Continuum effective Hamiltonian for graphene bilayers for an arbitrary smooth lattice deformation from microscopic theories

Oskar Vafek[1,2,*] and Jian Kang [ORCID][3,†]

[1]*National High Magnetic Field Laboratory, Tallahassee, Florida, 32310, USA*
[2]*Department of Physics, Florida State University, Tallahassee, Florida 32306, USA*
[3]*School of Physical Science and Technology, ShanghaiTech University, Shanghai 200031, China*

We provide a systematic real-space derivation of the continuum Hamiltonian for a graphene bilayer starting from a microscopic lattice theory, allowing for an arbitrary inhomogeneous smooth lattice deformation, including a twist. Two different microscopic models are analyzed: First, a Slater-Koster like model and, second, an *ab initio* derived model. We envision that our effective Hamiltonian can be used in conjunction with an experimentally determined atomic lattice deformation in twisted bilayer graphene in a specific device to predict and compare the electronic spectra with scanning tunneling spectroscopy measurements. As a byproduct, our approach provides electron-phonon couplings in the continuum Hamiltonian from microscopic models for any bilayer stacking. In the companion paper [J. Kang and O. Vafek, Phys. Rev. B **107**, 075408 (2023)], we analyze in detail the continuum models for relaxed atomic configurations of magic angle twisted bilayer graphene.

I. INTRODUCTION

Observation of the correlated electron phenomena [1], including superconductivity [2], in the vicinity of the first magic angle in twisted bilayer graphene [3–6] led to a large number of experimental [7–31] and theoretical studies [32–52] of this remarkable physical system. Although the main experimental findings [1,2,7–9] were reproduced by a number of experimental groups, there is a nagging lack of reproducibility in the finer details of the physical characteristics of devices, even when manufactured within the same laboratory and even within the same device. This is likely due to spatial inhomogeneity in the twist angle [16–18,20] and unintentional strain [11] produced during the device fabrication or, more generally, due to lattice deformations which vary over distances long compared to the microscopic spacing between neighboring carbon atoms.

It is thus being recognized that the twist angle is not the only parameter controlling the physics of a specific device [29]. This fact motivates the development of a theory whose input would be more than just the twist angle θ, Fermi velocity v_F, and the two interlayer tunneling constants through the AA (w_0) and AB (w_1) regions, as is the case for (the slight generalization of) the original Bistritzer-MacDonald (BM) model, but instead the input would be a smooth and possibly inhomogeneous configuration of the atomic displacement field. This configuration could in, principle, be extracted from topography measured using a scanning tunneling microscope [10,11,13,15,53] or from Bragg interferometry [54].

The goal of this paper is to provide a systematic derivation of such a continuum Hamiltonian for an arbitrary smooth

atomic displacement $\boldsymbol{u}_j(\boldsymbol{r})$ starting from a microscopic *ab initio* calibrated tight-binding model on the carbon lattice. Expanding in gradients of $\boldsymbol{u}_j(\boldsymbol{r})$ and of a slowly varying envelope of the graphene's $\mathbf{K}$ and $\mathbf{K}'$ Bloch functions, one can achieve any desired accuracy when comparing with the microscopic model, as we demonstrate in the companion paper for the relaxed atomic configurations of the magic angle twisted bilayer graphene [55]. Here, we provide the general formulas for two different microscopic models. For the first, we consider a microscopic hopping function which depends only on the separation between two carbon atoms, as is the case in the Slater-Koster (SK)-type models [4,33,56–59]. For the second, we allow for dependence of the interlayer tunneling terms on the relative orientation of the interatomic separation vector and the nearest-neighbor bonds, as is the case in the microscopic model derived from density functional theory (DFT) determined Wannier states of the monolayer (and untwisted bilayer) graphene's conduction and valance bands in Ref. [60], as well as for the configuration dependence of the on-site term.

The method which we develop here is inspired by the approach advanced by Balents [37], but strives to go beyond it in several ways.

First, the continuum Hamiltonian is derived entirely from the microscopic tight-binding models. As a consequence, all the parameters in the continuum Hamiltonian can be expressed by suitable moments of the hopping functions and the lattice distortions, yielding realistic values of the electron-phonon couplings as a byproduct. This allows for a direct comparison of the theory with the experiments if the deformed positions of atoms are measured by local probes.

Second, when applied to the twisted bilayer graphene, the continuum model derived here goes beyond the BM model [6,37] by systematically including higher order gradient terms i.e., gradients of both the slowly varying envelope of the

*vafek@magnet.fsu.edu
†kangjian@shanghaitech.edu.cn

fermion fields and of the atomic displacement fields. Because they are derived directly in real space via a gradient expansion, each term in our continuum theory is local [61] (although, of course, not necessarily just contact). In contrast, the continuum models in Refs. [6,62] are obtained in the momentum space which makes any treatment of spatial inhomogeneity inconvenient, and when translated to real space, the existing models include only (some of the) first-order gradient terms. The motivation for including higher order gradient terms is directly related to the physics of the magic angle. For a twist angle θ near the magic value $1.1°$, the estimate of the energy scale of the leading order terms constituting the BM model [6,37] can be obtained by multiplying the Fermi velocity v_F and the typical momentum deviation from the Dirac point, $\hbar v_F |\mathbf{K}|\theta \sim 200$ meV for the intralayer term, and ~ 100 meV for the contact interlayer term. The second-order intralayer derivative terms and the first-order derivative in fermion and atomic displacement fields interlayer terms are smaller by the factor of $\sim |\mathbf{K}|a\theta = 4\pi\theta/3 \sim 0.08$, seemingly justifying their omission (for definitions of various parameters mentioned, see the next section). As is well-known, however, at the magic angle the noninteracting bandwidth is anomalously smaller than the scale of the leading order terms by at least an order of magnitude, making the higher order terms comparable to the noninteracting narrow bandwidth [63]. Moreover, even if smaller, they can be of similar order to the scale of Coulomb interaction, and they break particle-hole symmetry [39–41,64], thus lifting degeneracy of the ground state manifold in strong coupling [43,47,51,52,65]. Therefore, it is desirable to study their effects systematically as we do here and the companion paper [55].

Finally, it was recognized [32] that atomic relaxation of twisted bilayer graphene near the first magic angle leads to an increase in the size of the AB-stacked regions of the moiré pattern at the expense of the AA-stacked regions, as compared to the structure resulting from a simple rigid twist. With few exceptions [62], such relaxation has been modeled as a simple change of AA and AB tunneling parameters w_0 and w_1, respectively. However, because the difference between w_0 and w_1 arises from lattice distortions, such relaxation must include pseudomagnetic vector potential terms——given by combinations of *first-order* spatial derivatives of the atomic displacement fields—in the intralayer Hamiltonian. Within the gradient expansion, such terms appear at the same order as the intralayer first-order gradient of the slow Fermi fields, i.e., same order as the massless Dirac terms. Indeed, we find that such terms are comparable to the interlayer tunneling terms $w_{0,1}$ included in the BM model, and therefore there is no *a priori* justification for neglecting them.

II. MICROSCOPIC DERIVATION OF THE CONTINUUM LOW-ENERGY MODEL FOR AN ARBITRARY SMOOTH LATTICE DEFORMATION

To derive the effective continuum Hamiltonian from the microscopic tight-binding model, we start by noting that, generally, the distorted position $\mathbf{X}_{j,S}$ of a carbon atom in the layer j and sublattice S can be expressed as

$$\mathbf{X}_{j,S} = \mathbf{r}_S + \mathbf{u}_{j,S}(\mathbf{r}_S) \equiv \mathbf{X}_{j,S}(\mathbf{r}_S), \qquad (1)$$

$$\mathbf{u}_{j,S}(\mathbf{r}_S) = \mathbf{u}_{j,S}^{\parallel}(\mathbf{r}_S) + \mathbf{u}_{j,S}^{\perp}(\mathbf{r}_S), \qquad (2)$$

where $\mathbf{r}_S = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + \boldsymbol{\tau}_S$ is the reference undistorted position of the carbon atom within a honeycomb lattice (see Fig. 1), with $n_{1,2}$ being integers. The basis vectors of the undistorted lattice are $\boldsymbol{\tau}_A = 0$ for sublattice A and $\boldsymbol{\tau}_B = (\mathbf{a}_1 + \mathbf{a}_2)/3$ for sublattice B, where $\mathbf{a}_1 = a(1, 0)$ and $\mathbf{a}_2 = a(\frac{1}{2}, \frac{\sqrt{3}}{2})$ are the two primitive lattice vectors and $a = 0.246$ nm being the lattice constant. The displacement $\mathbf{u}_{j,S}(\mathbf{r}_S)$ describes the deviation from the undistorted position of the carbon atoms. It is general enough to account for twist, in-plane relaxation, out-of-plane corrugation, and strain, as well as any possible difference between atomic displacements of the two sublattices. The vector $\mathbf{u}_{j,S}(\mathbf{r}_S)$ in Eq. (2) is decomposed into an in-plane component $\mathbf{u}_{j,S}^{\parallel}(\mathbf{r}_S)$ and an out-of-plane component $\mathbf{u}_{j,S}^{\perp}(\mathbf{r}_S)$. Its explicit dependence on the undistorted lattice point $\mathbf{r}_S$ is referred to as the Lagrangian coordinates [37,66].

Although we start with the Lagrangian formulation, we will reach a point in our derivation where we switch to the more convenient Eulerian coordinates [37,66], where the displacements are expressed in terms of the actual in-plane position of the atoms $\mathbf{X}_{j,S}^{\parallel}$ as opposed to the undistorted positions $\mathbf{r}_S$. Because each monolayer graphene sheet is assumed not to fold, there is a one-to-one mapping between $\mathbf{r}_S$ and $\mathbf{X}_{j,S}^{\parallel}$. If it folded, there would be overhangs for a sheet, and two different positions $\mathbf{r}_S$ would map onto the same $\mathbf{X}_{j,S}^{\parallel}$. Without overhangs, we can therefore adopt the Monge gauge [66] and use the Eulerian coordinates and write

$$\mathbf{X}_{j,S} = \mathbf{r}_S + \mathbf{U}_{j,S}^{\parallel}(\mathbf{X}_{j,S}^{\parallel}) + \mathbf{U}_{j,S}^{\perp}(\mathbf{X}_{j,S}^{\parallel}). \qquad (3)$$

The displacement functions $\mathbf{U}_{j,S}^{\parallel,\perp}$ now depend on the actual in-plane location of the distorted atoms which can be determined by solving Eq. (1) for $\mathbf{r}_S$ in terms of $\mathbf{X}_{j,S}^{\parallel}$ and then expressing the displacement fields in terms of $\mathbf{X}_{j,S}^{\parallel}$.

