Simple derivation of moiré-scale continuous models for twisted bilayer graphene

Éric Cancès,[1,*] Louis Garrigue,[1,†] and David Gontier[2,3,‡]
[1]*CERMICS, École des Ponts and Inria Paris, 6 and 8 av. Pascal, Marne-la-Vallée 77455, France*
[2]*CEREMADE, University of Paris-Dauphine, PSL University, Paris 75016, France*
[3]*ENS/PSL University, Département de Mathématiques et Applications, Paris F-75005, France*

(Received 12 July 2022; revised 22 December 2022; accepted 1 March 2023; published 5 April 2023)

We provide a formal derivation of a reduced model for twisted bilayer graphene (TBG) from Density Functional Theory. Our derivation is based on a variational approximation of the TBG Kohn-Sham Hamiltonian and asymptotic limit techniques. In contrast with other approaches, it does not require the introduction of an intermediate tight-binding model. The so-obtained model is similar to that of the Bistritzer-MacDonald (BM) model but contains additional terms. Its parameters can be easily computed from Kohn-Sham calculations on single-layer graphene and untwisted bilayer graphene with different stackings. It allows one in particular to estimate the parameters $w_{\rm AA}$ and $w_{\rm AB}$ of the BM model from first principles. The resulting numerical values, namely $w_{\rm AA} = w_{\rm AB} \simeq 126\,\text{meV}$ for the experimental interlayer mean distance are in good agreement with the empirical values $w_{\rm AA} = w_{\rm AB} = 110\,\text{meV}$ obtained by fitting to experimental data. We also show that if the BM parameters are set to $w_{\rm AA} = w_{\rm AB} \simeq 126\,\text{meV}$, the BM model is an accurate approximation of our reduced model.

DOI: 10.1103/PhysRevB.107.155403

I. INTRODUCTION

Moiré materials [1,2] have attracted a lot of interest in condensed matter physics since, notably, the experimental discovery of Mott insulating and nonconventional superconducting phases [3] in twisted bilayer graphene (TBG) for specific small twist angles θ. In particular, the experiments reported in Ref. [3] were done with $\theta \simeq 1.1^\circ$. For such small twist angles, the moiré pattern is quite large and typically contains on the order of 11 000 carbon atoms. In addition, the system is aperiodic (incommensurate), except for a countable set of values of θ. All this makes brute force first-principle calculations extremely challenging.

Most theoretical investigations on TBG rely on continuous models [4–7] such as the Bistritzer-MacDonald (BM) model [6], an effective continuous periodic model describing low-energy excitations in TBG close to half filling. The BM Hamiltonian is a self-adjoint operator on $L^2(\mathbb{R}^2;\mathbb{C}^4)$ given by

$$H_\theta^{\rm BM} = \begin{pmatrix} v_{\rm F}\boldsymbol{\sigma}_{-\theta/2}\cdot(-i\nabla) & \boldsymbol{V}(k_\theta \mathbf{x}) \\ \boldsymbol{V}(k_\theta \mathbf{x})^* & v_{\rm F}\boldsymbol{\sigma}_{\theta/2}\cdot(-i\nabla) \end{pmatrix}, \qquad (1)$$

where $v_{\rm F}$ is the Fermi velocity in single-layer graphene, $\boldsymbol{\sigma}_{\pm\theta/2} = e^{\mp i\frac{\theta}{4}\sigma_3}(\sigma_1,\sigma_2)e^{\pm i\frac{\theta}{4}\sigma_3}$ are rotated Pauli matrices, and $k_\theta = \frac{8\pi}{3a_0}\sin(\theta/2)$, with a_0 the single-layer graphene lattice constant. The function $\boldsymbol{V}:\mathbb{R}^2\to\mathbb{C}^2$ is quasiperiodic at the so-called moiré scale (see Sec. III D for details) and depends on two empirical parameters $w_{\rm AA}$ and $w_{\rm AB}$ describing interlayer coupling in AA and AB stacking respectively. A rigorous mathematical derivation of the BM model from a tight-binding Hamiltonian whose parameters satisfy suitable scaling laws, was recently proposed in Ref. [8]. A simplified chiral BM model, obtained by setting $w_{\rm AA} = 0$, was notably used in Refs. [9–11] to prove the existence of perfectly flat bands at the Fermi level for a sequence of so-called "magic" angles. From a physical point of view, the existence of partially occupied almost flat bands in the single-particle picture may reflect the presence of localized strongly correlated electrons, and provide a possible explanation of the experimentally observed superconducting behavior [12]. An alternative approach to using effective models at the moiré scale is to develop atomic-scale models and efficient computational methods adapted to incommensurate two-dimensional (2D) heterostructures [13–17].

This article is concerned with the derivation of BM-like effective models directly from Density Functional Theory (DFT). In contrast with other approaches [6,8,18–22], our derivation does not involve an intermediate tight-binding model and is based on real-space (not momentum-space) computations.

II. APPROXIMATION OF THE KOHN-SHAM HAMILTONIAN FOR TBG

A. Single-layer graphene

We denote the position variable by $\mathbf{r} = (\mathbf{x}, z) \in \mathbb{R}^3$ where $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ and $z \in \mathbb{R}$ are, respectively, the longitudinal (in-plane) and transverse (out-of-plane) position variables. Let V be the Kohn-Sham potential for a single-layer graphene in the horizontal plane $z = 0$. The space group of graphene is $\mathrm{Dg80} = \mathrm{D}_{6h} \ltimes \mathbb{L}$ (p6/mmm), so V has the honeycomb symmetry and is $\mathbb{L}$ periodic, where $\mathbb{L} = \mathbb{Z}\mathbf{a}_1 + \mathbb{Z}\mathbf{a}_2$ and

$$\mathbf{a}_1 = a_0\begin{pmatrix} 1/2 \\ -\sqrt{3}/2 \end{pmatrix} \quad \text{and} \quad \mathbf{a}_2 = a_0\begin{pmatrix} 1/2 \\ \sqrt{3}/2 \end{pmatrix}.$$

*eric.cances@enpc.fr
†louis.garrigue@enpc.fr
‡gontier@ceremade.dauphine.fr

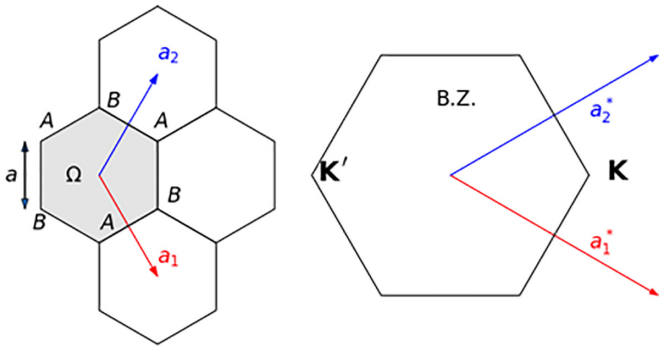

FIG. 1. Single-layer graphene. Left: atomic positions of the carbon atoms (A and B sublattices) in the physical space, and lattice vectors $\mathbf{a}_1$ and $\mathbf{a}_2$. Right: reciprocal lattice vectors $\mathbf{a}_1^*$ and $\mathbf{a}_2^*$, and first Brillouin zone in momentum space; the positions of the $\mathbf{K}$ and $\mathbf{K}'$ Dirac points are also indicated.

Here, $a_0 = \sqrt{3}a$ is the graphene lattice constant ($a \simeq 0.142$ nm $\simeq 2.68$ bohr is the carbon-carbon nearest-neighbor distance). We set $\Omega := \mathbb{R}^2/\mathbb{L}$ the Wigner-Seitz cell (see Fig. 1).

The single-layer graphene Kohn-Sham Hamiltonian reads

$$H^{(1)} := -\tfrac{1}{2}\Delta + V. \tag{2}$$

We denote by $H^{(1)}_{\mathbf{k}}$ its Bloch fibers. Recall that these operators have domains representing the $\mathbf{k}$–quasiperiodic boundary conditions $\Phi(\mathbf{x}-\mathbf{R}, z) = e^{i\mathbf{k}\cdot\mathbf{R}}\Phi(\mathbf{x}, z)$ for all $\mathbf{R} \in \mathbb{L}$. The map $\mathbf{k} \mapsto H^{(1)}_{\mathbf{k}}$ is $\mathbb{L}^*$ periodic, where $\mathbb{L}^*$ is the dual lattice of $\mathbb{L}$. Explicitly, $\mathbb{L}^* = \mathbf{a}_1^*\mathbb{Z} + \mathbf{a}_2^*\mathbb{Z}$ with

$$\mathbf{a}_1^* = \sqrt{3}k_D\begin{pmatrix}\sqrt{3}/2\\ -1/2\end{pmatrix} \quad \text{and} \quad \mathbf{a}_2^* = \sqrt{3}k_D\begin{pmatrix}\sqrt{3}/2\\ 1/2\end{pmatrix},$$

where $k_D := \frac{4\pi}{3\,a_0}$. At the special Dirac point

$$\mathbf{K} := \tfrac{1}{3}(\mathbf{a}_1^* + \mathbf{a}_2^*) = k_D(1,0)^T, \tag{3}$$

$H^{(1)}_{\mathbf{K}}$ has an eigenvalue of multiplicity 2 at the Fermi level μ_F. We denote by Φ_1 and Φ_2 two corresponding eigenvectors, oriented so that $\langle\Phi_1, (-i\nabla_{\mathbf{x}})\Phi_1\rangle = \langle\Phi_2, (-i\nabla_{\mathbf{x}})\Phi_2\rangle = (0,0)^T$, and

$$\langle\Phi_1, (-i\nabla_{\mathbf{x}})\Phi_2\rangle = v_F\begin{pmatrix}1\\ -i\end{pmatrix}, \tag{4}$$

with $v_F > 0$ the Fermi velocity. We refer to the Supplemental Material [23] for an explanation of how to achieve such an orientation. Finally, we denote by u_1 and u_2 the periodic parts of Φ_1 and Φ_2, *i.e.*, $u_j(\mathbf{r}) := \Phi_j(\mathbf{r})e^{-i\mathbf{K}\cdot\mathbf{x}}$.

B. Kohn-Sham model for TBG

We now consider two parallel layers of graphene, separated by a distance $d > 0$, and with a twist angle θ between the two. More precisely, we first place the top layer in the plane $z = \frac{d}{2}$ and the bottom layer in the plane $z = -\frac{d}{2}$ in AA stacking, and then rotate counterclockwise the top layer by $-\theta/2$ and the bottom layer by $\theta/2$ around the z axis, placing the origin at the center of a carbon hexagon. We set

$$c_\theta := \cos\frac{\theta}{2} \quad \text{and} \quad \varepsilon_\theta := 2\sin\frac{\theta}{2}.$$

Note that $\varepsilon_\theta \sim \theta$ in the small-angle limit. We denote by R_θ the 2D rotation matrix of angle $\theta \in \mathbb{R}$, specifically

$$R_\theta := c_\theta\,\mathbb{I}_2 - \varepsilon_\theta J, \quad \text{with} \quad J := \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},$$

and we introduce the unitary operator

$$(U_{d,\theta}f)(\mathbf{x}, z) := f(R^*_{-\theta/2}\mathbf{x}, z) = f\left(c_\theta\mathbf{x} - \frac{1}{2}\varepsilon_\theta J\mathbf{x}, z - \frac{d}{2}\right).$$

The inverse of $U_{d,\theta}$ is $U^{-1}_{d,\theta} = U^*_{d,\theta} = U_{-d,-\theta}$.

