \int_{\mathbf{k}, \mathbf{k}', \mathbf{q}} V(\mathbf{q}) \langle \Psi^\dagger(\mathbf{k}' + \mathbf{q}) \Psi(\mathbf{k}') \rangle_{\nu} \Psi^\dagger(\mathbf{k}) \Psi(\mathbf{k} - \mathbf{q}), \quad (7)$$

and the last term in Eq. (6) simply ensures there is no double counting when one calculates the total energy. In the above equation,  $V(\mathbf{q}) = 2\pi e^2 / (\epsilon |\mathbf{q}|)$  is the Fourier transform of the Coulomb interaction  $V(\mathbf{r})$  in Eq. (5), and the expectation value  $\langle \hat{\mathcal{O}} \rangle_{\nu} = \langle \hat{\mathcal{O}} \rangle_{\text{occ}} - \langle \hat{\mathcal{O}} \rangle_{\text{CN}}$  only includes states that are filled up to  $\nu$  relative to charge neutrality, as defined by diagonalizing the Hamiltonian  $H_{\text{MF}}^{(\nu)}$ . Typically, for a ‘‘jellium’’-like model, the expectation value vanishes save for  $\mathbf{q} = 0$ , which is subsequently cancelled by the background charge—allowing one to set  $V(\mathbf{q} = 0) = 0$  and completely ignore the Hartree interaction. However, because the moiré pattern breaks continuous translation symmetry, momentum is only conserved modulo a reciprocal lattice vector. We therefore obtain

$$H_{\text{H}}^{(\nu)} = \sum'_{\mathbf{G}} V(\mathbf{G}) \int_{\mathbf{k}'} \langle \Psi^\dagger(\mathbf{k}' + \mathbf{G}) \Psi(\mathbf{k}') \rangle_{\nu} \int_{\mathbf{k}} \Psi^\dagger(\mathbf{k}) \Psi(\mathbf{k} - \mathbf{G}), \quad (8)$$

where the prime above the summation over the moiré reciprocal lattice vectors indicates that  $\mathbf{G} = 0$  is excluded. The self-consistent procedure begins by assuming some initial value of  $H_{\text{H}}^{(\nu)}$  and diagonalizing the corresponding mean-field Hamiltonian  $H_{\text{MF}}^{(\nu)}$  to obtain the Bloch wavefunctions and energy eigenvalues. These quantities are then used re-compute the expectation values that define  $H_{\text{H}}^{(\nu)}$  and thus  $H_{\text{MF}}^{(\nu)}$ . This process is repeated until one obtains the quadratic Hamiltonian  $H_{\text{MF}}^{(\nu)}$  that yields the correlation functions  $\langle \cdot \rangle_{\nu}$  used in its definition.

It has further been shown<sup>S5,S11</sup> that the Hartree potential is dominated by the first ‘star’ of moiré reciprocal lattice vectors, which in our conventions corresponds to  $\mathbf{G}_n = R(2\pi(n - 1)/6) \frac{4\pi}{\sqrt{3}L_M} (1, 0)^T$  for  $n = 1, \dots, 6$ , with  $R(\phi)$  a rotation matrix. In this last approximation that we employ, the  $2\pi/6$  rotation symmetry of the continuum model greatly simplifies the calculation of the Hartree term. Notably,  $V(\mathbf{G}) \int_{\mathbf{k}'} \langle \Psi^\dagger(\mathbf{k}' + \mathbf{G}) \Psi(\mathbf{k}') \rangle_{\nu}$  must be the same for all  $\mathbf{G}_n$ , and, instead of Eq. (8), we use

$$H_{\text{H}}^{(\nu)} = V_{\text{H}}^{(\nu)} \sum_{n=1}^6 \int_{\mathbf{k}} \Psi^\dagger(\mathbf{k}) \Psi(\mathbf{k} - \mathbf{G}_n), \quad V_{\text{H}}^{(\nu)} = \frac{1}{6} \sum_{n=1}^6 V(\mathbf{G}_n) \int_{\mathbf{k}'} \langle \Psi^\dagger(\mathbf{k}' + \mathbf{G}) \Psi(\mathbf{k}') \rangle_{\nu}. \quad (9)$$

The self-consistent procedure in this case is identical to that described in the previous paragraph, but due to the reduced number of reciprocal lattice vectors that are included in the summation the calculation is computationally easier. Convergence is typically reached within  $\sim 6$  iterations.

For clarity, all bands corresponding to different fillings plotted in Fig. 2b have been shifted so that the Dirac points of the flat bands always occur at the zero of the energy scale; it follows

that the (independent) graphene-like Dirac cone is then displaced in energy relative to the fixed reference point of the flat bands for each filling. If this procedure was not performed for clarity purposes, then the Hartree calculation would yield band structures with a graphene-like Dirac cone fixed at one energy for all fillings, but with shifted flat bands relative to it, as predicted in ab-initio calculations<sup>S9</sup>.

### C. Hartree correction and estimate of dielectric constant

As discussed in the previous section, due to Hartree corrections, the Dirac cones shift downwards (upwards) in energy relative to the flat bands under electron (hole) doping, as seen in Fig. 2b-d. These relative shifts are measured to be rather large ( $\approx 70$  meV for  $\nu = +4$  and  $\approx 50$  meV for  $\nu = -4$ ), similar to the bandwidth of the MATTG flat bands (approximately 50 meV). These relative shifts allow us to estimate an effective dielectric constant  $\epsilon$  to be used in Hartree band-structure-renormalization calculations. In particular, we find that  $\epsilon = 12 - 13$  quantitatively reproduces the observed Dirac point shifts at  $\nu = \pm 4$ .