To illustrate the main idea, in this section we allow the hopping amplitude t to depend only on the separation of the two carbon atoms $\mathbf{X}_{j,S} - \mathbf{X}_{j',S'}'$. In general, t depends also on the orientation [60] of this vector relative to the nearest-neighbor sites of the atom at $\mathbf{X}_{j,S}$ and at $\mathbf{X}_{j',S'}'$ (Fig. 1). Moreover, the general on-site term acquires configuration dependence. We treat this more intricate case in Sec. II A. Thus, we start with a microscopic tight-binding model

$$H_{\text{tb}}^{\text{SK}} = \sum_{SS'} \sum_{jj'} \sum_{\mathbf{r}_S, \mathbf{r}_{S'}'} t(\mathbf{X}_{j,S} - \mathbf{X}_{j',S'}') c_{j,S,\mathbf{r}_S}^{\dagger} c_{j',S',\mathbf{r}_{S'}'}, \qquad (4)$$

where the fermion creation and annihilation operators satisfy the anticommutation relation $\{c_{j,S,\mathbf{r}_S}^{\dagger}, c_{j',S',\mathbf{r}_{S'}'}\} = \delta_{jj'}\delta_{SS'}\delta_{\mathbf{r}_S \mathbf{r}_{S'}'}$. Because $H_{\text{tb}}^{\text{SK}}$ is Hermitian,

$$t(\mathbf{X}) = t^*(-\mathbf{X}), \qquad (5)$$

and because (spinless) time-reversal symmetry is preserved:

$$t(\mathbf{X}) = t^*(\mathbf{X}). \qquad (6)$$

One example of a model with t depending only on the separation of the two carbon atoms is the often used SK-type model
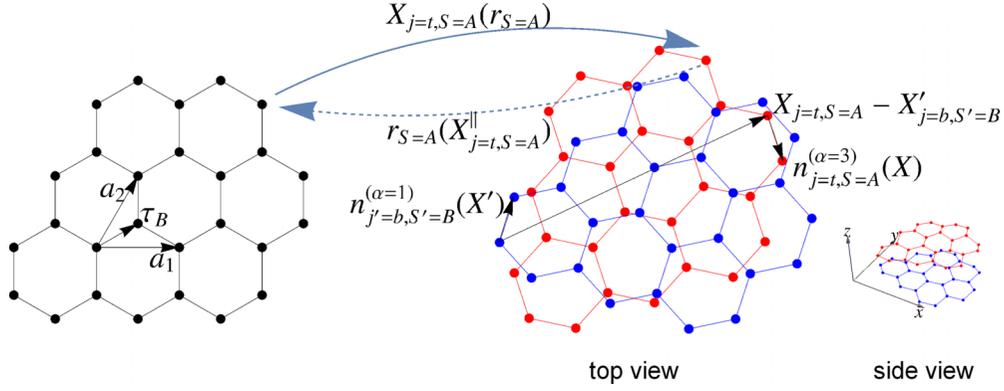
FIG. 1. Schematic illustration of the one-to-one mapping between (left) the undistorted atomic position r_S and (right, top view) the distorted atomic position $X_{j,S}$, where red is for the top layer and blue for the bottom layer. The separation between two carbon atoms $X_{j,S} - X'_{j',S'}$ and the corresponding nearest-neighbor vectors $n_{j,S}$ and $n_{j',S'}$ are labeled by black arrows. The distortion described by $X_{j,S}$ also includes the lattice corrugation, as shown in the side view.

[4,33,56–59] for the carbon p_z orbitals,

$$t(\boldsymbol{d}) = V_{pp\pi}^0 e^{-\frac{|\boldsymbol{d}|-a_0}{\Delta}}\left[1 - \left(\frac{\boldsymbol{d}\cdot\hat{z}}{|\boldsymbol{d}|}\right)^2\right] + V_{pp\sigma}^0 e^{-\frac{|\boldsymbol{d}|-d_0}{\Delta}}\left(\frac{\boldsymbol{d}\cdot\hat{z}}{|\boldsymbol{d}|}\right)^2,$$

(7)

where for concreteness $V_{pp\pi}^0 = -2.7$ eV, $V_{pp\sigma}^0 = 0.48$ eV, $a_0 = |\boldsymbol{\tau}_B| = 0.142$ nm is the distance between the two nearest-neighbor carbon atoms on the same layer; $d_0 = 0.335$ nm is the interlayer distance and the decay length for the hopping is $\Delta = 0.319 a_0$.

In this paper, we will not need to use the specific form of t in Eq. (7). As explained later, we only rely on its fast decay. To obtain the continuum effective Hamiltonian from the microscopic tight-binding model Eq. (4), we next write

$$H_{\text{tb}}^{\text{SK}} = \sum_{SS'}\sum_{jj'}\sum_{\boldsymbol{r}_S,\boldsymbol{r}'_{S'}}\int d^2\boldsymbol{r}\,d^2\boldsymbol{r}'\,\delta(\boldsymbol{r}-\boldsymbol{r}_S)\delta(\boldsymbol{r}'-\boldsymbol{r}'_{S'})$$
$$\times t(\boldsymbol{r}+\boldsymbol{u}_{j,S}(\boldsymbol{r})-\boldsymbol{r}'-\boldsymbol{u}_{j',S'}(\boldsymbol{r}'))c_{j,S,\boldsymbol{r}}^\dagger c_{j',S',\boldsymbol{r}'},$$

(8)

interchange the order of summation and integration, and apply the Dirac comb formula

$$\sum_{\boldsymbol{r}_S}\delta(\boldsymbol{r}-\boldsymbol{r}_S) = \frac{1}{A_{\text{mlg}}}\sum_{\boldsymbol{G}}e^{i\boldsymbol{G}\cdot(\boldsymbol{r}-\boldsymbol{\tau}_S)}.$$

Here $\boldsymbol{G} = 2\pi(m_1\boldsymbol{a}_2 - m_2\boldsymbol{a}_1)\times\hat{z}/A_{\text{mlg}}$ is the reciprocal lattice vector of the undistorted monolayer graphene, $m_{1,2}$ are integers, and $A_{\text{mlg}} = |\boldsymbol{a}_1\times\boldsymbol{a}_2| = \frac{\sqrt{3}}{2}a^2$ is the area of the undistorted monolayer graphene unit cell. Since the physically important states come from the vicinity of the Dirac points, we can decompose the fermion fields into two slowly spatially varying fields ψ and ϕ multiplied by the fast spatially varying functions from the valley $\boldsymbol{K} = 4\pi\boldsymbol{a}_1/(3a^2)$ and $\boldsymbol{K}' = -\boldsymbol{K}$ as

$$A_{\text{mlg}}^{-1/2}c_{j,S,\boldsymbol{r}} \simeq e^{i\boldsymbol{K}\cdot\boldsymbol{r}}\psi_{j,S}(\boldsymbol{r}) + e^{-i\boldsymbol{K}\cdot\boldsymbol{r}}\phi_{j,S}(\boldsymbol{r}).$$

(9)

The factor of $A_{\text{mlg}}^{-1/2}$ is included to satisfy the anticommutation relation

$$\{\psi_{j,S}(\boldsymbol{r}), \psi_{j',S'}^\dagger(\boldsymbol{r}')\} = \{\phi_{j,S}(\boldsymbol{r}), \phi_{j',S'}^\dagger(\boldsymbol{r}')\}$$
$$= \delta_{jj'}\delta_{SS'}\delta(\boldsymbol{r}-\boldsymbol{r}').$$

(10)

The effective Hamiltonian at the valley $\boldsymbol{K}$ can now be written as

$$H_{\text{SK,eff}}^{\boldsymbol{K}} = \frac{1}{A_{\text{mlg}}}\sum_{jj'}\sum_{SS'}\sum_{\boldsymbol{G},\boldsymbol{G}'}\int d^2\boldsymbol{r}\,d^2\boldsymbol{r}'\,e^{i\boldsymbol{G}\cdot(\boldsymbol{r}-\boldsymbol{\tau}_S)}$$
$$\times e^{-i\boldsymbol{G}'\cdot(\boldsymbol{r}'-\boldsymbol{\tau}_{S'})}t(\boldsymbol{r}+\boldsymbol{u}_{j,S}(\boldsymbol{r})-\boldsymbol{r}'-\boldsymbol{u}_{j',S'}(\boldsymbol{r}'))$$
$$\times e^{-i\boldsymbol{K}\cdot(\boldsymbol{r}-\boldsymbol{r}')}\psi_{j,S}^\dagger(\boldsymbol{r})\psi_{j',S'}(\boldsymbol{r}').$$

(11)

The effective Hamiltonian at the valley $\boldsymbol{K}'$ is related to $H_{\text{SK,eff}}^{\boldsymbol{K}}$ by spinless time-reversal symmetry.

The hopping amplitude t is a short-ranged function that decays exponentially fast as its argument increases beyond a few carbon lattice spacings, while ψ varies slowly over such length scales. To take advantage of this fact, it is convenient to switch to Eulerian coordinates [37]. By doing so, the locality of the effective theory will become manifest [37]. Thus, we now perform a coordinate transformation from $\boldsymbol{r}$ to $\boldsymbol{X}^\parallel$, where for each j and S we let the integration variable be $\boldsymbol{X}^\parallel = \boldsymbol{r} + \boldsymbol{u}_{j,S}^\parallel(\boldsymbol{r})$, and similarly for the primed variables. We also introduce new $\boldsymbol{X}^\parallel$-dependent fermion fields as

$$\Psi_{j,S}(\boldsymbol{X}^\parallel) = \left|J\left(\frac{\partial\boldsymbol{r}}{\partial\boldsymbol{X}^\parallel}\right)\right|^{\frac{1}{2}}\psi_{j,S}(\boldsymbol{r})$$
$$= \left|J\left(\frac{\partial(\boldsymbol{X}^\parallel - \boldsymbol{U}_{j,S}^\parallel(\boldsymbol{X}^\parallel))}{\partial\boldsymbol{X}^\parallel}\right)\right|^{\frac{1}{2}}$$
$$\times \psi_{j,S}(\boldsymbol{X}^\parallel - \boldsymbol{U}_{j,S}^\parallel(\boldsymbol{X}^\parallel)),$$

(12)

where J is the Jacobi determinant [67]. As emphasized in Ref. [37], $\boldsymbol{U}$ is not small for twisted structures and the formalism developed here does not make this assumption. However, we do assume that $\boldsymbol{U}$ is smooth, i.e., its gradients are small. Therefore, the Ψ fields are also slow. By the property of the Dirac delta functions under the change of variables (and because the transformation between $\boldsymbol{X}^\parallel$ and $\boldsymbol{r}$ is one-to-one), these fields satisfy the canonical fermion commutation relations:

$$\{\Psi_{j,S}(\boldsymbol{X}^\parallel), \Psi_{j',S'}^\dagger(\boldsymbol{X}'^\parallel)\} = \delta_{jj'}\delta_{SS'}\delta(\boldsymbol{X}^\parallel - \boldsymbol{X}'^\parallel).$$

(13)

For notational simplicity, we also introduce the symbol $\mathcal{J}_{j,S}(X^\parallel) \equiv |J(\frac{\partial r}{\partial X^\parallel})|^{\frac{1}{2}}$. The effective continuum Hamiltonian now becomes

$$
H^{\mathbf{K}}_{\text{SK,eff}} = \frac{1}{A_{\text{mlg}}} \sum_{jj'} \sum_{SS'} \sum_{\mathbf{G},\mathbf{G}'} e^{-i(\mathbf{G}\cdot\boldsymbol{\tau}_s - \mathbf{G}'\cdot\boldsymbol{\tau}_{s'})} \int d^2X^\parallel \, d^2X'^\parallel \, \mathcal{J}_{j,S}(X^\parallel)\mathcal{J}_{j',S'}(X'^\parallel) e^{-i(\mathbf{G}-\mathbf{K})\cdot U^\parallel_{j,S}(X^\parallel)} e^{i(\mathbf{G}'-\mathbf{K})\cdot U^\parallel_{j',S'}(X'^\parallel)}
$$

$$
\times \, t(X^\parallel + U^\perp_{j,S}(X^\parallel) - X'^\parallel - U^\perp_{j',S'}(X'^\parallel)) e^{i(\mathbf{G}\cdot X^\parallel - \mathbf{G}'\cdot X'^\parallel)} e^{-i\mathbf{K}\cdot(X^\parallel - X'^\parallel)} \Psi^\dagger_{j,S}(X^\parallel)\Psi_{j',S'}(X'^\parallel). \tag{14}
$$