For twist angles θ giving rise to a periodic structure at the moiré scale, the TBG Kohn-Sham potential is a well-defined moiré-periodic function. It is unclear how to define a mean-field potential for incommensurate twist angles. This problem actually occurs for all infinite aperiodic systems (see Ref. [24] for a mathematical definition of a mean-field model in an ergodic setting). Here, we assume that this potential exists, and can be approximated using the procedure in Ref. [25]. We consider an approximate Kohn-Sham potential for the TBG of the form

$$V^{(2)}_{d,\theta}(\mathbf{x}, z) := (U_{d,\theta}V)(\mathbf{x}, z) + \left(U^{-1}_{d,\theta}V\right)(\mathbf{x}, z) + V_{\mathrm{int},d}(z).$$

Each component $U^{\pm1}_{d,\theta}V$ represents a layer of graphene shifted by $\pm\frac{d}{2}\mathbf{e}_z$, and twisted by an angle $\mp\theta/2$. The last term $V_{\mathrm{int},d}$ is a correction which takes into account the relaxation of the Kohn-Sham potential due to interlayer coupling. This term is constructed as follows. For each disregistry vector $\mathbf{y} \in \Omega$, we denote by $V^{(2)}_{d,\mathbf{y}}$ the mean-field Kohn-Sham potential for the configuration where the two layers are aligned (no twist angle) with the top one shifted by $\mathbf{y}$ in the longitudinal direction. The potential $V_{\mathrm{int},d}(z)$ is defined as the average

$$V_{\mathrm{int},d}(z) = \fint_\Omega V_{\mathrm{int},d,\mathbf{y}}(z)\mathrm{d}\mathbf{y}$$

where $\fint_\Omega := |\Omega|^{-1}\int_\Omega$, and with $V_{\mathrm{int},d,\mathbf{y}}(z)$ defined by

$$\fint_\Omega \left(V^{(2)}_{d,\mathbf{y}}(\mathbf{x}, z) - V\left(\mathbf{x}, z + \frac{d}{2}\right) - V\left(\mathbf{x}-\mathbf{y}, z - \frac{d}{2}\right)\right)\mathrm{d}\mathbf{x}.$$

In other words, $V_{\mathrm{int,d}}$ is the mismatch between the interacting Kohn-Sham potential and the noninteracting one, averaged over all possible disregistries $\mathbf{y} \in \Omega$. Note that $V_{\mathrm{int,d}}(z)$ only depends on the z variable and satisfies $V_{\mathrm{int},d}(-z) = V_{\mathrm{int},d}(z)$.

In what follows, we study the approximate TBG Kohn-Sham Hamiltonian

$$H^{(2)}_{d,\theta} := -\tfrac{1}{2}\Delta + V^{(2)}_{d,\theta}(\mathbf{x}, z). \tag{5}$$

Our goal is to derive a 2D reduced model describing the low-energy band structure around the Fermi level μ_{F} in the limit of small twist angles.

III. REDUCED MODEL

The potential $V^{(2)}_{d,\theta}$ is of the form $V^{(2)}_{d,\theta}(\mathbf{x}, z) = v_d(c_\theta\mathbf{x}, \varepsilon_\theta\mathbf{x}, z)$, with

$$v_d(\mathbf{x}, \mathbf{X}, z) := V\left(\mathbf{x} - \frac{1}{2}J\mathbf{X}, z - \frac{d}{2}\right) + V\left(\mathbf{x} + \frac{1}{2}J\mathbf{X}, z + \frac{d}{2}\right) + V_{\mathrm{int},d}(z).$$

The potential v_d is $\mathbb{L}$ periodic in the first variable $\mathbf{x}$, and $2J\mathbb{L}$ periodic in its second variable $\mathbf{X}$. In the limit $\theta \to 0$, $v_d(c_\theta \mathbf{x}, \varepsilon_\theta \mathbf{x}, z)$ has a natural two-scale structure, so that $H^{(2)}_{d,\theta}$ could be studied using adiabatic theory, semiclassical analysis, and/or homogenization theory. In this article however, we take a different route, and present a simple approximation scheme to derive an effective Hamiltonian describing electronic transport around the Fermi level.

A. Variational approximation of low-energy wavepacket dynamics

The main idea of our approach is to project the time-dependent Schrödinger equation

$$i\partial_t \Psi(t,\mathbf{r}) = \big(H^{(2)}_{d,\theta} - \mu_{\mathrm{F}}\big)\Psi(t,\mathbf{r}), \quad \Psi(0,\mathbf{r}) = \Psi_0(\mathbf{r}) \tag{6}$$

onto the two-scale approximation space

$$\mathcal{X}_{d,\theta} := \{(\alpha : \Phi)_{d,\theta},\ \alpha : \mathbb{R}^2 \to \mathbb{C}^4\}, \tag{7}$$

where we set

$$(\alpha : \Phi)_{d,\theta}(\mathbf{x}, z) := \sum_{\substack{\eta\in\{\pm 1\}\\ j\in\{1,2\}}} \alpha_{\eta,j}(\varepsilon_\theta \mathbf{x})\big(U^{\eta}_{d,\theta}\Phi_j\big)(\mathbf{x}, z).$$

The trial functions in $\mathcal{X}_{d,\theta}$ are linear combinations of four wavepackets, each one consisting of an envelope function α oscillating at the moiré scale ε_θ^{-1} multiplied by one of the two (translated and rotated) Bloch waves Φ_1 or Φ_2 of one of the two layers. Note that both the TBG approximation subspace $\mathcal{X}_{d,\theta}$ and Hamiltonian $H^{(2)}_{d,\theta}$ depend on the small parameter θ.

Given an initial state of the form $\Psi_0 = (\alpha^0 : \Phi)_{d,\theta}$, the true solution $\Psi(t)$ of Eq. (6) is expected to be close to $(\alpha(\varepsilon_\theta t) : \Phi)_{d,\theta}$ up to times of order ε_θ^{-1}, where $\alpha(t)$ satisfies $\alpha(t=0) = \alpha^0$, and solves the projected equation

$$\begin{aligned} i\partial_t \langle (\widetilde{\alpha} : \Phi)_{d,\theta}, (\alpha(\varepsilon_\theta t) : \Phi)_{d,\theta}\rangle \\ = \big\langle (\widetilde{\alpha} : \Phi)_{d,\theta}, \big(H^{(2)}_{d,\theta} - \mu_F\big)(\alpha(\varepsilon_\theta t) : \Phi)_{d,\theta}\big\rangle \end{aligned} \tag{8}$$

for all $\widetilde{\alpha} : \mathbb{R}^2 \to \mathbb{C}^4$. The time variable has to be rescaled as $\tau := \varepsilon_\theta t$ in order to obtain wavepacket propagation with finite velocity at the moiré scale.

B. Formulation of the reduced model

It follows from tedious calculations detailed in Appendix A that if we let θ go to zero in Eq. (8) for fixed trial smooth functions $\widetilde{\alpha}$ and α, we obtain the asymptotic equality

$$\mathrm{i}\partial_\tau \langle \widetilde{\alpha}, \mathcal{S}_d \alpha(\tau)\rangle = \langle \widetilde{\alpha}, \mathcal{H}_{d,\theta}\alpha(\tau)\rangle + O\big(\varepsilon_\theta^{\infty}\big), \tag{9}$$

where the overlap operator $\mathcal{S}_d$ and the Hamiltonian operator $\mathcal{H}_{d,\theta}$ act on $L^2(\mathbb{R}^2;\mathbb{C}^4)$, and are defined by

$$\mathcal{S}_d := \begin{pmatrix} \mathbb{I}_2 & \Sigma_d(\mathbf{X}) \\ \Sigma_d^*(\mathbf{X}) & \mathbb{I}_2 \end{pmatrix}$$

and

$$\begin{aligned} \mathcal{H}_{d,\theta} := &\begin{pmatrix} v_{\mathrm{F}}\boldsymbol{\sigma}_{-\theta/2}\cdot(-\mathrm{i}\nabla_{\mathbf{X}}) & \varepsilon_\theta^{-1}\mathbb{V}_d(\mathbf{X}) \\ \varepsilon_\theta^{-1}\mathbb{V}_d(\mathbf{X})^* & v_{\mathrm{F}}\boldsymbol{\sigma}_{\theta/2}\cdot(-\mathrm{i}\nabla_{\mathbf{X}}) \end{pmatrix} + \begin{pmatrix} \varepsilon_\theta^{-1}\mathbb{W}_d^+ & 0 \\ 0 & \varepsilon_\theta^{-1}\mathbb{W}_d^- \end{pmatrix} \\ &+ \begin{pmatrix} 0 & c_\theta J(-\mathrm{i}\nabla\Sigma_d(\mathbf{X}))\cdot(-\mathrm{i}\nabla) \\ c_\theta J(-\mathrm{i}\nabla\Sigma_d^*)(\mathbf{X})\cdot(-\mathrm{i}\nabla) & 0 \end{pmatrix} - \frac{\varepsilon_\theta}{2}\nabla\cdot\left(\begin{pmatrix} \mathbb{I}_2 & \Sigma_d(\mathbf{X}) \\ \Sigma_d^*(\mathbf{X}) & \mathbb{I}_2 \end{pmatrix}\nabla\bullet\right). \end{aligned} \tag{10}$$

The 2×2 matrix-valued functions $\Sigma_d(\mathbf{X})$, $\mathbb{V}_d(\mathbf{X})$, and $\mathbb{W}_d(\mathbf{X})$ are given by

$$\begin{aligned} [\Sigma_d(\mathbf{X})]_{jj'} &:= \mathrm{e}^{-\mathrm{i}J\mathbf{K}\cdot\mathbf{X}}(\!(u_j, u_{j'})\!)_d^{+-}(\mathbf{X}), \\ [\mathbb{V}_d(\mathbf{X})]_{jj'} &:= \mathrm{e}^{-\mathrm{i}J\mathbf{K}\cdot\mathbf{X}}\left(\!\!\left(\left(V + V_{\mathrm{int},d}\left(\cdot + \frac{d}{2}\right)\right)u_j, u_{j'}\right)\!\!\right)_d^{+-}(\mathbf{X}), \\ [\mathbb{W}_d^{\pm}(\mathbf{X})]_{jj'} &:= (\!(u_j\overline{u_{j'}}, V)\!)_d^{\pm\mp}(\mathbf{X}) + \big(W^{\pm}_{\mathrm{int},d}\big)_{jj'}, \end{aligned}$$

where for all $\mathbb{L}$-periodic functions f and g

$$\begin{aligned} (\!(f,g)\!)_d^{\eta\eta'}(\mathbf{X}) := \int_{\Omega\times\mathbb{R}} &\overline{f\left(\mathbf{x} - \eta\frac{1}{2}J\mathbf{X}, z - \eta\frac{d}{2}\right)} \\ &\times g\left(\mathbf{x} - \eta'\frac{1}{2}J\mathbf{X}, z - \eta'\frac{d}{2}\right)\mathrm{d}\mathbf{x}\,\mathrm{d}z, \end{aligned} \tag{11}$$

and where

$$(W^{\pm}_{\mathrm{int},d})_{jj'} := \int_{\Omega\times\mathbb{R}} (\overline{u_j}u_{j'})\left(\mathbf{x}, z \mp \frac{d}{2}\right)V_{\mathrm{int},d}(z)\mathrm{d}\mathbf{x}\,\mathrm{d}z.$$

All these quantities can be computed from the single-layer potential V, the Bloch wavefunctions u_1 and u_2, and the Kohn-Sham correction potential $V_{\mathrm{int},d}$.

