

# Supplementary Material for "Twisted Bilayer Graphene: A Phonon Driven Superconductor"

Biao Lian,<sup>1</sup> Zhijun Wang,<sup>2</sup> and B. Andrei Bernevig<sup>2,3,4,5</sup>

<sup>1</sup>*Princeton Center for Theoretical Science, Princeton University, Princeton, New Jersey 08544, USA*

<sup>2</sup>*Department of Physics, Princeton University, Princeton, New Jersey 08544, USA*

<sup>3</sup>*Dahlem Center for Complex Quantum Systems and Fachbereich Physik,*

*Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany*

<sup>4</sup>*Max Planck Institute of Microstructure Physics, 06120 Halle, Germany*

<sup>5</sup>*Sorbonne Universités, UPMC Univ Paris 06, UMR 7589, LPTHE, F-75005, Paris, France*

(Dated: June 5, 2019)

## S1. DERIVATION OF ELECTRON-PHONON COUPLING AT MAGIC ANGLE

In this section we first briefly recall the continuum model in momentum space formulated in Ref. [1], and then use it to derive the coupling between electrons and interlayer phonons is derived.

Fig. S1 illustrates the real space configuration of two graphene layers which have a relative twist angle  $\theta$ . We denote the lattice vectors of layer  $j$  ( $j = 1, 2$ ) as  $\mathbf{a}_1^{(j)}$  and  $\mathbf{a}_2^{(j)}$ , and the reciprocal lattice vectors of layer  $j$  as  $\mathbf{G}_1^{(j)}$  and  $\mathbf{G}_2^{(j)}$ , which satisfy  $\mathbf{G}_a^{(j)} \cdot \mathbf{a}_b^{(j)} = 2\pi\delta_{ab}$ . The norms of the lattice vectors  $\mathbf{a}_i^{(j)}$  are equal to the lattice constant  $a_0 = 0.246\text{nm}$ . Each layer  $j$  consists of two sublattice positions  $A$  and  $B$ , which are located at  $\boldsymbol{\tau}_A^{(j)}$  and  $\boldsymbol{\tau}_B^{(j)}$  in the unit cell of layer  $j$ . Without loss of generality, we can choose sublattice  $A$  in layer  $j$  as the origin of the unit cell of layer  $j$ , so that  $\boldsymbol{\tau}_A^{(j)} = 0$ , and  $\boldsymbol{\tau}_B^{(j)} = \boldsymbol{\tau}^{(j)}$  as shown in Fig. S1. The vectors  $\mathbf{a}_1^{(j)}$  and  $\mathbf{a}_2^{(j)}$  are rotated by  $\theta$  from one another, and so do  $\boldsymbol{\tau}^{(1)}$  and  $\boldsymbol{\tau}^{(2)}$ . More explicitly, the above vectors in components are given by

$$\begin{aligned} \mathbf{a}_1^{(j)} &= a_0 R_{\theta/2}^{(j)} \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}^T, \quad \mathbf{a}_2^{(j)} = a_0 R_{\theta/2}^{(j)} \begin{pmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix}^T, \quad \boldsymbol{\tau}_A^{(j)} = (0, 0)^T, \quad \boldsymbol{\tau}_B^{(j)} = \frac{a_0}{\sqrt{3}} R_{\theta/2}^{(j)} (0, 1)^T, \\ \mathbf{G}_1^{(j)} &= \frac{4\pi}{\sqrt{3}a_0} R_{\theta/2}^{(j)} \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}^T, \quad \mathbf{G}_2^{(j)} = \frac{4\pi}{\sqrt{3}a_0} R_{\theta/2}^{(j)} \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}^T, \end{aligned} \quad (\text{S1})$$

for layer  $j = 1, 2$ , where

$$R_{\theta/2}^{(1)} = \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{pmatrix}, \quad R_{\theta/2}^{(2)} = \begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}, \quad (\text{S2})$$

which are the rotation matrices of angle  $\pm\theta/2$ , respectively.

The interlayer hopping  $t(r)$  between two atoms in different layers is generically a function of their in-plane distance  $r$ . When transformed into momentum space, the electron hopping from momentum  $\mathbf{p}'$  and sublattice  $\beta$  in layer 2 to momentum  $\mathbf{k}$  and sublattice  $\alpha$  in layer 1 takes the form

$$\begin{aligned} T_{\mathbf{k}\mathbf{p}'}^{\alpha\beta} &= \frac{1}{N} \sum_{\mathbf{R}^{(1)}, \mathbf{R}^{(2)}} t \left( \mathbf{R}^{(1)} + \boldsymbol{\tau}_\alpha^{(1)} - \mathbf{R}^{(2)} - \boldsymbol{\tau}_\beta^{(2)} \right) e^{i\mathbf{p}' \cdot (\mathbf{R}^{(2)} + \boldsymbol{\tau}_\beta^{(2)}) - i\mathbf{k} \cdot (\mathbf{R}^{(1)} + \boldsymbol{\tau}_\alpha^{(1)})} \\ &= \sum_{n_1, m_1} \sum_{n_2, m_2} \frac{t_{\mathbf{k} + \mathbf{G}_{n_1, m_1}^{(1)}}}{\Omega} \delta_{\mathbf{k} + \mathbf{G}_{n_1, m_1}^{(1)}, \mathbf{p}' + \mathbf{G}_{n_2, m_2}^{(2)}} e^{i\mathbf{G}_{n_2, m_2}^{(2)} \cdot \boldsymbol{\tau}_\beta^{(2)} - i\mathbf{G}_{n_1, m_1}^{(1)} \cdot \boldsymbol{\tau}_\alpha^{(1)}}, \end{aligned} \quad (\text{S3})$$

where  $\alpha, \beta = A, B$  are sublattice indices,  $\mathbf{k}$  and  $\mathbf{p}'$  are measured from  $\Gamma$  point of the graphene Brillouin zone (BZ),  $\Omega = \sqrt{3}a_0^2/2$  is the area of graphene unit cell,  $t_{\mathbf{k}}$  is the Fourier transform of  $t(\mathbf{r})$ , and  $\mathbf{G}_{n, m}^{(j)} = n\mathbf{G}_1^{(j)} + m\mathbf{G}_2^{(j)}$  runs over all  $n, m \in \mathbb{Z}$  reciprocal lattices in layer  $j$ .

We first consider the low energy physics near the  $K$  points of the graphene BZs of the two layers (see main text Fig. 2), namely,  $\mathbf{k}$  and  $\mathbf{p}'$  near the Dirac point momenta  $\mathbf{K}_D^{(1)} = -(\mathbf{G}_1^{(1)} + \mathbf{G}_2^{(1)})/3$  and  $\mathbf{K}_D^{(2)} = -(\mathbf{G}_1^{(2)} + \mathbf{G}_2^{(2)})/3$ , respectively. It is shown that the hopping  $t_{\mathbf{k}}$  decays exponentially with respect to  $|\mathbf{k}|$  [1], so a good approximation is to keep only the 3 leading nearest hopping terms  $t_{\mathbf{k} + \mathbf{G}_{n, m}^{(1)}}$  with  $|\mathbf{k} + \mathbf{G}_{n, m}^{(1)}|$  around the magnitude  $|\mathbf{K}_D^{(1)}|$ , and approximate them to  $t_{\mathbf{K}_D^{(1)}}$ .

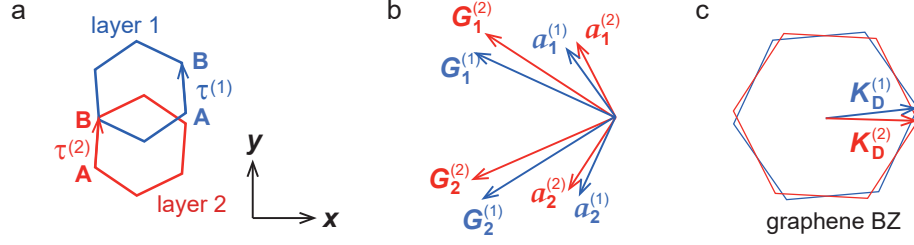


FIG. S1: **a.** Illustration of the real space TBG lattice, where the twist angle is  $\theta = 10^\circ$ . **b.** The graphene lattice vectors and reciprocal vectors in each layer. **c.** The graphene BZs of the two layers.

We then define the three vectors

$$\mathbf{q}_1 = \mathbf{K}_D^{(2)} - \mathbf{K}_D^{(1)}, \quad \mathbf{q}_2 = C_{3z}\mathbf{q}_1 = \mathbf{K}_D^{(2)} + \mathbf{G}_1^{(2)} - \mathbf{K}_D^{(1)} - \mathbf{G}_1^{(1)}, \quad \mathbf{q}_3 = C_{3z}^2\mathbf{q}_1 = \mathbf{K}_D^{(2)} + \mathbf{G}_2^{(2)} - \mathbf{K}_D^{(1)} - \mathbf{G}_2^{(1)},$$

where  $C_{3z}$  is the 3-fold rotation about  $z$  axis. Explicitly,  $\mathbf{q}_j$  in components are

$$\mathbf{q}_1 = k_\theta(0, -1)^T, \quad \mathbf{q}_2 = k_\theta\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right)^T, \quad \mathbf{q}_3 = k_\theta\left(-\frac{\sqrt{3}}{2}, \frac{1}{2}\right)^T,$$

where  $k_\theta = |\mathbf{q}_j| = (8\pi/3a_0)\sin(\theta/2)$ . Under the above nearest hopping approximation, an electron state with momentum  $\mathbf{p}'$  in layer 2 can hop to an electron state with momentum  $\mathbf{k}$  in layer 1 if  $\mathbf{k} - \mathbf{p}' = \mathbf{q}_j$ .

Since we are interested in the low energy band structure near the graphene BZ  $K$  (or  $K'$ ) point, hereafter we set the origin of  $\mathbf{k}$  and  $\mathbf{p}'$  at  $K$  (or  $K'$ ) point of each layer. The effective Hamiltonian of each layer is given by  $h^K(\mathbf{k}) = v(k_x\sigma_x - k_y\sigma_y) = \hbar v\boldsymbol{\sigma}^* \cdot \mathbf{k}$  at  $K$  point of each layer, and  $h^{K'}(\mathbf{k}) = -\hbar v\boldsymbol{\sigma} \cdot \mathbf{k}$  at  $K'$  point of each layer (related to  $h^K(\mathbf{k})$  via the time reversal transformation  $\mathcal{T}$ ), where  $v$  is the fermi velocity, and  $\sigma_{x,y,z}$  are the Pauli matrices for sublattice indices. Note that the Dirac fermions at graphene valleys  $K$  and  $K'$  have opposite helicities. Therefore, to the lowest order, the TBG Hamiltonian at  $K$  point in the vicinity of layer 1 momentum  $\mathbf{k} = 0$  (which is the  $K'_M$  point of the Moiré BZ) truncated at the nearest hoppings is [1]

$$H^{K,K'_M}(\mathbf{k}) = \begin{pmatrix} h_{\theta/2}^K(\mathbf{k}) & wT_1 & wT_2 & wT_3 \\ wT_1^\dagger & h_{-\theta/2}^K(\mathbf{k} - \mathbf{q}_1) & 0 & 0 \\ wT_2^\dagger & 0 & h_{-\theta/2}^K(\mathbf{k} - \mathbf{q}_2) & 0 \\ wT_3^\dagger & 0 & 0 & h_{-\theta/2}^K(\mathbf{k} - \mathbf{q}_3) \end{pmatrix}, \quad (\text{S4})$$

where the basis is  $(\psi_{0,\mathbf{k}}^T, \psi_{1,\mathbf{k}}^T, \psi_{2,\mathbf{k}}^T, \psi_{3,\mathbf{k}}^T)^T$ , and  $\psi_{0,\mathbf{k}}$  and  $\psi_{j,\mathbf{k}}$  ( $j = 1, 2, 3$ ) are the 2-component column spinors in the AB sublattice index basis at momentum  $\mathbf{k}$  in layer 1 and momentum  $\mathbf{k} - \mathbf{q}_j$  in layer 2, respectively.  $h_{\pm\theta/2}^K(\mathbf{k})$  is  $h^K(\mathbf{k})$  rotated by a  $\pm\theta/2$  angle, and  $w = t_{\mathbf{K}_D^{(1)}}/\Omega$  is the interlayer hopping which can be chosen as real. The hopping matrix  $T_j$  are defined by  $(T_j)_{\alpha\beta} = e^{i\mathbf{G}_{n_j,m_j}^{(2)} \cdot \boldsymbol{\tau}_\beta^{(2)} - i\mathbf{G}_{n_j,m_j}^{(1)} \cdot \boldsymbol{\tau}_\alpha^{(1)}}$  (which is nothing but the phase factor part of Eq. (S3)), where  $(n_1, m_1) = (0, 0)$ ,  $(n_2, m_2) = (1, 0)$  and  $(n_3, m_3) = (0, 1)$ , and  $\mathbf{G}_{n,m}^{(i)}$  in layer  $i$  is defined below Eq. (S3). They satisfy  $|\mathbf{K}_D^{(i)} + \mathbf{G}_{n_j,m_j}^{(i)}| = |\mathbf{K}_D^{(i)}| = 4\pi/3a_0$ , thus correspond to the nearest interlayer hoppings in the momentum space. Explicitly,  $T_j$  are given by

$$T_1 = 1 + \sigma_x, \quad T_2 = 1 - \frac{1}{2}\sigma_x - \frac{\sqrt{3}}{2}\sigma_y, \quad T_3 = 1 - \frac{1}{2}\sigma_x + \frac{\sqrt{3}}{2}\sigma_y.$$

Since the twist angle  $\theta$  is small, to the lowest order, we shall ignore the  $\pm\theta/2$  rotation of  $h_{\pm\theta/2}^K(\mathbf{k})$  in Eq. (S4). Under this approximation, the system has a particle-hole symmetry, and the low energy eigenstates are given by  $\psi_{j,\mathbf{k}} \approx -wh_j^{-1}T_j^\dagger\psi_{0,\mathbf{k}}$  ( $j = 1, 2, 3$ ), where  $h_j$  is short hand for  $h^K(-\mathbf{q}_j)$ . The low energy Hamiltonian is a  $2 \times 2$  Hamiltonian in the  $\psi_0$  space [1]:

$$\tilde{H}^{K,K'_M}(\mathbf{k}) = \frac{\langle \Psi | H^{K,K'_M}(\mathbf{k}) | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\hbar v}{1 + 6\alpha^2} \left( \boldsymbol{\sigma}^* \cdot \mathbf{k} + w^2 \sum_{j=1}^3 T_j h_j^{-1} \boldsymbol{\sigma}^* \cdot \mathbf{k} h_j^{-1} T_j^\dagger \right) = \frac{1 - 3\alpha^2}{1 + 6\alpha^2} \hbar v \boldsymbol{\sigma}^* \cdot \mathbf{k}, \quad (\text{S5})$$

where  $\alpha = w/\hbar vk_\theta$ . The first magic angle  $\theta = 1.05^\circ$  is given by  $\alpha^2 = 1/3$ , where the Dirac velocity of the above effective Hamiltonian vanishes. Note that  $\tilde{H}^{K,K'_M}(\mathbf{k})$  depends on the momenta  $\mathbf{q}_1, \mathbf{q}_2$  and  $\mathbf{q}_3$  (which implicitly appear in  $h_j$ ), so we shall also denote it as  $\tilde{H}^{K,K'_M}(\mathbf{k}, \mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$  later to emphasize its dependence on  $\mathbf{q}_j$ .

We note that the Dirac Hamiltonian  $\tilde{H}^{K,K'_M}(\mathbf{k})$  in Eq. (S5) around layer 1 momentum  $\mathbf{k} = 0$  is at the  $K'_M$  point of the Moiré BZ (MBZ) (see Fig. 2 of the main text), and comes from the  $K$  point of both layer 1 and layer 2 of TBG. Similarly, there is also a Dirac Hamiltonian around layer 2 momentum  $\mathbf{p}' = 0$  coming from the  $K$  point of both layer 1 and layer 2 of TBG, which is at the  $K_M$  point of the MBZ. In total, there are eight Dirac fermions labeled by graphene BZ valley  $K, K'$  (which is the same for both layers, no coupling between  $K$  and  $K'$  exists in the Hamiltonian of Eq. (S5)), Moiré valley  $K_M, K'_M$  and spin  $\uparrow, \downarrow$  indices. The band structure obtained in the above model is thus 4-fold degenerate everywhere in the MBZ with respect to graphene valley  $K, K'$  and spin  $\uparrow, \downarrow$ , and has 4 Dirac fermions at Moiré valley  $K_M$  and another 4 Dirac fermions at Moiré valley  $K'_M$ . In particular, the helicity of the Dirac fermions only depends on whether they come from graphene valley  $K$  or  $K'$ .

When small distortions are added to the graphene lattices, the above effective Hamiltonian will change. Since distortions can be expressed using phonon fields, the change in the effective Hamiltonian simply gives the low energy electron-phonon coupling term. The leading contribution to the change of Hamiltonian is due to the distortion of momentum vectors  $\mathbf{q}_j$ . In the main text we have argued that interlayer phonon waves change the Moiré pattern dramatically. To prove this from microscopics requires some more calculations. This can be shown explicitly as follows. Denote the in-plane displacement of atoms of layer  $j$  at  $\mathbf{r}$  as  $\mathbf{u}^{(j)}(\mathbf{r}) = \left( u_x^{(j)}(\mathbf{r}), u_y^{(j)}(\mathbf{r}) \right)^T$ . The displacement  $\mathbf{u}^{(j)}$  is nothing but the in-plane phonon field in layer  $j$ . In the continuum limit, the variation of the lattice vectors  $\mathbf{a}_1^{(j)}$  and  $\mathbf{a}_2^{(j)}$  under the displacement field  $\mathbf{u}^{(j)}$  are simply given by  $\delta \mathbf{a}_1^{(j)}(\mathbf{r}) = (\mathbf{a}_1^{(j)} \cdot \nabla) \mathbf{u}^{(j)}(\mathbf{r})$  and  $\delta \mathbf{a}_2^{(j)}(\mathbf{r}) = (\mathbf{a}_2^{(j)} \cdot \nabla) \mathbf{u}^{(j)}(\mathbf{r})$ . Explicitly, one has

$$\delta \mathbf{a}_b^{(j)} = (\mathbf{a}_b^{(j)} \cdot \nabla) \mathbf{u}^{(j)} = \left[ \mathbf{a}_b^{(j)T} \cdot \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} \right]^T \begin{pmatrix} u_x^{(j)} \\ u_y^{(j)} \end{pmatrix} = a_0 \begin{pmatrix} \partial_x u_x^{(j)} & \partial_y u_x^{(j)} \\ \partial_x u_y^{(j)} & \partial_y u_y^{(j)} \end{pmatrix} R_{\theta/2}^{(j)} \begin{pmatrix} -\frac{1}{2} \\ \pm \frac{\sqrt{3}}{2} \end{pmatrix}, \quad (\text{S6})$$

where the  $\pm$  signs are for  $b = 1$  (lattice vector  $\mathbf{a}_1^{(j)}$ ) and  $b = 2$  (lattice vector  $\mathbf{a}_2^{(j)}$ ), respectively. Accordingly, to linear order the distortion of reciprocal vectors of layer  $j$  satisfies  $\delta \mathbf{G}_a^{(j)} \cdot \mathbf{a}_b^{(j)} + \mathbf{G}_a^{(j)} \cdot \delta \mathbf{a}_b^{(j)} = 0$  ( $a, b = 1, 2$ ), which has a solution

$$\delta \mathbf{G}_a^{(j)} = -\nabla \left( \mathbf{u}^{(j)} \cdot \mathbf{G}_a^{(j)} \right) = -\begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} \left( \mathbf{u}^{(j)} \cdot \mathbf{G}_a^{(j)} \right) = -\frac{4\pi}{\sqrt{3}a_0} \begin{pmatrix} \partial_x u_x^{(j)} & \partial_x u_y^{(j)} \\ \partial_y u_x^{(j)} & \partial_y u_y^{(j)} \end{pmatrix} R_{\theta/2}^{(j)} \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ \pm \frac{1}{2} \end{pmatrix},$$

where the  $\pm$  signs are for index  $a = 1$  (reciprocal vector  $\mathbf{G}_1^{(j)}$ ) and  $a = 2$  (reciprocal vector  $\mathbf{G}_2^{(j)}$ ), respectively. In particular, since  $\boldsymbol{\tau}^{(j)} = (\mathbf{a}_1^{(j)} - \mathbf{a}_2^{(j)})/3$ , one has  $\delta \mathbf{G}_a^{(j)} \cdot \boldsymbol{\tau}^{(j)} + \mathbf{G}_a^{(j)} \cdot \delta \boldsymbol{\tau}^{(j)} = 0$ , which implies  $\mathbf{G}_{1,2}^{(j)} \cdot \boldsymbol{\tau}^{(j)} = \pm 2\pi/3$  remains invariant, so the interlayer hopping matrix  $T_j = e^{i\mathbf{G}_{n_j, m_j}^{(2)} \cdot \boldsymbol{\tau}_\beta^{(2)} - i\mathbf{G}_{n_j, m_j}^{(1)} \cdot \boldsymbol{\tau}_\alpha^{(1)}}$  remains unchanged under the deformation. Only  $\mathbf{q}_j$  in the continuum model Hamiltonian change.