## 4. TUNNELING CONDUCTANCE NORMALIZATION AND FITTING PROCEDURE

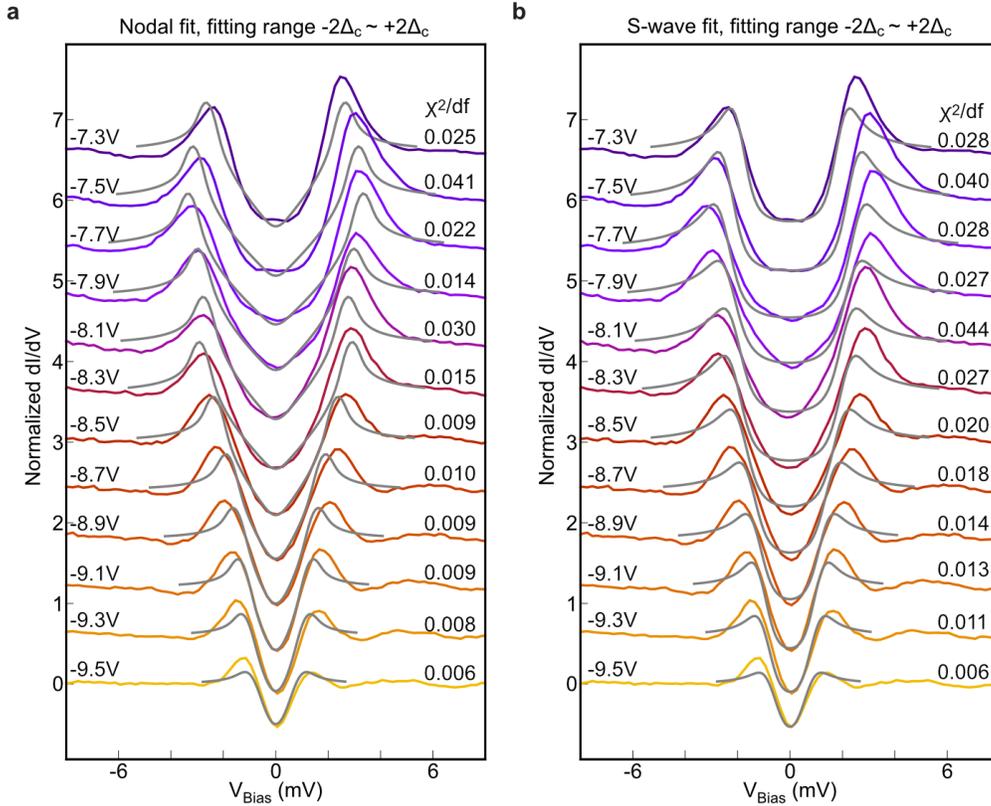
In Fig. 3b,c the tunneling conductance has been normalized by dividing the spectra with a sixth-order polynomial fit that preserves the area of the spectrum<sup>S12</sup> (see also Extended Data Fig. 1). This procedure returns normalized  $dI/dV$  curves that approach unity outside of the spectroscopic gap and removes in part the large asymmetry between electrons and holes near  $\nu = -2$  and above  $V_{\text{Bias}} = 5$  meV. We emphasize that the regimes displaying U- and V-shaped tunneling spectra are clearly visible both before and after this normalization procedure. The dip-hump structure persists after this step as well (see black arrow in Extended Data Fig. 1).

The normalized  $dI/dV$  curves are fitted with the Dynes formula<sup>S13</sup>,

$$\frac{dI}{dV} \propto \int_{-\infty}^{\infty} d\omega \int_0^{2\pi} d\theta \operatorname{Re} \left[ \frac{\omega + i\Gamma}{\sqrt{(\omega + i\Gamma)^2 - \Delta(\theta)^2}} \right] \left( -\frac{df}{d\omega} \right) \Big|_{\omega=\omega+eV}, \quad (10)$$

where  $f(\omega) = 1/(e^{\omega/k_B T} + 1)$  ( $k_B$  is a Boltzmann constant and  $T = 400$  mK in our measurements);  $\Delta(\theta)$  is the superconducting pairing potential and; spectral broadening coming from disorder and finite lifetime of Cooper pairs are incorporated by the parameter  $\Gamma$ . We consider isotropic  $s$ -wave pairing, a pairing with a nodal order parameter, and a combination of the two. For the nodal case we use  $\Delta(\theta) = \Delta_0 \cos(2\theta)$  (i.e., a  $d$ -wave profile), though any  $\Delta(\theta) = \Delta_0 \cos(N\theta)$  with integer  $N \neq 0$  gives the same spectrum. We therefore do not distinguish between different nodal order parameters, e.g.,  $p$ - versus  $d$ - versus  $f$ -wave. In the plots, we also took into account the broadening due to finite lock-in modulation excitation  $V_{\text{mod}} = 200\mu$  V. Overview of the fitting procedures and various fits is presented in Extended Data Fig. 9, Extended Data Fig. 10, Extended Data Fig. 11 and SI Fig. 3. In general, the quality of the fit can be quantified by the  $\chi^2$  value normalized with the number of different  $V_{\text{Bias}}$  points ('degrees of freedom', i.e. df) used in the fit.

The existence of the U and V regions can be showed by undergoing detailed analysis of the tunneling profile fits for different backgate voltages  $V_{\text{Gate}}$ , see Extended Data Fig. 10 and Extended Data Fig. 11. In the V-shaped region (Extended Data Fig. 10 and Extended Data Fig. 9f-k), the nodal fits match the data much better compared to the  $s$ -wave fit regardless of the range of bias used in the fit. This is both evident by visually inspecting the matching between the fits and the

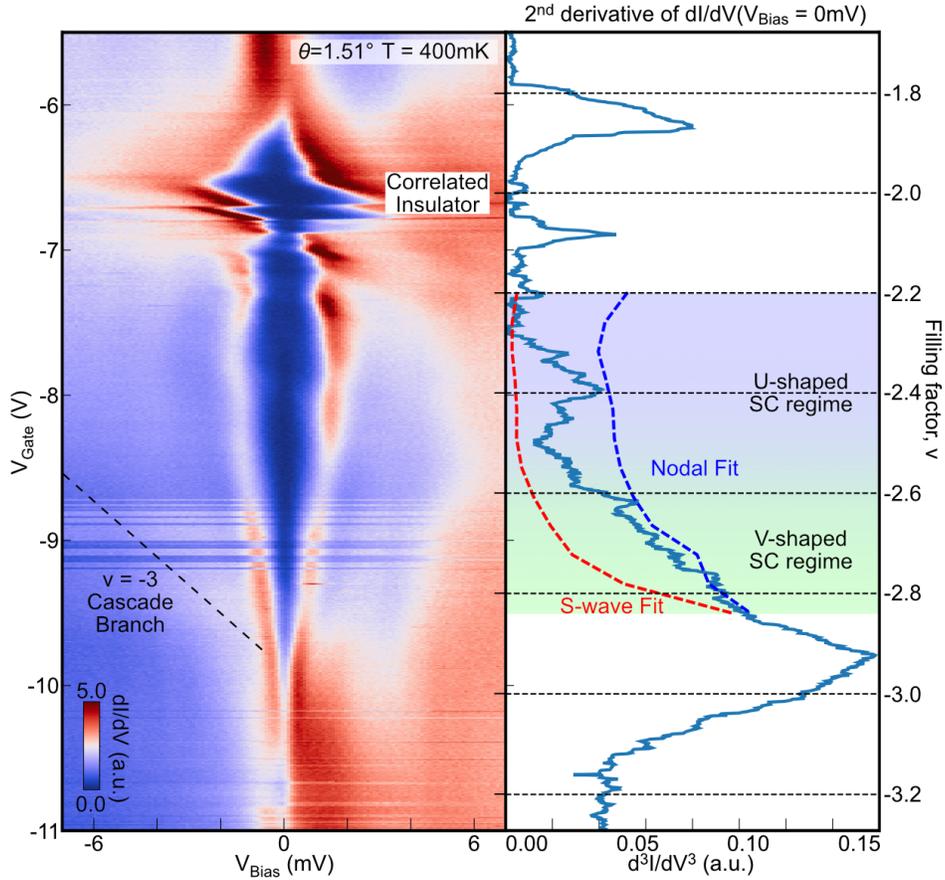


**Supplementary Information Fig. 3. Overview of the best fits to Dynes formula fit using s-wave and nodal gap for different  $V_{\text{Gate}}$ .** **a**, Line traces of  $dI/dV$  spectra for  $V_{\text{Gate}}$  within  $-2\Delta_c$  to  $2\Delta_c$  range is used to obtain the fit parameters  $\Delta$  and  $\Gamma$  where  $\Delta_c$  is defined as half of coherence peak separation for each  $V_{\text{Gate}}$ .  $\chi^2/df$  value for each fit is displayed. **b**, Same traces as (a) with s-wave pairing symmetry fit with same range of  $V_{\text{Bias}}$  was used for fit parameter extraction.