It is not obvious that the operator $\mathcal{H}_{d,\theta}$ is Hermitian. It is however the case. For instance, one can check that $\overline{(\!(u,v)\!)_d^{+-}}(\mathbf{X}) = (\!(\overline{u},\overline{v})\!)_d^{+-}(\mathbf{X})$, which proves that the matrices $\mathbb{W}_d^+(\mathbf{X})$ and $\mathbb{W}_d^-(\mathbf{X})$ are Hermitian. In addition, from the equality $\mathrm{div}J\nabla = -\partial^2_{x_1x_2} + \partial^2_{x_2x_1} = 0$, we see that the third matrix in the definition Eq. (10) of $\mathcal{H}_{d,\theta}$ also defines an Hermitian operator.

The asymptotic equality Eq. (9) leads us to introduce the following effective model for the propagation of low-energy wavepackets in TBG:

$$\mathrm{i}\partial_\tau(\mathcal{S}_d\alpha)(\tau,\mathbf{X}) = (\mathcal{H}_{d,\theta}\alpha)(\tau,\mathbf{X}). \tag{12}$$

At this stage, we have kept all the terms in $\mathcal{S}_d$ and $\mathcal{H}_{d,\theta}$, and only thrown away the remainders of order $\varepsilon_\theta^{\infty}$. Indeed, ε_θ should not be considered as the only small parameter in this problem. The interlayer coupling is also small, hence the operators $\mathbb{W}_d^{\pm}$ can be considered to be small as well. The

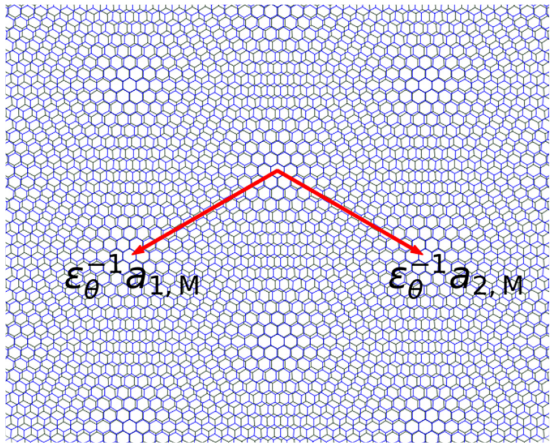

FIG. 2. Moiré lattice vectors for twist angle $\theta \simeq 2.2^\circ$.

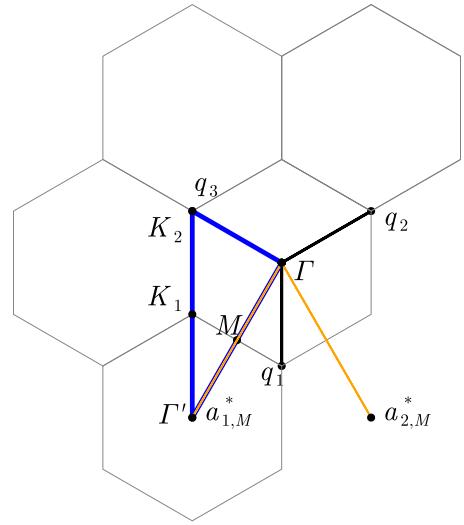

FIG. 3. Path $\mathbf{K}_2 \to \mathbf{K}_1 \to \Gamma' \to M \to \Gamma \to \mathbf{K}_2$ in momentum space, the hexagon centered at Γ corresponds to the mini Brillouin zone.

interplay between the two parameters ε_θ and w (the interlayer characteristic interaction energy) is subtle and will be explored in a future work.

Remark 1. A similar approach can be used to derive an effective model for wavepacket propagation in monolayer graphene in an energy range close to the Fermi level and localized in the K valley in momentum space. One obtains the effective Hamiltonian

$$\mathcal{H}_\varepsilon^{(1)} = v_F \boldsymbol{\sigma} \cdot (-\mathrm{i}\nabla) + \tfrac{1}{2}\varepsilon(-\mathrm{i}\nabla)^2$$

acting on $L^2(\mathbb{R}^2;\mathbb{C}^2)$. Neglecting the second term, we recover the usual massless Dirac operator. This effective model was rigorously derived in Ref. [26] using other methods.

C. Translational covariance

For any $\mathbb{L}$-periodic functions f and g, the maps $((f,g))_d^{\pm\mp}(\mathbf{X})$ defined in Eq. (11) are $J\mathbb{L}$ periodic. This suggests to introduce the rescaled moiré lattice

$$\mathbb{L}_\mathrm{M} := J\mathbb{L},$$

which satisfies $\mathbb{L}_\mathrm{M} = \mathbf{a}_{1,\mathrm{M}}\mathbb{Z} + \mathbf{a}_{2,\mathrm{M}}\mathbb{Z}$ with lattice vectors $\mathbf{a}_{1,\mathrm{M}} = J\mathbf{a}_1$ and $\mathbf{a}_{2,\mathrm{M}} = J\mathbf{a}_2$. Explicitly,

$$\mathbf{a}_{1,\mathrm{M}} = a_0\begin{pmatrix} -\sqrt{3}/2 \\ -1/2 \end{pmatrix} \quad \text{and} \quad \mathbf{a}_{2,\mathrm{M}} = a_0\begin{pmatrix} \sqrt{3}/2 \\ -1/2 \end{pmatrix}.$$

The corrresponding Wigner–Seitz cell is $\Omega_\mathrm{M} := \mathbb{R}^2/\mathbb{L}_\mathrm{M}$, and its dual basis is given by $\mathbf{a}^*_{1,\mathrm{M}} = J\mathbf{a}_1^*$ and $\mathbf{a}^*_{2,\mathrm{M}} = J\mathbf{a}_2^*$, that is

$$\mathbf{a}^*_{1,\mathrm{M}} = \sqrt{3}k_\mathrm{D}\begin{pmatrix} -1/2 \\ -\sqrt{3}/2 \end{pmatrix} \quad \text{and} \quad \mathbf{a}^*_{2,\mathrm{M}} = \sqrt{3}k_\mathrm{D}\begin{pmatrix} 1/2 \\ -\sqrt{3}/2 \end{pmatrix}.$$

We also introduce the vectors (see Fig. 3)

$$\mathbf{q}_1 := k_\mathrm{D}\begin{pmatrix} 0 \\ -1 \end{pmatrix} = \frac{1}{3}(\mathbf{a}^*_{1,\mathrm{M}} + \mathbf{a}^*_{2,\mathrm{M}}),$$

$$\mathbf{q}_2 := k_\mathrm{D}\begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix} = \frac{1}{3}(-2\mathbf{a}^*_{1,\mathrm{M}} + \mathbf{a}^*_{2,\mathrm{M}}) = R_{\frac{2\pi}{3}}\mathbf{q}_1,$$

$$\mathbf{q}_3 := k_\mathrm{D}\begin{pmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix} = \frac{1}{3}(\mathbf{a}^*_{1,\mathrm{M}} - 2\mathbf{a}^*_{2,\mathrm{M}}) = R_{\frac{2\pi}{3}}\mathbf{q}_2.$$

These vectors satisfy $\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 = \mathbf{0}$ and correspond to the $\mathbf{K}$ valley of the moiré Brillouin zone. Actually, we have $\mathbf{q}_1 = J\mathbf{K}$.

Going back to our reduced model, we see that the diagonal elements $\mathbb{W}_d^\pm$ are $\mathbb{L}_M$ periodic, while for $\mathbf{R}_M \in \mathbb{L}_M$,

$$\mathbb{V}_d(\mathbf{X} - \mathbf{R}_M) = \mathrm{e}^{\mathrm{i}\mathbf{q}_1\cdot\mathbf{R}_M}\mathbb{V}_d(\mathbf{X}), \tag{13}$$

and similarly for Σ_d. Writing $\mathbf{R}_M = m_1\mathbf{a}_{1,M} + m_2\mathbf{a}_{2,M}$, we have $\mathrm{e}^{\mathrm{i}\mathbf{q}_1\cdot\mathbf{R}_M} = \omega^{(m_1+m_2)}$, where we set $\omega := \mathrm{e}^{\mathrm{i}\frac{2\pi}{3}}$. Since $\omega^3 = 1$, our model is $3\mathbb{L}_\mathrm{M}$ periodic.

Thus, although the true moiré pattern generated by the superposition of two twisted honeycomb lattices is not periodic for a generic twist angle θ, our reduced model is. In some sense, the moiré pattern looks $\varepsilon_\theta^{-1}\mathbb{L}_\mathrm{M}$ periodic at the mesoscopic scale ε_θ^{-1} for $\theta \ll 1$ (see Fig. 2).

D. Comparison with the Bistritzer-MacDonald model

The BM Hamiltonian [Eq. (1)] can be written more explicitly [6,8,9] as

$$H_\theta^\mathrm{BM} = \begin{pmatrix} v_\mathrm{F}\boldsymbol{\sigma}_{-\theta/2}\cdot(-\mathrm{i}\nabla_\mathbf{x}) & \boldsymbol{V}(\varepsilon_\theta\mathbf{x}) \\ \boldsymbol{V}(\varepsilon_\theta\mathbf{x})^* & v_\mathrm{F}\boldsymbol{\sigma}_{\theta/2}\cdot(-\mathrm{i}\nabla_\mathbf{x}) \end{pmatrix},$$

with

$$\boldsymbol{V}(\mathbf{X}) := \begin{pmatrix} w_\mathrm{AA}\,G(\mathbf{X}) & w_\mathrm{AB}\,\overline{F(-\mathbf{X})} \\ w_\mathrm{AB}\,F(\mathbf{X}) & w_\mathrm{AA}\,G(\mathbf{X}) \end{pmatrix}, \tag{14}$$

where w_AA and w_AB are the two real parameters describing the interlayer coupling in AA and AB stacking, and

$$F(\mathbf{X}) := \mathrm{e}^{-\mathrm{i}\mathbf{q}_1\cdot\mathbf{X}} + \omega\mathrm{e}^{-\mathrm{i}\mathbf{q}_2\cdot\mathbf{X}} + \omega^2\mathrm{e}^{-\mathrm{i}\mathbf{q}_3\cdot\mathbf{X}}, \tag{15}$$

$$G(\mathbf{X}) := \mathrm{e}^{-\mathrm{i}\mathbf{q}_1\cdot\mathbf{X}} + \mathrm{e}^{-\mathrm{i}\mathbf{q}_2\cdot\mathbf{X}} + \mathrm{e}^{-\mathrm{i}\mathbf{q}_3\cdot\mathbf{X}}. \tag{16}$$

The BM potential satisfies, for all $\mathbf{R}_M \in \mathbb{L}_M$,

$$\boldsymbol{V}(\mathbf{X} - \mathbf{R}_M) = \mathrm{e}^{\mathrm{i}\mathbf{q}_1\cdot\mathbf{R}_M}\boldsymbol{V}(\mathbf{X}) = \omega^{(m_1+m_2)}\boldsymbol{V}(\mathbf{X}), \tag{17}$$

so that the BM potential $\boldsymbol{V}$ and our reduced potential $\mathbb{V}$ have the same covariance symmetries. They actually share many

other symmetries (see Sec. II in the Supplemental Materials [23] for a comprehensive analysis of the symmetries of our reduced model).