With the expressions for  $\delta \mathbf{G}_a^{(j)}$ , it is straightforward to derive the variation of  $\mathbf{q}_j$  from their definitions:

$$\begin{aligned} \delta \mathbf{q}_1 &= \frac{\delta \mathbf{G}_1^{(1)} + \delta \mathbf{G}_2^{(1)} - \delta \mathbf{G}_1^{(2)} - \delta \mathbf{G}_2^{(2)}}{3} = \frac{4\pi}{3a_0} \left[ \begin{pmatrix} \partial_x u_x^{(1)} & \partial_x u_y^{(1)} \\ \partial_y u_x^{(1)} & \partial_y u_y^{(1)} \end{pmatrix} \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix} - \begin{pmatrix} \partial_x u_x^{(2)} & \partial_x u_y^{(2)} \\ \partial_y u_x^{(2)} & \partial_y u_y^{(2)} \end{pmatrix} \begin{pmatrix} \cos \frac{\theta}{2} \\ -\sin \frac{\theta}{2} \end{pmatrix} \right] \\ &= k_\theta \left( \gamma \partial_x u_x + \partial_x u_y^c, \gamma \partial_y u_x + \partial_y u_y^c \right)^T, \\ \delta \mathbf{q}_2 &= \frac{-2\delta \mathbf{G}_1^{(1)} + \delta \mathbf{G}_2^{(1)} + 2\delta \mathbf{G}_1^{(2)} - \delta \mathbf{G}_2^{(2)}}{3} \\ &= k_\theta \left[ \frac{\sqrt{3}}{2} (\gamma \partial_x u_y - \partial_x u_x^c) - \frac{1}{2} (\gamma \partial_x u_x + \partial_x u_y^c), -\frac{1}{2} (\gamma \partial_y u_x + \partial_y u_y^c) + \frac{\sqrt{3}}{2} (\gamma \partial_y u_y - \partial_y u_x^c) \right]^T, \\ \delta \mathbf{q}_3 &= \frac{\delta \mathbf{G}_1^{(1)} - 2\delta \mathbf{G}_2^{(1)} - \delta \mathbf{G}_1^{(2)} + 2\delta \mathbf{G}_2^{(2)}}{3} \\ &= k_\theta \left[ -\frac{\sqrt{3}}{2} (\gamma \partial_x u_y - \partial_x u_x^c) - \frac{1}{2} (\gamma \partial_x u_x + \partial_x u_y^c), -\frac{1}{2} (\gamma \partial_y u_x + \partial_y u_y^c) - \frac{\sqrt{3}}{2} (\gamma \partial_y u_y - \partial_y u_x^c) \right]^T, \end{aligned} \quad (\text{S7})$$

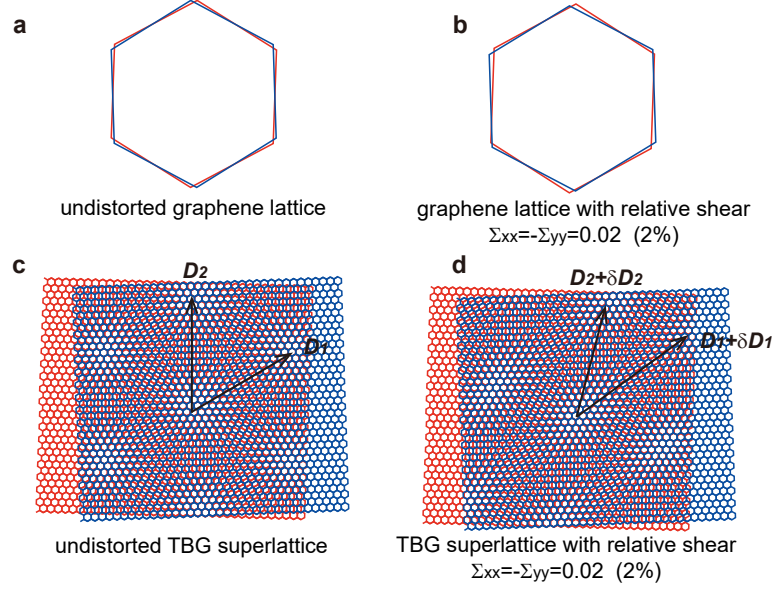


FIG. S2: **a.** Graphene lattice plaquettes of two layers of TBG (at AA stacking center) before deformation (plotted for  $\theta = 5^\circ$ ). **b.** Graphene lattice plaquettes of two layers of TBG (at AA stacking center) with a relative shear deformation  $\Sigma_{xx} = -\Sigma_{yy} = \partial_x u_x - \partial_y u_y = 0.02$ , while expansion and rotation deformations are zero. **c.** The Moiré pattern superlattice of TBG before deformation. **d.** The Moiré pattern superlattice of TBG after the shear deformation  $\Sigma_{xx} = -\Sigma_{yy} = 0.02$ . One can see the superlattice is greatly affected, although the graphene lattices only have a 2% shear deformation.

where we have defined  $\gamma = [2 \tan(\theta/2)]^{-1}$ , the relative displacement  $\mathbf{u} = \mathbf{u}^{(1)} - \mathbf{u}^{(2)}$ , and the center of mass displacement  $\mathbf{u}^c = (\mathbf{u}^{(1)} + \mathbf{u}^{(2)})/2$ . For small angles  $\theta$ , we have  $\gamma \approx 1/\theta \gg 1$ , so  $\delta \mathbf{q}_j$  is dominated by the relative displacement phonon field  $\mathbf{u}$ . We therefore will ignore the contribution of center of mass displacement  $\mathbf{u}^c$  hereafter.

Fig. S2 shows the graphene lattices and Moiré patterns for  $\theta = 5^\circ$  before and after a relative shear deformation  $\Sigma_{xx} = -\Sigma_{yy} = 0.02$ , where  $\Sigma_{ab} = (\partial_a u_b + \partial_b u_a)/2 - (\sum_l \partial_l u_l) \delta_{ab}/2$  ( $a, b, l = 1, 2$ ) is the relative shear tensor. One can see that a small relative deformation greatly affects the Moiré pattern.

With the deformed vectors  $\mathbf{q}_j$ , the electron-phonon coupling Hamiltonian is simply given by the variation of the effective  $2 \times 2$  Hamiltonian  $\tilde{H}^{K, K'_M}$  in Eq. (S5), namely,  $H_{\text{ep}}^{K, K'_M}(\mathbf{k}) = \tilde{H}^{K, K'_M}(\mathbf{k}, \mathbf{q}_1 + \delta \mathbf{q}_1, )$

$$\begin{aligned} H_{\text{ep}}^{K, K'_M}(\mathbf{k}) &= \tilde{H}^{K, K'_M}(\mathbf{k}, \mathbf{q}_1 + \delta \mathbf{q}_1, \mathbf{q}_2 + \delta \mathbf{q}_2, \mathbf{q}_3 + \delta \mathbf{q}_3) - \tilde{H}^{K, K'_M}(\mathbf{k}, \mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \\ &= \frac{\hbar v w^2}{1 + 6\alpha^2} \sum_{j=1}^3 \left( T_j h_j^{-1} \boldsymbol{\sigma}^* \cdot \mathbf{k} \delta h_j^{-1} T_j^\dagger + h.c. \right) - \frac{1 - 3\alpha^2}{(1 + 6\alpha^2)^2} \hbar v w^2 \boldsymbol{\sigma}^* \cdot \mathbf{k} \sum_{j=1}^3 \psi_0^\dagger T_j \delta(h_j^{-2}) T_j^\dagger \psi_0, \end{aligned} \quad (\text{S8})$$

where we have defined  $h_j = h^K(-\mathbf{q}_j)$ , and  $\delta h_j = h^K(-\mathbf{q}_j - \delta \mathbf{q}_j) - h_j$ . The first term and the second term in the above result come from the variations  $\delta \langle \Psi | H^{K, K'_M} | \Psi \rangle$  and  $\delta \langle \Psi | \Psi \rangle$  of Eq. (S5), respectively. Explicitly, one can show that

$$\begin{aligned} \delta h_1^{-1} &= \gamma (\hbar v k_\theta)^{-1} (\partial_x u_x \sigma_x + \partial_y u_x \sigma_y), \\ \delta h_2^{-1} &= \gamma (\hbar v k_\theta)^{-1} \left\{ \left[ \frac{\sqrt{3}}{4} (\partial_y u_x - \partial_x u_y) + \frac{1}{4} \partial_x u_x - \frac{1}{4} \partial_y u_y \right] \sigma_x + \left[ \frac{1}{4} \partial_y u_x + \frac{3}{4} \partial_x u_y - \frac{\sqrt{3}}{4} (\partial_x u_x + \partial_y u_y) \right] \sigma_y \right\}, \\ \delta h_3^{-1} &= \gamma (\hbar v k_\theta)^{-1} \left\{ \left[ \frac{\sqrt{3}}{4} (\partial_x u_y - \partial_y u_x) + \frac{1}{4} \partial_x u_x - \frac{1}{4} \partial_y u_y \right] \sigma_x + \left[ \frac{1}{4} \partial_y u_x + \frac{3}{4} \partial_x u_y + \frac{\sqrt{3}}{4} (\partial_x u_x + \partial_y u_y) \right] \sigma_y \right\}. \end{aligned} \quad (\text{S9})$$

If we denote the above expression as  $\delta h_j^{-1} = (\hbar v k_\theta)^{-1} (\sigma_x A_{j,x} + \sigma_y A_{j,y})$ , we find  $w^2 T_j h_j^{-1} \boldsymbol{\sigma}^* \cdot \mathbf{k} \delta h_j^{-1} T_j^\dagger = -2\alpha^2 T_j (k_x A_{j,y} + k_y A_{j,x})$ . Besides, we note that  $h_j^{-2} = (\hbar v)^{-2} |\mathbf{q}_j|^{-2} I_2$  (where  $I_2$  is the  $2 \times 2$  identity matrix), so one has  $\delta(h_j^{-2}) =$

$(\hbar v)^{-2} \delta|\mathbf{q}_j|^{-2} I_2$ , and one can show that

$$\begin{aligned} \delta|\mathbf{q}_1|^{-2} &= 2\gamma k_\theta^{-2} \partial_y u_x, & \delta|\mathbf{q}_2|^{-2} &= \gamma k_\theta^{-2} \left[ \frac{1}{2} \partial_y u_x - \frac{3}{2} \partial_x u_y + \frac{\sqrt{3}}{2} (\partial_x u_x - \partial_y u_y) \right], \\ \delta|\mathbf{q}_3|^{-2} &= \gamma k_\theta^{-2} \left[ \frac{1}{2} \partial_y u_x - \frac{3}{2} \partial_x u_y - \frac{\sqrt{3}}{2} (\partial_x u_x - \partial_y u_y) \right]. \end{aligned}$$

Further, we set  $\psi_0$  in Eq. (S8) to be an eigenstate of  $\boldsymbol{\sigma}^* \cdot \mathbf{k}$  (with eigenvalue either  $+|\mathbf{k}|$  or  $-|\mathbf{k}|$ ). With these results, one then finds the electron-phonon coupling in Eq. (S8) to be

$$\begin{aligned} H_{\text{ep}}^{K,K'}(\mathbf{k}) &= -\gamma \hbar v \left\{ \frac{9\alpha^2(1+3\alpha^2)}{(1+6\alpha^2)^2} [k_x(\partial_y u_x + \partial_x u_y) + k_y(\partial_x u_x - \partial_y u_y)] \right. \\ &\quad \left. + \frac{9\alpha^2}{(1+6\alpha^2)^2} \boldsymbol{\sigma}^* \cdot \mathbf{k} (\partial_y u_x - \partial_x u_y) + \frac{3\alpha^2}{1+6\alpha^2} \hat{\mathbf{z}} \cdot (\boldsymbol{\sigma}^* \times \mathbf{k}) (\partial_x u_x + \partial_y u_y) \right\}, \end{aligned} \quad (\text{S10})$$

where the momentum  $\mathbf{k}$  in the above electron-phonon coupling should be understood as the average momentum of the electron before and after phonon emission (absorption), and explicitly one has  $\hat{\mathbf{z}} \cdot (\boldsymbol{\sigma}^* \times \mathbf{k}) = \sigma_x k_y + \sigma_y k_x$ . Note that the three terms

$$k_x(\partial_y u_x + \partial_x u_y) + k_y(\partial_x u_x - \partial_y u_y), \quad \boldsymbol{\sigma}^* \cdot \mathbf{k} (\partial_y u_x - \partial_x u_y), \quad \hat{\mathbf{z}} \cdot (\boldsymbol{\sigma}^* \times \mathbf{k}) (\partial_x u_x + \partial_y u_y)$$

are contributed to by the relative shear tensor  $\Sigma_{ab}$ , relative rotation  $R = \partial_x u_y - \partial_y u_x$ , and relative expansion  $\Theta = \partial_x u_x + \partial_y u_y$ , respectively. Since the shear tensor is a rank 2 tensor in the 2D space, the first term contributed by shear ( $\mathbf{k}$  times the shear tensor) is a rank 3 tensor. In particular, the first term can be rewritten as

$$k_x(\partial_y u_x + \partial_x u_y) + k_y(\partial_x u_x - \partial_y u_y) = \text{Im}[(k_x + ik_y)(\partial_x + i\partial_y)(u_x + iu_y)].$$

Under a rotation of angle  $\phi$  about  $z$  axis, each of the three terms  $k_x + ik_y$ ,  $\partial_x + i\partial_y$  and  $u_x + iu_y$  gains a phase factor  $e^{i\phi}$ . Therefore, this term is clearly only  $C_{3z}$  rotationally invariant about  $z$  axis (i.e.,  $\phi = 2\pi/3$ ). In contrast, both  $\boldsymbol{\sigma}^* \cdot \mathbf{k}$  and  $\hat{\mathbf{z}} \cdot (\boldsymbol{\sigma}^* \times \mathbf{k})$  are 2D scalars, and both the relative rotation  $R$  and the relative expansion  $\Theta$  are also 2D scalars, so the second and the third terms contributed by relative rotation and expansion are  $\text{SO}(2)$  rotationally symmetric about  $z$  axis.

One may ask why we only have term  $\text{Im}[(k_x + ik_y)(\partial_x + i\partial_y)(u_x + iu_y)]$  but not  $\text{Re}[(k_x + ik_y)(\partial_x + i\partial_y)(u_x + iu_y)]$ , both of which respect  $C_{3z}$  symmetry. This is because the TBG also has a 2-fold rotation symmetry  $C_{2x}$  about the  $x$  axis. Under  $C_{2x}$ , we have  $(\partial_x, \partial_y) \rightarrow (\partial_x, -\partial_y)$  and  $(k_x, k_y) \rightarrow (k_x, -k_y)$ , while the relative displacement field transforms differently as  $(u_x, u_y) \rightarrow (-u_x, u_y)$  due to the exchange of two layers. Since  $\text{Re}[(k_x + ik_y)(\partial_x + i\partial_y)(u_x + iu_y)]$  flips sign under  $C_{2x}$ , it is forbidden in the TBG electron phonon coupling. Similarly, one can check that terms like  $\boldsymbol{\sigma}^* \cdot \mathbf{k} (\partial_x u_x + \partial_y u_y)$  and  $\hat{\mathbf{z}} \cdot (\boldsymbol{\sigma}^* \times \mathbf{k}) (\partial_y u_x - \partial_x u_y)$  are also forbidden by  $C_{2x}$ .

When the TBG is at the magic angle, namely  $\alpha^2 = 1/3$ , the second term in Eq. (S8) vanishes, and Eq. (S10) becomes

$$H_{\text{ep}}^{K,K'}(\mathbf{k}) = -\frac{\gamma \hbar v}{3} [2k_x(\partial_y u_x + \partial_x u_y) + 2k_y(\partial_x u_x - \partial_y u_y) + \boldsymbol{\sigma}^* \cdot \mathbf{k} (\partial_y u_x - \partial_x u_y) + (\sigma_x k_y + \sigma_y k_x)(\partial_x u_x + \partial_y u_y)]. \quad (\text{S11})$$

In the same way one could obtain the electron-phonon couplings at other valley/Moiré valley/spin indices, which we shall not repeat here. Instead, we give a symmetry analysis derivation of electron-phonon coupling at other valley/Moiré valley/spin indices in the next section above Eq. (S21). The results are as shown in Eq. (S21) as well as in the main text Eq. (3).

## S2. THE PHONON HAMILTONIAN AND PHONON MEDIATED ELECTRON-ELECTRON INTERACTION

In this section, we first describe the Hamiltonian of the relative in-plane displacement phonon mode, and then consider the phonon mediated electron-electron interaction.