data in the whole range of bias voltages within the gap and in particular focusing on the regions around  $V_{\text{Bias}} = 0$  V (see insets in Extended Data Fig. 10a,b) as well as comparing the  $\chi^2/df$  values for the s-wave and nodal fits. This comparison directly points towards the nodal nature of the gap within this gate voltage range.

However, in the U-shaped region this simple comparison is no longer valid (Extended Data Fig. 11 and Extended Data Fig. 9c-e). Here the s-wave fits match the data much better (in comparison to nodal fit), especially in the region within the coherence peaks (Extended Data Fig. 11a,b). Moreover, when we narrow the range of bias to half of the width of the coherence peak separation, the s-wave fit reproduces the U-shaped profile remarkably well (See Extended Data Fig. 11c). This distinction between V-shaped and U-shaped region is further corroborated by comparison of the  $\chi^2$  values for the nodal and s-wave fits taken in different ranges of bias voltage. A comprehensive summary of the respective best fits are shown in SI Fig. 3. From comparison of  $\chi^2$  values (and visual inspection), it is clear that for  $V_{\text{Gate}} \lesssim -8.4$  V (the V-shaped regime), the nodal pairing profile produces a consistently better fits.

We note that in the U-shaped regime and for the larger fitting range,  $\chi^2$  values are relatively



**Supplementary Information Fig. 4. Distinction of U-shaped and V-shaped SC gapped region from second derivative of tunneling conductance.**  $d^3I/dV^3(V_{\text{Bias}} = 0 \text{ mV})$  plotted side-by-side to the SC gatemap to quantitatively evaluate how much the spectrum is close to V-shape (high curvature, high  $d^3I/dV^3$ ) or U-shape (low curvature, low  $d^3I/dV^3$ ). Because the gap size  $\Delta$  and quasiparticle broadening  $\Gamma$  also affects the  $d^3I/dV^3$ , we plotted expected  $d^3I/dV^3$  as a function of  $V_{\text{Gate}}$  by using the  $\Delta$  and  $\Gamma$  extracted from Extended Data Fig. 10. Within  $V_{\text{Gate}}$  range we are interested in, s-wave fit gives much lower  $d^3I/dV^3$  compared to Nodal fit which agrees with our intuition.

large for both s-wave and node fits. This observation likely indicates that in the U-regime regions around the coherence peaks are probably described by physics that goes beyond simple Dynes models regardless of gap order parameter. We also note that Dynes formula fit parameter  $\Gamma$  extracted from s-wave gap is approximately  $150 \mu\text{eV}$  over the whole range of gate voltages while nodal gap gives values in the range of  $\Gamma \approx 10 \mu\text{eV}$ , order of magnitude smaller. These values are similar to the ones found in Ref. S14, where unusually large values of  $\Gamma$  for the s-wave gap fitting was argued to indicate that nature of superconductivity in twisted graphene bilayers is unlikely to be s-wave. In this context, much better fits could be obtained by using more involved order parameter such as  $s + id$  or  $d + id$  (see Extended Data Fig. 9d,e).

We emphasize that regardless of the fitting details, direct comparison focusing on the narrow range  $V_{\text{Bias}} = 0 \text{ V}$ , clearly contrasts the behaviour of the data and quality of the fits between

V-shaped and U-shaped regions and gives a clue that within the U-shaped region, one should invoke a gapped superconducting spectrum over the gapless one in order to match better with the experimental tunneling profiles. This qualitatively distinct nature of the two tunneling regimes is substantiated by carrying out a numerical derivative of the tunneling conductance. In particular, the second derivative of  $dI/dV$  ( $d^3I/dV^3$ ) versus  $V_{\text{Bias}}$  gives a quantitative measure of the curvature of the tunneling profile around zero bias. A linecut at  $V_{\text{bias}} = 0$  V of  $d^3I/dV^3$  is shown on the right of SI Fig. 4 alongside the  $dI/dV$  colormap. We emphasize that, in the V-shaped region, (starting around  $V_{\text{Gate}} \lesssim -8.4$  V), the curvature signal ( $d^3I/dV^3$ ) is much larger than one would expect from an s-wave tunneling profile obtained by fitting; by comparison, it follows well with the curvature obtained from the nodal gap profile. However, in the U-shaped region ( $V_{\text{Gate}} \gtrsim -8.4$  V), the curvature is intermediate between the values predicted by nodal and s-wave fits, a clear deviation from the behavior in the V-shaped region.

## 5. POSSIBLE SCENARIOS OF U-SHAPED TO V-SHAPED SPECTRAL EVOLUTION

In the main text, we introduced the experimental observation that the tunneling conductance exhibits two qualitatively different tunneling profiles (U- vs. V-shaped) as a function of filling. We now discuss the details of two possible scenarios for this outcome: (i) a BCS-like superconductor with filling-dependent order parameter symmetry and (ii) a BEC-to-BCS transition with a common nodal order parameter. As noted in the main text, we emphasize that ‘BCS’ in this context does *not* imply any assumptions regarding the pairing mechanism or coupling strength, but simply refers to a pairing scenario wherein the chemical potential lies inside the band. Finally, we discuss the Ginzburg-Landau coherence length in the BEC-BCS transition scenario and argue that it is consistent with the results of Ref. S15.

### A. BCS-like superconductor with filling-dependent order parameter symmetry

The existence of U- and V-shaped tunneling spectra suggests that superconductivity evolves with doping from a fully gapped to a gapless state. Here we address the possibility that these two regimes both arise from Cooper pairing a partially filled band with a Fermi surface, but with qualitatively different superconducting order parameters. This scenario *a priori* does not address the different behaviors of the Ginzburg-Landau coherence length  $\xi_{\text{GL}}$  seen in Ref. S15, e.g., the scaling of  $\xi_{\text{GL}}$  with the interparticle spacing (see section 5 B 2). Nevertheless, whatever mechanism underlies the putative change in order parameter could potentially conspire to yield such dependence.