Rescaling lengths and energies as $\mathbf{X} = \varepsilon_\theta \mathbf{x}$ and $\mathcal{E} = \varepsilon_\theta^{-1} E$ (in the BM model, the Fermi level is set to zero), we obtain the rescaled BM Hamiltonian

$$\mathcal{H}_\theta^{\mathrm{BM}} = \begin{pmatrix} v_{\mathrm{F}} \boldsymbol{\sigma}_{-\theta/2} \cdot (-\mathrm{i}\nabla_{\mathbf{X}}) & \varepsilon_\theta^{-1} \boldsymbol{V}(\mathbf{X}) \\ \varepsilon_\theta^{-1} \boldsymbol{V}(\mathbf{X})^* & v_{\mathrm{F}} \boldsymbol{\sigma}_{\theta/2} \cdot (-\mathrm{i}\nabla_{\mathbf{X}}) \end{pmatrix}. \tag{18}$$

Going back to our model in Eq. (10), we see that the two models are similar under the following assumptions (which will be justified numerically in Sec. IV): (1) the matrix $\Sigma_d(\mathbf{X})$ and its gradient can be neglected; (2) the term $-\frac{1}{2}\varepsilon_\theta \Delta$ can be neglected, which is the case if the oscillations of the envelope functions $\alpha_{\eta,j}$ at the moiré scale contribute more than those at the atomic scale; (3) the functions $\mathbb{W}_d^{\pm}$ are almost proportional to the identity matrix (and thus only induce a global energy shift); (4) the function $\mathbb{V}_d$ is close to Eq. (14) for some well-chosen parameters w_{AA} and w_{AB}.

As for the last point, first-principle values of the BM parameters can be inferred from our reduced model by setting

$$w_{\mathrm{AA}}^d := \frac{1}{3|\Omega_{\mathrm{M}}|} \int_{\Omega_{\mathrm{M}}} [\mathbb{V}_d]_{11}(\mathbf{X}) \overline{G(\mathbf{X})}\, \mathrm{d}\mathbf{X}, \tag{19}$$

$$w_{\mathrm{AB}}^d := \frac{1}{3|\Omega_{\mathrm{M}}|} \int_{\Omega_{\mathrm{M}}} [\mathbb{V}_d]_{21}(\mathbf{X}) \overline{F(\mathbf{X})}\, \mathrm{d}\mathbf{X}, \tag{20}$$

where we used the fact that $\int_{\Omega_{\mathrm{M}}} |F|^2 = \int_{\Omega_{\mathrm{M}}} |G|^2 = 3|\Omega_{\mathrm{M}}|$. Note that from Eqs. (13) and (17), the integrands are $\mathbb{L}_{\mathrm{M}}$-periodic functions, and that $|\Omega_{\mathrm{M}}| = |\Omega| = \frac{\sqrt{3}}{2} a_0^2$. We prove in Appendix B that $w_{AA}^d = w_{AB}^d$.

IV. NUMERICAL RESULTS

In this section, we numerically study the generalized spectral problem associated with the operators $(\mathcal{H}_{d,\theta}, \mathcal{S}_d)$. Due to the energy shift in Eq. (6) and the rescaling by ε_θ, the spectrum of the operator $H_{d,\theta}^{(2)}$ close to μ_{F} is related to the spectrum of $(\mathcal{H}_{d,\theta}, \mathcal{S}_d)$ around 0 by

$$\sigma\left(H_{d,\theta}^{(2)}\right) \simeq \mu_{\mathrm{F}} + \varepsilon_\theta \sigma(\mathcal{H}_{d,\theta}, \mathcal{S}_d) = \mu_{\mathrm{F}} + \varepsilon_\theta \sigma(\widetilde{\mathcal{H}}_{d,\theta}, \widetilde{\mathcal{S}}_d),$$

where $\widetilde{\mathcal{H}}_{d,\theta}$ and $\widetilde{\mathcal{S}}_d$ are $\mathbb{L}_{\mathrm{M}}$-periodic operators obtained from $\mathcal{H}_{d,\theta}$ and $\mathcal{S}_d$ by the gauge transformation specified in the next section.

A. Gauge transformation

First, we perform a gauge transformation in order to remove the phase factors in Eq. (13) and end up with an $\mathbb{L}_{\mathrm{M}}$-periodic model. The same arguments can be used for the BM model. Let $\mathbf{K}_1$ and $\mathbf{K}_2$ be two vectors such that $\mathbf{K}_1 - \mathbf{K}_2 = \mathbf{q}_1$, *e.g.*, $\mathbf{K}_2 = \mathbf{q}_3$, and $\mathbf{K}_1 = -\mathbf{q}_2$ (recall that $\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 = \mathbf{0}$). We introduce the unitary multiplication operator

$$P(\mathbf{X}) := \begin{pmatrix} \mathrm{e}^{\mathrm{i}\mathbf{K}_1 \cdot \mathbf{X}} \mathbb{I}_2 & \mathbf{0} \\ \mathbf{0} & \mathrm{e}^{\mathrm{i}\mathbf{K}_2 \cdot \mathbf{X}} \mathbb{I}_2 \end{pmatrix},$$

with inverse $P^{-1}(\mathbf{X}) = P^*(\mathbf{X}) = P(-\mathbf{X})$. First, we have

$$\widetilde{\mathcal{S}}_d := P \mathcal{S}_d P^* = \begin{pmatrix} \mathbb{I}_2 & \widetilde{\Sigma_d}(\mathbf{X}) \\ \widetilde{\Sigma_d}^*(\mathbf{X}) & \mathbb{I}_2 \end{pmatrix},$$

with $\widetilde{\Sigma_d}(\mathbf{X}) = \mathrm{e}^{\mathrm{i}(\mathbf{K}_1 - \mathbf{K}_2)\cdot \mathbf{X}} \Sigma_d(\mathbf{X})$. Using the definition of Σ_d and the fact that $\mathbf{q}_1 = J\mathbf{K}$, we obtain

$$\widetilde{\Sigma_d}(\mathbf{X}) = \mathrm{e}^{\mathrm{i}\mathbf{q}_1 \cdot \mathbf{X}} \Sigma_d(\mathbf{X}) = (\!(u_j, u_{j'})\!)_d^{+-}(\mathbf{X}).$$

The $\widetilde{\mathcal{S}}_d$ matrix-valued function is now $\mathbb{L}_M$ periodic. Similarly, we find, using the notation $\nabla_{\mathbf{k}} := \nabla - \mathrm{i}\mathbf{k}$,

$$\widetilde{\mathcal{H}}_{d,\theta} = \begin{pmatrix} v_{\mathrm{F}} \boldsymbol{\sigma}_{-\theta/2} \cdot (-\mathrm{i}\nabla_{\mathbf{K}_1}) & \varepsilon_\theta^{-1} \widetilde{\mathbb{V}}_d(\mathbf{X}) \\ \varepsilon_\theta^{-1} \widetilde{\mathbb{V}}_d(\mathbf{X})^* & v_{\mathrm{F}} \boldsymbol{\sigma}_{\theta/2} \cdot (-\mathrm{i}\nabla_{\mathbf{K}_2}) \end{pmatrix} + \begin{pmatrix} \varepsilon_\theta^{-1} \widetilde{\mathbb{W}}_d^+ & 0 \\ 0 & \varepsilon_\theta^{-1} \widetilde{\mathbb{W}}_d^- \end{pmatrix}$$
$$+ \begin{pmatrix} 0 & c_\theta J \widetilde{A}_d(\mathbf{X}) \cdot (-\mathrm{i}\nabla_{\mathbf{K}_2}) \\ c_\theta J \widetilde{A}_d^*(\mathbf{X}) \cdot (-\mathrm{i}\nabla_{\mathbf{K}_1}) & 0 \end{pmatrix} + \frac{\varepsilon_\theta}{2} \begin{pmatrix} (-\mathrm{i}\nabla_{\mathbf{K}_1})^2 & (-\mathrm{i}\nabla_{\mathbf{K}_1}) \cdot [\widetilde{\Sigma_d}(-\mathrm{i}\nabla_{\mathbf{K}_2})\bullet] \\ (-\mathrm{i}\nabla_{\mathbf{K}_2}) \cdot [\widetilde{\Sigma_d}^*(-\mathrm{i}\nabla_{\mathbf{K}_1})\bullet] & (-\mathrm{i}\nabla_{\mathbf{K}_2})^2 \end{pmatrix}, \tag{21}$$

with components given by

$$[\widetilde{\Sigma_d}(\mathbf{X})]_{jj'} := (\!(u_j, u_{j'})\!)_d^{+-}(\mathbf{X}),$$

$$[\widetilde{\mathbb{V}}_d(\mathbf{X})]_{jj'} := \left(\left(\left(V + V_{\mathrm{int},d}\left(\cdot + \frac{d}{2} \right) \right) u_j, u_{j'} \right) \right)_d^{+-}(\mathbf{X}),$$

$$[\widetilde{\mathbb{W}}_d^{\pm}(\mathbf{X})]_{jj'} := (\!(u_j \overline{u_{j'}}, V)\!)_d^{\pm\mp}(\mathbf{X}) + (W_{\mathrm{int},d}^{\pm})_{jj'},$$

$$\widetilde{A}_d := e^{i\mathbf{q}_1 \mathbf{x}}(-i\nabla \Sigma_d) = (-i\nabla_{\mathbf{q}_1})\widetilde{\Sigma}_d.$$

In this gauge, the model is $\mathbb{L}_{\mathrm{M}}$ periodic, and we can apply the usual Bloch transform to compute its band diagram. For the sake of illustration, we display on Figs. 4 and 5 the band diagrams of the BM model (black) and of our continuous model (red). The path in momentum space used to produce the bands diagrams is displayed in Fig. 3.

Quantities are computed with $d = 6.45$ bohr for our effective model. There are at least two ways of characterizing special angles associated to almost flat bands: the standard one is to consider the local minimizers of the Fermi velocity (called magic angles); an alternative consists of considering the local minimizers of the almost flat bands bandwidth. We choose here the second way. The first minimizing angle is $\theta \simeq 1.175^\circ$ for the BM model (black lines) with $w_{\mathrm{AA}} = w_{\mathrm{AB}} = 110\,\mathrm{meV}$, and $\theta \simeq 1.164^\circ$ for our effective model (red lines). A noticeable difference between our effective model and the BM model is that for both definitions of the special

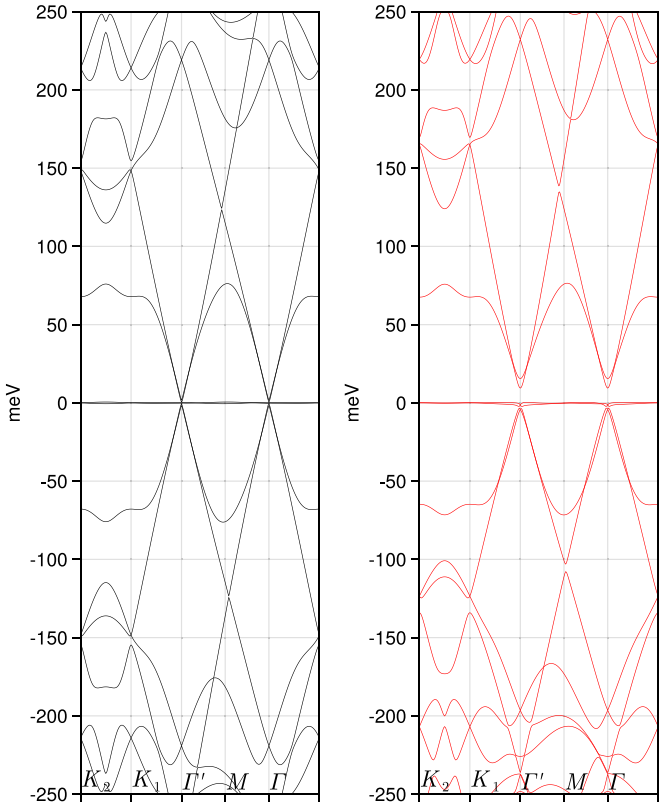

FIG. 4. Left: band diagram of $\mathcal{H}_\theta^{\mathrm{BM}}$ (black) with $w_{\mathrm{AA}} = w_{\mathrm{AB}} = 110\,\mathrm{meV}$ for the first bandwidth minimizing angle $\theta \simeq 1.175^\circ$ (bandwidth 1.1 meV). Right: band diagram of $(\widetilde{\mathcal{H}}_{d,\theta}, \widetilde{\mathcal{S}}_d)$ (red) for the first minimizing bandwidth angle $\theta \simeq 1.164^\circ$ (bandwidth 2.3 meV). In both cases, the zero of the energy scale is the center of the almost flat band which is obtained by shifting the middle band diagram (red) of 957 meV.