### A. Phonon Hamiltonian

In principle, the relative in-plane displacement phonons are optical phonons. However, if the two graphene lattices twisted by angle  $\theta$  are not commensurate, a uniform relative in-plane displacement  $\mathbf{u}$  does not cost energy, so these phonons are still acoustic. In fact, since the interlayer Van der Waals interaction between atoms in two graphene layers is much weaker than the intralayer atomic interaction, the in-plane polarized phonons of the two layers are nearly decoupled, and thus the relative in-plane displacement phonons are almost acoustic [2]. As a good approximation, we shall ignore the coupling between phonons in different layers. From elastic dynamics [3], we know the in-plane deformation (i.e., in-plane phonon) energy of the TBG in the continuum limit can be expressed as a quadratic function of the expansion scalar  $\Theta^{(j)} = \sum_l \partial_l u_l^{(j)}$  and the traceless shear tensor  $\Sigma_{ab}^{(j)} = (\partial_a u_b^{(j)} + \partial_b u_a^{(j)})/2 - (\sum_l \partial_l u_l^{(j)})\delta_{ab}/2$  of layer  $j = 1, 2$ , where  $a, b, l = x, y$ . In particular,  $\Theta^{(j)}$  and  $\Sigma_{ab}^{(j)}$  occupy the spin 0 and spin 2 representations of the SO(2) rotation group about  $z$  axis, respectively (one can easily verify  $\Sigma_{xx} \pm i\Sigma_{xy}$  have spin  $\pm 2$  under SO(2), respectively, where  $\Sigma_{xx}$  and  $\Sigma_{xy}$  are the two independent components of  $\Sigma_{ab}$ ). Due to the  $C_{6z}$  rotation symmetry of monolayer graphene, the symmetry allowed terms (scalars) in the deformation energy of layer  $j$  must occupy the spin 0 (mod 6) representations of SO(2). Therefore, the only allowed quadratic terms (scalars) are  $\Theta^{(j)^2}$  and  $\Sigma_{ab}^{(j)}\Sigma_{ab}^{(j)}$  which have spin 0 (there are no interlayer terms since we have approximated the two layers as decoupled). Therefore, the Hamiltonian (to quadratic order) for in-plane displacement  $\mathbf{u}^{(j)}$  (with two layers decoupled) can be written as

$$H_{ph} = \sum_{j=1}^2 \int d^2\mathbf{r} \left[ \frac{M}{\Omega} (\partial_t \mathbf{u}^{(j)})^2 + \frac{1}{2} K_{el} \Theta^{(j)^2} + \mu_{el} \Sigma_{ab}^{(j)} \Sigma_{ab}^{(j)} \right], \quad (\text{S12})$$

where  $M$  is the Carbon atom mass,  $\Omega = \sqrt{3}a_0^2/2$  is the graphene unit cell area, while  $K_{el}$  and  $\mu_{el}$  are the bulk modulus and shear modulus of monolayer graphene, respectively. The first term is the kinetic energy of the Carbon atoms (note that there are 2 atoms in each graphene unit cell), while the second and the third terms are the elastic potential energy. To separate the relative deformation  $\mathbf{u} = \mathbf{u}^{(1)} - \mathbf{u}^{(2)}$  and the center-of-mass deformation  $\mathbf{u}^c = (\mathbf{u}^{(1)} + \mathbf{u}^{(2)})/2$ , we define the relative expansion scalar and shear tensor as  $\Theta = \sum_l \partial_l u_l$  and  $\Sigma_{ab} = (\partial_a u_b + \partial_b u_a)/2 - (\sum_l \partial_l u_l)\delta_{ab}/2$ , and the center-of-mass expansion and shear  $\Theta^c = \sum_l \partial_l u_l^c$  and  $\Sigma_{ab}^c = (\partial_a u_b^c + \partial_b u_a^c)/2 - (\sum_l \partial_l u_l^c)\delta_{ab}/2$ , respectively. The phonon Hamiltonian can then be written as two parts  $H_{ph} = H_{ph}^r + H_{ph}^c$ , where the relative phonon wave part is

$$\begin{aligned} H_{ph}^r &= \int d^2\mathbf{r} \left[ \frac{M}{2\Omega} (\partial_t \mathbf{u})^2 + \frac{1}{4} K_{el} \Theta^2 + \frac{\mu_{el}}{2} \Sigma_{ab} \Sigma_{ab} \right] \\ &= \int d^2\mathbf{r} \left\{ \frac{M}{2\Omega} (\partial_t \mathbf{u})^2 + \frac{K_{el} + \mu_{el}}{4} [(\partial_x u_x)^2 + (\partial_y u_y)^2] + \frac{\mu_{el}}{4} [(\partial_x u_y)^2 + (\partial_y u_x)^2] + \frac{K_{el}}{2} \partial_x u_x \partial_y u_y \right\}, \end{aligned} \quad (\text{S13})$$

and the center-of-mass phonon wave part is

$$H_{ph}^c = \int d^2\mathbf{r} \left[ \frac{2M}{\Omega} (\partial_t \mathbf{u}^c)^2 + K_{el} (\Theta^c)^2 + 2\mu_{el} \Sigma_{ab}^c \Sigma_{ab}^c \right]. \quad (\text{S14})$$

Since the relative phonon wave dominates the electron-phonon coupling, we shall consider only the relative phonon part  $H_{ph}^r$  in Eq. (S13). Explicitly, the equation of motion of  $H_{ph}^r$  is

$$\frac{M}{\Omega} \partial_t^2 \begin{pmatrix} u_x \\ u_y \end{pmatrix} = \frac{1}{2} \begin{pmatrix} K_{el} \partial_x^2 + \mu_{el} (\partial_x^2 + \partial_y^2) & K_{el} \partial_x \partial_y \\ K_{el} \partial_x \partial_y & K_{el} \partial_y^2 + \mu_{el} (\partial_x^2 + \partial_y^2) \end{pmatrix} \begin{pmatrix} u_x \\ u_y \end{pmatrix} = \frac{K_{el}}{2} \nabla (\nabla \cdot \mathbf{u}) + \frac{\mu_{el}}{2} \nabla^2 \mathbf{u}. \quad (\text{S15})$$

Given the momentum  $-i\nabla = \mathbf{p} = (p_x, p_y)$ , there are two eigenvectors:  $\mathbf{u}_{\mathbf{p},L} \propto \mathbf{p}$  and  $\mathbf{u}_{\mathbf{p},T} \propto \hat{\mathbf{z}} \times \mathbf{p}$ , which satisfies  $\nabla (\nabla \cdot \mathbf{u}_{\mathbf{p},L}) = \nabla^2 \mathbf{u}_{\mathbf{p},L}$  and  $\nabla (\nabla \cdot \mathbf{u}_{\mathbf{p},T}) = 0$ , respectively. Accordingly, their eigenfrequencies are  $\omega_{\mathbf{p},L} = c_L p$  and  $\omega_{\mathbf{p},T} = c_T p$ , respectively, where  $c_L = \sqrt{(K_{el} + \mu_{el})\Omega/2M}$  and  $c_T = \sqrt{\mu_{el}\Omega/2M}$  being the longitudinal sound speed and transverse sound speed of monolayer graphene, and  $p = |\mathbf{p}|$ .

The Hamiltonian of relative phonon waves can then be quantized following the standard canonical method. In real space, the canonical momentum  $\boldsymbol{\pi}(\mathbf{r}) = \frac{M}{\Omega} \partial_t \mathbf{u}(\mathbf{r})$  satisfies  $[u_a(\mathbf{r}), \pi_b(\mathbf{r}')] = i\hbar \delta_{ab} \delta(\mathbf{r} - \mathbf{r}')$  (where  $a, b = x, y$ ). When transformed into the momentum space, the canonical momentum  $\boldsymbol{\pi}_{\mathbf{p}} = \int \frac{d^2\mathbf{r}}{\sqrt{A_s}} e^{-i\mathbf{p} \cdot \mathbf{r}} \boldsymbol{\pi}(\mathbf{r})$  and displacement  $\mathbf{u}_{\mathbf{p}} = \int \frac{d^2\mathbf{r}}{\sqrt{A_s}} e^{-i\mathbf{p} \cdot \mathbf{r}} \mathbf{u}(\mathbf{r})$  satisfies  $[u_{\mathbf{p},a}, \pi_{\mathbf{p}',b}] = i\hbar \delta_{ab} \delta_{\mathbf{p},-\mathbf{p}'}$ , where  $A_s$  is the total area of the sample. We can decompose  $\mathbf{u}_{\mathbf{p}} = \mathbf{u}_{\mathbf{p},L} + \mathbf{u}_{\mathbf{p},T}$  into a longitudinal part  $\mathbf{u}_{\mathbf{p},L}$  and a transverse part  $\mathbf{u}_{\mathbf{p},T}$  (which are perpendicular to each other), and similarly we can also do so for the canonical momentum  $\boldsymbol{\pi}_{\mathbf{p}} = \boldsymbol{\pi}_{\mathbf{p},L} + \boldsymbol{\pi}_{\mathbf{p},T}$  (which are also perpendicular to each

other,  $\mathbf{u}_{\mathbf{p},T} \cdot \mathbf{u}_{\mathbf{p},L} = 0$ ). One can then show the phonon Hamiltonian in the momentum space becomes

$$H_{\text{ph}}^r = \sum_{\mathbf{p}} \left[ \frac{\Omega}{2M} \boldsymbol{\pi}_{\mathbf{p}} \cdot \boldsymbol{\pi}_{-\mathbf{p}} + \frac{K_{el}}{2} (\mathbf{p} \cdot \mathbf{u}_{\mathbf{p}})(\mathbf{p} \cdot \mathbf{u}_{-\mathbf{p}}) + \frac{\mu_{el}}{2} p^2 \mathbf{u}_{\mathbf{p}} \cdot \mathbf{u}_{-\mathbf{p}} \right] = \sum_{\mathbf{p}} \left( \hbar \omega_{\mathbf{p},L} a_{\mathbf{p},L}^\dagger a_{\mathbf{p},L} + \hbar \omega_{\mathbf{p},T} a_{\mathbf{p},T}^\dagger a_{\mathbf{p},T} \right), \quad (\text{S16})$$

where  $a_{\mathbf{p},\chi}$ ,  $a_{\mathbf{p},\chi}^\dagger$  are the phonon annihilation and creation operators with momentum  $\mathbf{p}$  and polarization  $\chi$  satisfying  $[a_{\mathbf{p},\chi}, a_{\mathbf{p},\chi}^\dagger] = 1$ . The relative displacement phonon field  $\mathbf{u}$  in the Shrödinger picture is given by

$$\mathbf{u}(\mathbf{r}) = \sum_{\mathbf{p}} \frac{e^{i\mathbf{p} \cdot \mathbf{r}}}{\sqrt{N_s \Omega_s}} (i \hat{\mathbf{p}} u_{\mathbf{p},L} + i \hat{\mathbf{z}} \times \hat{\mathbf{p}} u_{\mathbf{p},T}), \quad u_{\mathbf{p},L} = \sqrt{\frac{\hbar \Omega}{2M \omega_{\mathbf{p},L}}} (a_{\mathbf{p},L} + a_{-\mathbf{p},L}^\dagger), \quad u_{\mathbf{p},T} = \sqrt{\frac{\hbar \Omega}{2M \omega_{\mathbf{p},T}}} (a_{\mathbf{p},T} + a_{-\mathbf{p},T}^\dagger), \quad (\text{S17})$$

where  $\hat{\mathbf{p}} = \mathbf{p}/p$  is the unit vector along momentum  $\mathbf{p}$ , while  $\hat{\mathbf{z}}$  is the unit vector along the out of plane direction ( $\hat{\mathbf{z}} \times \hat{\mathbf{p}}$  is the transverse direction), and for later convenience we have rewritten the total area of the sample  $A_s$  as  $N_s \Omega_s$ , with  $N_s$  being the number of superlattice unit cells, and  $\Omega_s = \Omega/[4 \sin^2(\theta/2)]$  being the superlattice unit cell area. In the Heisenberg picture, the phonon field  $\mathbf{u}$  is time  $t$  dependent, namely,  $\mathbf{u}(\mathbf{r}, t) = e^{iH_{\text{ph}}^r t} \mathbf{u}(\mathbf{r}) e^{-iH_{\text{ph}}^r t}$ , with  $\mathbf{u}(\mathbf{r})$  defined in Eq. (S17) above. Accordingly, one can obtain the canonical momentum as  $\boldsymbol{\pi}(\mathbf{r}, t) = \frac{M}{\Omega} \partial_t \mathbf{u}(\mathbf{r}, t)$ .

Since the interlayer phonon coupling is ignored, the phonon bands in the MBZ is simply obtained by folding  $\mathbf{p}$  into the superlattice MBZ. Among these phonon bands, the lowest two bands (one longitudinal and one transverse) are acoustic (under the approximation that the two layers are decoupled) and cause the long wavelength deformation of the Moiré pattern superlattice, while the higher bands are optical and mainly lead to short wavelength deformations within each unit cell of the superlattice. Since our electron phonon coupling is derived in the long wavelength deformation limit, for now we shall restrict ourselves to the lowest acoustic phonon band in the MBZ. We will briefly discuss the contribution of optical phonon bands at the end of this supplementary section.

## B. Electron-Phonon Coupling for other valley/Moiré valley/spin

For convenience, hereafter we define the graphene BZ valley  $K, K'$  as index  $\eta = \pm 1$ , the Moiré valley  $K_M, K'_M$  as index  $\zeta = \pm 1$ , and the spin up and down as index  $s = \pm 1$ . Since there is no spin-orbit coupling, the electron-phonon coupling does not flip spin and is independent of spin index  $s$ . In last section we derived the electron-phonon coupling for  $(\eta, \zeta) = (+1, -1)$ . Now we use symmetry arguments to obtain the expression of electron-phonon coupling at other indices  $(\eta, \zeta)$ . Identical results can be obtained by performing brute force calculations.

First, TBG has the time-reversal symmetry. A time-reversal transformation  $\mathcal{T}$  changes  $(\eta, \zeta, s) \rightarrow (-\eta, -\zeta, -s)$  and  $\mathbf{k} \rightarrow -\mathbf{k}$ . Since for spinless (no spin-orbit coupling) fermions the time reversal  $\mathcal{T}$  is simply complex conjugation, one finds the electron-phonon coupling at  $(-\eta, -\zeta)$  is given by

$$H_{\text{ep}}^{-\eta, -\zeta}(\mathbf{k}) = [H_{\text{ep}}^{\eta, \zeta}(-\mathbf{k})]^*, \quad (\text{S18})$$

where we have found  $H_{\text{ep}}^{K, K'_M}(\mathbf{k})$  in Eq. (S11), and  $z^*$  stands for the complex conjugate of  $z$ .

Secondly, TBG also has a  $C_{2x}$  symmetry, which is the 2-fold rotation about  $x$  axis (see Fig. S1a for definitions of  $x$  and  $y$  axes). From the main text Fig. 2a and supplementary Fig. S1, one can see the  $C_{2x}$  transformation interchanges Moiré valley  $K_M$  and  $K'_M$ , changes momentum  $(k_x, k_y)$  (measured from the  $K_M$  point) to  $(k_x, -k_y)$  (measured from the  $K'_M$  point) in momentum space, interchanges layer 1 with layer 2, and interchanges sublattice indices A and B. Meanwhile, the valley  $K$  or  $K'$  remains invariant. Since the Pauli matrices in Eq. (S4) (and afterwards) are in the sublattice basis, the interchange of sublattice indices A and B leads to a transformation of Pauli matrices  $\boldsymbol{\sigma} \rightarrow \sigma_x \boldsymbol{\sigma} \sigma_x^{-1}$ , namely,  $(\sigma_x, \sigma_y) \rightarrow (\sigma_x, -\sigma_y)$ . Therefore, the band Hamiltonian of the continuum model at Moiré valley  $K_M$  and valley  $K$  is given by

$$H^{K, K_M}(\mathbf{k}) = U \left[ H^{K, K'_M}(k_x, -k_y) \right] U^{-1} = \begin{pmatrix} h_{-\theta/2}^K(\mathbf{k}) & wT_1 & wT_3 & wT_2 \\ wT_1^\dagger & h_{\theta/2}^K(\mathbf{k} + \mathbf{q}_1) & 0 & 0 \\ wT_3^\dagger & 0 & h_{\theta/2}^K(\mathbf{k} + \mathbf{q}_3) & 0 \\ wT_2^\dagger & 0 & 0 & h_{\theta/2}^K(\mathbf{k} + \mathbf{q}_2) \end{pmatrix}, \quad (\text{S19})$$

where  $H^{K, K'_M}(\mathbf{k})$  is given by Eq. (S4), and  $U = \sigma_x \otimes I_4$  is the transformation matrix for the interchange of sublattices A and B under  $C_{2x}$  (i.e.,  $\psi_{j,\mathbf{k}} \rightarrow \sigma_x \psi_{j,\mathbf{k}}$  for each  $\psi_{j,\mathbf{k}}$  ( $1 \leq j \leq 4$ ) in the basis of Hamiltonian  $(\psi_{0,\mathbf{k}}^T, \psi_{1,\mathbf{k}}^T, \psi_{2,\mathbf{k}}^T, \psi_{3,\mathbf{k}}^T)^T$ , recall that  $\psi_{j,\mathbf{k}}$  are defined in the 2D Hilbert space of sublattice A and B). If one approximate  $\theta/2 = 0$  in  $h_{\pm\theta/2}^K$  (as we

have assumed in the first supplementary section), and interchanges electron basis  $\psi_{2,\mathbf{k}} \leftrightarrow \psi_{3,\mathbf{k}}$ , one finds  $H^{K,K_M}(\mathbf{k})$  and  $H^{K,K'_M}(\mathbf{k})$  only differ by a sign flip  $\mathbf{q}_j \rightarrow -\mathbf{q}_j$  ( $j = 1, 2, 3$ ). From Eq. (S8) one sees the expression of electron-phonon coupling  $H_{\text{ep}}^{K,K'_M}(\mathbf{k})$  is quadratic in  $\mathbf{q}_j$  (thus invariant under  $\mathbf{q}_j \rightarrow -\mathbf{q}_j$ ), so we conclude  $H_{\text{ep}}^{K,K'_M}(\mathbf{k}) = H_{\text{ep}}^{K,K_M}(\mathbf{k})$  (where  $\mathbf{k}$  of  $H_{\text{ep}}^{K,K'_M}(\mathbf{k})$  and  $H_{\text{ep}}^{K,K_M}(\mathbf{k})$  are measured from  $K'_M$  and  $K_M$ , respectively), and  $H_{\text{ep}}^{\eta,\zeta}(\mathbf{k})$  is invariant under  $\zeta \rightarrow -\zeta$ . Alternatively, one can achieve this conclusion by directly applying  $C_{2x}$  on  $H_{\text{ep}}^{K,K'_M}(\mathbf{k})$  we derived in Eq. (S10). In particular,  $C_{2x}$  changes the displacement field in layer  $j$  as  $(u_x^{(j)}, u_y^{(j)}) \rightarrow (u_x^{(j)}, -u_y^{(j)})$ , and interchanges layer 1 with layer 2, so the relative displacement field  $\mathbf{u} = \mathbf{u}^{(1)} - \mathbf{u}^{(2)}$  changes as  $(u_x, u_y) \rightarrow (-u_x, u_y)$ . Besides, it changes  $(\partial_x, \partial_y) \rightarrow (\partial_x, -\partial_y)$ . Therefore, one finds

$$H_{\text{ep}}^{K,K_M}(\mathbf{k}) = \sigma_x \left[ H_{\text{ep}}^{K,K'_M}(k_x, -k_y) \right]_{\partial_y \rightarrow -\partial_y, u_x \rightarrow -u_x} \sigma_x^{-1} = H_{\text{ep}}^{K,K'_M}(\mathbf{k}), \quad (\text{S20})$$

where  $H_{\text{ep}}^{K,K'_M}(\mathbf{k})$  is given in Eq. (S10) (obtained under the approximation  $\theta/2 = 0$  in  $h_{\pm\theta/2}^K$  in the first supplementary section). This indicates the electron-phonon coupling  $H_{\text{ep}}^{\eta,\zeta}(\mathbf{k})$  is also independent of the Moiré valley  $\zeta$ . We can then obtain the electron-phonon coupling Hamiltonian for all indices, which in the second quantized language takes the form