The V-shaped spectra can be adequately fit by postulating a nodal order parameter, as described in the main text and in section 4. In the present scenario, the U-shaped spectra are best fit by invoking multiple co-existing order parameters: either an s-wave gap together with a nodal order parameter or a combination of two nodal order parameters (e.g.,  $d_{x^2-y^2} + id_{xy}$ ) that together produce a gap in the tunneling conductance. Extended Data Fig. 9e displays the relevant fits. As noted in the main text, a similar change in pairing order with doping has been proposed in cuprates<sup>S16</sup> (albeit with a less pronounced U-to-V evolution). Moreover, it has been argued that have argued that a  $d_{x^2-y^2} + id_{xy}$  spin-fluctuation-mediated pairing is energetically unfavourable compared to a real superposition of the two order parameters.<sup>S9</sup>

## B. BEC-to-BCS transition

### 1. Tunnelling current

To describe the tunneling current expected in the BEC-BCS transition scenario and demonstrate qualitative consistency with experiment, we consider a phenomenological two-parabolic-band model. Specifically, we model the system near filling  $\nu = -2$  with two bands of energy (in these two sections we set  $\hbar = 1$ )

$$\xi_{\pm, \mathbf{k}} = \pm \left( \frac{k^2}{2m} + \Delta_{\text{CI}} \right) - \mu, \quad (11)$$

separated by a  $2\Delta_{\text{CI}}$  correlated-insulator gap. Each band admits a two-fold ‘spin’ degeneracy—which need not coincide exactly with physical spin, but could, e.g., represent some combination of spin and valley. In the absence of pairing,  $\mu$  residing in the electron band  $\xi_+$  (hole band  $\xi_-$ ) corresponds to filling  $\nu = -2 + \delta n$  with  $\delta n > 0$  ( $\delta n < 0$ ). We focus primarily on the hole-doping case relevant for experiment.

For simplicity, we assume a ‘spin’-singlet, nodal  $d$ -wave pairing with a pair field  $\Delta_{\mathbf{k}}$  that is the same in the electron and hole bands; inter-band pairing is neglected. (We anticipate that triplet pairing would yield similar results, as would other nodal order parameters.) The standard Bogoliubov–de Gennes formalism yields

$$E_{\pm, \mathbf{k}} = \sqrt{\xi_{\pm, \mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}, \quad u_{\pm, \mathbf{k}}^2 = 1 + \frac{\xi_{\pm, \mathbf{k}}}{E_{\pm, \mathbf{k}}}, \quad v_{\pm, \mathbf{k}}^2 = 1 - \frac{\xi_{\pm, \mathbf{k}}}{E_{\pm, \mathbf{k}}} \quad (12)$$

with  $u_{\pm, \mathbf{k}}^2, v_{\pm, \mathbf{k}}^2$  coherence factors describing overlap of the bare electron/hole wavefunctions with those of quasiparticles with dispersion  $E_{\pm, \mathbf{k}}$ . The BEC phase corresponds to  $|\mu| < \Delta_{\text{CI}}$ . Here  $\Delta_{\text{CI}}$  renders the quasiparticles fully gapped despite the assumed nodal  $d$ -wave order parameter, and population of the electron and hole bands arises solely from pairing. (At  $\mu = 0$ , the symmetry built into the electron and hole bands implies that the system remains undoped, corresponding to  $\nu = -2$ , even with  $\Delta_{\mathbf{k}} \neq 0$ .) The regime  $|\mu| > \Delta_{\text{CI}}$  corresponds to a BCS phase wherein an electron- or hole-like Fermi surface undergoes Cooper pairing, yielding gapless quasiparticle excitations due to nodes in  $\Delta_{\mathbf{k}}$ . Figure 3i,j schematically depicts the chemical potential associated with these two phases.

The tunneling current follows from

$$I(eV, \mu) \propto \sum_{s=\pm} \int d^2 \mathbf{k} \left\{ u_{s, \mathbf{k}}^2 [f(E_{s, \mathbf{k}} - eV) - f(E_{s, \mathbf{k}})] - v_{s, \mathbf{k}}^2 [1 - f(-E_{s, \mathbf{k}} - eV) - f(E_{s, \mathbf{k}})] \right\}, \quad (13)$$

where  $f(E) = 1/(e^{E/k_B T} + 1)$  is the Fermi-Dirac distribution; the differential tunneling conductance  $dI/dV$  is obtained by numerically differentiating the current after the integral is evaluated. Below we will use this general formula to evaluate the tunneling conductance across the BEC-BCS transition. As an example, however, it is instructive to examine limiting cases.

Consider first the conductance deep in the BCS phase. Here the current simplifies dramatically for relevant voltages. First, focusing on the hole-doping case with  $\mu \ll -\Delta_{\text{CI}}$ , we can neglect the electron band to an excellent approximation and focus solely on momenta near the Fermi surface for the hole band. The remaining quasiparticle dispersion  $E_{-, \mathbf{k}}$  then has two ‘branches’ with the same energy—corresponding to excitations above and below the hole-like Fermi surface (i.e., with

$\xi_{-,k} > 0$  and  $\xi_{-,k} < 0$ ). That is, for each momentum  $k^+ > k_F$  ( $k_F$  is the Fermi momentum), there exists a momentum  $k^- < k_F$  such that  $\xi_{-,k^+} = -\xi_{-,k^-}$ , but  $E_{-,k^+} = E_{-,k^-}$ . The momentum-dependent part of the coherence factors therefore cancels, yielding a tunneling current

$$I(eV, \mu) \propto \int d^2\mathbf{k} \left\{ [f(E_{-,k} - eV) - f(E_{-,k})] - [1 - f(-E_{-,k} - eV) - f(E_{-,k})] \right\} \quad (14)$$

that depends on the quasiparticle dispersion but not the coherence factors. Upon taking  $d^2\mathbf{k} \approx k_F dk d\theta$ , carrying out a variable change  $\omega = \sqrt{\xi_{-,k}^2 + \Delta_k}$ , and assuming no  $|\mathbf{k}|$  dependence in the pairing gap evaluated at the Fermi surface [ $\Delta_{\mathbf{k}} \rightarrow \Delta(\theta)$ ], we arrive at the conventional BCS expression:

$$\begin{aligned} I(eV, \mu) &\propto \int_0^{2\pi} d\theta \int_{\Delta(\theta)}^{\infty} d\omega \frac{\omega}{\sqrt{\omega^2 - \Delta(\theta)^2}} \left\{ [f(\omega - eV) - f(\omega)] - [1 - f(-\omega - eV) - f(\omega)] \right\} \\ &\propto \int_0^{2\pi} d\theta \int_{\Delta(\theta)}^{\infty} d\omega \frac{\omega}{\sqrt{\omega^2 - \Delta(\theta)^2}} \left( -\frac{df}{d\omega} eV \right) \\ \implies \frac{dI}{dV} &\propto \int_0^{2\pi} d\theta \int_{\Delta(\theta)}^{\infty} d\omega \frac{\omega}{\sqrt{\omega^2 - \Delta(\theta)^2}} \left( -\frac{df}{d\omega} \right). \end{aligned} \quad (15)$$

Implementing the Dynes substitution<sup>S13</sup>  $\omega \rightarrow \omega + i\Gamma$  then recovers the expression from Eq. (10). The square-root factor in the denominator underlies coherence peaks associated with pairing-induced density-of-states rearrangement.