To exploit the short-range nature of t, we switch to the center-of-mass $\mathbf{x} = \frac{1}{2}(X^\parallel + X'^\parallel)$ and relative coordinates $\mathbf{y} = X^\parallel - X'^\parallel$. Thus, $\int d^2X^\parallel \, d^2X'^\parallel \ldots = \int d^2\mathbf{x} \, d^2\mathbf{y} \ldots$, and $e^{i(\mathbf{G}\cdot X^\parallel - \mathbf{G}'\cdot X'^\parallel)} = e^{i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{x}} e^{i\frac{1}{2}(\mathbf{G}+\mathbf{G}')\cdot\mathbf{y}}$. The integral over $\mathbf{x}$ contains the phase factor $e^{i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{x}}$ that oscillates strongly over the scale of the monolayer graphene lattice constant a when $\mathbf{G} \neq \mathbf{G}'$, whereas all other factors are smooth functions of $\mathbf{x}$. As a consequence, the integral over $\mathbf{x}$ is negligible as long as $\mathbf{G} \neq \mathbf{G}'$; this collapses the double sum over $\mathbf{G}, \mathbf{G}'$ to a single sum. Moreover, the remaining fields, whether Ψ or U, are smooth and can now be expanded in powers of gradients, e.g., $\Psi(\mathbf{x} - \frac{1}{2}\mathbf{y}) \simeq \Psi(\mathbf{x}) - \frac{1}{2}\mathbf{y}\cdot\nabla_{\mathbf{x}}\Psi(\mathbf{x}) + \frac{1}{8}(\mathbf{y}\cdot\nabla_{\mathbf{x}})^2\Psi(\mathbf{x})\ldots$, because powers of $\mathbf{y}$ are compensated by the exponential decay of t at large $\mathbf{y}$, effectively confining $\mathbf{y}$ to small values. Changing $\mathbf{G}$ to $-\mathbf{G}$ and to the first order in gradients, we obtain the main result of this section:

$$
H^{\mathbf{K}}_{\text{SK,eff}} \simeq \frac{1}{A_{\text{mlg}}} \sum_{S,S'} \sum_{jj'} \sum_{\mathbf{G}} e^{i\mathbf{G}\cdot(\boldsymbol{\tau}_S - \boldsymbol{\tau}_{S'})} \int d^2\mathbf{x}\, \mathcal{J}_{j,S}(\mathbf{x})\mathcal{J}_{j',S'}(\mathbf{x}) e^{i(\mathbf{G}+\mathbf{K})\cdot(U^\parallel_{j,S}(\mathbf{x}) - U^\parallel_{j',S'}(\mathbf{x}))}
$$

$$
\times \int d^2\mathbf{y}\, e^{-i(\mathbf{G}+\mathbf{K})\cdot\mathbf{y}} e^{i\frac{\mathbf{y}}{2}\cdot\nabla_{\mathbf{x}}(U^\parallel_{j,S}(\mathbf{x})+U^\parallel_{j',S'}(\mathbf{x}))\cdot(\mathbf{G}+\mathbf{K})} t[\mathbf{y} + U^\perp_{j,S}(\mathbf{x}) - U^\perp_{j',S'}(\mathbf{x})w]
$$

$$
\times \left[\Psi^\dagger_{j,S}(\mathbf{x})\Psi_{j',S'}(\mathbf{x}) + \frac{\mathbf{y}}{2}\cdot((\nabla_{\mathbf{x}}\Psi^\dagger_{j,S}(\mathbf{x}))\Psi_{j',S'}(\mathbf{x}) - \Psi^\dagger_{j,S}(\mathbf{x})\nabla_{\mathbf{x}}\Psi_{j',S'}(\mathbf{x})) \right]. \tag{15}
$$

Extension to higher order gradients is straightforward (see Appendix A). We analyze the accuracy of this formula for the SK-like models of twisted bilayer graphene [4,59], including in-plane lattice relaxation, by comparing the low-energy continuum and tight-binding spectra in the companion paper [55] by also including a second-order gradient term in the intralayer part of the effective Hamiltonian. Analogous formula is derived in the next two sections for the model in Ref. [60] that includes dependence of t on the relative orientation of the intralayer nearest-neighbor sites and the vector connecting two interlayer sites $X_{j,S} - X'_{j',S'}$.

The intervalley scattering terms are negligibly small. This can be seen by a direct substitution of Eq. (9), following the analysis above, and noticing that the vector $2\mathbf{K}$ is *not* a reciprocal lattice vector $\mathbf{G}$, while $3\mathbf{K}$ is. For the intervalley scattering, we therefore need to compensate for the missing $\mathbf{K}$ using terms of order $\sim\mathbf{G}\cdot\partial_\mu U$. Because for rigid twist angle θ, $\partial_\mu U \sim \theta$ and because the relaxed atomic configuration is smooth on the moiré scale L_m, more generally $\partial_\mu U \sim a/L_m \ll 1$. This forces us to either go to very high $\mathbf{G} \sim \mathbf{K}L_m/a$ for which the Fourier transform of t is exponentially small or, for smaller $\mathbf{G}$, to extract the Fourier component of terms of the form $e^{i(\mathbf{G}+\mathbf{K})\cdot U(\mathbf{x})}$ at $\mathbf{K}$. Upon Fourier expanding $U(\mathbf{x})$, the function $e^{i(\mathbf{G}+\mathbf{K})\cdot U(\mathbf{x})}$ can be thought of as a product of generating functions for the Bessel functions. While nonzero, the Fourier component of $e^{i(\mathbf{G}+\mathbf{K})\cdot U(\mathbf{x})}$ at $\mathbf{K}$ corresponds to Bessel functions at high indices with arguments set by the $\partial_\mu U(\mathbf{x})$, which are exponentially small. Inspecting the tight-binding spectra analyzed in the companion paper [55], which contain the intervalley scattering terms, and comparing them with the spectra obtained from the continuum models which neglect them, indeed justifies neglecting the intervalley scattering terms over the experimentally relevant energy scale.

Different from the earlier works [62], our derivation does not distinguish the intralayer and interlayer Hamiltonians of

the continuum theory. Consequently, the interlayer tunneling obtained from Eq. (15) depends not only on the asymmetric lattice displacement $U^\parallel_{t,S} - U^\parallel_{b,S'}$ but also on the gradient of the symmetric part $\nabla(U^\parallel_{t,S} + U^\parallel_{b,S'})$.

A. Bond orientation dependent hopping

In the previous section, we derived the effective continuum Hamiltonian when the hopping depends only on the separation of two carbon atoms $X_{j,S} - X'_{j',S'}$, as is the case for SK-type models [4,33,56–59]. In such models, the Wannier states are essentially the atomic p_z orbitals on each carbon atom and therefore the full azimuthal symmetry is retained, making the interlayer hoppings in-plane isotropic [see Eq. (7)], with no dependence on the three nearest-neighbor bond vectors $\mathbf{n}^{(\alpha)}_{j,S}(X)$ at the position of $X_{j,S}$ where $\alpha = 1, 2$, or 3 (see Fig. 1); and, similarly, no dependence on $\mathbf{n}^{(\alpha')}_{j',S'}(X')$. In the more detailed microscopic model derived from DFT determined Wannier states of the monolayer (and untwisted bilayer) graphene's conduction and valence bands [60], the localized state indeed has a dominant p_z character, but the azimuthal symmetry is lost due to the trigonal crystal field of the neighboring atoms. The localized state is therefore a superposition of several lattice harmonics with angular momenta $L_z = 0, 3, 6$, etc. As a consequence, the interlayer hopping part of t acquires the dependence on the relative orientation of the atomic separation vector $X_{j,S} - X'_{j',S'}$ and $\mathbf{n}^{(\alpha)}_{j,S}(X)$, $\mathbf{n}^{(\alpha')}_{j',S'}(X')$.

Here we generalize Eq. (15) to include such effects on the effective continuum Hamiltonian. In Eq. (4), we therefore replace

$$
t(X_{j,S} - X'_{j',S'}) \to t\left(X_{j,S} - X'_{j',S'}, \{\mathbf{n}^{(\alpha)}_{j,S}(X)\}, \{\mathbf{n}^{(\alpha)}_{j',S'}(X')\}\right), \tag{16}
$$

where the {} denotes the dependence on each term in the set, i.e., $\alpha = 1, 2, 3$. Next, from the definition of the nearest-neighbor vectors, we can write

$$\boldsymbol{n}_{j,S}^{(\alpha)}(X) = \boldsymbol{X}_{j,\bar{S}}^{\parallel}(\boldsymbol{r}_S + \boldsymbol{\delta}_S^{(\alpha)}) - \boldsymbol{X}_{j,S}^{\parallel}(\boldsymbol{r}_S), \qquad (17)$$

where $\boldsymbol{\delta}_S^{(\alpha)}$ are the three nearest-neighbor bond vectors of the undistorted lattice, that can be expressed as

$$\boldsymbol{\delta}_S^{(\alpha)} = R(2\pi(\alpha - 1)/3)\boldsymbol{\delta}_S^{(1)}, \qquad (18)$$

$$\boldsymbol{\delta}_S^{(1)} = \boldsymbol{\tau}_{\bar{S}} - \boldsymbol{\tau}_S \equiv \boldsymbol{\delta}_S, \qquad (19)$$

where $R(\omega)$ is the two-dimensional rotation matrix with the angle ω:

$$R(\omega) = \begin{pmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{pmatrix}. \qquad (20)$$

With our choice of the coordinate system, $\boldsymbol{\delta}_A^{(1)} = \boldsymbol{\tau}_B = \frac{1}{3}(\boldsymbol{a}_1 + \boldsymbol{a}_2)$, $\boldsymbol{\delta}_A^{(2)} = \frac{1}{3}(\boldsymbol{a}_2 - 2\boldsymbol{a}_1)$, $\boldsymbol{\delta}_A^{(3)} = \frac{1}{3}(\boldsymbol{a}_1 - 2\boldsymbol{a}_2)$, and $\boldsymbol{\delta}_B^{(\alpha)} = -\boldsymbol{\delta}_A^{(\alpha)}$.