$$\begin{aligned} H_{\text{ep}}^{\eta,\zeta,s} &= H_{C3}^{\eta,\zeta,s} + H_{SO(2)}^{\eta,\zeta,s}, \\ H_{C3}^{\eta,\zeta,s} &= g_{1\alpha} \frac{\gamma\hbar v}{\sqrt{N_s\Omega_s}} \sum_{\mathbf{k},\mathbf{k}'} p \psi_{\mathbf{k}}^{\eta,\zeta,s\dagger} \left\{ [2\bar{k}_x \hat{p}_x \hat{p}_y + \bar{k}_y (\hat{p}_x^2 - \hat{p}_y^2)] u_{\mathbf{p},L} + [2\bar{k}_y \hat{p}_x \hat{p}_y - \bar{k}_x (\hat{p}_x^2 - \hat{p}_y^2)] u_{\mathbf{p},T} \right\} \psi_{\mathbf{k}'}^{\eta,\zeta,s}, \\ H_{SO(2)}^{\eta,\zeta,s} &= -\frac{\gamma\hbar v}{\sqrt{N_s\Omega_s}} \sum_{\mathbf{k},\mathbf{k}'} p \psi_{\mathbf{k}}^{\eta,\zeta,s\dagger} [g_{2\alpha} (\eta\sigma_x \bar{k}_x - \sigma_y \bar{k}_y) u_{\mathbf{p},T} + g_{3\alpha} (\sigma_y \bar{k}_x + \eta\sigma_x \bar{k}_y) u_{\mathbf{p},L}] \psi_{\mathbf{k}'}^{\eta,\zeta,s}, \end{aligned} \quad (\text{S21})$$

where the numerical factors are from Eq. (S10):

$$g_{1\alpha} = \frac{9\alpha^2(1+3\alpha^2)}{(1+6\alpha^2)^2}, \quad g_{2\alpha} = \frac{9\alpha^2}{(1+6\alpha^2)^2}, \quad g_{3\alpha} = \frac{3\alpha^2}{1+6\alpha^2}, \quad (\text{S22})$$

while  $\psi_{\mathbf{k}'}^{\eta,\zeta,s}$  is the electron annihilation operator at momentum  $\mathbf{k}$  with  $\eta, \zeta, s$  indices,  $\mathbf{k} - \mathbf{k}' = \mathbf{p}$  (the momentum of the  $\partial_a u_b$  terms), and we have defined  $\mathbf{k} = (\mathbf{k} + \mathbf{k}')/2$  is the average electron momentum before and after phonon emission (absorption), which comes from Eq. (S10). Physically, this comes from the Fourier transformation of the real space hopping amplitude  $t(\mathbf{r}, \mathbf{r}') c^\dagger(\mathbf{r}) c(\mathbf{r}')$ , where  $t(\mathbf{r}, \mathbf{r}')$  is induced by the displacement field  $\mathbf{u}(\mathbf{r})$ . In particular, for the magic angle  $\alpha^2 = 1/3$ , we have  $g_{1\alpha} = 2/3$  and  $g_{2\alpha} = g_{3\alpha} = 1/3$ . Besides, as we have shown below Eq. (S10), the term  $H_{C3}^{\eta,\zeta,s}$  is only  $C_{3z}$  rotationally invariant, while the term  $H_{SO(2)}^{\eta,\zeta,s}$  is  $SO(2)$  rotationally invariant. Finally, we note that  $H_{C3}^{\eta,\zeta,s}$  is odd in  $K, K'$  valley index  $\eta$ , since the electron band Hamiltonian and thus the electron-phonon coupling undergoes a sign flip and complex conjugate under  $\eta \rightarrow -\eta$ .

The electron-phonon coupling  $H_{\text{ep}}^{\eta,\zeta,s}$  we obtained is independent of the Moiré valley  $\zeta$  and the spin  $s$ . However, we note that the independence of Moiré valley holds only under the approximation  $\theta/2 = 0$  in  $h_{\pm\theta/2}^K(\mathbf{k})$ . If we do not make this approximation,  $H_{\text{ep}}^{\eta,\zeta,s}$  will be weakly  $\zeta$  dependent, where the  $\zeta$  dependent part will be of a factor  $\mathcal{O}(\theta)$  smaller than the  $\zeta$  independent part given in Eq. (S10). Since we are interested in small angles  $\theta \sim 1^\circ \sim 0.02$  rad, we shall ignore the weak  $\zeta$  dependence of  $H_{\text{ep}}^{\eta,\zeta,s}$ . Besides, since the continuum model of TBG has no coupling between valleys  $K$  and  $K'$ , the electron-phonon coupling we obtained has a definite  $K, K'$  valley index  $\eta$  (i.e., initial and final states of the electron are in the same valley  $\eta$ ).

### C. Phonon-Mediated Electron-Electron Interaction

We now calculate the phonon mediated electron-electron interaction near the Fermi surface. To be concrete, we shall assume the Fermi surface is at  $|\mathbf{k}| = k_F$  in the hole Dirac bands (where superconductivity is observed), and project the electron operators to the vicinity of the Fermi surface. Such a Fermi surface might not be very accurate since the band is quite flat, but we shall simply model the Fermi surface by a circle with a large  $k_F$  of a low velocity 2D Dirac fermion. The wave function  $\phi_{\mathbf{k}}^\eta$  of the Dirac hole band state  $|\mathbf{k}_{\eta,\zeta,s}\rangle$  is assumed to be the negative eigenvalue eigenstate of  $(\eta\sigma_x k_x - \sigma_y k_y)$  (which is proportional to the Dirac Hamiltonian  $\tilde{H}^{\eta,\zeta}(\mathbf{k}) = \frac{1-3\alpha^2}{1+6\alpha^2} \hbar v (\eta\sigma_x k_x - \sigma_y k_y)$ ), namely,  $\phi_{\mathbf{k}}^\eta = (1, -\eta e^{-i\eta\varphi_{\mathbf{k}}})^T / \sqrt{2}$  at valley  $\eta$ , where  $\varphi_{\mathbf{k}} = \arg(k_x + ik_y)$  is the polar angle of momentum  $\mathbf{k}$  (we note



that for  $\alpha^2 < 1/3$  this is the valence band of the Dirac Hamiltonian, while for  $\alpha^2 > 1/3$  this in fact becomes the conduction band of the Dirac Hamiltonian, and for  $\alpha^2 = 1/3$  the conduction or valence band becomes ill-defined unless higher order terms in  $\mathbf{k}$  are included. Here we shall ignore these complications and take the wave function  $\phi_{\mathbf{k}}^\eta$ . One can then rewrite the Dirac annihilation and creation operators  $\psi_{\mathbf{k}}^{\eta,\zeta,s}, \psi_{\mathbf{k}}^{\eta,\zeta,s\dagger}$  in Eq. (S21) as  $\psi_{\mathbf{k}}^{\eta,\zeta,s} = \phi_{\mathbf{k}}^\eta c_{\mathbf{k},\eta,\zeta,s}$  and  $\psi_{\mathbf{k}}^{\eta,\zeta,s\dagger} = \phi_{\mathbf{k}}^{\eta\dagger} c_{\mathbf{k},\eta,\zeta,s}^\dagger$ , where  $c_{\mathbf{k},\eta,\zeta,s}$  and  $c_{\mathbf{k},\eta,\zeta,s}^\dagger$  are the electron annihilation and creation operators in the hole Dirac band with indices  $\eta, \zeta, s$ . The Dirac hole band Hamiltonian can then be written as

$$\tilde{H}^{\eta,\zeta,s}(\mathbf{k}) = \xi_{\mathbf{k}} c_{\mathbf{k},\eta,\zeta,s}^\dagger c_{\mathbf{k},\eta,\zeta,s},$$

where  $\xi_{\mathbf{k}} = -\frac{1-3\alpha^2}{1+6\alpha^2} \hbar v(|\mathbf{k}| - k_F)$  is the band energy relative to the Fermi level.

We then make the approximation  $|\mathbf{k}| \approx |\mathbf{k}'| \approx k_F$ , based on which we find the two terms in  $H_{SO(2)}^{\eta,\zeta,s}$  are approximately

$$\phi_{\mathbf{k}}^{\eta\dagger} (\eta \sigma_x \bar{k}_x - \sigma_y \bar{k}_y) \phi_{\mathbf{k}'}^\eta = \phi_{\mathbf{k}}^{\eta\dagger} \left[ \frac{(\eta \sigma_x k_x - \sigma_y k_y)}{2} + \frac{(\eta \sigma_x k'_x - \sigma_y k'_y)}{2} \right] \phi_{\mathbf{k}'}^\eta = \phi_{\mathbf{k}}^{\eta\dagger} \left( -\frac{|\mathbf{k}|}{2} - \frac{|\mathbf{k}'|}{2} \right) \phi_{\mathbf{k}'}^\eta \approx -k_F \phi_{\mathbf{k}}^{\eta\dagger} \phi_{\mathbf{k}'}^\eta,$$

$$\phi_{\mathbf{k}}^{\eta\dagger} (\eta \sigma_x \bar{k}_y + \sigma_y \bar{k}_x) \phi_{\mathbf{k}'}^\eta = \phi_{\mathbf{k}}^{\eta\dagger} \left[ \frac{(\eta \sigma_x k_y + \sigma_y k_x)}{2} + \frac{(\eta \sigma_x k'_y + \sigma_y k'_x)}{2} \right] \phi_{\mathbf{k}'}^\eta = i(|\mathbf{k}'| - |\mathbf{k}|) \phi_{\mathbf{k}}^{\eta\dagger} \sigma_z \phi_{\mathbf{k}'}^\eta \approx 0,$$

where we have used the definition  $(\eta \sigma_x k_x - \sigma_y k_y) \phi_{\mathbf{k}}^\eta = -|\mathbf{k}| \phi_{\mathbf{k}}^\eta$ , and  $(\eta \sigma_x k_y + \sigma_y k_x) \phi_{\mathbf{k}}^\eta = -i \sigma_z (\eta \sigma_x k_x - \sigma_y k_y) \phi_{\mathbf{k}}^\eta = i|\mathbf{k}| \sigma_z \phi_{\mathbf{k}}^\eta$ . Besides, we note that the two terms in  $H_{C_3}^{\eta,\zeta,s}$  of Eq. (S21) satisfy

$$\begin{aligned} & -[2\bar{k}_y \hat{p}_x \hat{p}_y - \bar{k}_x (\hat{p}_x^2 - \hat{p}_y^2)] + i[2\bar{k}_x \hat{p}_x \hat{p}_y + \bar{k}_y (\hat{p}_x^2 - \hat{p}_y^2)] = (\hat{p}_x + i\hat{p}_y)^2 (\bar{k}_x + i\bar{k}_y) = \left( \frac{p_x + ip_y}{p_x - ip_y} \right) (\bar{k}_x + i\bar{k}_y) \\ & = \left( \frac{|\mathbf{k}| e^{i\varphi_{\mathbf{k}}} - |\mathbf{k}'| e^{i\varphi_{\mathbf{k}'}}}{|\mathbf{k}| e^{-i\varphi_{\mathbf{k}}} - |\mathbf{k}'| e^{-i\varphi_{\mathbf{k}'}}} \right) \times \left( \frac{|\mathbf{k}| e^{i\varphi_{\mathbf{k}}} + |\mathbf{k}'| e^{i\varphi_{\mathbf{k}'}}}{2} \right) \approx \frac{k_F}{2} \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}}, \end{aligned} \quad (\text{S23})$$

where we have used  $\mathbf{p} = \mathbf{k} - \mathbf{k}'$  and  $\bar{\mathbf{k}} = (\mathbf{k} + \mathbf{k}')/2$ . Therefore, they can be written as the real part and imaginary part of the complex quantity  $\frac{k_F}{2} \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}}$ , respectively. This enables us to approximate the electron-phonon interaction projected in the Dirac hole band as

$$\begin{aligned} H_{\text{ep}}^{\eta,\zeta,s} & \approx \frac{\gamma \hbar v k_F}{\sqrt{N_s \Omega_s}} \sum_{\mathbf{k}, \mathbf{k}'} p \left\{ \left[ g_{2\alpha} - \eta \frac{g_{1\alpha}}{2} \text{Re} \left( \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right) \right] u_{\mathbf{p},T} + \eta \frac{g_{1\alpha}}{2} \text{Im} \left( \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right) u_{\mathbf{p},L} \right\} \\ & \times \phi_{\mathbf{k}}^{\eta\dagger} \phi_{\mathbf{k}'}^\eta c_{\mathbf{k},\eta,\zeta,s}^\dagger c_{\mathbf{k}',\eta,\zeta,s}. \end{aligned} \quad (\text{S24})$$

In particular, under momentum reversal  $\mathbf{k}, \mathbf{k}' \rightarrow -\mathbf{k}, -\mathbf{k}'$ , the polar angles  $\varphi_{\mathbf{k}} \rightarrow \varphi_{\mathbf{k}} + \pi$  and  $\varphi_{\mathbf{k}'} \rightarrow \varphi_{\mathbf{k}'} + \pi$ , so we have  $\frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \rightarrow \frac{e^{2i\varphi_{\mathbf{k}}+2i\pi} - e^{2i\varphi_{\mathbf{k}'}+2i\pi}}{e^{-i\varphi_{\mathbf{k}}-i\pi} - e^{-i\varphi_{\mathbf{k}'}-i\pi}} = -\frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}}$ . Therefore, the terms in the projected electron-phonon coupling contributed by  $H_{C_3}^{\eta,\zeta,s}$  (proportional to  $g_{1\alpha}$ ) and by  $H_{SO(2)}^{\eta,\zeta,s}$  (proportional to  $g_{2\alpha}$ ) are odd and even under momentum reversal  $\mathbf{k}, \mathbf{k}' \rightarrow -\mathbf{k}, -\mathbf{k}'$ , respectively, which will be useful in later calculations. We note that this odd/evenness results from the projection of electron-phonon coupling onto a single band. In the original expression of Eq. (S21) before projection, one may thought both  $H_{C_3}^{\eta,\zeta,s}$  and  $H_{SO(2)}^{\eta,\zeta,s}$  are odd in  $\mathbf{k}$  since they are both linear in  $\mathbf{k}$ . However, this naive expectation ignores the effect of the Dirac wave functions  $\psi_{\mathbf{k}}^{\eta,\zeta,s}$  in the expression of Eq. (S21), thus is incorrect.

One can then use the standard second order perturbation theory to calculate the phonon-mediated electron electron interaction. We treat the electron energy  $\tilde{H}^{\eta,\zeta,s}(\mathbf{k})$  and the phonon energy  $H_{ph}^r$  as the unperturbed Hamiltonian, and the electron-phonon interaction  $H_{\text{ep}}^{\eta,\zeta,s}$  as the perturbation. The electron electron interaction is then induced by the emission and absorption of a phonon between two electrons. Consider the initial state  $|\Psi_0\rangle = |\mathbf{k}_{1,I}, \mathbf{k}_{2,I'}\rangle = c_{\mathbf{k}_{1,I}}^\dagger c_{\mathbf{k}_{2,I'}}^\dagger |0\rangle$  of two electrons which have momenta  $\mathbf{k}_1$  and  $\mathbf{k}_2$ , respectively, where  $I = (\eta, \zeta, s)$  and  $I' = (\eta', \zeta', s')$  are short hand for their Dirac cone indices, and  $|0\rangle$  is the particle vacuum. The initial state energy is  $E_0 = \xi_{\mathbf{k}_1} + \xi_{\mathbf{k}_2}$ . Assume the final state is  $|\Psi_f\rangle = |\mathbf{k}_{3,I}, \mathbf{k}_{4,I'}\rangle$ , where the total momentum  $\mathbf{k}_3 + \mathbf{k}_4 = \mathbf{k}_1 + \mathbf{k}_2$  is conserved, and the final state energy  $E_f = \xi_{\mathbf{k}_3} + \xi_{\mathbf{k}_4} = E_0$  remains unchanged. There are two intermediate states with an emitted phonon with polarization  $\chi$ : one is  $|\Psi_{1,\chi}\rangle = |\mathbf{k}_{3,I}, \mathbf{k}_{2,I'}, \mathbf{p}, \chi\rangle = c_{\mathbf{k}_{3,I}}^\dagger c_{\mathbf{k}_{2,I'}}^\dagger a_{\mathbf{p},\chi}^\dagger |0\rangle$  where a phonon with momentum  $\mathbf{p}$  and polarization  $\chi$  is emitted from the first electron, while the other is  $|\Psi_{2,\chi}\rangle = |\mathbf{k}_{1,I}, \mathbf{k}_{4,I'}, -\mathbf{p}, \chi\rangle$  where a phonon with

momentum  $-\mathbf{p}$  and polarization  $\chi$  is emitted from the second electron, with the momentum  $\mathbf{p} = \mathbf{k}_1 - \mathbf{k}_3 = \mathbf{k}_4 - \mathbf{k}_2$ . The energy of the two intermediate states are  $E_{1,\chi} = \xi_{\mathbf{k}_3} + \xi_{\mathbf{k}_2} + \hbar\omega_{\mathbf{p},\chi}$  and  $E_{2,\chi} = \xi_{\mathbf{k}_1} + \xi_{\mathbf{k}_4} + \hbar\omega_{-\mathbf{p},\chi}$ , respectively, where  $\hbar\omega_{\mathbf{p},\chi}$  is the phonon energy. The phonon of the intermediate state is then absorbed by the other electron, resulting in the final state  $|\Psi_f\rangle$ . According to the second order perturbation theory, the interaction between the two electrons is given by

$$\begin{aligned} \frac{V_{\mathbf{k}_3\mathbf{k}_4,\mathbf{k}_1\mathbf{k}_2}^{II'}}{N_s\Omega_s} &= \sum_{\chi} \left( \frac{\langle\Psi_f|H_{\text{ep}}^{I'}|\Psi_{1,\chi}\rangle\langle\Psi_{1,\chi}|H_{\text{ep}}^I|\Psi_0\rangle}{E_0 - E_{1,\chi}} + \frac{\langle\Psi_f|H_{\text{ep}}^I|\Psi_{2,\chi}\rangle\langle\Psi_{2,\chi}|H_{\text{ep}}^{I'}|\Psi_0\rangle}{E_0 - E_{2,\chi}} \right) \\ &= \sum_{\chi} \left( \frac{\langle\Psi_f|H_{\text{ep}}^{I'}|\Psi_1\rangle\langle\Psi_1|H_{\text{ep}}^I|\Psi_0\rangle}{\xi_{\mathbf{k}_3} - \xi_{\mathbf{k}_1} - \hbar\omega_{\mathbf{p},\chi}} + \frac{\langle\Psi_f|H_{\text{ep}}^I|\Psi_2\rangle\langle\Psi_2|H_{\text{ep}}^{I'}|\Psi_0\rangle}{\xi_{\mathbf{k}_4} - \xi_{\mathbf{k}_2} - \hbar\omega_{-\mathbf{p},\chi}} \right), \end{aligned} \quad (\text{S25})$$

where  $I = (\eta, \zeta, s)$  and  $I' = (\eta', \zeta', s')$  are notations for the Dirac cone indices of the two electrons. Since the electron phonon interaction  $H_{\text{ep}}^I$  is independent of spin  $s$ , the above electron-electron interaction is independent of  $s$  and  $s'$ . Besides, in the continuum model  $H_{\text{ep}}^I$  is also independent of Moiré valley  $\zeta$  (see Eq. (S21)). Since our goal is to study the BCS superconductivity, we shall focus on the Cooper channel of the interaction, namely, we shall set the momenta of the two electrons as opposite to each other,  $\mathbf{k}_1 = -\mathbf{k}_2 = \mathbf{k}$  and  $\mathbf{k}_3 = -\mathbf{k}_4 = \mathbf{k}'$ . The Cooper channel electron-electron interaction is then

$$\begin{aligned} \frac{V_{\mathbf{k}\mathbf{k}'}^{\eta\eta'}(\omega)}{N_s\Omega_s} &= \sum_{\chi} \left[ \frac{\langle\mathbf{k}'_{\eta,\zeta,s}, -\mathbf{k}'_{\eta',\zeta',s'}|H_{\text{ep}}^{\eta',\zeta',s'}|\mathbf{k}'_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}, \mathbf{p}, \chi\rangle\langle\mathbf{k}'_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}, \mathbf{p}, \chi|H_{\text{ep}}^{\eta,\zeta,s}|\mathbf{k}_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}\rangle}{\hbar\omega - \hbar\omega_{\mathbf{p},\chi}} \right. \\ &\quad \left. + \frac{\langle\mathbf{k}'_{\eta,\zeta,s}, -\mathbf{k}'_{\eta',\zeta',s'}|H_{\text{ep}}^{\eta,\zeta,s}|\mathbf{k}_{\eta,\zeta,s}, -\mathbf{k}'_{\eta',\zeta',s'}, -\mathbf{p}, \chi\rangle\langle\mathbf{k}_{\eta,\zeta,s}, -\mathbf{k}'_{\eta',\zeta',s'}, -\mathbf{p}, \chi|H_{\text{ep}}^{\eta',\zeta',s'}|\mathbf{k}_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}\rangle}{-\hbar\omega - \hbar\omega_{\mathbf{p},\chi}} \right], \end{aligned} \quad (\text{S26})$$

where  $\omega = (\xi_{\mathbf{k}'} - \xi_{\mathbf{k}})/\hbar$ , and we have replaced  $\omega_{-\mathbf{p},\chi}$  by  $\omega_{\mathbf{p},\chi}$  since they equal to each other. Accordingly, in the second quantized language, the phonon-mediated electron-electron interaction in the Cooper channel is

$$H_{\text{int}}^{(\text{ph})} = \frac{1}{N_s\Omega_s} \sum_{\mathbf{k}, \mathbf{k}'} \sum_{\eta, \eta'} \sum_{\zeta, \zeta'} \sum_{s, s'} V_{\mathbf{k}\mathbf{k}'}^{\eta\eta'}(\omega) c_{\mathbf{k}', \eta, \zeta, s}^\dagger c_{-\mathbf{k}', \eta', \zeta', s'}^\dagger c_{-\mathbf{k}, \eta', \zeta', s'} c_{\mathbf{k}, \eta, \zeta, s}. \quad (\text{S27})$$