By contrast, in the BEC phase ( $|\mu| < \Delta_{\text{CI}}$ ), or sufficiently close to the BEC-BCS transition, the simplifying procedure above breaks down. Both electron and hole bands need to be retained;  $\Delta_{\mathbf{k}}$  can not be simply evaluated at a Fermi surface, and hence dependence on the orientation *and* magnitude of  $\mathbf{k}$  become important; and since the minimum of the quasiparticle dispersion  $E_{\pm, \mathbf{k}}$  occurs at or near  $\mathbf{k} = 0$ , the momentum-dependent part of the coherence factors no longer perfectly cancels. Together, these details manifest both through a ‘softening’ of the coherence peaks in the tunneling conductance and the generation of a tunneling gap for *any* pairing function  $\Delta_{\mathbf{k}}$ , *d*-wave or otherwise, in the BEC state; cf. Fig. 3k,l.

Returning to the general current formula in Eq. (13), in simulations of Fig. 3k,l and supplemental simulations below, we employ a *d*-wave pairing potential with

$$\Delta_{\mathbf{k}} = \Delta_0 h(k) \cos(2\theta). \quad (16)$$

Here  $k$  and  $\theta$  are the magnitude and polar angle of  $\mathbf{k}$ , while  $\Delta_0$  sets the pairing energy scale. We take the  $k$ -dependent prefactor to be  $h(k) = \tanh(k^2 \ell^2)$ , where  $\ell$  is roughly the real-space distance over which the *d*-wave pairing potential acts. This choice results in  $\Delta_{\mathbf{k}}$  vanishing at  $k = 0$  as required for *d*-wave pairing, and regularizes the unphysical divergence that would appear with a simple  $h(k) \propto k^2$  profile in a manner that preserves locality in real-space. Let  $\eta \equiv 2m\Delta_0\ell^2$  be a dimensionless quantity involving  $\ell$ . In the regime of the BCS phase with  $k_F\ell \gg 1$ , near the Fermi surface we have  $\Delta_{\mathbf{k}} \approx \Delta_0 \cos(2\theta)$ ; hence the value of  $\eta$  is largely irrelevant provided  $k_F^2/2m$  remains sufficiently large compared to  $\Delta_0$ . In both the BCS regime with  $k_F\ell \lesssim 1$  and throughout the BEC phase, the choice of  $\eta$  is more significant. Here, for the physically important ‘small’ momenta, the pairing behaves like  $\Delta_{\mathbf{k}} \approx \Delta_0 k^2 \ell^2 \cos(2\theta)$  and should be compared to the  $k^2/2m$  kinetic energy scale. With  $\eta \lesssim 1$ , pairing effects are suppressed since the latter scale dominates over the former. By contrast, with  $\eta \gtrsim 1$  the pairing scale dominates and correspondingly

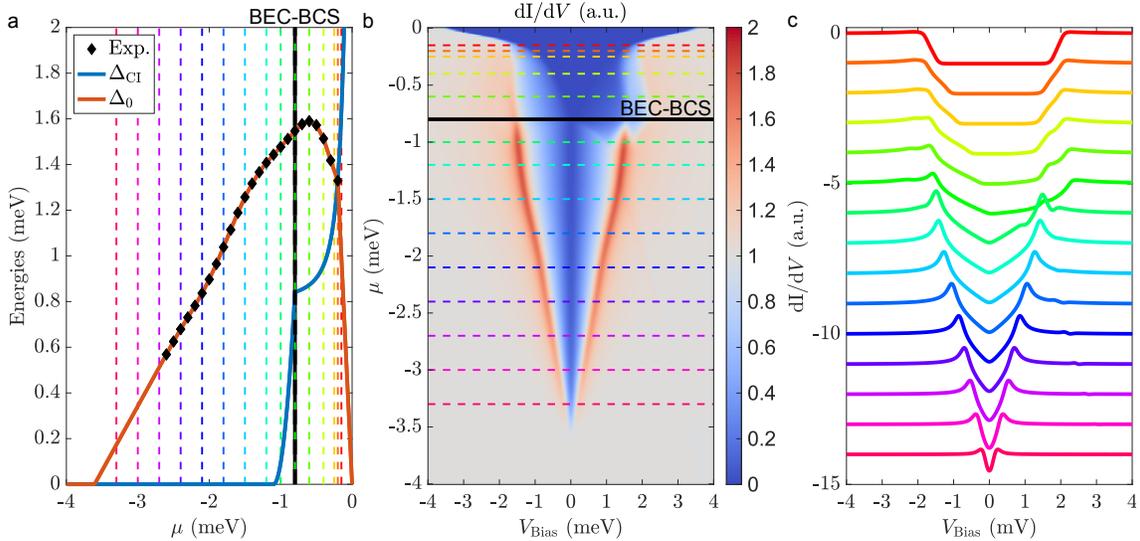
yields more dramatic signatures in density of states and tunneling conductance. In particular, the coherence peaks appear most prominently in the BEC phase at  $\eta \gg 1$ .

The tunneling conductance in the BEC and BCS phases can be studied as a function of chemical potential or as a function of filling. In our formalism, treating  $\mu$  as the tuning parameter is more convenient since all  $\mu$  dependence is contained in the quasiparticle dispersion  $E_{\pm, \mathbf{k}}$  and the relation between filling and  $\mu$  evolves nontrivially between the BEC and BCS phases. In experiment, however, the gate-controlled filling  $\nu$  is the natural tuning parameter. Additionally, the pairing strength and  $\nu = -2$  gap, modeled here by  $\Delta_0$  and  $\Delta_{\text{CI}}$ , certainly depend on  $\nu$ —which further complicates the relation between filling and  $\mu$ . We defer a careful examination of this relation to future work. Instead, here we will simply explore the tunneling conductance as a function of  $\mu$ , with  $\mu$ -dependent  $\Delta_0$  and  $\Delta_{\text{CI}}$  input parameters extracted (crudely) from the experiment as follows.