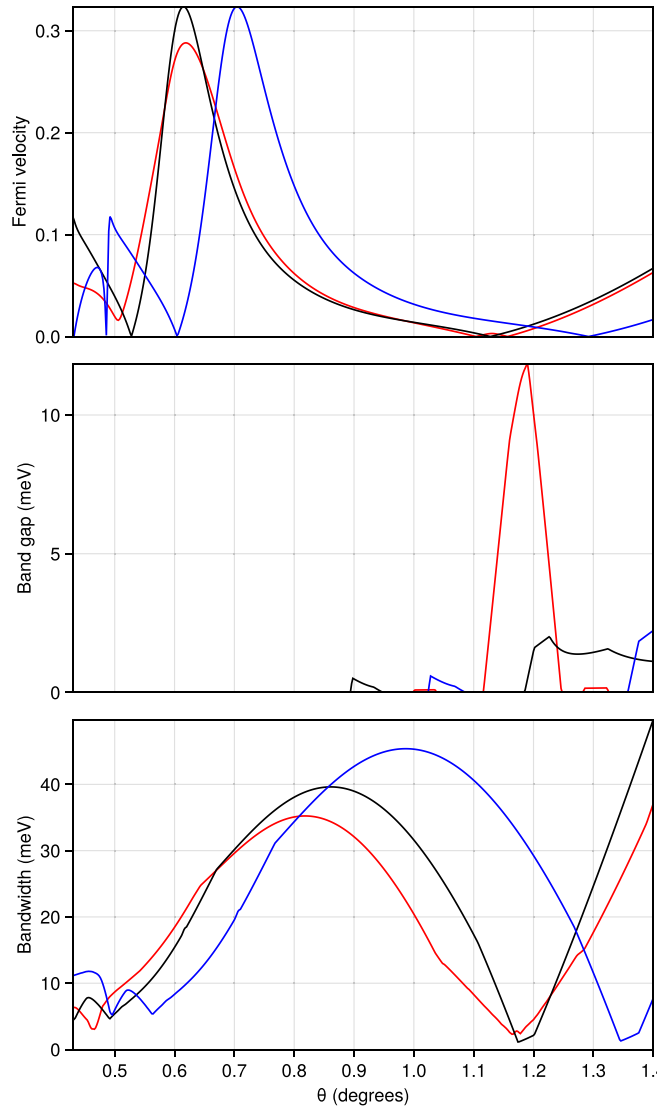

FIG. 5. Normalized Fermi velocities $v_{\mathrm{F},\theta}/v_{\mathrm{F}}$, band gaps, and bandwidths of the almost flat bands, where v_{F} is the Fermi velocity of the monolayer of $\mathcal{H}_\theta^{\mathrm{BM}}$ (black) with $w_{\mathrm{AA}} = w_{\mathrm{AB}} = 110\,\mathrm{meV}$, of $\mathcal{H}_\theta^{\mathrm{BM}}$ (blue) with $w_{\mathrm{AA}} = w_{\mathrm{AB}} = 126\,\mathrm{meV}$, and of $(\widetilde{\mathcal{H}}_{d,\theta}, \widetilde{\mathcal{S}}_d)$ (red) as functions of θ.

angles (minimal Fermi velocity vs minimal bandwidth), the BM Hamiltonian is not gapped at the flat bands while ours is so. Finally, we remark that the BM model with $w_{\mathrm{AA}} = w_{\mathrm{AB}} = 110\,\mathrm{meV}$ (the empirical values used in Ref. [6]) is a better approximation of our model than the BM model with $w_{\mathrm{AA}} = w_{\mathrm{AB}} = 126\,\mathrm{meV}$ (the values derived from DFT, see next section).

A thorough comparison of the band diagrams of various continuous models (including atomic relaxation) will be the matter of a forthcoming paper.

B. Numerical details

In order to numerically compute the Fourier coefficients of $[\widetilde{\Sigma}_d]_{jj'}$, $[\widetilde{\mathbb{V}}_d]_{jj'}$, and $[\widetilde{\mathbb{W}}_d^{\pm}]_{jj'}$, we have developed a code in Julia [27], interfaced with the DFTK planewave DFT package [28].

The single-layer graphene Kohn-Sham model is solved with DFTK using the PBE exchange-correlation functional with Goedecker-Teter-Hutter (GTH) pseudopotential, a unit cell of height $L_z = 110$ bohr, an energy cutoff of $E_{\mathrm{cut}} = 900\,\mathrm{eV}$, and a $(5 \times 5 \times 1)$ k-point grid. We extract from the DFTK computation $\mathbb{L}$-periodic functions u_1 and u_2 such that $\Phi_1(\mathbf{r}) = e^{i\mathbf{K}\cdot\mathbf{x}}u_1(\mathbf{r})$ and $\Phi_2(\mathbf{r}) = e^{i\mathbf{K}\cdot\mathbf{x}}u_2(\mathbf{r})$ form an orthogonal basis of $\mathrm{Ker}(H_{\mathbf{K}}^{(1)} - \mu_{\mathrm{F}})$, where $H_{\mathbf{K}}^{(1)}$ is the Bloch fiber of the single-layer graphene Kohn-Sham Hamiltonian at the Dirac point $\mathbf{K}$. We also extract the local component of the single-layer graphene Kohn-Sham potential V (as well as the required information on the nonlocal component of the carbon atom GTH pseudopotential, see Supplemental Materials [23] for details). More precisely, DFTK returns the Fourier coefficients of the $\mathbb{L}$-periodic functions u_j [assumed to be well-oriented, see Eq. (4)] and V, of the form

$$f(\mathbf{r}) = \sum_{m_1,m_2,m_z \in \mathbb{Z}} [f]_{m_1,m_2,m_z} \frac{\mathrm{e}^{\mathrm{i}\left((m_1\mathbf{a}_1^* + m_2\mathbf{a}_2^*)\cdot\mathbf{x} + m_z\frac{2\pi z}{L_z}\right)}}{|\Omega|^{1/2}L_z^{1/2}}. \quad (22)$$

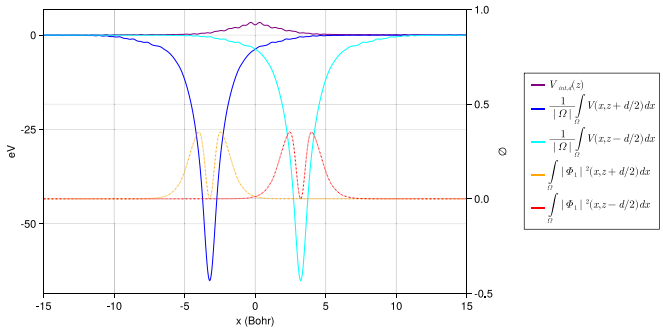

FIG. 6. Effective potential $V_{\mathrm{int},d}$ and in-plane averages of the vertically shifted single-layer Kohn-Sham potential V and Bloch wave densities $|\Phi_1|^2$ for interlayer distance $d = 6.45$ bohr.

At the discrete level, these sums are finite and run over the triplets of integers $(m_1, m_2, m_z) \in \mathbb{Z}^3$ such that

$$\frac{|m_1\mathbf{a}_1^* + m_2\mathbf{a}_2^* + \mathbf{K}|^2 + m_z^2 \frac{4\pi^2}{L_z^2}}{2} \leqslant E_{\mathrm{cut}}.$$

The Kohn-Sham potential $V^{(2)}_{\mathrm{int},d}$ is computed by averaging the disregistries $\mathbf{y}$ in a (5×5) uniform grid of the graphene unit cell. In the results reported below, we used the experimental interlayer mean distance $d = 6.45$ bohr. In accordance with the results in Ref. [25], the potential $V_{\mathrm{int},d,\mathbf{y}}(z)$ only slightly depends on $\mathbf{y}$:

$$\delta_{V_{\mathrm{int},d}} := \frac{\int_{\Omega\times\mathbb{R}} \left|V_{\mathrm{int,d},\mathbf{y}}(z) - V_{\mathrm{int,d}}(z)\right|^2 \mathrm{d}\mathbf{y}\mathrm{d}z}{|\Omega| \int_{\mathbb{R}} V_{\mathrm{int}}(z)^2 \, \mathrm{d}z} \simeq 1 \times 10^{-4}.$$

The effective potential $V_{\mathrm{int},d}$ and in-plane averages of the vertically shifted single-layer Kohn-Sham potential V and Bloch wave densities $|\Phi_1|^2$ are plotted in Fig. 6.

For f and g of the form Eq. (22), we find that

$$(\!(f, g)\!)^{+-}_d(\mathbf{X}) = \sum_{m_1,m_2\in\mathbb{Z}} [f|g]_{d,m_1,m_2} \frac{\mathrm{e}^{\mathrm{i}(m_1\mathbf{a}^*_{1,M} + m_2\mathbf{a}^*_{2,M})\cdot\mathbf{X}}}{|\Omega_M|^{1/2}},$$

where the coefficient $[f|g]_{d,m_1,m_2}$ is given by

$$|\Omega_{\mathrm{M}}|^{1/2} \sum_{m_z\in\mathbb{Z}} [\overline{f}]_{-m_1,-m_2,m_z} [g]_{-m_1,-m_2,m_z} \mathrm{e}^{\mathrm{i}m_z \frac{2\pi}{L_z} d}.$$

For $d = 6.45$ bohr, we obtain the BM parameters

$$w^{d=6.45\,a.u.}_{\mathrm{AA}} = w^{d=6.45\,a.u.}_{\mathrm{AB}} \simeq 126 \text{ meV},$$

in good agreement with the value $w_{\mathrm{AA}} = w_{\mathrm{AB}} = 110$ meV chosen in Ref. [6] to fit experimental data.