We now calculate  $V_{\mathbf{k}\mathbf{k}'}^{\eta\eta'}$  using the projected electron-phonon Hamiltonian  $H_{\text{ep}}^{\eta,\zeta,s}$  which we obtained in Eq. (S24). By Eqs. (S17) and (S24), we find the matrix elements for intermediate states (defined as  $|\mathbf{k}_I, \mathbf{k}'_{I'}, \mathbf{p}, \chi\rangle = c_{\mathbf{k}, I}^\dagger c_{\mathbf{k}', I'}^\dagger a_{\mathbf{p}, \chi}^\dagger |0\rangle$ ) with phonon polarizations  $\chi = L$  and  $\chi = T$  are:

$$\begin{aligned} \langle\mathbf{k}'_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}, \mathbf{p}, T|H_{\text{ep}}^{\eta,\zeta,s}|\mathbf{k}_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}\rangle &= \langle\mathbf{k}'_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}|H_{\text{ep}}^{\eta,\zeta,s}|\mathbf{k}_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}, -\mathbf{p}, T\rangle \\ &= \frac{\gamma\hbar vk_F}{\sqrt{N_s\Omega_s}} \sqrt{\frac{\hbar p\Omega}{2Mc_T}} \phi_{\mathbf{k}'}^{\eta\dagger} \phi_{\mathbf{k}}^\eta \left[ g_{2\alpha} - \eta \frac{g_{1\alpha}}{2} \text{Re} \left( \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right) \right], \\ \langle\mathbf{k}'_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}, \mathbf{p}, L|H_{\text{ep}}^{\eta,\zeta,s}|\mathbf{k}_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}\rangle &= \langle\mathbf{k}'_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}|H_{\text{ep}}^{\eta,\zeta,s}|\mathbf{k}_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}, -\mathbf{p}, L\rangle \\ &= \frac{\gamma\hbar vk_F}{\sqrt{N_s\Omega_s}} \sqrt{\frac{\hbar p\Omega}{2Mc_L}} \phi_{\mathbf{k}'}^{\eta\dagger} \phi_{\mathbf{k}}^\eta \times \eta \frac{g_{1\alpha}}{2} \text{Im} \left( \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right), \\ \langle\mathbf{k}_{\eta,\zeta,s}, -\mathbf{k}'_{\eta',\zeta',s'}, -\mathbf{p}, T|H_{\text{ep}}^{\eta',\zeta',s'}|\mathbf{k}_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}\rangle &= \langle\mathbf{k}'_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}|H_{\text{ep}}^{\eta',\zeta',s'}|\mathbf{k}'_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}, \mathbf{p}, T\rangle \\ &= \frac{\gamma\hbar vk_F}{\sqrt{N_s\Omega_s}} \sqrt{\frac{\hbar p\Omega}{2Mc_T}} \phi_{-\mathbf{k}'}^{\eta'\dagger} \phi_{-\mathbf{k}}^{\eta'} \left[ g_{2\alpha} + \eta' \frac{g_{1\alpha}}{2} \text{Re} \left( \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right) \right], \\ \langle\mathbf{k}_{\eta,\zeta,s}, -\mathbf{k}'_{\eta',\zeta',s'}, -\mathbf{p}, L|H_{\text{ep}}^{\eta',\zeta',s'}|\mathbf{k}_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}\rangle &= \langle\mathbf{k}'_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}|H_{\text{ep}}^{\eta',\zeta',s'}|\mathbf{k}'_{\eta,\zeta,s}, -\mathbf{k}_{\eta',\zeta',s'}, \mathbf{p}, L\rangle \\ &= -\frac{\gamma\hbar vk_F}{\sqrt{N_s\Omega_s}} \sqrt{\frac{\hbar p\Omega}{2Mc_L}} \phi_{-\mathbf{k}'}^{\eta'\dagger} \phi_{-\mathbf{k}}^{\eta'} \times \eta' \frac{g_{1\alpha}}{2} \text{Im} \left( \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right), \end{aligned} \quad (\text{S28})$$

where we have used the fact that  $\varphi_{-\mathbf{k}} = \varphi_{\mathbf{k}} + \pi$ , and recall that  $\Omega$  and  $\Omega_s$  are the graphene unit cell area and superlattice unit cell area, respectively. Since  $H_{\text{ep}}^{\eta,\zeta,s}$  is proportional to  $u_{\mathbf{p},\chi} \propto (a_{\mathbf{p},\chi} + a_{-\mathbf{p},\chi}^\dagger)$  (see Eq. (S17)), the

amplitude of creating a phonon state  $|\mathbf{p}, \chi\rangle$  is always the same as that of annihilating a phonon state  $|\mathbf{p}, \chi\rangle$  (when the other quantum numbers are the same), so we have the above equal relations between every two matrix amplitudes. To further simplify the result, we approximate  $c_L \approx c_T$ , both of which are of the order of magnitude  $10^4 \text{m/s}$ , so that  $\omega_{\mathbf{p},L} \approx \omega_{\mathbf{p},T} = c_T p$ . Under this approximation, the prefactors of all the matrix elements in Eq. (S28) become identical, i.e.,  $\sqrt{\frac{\hbar p \Omega}{2M c_T}} = \sqrt{\frac{\hbar p \Omega}{2M c_L}}$ . By defining

$$\varpi_{\mathbf{k}\mathbf{k}'}^{\eta\eta'} = \phi_{\mathbf{k}}^{\eta\dagger} \phi_{\mathbf{k}}^{\eta} \phi_{-\mathbf{k}}^{\eta'\dagger} \phi_{-\mathbf{k}}^{\eta'} = e^{i(\frac{\eta+\eta'}{2})(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})} \left[ \frac{1 + \cos(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})}{2} \right], \quad (\text{S29})$$

one can then show the interaction in Eq. (S26) is

$$\begin{aligned} \frac{V_{\mathbf{k}\mathbf{k}'}^{\eta\eta'}(\omega)}{\Omega_s} &= \frac{\gamma^2 \hbar^2 v^2 \Omega k_F^2 p \varpi_{\mathbf{k}\mathbf{k}'}^{\eta\eta'}}{2M c_T \Omega_s} \left( \frac{1}{\omega - \omega_{\mathbf{p},T}} + \frac{1}{-\omega - \omega_{\mathbf{p},T}} \right) \times \\ &\left\{ \left[ g_{2\alpha} - \eta \frac{g_{1\alpha}}{2} \text{Re} \left( \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right) \right] \left[ g_{2\alpha} + \eta' \frac{g_{1\alpha}}{2} \text{Re} \left( \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right) \right] - \eta\eta' \frac{g_{1\alpha}^2}{4} \left[ \text{Im} \left( \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right) \right]^2 \right\} \\ &= \frac{\gamma^2 \hbar^2 v^2 \Omega k_F^2 \varpi_{\mathbf{k}\mathbf{k}'}^{\eta\eta'}}{2M c_T^2 \Omega_s} \frac{\omega_{\mathbf{p},T}^2}{\omega^2 - \omega_{\mathbf{p},T}^2} \left\{ g_{2\alpha}^2 + \frac{g_{1\alpha} g_{2\alpha}}{2} (\eta' - \eta) \text{Re} \left( \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right) - \eta\eta' \frac{g_{1\alpha}^2}{4} \left| \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right|^2 \right\}. \end{aligned} \quad (\text{S30})$$

Therefore, the intra valley interaction with  $\eta = \eta'$  is

$$\begin{aligned} \frac{V_{\mathbf{k}\mathbf{k}'}^{\eta\eta}}{\Omega_s} &= \frac{\gamma^2 \hbar^2 v^2 \Omega k_F^2 \varpi_{\mathbf{k}\mathbf{k}'}^{\eta\eta}}{M c_T^2 \Omega_s} \frac{\omega_{\mathbf{p},T}^2}{\omega^2 - \omega_{\mathbf{p},T}^2} \left\{ g_{2\alpha}^2 - \frac{g_{1\alpha}^2}{2} [1 + \cos(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})] \right\} \\ &= -\mathcal{N}_0(\omega) e^{i\eta(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})} [1 + \cos(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})] \left\{ g_{2\alpha}^2 - \frac{g_{1\alpha}^2}{2} [1 + \cos(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})] \right\}, \end{aligned} \quad (\text{S31})$$

where we have used the fact that  $\left| \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right|^2 = \frac{2-2\cos 2(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})}{2-2\cos(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})} = 2[1 + \cos(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})]$ , and we have defined  $\mathcal{N}_0(\omega) = \frac{\gamma^2 \hbar^2 v^2 \Omega k_F^2}{M c_T^2 \Omega_s} \frac{\omega_{\mathbf{p},T}^2}{\omega_{\mathbf{p},T}^2 - \omega^2}$ , and used Eq. (S29) for the expression of  $\varpi_{\mathbf{k}\mathbf{k}'}^{\eta\eta'}$ . We note that for low frequency processes with  $|\omega| < \omega_{\mathbf{p},T}$ , we have the coefficient  $\mathcal{N}_0(\omega) > 0$ . In a similar way, we find the inter valley interaction with  $\eta = -\eta' = 1$  is

$$\begin{aligned} \frac{V_{\mathbf{k}\mathbf{k}'}^{KK'}}{\Omega_s} &= \frac{\gamma^2 \hbar^2 v^2 \Omega k_F^2 \varpi_{\mathbf{k}\mathbf{k}'}^{KK'}}{M c_T^2 \Omega_s} \frac{\omega_{\mathbf{p},T}^2}{\omega^2 - \omega_{\mathbf{p},T}^2} \left| g_{2\alpha} - \frac{g_{1\alpha}}{2} \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right|^2 \\ &= -\mathcal{N}_0(\omega) [1 + \cos(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})] \left| g_{2\alpha} - \frac{g_{1\alpha}}{2} \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right|^2. \end{aligned} \quad (\text{S32})$$

For  $\theta$  near magic angle  $\alpha^2 \approx 1/3$ , one has  $g_{1\alpha} \approx 2/3$  and  $g_{2\alpha} \approx 1/3$ , and the interaction takes the form shown in main text Eq. (8).

In particular, for low frequencies  $|\omega| < \omega_{\mathbf{p},T}$ , one finds that the intervalley interaction is attractive, while the intravalley interaction is repulsive. To see this, consider two electrons (near the Fermi surface) at valley  $\eta$  and  $\eta'$ , whose 2-body wave function (with total momentum zero assumed) can generically be written as

$$|\Psi_{\eta\eta'}\rangle = \sum_{l \in \mathbb{Z}} \sum_{|\mathbf{k}|=k_F} \beta_l e^{il\varphi_{\mathbf{k}}} c_{\mathbf{k},\eta,\zeta,s}^\dagger c_{-\mathbf{k},\eta',\zeta',s'}^\dagger |0\rangle,$$

where the physical meaning of  $l \in \mathbb{Z}$  is the relative angular momentum between the two electrons. For two electrons

in the same valley  $\eta = \eta' = K$  (with wave function  $|\Psi_{KK}\rangle$ ), the intravalley interaction energy is given by

$$\begin{aligned}
E_{KK} &= \langle \Psi_{KK} | H_{\text{int}}^{(\text{ph})} | \Psi_{KK} \rangle = \int_0^{2\pi} \frac{d\varphi_{\mathbf{k}}}{2\pi} \int_0^{2\pi} \frac{d\varphi_{\mathbf{k}'}}{2\pi} \sum_{l \in \mathbb{Z}} \sum_{l' \in \mathbb{Z}} \frac{V_{\mathbf{k}\mathbf{k}'}^{KK}}{\Omega_s} \beta_l^* \beta_{l'} e^{il\varphi_{\mathbf{k}} - il'\varphi_{\mathbf{k}'}} \\
&= -\mathcal{N}_0(\omega) e^{i(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})} \int_0^{2\pi} \frac{d\varphi_{\mathbf{k}}}{2\pi} \int_0^{2\pi} \frac{d\varphi_{\mathbf{k}'}}{2\pi} [1 + \cos(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})] \left\{ g_{2\alpha}^2 - \frac{g_{1\alpha}^2}{2} [1 + \cos(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})] \right\} \beta_l^* \beta_{l'} e^{il\varphi_{\mathbf{k}} - il'\varphi_{\mathbf{k}'}} \\
&= \mathcal{N}_0(\omega) \left[ \left( \frac{3}{4} g_{1\alpha}^2 - g_{2\alpha}^2 \right) |\beta_{-1}|^2 + \frac{1}{2} (g_{1\alpha}^2 - g_{2\alpha}^2) (|\beta_{-2}|^2 + |\beta_0|^2) + \frac{g_{1\alpha}^2}{4} (|\beta_{-3}|^2 + |\beta_1|^2) \right] \\
&\approx \mathcal{N}_0(\omega) \left[ \frac{2}{9} |\beta_{-1}|^2 + \frac{1}{6} (|\beta_{-2}|^2 + |\beta_0|^2) + \frac{1}{9} (|\beta_{-3}|^2 + |\beta_1|^2) \right],
\end{aligned} \tag{S33}$$

where in the last line we have assumed  $\theta$  is near the first magic angle so that  $g_{1\alpha} \approx 2/3$  and  $g_{2\alpha} \approx 1/3$ . For  $|\omega| < \omega_{\mathbf{p},T}$ , one has  $\mathcal{N}_0(\omega) > 0$ , and one finds the intravalley interaction energy is in fact always positive, i.e.,  $E_{KK} > 0$ . Therefore, two electrons in the same valley always repulse each other (similarly one can show this for  $\eta = \eta' = K'$ ).

In contrast, one notes that the intervalley interaction  $V_{\mathbf{k}\mathbf{k}'}^{KK'}$  in Eq. (S32) is always real and negative for  $|\omega| < \omega_{\mathbf{p},T}$ , i.e.,  $\mathcal{N}_0(\omega) > 0$ . Therefore, if one considers two electrons in valley  $\eta = K$  and  $\eta' = K'$ , respectively, one has the intervalley interaction energy

$$E_{KK'} = \langle \Psi_{KK'} | H_{\text{int}}^{(\text{ph})} | \Psi_{KK'} \rangle < 0, \tag{S34}$$

which is attractive (for  $|\omega| < \omega_{\mathbf{p},T}$ ). Therefore, we conclude the intervalley Cooper pairing will be favored.

Physically, the fact that the intervalley interaction is more attractive than intravalley interaction can be understood as follows (see also the paragraph below main text Eq. (8)). Recall that the electron-phonon interaction projected to the vicinity of the Fermi surface contains two parts,  $H_{\text{ep}}^{\eta,\zeta,s}(\mathbf{k}) = H_{C3}^{\eta,\zeta,s}(\mathbf{k}) + H_{SO(2)}^{\eta,\zeta,s}(\mathbf{k})$ , where  $H_{C3}^{\eta,\zeta,s}$  and  $H_{SO(2)}^{\eta,\zeta,s}$  are odd and even with respect to the electron momentum  $\bar{\mathbf{k}}$  (average momentum before and after phonon emission/absorption), respectively. Now we consider an electron wave packet state with average momentum  $\bar{\mathbf{k}}$  as  $|\bar{\mathbf{k}}_{\eta,\zeta,s}\rangle = \sum_{\mathbf{k}} \beta(\mathbf{k} - \bar{\mathbf{k}}) c_{\mathbf{k},\eta,\zeta,s}^\dagger |0\rangle$ , where  $\beta(\mathbf{k} - \bar{\mathbf{k}})$  is a wave packet function peaked at  $\bar{\mathbf{k}}$ . Under a certain lattice deformation (i.e., given a nonzero configuration of the displacement field  $\mathbf{u}(\mathbf{r})$ ), the electron-phonon interaction will generate a background lattice potential, which we assume is  $U_{\text{ep}} = U_{C3} + U_{SO(2)}$  for an electron  $|\bar{\mathbf{k}}_{K,\zeta,s}\rangle$  at valley  $K$ , where  $U_{C3} = \langle \bar{\mathbf{k}}_{K,\zeta,s} | H_{C3}^{K,\zeta,s} | \bar{\mathbf{k}}_{K,\zeta,s} \rangle$ , and  $U_{SO(2)} = \langle \bar{\mathbf{k}}_{K,\zeta,s} | H_{SO(2)}^{K,\zeta,s} | \bar{\mathbf{k}}_{K,\zeta,s} \rangle$ . Since  $H_{C3}^{\eta,\zeta,s}$  and  $H_{SO(2)}^{\eta,\zeta,s}$  are odd and even with respect to  $\bar{\mathbf{k}}$ , an electron state  $|\bar{\mathbf{k}}_{K,\zeta,s}\rangle$  in the same valley  $K$  with opposite momentum would feel a background lattice potential  $U'_{\text{ep}} = -U_{C3} + U_{SO(2)}$ . On the other hand, an electron state  $|\bar{\mathbf{k}}_{K',\zeta,s}\rangle$  in the other valley  $K'$  would feel the same potential  $U_{\text{ep}} = U_{C3} + U_{SO(2)}$  as that of the electron  $|\bar{\mathbf{k}}_{K,\zeta,s}\rangle$ .