First, for each filling we fix  $\Delta_0$  to the measured location of coherence peaks in Fig. 3h (and linearly extrapolate to continue to more negative  $\mu$  values). In the V-shaped regime this assignment is expected to be quantitatively reliable, given our interpretation of that regime as a BCS phase (which would indeed have coherence peaks set by  $\Delta_0$ ). However, the U-shaped regime, interpreted as a BEC phase, would have coherence peaks at an energy determined by multiple parameters including  $\mu$ ,  $\Delta_{\text{CI}}$ , and  $\Delta_0$ ; thus here the assignment becomes an approximation that we invoke for simplicity. We then obtain a  $\Delta_0$  vs.  $\mu$  profile by naively replacing filling (or gate voltage) with  $\mu$ ; i.e., we ignore the nontrivial relation linking these quantities. To determine  $\Delta_{\text{CI}}$  vs.  $\mu$ , we first fix the value at  $\mu = 0$  to be  $\Delta_{\text{CI},0} = 2.7$  meV, corresponding to the  $\nu = -2$  spectral gap seen in Extended Data Fig. 4. We also fix the chemical potential  $\mu_*$  corresponding to the BEC-BCS transition, which in our model occurs when  $-\mu_* = \Delta_{\text{CI}}(\mu_*)$ . We specifically set  $\mu_* = -0.8$  meV so that the transition coincides roughly with the experimentally observed U-to-V change in Fig. 3 (after replacing density as  $\mu$  as described above). We phenomenologically model the remaining  $\mu$  dependence of  $\Delta_{\text{CI}}$  as

$$\Delta_{\text{CI}}(\mu) = \begin{cases} \Delta_{\text{CI},0} \frac{\gamma_{\text{CI}}^2}{\mu^2 + \gamma_{\text{CI}}^2} & \mu \geq \mu_* \\ \alpha_2 \mu^2 + \alpha_1 \mu + \alpha_0 & \mu_* \geq \mu \end{cases} \quad (17)$$

with  $\alpha_2 = \Delta_{\text{CI}}(\mu^+)/(\mu_* - \mu_{**})^2$ ,  $\alpha_1 = -2\Delta_{\text{CI}}(\mu^+)\mu_{**}/(\mu_* - \mu_{**})^2$ ,  $\alpha_0 = \Delta_{\text{CI}}(\mu^+)\mu_{**}^2/(\mu_* - \mu_{**})^2$  and  $\mu_{**} = -1.1$  meV. We further choose small enough  $\gamma_{\text{CI}} = 0.1$  meV to ensure coherence peak separation comparable with the experiment. The parametrization above causes  $\Delta_{\text{CI}}$  to decrease upon hole doping and eventually vanish at a chemical potential  $\mu_{**}$  (we fix  $\Delta_{\text{CI}}$  to zero beyond this point rather than allowing it to become negative). This collapse of  $\Delta_{\text{CI}}$  is invoked to emulate experiment;  $\mu$ -independent  $\Delta_{\text{CI}}$  would produce additional structure in the tunneling conductance that is not resolved in measurements. SI Fig. 5a illustrates the resulting  $\mu$  dependence of  $\Delta_0$  and  $\Delta_{\text{CI}}$ .

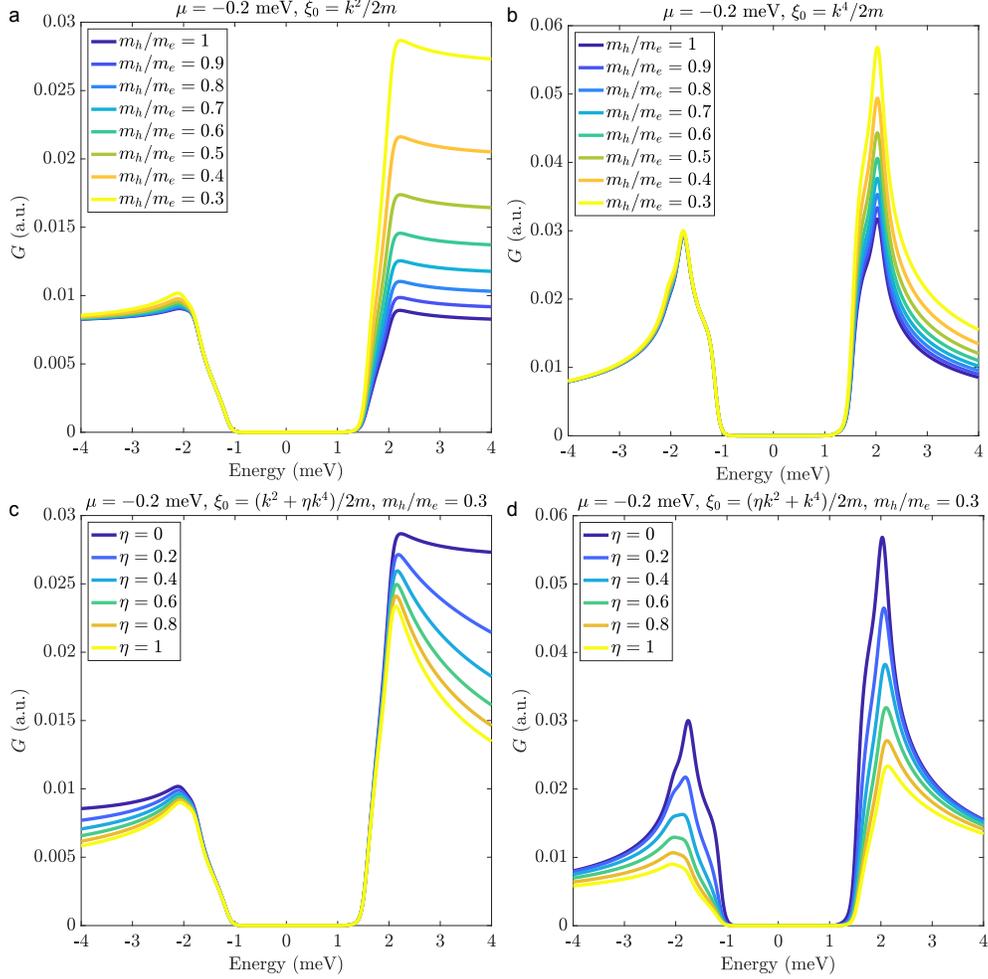


**Supplementary Information Fig. 5. Simulated tunneling conductance across the BEC-BCS transition.** **a**, Chemical potential dependence of  $\Delta_0$  and  $\Delta_{C1}$  used in simulations. Black data points represent coherence-peak locations crudely extracted from experiment, as detailed in the text. **b,c**, Color map and linecuts of differential conductance  $dI/dV$  as a function of  $\mu$ . Here and in Fig. 3k,j, we set  $T = 0.05$  meV and employed a nodal  $d$ -wave gap with  $1/2m\ell^2 = 6.25$   $\mu\text{eV}$ . The BEC-BCS transition manifests as a clear evolution from U- to V-shaped spectra as observed experimentally. We nevertheless stress, as in the text, that panels **b,c** do not correspond directly to Fig. 3a due in part to the nontrivial relation between chemical potential  $\mu$  and filling that has not been incorporated.