C. Numerical justification of the Bistritzer-MacDonald model

As discussed in Sec. III D, the BM model can be deduced from our reduced model by assuming that $\Sigma_d(\mathbf{X})$ and its gradient can be neglected, that $\mathbb{W}^{\pm}_d(\mathbf{X})$ is proportional to the identity matrix, and that $\mathbb{V}_d(\mathbf{X})$ is of the form Eq. (14) for some well-chosen parameters w_{AA} and w_{AB}. To test these assumptions, we first plot in Figs. 7–10 the real-space structures and magnitudes of the functions $[\Sigma_d]_{jj'}(\mathbf{X})$, $|\nabla\Sigma_d(\mathbf{X})|$, $[\mathbb{W}^{\pm}_d]_{jj'}(\mathbf{X}) - \fint_\Omega [\mathbb{W}^{\pm}_d]_{jj'}$, and $[\mathbb{V}_d]_{jj'}(\mathbf{X}) - \boldsymbol{V}_{jj'}(\mathbf{X})$ for $d = 6.45$ bohr and $w_{\mathrm{AA}} = w_{\mathrm{BB}} = 126$ meV. We can see these

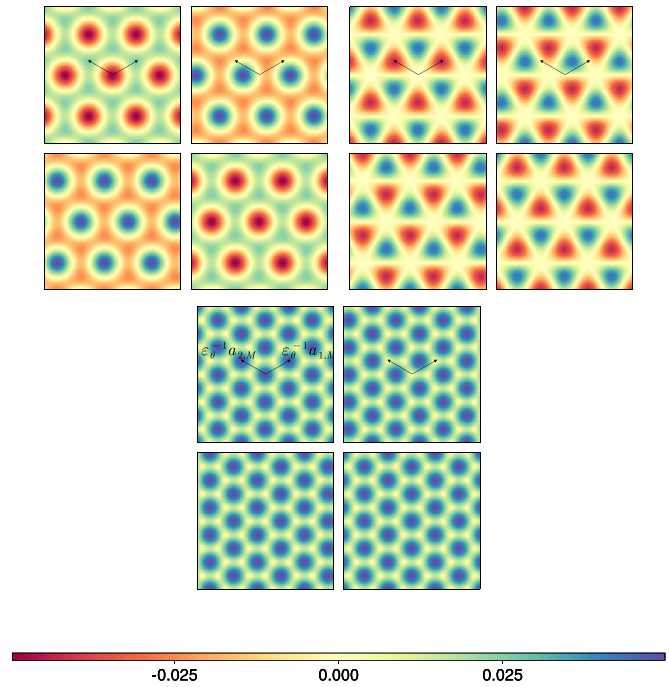

FIG. 7. The 4 entries of, respectively, the real part, the imaginary part, and the modulus of the matrix-valued function $\Sigma_d(\mathbf{X})$ for $d = 6.45$ bohr.

fields are indeed small in the relevant units: $\Sigma_d(\mathbf{X})$ is small ($\sim$0.03 compared to 1), $|\nabla\Sigma_d(\mathbf{X})|$ is small ($\sim$0.03 compared to the Fermi velocity $v_{\mathrm{F}} \simeq 0.380$), and $\mathbb{V}_d(\mathbf{X}) - \boldsymbol{V}(\mathbf{X})$ and $\mathbb{W}^{\pm}_d(\mathbf{X}) - \fint_\Omega \mathbb{W}^{\pm}_d$ are small (respectively, $\sim$1 meV and $\sim$40 meV compared to the interlayer characteristic interaction energy 126 meV). This provides a new argument supporting the validity of the BM model.

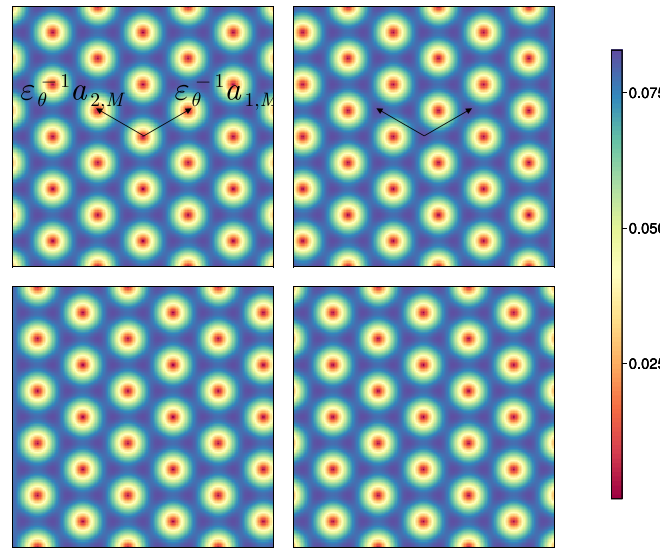

FIG. 8. The 4 entries of the matrix-valued function $|\nabla\Sigma_d(\mathbf{X})|/v_{\mathrm{F}}$ for $d = 6.45$ bohr.

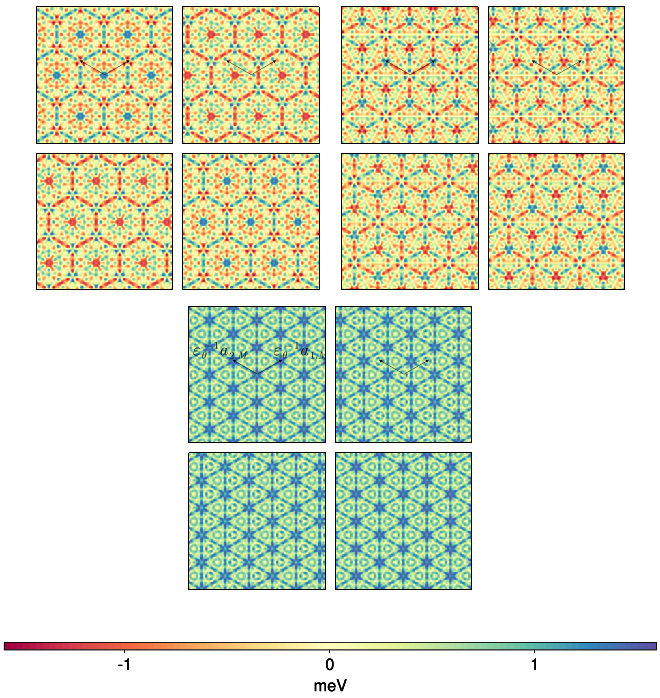

FIG. 9. The 4 entries of, respectively, the real part, the imaginary part, and the modulus of the matrix-valued function $\mathbb{V}_d(\mathbf{X}) - \boldsymbol{V}(\mathbf{X})$ for $d = 6.45$ bohr and $w_{\mathrm{AA}} = w_{\mathrm{BB}} = 126$ meV.

V. CONCLUSION AND PERSPECTIVES

We have proposed a simple direct space construction of an effective model for TBG at the moiré scale from DFT based on

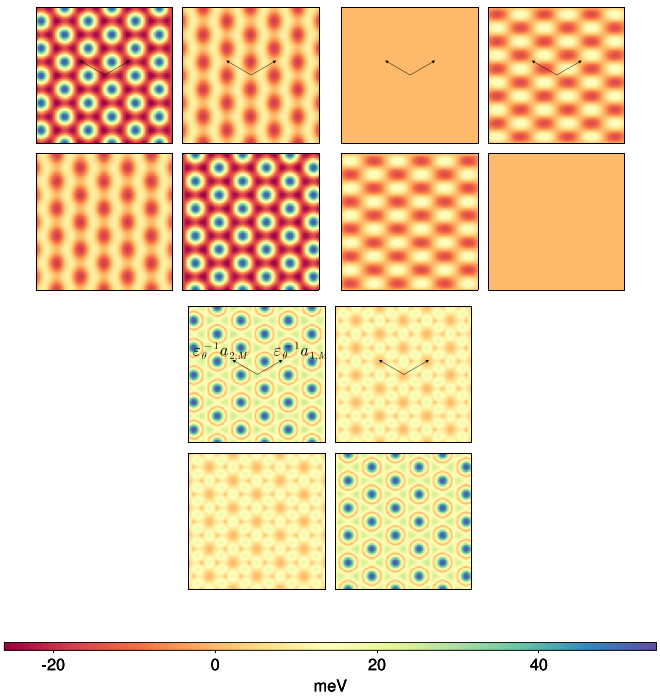

FIG. 10. The 4 entries of, respectively, the real part, the imaginary part, and the modulus of the matrix-valued function $\mathbb{W}_d^+(\mathbf{X}) - \fint_\Omega \mathbb{W}_d^+$ for $d = 6.45$ bohr.

(1) an approximation of the TBG Kohn-Sham Hamiltonian following the method introduced in Ref. [25];

(2) a variational approximation of the so-obtained Hamiltonian at the Dirac point $\mathbf{K}$ around the Fermi level;

(3) an asymptotic expansion valid for small twist angles θ.

This effective model has a structure similar to the one of the Bistritzer-MacDonald model, but contains additional terms: a nontrivial overlap operator $\mathcal{S}_d$, intralayer effective potentials $\mathbb{W}_d^{\pm}$, more complicated interlayer effective potentials $\mathbb{V}_d$, and higher-order corrections. We show numerically that both models give rise to similar band diagrams, the main difference being that at the first magic angle, the almost flat bands are separated from the rest of the spectrum in our model, which is not the case in the BM model.

It is well established that atomic relaxation plays a key role in the electronic properties of TBG (see, e.g., Ref. [29] and references therein), especially at very small twist angles $\theta < 1^\circ$. We derived our reduced model Eq. (12) for an unrelaxed configuration, but it can be extended to take both in-plane and out-of-plane relaxation [19,30,31] into account. The derivation and numerical simulation of such a model is work in progress.

The same methodology can be applied to study the propagation of low-energy wavepackets localized in momentum space in both the $\mathbf{K}$ and $\mathbf{K}'$ valleys. The approximation space then contains functions of the form

$$(\alpha : \Phi)_{d,\theta}(\mathbf{x}, z) := \sum_{p\in\{\mathbf{K},\mathbf{K}'\}} \sum_{\substack{\eta\in\{\pm1\}\\ j\in\{1,2\}}} \alpha_{\eta,j,p}(\varepsilon_\theta \mathbf{x}) \big(U_{d,\theta}^{\eta} \Phi_j^p\big)(\mathbf{x}, z).$$

For the unrelaxed configuration, it can be checked that the two valleys are uncoupled. Intervalley coupling may appear, however, when taking atomic relaxations into account.

ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement EMC2 No. 810367) and from the Simons Targeted Grant Award No. 896630. The authors thank Simon Becker, Antoine Levitt, Mitchell Luskin, Allan MacDonald, Etienne Polack, and Alexander Watson for useful discussions and comments. Part of this work was done during the IPAM program, Advancing quantum mechanics with mathematics and statistics.

APPENDIX A: DERIVATION OF THE EFFECTIVE MODEL

Our goal is to identify the leading orders terms in Eq. (8) in the limit of small twist angle. For that, we use the following elementary lemmas. Here $\mathcal{S}$ is the Schwartz class of smooth function decaying faster than any polynomial.

Lemma 1. Let $\beta \in \mathcal{S}(\mathbb{R}^2, \mathbb{C})$ and $u \in L^1_{\mathrm{per}}(\Omega \times \mathbb{R}; \mathbb{C})$. Then, as $\varepsilon \to 0$, we have

$$\int_{\mathbb{R}^3} \beta(\varepsilon \mathbf{x}) u(\mathbf{x}, z)\, \mathrm{d}\mathbf{x}\, \mathrm{d}z = \varepsilon^{-2} \left(\frac{1}{|\Omega|} \int_{\Omega\times\mathbb{R}} u \right) \left(\int_{\mathbb{R}^2} \beta \right) + O(\varepsilon^\infty).$$

Proof. Expanding u as $u(\mathbf{x}, z) = \sum_{\mathbf{G}\in\mathbb{L}^*} u_{\mathbf{G}}(z) \mathrm{e}^{\mathrm{i}\mathbf{G}\cdot\mathbf{x}}$ and making the change of variable $\mathbf{X} = \varepsilon\mathbf{x}$, we obtain that the

left-hand side equals

$$\varepsilon^{-2}\sum_{\mathbf{G}\in\mathbb{L}^*}\left(\int_{\mathbb{R}} u_{\mathbf{G}}(z)\mathrm{d}z\right)\int_{\mathbb{R}}\beta(\mathbf{X})\mathrm{e}^{\mathrm{i}\varepsilon^{-1}\mathbf{G}\cdot\mathbf{X}}\,\mathrm{d}\mathbf{X}$$
$$=\varepsilon^{-2}\sum_{\mathbf{G}\in\mathbb{L}^*}\left(\int_{\mathbb{R}} u_{\mathbf{G}}(z)\,\mathrm{d}z\right)\widehat{\beta}(\varepsilon^{-1}\mathbf{G}).$$

As $\beta\in\mathcal{S}(\mathbb{R}^2;\mathbb{C})$, $\widehat{\beta}$ decays faster than any polynomial. Isolating the term $\mathbf{G}=\mathbf{0}$ gives the result. ■

We now denote by $L^2_{\mathbf{K}}$ the set of locally square integrable functions which are $\mathbf{K}$ quasiperiodic.