In general, if two electrons feel the same background potential due to deformed lattice, they will tend to get closer in the space, thus leading to a phonon-mediated electron-electron attraction. In the above, we have shown that two electrons  $|\bar{\mathbf{k}}_{K,\zeta,s}\rangle$  and  $|\bar{\mathbf{k}}_{K',\zeta,s}\rangle$  in opposite valleys  $K$  and  $K'$  feel the same background potential  $U_{C3} + U_{SO(2)}$ , thus an effective attraction will be produced between them. Instead, two electrons  $|\bar{\mathbf{k}}_{K,\zeta,s}\rangle$  and  $|\bar{\mathbf{k}}_{K,\zeta,s}\rangle$  in the same valley feel different background potentials  $U_{C3} + U_{SO(2)}$  and  $-U_{C3} + U_{SO(2)}$ , thus the effective attraction between them will be weaker or even absent. Therefore, the intervalley interaction is more attractive than the intravalley interaction.

By substituting the realistic system parameters  $v \approx 10^6$  m/s,  $c_T \approx 10^4$  m/s,  $w = 110$  meV,  $a_0 = 0.246$  nm, and  $\theta \approx 1.05^\circ$  into the phonon-induced inter valley interaction (Eq. (S32)), and take  $k_F \approx k_\theta$  as an order of magnitude estimation, we find the interaction contributed by the acoustic MBZ phonon bands near magic angle is

$$\frac{V_{\mathbf{k}\mathbf{k}'}^{KK'}(0)}{\Omega_s} \sim -\frac{4\gamma^2 \hbar^2 v^2 \Omega k_\theta^2}{9Mc_T^2 \Omega_s} \approx -\frac{4\hbar^2 v^2 k_\theta^2}{9Mc_T^2} \sim -1\text{meV}, \tag{S35}$$

which is comparable to both the electron band width around the first magic angle (of order  $1 \sim 10\text{meV}$ ) and the band width of the lowest acoustic MBZ phonon bands  $\hbar\omega_D \sim \hbar c_T k_\theta \approx 2\text{meV}$ . Therefore, the electron phonon coupling is relatively strong.

Furthermore, when the optical phonon bands are taken into account, the total phonon induced attractive interaction should be further enhanced. Since our electron phonon coupling  $H_{\text{ep}}$  is derived in the long wavelength phonon limit, it does not apply for optical phonon bands. Nevertheless, here we give a very rough discussion of the optical phonon band contributions to the electron-electron interaction. A very crude approximation is to assume the electron phonon coupling matrix elements  $\mathcal{M}_{\text{ep},\chi}$  of all optical phonon bands  $\chi$  are roughly the same as that of the lowest acoustic

bands. As we have discussed, the optical phonon bands are obtained by folding the phonon bands in the graphene BZ of the two layers into the MBZ. Equivalently, before folding, we can say that the lowest acoustic phonon bands are in the first MBZ, while the post-folding optical phonon bands are in the second, third and higher MBZ. Since the  $n$ -th MBZ is roughly  $nk_\theta$  away from  $\Gamma$  point of the first MBZ, the energy of the optical phonon band in the  $n$ -th MBZ is roughly  $n\hbar c_T k_\theta$ . Besides, the number of  $n$ -th MBZs is roughly  $2\pi n$  (roughly speaking, all the  $n$ -th MBZs together form a circle at radius  $nk_\theta$ ), so there are roughly  $2\pi n$  transverse and longitudinal polarized optical phonon bands with energy  $n\hbar c_T k_\theta$ . The upper limit of  $n$  is around  $1/\theta$ , where  $nk_\theta$  reaches the boundary of the original graphene BZ. The modified total electron-electron interaction will then be enhanced to

$$V_{\mathbf{k}\mathbf{k}'}^{\text{tot}} \sim \sum_{\chi} \frac{|\mathcal{M}_{\text{ep},\chi}|^2}{-\hbar\omega_{\mathbf{p},\chi}} \sim \sum_{n=1}^{1/\theta} \frac{2\pi n}{n} \frac{|\mathcal{M}_{\text{ep},T}|^2 + |\mathcal{M}_{\text{ep},L}|^2}{-\hbar c_T k_\theta} \approx \frac{2\pi}{\theta} V_{\mathbf{k}\mathbf{k}'}^{\text{ac}},$$

where  $\chi$  here denotes all the MBZ phonon bands (acoustic and optical),  $\omega_{\mathbf{p},\chi}$  is the eigenfrequency of phonon band  $\chi$ , and  $V_{\mathbf{k}\mathbf{k}'}^{\text{ac}}$  is the interaction contributed solely by the acoustic phonon bands we derived earlier in this section. Namely, the optical phonon bands contribute an additional factor  $2\pi/\theta$  to the electron-electron interaction if all the optical phonon bands contribute, which is a large number for small angles  $\theta$ . This is clearly an overestimation, since high energy optical phonon bands are expected not to participate in the low energy superlattice physics. In practice, we expect this factor of interaction enhancement by optical phonons to be much smaller than  $2\pi/\theta$  (but obviously greater than 1).

### S3. SCREENED COULOMB POTENTIAL AND BCS SUPERCONDUCTIVITY AT MAGIC ANGLE

The phonon-mediated electron-electron interaction can induce conventional BCS superconductivity. The density of states of the lowest bands at magic angle is around  $N_D \gtrsim 1\text{meV}^{-1} \cdot \Omega_s^{-1}$  for a 1 meV band width, which gives a BCS coupling strength  $\lambda \approx N_D |V_{\mathbf{k}\mathbf{k}'}(0)| \gtrsim 1$ , which is relatively strong.

To determine the superconductor critical temperature, we also need to estimate the screened Coulomb potential between electrons. Here we shall simply adopt the Thomas-Fermi approximation in two dimensions (2D) for an order of magnitude estimation, and do not discuss the accuracy of the approximation. The Coulomb potential without screening is given by  $V_e(r) = e^2/\epsilon_I r$ , whose Fourier transform is  $\mathcal{V}_e(q) = 2\pi e^2/\epsilon_I q$ , where  $q = |\mathbf{q}|$  is the Fourier wave vector, and  $\epsilon_I \approx 2 \sim 10$  is the dielectric constant of a charge neutral graphene. The 2D Poisson's equation for the bare Coulomb potential can then be written as

$$q\epsilon_I \mathcal{V}_e(q) = 2\pi\rho(q), \quad (\text{S36})$$

where  $\rho(q)$  is the bare charge density. In the presence of free electrons, a Coulomb potential  $V_e$  will induce a local charge density  $\rho_f(r) \approx -e^2(\partial n_e/\partial\mu)V_e(r)$ , or in Fourier space  $\rho_f(q) \approx -e^2(\partial n_e/\partial\mu)\mathcal{V}_e(q)$ , where  $n_e$  is the electron number density,  $\mu$  is the chemical potential, and thus  $\partial n_e/\partial\mu$  is the density of states  $N_D$ . One can then write the charge density as  $\rho = \rho_s + \rho_f$ , where  $\rho_s$  is the screened charge. From the Poisson's equation we have

$$2\pi\rho_s(q) = q\epsilon_I \mathcal{V}_e(q) - 2\pi\rho_f(q) = [q\epsilon_I + 2\pi e^2(\partial n_e/\partial\mu)]\mathcal{V}_e(q) = q\epsilon(q)\mathcal{V}_e(q), \quad (\text{S37})$$

so we find the screened dielectric function of the form

$$\epsilon(q) \approx \epsilon_I \left( 1 + \frac{q_{\text{TF}}}{q} \right), \quad (\text{S38})$$

where  $q_{\text{TF}} = 2\pi e^2(\partial n_e/\partial\mu)/\epsilon_I = 2\pi e^2 N_D/\epsilon_I$  is the Thomas-Fermi screening momentum. Near the magic angle,  $N_D \gtrsim 1\text{meV}^{-1} \cdot \Omega_s^{-1}$ , which yields a  $q_{\text{TF}} \gtrsim 50k_\theta \gg q$ , so the screened Coulomb potential is approximately  $\mathcal{V}_e(q) \approx 2\pi e^2/\epsilon_I q_{\text{TF}} \sim N_D^{-1}$ , and the Coulomb coupling strength  $\mu_c \approx N_D \mathcal{V}_e(q) \sim 1$ . We note that due to large density of states  $N_D$ , the screened Coulomb potential is much smaller than the bare Coulomb potential. In particular, the screened Coulomb potential is comparable to  $V_{\mathbf{k}\mathbf{k}'}$ . However, the phonon induced attraction  $V_{\mathbf{k}\mathbf{k}'}(\omega)$  is frequency  $\omega$  dependent and large as  $\omega$  approaches  $\omega_{T,\mathbf{p}}$ , while the screened Coulomb potential can be approximated as frequency independent. This is because the 2D plasma frequency  $\hbar\omega_{pe} = \sqrt{(4\pi n_e)^{1/2} e^2 \epsilon_F/\epsilon_I} \gtrsim 20\text{meV}$  at the magic angle is much larger than the Debye frequency  $\hbar\omega_D \approx \hbar c_T k_\theta \approx 2\text{meV}$ , where  $\epsilon_F$  is the Fermi energy (of order  $1 \sim 10\text{meV}$  around magic angle) [4]. We then adopt the McMillan formula [5, 6] to give a proper estimation of the BCS superconductivity critical temperature taking into account both the phonon induced attraction and the Coulomb repulsion:

$$k_B T_c = \frac{\hbar\omega_D}{1.45} \exp \left[ -\frac{1.04(1+\lambda)}{\lambda - \mu_c^*(1+0.62\lambda)} \right], \quad (\text{S39})$$

where  $\mu_c^* = \mu_c/[1 + \mu_c \ln(\omega_{pe}/\omega_D)]$  is the reduced Coulomb coupling strength, and  $\omega_{pe}/\omega_D \gtrsim 10$  around the magic angle. If we take the BCS coupling strength  $\lambda \approx N_D|V_{\mathbf{k}\mathbf{k}'}| \approx 1.5$ , and  $\mu_c \approx 1$ , we find the superconductivity critical temperature from the McMillan formula is  $T_c \approx 0.9K$ , close to the experimentally measured value.

Next, we discuss the pairing amplitude of the inter valley pairing. The pairing amplitude is defined as

$$\Delta_{\mathbf{k}}^{\eta\eta',\zeta\zeta',ss'} = \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'}^{\eta\eta'} \langle c_{-\mathbf{k}',\eta',\zeta',s'} c_{\mathbf{k}',\eta,\zeta,s} \rangle, \quad (\text{S40})$$

where  $V_{\mathbf{k}\mathbf{k}'}^{\eta\eta'}$  is as given in Eq. (S30). At zero temperature, the BCS self-consistency gap equation is given by

$$\Delta_{\mathbf{k}}^{\eta\eta',\zeta\zeta',ss'} = -\frac{1}{2} \sum_{|\xi_{\mathbf{k}'}| < \hbar\omega_D} \frac{V_{\mathbf{k}\mathbf{k}'}^{\eta\eta'} \Delta_{\mathbf{k}'}^{\eta\eta',\zeta\zeta',ss'}}{E_{\mathbf{k}'}^{BdG}}, \quad (\text{S41})$$

where  $E_{\mathbf{k}'}^{BdG}$  is the Bogoliubov-de Gennes (BdG) band energy, while  $\xi_{\mathbf{k}} = \hbar v_F(|\mathbf{k}| - k_F)$  is the band energy with the Fermi velocity  $v_F = -\frac{1-3\alpha^2}{1+6\alpha^2}v$ . Since we have shown the intervalley pairing is favored, we shall ignore the intravalley interaction (which is in fact repulsive as we have shown in Eq. (S33)), and only keep the intervalley interaction  $V_{\mathbf{k}\mathbf{k}'}^{KK'}$ . Furthermore, for simplicity we shall adopt the BCS approximation which assumes the electron-electron interaction is a constant function of  $\omega$  for frequency  $|\omega| < \omega_D$  and zero otherwise ( $\omega_D$  is the Debye frequency), namely, we approximate the frequency dependence factor  $\frac{\omega_{\mathbf{p},T}^2}{\omega^2 - \omega_{\mathbf{p},T}^2}$  in Eq. (S32) to  $-1$  for  $|\omega| < \omega_D$ , and  $0$  for  $|\omega| \geq \omega_D$ . Qualitatively, this does not affect the shape of the pairing function. We shall also assume  $\theta$  is near the magic angle, so that  $\alpha^2 \approx 1/3$ , and accordingly  $g_{1\alpha} \approx 2/3$  and  $g_{2\alpha} \approx 1/3$ . The intervalley interaction in Eq. (S32) is then approximated as

$$\begin{aligned} V_{\mathbf{k}\mathbf{k}'}^{KK',\zeta\zeta',ss'} &\approx -V_0 \varpi_{\mathbf{k}\mathbf{k}'}^{KK'} \left| 1 - \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right|^2 \\ &= -V_0 \left[ \frac{1 + \cos(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})}{2} \right] \left| 1 - \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right|^2, \end{aligned} \quad (\text{S42})$$

where  $V_0 \approx \frac{\gamma^2 \hbar^2 v^2 \Omega k_F^2}{9Mc_F^2 \Omega_s}$ , and we have used Eq. (S29).

With the 4-fold degeneracy from Moiré valley  $\zeta$  and spin  $s$ , the pairing can be either Moiré valley triplet spin singlet, or Moiré valley singlet spin triplet. Here we shall simply assume the pairing is time-reversal invariant (which is more robust than time-reversal violating pairings in the presence of disorder [7]). Since the time-reversal symmetry  $\mathcal{T}$  brings valley  $K \rightarrow K'$ , Moiré valley  $K_M \rightarrow K'_M$ , and spin  $s \rightarrow -s$  (see Eq. (S18)), a time reversal invariant pairing will be between opposite spins and opposite Moiré valleys. Such a time reversal invariant intervalley pairing then takes the form

$$\Delta_{\mathbf{k}}^{\eta\eta',\zeta\zeta',ss'} \sim s \delta_{s,-s'} \delta_{\zeta,-\zeta'} \delta_{\eta,-\eta'} \tilde{\Delta}(\eta\mathbf{k}), \quad (\text{S43})$$

where  $\tilde{\Delta}(\mathbf{k})$  satisfies

$$\begin{aligned} \tilde{\Delta}(\mathbf{k}) &= N_D V_0 \int_{-\hbar\omega_D}^{\hbar\omega_D} d\xi \int_0^{2\pi} \frac{d\varphi_{\mathbf{k}'}}{2\pi} \left[ \frac{1 + \cos(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})}{2} \right] \left| 1 - \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right|^2 \frac{\tilde{\Delta}(\mathbf{k}')}{2\sqrt{|\tilde{\Delta}(\mathbf{k}')|^2 + \xi^2}} \\ &= N_D V_0 \int_0^{2\pi} \frac{d\varphi_{\mathbf{k}'}}{2\pi} \left[ \frac{1 + \cos(\varphi_{\mathbf{k}} - \varphi_{\mathbf{k}'})}{2} \right] \left| 1 - \frac{e^{2i\varphi_{\mathbf{k}}} - e^{2i\varphi_{\mathbf{k}'}}}{e^{-i\varphi_{\mathbf{k}}} - e^{-i\varphi_{\mathbf{k}'}}} \right|^2 \tilde{\Delta}(\mathbf{k}') \sinh^{-1} \frac{\hbar\omega_D}{|\tilde{\Delta}(\mathbf{k}')|}, \end{aligned} \quad (\text{S44})$$

where in this case the BdG band energy is  $E_{\mathbf{k}'}^{BdG} = \sqrt{|\tilde{\Delta}(\mathbf{k}')|^2 + \xi_{\mathbf{k}'}^2}$ .

Eq. (S44) can then be numerically solved by iteration. As an example, for  $N_D V_0 = 0.5$ , we find the pairing amplitude  $\tilde{\Delta}(\mathbf{k})$  is real and has the shape as shown in the main text Fig. [3c]. For generic values of  $N_D V_0 > 0$ , we find the pairing amplitude is always real and nodeless, thus is dominated by  $s$ -wave pairing and is topologically trivial. Finally, we note that in the absence of disorders, the pairing is not necessarily time-reversal invariant, and the pairing amplitude could be either Moiré valley singlet spin triplet or Moiré valley singlet spin triplet, which are degenerate.

#### S4. NUMERICAL CALCULATION FOR OTHER ANGLES AND ELECTRON DENSITIES: PREDICTION OF OTHER SUPERCONDUCTING ANGLES

The above calculations of electron-phonon coupling can be generically applied to any twist angle  $\theta$  and electron density. We still adopt the continuum model with nearest hoppings [1], namely, an electron with momentum  $\mathbf{k}$  in layer 1 can hop with an electron with momentum  $\mathbf{p}'$  in layer 2 iff  $\mathbf{k} - \mathbf{p}' = \mathbf{q}_j$  ( $j = 1, 2, 3$ ). However, instead of truncating at the smallest four momenta ( $\mathbf{k}$  and  $\mathbf{k} - \mathbf{q}_j$ ) in Eq. (S4), we truncate the momentum of the Hamiltonian to sufficiently high momenta (momentum shells, i.e., second, third and higher MBZs), so that the band structure (which can be solved numerically) is more accurate. Besides, in the calculation we do not approximate  $\theta/2$  in  $h_{\pm\theta/2}^K(\mathbf{k})$  to zero, and this leads to a small particle-hole asymmetry to the band structure [8]. Fig. S4a and S4b show two examples of Moiré BZ band structure calculated from the continuum model for  $\theta = 1.05^\circ$  and  $\theta = 1.20^\circ$ , respectively. We can then calculate the density of states  $N_D$  from the band structure. In the main text Fig. [4a] we have shown  $N_D$  at  $\theta = 1.05^\circ$  as a function of number of electrons per superlattice unit cell  $n = n_e \Omega_s$ . Here in the upper panel of Fig. S4c, we show  $N_D$  at  $\theta = 1.20^\circ$  as another example. Fig. S4d shows the  $\text{Log}_{10}(N_D \cdot \text{meV} \cdot \Omega_s)$  in a wide range of  $\theta$  and  $n$ .

We then add small deformations to  $\mathbf{q}_j$  induced by  $\partial_a u_b$  ( $a, b = x, y$ ) (according to supplemental Eq. (S7)), and numerically calculate the band structure under different deformations. To first order, the change of band energy of a band  $E_m$  generically takes the form

$$\delta E_m(\mathbf{k}) = b_1(\mathbf{k}) \left( \frac{\partial_x u_y + \partial_y u_x}{2} \right) + b_2(\mathbf{k}) \left( \frac{\partial_x u_x - \partial_y u_y}{2} \right) + b_3(\mathbf{k}) \left( \frac{\partial_x u_y - \partial_y u_x}{2} \right) + b_4(\mathbf{k}) \left( \frac{\partial_x u_x + \partial_y u_y}{2} \right), \quad (\text{S45})$$

which is contributed by the band energy response to relative shear (first two terms), rotation and expansion between the two layers, respectively. This is nothing but the electron-phonon coupling of band  $E_m$  in the long wavelength limit. The four coefficients  $b_j$  ( $1 \leq j \leq 4$ ) are real and can be numerically extracted out from band structure variation under small deformations.

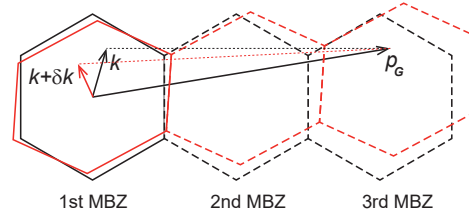


FIG. S3: A higher Moiré band state with Moiré momentum  $\mathbf{k}$  in the 1st Moiré BZ (MBZ, black solid hexagon) originates from the folding of the state with graphene momentum  $\mathbf{p}_G$  in the graphene BZ. Under the deformation  $\mathbf{q}_j \rightarrow \mathbf{q}_j + \delta \mathbf{q}_j$ , the MBZ is deformed (into the red hexagon); accordingly, the graphene momentum  $\mathbf{p}_G$  will be folded to a different Moiré momentum  $\mathbf{k} + \delta \mathbf{k}$ . This implies a redefinition of  $\mathbf{k}$  in higher Moiré bands.