We will consider only the atomic configurations which are varying smoothly not only within a sublattice but also between the two sublattices. All configurations examined in the companion paper are of this type [55]. Therefore, we can drop the S subscript in Eq. (3):

$$\boldsymbol{X}_{j,S}(\boldsymbol{r}) \equiv \boldsymbol{X}_j(\boldsymbol{r}) = \boldsymbol{r} + \boldsymbol{U}_j^{\parallel}(\boldsymbol{X}_j) + \boldsymbol{U}_j^{\perp}(\boldsymbol{X}_j). \qquad (21)$$

Correspondingly, the bond vectors in Eq. (17) become $\boldsymbol{n}_{j,S}^{(\alpha)}(X) = \boldsymbol{X}_j^{\parallel}(\boldsymbol{r}_S + \boldsymbol{\delta}_S^{\alpha}) - \boldsymbol{X}_j^{\parallel}(\boldsymbol{r}_S)$. Introducing a continuum variable $\boldsymbol{r}$ and changing the integration variable to $\boldsymbol{X}^{\parallel} = \boldsymbol{r} + \boldsymbol{u}_j^{\parallel}(\boldsymbol{r})$ for each j and S, as in the previous section makes the bond vectors $\boldsymbol{n}_{j,S}^{(\alpha)}$ a function of $\boldsymbol{X}^{\parallel}$. Therefore, t in Eq. (14) gains an additional dependence on $\boldsymbol{n}_{j,S}^{(\alpha)}(\boldsymbol{X}^{\parallel})$ and $\boldsymbol{n}_{j',S'}^{(\alpha)}(\boldsymbol{X}'^{\parallel})$. For a smooth atomic displacement field, we can then write

$$\boldsymbol{n}_{j,S}^{(\alpha)}(\boldsymbol{X}^{\parallel}) = \boldsymbol{\delta}_S^{\alpha} + \boldsymbol{u}_j^{\parallel}(\boldsymbol{X}^{\parallel} - \boldsymbol{U}_j^{\parallel}(\boldsymbol{X}^{\parallel}) + \boldsymbol{\delta}_S^{\alpha}) - \boldsymbol{U}_j(\boldsymbol{X}^{\parallel})$$

$$\simeq \boldsymbol{\delta}_S^{\alpha} + \delta_{S,\mu}^{\alpha} \frac{\partial \boldsymbol{u}_j^{\parallel}}{\partial r_\mu} \simeq \boldsymbol{\delta}_S^{\alpha} + \delta_{S,\mu}^{\alpha} \frac{\partial \boldsymbol{U}_j^{\parallel}(\boldsymbol{X}^{\parallel})}{\partial X_\mu^{\parallel}} \qquad (22)$$

because

$$\frac{\partial u_{j,\nu}^{\parallel}}{\partial r_\mu} = \frac{\partial (X_\nu^{\parallel} - r_\nu)}{\partial r_\mu} = \frac{\partial X_\nu^{\parallel}}{\partial r_\mu} - \delta_{\mu\nu} = \left(\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{X}}\right)_{\nu\mu}^{-1} - \delta_{\mu\nu}$$

$$= \left(\frac{\partial (\boldsymbol{X}^{\parallel} - \boldsymbol{U}_j^{\parallel})}{\partial \boldsymbol{X}^{\parallel}}\right)_{\nu\mu}^{-1} - \delta_{\mu\nu} \simeq \frac{\partial U_{j,\nu}^{\parallel}}{\partial X_\mu^{\parallel}}. \qquad (23)$$

By going to center-of-mass and relative coordinates, and keeping only the term up to the first-order derivative of $\boldsymbol{U}^{\parallel}$, we find

$$\boldsymbol{n}_{j,S}^{(\alpha)}\left(\boldsymbol{x} \pm \frac{1}{2}\boldsymbol{y}\right) \simeq \boldsymbol{\delta}_S^{(\alpha)} + \delta_{S,\mu}^{(\alpha)} \frac{\partial \boldsymbol{U}_j^{\parallel}(\boldsymbol{x})}{\partial x_\mu}. \qquad (24)$$

Following the arguments that led to Eq. (14), we find that $H_{\text{eff}}^{\mathbf{K}}$ can be obtained from Eq. (15) if for each layer index j, j', we drop the sublattice index on $\boldsymbol{U}$, i.e., we replace $\boldsymbol{U}_{j,S}^{\parallel,\perp}(\boldsymbol{x}) \rightarrow$

$\boldsymbol{U}_j^{\parallel,\perp}(\boldsymbol{x})$ and similarly for j', S', and we replace

$$t[\boldsymbol{d}_{S,S'}] \rightarrow t\left[\boldsymbol{d}, \left\{\boldsymbol{\delta}_S^{(\alpha)} + \delta_{S,\mu}^{(\alpha)} \frac{\partial \boldsymbol{U}_j^{\parallel}(\boldsymbol{x})}{\partial x_\mu}\right\}, \right.$$

$$\left. \left\{\boldsymbol{\delta}_{S'}^{(\alpha)} + \delta_{S',\mu}^{(\alpha)} \frac{\partial \boldsymbol{U}_{j'}^{\parallel}(\boldsymbol{x})}{\partial x_\mu}\right\}\right]. \qquad (25)$$

With these replacements, Eq. (15) gives the effective continuum Hamiltonian for the bond orientation dependent interlayer hopping for an arbitrary, sublattice-independent, smooth atomic deformation. The additional configuration dependent on-site term is discussed in the next subsection.

B. Bond-dependent on-site energy

The on-site terms in the tight-binding model need to be considered separately because in practice they may not be accounted for accurately by the continuous interpolation function t in the expression Eq. (25). We assume that the difference between the full configuration dependence of the on-site term and the contribution from t at $\boldsymbol{d} = 0$ can be approximated by the form

$$H_{\text{on-site}} = \sum_{j,S} \sum_{\boldsymbol{r}_S} \epsilon(\{|\boldsymbol{n}_{j,S}^{(\alpha)}(\boldsymbol{X}_{j,S})|\}) c_{j,S,\boldsymbol{r}_S}^{\dagger} c_{j,S,\boldsymbol{r}_S}, \qquad (26)$$

where the on-site energy ϵ is assumed to depend on the length of the three nearest bonds $\boldsymbol{n}_{j,S}^{(\alpha)}(X)$, defined in Eq. (22). Applying the same methods, we can write

$$H_{\text{on-site}} = \frac{1}{A_{\text{mlg}}} \sum_{j,S,\boldsymbol{G}} \int d^2\boldsymbol{r}\, e^{i\boldsymbol{G}\cdot(\boldsymbol{r}-\boldsymbol{\tau}_S)}$$

$$\times \epsilon(\{\boldsymbol{n}_{j,S}^{(\alpha)}(\boldsymbol{X}_{j,S})\}) c_{j,S,\boldsymbol{r}}^{\dagger} c_{j,S,\boldsymbol{r}}. \qquad (27)$$

Next, we introduce the field operator $\psi_{j,S}(\boldsymbol{r})$ via Eq. (9) to obtain the correction to the effective Hamiltonian at the valley $\boldsymbol{K}$ from the on-site term. Changing the integration variable $\boldsymbol{r}$ to $\boldsymbol{X}^{\parallel}$ introduces the Jacobi determinant $|J(\partial \boldsymbol{r}/\partial \boldsymbol{X}^{\parallel})|$. As shown in Eq. (12), this factor is absorbed by the redefinition of the field operator $\Psi_{j,S}(\boldsymbol{X}^{\parallel})$. In addition, $e^{i\boldsymbol{G}\cdot\boldsymbol{r}} = e^{i\boldsymbol{G}\cdot(\boldsymbol{X}^{\parallel}-\boldsymbol{U}^{\parallel}(\boldsymbol{X}^{\parallel}))}$. Thus, the on-site term at valley $\boldsymbol{K}$ can be written as

$$H_{\text{on-site}}^{\mathbf{K}} = \sum_{j,S,\boldsymbol{G}} e^{-i\boldsymbol{G}\cdot\boldsymbol{\tau}_S} \int d^2\boldsymbol{X}^{\parallel}\, e^{i\boldsymbol{G}\cdot(\boldsymbol{X}^{\parallel}-\boldsymbol{U}^{\parallel}(\boldsymbol{X}^{\parallel}))}$$

$$\times \epsilon(\{|\boldsymbol{n}_{j,S}^{(\alpha)}(\boldsymbol{X}^{\parallel})|\}) \Psi_{j,S}^{\dagger}(\boldsymbol{X}^{\parallel})\Psi_{j,S}(\boldsymbol{X}^{\parallel}). \qquad (28)$$

If $\boldsymbol{G} \neq 0$, the factor $e^{i\boldsymbol{G}\cdot\boldsymbol{X}^{\parallel}}$ oscillates around zero on the scale of the carbon-carbon distance and because it multiplies much more slowly, varying functions of $\boldsymbol{X}^{\parallel}$ the integral vanishes. Therefore, we can keep only the term with $\boldsymbol{G} = 0$ in the above sum and obtain

$$H_{\text{eff,on-site}}^{\mathbf{K}} = \sum_{j,S} \int d^2\boldsymbol{x}\, \epsilon(\{|\boldsymbol{n}_{j,S}^{(\alpha)}(\boldsymbol{x})|\}) \Psi_{j,S}^{\dagger}(\boldsymbol{x})\Psi_{j,S}(\boldsymbol{x}). \qquad (29)$$

To the linear order of gradients of $\boldsymbol{U}^{\parallel}$, the length of the distorted nearest-neighbor bond is

$$|\boldsymbol{n}_{j,S}^{(\alpha)}(\boldsymbol{x})| \approx |\boldsymbol{\delta}_S^{\alpha}| + \delta_{S,\mu}^{\alpha} \frac{\partial U_{j,\mu}^{\parallel}}{\partial x_\nu} \delta_{S,\nu}^{\alpha} / |\boldsymbol{\delta}_S^{\alpha}|, \qquad (30)$$

TABLE I. Parameters of the formula for the intralayer hopping with (left) $j = j'$ in Eq. (35) and interlayer hopping (right, and from Ref. [60]) $j \neq j'$ in Eq. (37).

i	0	1	i	0	3	6
$\tilde{\lambda}_i/\text{eV}$	-18.4295	-3.7183	λ_i/eV	0.3155	-0.0688	-0.0083
$\tilde{\xi}_i$	1.2771	6.2194	ξ_i	1.7543	3.4692	2.8764
$\tilde{x}_i$		0.9071	x_i		0.5212	1.5206
$\tilde{\kappa}_i$	2.3934		κ_i	2.0010		1.5731

where the length of the undistorted nearest-neighbor bond vectors is the same: $|\boldsymbol{\delta}_S^\alpha| = a/\sqrt{3}$. Thus, to leading order gradient expansion, the on-site energy ϵ is

$$\epsilon\left(\{|\boldsymbol{n}_{j,S}^{(\alpha)}(\boldsymbol{x})|\}\right) \approx \epsilon\left(\frac{a}{\sqrt{3}}\right) + \frac{\sqrt{3}}{a}\sum_{\alpha=1}^{3}\frac{\partial\epsilon_0}{\partial|\boldsymbol{\delta}_S^\alpha|}\delta_{S,\mu}^\alpha\frac{\partial U_{j,\mu}^\parallel}{\partial x_\nu}\delta_{S,\nu}^\alpha. \tag{31}$$

Due to C_3 and space inversion symmetries, $\partial\epsilon/\partial|\boldsymbol{\delta}_S^\alpha|$ is independent of α and S. In addition, $\sum_\alpha \delta_{S,\mu}^\alpha \delta_{S,\nu}^\alpha = \delta_{\mu\nu}a^2/2$. Introducing $\epsilon_0 = \epsilon(a/\sqrt{3})$ and $\kappa = \frac{\sqrt{3}}{2}a(\partial\epsilon/\partial|\boldsymbol{\delta}_S^\alpha|)$, we obtain

$$\epsilon\left(\{|\boldsymbol{n}_{j,S}^{(\alpha)}(\boldsymbol{x})|\}\right) \approx \epsilon_0 + \kappa\boldsymbol{\nabla}\cdot\boldsymbol{U}_j^\parallel. \tag{32}$$

Therefore, the contribution of the on-site term to the effective continuum Hamiltonian at $\boldsymbol{K}$ is

$$H_{\text{eff,on-site}}^{\mathbf{K}} = \sum_{j,S}\int d^2\boldsymbol{x}\,(\epsilon_0 + \kappa\boldsymbol{\nabla}\cdot\boldsymbol{U}_j^\parallel)\Psi_{j,S}^\dagger(\boldsymbol{x})\Psi_{j,S}(\boldsymbol{x}), \tag{33}$$

thus correcting the value of the deformation potential obtained from t alone.