Given these parameters, we evaluate the bias voltage and  $\mu$  dependence of the tunneling conductance assuming  $1/2m\ell^2 = 6.25$   $\mu\text{eV}$ , which yields values of  $\eta$  as large as  $\sim 250$ . SI Fig. 5b,c presents tunneling conductance color maps and linecuts; data from Fig. 3k,l were generated from the same parameter set. While we caution against direct comparison of Fig. 3a and SI Fig. 5b given the crude model and parameter extraction used for the latter, our simulations do robustly capture the observed U- to V-shaped evolution. Improved modeling of experiment could be pursued in several ways, e.g., by self-consistently relating  $\mu$  and filling, and by employing more sophisticated band-structure modeling that accounts for density of states features at  $\nu = -2$ . The latter in particular may be required to obtain more refined agreement with experimental details such as the relative coherence peak heights in the U- and V-shaped regimes, see SI Fig. 6.

## 2. Connection to coherence length measurements

Finally, we discuss the behaviour of the Ginzburg-Landau coherence length  $\xi_{GL}$  in the proposed BEC-BCS transition scenario. The primary intent of this analysis is to emphasize that this scenario is consistent with the transport-based observations of Ref. S15, which found that  $\xi_{GL}$  admits two distinct regimes. First, in the part of the superconducting dome with  $\nu \lesssim -2.5$ —roughly our V-shaped region— $\xi_{GL}$  significantly exceeds the inter-particle spacing  $d = 1/\sqrt{|\delta n|}$  (where  $\delta n$  is measured relative to  $\nu = -2$ ). In this regime, the coherence length can be well captured by a standard form  $\xi_{GL} = cv_F/\Delta$  expected from dimensional analysis in a BCS phase, where  $v_F$  is the Fermi velocity,  $\Delta$  is the characteristic pairing energy, and  $c$  is a (presumably order-one) con-



**Supplementary Information Fig. 6. Modeling of coherence peak asymmetry in the BEC regime.**

Tunneling conductance expected in the BEC regime using two generalizations of the toy-model introduced in the manuscript. As in the manuscript two bands are considered, an electron- and hole-like, each with a different carrier dispersion. In panels **a** and **b** we consider two power-law dispersions with variable masses. In panels **c** and **d** we keep mass constant and instead vary the dispersion as indicted in the legend. An asymmetry of the coherence peak height can be obtained by varying quasiparticle mass and coherence peak sharpness depends on the details of the dispersion. We note that in moiré systems, the bandstructure<sup>S4,S5,S9</sup> is highly filling dependent and deviates strongly from a simple  $\propto k^2$  dispersion. For all of the considered models in the BCS regime a V-like profile signaling presence of gapless excitations is always present confirming the robustness of the U-shape to V-shape transition.

stant. Using  $v_F \sim 10^5$  m/s (comparable to the flat-band velocity extracted from previous MATBG measurements<sup>S17</sup>), our measured spectroscopic gaps  $\Delta$  (see above in section 4), and  $c \approx 2/3$  indeed yields coherence lengths that quantitatively agree with Ref. S15 over this filling range. For example, our measured  $\Delta$  at  $\nu = -2.5$  yields  $\xi_{GL} \approx 30$  nm. This agreement supports the emergence of a ‘BCS’ regime—albeit of a strongly coupled nature as confirmed by the anomalously large  $2\Delta/(k_B T_C)$  ratio reported in the main text.

By contrast, in the complementary part of the superconducting dome with  $\nu \gtrsim -2.5$ —

coinciding roughly with our U-shaped region—Ref. S15 measured  $\xi_{\text{GL}}$  values that closely track the relative inter-particle spacing  $d$  and become as small as  $\sim 12$  nm. The deviation from the form  $\xi_{\text{GL}} \propto v_F/\Delta$  can be accounted for by the presence of an additional energy scale, the gap for dissociating the Cooper-pair molecules, as well as the fact that  $v_F$  has no meaningful definition in the absence of a Fermi surface. Instead, the scaling relation  $\xi_{\text{GL}} \propto d$  is predicted for a BEC regime in related contexts<sup>S18–S20</sup>, and we briefly sketch how the pertinent scaling may be obtained using the results of Ref. S18. We emphasize, however, that direct use of this reference requires a number of simplifying assumptions that limit the scope and applicability of the analysis. Although the arguments outlined in the previous subsection hinge on the assumption of a nodal order parameter, we specialize here to nodeless  $s$ -wave pairing. Nevertheless, because the BEC phase is gapped regardless of the function form of the gap, we do not expect this distinction to alter the functional relationship of  $\xi_{\text{GL}}$  vis-à-vis the interparticle spacing  $d = 1/\sqrt{|\delta n|}$ . We also restrict our attention to the hole band,  $\xi_{-,k}$ , which can be viewed as taking the  $\Delta_{\text{CI}} \rightarrow \infty$  limit in the model presented in the previous subsection. For convenience, we drop the subscript ‘-’ as well as the reference to  $\Delta_{\text{CI}}$ , simply expressing the dispersion as  $\xi_{\mathbf{k}} \equiv \xi_k = -k^2/(2m) - \mu$ , where  $k$  is the magnitude of the vector  $\mathbf{k}$ . It follows that  $\mu > 0$  corresponds to the BEC regime, while  $\mu < 0$  is the BCS regime (which we do not consider here). As in the previous subsection, details of the symmetry breaking leading to the  $\nu = -2$  insulator are neglected, and a generic two-fold ‘spin’ symmetry with quantum numbers labelled by  $a = 1, 2$  is assumed to remain. A filling  $\delta n$  of the hole bands corresponds to a filling  $\nu = -2 + \delta n$  of the TTG system with  $\delta n < 0$ .

We start with a Hamiltonian

$$H = \sum_{\mathbf{k}, a} c_a^\dagger(\mathbf{k}) \xi_{\mathbf{k}} c_a(\mathbf{k}) + \sum_{\mathbf{k}, \mathbf{k}', \mathbf{q}} U c_1^\dagger(\mathbf{k} + \mathbf{q}/2) c_2^\dagger(-\mathbf{k} + \mathbf{q}/2) c_2(-\mathbf{k}' + \mathbf{q}/2) c_1(\mathbf{k}' + \mathbf{q}/2), \quad (18)$$

where  $U$  characterizes the interaction strength and  $c_{a=1,2}(\mathbf{k})$  are electron annihilation operators. The superconducting gap  $\Delta$  that develops should be obtained from  $H$  via a self-consistent equation, but for simplicity, we instead consider  $\Delta$  as a constant, implying a superconducting spectrum given by  $E_k = \sqrt{\xi_k^2 + \Delta^2}$ . The macroscopically based coherence length  $\xi_{\text{GL}}$  is proportional to the microscopically derived  $\xi_{\text{phase}}$ , which is identified with the inverse mass of the canonical boson  $\phi(\mathbf{r}) \sim c_1(\mathbf{r})c_2(\mathbf{r})$  in the effective action determined in Ref. S18. They find that  $\xi_{\text{phase}} = \sqrt{b/a}$  where

$$a = \frac{\Delta^2}{4\pi} \int_0^\infty dk k \frac{1}{E_k^3}, \quad b = \frac{1}{32\pi m} \int_0^\infty dk k \frac{\xi_k^2}{E_k^5} \left[ -\frac{\xi_k^2 - 2\Delta^2}{\xi_k} + \frac{5\Delta^2 k^2}{2m E_k^2} \right]. \quad (19)$$