One effect we need to exclude in calculating  $\delta E_m(\mathbf{k})$  is the following: when one makes a change  $\mathbf{q}_j \rightarrow \mathbf{q}_j + \delta \mathbf{q}_j$ , the Moiré BZ is deformed, so the folding of the original graphene BZ momentum into the Moiré BZ is also changed (Fig. S3). Therefore, for higher Moiré bands (which originates from folding of the graphene BZ), the definition of momentum  $\mathbf{k}$  in the Moiré BZ is changed under deformations. This additional change of electron momentum is simply a redefinition and is unphysical, thus should be eliminated. To be precise, consider the electron state in the  $m$ -th Moiré band (at graphene valley  $K$ ) which takes a generic form

$$|\mathbf{k}, m\rangle_M = \sum_{l_1, l_2 \in \mathbb{Z}} \sum_{j=1,2} W_{\mathbf{k}, l_1, l_2, j}^m |\mathbf{k} + l_1 \mathbf{g}_1 + l_2 \mathbf{g}_2 - (j-1) \mathbf{q}_1\rangle_j, \quad (\text{S46})$$

where  $\mathbf{k}$  is the momentum in the Moiré BZ,  $j$  is the layer index,  $\mathbf{g}_1 = \mathbf{q}_2 - \mathbf{q}_3$  and  $\mathbf{g}_2 = \mathbf{q}_3 - \mathbf{q}_1$  are the reciprocal vectors of the Moiré superlattice, and  $|\mathbf{p}\rangle_j$  is the basis of Dirac electron in layer  $j$  with an original graphene BZ momentum  $\mathbf{p}$  measured from graphene valley  $K$ . Besides,  $W_{\mathbf{k}, l_1, l_2, j}^m$  denotes the coefficient of basis  $|\mathbf{k} + l_1 \mathbf{g}_1 + l_2 \mathbf{g}_2 - (j-1) \mathbf{q}_1\rangle_j$ . Therefore, the state  $|\mathbf{k}, m\rangle_M$  carries an average graphene momentum

$$\langle \mathbf{p}_G \rangle = \langle \mathbf{k}, m | \hat{\mathbf{p}}_G | \mathbf{k}, m \rangle_M = \sum_{l_1, l_2 \in \mathbb{Z}} \sum_{j=1,2} |W_{\mathbf{k}, l_1, l_2, j}^m|^2 \left( \mathbf{k} + l_1 \mathbf{g}_1 + l_2 \mathbf{g}_2 - (j-1) \mathbf{q}_1 \right). \quad (\text{S47})$$

Upon the uniform deformation  $\mathbf{q}_j \rightarrow \mathbf{q}_j + \delta\mathbf{q}_j$ , one should keep the electron's graphene momentum  $\langle \mathbf{p}_G \rangle$  invariant and then examine the energy variation, because the absorption or emission of a long wave length phonon should not change the electron momentum. Namely, one need to shift the Moiré BZ momentum  $\mathbf{k}$  to  $\mathbf{k} + \delta\mathbf{k}$  so that

$$0 = \delta\langle \mathbf{p}_G \rangle \approx \delta\mathbf{k} + \sum_{l_1, l_2 \in \mathbb{Z}} \sum_{j=1,2} |W_{\mathbf{k}, l_1, l_2, j}^m|^2 \left( l_1 \delta\mathbf{g}_1 + l_2 \delta\mathbf{g}_2 - (j-1) \delta\mathbf{q}_1 \right), \quad (\text{S48})$$

where  $\delta\mathbf{g}_1 = \delta\mathbf{q}_2 - \delta\mathbf{q}_3$  and  $\delta\mathbf{g}_2 = \delta\mathbf{q}_3 - \delta\mathbf{q}_1$  (see Fig. S3). Since the wave function  $W_{\mathbf{k}, l_1, l_2, j}^m$  can be calculated numerically, we are able to compute  $\delta\mathbf{k}$ . Accordingly, one should evaluate the energy variation in Eq. (S45) as (Fig. S3)

$$\delta E_m(\mathbf{k}) = E'_m(\mathbf{k} + \delta\mathbf{k}) - E_m(\mathbf{k}), \quad (\text{S49})$$

where  $E_m(\mathbf{k})$  and  $E'_m(\mathbf{k})$  are the energy dispersion of the  $m$ -th band before and after deformation, respectively. We note that in our analytical calculations of electron-phonon coupling near the Dirac points of the lowest two Moiré flat bands (Eq. (S21)), one has  $\delta\mathbf{k} \approx -\frac{2\alpha^2}{1+6\alpha^2}(\delta\mathbf{q}_1 + \delta\mathbf{q}_2 + \delta\mathbf{q}_3) = 0$ , so we need not consider this momentum shift  $\delta\mathbf{k}$  problem there. For numerical calculations of higher bands, however,  $\delta\mathbf{k}$  is in general nonzero.

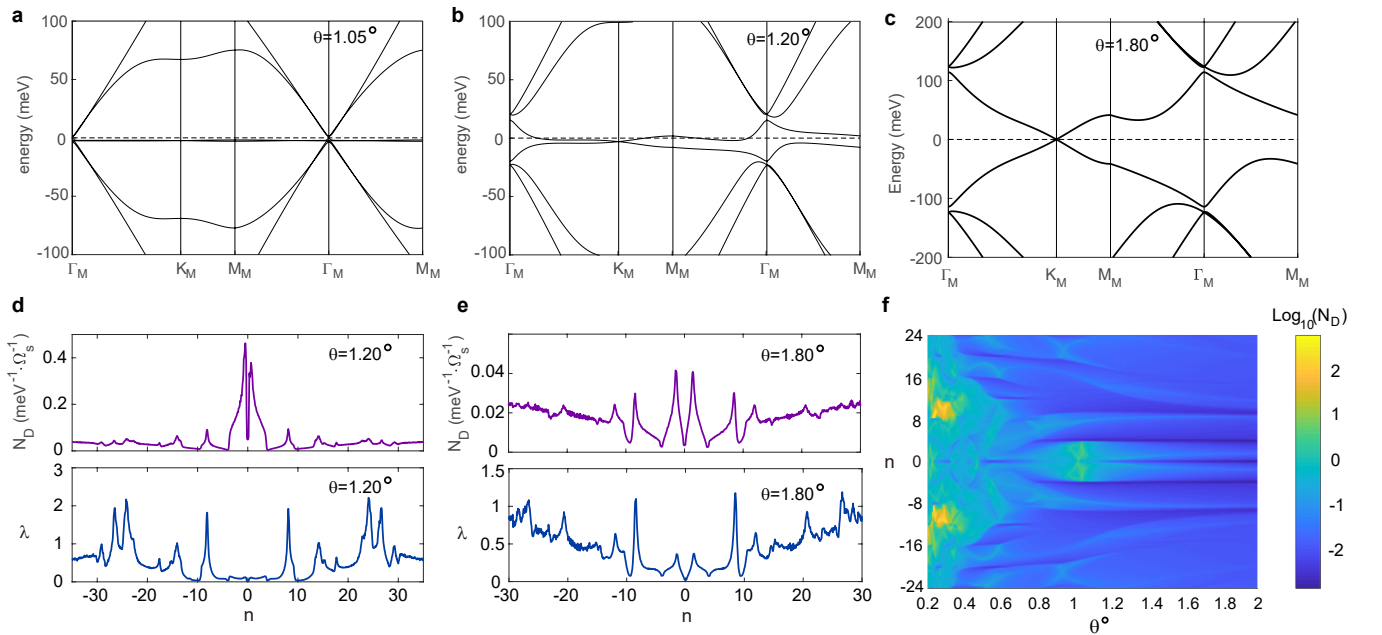


FIG. S4: **a.** TBG band structure at  $\theta = 1.05^\circ$  from the continuum model. **b.** TBG band structure at  $\theta = 1.20^\circ$  from the continuum model. **c.** TBG band structure at  $\theta = 1.80^\circ$  from the continuum model. **d.** Density of states  $N_D$  and BCS coupling strength  $\lambda$  with respect to  $n$  calculated for  $\theta = 1.20^\circ$ . In particular, the peaks in  $N_D$  around  $|n| = 8$  indicates a Van Hove singularity in the second or third conduction (valence) bands (the second band and third band have a significant overlap in energy, as shown in panels **a** and **b** here, so  $|n| = 8$  is not an in-gap filling factor). **e.** Density of states  $N_D$  and BCS coupling strength  $\lambda$  with respect to  $n$  calculated for  $\theta = 1.80^\circ$ . The shape of the  $N_D$  function agrees well with the STM measurement of large angle TBG [9]. **f.**  $\text{Log}_{10}(N_D \cdot \text{meV} \cdot \Omega_s)$  plotted as a function of  $\theta$  and  $n$ , where one can clearly see  $N_D$  is large at the magic angle, and is generically larger for smaller angles.

From Eq. (S17) we have the displacement field after quantization  $\mathbf{u}(\mathbf{r}) = \sum_{\mathbf{p}, \chi} \frac{e^{i\mathbf{p} \cdot \mathbf{r}} \boldsymbol{\epsilon}_\chi u_{\mathbf{p}, \chi}}{\sqrt{N_s \Omega_s}}$ , where  $\mathbf{p}$  is the phonon momentum,  $u_{\mathbf{p}, \chi} = \sqrt{\frac{\hbar \Omega}{2M\omega_{\mathbf{p}, \chi}}} (a_{\mathbf{p}, \chi} + a_{-\mathbf{p}, \chi}^\dagger) = \sqrt{\frac{\hbar \Omega}{2Mc_\chi p}} (a_{\mathbf{p}, \chi} + a_{-\mathbf{p}, \chi}^\dagger)$  for polarization  $\chi$ , and  $\boldsymbol{\epsilon}_\chi$  is the polarization vector. The electron-phonon coupling in band  $E_m$  is then simply the expression of  $\delta E_m(\bar{\mathbf{k}})$  in Eq. (S45) with the displacement field  $\mathbf{u}(\mathbf{r})$  expressed in terms of phonon operators (as above), which in the momentum space takes the form

$$H_{\text{ep}}(\bar{\mathbf{k}}) \approx \sum_{j=1}^4 \sum_{\chi=L,T} \mathcal{N}_{j, \chi}(\hat{\mathbf{p}}) b_j(\bar{\mathbf{k}}) \sqrt{\frac{\hbar \Omega p}{2N_s \Omega_s M c_\chi}} (a_{\mathbf{p}, \chi} + a_{-\mathbf{p}, \chi}^\dagger), \quad (\text{S50})$$



where  $\bar{\mathbf{k}}$  should be understood as the average of initial momentum  $\mathbf{k}$  and final momentum  $\mathbf{k}'$  of the electron,  $\Omega$  and  $\Omega_s$  are the graphene unit cell area and the superlattice unit cell area, and  $\mathcal{N}_{j,\chi}(\hat{\mathbf{p}})$  are dimensionless coefficients depending on the unit vector  $\hat{\mathbf{p}} = \mathbf{p}/p$  (direction of phonon momentum) and are of order 1. These coefficients  $\mathcal{N}_{j,\chi}(\hat{\mathbf{p}})$  can be obtained by substituting the expression of  $\mathbf{u}(\mathbf{r})$  in Eq. (S17) into Eq. (S45), and then rewrite the results into the form of Eq. (S50).

The electron-electron interaction  $V_{\mathbf{k}\mathbf{k}'}(\omega)$  can then be estimated from Eq. (S25). For zero frequency  $\omega = 0$ , the interaction  $V_{\mathbf{k}\mathbf{k}'}$  can be crudely estimated as

$$\frac{V_{\mathbf{k}\mathbf{k}'}}{\Omega_s} = \sum_{\chi=L,T} \frac{2N_s}{-\hbar\omega_{\mathbf{p},\chi}} \frac{\hbar\Omega p}{2N_s\Omega_s M c_\chi} \sum_{j,j'} \mathcal{N}_{j,\chi}(\hat{\mathbf{p}}) \mathcal{N}_{j',\chi}(\hat{\mathbf{p}}) b_j(\bar{\mathbf{k}}) b_{j'}(\bar{\mathbf{k}}) \approx -\frac{\Omega}{M c_T^2 \Omega_s} \left( \sum_{j=1}^4 b_j(\bar{\mathbf{k}})^2 \right), \quad (\text{S51})$$

where for order of magnitude estimation we have simply approximated  $\mathcal{N}_{j',\chi} \sim 1$  and ignored the cross terms between  $b_j(\bar{\mathbf{k}})$  and  $b_{j'}(\bar{\mathbf{k}})$ . We also approximated  $c_L = c_T$ . Given a Fermi energy  $\epsilon_F$ , the BCS coupling strength  $\lambda \approx N_D |V_{\mathbf{k}\mathbf{k}'}|$  can be numerically estimated as

$$\lambda \sim \frac{1}{2\hbar\omega_D} \int_{|\xi_{\mathbf{k}} - \epsilon_F| \leq \hbar\omega_D} \frac{d^2\mathbf{k}}{(2\pi)^2} [-V(\mathbf{k})], \quad (\text{S52})$$

where  $\omega_D$  is the Debye frequency, and we have written  $V_{\mathbf{k}\mathbf{k}'} = V(\bar{\mathbf{k}})$  for short. Note that  $\lambda$  is well-defined in the limit  $\omega_D \rightarrow 0$ .

The main text Fig. 4b shows the BCS coupling strength  $\lambda$  we estimated for  $\theta = 1.05^\circ$ , while the lower panels of Fig. S4d and Fig. S4e here shows  $\lambda$  for  $\theta = 1.20^\circ$  and  $\theta = 1.80^\circ$ , respectively. We find the BCS coupling strength  $\lambda$  in higher bands  $|n| > 4$  can also be generically as large as order 1, although the density of states therein is much lower ( $\sim 0.05 \text{ meV}^{-1} \cdot \Omega_s^{-1}$ ). Here we give a heuristic understanding and estimation. First, for angles  $\theta$  near  $1^\circ$ , numerical calculation shows the second and higher MBZ bands have band widths  $\sim 100$  meV, and this is determined by the two characteristic TBG energy scales which are of the same order: the interlayer hopping  $w = 110$  meV and the graphene kinetic energy  $\hbar v k_\theta \sim 200$  meV. Roughly speaking, the Moiré bands are obtained by folding the original graphene band structure into the MBZ, while the interlayer hopping  $w$  couples different Moiré bands. If we treat  $w$  as a perturbation, the energy of the  $m$ -th band  $E_m(\mathbf{k})$  in the 2nd perturbation theory is roughly

$$E_m(\mathbf{k}) \sim \hbar v k_G^{(m)} + z \frac{w^2}{E_W}, \quad (\text{S53})$$

where  $k_G^{(m)}$  is the original graphene momentum,  $z$  is a numerical factor, and  $E_W \sim 100$  meV is the energy scale of the band width as well as the energy separation between neighbouring bands. Since each Moiré band is coupled to three other bands via  $\mathbf{q}_j$  ( $j = 1, 2, 3$ ), we can estimate the numerical factor  $z$  as  $|z| \sim 3$ .

Under a lattice deformation, the three vectors  $\mathbf{q}_j$  changes to  $\mathbf{q}_j + \delta\mathbf{q}_j$  ( $j = 1, 2, 3$ ), where  $\delta\mathbf{q}_j \sim \gamma k_\theta \partial_a u_b$  ( $a, b = x, y$ ) as given in Eq. (S7). This changes the energy separation between neighbouring bands  $E_W$  by an amount  $\sim \hbar v \delta\mathbf{q}_j$ . Keeping the graphene momentum  $k_G^{(m)}$  invariant in Eq. (S53) (which is the requirement of Eq. (S48)), we have the energy variation of band  $E_m(\mathbf{k})$  to be roughly

$$\delta E_m(\mathbf{k}) \sim z \frac{w^2}{E_W^2} \hbar v \delta\mathbf{q}_j \sim z \frac{w^2}{E_W^2} \hbar v \gamma k_\theta \partial_a u_b. \quad (\text{S54})$$

This gives an estimation of the coefficients in Eq. (S45) as  $b_j(\mathbf{k}) \sim z \frac{w^2}{E_W^2} \hbar v \gamma k_\theta$ . By Eq. (S51), we then find the phonon mediated interaction is of order

$$\frac{V_{\mathbf{k}\mathbf{k}'}}{\Omega_s} \sim -\frac{\Omega z^2 \gamma^2 (\hbar v k_\theta)^2}{M c_T^2 \Omega_s} \frac{w^4}{E_W^4} \sim -\frac{z^2}{M c_T^2} \frac{w^4}{E_W^2}, \quad (\text{S55})$$

where we have used  $\hbar v k_\theta \sim E_W$ , and  $\gamma^2 \approx \Omega_s/\Omega$  (recall that  $\gamma \sim 1/\theta$ , while  $\Omega$  and  $\Omega_s$  are the graphene unit cell area and the Moiré supercell area, respectively). On the other hand, the density of states  $N_D$  also has a generical grow trend as the Fermi energy increases (for electron densities  $|n| > 4$ ), as can be seen in the upper panels of Fig. S4d and S4e. This is because as the energy increases, there are more and more Moiré bands falling into the same energy interval. This can also be understood in the  $w \rightarrow 0$  limit, in which case the TBG becomes two decoupled monolayer graphene, and the density of state  $N_D$  will be that of the graphene Dirac electrons, which grows as the energy increases. From the numerical results of the upper panel of Fig. S4d, we simply estimate  $N_D$  as  $N_D \sim 5 E_W^{-1} \cdot \Omega_s^{-1} \sim 0.05 \text{ meV}^{-1} \cdot \Omega_s^{-1}$

for second or higher bands near the magic angle. Therefore, if we take  $|z| \approx 3$  (for the three nearest neighbour couplings of momenta  $\mathbf{q}_j$ ), we find the BCS coupling strength in higher bands to be around

$$\lambda \sim -N_D V_{\mathbf{k}\mathbf{k}'} \sim \frac{N_D z^2}{M c_T^2} \frac{w^4}{E_W^2} \sim \frac{5z^2}{M c_T^2} \frac{w^4}{E_W^3} \sim 0.5. \quad (\text{S56})$$

The density of states  $N_D$  could be higher near von Hove singularities, where the BCS coupling strength could be higher. This estimation indicates that the electron-phonon coupling is generically enhanced by the Moiré pattern (by the amplification factor  $\gamma$ ), and does not necessarily require the presence of flat bands.

Based on  $\lambda$  estimated in the above, we use the McMillan formula to estimate  $T_c$  in a wide range of  $\theta$  and  $n$ , as is shown in the main text Fig. 4c. Despite the fact that  $T_c$  is estimated in a very rough way, the parameter space where we predict superconductivity may occur at large  $|n|$  is stable.

## S5. AB INITIO CALCULATIONS

To verify the continuum model gives the correct order of magnitude of electron-phonon coupling, we also run *ab initio* calculations of the variation of band structure of TBG under deformations where the lattices of the two layers remain commensurate (which we shall explain below), and compare it with the spectrum calculated from the continuum model. We performed *ab initio* calculations based on the density functional theory with the projector augmented wave (PAW) method [10, 11] as implemented in VASP package [12, 13]. The local density approximation (LDA) was adopted for the exchange-correlation functional [14]. The kinetic energy cutoff of the plane wave basis was set to 300 eV. Only the convergence of band energy at  $\Gamma$  point in the MBZ is used as the criteria for the convergence of self-consistent calculations.