Lemma 2. Let $\beta\in\mathcal{S}(\mathbb{R}^2,\mathbb{C})$, $\Phi,\Phi'\in L^2_{\mathbf{K}}(\Omega\times\mathbb{R};\mathbb{C})$, and $\eta,\eta'\in\{\pm1\}$. Then, as $\varepsilon\to0$, we have

$$\int_{\mathbb{R}^3}\beta(\varepsilon_\theta\mathbf{x})\big(\overline{U^{\eta}_{d,\theta}\Phi}\cdot U^{\eta'}_{d,\theta}\Phi'\big)(\mathbf{x},z)\,\mathrm{d}\mathbf{x}\,\mathrm{d}z$$
$$=\frac{\varepsilon_\theta^{-2}}{|\Omega|}\int_{\mathbb{R}^2}\beta(\mathbf{X})\langle\!\langle\Phi,\Phi'\rangle\!\rangle_d^{\eta\eta'}(\mathbf{X})\mathrm{d}\mathbf{X}+O\big(\varepsilon_\theta^{\infty}\big),$$

where $\langle\!\langle\Phi,\Phi'\rangle\!\rangle_d^{\eta\eta'}(\mathbf{X})$ is defined by

$$\int_{\Omega\times\mathbb{R}}\overline{\Phi\left(\mathbf{x}-\eta\frac{1}{2}J\mathbf{X},z-\eta\frac{d}{2}\right)}$$
$$\times\,\Phi'\left(\mathbf{x}-\eta'\frac{1}{2}J\mathbf{X},z-\eta'\frac{d}{2}\right)\mathrm{d}\mathbf{x}\,\mathrm{d}z.$$

In the case $\eta=\eta'$, we have $\langle\!\langle\Phi,\Phi'\rangle\!\rangle_d^{\eta\eta'}(\mathbf{X})=\langle\Phi,\Phi'\rangle_{L^2(\Omega\times\mathbb{R})}$, independent of $\mathbf{X}$. When $\eta\neq\eta'$, the function $\langle\!\langle\Phi,\Phi'\rangle\!\rangle_d^{\eta\eta'}(\mathbf{X})$ does depend on $\mathbf{X}$ in general and is $J\mathbb{L}$ periodic.

Comparing $\langle\!\langle\cdot,\cdot\rangle\!\rangle_d$ and $(\!(\cdot,\cdot)\!)_d$ in Eq. (11), we see that with $\Phi=\mathrm{e}^{\mathrm{i}\mathbf{K}\cdot\mathbf{x}}u$,

$$\langle\!\langle\Phi,\Phi'\rangle\!\rangle_d^{+-}(\mathbf{X})=\mathrm{e}^{-\mathrm{i}J\mathbf{K}\cdot\mathbf{X}}(\!(u,u')\!)_d^{+-}(\mathbf{X}),$$

and

$$\langle\!\langle\Phi,\Phi'\rangle\!\rangle_d^{++}=(\!(u,u')\!)_d^{++}=\langle\Phi,\Phi'\rangle=\langle u,u'\rangle.$$

In what follows, we express our quantities with $\langle\!\langle\Phi,\Phi'\rangle\!\rangle_d$, but we translated our results with $(\!(u,u')\!)_d$ to present our reduced model.

Proof of Lemma 2. In the case $\eta=\eta'$, the left-hand side equals

$$\int_{\mathbb{R}^3}\beta(\varepsilon_\theta\mathbf{x})(\overline{\Phi}\Phi')\left(R_{\eta\frac{\theta}{2}}\mathbf{x},z-\eta\frac{d}{2}\right)\mathrm{d}\mathbf{x}\,\mathrm{d}z$$
$$=\varepsilon_\theta^{-2}\int_{\mathbb{R}^3}\beta(\mathbf{X})(\overline{\Phi}\Phi')\big(\varepsilon_\theta^{-1}R_{\eta\frac{\theta}{2}}\mathbf{X},z\big)\,\mathrm{d}\mathbf{X}\,\mathrm{d}z.$$

As the function $\overline{\Phi}\Phi'$ is $\mathbb{L}$ periodic, the result can be obtained by applying the same arguments as in the proof of Lemma 1.

Let us now focus on the case when $\eta\neq\eta'$, and prove the result for $\eta'=+1$ and $\eta=-1$, the other case being similar. Let $u(\mathbf{r}):=\mathrm{e}^{-\mathrm{i}\mathbf{K}\cdot\mathbf{x}}\Phi(\mathbf{r})$ and $u'(\mathbf{r}):=\mathrm{e}^{-\mathrm{i}\mathbf{K}\cdot\mathbf{x}}\Phi'(\mathbf{x})$ be the periodic components of the Bloch waves Φ and Φ' respectively. We have

$$\big(\overline{U_{d,\theta}\Phi}\cdot U^{-1}_{d,\theta}\Phi'\big)(\mathbf{x},z)$$
$$=\mathrm{e}^{\mathrm{i}\mathbf{K}\varepsilon_\theta\cdot J\mathbf{x}}\overline{u\left(c_\theta\mathbf{x}-\frac{1}{2}\varepsilon_\theta J\mathbf{x},z-\frac{d}{2}\right)}$$
$$\times\,u'\left(c_\theta\mathbf{x}+\frac{1}{2}\varepsilon_\theta J\mathbf{x},z+\frac{d}{2}\right).$$

Introducing the Fourier expansions of u and u', this is also

$$\frac{1}{|\Omega|}\sum_{\mathbf{G},\mathbf{G}'\in\mathbb{L}^*}\left[\overline{u_{\mathbf{G}}}\left(z-\frac{d}{2}\right)u'_{\mathbf{G}'}\left(z+\frac{d}{2}\right)e^{\mathrm{i}(\mathbf{G}'-\mathbf{G})\cdot c_\theta\mathbf{x}}\right]$$
$$\times\,e^{\mathrm{i}\frac{1}{2}(\mathbf{G}+\mathbf{G}'+2\mathbf{K})\varepsilon_\theta\cdot J\mathbf{x}}.$$

Note that the last phase factor varies at the moiré scale. The term in brackets is a $c_\theta^{-1}\mathbb{L}$-periodic function with zero mean unless $\mathbf{G}=\mathbf{G}'$. Reasoning as above, we obtain

$$\int_{\mathbb{R}^3}\beta(\varepsilon_\theta\mathbf{x})\big(\overline{U^{\eta}_{d,\theta}\Phi}\cdot U^{\eta'}_{d,\theta}\Phi'\big)(\mathbf{x},z)\,\mathrm{d}\mathbf{x}\,\mathrm{d}z$$
$$=\frac{\varepsilon_\theta^{-2}}{|\Omega|}\int_{\mathbb{R}^2}\beta(\mathbf{X})\left[\mathrm{e}^{\mathrm{i}\mathbf{K}\cdot J\mathbf{X}}\sum_{\mathbf{G}\in\mathbb{L}^*}\left(\int_{\mathbb{R}}\overline{u_{\mathbf{G}}}\left(z-\frac{d}{2}\right)\right.\right.$$
$$\left.\left.\times\,u'_{\mathbf{G}}\left(z+\frac{d}{2}\right)\mathrm{d}z\right)\mathrm{e}^{\mathrm{i}\mathbf{G}\cdot J\mathbf{X}}\right]\mathrm{d}\mathbf{X}+O\big(\varepsilon_\theta^{\infty}\big).$$

Finally, by Parseval theorem, the term in brackets is equal to $\langle\!\langle\Phi,\Phi'\rangle\!\rangle_d^{\eta\eta'}(\mathbf{X})$, which proves the result. ■

1. Effective overlap operator

Let us first focus on the left-hand side of Eq. (8). Setting $\Phi_{\eta j}:=U^{\eta}_{d,\theta}\Phi_j$, we have (all quantities are summed over $\eta,\eta'\in\{\pm1\}$ and $j,j'\in\{1,2\}$)

$$\langle(\widetilde{\alpha}:\Phi)_{d,\theta},(\alpha(\varepsilon_\theta t):\Phi)_{d,\theta}\rangle$$
$$=\int_{\mathbb{R}^3}(\overline{\widetilde{\alpha}_{\eta j}}\alpha_{\eta' j'}(t))(\varepsilon_\theta\mathbf{x})\times(\overline{\Phi_{\eta' j'}}\Phi_{\eta j})(\mathbf{x},z)\,\mathrm{d}\mathbf{x}\,\mathrm{d}z.$$

Using directly Lemma 2 (with $\beta=\overline{\widetilde{\alpha}_{\eta j}}\alpha_{\eta' j'}$), we obtain that this term equals

$$\varepsilon_\theta^{-2}\int_{\mathbb{R}^2}(\overline{\widetilde{\alpha}_{\eta j}}\alpha_{\eta' j'}(\varepsilon_\theta t))(\mathbf{X})\langle\!\langle\Phi_j,\Phi_{j'}\rangle\!\rangle_d^{\eta\eta'}(\mathbf{X})\mathrm{d}\mathbf{X}$$

up to errors of order $O(\varepsilon^\infty)$. Ranking the components of α as $\alpha=(\alpha_{+,1},\alpha_{+,2},\alpha_{-,1},\alpha_{-,2})^T$ (the first two entries correspond to the top layer, the last two to the bottom one), we obtain

$$\partial_t\langle(\widetilde{\alpha}:\Phi)_{d,\theta},(\alpha(\varepsilon_\theta t):\Phi)_{d,\theta}\rangle=\frac{\varepsilon_\theta^{-1}}{|\Omega|}\langle\widetilde{\alpha},\mathcal{S}_d\partial_\tau\alpha(\varepsilon_\theta t)\rangle$$

up to errors of order $O(\varepsilon_\theta^\infty)$, with

$$\mathcal{S}_d=\begin{pmatrix}\mathbb{I}_2 & \Sigma_d(\mathbf{X})\\ \Sigma_d^*(\mathbf{X}) & \mathbb{I}_2\end{pmatrix},\quad[\Sigma_d]_{jj'}(\mathbf{X}):=\langle\!\langle\Phi_j,\Phi_{j'}\rangle\!\rangle_d^{+-}(\mathbf{X}).$$

2. Effective Hamiltonian operator

We now focus on the terms on the right-hand side. First, we record that

$$\begin{aligned}&\big(-\tfrac{1}{2}\Delta_{\mathbf{r}}\big)[\beta(\varepsilon_\theta\mathbf{x})\varphi(\mathbf{r})]\\&=\varepsilon_\theta^2\big(-\tfrac{1}{2}\Delta\beta\big)(\varepsilon_\theta\mathbf{x})\varphi(\mathbf{r})+\varepsilon_\theta(-\mathrm{i}\nabla\beta)(\varepsilon_\theta\mathbf{x})\cdot(-\mathrm{i}\nabla_{\mathbf{x}}\varphi)(\mathbf{r})\\&\quad+\beta(\varepsilon_\theta\mathbf{x})\big(-\tfrac{1}{2}\Delta\varphi\big)(\mathbf{r}).\end{aligned}\tag{A1}$$