The model is analytically tractable, returning

$$\xi_{\text{phase}} = \sqrt{\frac{1}{12m} \frac{1}{x - \mu} \left( \frac{\mu^2}{x^2} + \frac{x}{x + \mu} \right)}, \quad x = \sqrt{\mu^2 + \Delta^2}. \quad (20)$$

This expression is explicitly a function of  $\mu$  and not of the density  $\delta n$  of the bands. We relate the two via

$$\delta n = -\frac{1}{2\pi} \int_0^\infty dk k \left( 1 + \frac{\xi_k}{E_k} \right), \quad (21)$$

which can be solved and inverted to obtain  $\mu$  as a function of  $\delta n$ :

$$\mu = \frac{(2\pi\delta n/m)^2 - \Delta^2}{4\pi\delta n/m}. \quad (22)$$

Deep in the BEC regime with  $\delta n \rightarrow 0^-$ , we find

$$\xi_{\text{phase}} \xrightarrow{\delta n \rightarrow 0^-} \frac{1}{4\sqrt{-\pi\delta n}} \propto d, \quad (23)$$

consistent with the observations of Ref. [S15](#). Hence, when comparing with experiment,  $\xi_{\text{phase}}$  has the same functional dependence on  $d = 1/\sqrt{|\delta n|}$  in the BEC regime. Again, we emphasize that while the coefficient may differ, we do not expect the presence of nodes in the superconducting order parameter to alter our conclusions in this limit.

We now turn to the intermediate regime between the BCS and BEC limits. Based on transport measurements, Ref. [S15](#) proposed that MATTG can be tuned close to the BEC-BCS crossover (see also Ref. [S21](#)). We advocate for a complementary scenario, wherein the presence of gapless modes in the BCS regime implies that the system undergoes a BEC to BCS *phase transition*. This distinction was explicitly emphasized in Refs. [S22](#) in the context of the cuprates, and the corresponding transition was also explored in Refs. [S20](#) and [S23](#). The prospect of a gate-tuned transition within the superconducting dome is especially encouraging since it may be consistent with the apparent discontinuity in the coherence length measured in Ref. [S15](#). We leave the determination of the coherence length across the transition for future work.

- 
- [S1] Carr, S. *et al.* Ultraheavy and Ultrarelativistic Dirac Quasiparticles in Sandwiched Graphenes. *Nano Letters* **20**, 3030–3038 (2020).
  - [S2] Khalaf, E., Kruchkov, A. J., Tarnopolsky, G. & Vishwanath, A. Magic angle hierarchy in twisted graphene multilayers. *Physical Review B* **100**, 085109 (2019).
  - [S3] Bistritzer, R. & MacDonald, A. H. Moiré bands in twisted double-layer graphene. *Proceedings of the National Academy of Sciences* **108**, 12233–12237 (2011).
  - [S4] Choi, Y. *et al.* Interaction-driven band flattening and correlated phases in twisted bilayer graphene. *Nature Physics* **17**, 1375–1381 (2021).
  - [S5] 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).
  - [S6] 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).
  - [S7] 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). 2004.14784.
  - [S8] Calderón, M. J. & Bascones, E. Interactions in the 8-orbital model for twisted bilayer graphene. *Physical Review B* **102**, 155149 (2020). 2007.16051.
  - [S9] Fischer, A. *et al.* Unconventional superconductivity in magic-angle twisted trilayer graphene. *npj Quantum Materials* **7**, 1–10 (2022).
  - [S10] Phong, V. T., Pantaleón, P. A., Cea, T. & Guinea, F. Band structure and superconductivity in twisted trilayer graphene. *Physical Review B* **104**, L121116 (2021).
  - [S11] 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).
  - [S12] Zasadzinski, J. F., Coffey, L., Romano, P. & Yusof, Z. Tunneling spectroscopy of Bi2Sr2CaCu2O8+ $\delta$  Eliashberg analysis of the spectral dip feature. *Physical Review B* **68**, 180504 (2003).

- [S13] Dynes, R. C., Narayanamurti, V. & Garno, J. P. Direct Measurement of Quasiparticle-Lifetime Broadening in a Strong-Coupled Superconductor. *Physical Review Letters* **41**, 1509–1512 (1978).
- [S14] Oh, M. *et al.* Evidence for unconventional superconductivity in twisted bilayer graphene. *Nature* **600**, 240–245 (2021).
- [S15] Park, J. M., Cao, Y., Watanabe, K., Taniguchi, T. & Jarillo-Herrero, P. Tunable strongly coupled superconductivity in magic-angle twisted trilayer graphene. *Nature* **590**, 249–255 (2021).
- [S16] Yeh, N.-C. *et al.* Evidence of Doping-Dependent Pairing Symmetry in Cuprate Superconductors. *Physical Review Letters* **87**, 087003 (2001).
- [S17] Choi, Y. *et al.* Correlation-driven topological phases in magic-angle twisted bilayer graphene. *Nature* **589**, 536–541 (2021).
- [S18] Pistolesi, F. & Strinati, G. C. Evolution from BCS superconductivity to Bose condensation: Calculation of the zero-temperature phase coherence length. *Physical Review B* **53**, 15168–15192 (1996).
- [S19] Randeria, M. & Taylor, E. Crossover from Bardeen-Cooper-Schrieffer to Bose-Einstein Condensation and the Unitary Fermi Gas. *Annual Review of Condensed Matter Physics* **5**, 209–232 (2014).
- [S20] Stintzing, S. & Zwerger, W. Ginzburg-Landau theory of superconductors with short coherence length. *Physical Review B* **56**, 9004–9014 (1997).
- [S21] Hao, Z. *et al.* Electric field-tunable superconductivity in alternating-twist magic-angle trilayer graphene. *Science* **371**, 1133–1138 (2021).
- [S22] Botelho, S. S. & Sá de Melo, C. A. R. Lifshitz transition in  $d$ -wave superconductors. *Physical Review B* **71**, 134507 (2005).
- [S23] Cao, G., He, L. & Zhuang, P. BCS-BEC quantum phase transition and collective excitations in two-dimensional Fermi gases with  $p$ - and  $d$ -wave pairings. *Physical Review A* **87**, 013613 (2013).