C. Bond orientation dependent microscopic model of Ref. [60]

In the derivation above, we allow for a general form of the hopping, depending on all the nearest-neighbor bond vectors $\boldsymbol{n}_{j,S}^{(\alpha)}$ and $\boldsymbol{n}_{j',S'}^{(\alpha)}$. The model of Ref. [60] was derived for configurations which are locally C_3 symmetric, i.e., all three bond vectors $\boldsymbol{n}_{j,S}^{(\alpha)}$ are equivalent to each other, as are the three bond vectors $\boldsymbol{n}_{j',S'}^{(\alpha)}$. In this case, the bond dependence can be simplified because the hopping is the same for each one of the three bond vectors. With an eye toward generalizing to smooth lattice distortions which lead to a (small) violation of the local C_3 symmetry, we write the formula for the hoppings in Eq. (16) as

$$t\left(\boldsymbol{X}_{j,S} - \boldsymbol{X}_{j',S'}', \{\boldsymbol{n}_{j,S}^{(\alpha)}(\boldsymbol{X})\}, \{\boldsymbol{n}_{j',S'}^{(\alpha)}(\boldsymbol{X}')\}\right)$$
$$= \frac{1}{9}\sum_{\alpha=1}^{3}\sum_{\alpha'=1}^{3}t_{\text{sym}}^{jj'}\left(\boldsymbol{X}_{j,S} - \boldsymbol{X}_{j',S'}', \boldsymbol{n}_{j,S}^{(\alpha)}(\boldsymbol{X}), \boldsymbol{n}_{j',S'}^{(\alpha')}(\boldsymbol{X}')\}\right), \tag{34}$$

where $t_{\text{sym}}^{jj'}$ is the hopping function of Ref. [60] when the configuration is locally C_3 symmetric. For the intralayer hopping,

$$t_{\text{sym}}^{j=j'}\left(\boldsymbol{X}_{j,S} - \boldsymbol{X}_{j',S'}', \boldsymbol{n}_{j,S}^{(\alpha)}(\boldsymbol{X}), \boldsymbol{n}_{j',S'}^{(\alpha')}(\boldsymbol{X}')\right) = \tilde{V}_0(y), \tag{35}$$

where

$$\tilde{V}_0(y) = \tilde{\lambda}_0 e^{-\tilde{\xi}_0(y/a)^2}\cos\left(\tilde{\kappa}_0\frac{y}{a}\right) + \tilde{\lambda}_1\frac{y^2}{a^2}e^{-\tilde{\xi}_1(y/a-\tilde{x}_1)^2}, \tag{36}$$

where $\boldsymbol{y} = \boldsymbol{X}_{j,S}^\parallel - \boldsymbol{X}_{j',S'}'^\parallel$ is the in-plane projected separation vector, $y = |\boldsymbol{y}|$ is its magnitude. The intralayer hopping with $j = j'$ is rotationally isotropic, depending only on y. Note that its explicit formula is not provided by Ref. [60], in which the hopping constants are listed only for discrete values of y, i.e., for distances of several pairs of carbon atoms on the undistorted monolayer graphene lattice. To obtain the values of the hopping constants with arbitrary y, we fit these hopping constants with the formula in Eq. (35) and extract the parameters that are listed in the left table of Table I.

In the locally C_3 symmetric case, the interlayer part of the $t_{\text{sym}}^{jj'}$ depends only on two bond vectors, one at $\boldsymbol{X}_{j,S}$ and another one at $\boldsymbol{X}_{j',S'}'$, as

$$t_{\text{sym}}^{j\neq j'}(\boldsymbol{X}_{j,S} - \boldsymbol{X}_{j',S'}', \boldsymbol{n}_{j,S}, \boldsymbol{n}_{j',S'})$$
$$= V_0(y) + V_3(y)(\cos(3\theta_{12}) + \cos(3\theta_{21}))$$
$$+ V_6(y)(\cos(6\theta_{12}) + \cos(6\theta_{21})). \tag{37}$$

The explicit formulas for $V_i(y)$ are presented in Ref. [60] and we include them here for completeness:

$$V_0(y) = \lambda_0 e^{-\xi_0(y/a)^2}\cos\left(\kappa_0\frac{y}{a}\right), \tag{38}$$

$$V_3(y) = \lambda_3\frac{y^2}{a^2}e^{-\xi_3(y/a-x_3)^2}, \tag{39}$$

$$V_6(y) = \lambda_6 e^{-\xi_6(y/a-x_6)^2}\sin\left(\kappa_6\frac{y}{a}\right). \tag{40}$$

The parameters are specified in Table I.

The variables θ_{12} and θ_{21} in Eq. (37) are the angles between $\boldsymbol{y}$ and the nearest-neighbor bond vectors on two layers, i.e.,

$$\theta_{12} = \cos^{-1}\left(-\frac{\boldsymbol{y}\cdot\boldsymbol{n}_{j,S}}{y|\boldsymbol{n}_{j,S}|}\right) = \theta_{\boldsymbol{y}} - \theta_{j,S} + \pi, \tag{41}$$

$$\theta_{21} = \cos^{-1}\left(\frac{\boldsymbol{y}\cdot\boldsymbol{n}_{j',S'}}{y|\boldsymbol{n}_{j',S'}|}\right) = \theta_{\boldsymbol{y}} - \theta_{j',S'}. \tag{42}$$

In the above, we defined $\theta_{\boldsymbol{y}}$ to be the angle between the separation vector $\boldsymbol{y}$ and the x axis, and $\theta_{j,S}$ ($\theta_{j',S'}$) to be the angle between the bond vector $\boldsymbol{n}_{j,S}$ ($\boldsymbol{n}_{j',S'}$) and the x axis. $\theta_{j,S}^{(\alpha)}$ ($\theta_{j',S'}^{(\alpha)}$) is introduced similarly but with the superscript α to distinguish the angle of different bond vectors. In the absence of the lattice distortion (e.g., as for a rigid twist), the three in-plane nearest neighbors of a carbon atom are C_3 symmetric about the carbon atom, and $\theta_{j,S}^{(\alpha)} = \theta_{j,S}^{(1)} + 2\pi(\alpha-1)/3$. Therefore, the angles θ_{12} and θ_{21} could differ by $2\pi/3$ if choosing a different nearest-neighbor bond, leading to the

same $\cos(3m\theta_{12})$ and $\cos(3m\theta_{21})$, with m being an integer. Therefore, without distortions, each term in the sum on the right-hand side of Eq. (34) is the same and the sum is redundant. In the presence of the lattice relaxation, however, the local C_3 symmetry is, in general, broken and the bond vectors become inequivalent. To generalize t to include such slowly varying atomic displacements, we use the formula Eq. (34). With the local C_3 symmetry broken, the difference between the angles $\theta_{j,S}^{(\alpha)}$ deviates from $\pm 2\pi/3$. For smooth lattice deformation, the deviation is small. To obtain this deviation, we write $\boldsymbol{n}_{j,S}^{(\alpha)} = \boldsymbol{\delta}_S^{(\alpha)} + \delta\boldsymbol{n}_{j,S}^{(\alpha)}$ with $\delta\boldsymbol{n}_{j,S}^{(\alpha)} = \delta_{S,\mu}^{(\alpha)} \frac{\partial U_j^{\parallel}}{\partial x_\mu}$ and expand the angle $\theta_{j,S}^{(\alpha)}$ to the linear order of the derivatives of $\boldsymbol{U}_j^{\parallel}$ as

$$\theta_{j,S}^{(\alpha)} = \theta_{\boldsymbol{\delta}_S^{(\alpha)}} + \delta\theta_{j,S}^{(\alpha)}, \tag{43}$$

$$\delta\theta_{j,S}^{(\alpha)} = \frac{(\hat{z} \times \boldsymbol{\delta}_S^{(\alpha)}) \cdot \delta\boldsymbol{n}_{j,S}^{(\alpha)}}{\left|\boldsymbol{\delta}_S^{(\alpha)}\right|^2} = \frac{\epsilon_{\mu\nu}}{\left|\boldsymbol{\delta}_S^{(\alpha)}\right|^2} \delta_{S,\mu}^{(\alpha)} \frac{\partial U_{j,\nu}^{\parallel}}{\partial x_\rho} \delta_{S,\rho}^{(\alpha)}. \tag{44}$$

$\theta_{\boldsymbol{\delta}_B^{(\alpha)}} = \theta_{\boldsymbol{\delta}_A^{(\alpha)}} + \pi$, and for our choice of the coordinate system, $\theta_{\boldsymbol{\delta}_A^{(1)}} = \pi/6$, $\theta_{\boldsymbol{\delta}_A^{(2)}} = \pi/6 + 2\pi/3$, $\theta_{\boldsymbol{\delta}_A^{(3)}} = \pi/6 - 2\pi/3$.