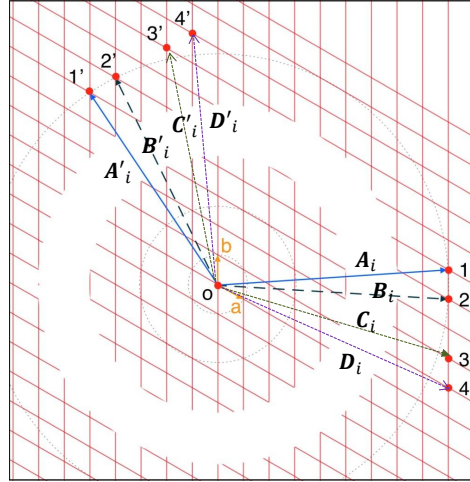


FIG. S5: Illustration of the commensurate Moiré pattern constructed from supercells, and the construction of slightly deformed Moiré pattern. The numbers 1, 2, 3, 4 and 1', 2', 3', 4' label several particular sites (red solid points) useful in defining the commensurate configuration.

Fig. S5 illustrates a graphene lattice, where each site here denotes the center of a hexagon plaquette of graphene. The primitive lattice vectors are denoted as  $\mathbf{a}$  and  $\mathbf{b}$ , and the lattice constant is set to be  $a_0 = 2.456 \text{ \AA}$ . We define the lattice vector  $\mathbf{R}_{m,n} = m\mathbf{a} + n\mathbf{b}$ . Now assume this graphene lattice is the lattice of layer 1 of TBG. By stacking the graphene lattice of layer 2 on top of it (which is not shown in Fig. S5), we arrive at a TBG. To construct a commensurate TBG (undistorted), we first assume the lattices of layer 1 and layer 2 differ by a rotation of angle  $\theta$  about point  $o$  in Fig. S5 (which we define as the origin), namely, the hexagon plaquette centers of layer 1 and layer 2 coincide at point  $o$ . Now consider two lattice vectors  $\mathbf{A}_i = \mathbf{R}_{2i+1,i+1}$  and  $\mathbf{B}_i = \mathbf{R}_{2i+1,i}$  away from the origin  $o$  with  $i \geq 0$  being an integer, which correspond to points 1 and 2 as illustrated in Fig. S5. The two vectors have equal lengths

$$|\mathbf{A}_i| = |\mathbf{B}_i| = L_i^A = \frac{\sqrt{3(2i+1)^2 + 1}}{2} a_0.$$

Similarly, the vectors  $\mathbf{A}'_i = \mathbf{R}_{-i-1,i}$  and  $\mathbf{B}'_i = \mathbf{R}_{-i,i+1}$  from the origin  $o$  correspond to points  $1'$  and  $2'$  in Fig. S5, which also have equal lengths and are related to  $\mathbf{A}_i$  and  $\mathbf{B}_i$  by a  $2\pi/3$  rotation. We then assume the layer 2 lattice is given by rotating the layer 1 lattice so that  $\mathbf{B}_i$  ( $\mathbf{B}'_i$ ) is rotated to  $\mathbf{A}_i$  ( $\mathbf{A}'_i$ ). Such a TBG then forms a commensurate Moiré pattern superlattice which is periodic in real space. As shown in Ref. [15], the twist angle  $\theta = \theta_i$  of such a commensurate configuration satisfies  $\cos \theta_i = \frac{3i^2+3i+0.5}{3i^2+3i+1}$ , or equivalently,

$$\theta_i = 2 \arctan \frac{1}{\sqrt{3}(2i+1)} .$$

The spacial periods, namely, the superlattice vectors, are then given by  $\mathbf{A}_i$  and  $\mathbf{A}'_i$ , and the superlattice unit cell is the parallelogram with edges  $\mathbf{A}_i$  and  $\mathbf{A}'_i$ . We call such a Moiré pattern the  $(11'22')$  configuration, of which the definition involves the four points 1,  $1'$  and 2,  $2'$ .

As an example, we consider the commensurate undistorted TBG configuration  $(11'22')$  with  $i = 10$ , and calculate the band structure with *ab initio*. The distance between the two layers is  $d_0$  ( $\sim 3.35$  Å). The electronic band structure from *ab initio* for  $i = 10$  is shown as red solid lines in both Fig. S6c and S6d (identical between these two figures), where the twist angle  $\theta = \theta_{10} = 3.15^\circ$ , and the high symmetry points in the figure should be understood as those of the MBZ.

To generate an example of deformed Moiré pattern, we consider two different lattice vectors  $\mathbf{C}_i = \mathbf{R}_{2i+1,i-2}$  and  $\mathbf{D}_i = \mathbf{R}_{2i+1,i-3}$  away from the origin  $o$ , which correspond to points 3 and 4 in Fig. S5, respectively (The reason we choose points 3 and 4 will be explained later). We also define another two vectors  $\mathbf{C}'_i = \mathbf{R}_{-i+2,i+3}$  and  $\mathbf{D}'_i = \mathbf{R}_{-i+3,i+4}$ , which are  $\mathbf{C}_i$  and  $\mathbf{D}_i$  rotated by  $2\pi/3$ , and correspond to points  $3'$  and  $4'$ , respectively. We shall still assume the lattice plotted in Fig. S5 is the lattice of layer 1. Next, we assume the lattice of layer 2 is *rotated and deformed* relative to the lattice of layer 1, so that vector  $\mathbf{D}_i$  and  $\mathbf{D}'_i$  of layer 2 coincide with  $\mathbf{C}_i$  and  $\mathbf{C}'_i$  of layer 1, respectively. This is again a commensurate configuration, and the superlattice unit cell are the parallelogram with edges  $\mathbf{C}_i$  and  $\mathbf{C}'_i$  in layer 1. (which coincide with the parallelogram with edges  $\mathbf{D}_i$  and  $\mathbf{D}'_i$  in layer 2, after layer 2 is rotated and deformed). We call this configuration  $(33'44')$ . In particular, the lengths of vectors are

$$|\mathbf{C}_i| = |\mathbf{C}'_i| = L_i^C = \frac{\sqrt{3(2i+1)^2 + 5^2}}{2} a_0 , \quad |\mathbf{D}_i| = |\mathbf{D}'_i| = L_i^D = \frac{\sqrt{3(2i+1)^2 + 7^2}}{2} a_0 ,$$

which are not equal, and the twist angle is now the angle between  $\mathbf{C}_i$  and  $\mathbf{D}_i$ , which is

$$\theta'_i = \arctan \frac{7}{\sqrt{3}(2i+1)} - \arctan \frac{5}{\sqrt{3}(2i+1)} .$$

Therefore, this configuration  $(33'44')$  involves both a relative rotation  $\delta\theta = \theta'_i - \theta_i$  and a relative expansion  $\Theta \approx 2(L_i^D - L_i^C)/L_i^C$  compared to the undistorted configuration  $(11'22')$ . The relative shear  $\Sigma_{ab}$  is, however, zero, since the deformation in this case is isotropic. We can then write down the relative deformation field  $\mathbf{u}$  for  $(33'44')$  (compared to the undistorted TBG configuration) as

$$\begin{pmatrix} \partial_x u_x & \partial_x u_y \\ \partial_y u_x & \partial_y u_y \end{pmatrix} \approx \begin{pmatrix} \frac{L_i^D - L_i^C}{L_i^C} & \theta_i - \theta'_i \\ \theta'_i - \theta_i & \frac{L_i^D - L_i^C}{L_i^C} \end{pmatrix} . \quad (\text{S57})$$

We can also construct a configuration with nonzero relative shear. To do this, we can assume the lattice of layer 2 is *rotated and deformed* relative to the lattice of layer 1 so that vector  $\mathbf{D}_i$  and  $\mathbf{B}'_i$  of layer 2 coincide with  $\mathbf{C}_i$  and  $\mathbf{A}'_i$  of layer 1, respectively, which we shall name as configuration  $(31'42')$ . The superlattice unit cell is then given by the parallelogram with edges  $\mathbf{C}_i$  and  $\mathbf{A}'_i$  in layer 1 (which coincide with the parallelogram with edges  $\mathbf{D}_i$  and  $\mathbf{B}'_i$  in layer 2 after relative deformation). This configuration then has relative rotation, shear and expansion, which can be seen in the following calculation of relative displacement field  $\mathbf{u}$ . The relative displacement field  $\mathbf{u}$  of  $(31'42')$  relative to the undistorted TBG  $(11'22')$  can be solved as follows: Define  $\mathbf{U}(\mathbf{r}) = \mathbf{u}(\mathbf{r}) + \theta_i \mathbf{r} \times \hat{\mathbf{z}}$ , which the relative displacement compared to the untwisted (AA stacking) bilayer graphene (The undistorted TBG  $(11'22')$  has displacement  $\theta_i \mathbf{r} \times \hat{\mathbf{z}}$  relative to the untwisted bilayer graphene). The deformation field  $\mathbf{U}$  then satisfies  $(\mathbf{C}_i \cdot \nabla) \mathbf{U} \approx \mathbf{D}_i - \mathbf{C}_i$ , and  $(\mathbf{A}_i \cdot \nabla) \mathbf{U} \approx \mathbf{B}_i - \mathbf{A}_i$ . We then find the relative displacement field compared to undistorted TBG  $(11'22')$  for large  $i$  is

$$\begin{pmatrix} \partial_x u_x & \partial_x u_y \\ \partial_y u_x & \partial_y u_y \end{pmatrix} \approx \begin{pmatrix} \frac{L_i^D - L_i^C}{L_i^C} & \theta_i - \theta'_i \\ \frac{L_i^D - L_i^C}{\sqrt{3}L_i^C} & \frac{\theta_i - \theta'_i}{\sqrt{3}} \end{pmatrix} . \quad (\text{S58})$$

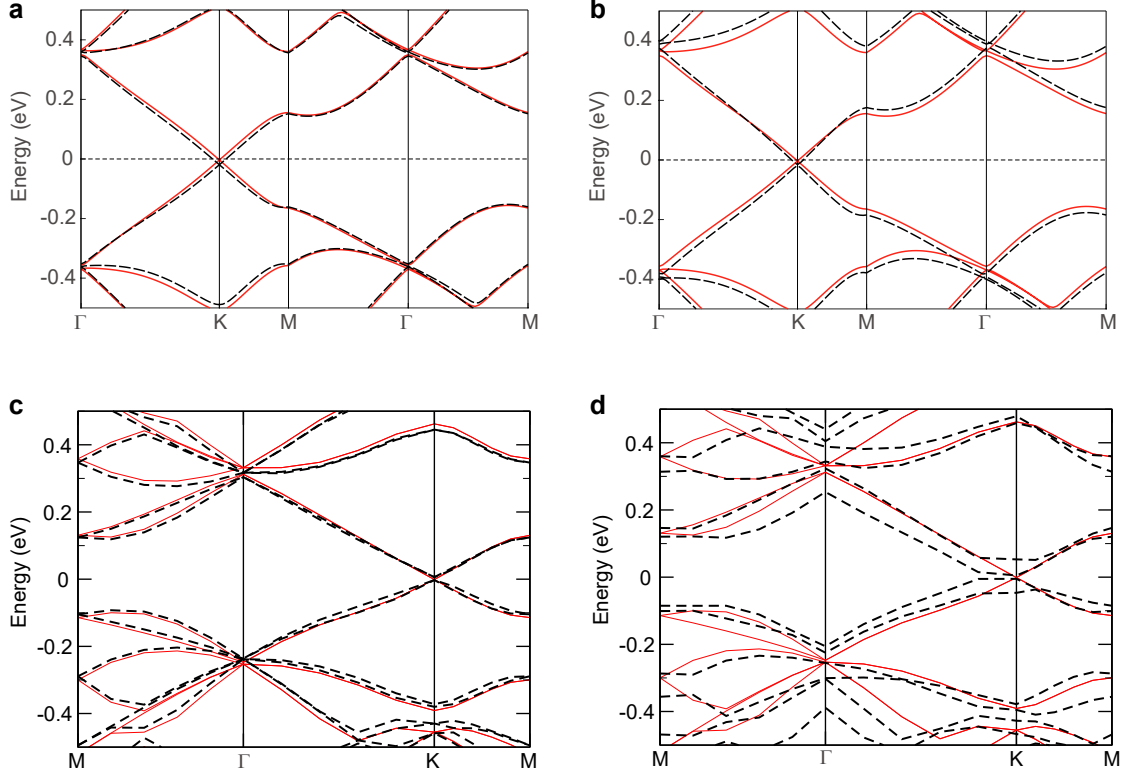


FIG. S6: Variation of band energies under two deformation configurations “3344” and “3142” at  $\theta = 3.15^\circ$  from the continuum model (a and b), and from commensurate TBG *ab initio* calculations (c and d). The energy variations are comparable in order of magnitude in the continuum model and in *ab initio* (slightly larger in *ab initio*).

We now explain why we choose lattice points 3 and 4 (3' and 4') to define the deformed TBG. In our *ab initio* calculations, this is to ensure the  $K$  point of graphene BZ of layers 1 and 2 coincide with the  $K'_M$  and  $K_M$  points of MBZ when folded into the deformed MBZ, respectively. This allows the band structure to be compared with that of the continuum model (in which  $K$  point of two layers always coincide with  $K'_M$  and  $K_M$ ). To see this, consider a configuration that has superlattice vectors (which spans the superlattice unit cell parallelogram)  $\mathbf{R}_1 = m_1\mathbf{a} + n_1\mathbf{b}$  and  $\mathbf{R}_2 = m_2\mathbf{a} + n_2\mathbf{b}$ , where  $\mathbf{a}$  and  $\mathbf{b}$  are the lattice vectors of the lattice of layer 1. The reciprocal vectors of the superlattice  $\mathbf{g}_i$  ( $i = 1, 2$ ) satisfies  $\mathbf{g}_i \cdot \mathbf{R}_j = 2\pi\delta_{ij}$ . On the other hand, the momentum of the  $K$  point of layer 1 is  $\mathbf{K}_D$  as we defined in the first section, which satisfies  $\mathbf{K}_D \cdot \mathbf{a} = \mathbf{K}_D \cdot \mathbf{b} = 2\pi/3$ . Therefore, we find

$$\mathbf{K}_D \cdot \mathbf{R}_1 = 2\pi(m_1 + n_1)/3, \quad \mathbf{K}_D \cdot \mathbf{R}_2 = 2\pi(m_2 + n_2)/3.$$

This shows  $\mathbf{K}_D = [(m_1 + n_1)\mathbf{g}_1 + (m_2 + n_2)\mathbf{g}_2]/3$ . Note that  $K'_M$  point of the MBZ is located at momentum  $K'_M = -(\mathbf{g}_1 + \mathbf{g}_2)/3$ . Therefore, in order for  $K_D$  to coincide with  $K'_M$  point, one has to have  $m_1 + n_1 \equiv m_2 + n_2 \equiv -1 \pmod{3}$  (so that  $K_D$  and  $K'_M$  differ by integer multiples of superlattice reciprocal vectors). One can easily see this is satisfied for configuration (11'22') where  $\mathbf{R}_1 = \mathbf{A}_i$  and  $\mathbf{R}_2 = \mathbf{A}'_i$ . In order to find another configuration satisfying this condition, one has to change  $m_1 + n_1$  and  $m_2 + n_2$  by multiples of 3. The configurations (33'44') and (31'42') are two such configurations which have small deformations.

We then use *ab initio* to calculate the band structure for configurations (33'44') and (31'42') with  $i = 10$  (see the black dashed lines in Fig. S6c and S6d), respectively, and compare with that of the undistorted configuration (11'22') (the red solid lines in Fig. S4c and S4d). Accordingly, we calculate the deformation of TBG band structure from the continuum model (at graphene valley  $K$ ) for configurations (33'44') and (31'42') (using deformations in Eqs. (S57) and (S58)) at  $\theta = \theta_{10} = 3.15^\circ$ , respectively, and the results are shown in Fig. S6a and S6b, where the red solid lines are the original band structure, and the black dashed lines are the deformed band structure. In the figure, the high symmetry points should be understood as those of the MBZ.

First, we note the band structures before deformations match well between the continuum model and the *ab initio*. The continuum model bands appear to be less in number than that of *ab initio*, which is because we have only plotted the bands of continuum model at valley  $K$ . Besides, for both configurations, one can see that the deformed band energies in the continuum model also have the same order of magnitude as that in *ab initio*. Therefore, our estimation of electron-phonon coupling strength from the continuum model in the previous sections has the correct order of magnitude.

- 
- [1] Rafi Bistritzer and Allan H. MacDonald, “Moiré bands in twisted double-layer graphene,” *Proceedings of the National Academy of Sciences* **108**, 12233–12237 (2011), <http://www.pnas.org/content/108/30/12233.full.pdf>.
  - [2] Alexandr I. Cocemasov, Denis L. Nika, and Alexander A. Balandin, “Phonons in twisted bilayer graphene,” *Phys. Rev. B* **88**, 035428 (2013).
  - [3] L. D. Landau, L. P. Pitaevskii, A. M. Kosevich, and E. M. Lifshitz, “Theory of elasticity,” (2012).
  - [4] S. Das Sarma and E. H. Hwang, “Collective modes of the massless dirac plasma,” *Phys. Rev. Lett.* **102**, 206412 (2009).
  - [5] W. L. McMillan, “Transition temperature of strong-coupled superconductors,” *Phys. Rev.* **167**, 331–344 (1968).
  - [6] P. B. Allen and R. C. Dynes, “Transition temperature of strong-coupled superconductors reanalyzed,” *Phys. Rev. B* **12**, 905–922 (1975).
  - [7] P.W. Anderson, “Theory of dirty superconductors,” *Journal of Physics and Chemistry of Solids* **11**, 26 – 30 (1959).
  - [8] Z. Song, Z. Wang, W. Shi, G. Li, C. Fang, and B. A. Bernevig, “All “Magic Angles” Are “Stable” Topological,” *ArXiv e-prints* (2018), arXiv:1807.10676 [cond-mat.mes-hall].
  - [9] Dillon Wong, Yang Wang, Jeil Jung, Sergio Pezzini, Ashley M. DaSilva, Hsin-Zon Tsai, Han Sae Jung, Ramin Khajeh, Youngkyou Kim, Juwon Lee, Salman Kahn, Sajjad Tollabimazraehno, Haider Rasool, Kenji Watanabe, Takashi Taniguchi, Alex Zettl, Shaffique Adam, Allan H. MacDonald, and Michael F. Crommie, “Local spectroscopy of moiré-induced electronic structure in gate-tunable twisted bilayer graphene,” *Phys. Rev. B* **92**, 155409 (2015).
  - [10] P. E. Blöchl, “Projector augmented-wave method,” *Phys. Rev. B* **50**, 17953–17979 (1994).
  - [11] G. Kresse and D. Joubert, “From ultrasoft pseudopotentials to the projector augmented-wave method,” *Phys. Rev. B* **59**, 1758–1775 (1999).
  - [12] G. Kresse and J. Hafner, “Ab initio molecular dynamics for liquid metals,” *Phys. Rev. B* **47**, 558–561 (1993).
  - [13] G. Kresse and J. Furthmüller, “Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set,” *Computational Materials Science* **6**, 15 – 50 (1996).
  - [14] P. Hohenberg and W. Kohn, “Inhomogeneous electron gas,” *Phys. Rev.* **136**, B864–B871 (1964).
  - [15] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, “Graphene bilayer with a twist: Electronic structure,” *Phys. Rev. Lett.* **99**, 256802 (2007).