For the interlayer part, we therefore introduce derivatives of $t_{\text{sym}}^{jj'}$ with respect to the angles as

$$t_{j \neq j',S}^{(1)}(\boldsymbol{y}) = \left. \frac{\partial t_{\text{sym}}^{j \neq j'}}{\partial \theta_{j,S}} \right|_{\theta_{j,S} = \theta_{\delta_S}}$$

$$= -3V_3(y)\sin(3(\theta_{\boldsymbol{y}} - \theta_{\delta_S}))$$
$$+ 6V_6(y)\sin(6(\theta_{\boldsymbol{y}} - \theta_{\delta_S})), \tag{45}$$

$$t_{j \neq j',S'}^{(2)}(\boldsymbol{y}) = \left. \frac{\partial t_{\text{sym}}^{j \neq j'}}{\partial \theta_{j',S'}} \right|_{\theta_{j',S'} = \theta_{\delta_{S'}}}$$

$$= 3V_3(y)\sin(3(\theta_{\boldsymbol{y}} - \theta_{\delta_{S'}}))$$
$$+ 6V_6(y)\sin(6(\theta_{\boldsymbol{y}} - \theta_{\delta_{S'}})), \tag{46}$$

and vanishing for the intralayer part:

$$t_{j=j',S}^{(1)}(\boldsymbol{y}) = t_{j=j',S'}^{(2)}(\boldsymbol{y}) = 0. \tag{47}$$

The above expressions are clearly independent under $\theta_{\delta_S} \to \theta_{\delta_S} \pm 2\pi/3$, and therefore it does not matter which $\theta_{\boldsymbol{\delta}_S^{(\alpha)}}$ is substituted for θ_{δ_S}. Thus, combining with Eq. (33), for an arbitrary smooth lattice deformation, the effective continuum Hamiltonian for the lattice model of Ref. [60] is

$$H_{\text{eff}}^{\mathbf{K}} \simeq \frac{1}{A_{\text{mlg}}} \sum_{S,S'} \sum_{jj'} \sum_{\boldsymbol{G}} e^{i\boldsymbol{G}\cdot(\boldsymbol{\tau}_S - \boldsymbol{\tau}_{S'})} \int d^2\boldsymbol{x}\, \mathcal{J}_j(\boldsymbol{x})\mathcal{J}_{j'}(\boldsymbol{x}) e^{i(\boldsymbol{G}+\boldsymbol{K})\cdot(\boldsymbol{U}_j^{\parallel}(\boldsymbol{x}) - \boldsymbol{U}_{j'}^{\parallel}(\boldsymbol{x}))} \int d^2\boldsymbol{y}\, e^{-i(\boldsymbol{G}+\boldsymbol{K})\cdot\boldsymbol{y}},\ e^{i\frac{\boldsymbol{y}}{2}\cdot\nabla_{\boldsymbol{x}}(\boldsymbol{U}_j^{\parallel}(\boldsymbol{x}) + \boldsymbol{U}_{j'}^{\parallel}(\boldsymbol{x}))\cdot(\boldsymbol{G}+\boldsymbol{K})}$$

$$\times \left(t_{\text{sym}}^{jj'}[\boldsymbol{y} + \boldsymbol{U}_j^{\perp}(\boldsymbol{x}) - \boldsymbol{U}_{j'}^{\perp}(\boldsymbol{x}), \boldsymbol{\delta}_S, \boldsymbol{\delta}_{S'}] + t_{jj',S}^{(1)}(\boldsymbol{y})\frac{1}{3}\sum_{\alpha=1}^{3}\delta\theta_{j,S}^{(\alpha)} + t_{jj',S'}^{(2)}(\boldsymbol{y})\frac{1}{3}\sum_{\alpha'=1}^{3}\delta\theta_{j',S'}^{(\alpha')} \right) \left[\Psi_{j,S}^{\dagger}(\boldsymbol{x})\Psi_{j',S'}(\boldsymbol{x}) \right.$$

$$\left. + \frac{\boldsymbol{y}}{2}\cdot((\nabla_{\boldsymbol{x}}\Psi_{j,S}^{\dagger}(\boldsymbol{x}))\Psi_{j',S'}(\boldsymbol{x}) - \Psi_{j,S}^{\dagger}(\boldsymbol{x})\nabla_{\boldsymbol{x}}\Psi_{j',S'}(\boldsymbol{x})) \right] + \sum_{j,S}\int d^2\boldsymbol{x}(\epsilon_0 + \kappa\nabla\cdot\boldsymbol{U}_j^{\parallel}(\boldsymbol{x}))\Psi_{j,S}^{\dagger}(\boldsymbol{x})\Psi_{j,S}(\boldsymbol{x}). \tag{48}$$

The comparison between the continuum and tight-binding spectra for the model of Ref. [60] for rigid twist as well as for the (relaxed) atomic configurations obtained from solving continuum elastic theory for twisted bilayer are shown in the companion paper [55].

III. DISCUSSION

In our derivation of the continuum effective Hamiltonians $H_{\text{eff}}^{\mathbf{K}}$ for graphene bilayers, we have not made use of symmetries. Although this might seem reasonable given that we are considering arbitrary smooth inhomogeneous atomic configurations which would remove any remaining symmetries, as pointed out by Balents [37], the form of the leading order terms in the effective Hamiltonian can nevertheless be further constrained. That is because $H_{\text{eff}}^{\mathbf{K}}$ must be invariant under symmetry operations of the undistorted lattice (i.e., the AA-stacked bilayer) that leave a valley invariant if we *simultaneously* transform the fermion operators *and* the atomic displacement fields [37].

Although we postpone the detailed analysis of the symmetry, here and in Appendix B we would like to highlight some of its consequences. The symmetries of interest to us will be C_3, $C_2\mathcal{T}$, C_{2x}, and $\mathcal{R}_y$. Here C_3 is the threefold rotation along z

axis. $C_2\mathcal{T}$ is the time reversal followed by the twofold rotation along z axis with the two sublattices interchanged. $\mathcal{R}_y$ is the mirror reflection along xz plane bisecting the nearest-neighbor carbon bond so $(x, y, z) \to (x, -y, z)$ and the sublattice index $A \leftrightarrow B$. And C_{2x} is $\mathcal{R}_y$ followed by the interchange of the two layers, i.e., followed by the xy plane mirror reflection halfway between the layers $\mathcal{R}_z$.

The consequences of $C_2\mathcal{T}$ and $\mathcal{R}_y$ at $\mathbf{G} = 0$ for the contact interlayer tunneling term—independent of the spatial gradients of the atomic displacement, i.e., to zeroth order in $\nabla_x U$—were worked out in Ref. [37]. There it was shown that, when combined with C_3, only two independent real parameters are allowed for the first shell of wave vectors $\mathbf{G} = 0, -4\pi \boldsymbol{a}_2 \times \hat{z}/\sqrt{3}a^2, 4\pi(\boldsymbol{a}_1 - \boldsymbol{a}_2) \times \hat{z}/\sqrt{3}a^2$. Physically, these correspond to the interlayer tunneling through the AA region and the AB region, and are the only interlayer tunneling terms kept in the BM model [6,37].

As mentioned in the Introduction, the anomalous decrease of the bandwidth near the magic twist angle promotes the importance of the next-to-leading order terms in setting the anisotropies, thus selecting from the nearly degenerate manifold of correlated states that are obtained if only the leading order terms are kept. Instead of listing all the consequences of the above symmetries on such higher order terms, here we

only mention in passing that $C_2\mathcal{T}$ and the combined operation $C_{2x}\mathcal{R}_y$ will be seen to allow for a particularly interesting inter-layer tunneling contact term which, as shown in the companion paper [55], is the main source of the particle-hole symmetry breaking in the model of Ref. [60], but which is altogether absent in SK-type models [4,33,59].

APPENDIX A: QUADRATIC ORDER

In Sec. II, we expand the effective continuum model to the first order of $\boldsymbol{y} = \boldsymbol{X}^{\parallel} - \boldsymbol{X}'^{\parallel}$. Numerically, we found that the intralayer part needs to be expanded to the second order of $\boldsymbol{y}$ to achieve the agreement between the two dispersion produced by H_{eff}^K and the microscopic tight-binding model H_{tb}. To the second order of $\boldsymbol{y}$, we have

$$\Psi_{j,S}^{\dagger}\left(\boldsymbol{x} + \frac{\boldsymbol{y}}{2}\right)\Psi_{j',S'}\left(\boldsymbol{x} - \frac{\boldsymbol{y}}{2}\right) \simeq \Psi_{j,S}^{\dagger}(\boldsymbol{x})\Psi_{j',S'}(\boldsymbol{x}) + \frac{\boldsymbol{y}}{2} \cdot [(\nabla\Psi_{j,S}^{\dagger}(\boldsymbol{x}))\Psi_{j',S'}(\boldsymbol{x}) - \Psi_{j,S}^{\dagger}(\boldsymbol{x})(\nabla\Psi_{j',S'}(\boldsymbol{x}))]$$

$$+ \frac{1}{8}\boldsymbol{y}^{\mu}\boldsymbol{y}^{\nu}[(\partial_{\mu}\partial_{\nu}\Psi_{j,S}^{\dagger}(\boldsymbol{x}))\Psi_{j',S'}(\boldsymbol{x}) - 2(\partial_{\mu}\Psi_{j,S}^{\dagger}(\boldsymbol{x}))(\partial_{\nu}\Psi_{j',S'}(\boldsymbol{x})) + \Psi_{j,S}^{\dagger}(\boldsymbol{x})(\partial_{\mu}\partial_{\nu}\Psi_{j',S'}(\boldsymbol{x}))]. \quad \text{(A1)}$$

APPENDIX B: SYMMETRY

For a smooth but otherwise arbitrary $\boldsymbol{U}_{j,S}^{\parallel,\perp}(\boldsymbol{x})$, the effective Hamiltonian $H_{\text{eff}}^{\mathbf{K}}$ in Eq. (15) is invariant under $C_2\mathcal{T}$ if we simultaneously transform the fermion operators *and* the atomic displacement fields [37] as

$$C_2\mathcal{T}: \ \Psi_{j,S}(\boldsymbol{x}) \longrightarrow \Psi_{j,\bar{S}}(-\boldsymbol{x}), \quad \text{(B1)}$$

$$\boldsymbol{U}_{j,S}^{\parallel}(\boldsymbol{x}) \longrightarrow -\boldsymbol{U}_{j,\bar{S}}^{\parallel}(-\boldsymbol{x}), \quad \text{(B2)}$$

$$\boldsymbol{U}_{j,S}^{\perp}(\boldsymbol{x}) \longrightarrow \boldsymbol{U}_{j,\bar{S}}^{\perp}(-\boldsymbol{x}), \quad \text{(B3)}$$

where $\bar{S}$ is the sublattice index different from S, provided the microscopic hopping function $t(\boldsymbol{y} + \boldsymbol{U}_{j,S}^{\perp} - \boldsymbol{U}_{j',S'}^{\perp}) = t^*(-\boldsymbol{y} + \boldsymbol{U}_{j,S}^{\perp} - \boldsymbol{U}_{j',S'}^{\perp})$. This is certainly satisfied for the SK-type model Eq. (7); $C_2\mathcal{T}$ is also satisfied for the orientation-dependent hopping function of Ref. [60].

The consequences of this symmetry for the contact term of the interlayer tunneling part of $H_{\text{eff}}^{\mathbf{K}}$ [i.e., the first term in the third line of Eq. (15)] can be seen if we assume that the $\boldsymbol{U}_{j,S}^{\parallel,\perp}(\boldsymbol{x})$ is independent of S. Then the said term can be expressed as

$$\sum_{\boldsymbol{G}} \int d^2\boldsymbol{x} \ e^{i(\boldsymbol{G}+\boldsymbol{K})\cdot(\boldsymbol{U}_t^{\parallel}(\boldsymbol{x})-\boldsymbol{U}_b^{\parallel}(\boldsymbol{x}))}\Psi_{t,S}^{\dagger}(\boldsymbol{x})\Psi_{b,S'}(\boldsymbol{x})T_{SS'}^{\boldsymbol{G}}(\nabla_{\boldsymbol{x}}\boldsymbol{U}_t^{\parallel}(\boldsymbol{x}),$$

$$\nabla_{\boldsymbol{x}}\boldsymbol{U}_b^{\parallel}(\boldsymbol{x}), \boldsymbol{U}_t^{\perp}(\boldsymbol{x}) - \boldsymbol{U}_b^{\perp}(\boldsymbol{x})) + \text{H.c.} \quad \text{(B4)}$$

Note that Eqs. (5) and (6) guarantee that $H_{\text{eff}}^{\mathbf{K}}$ is Hermitian. Now, $C_2\mathcal{T}$ forces

$$T_{SS'}^{\boldsymbol{G}}(\nabla_{\boldsymbol{x}}\boldsymbol{U}_t^{\parallel}(\boldsymbol{x}), \nabla_{\boldsymbol{x}}\boldsymbol{U}_b^{\parallel}(\boldsymbol{x}), \boldsymbol{U}_t^{\perp}(\boldsymbol{x}) - \boldsymbol{U}_b^{\perp}(\boldsymbol{x}))$$

$$= \sigma_{SS_1}^x T_{S_1 S_2}^{\boldsymbol{G}*}(\nabla_{\boldsymbol{x}}\boldsymbol{U}_t^{\parallel}(\boldsymbol{x}), \nabla_{\boldsymbol{x}}\boldsymbol{U}_b^{\parallel}(\boldsymbol{x}), \boldsymbol{U}_t^{\perp}(\boldsymbol{x}) - \boldsymbol{U}_b^{\perp}(\boldsymbol{x}))\sigma_{S_2 S'}^x. \quad \text{(B5)}$$

This implies that (temporarily suppressing its arguments)

$$T_{SS'}^{\boldsymbol{G}} = 1_{SS'}W_0^{\boldsymbol{G}} + \sigma_{SS'}^x W_1^{\boldsymbol{G}} + \sigma_{SS'}^y W_2^{\boldsymbol{G}} + i\sigma_{SS'}^z W_3^{\boldsymbol{G}}, \quad \text{(B6)}$$

where $W_j^{\boldsymbol{G}}(\nabla_{\boldsymbol{x}}\boldsymbol{U}_t^{\parallel}(\boldsymbol{x}), \nabla_{\boldsymbol{x}}\boldsymbol{U}_b^{\parallel}(\boldsymbol{x}), \boldsymbol{U}_t^{\perp}(\boldsymbol{x}) - \boldsymbol{U}_b^{\perp}(\boldsymbol{x}))$ are purely real functions.

Another useful constraint can be obtained from the combination of C_{2x} and $\mathcal{R}_y$,

$$C_{2x}\mathcal{R}_y: \ \Psi_{j,S}(\boldsymbol{x}) \longrightarrow \Psi_{\bar{j},S}(\boldsymbol{x}), \quad \text{(B7)}$$

$$\boldsymbol{U}_{j,S}^{\parallel}(\boldsymbol{x}) \longrightarrow \boldsymbol{U}_{\bar{j},S}^{\parallel}(\boldsymbol{x}), \quad \text{(B8)}$$

$$\boldsymbol{U}_{j,S}^{\perp}(\boldsymbol{x}) \longrightarrow -\boldsymbol{U}_{\bar{j},S}^{\perp}(\boldsymbol{x}), \quad \text{(B9)}$$

under which $H_{\text{eff}}^{\mathbf{K}}$ in Eq. (15) is also invariant if $t(\boldsymbol{y} + \boldsymbol{U}_{j,S}^{\perp} - \boldsymbol{U}_{j',S'}^{\perp}) = t(-\boldsymbol{y} + \boldsymbol{U}_{j,S}^{\perp} - \boldsymbol{U}_{j',S'}^{\perp})$, as satisfied by the SK-type model Eq. (7); $C_{2x}\mathcal{R}_y$ is also a symmetry of the orientation-dependent hopping function of Ref. [60]. Then, $C_{2x}\mathcal{R}_y$ forces

$$T_{SS'}^{\boldsymbol{G}}(\nabla_{\boldsymbol{x}}\boldsymbol{U}_t^{\parallel}(\boldsymbol{x}), \nabla_{\boldsymbol{x}}\boldsymbol{U}_b^{\parallel}(\boldsymbol{x}), \boldsymbol{U}_t^{\perp}(\boldsymbol{x}) - \boldsymbol{U}_b^{\perp}(\boldsymbol{x}))$$

$$= T_{S'S}^{\boldsymbol{G}*}(\nabla_{\boldsymbol{x}}\boldsymbol{U}_b^{\parallel}(\boldsymbol{x}), \nabla_{\boldsymbol{x}}\boldsymbol{U}_t^{\parallel}(\boldsymbol{x}), \boldsymbol{U}_t^{\perp}(\boldsymbol{x}) - \boldsymbol{U}_b^{\perp}(\boldsymbol{x})). \quad \text{(B10)}$$

The anti-Hermiticity of $i\sigma^z$ implies that $W_3^{\mathbf{G}}$ must be *odd* under $\nabla_{\boldsymbol{x}}\boldsymbol{U}_t^{\parallel}(\boldsymbol{x}) \leftrightarrow \nabla_{\boldsymbol{x}}\boldsymbol{U}_b^{\parallel}(\boldsymbol{x})$. But, the contact interlayer tunneling term in Eq. (15) is clearly *even* under $\nabla_{\boldsymbol{x}}\boldsymbol{U}_t^{\parallel}(\boldsymbol{x}) \leftrightarrow \nabla_{\boldsymbol{x}}\boldsymbol{U}_b^{\parallel}(\boldsymbol{x})$. Therefore, for the SK-type models, $W_3^{\mathbf{G}} = 0$. This is indeed what we find from the detailed analysis presented in the companion paper [55]. On the other hand, for the model based on the *ab initio* hopping integrals [60], there is an additional dependence of the hoppings on the orientation of hopping vector to the nearest-neighbor vectors. In this case, $W_3^{\mathbf{G}} \neq 0$ even for a rigid twist. As we show in the companion paper [55], this term gives the largest contribution to the particle-hole asymmetry.

[1] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, R. C. Ashoori, and P. Jarillo-Herrero, Correlated insulator behaviour at half-filling in magic-angle graphene superlattices, Nature (London) **556**, 43 (2018).

[2] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Unconventional superconductivity in magic-angle graphene superlattices, Nature (London) **556**, 80 (2018).

[3] 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).

[4] G. Trambly de Laissardiére, D. Mayou, and L. Magaud, Localization of Dirac electrons in rotated graphene bilayers, Nano Lett. **10**, 804 (2010).

[5] E. Suárez Morell, J. D. Correa, P. Vargas, M. Pacheco, and Z. Barticevic, Flat bands in slightly twisted bilayer graphene: Tight-binding calculations, Phys. Rev. B **82**, 121407(R) (2010).

[6] R. Bistritzer and A. H. MacDonald, Moire bands in twisted double-layer graphene, Proc. Natl. Acad. Sci. USA **108**, 12233 (2011).

[7] M. Yankowitz, S. Chen, H. Polshyn, Y. Zhang, K. Watanabe, T. Taniguchi, D. Graf, A. F. Young, and C. R. Dean, Tuning superconductivity in twisted bilayer graphene, Science **363**, 1059 (2019).

[8] A. L. Sharpe, E. J. Fox, A. W. Barnard, J. Finney, K. Watanabe, T. Taniguchi, M. A. Kastner, and D. Goldhaber-Gordon, Emergent ferromagnetism near three-quarters filling in twisted bilayer graphene, Science **365**, 605 (2019).

[9] M. Serlin, C. L. Tschirhart, H. Polshyn, Y. Zhang, J. Zhu, K. Watanabe, T. Taniguchi, L. Balents, and A. F. Young, Intrinsic quantized anomalous Hall effect in a moire heterostructure, Science **367**, 900 (2019).

[10] A. Kerelsky, L. J. McGilly, D. M. Kennes, L. Xian, M. Yankowitz, S. Chen, K. Watanabe, T. Taniguchi, J. Hone, C. Dean, A. Rubio, and A. N. Pasupathy, Maximized electron interactions at the magic angle in twisted bilayer graphene, Nature (London) **572**, 95 (2019).

[11] Y. Xie, B. Lian, B. Jack, X. Liu, C.-L. Chiu, K. Watanabe, T. Taniguchi, B. A. Bernevig, and A. Yazdani, Spectroscopic signatures of many body correlations in magic-angle twisted bilayer graphene, Nature (London) **572**, 101 (2019).

[12] S. L. Tomarken, Y. Cao, A. Demir, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, and R. C. Ashoori, Electronic Compressibility of Magic-Angle Graphene Superlattices, Phys. Rev. Lett. **123**, 046601 (2019).

[13] Y. Jiang, X. Lai, K. Watanabe, T. Taniguchi, K. Haule, J. Mao, and E. Y. Andrei, Charge order and broken rotational symmetry in magic-angle twisted bilayer graphene, Nature (London) **573**, 91 (2019).

[14] X. Lu, P. Stepanov, W. Yang, M. Xie, M. A. Aamir, I. Das, C. Urgell, K. Watanabe, T. Taniguchi, G. Zhang, A. Bachtold, A. H. MacDonald, and D. K. Efetov, Superconductors, orbital magnets and correlated states in magic-angle bilayer graphene, Nature (London) **574**, 653 (2019).

[15] Y. Choi, J. Kemmer, Y. Peng, A. Thomson, H. Arora, R. Polski, Y. Zhang, H. Ren, J. Alicea, G. Refael, F. von Oppen, K. Watanabe, T. Taniguchi, and S. Nadj-Perge, Electronic correlations in twisted bilayer graphene near the magic angle, Nat. Phys. **15**, 1174 (2019).

[16] A. Uri, S. Grover, Y. Cao, J. A. Crosse, K. Bagani, D. Rodan-Legrain, Y. Myasoedov, K. Watanabe, T. Taniguchi, P. Moon, M. Koshino, P. Jarillo-Herrero, and E. Zeldov, Mapping the twist-angle disorder and Landau levels in magic-angle graphene, Nature (London) **581**, 47 (2020).

[17] D. Wong, K. P. Nuckolls, M. Oh, B. Lian, Y. Xie, S. Jeon, K. Watanabe, T. Taniguchi, B. A. Bernevig, and A. Yazdani, Cascade of electronic transitions in magic-angle twisted bilayer graphene, Nature (London) **582**, 198 (2020).

[18] U. Zondiner, A. Rozen, D. Rodan-Legrain, Y. Cao, R. Queiroz, T. Taniguchi, K. Watanabe, Y. Oreg, F. von Oppen, A. Stern, E. Berg, P. Jarillo-Herrero, and S. Ilani, Cascade of phase transitions and dirac revivals in magic angle graphene, Nature (London) **582**, 203 (2020).

[19] P. Stepanov, I. Das, X. Lu, A. Fahimniya, K. Watanabe, T. Taniguchi, F. H. L. Koppens, J. Lischner, L. Levitov, D. K. Efetov, Untying the insulating and superconducting orders in magic-angle graphene, Nature (London) **583**, 375 (2020).

[20] Y. Saito, J. Ge, K. Watanabe, T. Taniguchi, and A. F. Young, Independent superconductors and correlated insulators in twisted bilayer graphene, Nat. Phys. **16**, 926 (2020).

[21] X. Liu, Z. Wang, K. Watanabe, T. Taniguchi, O. Vafek, and J. I. A. Li, Tuning electron correlation in magic-angle twisted bilayer graphene using Coulomb screening, Science **371**, 1261 (2021).

[22] Y. Cao, D. Rodan-Legrain, J. M. Park, F. N. Yuan, K. Watanabe, T. Taniguchi, R. M. Fernandes, L. Fu, and P. Jarillo-Herrero, Nematicity and competing orders in superconducting magic-angle graphene, Science **372**, 264 (2021).

[23] C. L. Tschirhart, M. Serlin, H. Polshyn, A. Shragai, Z. Xia, J. Zhu, Y. Zhang, K. Watanabe, T. Taniguchi, M. E. Huber, A. F. Young, Imaging orbital ferromagnetism in a moire Chern insulator, Science **372**, 1323 (2021).

[24] J. M. Park, Y. Cao, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, Flavour Hund's coupling, correlated Chern gaps, and diffusivity in Moiré flat bands, Nature (London) **592**, 43 (2021).

[25] A. Rozen, J. M. Park, U. Zondiner, Y. Cao, D. Rodan-Legrain, T. Taniguchi, K. Watanabe, Y. Oreg, A. Stern, E. Berg, P. Jarillo-Herrero, and S. Ilani, Entropic evidence for a Pomeranchuk effect in magic angle graphene, Nature (London) **592**, 214 (2021).

[26] Y. Saito, F. Yang, J. Ge, X. Liu, K. Watanabe, T. Taniguchi, J. I. A. Li, E. Berg, and A. F. Young, Isospin Pomeranchuk effect and the entropy of collective excitations in twisted bilayer graphene, Nature (London) **592**, 220 (2021).

[27] M. Oh, K. P. Nuckolls, D. Wong, R. L. Lee, X. Liu, K. Watanabe, T. Taniguchi, and A. Yazdani, Evidence for unconventional superconductivity in twisted bilayer graphene, Nature (London) **600**, 240 (2021).

[28] Y. Xie, A. T. Pierce, J. M. Park, D. E. Parker, E. Khalaf, P. Ledwith, Y. Cao, S. H. Lee, S. Chen, P. R. Forrester, K. Watanabe, T. Taniguchi, A. Vishwanath, P. Jarillo-Herrero, and A. Yacoby, Fractional Chern insulators in magic-angle twisted bilayer graphene, Nature (London) **600**, 439 (2021).

[29] A. T. Pierce, Y. Xie, J. M. Park, E. Khalaf, S. Hwan Lee, Y. Cao, D. E. Parker, P. R. Forrester, S. Chen, K. Watanabe, T. Taniguchi, A. Vishwanath, P. Jarillo-Herrero, and A. Yacoby, Unconventional sequence of correlated Chern insulators in

magic-angle twisted bilayer graphene, Nat. Phys. **17**, 1210 (2021).

[30] J.-X. Lin, Y.-H. Zhang, E. Morissette, Z. Wang, S. Liu, D. Rhodes, K. Watanabe, T. Taniguchi, J. Hone, and JIA Li, Spin-orbit driven ferromagnetism at half moiré filling in magic-angle twisted bilayer graphene, Science **375**, 437 (2022).

[31] H. Polshyn, Y. Zhang, M. A. Kumar, T. Soejima, P. Ledwith, K. Watanabe, T. Taniguchi, A. Vishwanath, M. P. Zaletel, and A. F. Young, Topological charge density waves at half-integer filling of a moiré superlattice, Nat. Phys. **18**, 42 (2022).

[32] M. Koshino, N. F. Q. Yuan, T. Koretsune, M. Ochi, K. Kuroki, and L. Fu, Maximally Localized Wannier Orbitals and the Extended Hubbard Model for Twisted Bilayer Graphene, Phys. Rev. X **8**, 031087 (2018).

[33] J. Kang and O. Vafek, Symmetry, maximally localized Wannier states, and a low-energy model for twisted bilayer graphene narrow bands, Phys. Rev. X **8**, 031088 (2018).

[34] H. C. Po, L. Zou, A. Vishwanath, and T. Senthil, Origin of Mott insulating behavior and superconductivity in twisted bilayer graphene, Phys. Rev. X **8**, 031089 (2018).

[35] F. Wu, A. H. MacDonald, and I. Martin, Theory of Phonon-Mediated Superconductivity in Twisted Bilayer Graphene, Phys. Rev. Lett. **121**, 257001 (2018).

[36] F. Guinea and N. R. Walet, Electrostatic effects, band distortions, and superconductivity in twisted graphene bilayers, Proc. Natl. Acad. Sci. USA **115**, 13174 (2018).

[37] L. Balents, General continuum model for twisted bilayer graphene and arbitrary smooth deformations, SciPost Phys. **7**, 048 (2019).

[38] J. Ahn, S. Park, and B.-J. Yang, Failure of Nielsen-Ninomiya theorem and fragile topology in two-dimensional systems with space-time inversion symmetry: Application to twisted bilayer graphene at magic angle, Phys. Rev. X **9**, 021013 (2019).

[39] Z. Song, Z. Wang, W. Shi, G. Li, C. Fang, and B. A. Bernevig, All Magic Angles in Twisted Bilayer Graphene are Topological, Phys. Rev. Lett. **123**, 036401 (2019).

[40] K. Hejazi, C. Liu, H. Shapourian, X. Chen, and L. Balents, Multiple topological transitions in twisted bilayer graphene near the first magic angle, Phys. Rev. B **99**, 035111 (2019).

[41] J. Liu, J. Liu, and X. Dai, The pseudo-Landau-level representation of twisted bilayer graphene: Band topology and the implications on the correlated insulating phase, Phys. Rev. B **99**, 155415 (2019).

[42] G. Tarnopolsky, A. J. Kruchkov, and A. Vishwanath, Origin of Magic Angles in Twisted Bilayer Graphene, Phys. Rev. Lett. **122**, 106405 (2019).

[43] J. Kang and O. Vafek, Strong Coupling Phases of Partially Filled Twisted Bilayer Graphene Narrow Bands, Phys. Rev. Lett. **122**, 246401 (2019).

[44] Y.-H. Zhang, D. Mao, Y. Cao, P. Jarillo-Herrero, and T. Senthil, Nearly flat Chern bands in moiré superlattices, Phys. Rev. B **99**, 075127 (2019).

[45] M. Xie, and A. H. MacDonald, Nature of the Correlated Insulator States in Twisted Bilayer Graphene, Phys. Rev. Lett. **124**, 097601 (2020).

[46] N. Bultinck, S. Chatterjee, and M. P. Zaletel, Mechanism for Anomalous Hall Ferromagnetism in Twisted Bilayer Graphene, Phys. Rev. Lett. **124**, 166601 (2020).

[47] N. Bultinck, E. Khalaf, S. Liu, S. Chatterjee, A. Vishwanath, and M. P. Zaletel, Ground state and hidden symmetry of magic angle graphene at even integer filling, Phys. Rev. X **10**, 031034 (2020).

[48] J. Kang and O. Vafek, Non-Abelian Dirac node braiding and near-degeneracy of correlated phases at odd integer filling in magic angle twisted bilayer graphene, Phys. Rev. B **102**, 035161 (2020).

[49] T. Soejima, D. E. Parker, N. Bultinck, J. Hauschild, and M. P. Zaletel, Efficient simulation of moire materials using the density matrix renormalization group, Phys. Rev. B **102**, 205111 (2020).

[50] J. Liu and X. Dai, Theories for the correlated insulating states and quantum anomalous Hall effect phenomena in twisted bilayer graphene, Phys. Rev. B **103**, 035427 (2021).

[51] Y. H. Kwan, G. Wagner, T. Soejima, M. P. Zaletel, S. H. Simon, S. A. Parameswaran, and Nick Bultinck, Kekulé spiral order at all nonzero integer fillings in twisted bilayer graphene, Phys. Rev. X **11**, 041063 (2021).

[52] B. A. Bernevig, Z. D. Song, N. Regnault, and B. Lian, Twisted bilayer graphene. I. Matrix elements, approximations, perturbation theory, and a two-band model, Phys. Rev. B **103**, 205411 (2021); Z. D. Song, B. Lian, N. Regnault, and B. A. Bernevig, Twisted bilayer graphene. II. Stable symmetry anomaly, *ibid.* **103**, 205412 (2021); B. A. Bernevig, Z. D. Song, N. Regnault, and B. Lian, Twisted bilayer graphene. III. Interacting Hamiltonian and exact symmetries, *ibid.* **103**, 205413 (2021); B. Lian, Z. D. Song, N. Regnault, D. K. Efetov, A. Yazdani, and B. A. Bernevig, Twisted bilayer graphene. IV. Exact insulator ground states and phase diagram, *ibid.* **103**, 205414 (2021); B. A. Bernevig, B. Lian, A. Cowsik, F. Xie, N. Regnault, and Z. D. Song, Twisted bilayer graphene. V. Exact analytic many-body excitations in Coulomb Hamiltonians: Charge gap, Goldstone modes, and absence of Cooper pairing, *ibid.* **103**, 205415 (2021); F. Xie, A. Cowsik, Z. D. Song, B. Lian, B. A. Bernevig, and N. Regnault, Twisted bilayer graphene. VI. An exact diagonalization study at nonzero integer filling, *ibid.* **103**, 205416 (2021).

[53] T. Benschop, T. A. de Jong, P. Stepanov, Xiaobo Lu, Vincent Stalman, S. Jan van der Molen, D. K. Efetov, and M. P. Allan, Measuring local moiré lattice heterogeneity of twisted bilayer graphene, Phys. Rev. Res. **3**, 013153 (2021).

[54] N. P. Kazmierczak, M. Van Winkle, C. Ophus, K. C. Bustillo, S. Carr, H. G. Brown, J. Ciston, T. Taniguchi, K. Watanabe, and D. K. Bediako, Strain fields in twisted bilayer graphene, Nat. Mater. **20**, 956 (2021).

[55] J. Kang and O. Vafek, Pseudomagnetic fields, particle-hole asymmetry, and microscopic effective continuum Hamiltonians of twisted bilayer graphene, Phys. Rev. B **107**, 075408 (2023).

[56] T. Nakanishi and T. Ando, Conductance of crossed carbon nanotubes, J. Phys. Soc. Jpn. **70**, 1647 (2001).

[57] S. Uryu, Electronic states and quantum transport in double-wall carbon nanotubes, Phys. Rev. B **69**, 075402 (2004).

[58] J. C. Slater and G. F. Koster, Simplified LCAO method for the periodic potential problem, Phys. Rev. **94**, 1498 (1954).

[59] P. Moon and M. Koshino, Energy spectrum and quantum Hall effect in twisted bilayer graphene, Phys. Rev. B **85**, 195458 (2012).

[60] S. Fang and E. Kaxiras, Electronic structure theory of weakly interacting bilayers, Phys. Rev. B **93**, 235153 (2016).

[61] M. Xie, A. H. MacDonald, Weak-Field Hall Resistivity and Spin/Valley Flavor Symmetry Breaking in Magic-Angle Twisted Bilayer Graphene, Phys. Rev. Lett. **127**, 196401 (2021).

[62] M. Koshino and N. N. T. Nam, Continuum model for relaxed twisted bilayer graphenes and moiré electron-phonon interaction, Phys. Rev. B **101**, 195425 (2020).

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

[64] S. Fang, S. Carr, Z. Zhu, D. Massatt, and E. Kaxiras, Angle-dependent *ab initio* low-energy Hamiltonians for a relaxed twisted bilayer graphene heterostructure, arXiv:1908.00058.

[65] O. Vafek and J. Kang, Renormalization Group Study of Hidden Symmetry in Twisted Bilayer Graphene With Coulomb Interactions, Phys. Rev. Lett. **125**, 257602 (2020).

[66] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, 1995).

[67] Note that this differs from the choice made in Ref. [37] by a phase factor.