This gives

$$\big\langle(\widetilde{\alpha}:\Phi)_{d,\theta},\big(H^{(2)}_{d,\theta}-\mu_F\big)(\alpha:\Phi)_{d,\theta}\big\rangle=g_1+g_2+g_3,$$

with

$$g_1:=\varepsilon_\theta^2\int_{\mathbb{R}^3}\left[\overline{\widetilde{\alpha}_{\eta,j}\Big(-\frac{1}{2}\Delta\alpha_{\eta',j'}\Big)}\right](\varepsilon_\theta\mathbf{x})(\overline{\Phi}_{\eta,j}\Phi_{\eta',j'})(\mathbf{r})\mathrm{d}\mathbf{r},$$

$$g_2:=\varepsilon_\theta\int_{\mathbb{R}^3}[\overline{\widetilde{\alpha}_{\eta,j}(-\mathrm{i}\nabla\alpha_{\eta',j'})}](\varepsilon_\theta\mathbf{x})\cdot[\overline{\Phi}_{\eta,j}(-\mathrm{i}\nabla_{\mathbf{x}}\Phi_{\eta',j'})](\mathbf{r})\mathrm{d}\mathbf{r},$$

$$g_3:=\int_{\mathbb{R}^3}[\overline{\widetilde{\alpha}_{\eta,j}}\alpha_{\eta',j'}(\varepsilon_\theta\mathbf{x})]\cdot\big[\overline{\Phi}_{\eta,j}\times\big(H^{(2)}_{d,\theta}-\mu_F\big)\Phi_{\eta',j'}\big](\mathbf{r})\mathrm{d}\mathbf{r}.$$

For the first term g_1, we apply directly Lemma 2, which gives

$$g_1=\frac{1}{|\Omega|}\int_{\mathbb{R}^2}\left(\overline{\widetilde{\alpha}_{\eta j}\Big(-\frac{1}{2}\Delta\alpha_{\eta' j'}\Big)}\right)(\mathbf{X})\langle\!\langle\Phi_j,\Phi_{j'}\rangle\!\rangle_d^{\eta\eta'}(\mathbf{X})\mathrm{d}\mathbf{X}$$

up to errors of order $O(\varepsilon_\theta^\infty)$. This term can be written as $g_1=|\Omega|^{-1}\langle\widetilde{\alpha},G_1\alpha\rangle+O(\varepsilon_\theta^\infty)$ with

$$G_1:=\mathcal{S}_d(\mathbf{X})\big(-\tfrac{1}{2}\Delta\big).$$

For the second term g_2, we first notice that

$$\begin{aligned}(-\mathrm{i}\nabla_{\mathbf{x}}\Phi_{\eta',j'})(\mathbf{r})&=(-\mathrm{i}\nabla_{\mathbf{x}})\left[\Phi_{j'}\left(R^{-\eta'}_{-\theta/2}\mathbf{x},z-\eta'\frac{d}{2}\right)\right]\\&=R^{\eta'}_{-\theta/2}\big[U^{\eta'}_{d,\theta}(-\mathrm{i}\nabla_{\mathbf{x}}\Phi_{j'})\big](\mathbf{r}).\end{aligned}$$

This gives, using arguments similar to the ones of Lemma 2, that g_2 equals

$$\begin{aligned}&\frac{\varepsilon_\theta^{-1}}{|\Omega|}\int_{\mathbb{R}^2}\big(\overline{\widetilde{\alpha}_{\eta j}}(-\mathrm{i}\nabla\alpha_{\eta' j'})\big)(\mathbf{X})\cdot R^{\eta'}_{-\theta/2}\langle\!\langle\Phi_j,(-\mathrm{i}\nabla_{\mathbf{x}}\Phi_{j'})\rangle\!\rangle_d^{\eta\eta'}\\&\quad\times(\mathbf{X})\mathrm{d}\mathbf{X}\end{aligned}$$

up to errors of order $O(\varepsilon^\infty)$. To deal with the diagonal terms $\eta=\eta'$, we use our orientation Eq. (4). Regarding the off-diagonal terms (here for $\eta=+1$ and $\eta'=-1$), we have, using that $J^T=-J$,

$$\begin{aligned}\nabla\langle\!\langle\Phi,\Phi'\rangle\!\rangle_d^{+-}(\mathbf{X})&=\frac{1}{2}J\int_{\Omega\times\mathbb{R}}\overline{(\nabla_{\mathbf{x}}\Phi)\Big(\mathbf{x}-\frac{1}{2}J\mathbf{X},z-\frac{d}{2}\Big)}\\&\quad\times\Phi'\Big(\mathbf{x}+\frac{1}{2}J\mathbf{X},z+\frac{d}{2}\Big)\mathrm{d}\mathbf{x}\,\mathrm{d}z\\&\quad-\frac{1}{2}J\int_{\Omega\times\mathbb{R}}\overline{\Phi\Big(\mathbf{x}-\frac{1}{2}J\mathbf{X},z-\frac{d}{2}\Big)}(\nabla_{\mathbf{x}}\Phi')\\&\quad\times\Big(\mathbf{x}+\frac{1}{2}J\mathbf{X},z+\frac{d}{2}\Big)\mathrm{d}\mathbf{x}\,\mathrm{d}z.\end{aligned}$$

Integrating by part the first term in the RHS, and multiplying by $(-i)$ shows that

$$\langle\!\langle\Phi,(-\mathrm{i}\nabla_{\mathbf{x}})\Phi'\rangle\!\rangle_d^{+-}=J(-\mathrm{i}\nabla)\langle\!\langle\Phi,\Phi'\rangle\!\rangle_d^{+-}.$$

Similarly, in the case $\eta=-1$ and $\eta'=1$, we have

$$\langle\!\langle\Phi,(-\mathrm{i}\nabla_{\mathbf{x}})\Phi'\rangle\!\rangle_d^{-+}=-J(-\mathrm{i}\nabla)\langle\!\langle\Phi,\Phi'\rangle\!\rangle_d^{-+}.$$

This gives $g_2=|\Omega|^{-1}\langle\widetilde{\alpha},G_2\alpha\rangle+O(\varepsilon_\theta^\infty)$ with G_2 of the form (we use that $R_{-\theta/2}=c_\theta\mathbb{I}_2+\frac{1}{2}\varepsilon_\theta J$)

$$\begin{aligned}&\frac{c_\theta}{\varepsilon_\theta}\begin{pmatrix}v_F\boldsymbol{\sigma}\cdot(-\mathrm{i}\nabla)&J(-\mathrm{i}\nabla\Sigma_d)(\mathbf{X})\cdot(-\mathrm{i}\nabla)\\J(-\mathrm{i}\nabla\Sigma_d^*)(\mathbf{X})\cdot(-\mathrm{i}\nabla)&v_F\boldsymbol{\sigma}\cdot(-\mathrm{i}\nabla)\end{pmatrix}\\&+\frac{1}{2}\begin{pmatrix}v_F\boldsymbol{\sigma}\cdot[J(-\mathrm{i}\nabla)]&(-\mathrm{i}\nabla\Sigma_d)(\mathbf{X})\cdot(-\mathrm{i}\nabla)\\(-\mathrm{i}\nabla\Sigma_d^*)(\mathbf{X})\cdot(-\mathrm{i}\nabla)&-v_F\boldsymbol{\sigma}\cdot[J(-\mathrm{i}\nabla)]\end{pmatrix}.\end{aligned}$$

Finally, for the term g_3, we recall that Φ_j is an eigenvector of the single-layer graphene Hamiltonian $H^{(1)}$ associated with the eigenvalue μ_{F} and get

$$\big[\big(H^{(2)}_{d,\theta}-\mu_F\big)\Phi_{\eta,j}\big]=\big(U^{-\eta}_{d,\theta}V\big)(\mathbf{r})\Phi_{\eta,j}(\mathbf{r})+V_{\mathrm{int},d}(z)\Phi_{\eta,j}(\mathbf{r}),$$

and

$$\begin{aligned}g_3=\int_{\mathbb{R}^3}&[(\overline{\widetilde{\alpha}_{\eta',j'}}\alpha_{\eta,j})(\varepsilon_\theta\mathbf{x})]\\&\times\big[\big(U^{-\eta}_{d,\theta}V\big)\overline{\Phi}_{\eta',j'}\Phi_{\eta,j}+V_{\mathrm{int},d}\overline{\Phi}_{\eta',j'}\Phi_{\eta,j}\big](\mathbf{r})\mathrm{d}\mathbf{r}.\end{aligned}$$

Using reasoning similar to the proof of Lemma 2, we obtain that $g_3=\langle\widetilde{\alpha},G_3\alpha\rangle+O(\varepsilon_\theta^\infty)$, with

$$G_3=\varepsilon_\theta^{-2}\begin{pmatrix}\mathbb{W}_d^+(\mathbf{X})&\mathbb{V}_d(\mathbf{X})\\\mathbb{V}_d(\mathbf{X})^*&\mathbb{W}_d^-(\mathbf{X})\end{pmatrix},$$

where [recall that the notation $(\!(f,g)\!)_d$ was defined in Eq. (11), and is used when f and g are periodic]

$$[\mathbb{V}_d(\mathbf{X})]_{jj'}=\langle\!\langle V\Phi_j,\Phi_{j'}\rangle\!\rangle_d^{+-}(\mathbf{X})+\int_{\Omega\times\mathbb{R}}\overline{\Phi_j\Big(\mathbf{x}-\frac{1}{2}J\mathbf{X},z-\frac{d}{2}\Big)}\Phi_{j'}\Big(\mathbf{x}+\frac{1}{2}J\mathbf{X},z+\frac{d}{2}\Big)V_{\mathrm{int},d}(z)\,\mathrm{d}\mathbf{r},$$

$$[\mathbb{W}_d^\pm(\mathbf{X})]_{jj'}=(\!(\Phi_j\overline{\Phi_{j'}},V)\!)_d^{\pm\mp}(\mathbf{X})+\int_{\Omega\times\mathbb{R}}(\overline{\Phi_j}\Phi_{j'})\Big(\mathbf{x},z\mp\frac{d}{2}\Big)V_{\mathrm{int},d}(z)\mathrm{d}\mathbf{r}.$$

The second term of $\mathbb{W}_d^\pm$ is a constant matrix (independent of $\mathbf{X}$). We prove in the Supplemental Material [23] that this term is proportional to $\mathbb{I}_2$. Finally, since $V_{\mathrm{int},d}(-z)=V_{\mathrm{int},d}(z)$, this matrix is the same for the $\mathbb{W}_d^+$ and the $\mathbb{W}_d^-$ terms.

To conclude and obtain the expression in Eq. (10), we have used the equality

$$\Sigma_d\big(-\tfrac{1}{2}\Delta\big)+\tfrac{1}{2}(-\mathrm{i}\nabla\Sigma_d)\cdot(-\mathrm{i}\nabla)=-\tfrac{1}{2}\mathrm{div}(\Sigma_d(\mathbf{X})\nabla\bullet).$$

APPENDIX B: PROOF THAT $w^d_{\mathrm{AA}} = w^d_{\mathrm{AB}}$

Recall that w^d_{AA} and w^d_{AB} are defined, respectively, in Eqs. (19)-(20). We prove in the Supplemental Material [23] that $[\mathbb{V}_d]_{11}$ satisfies $[\mathbb{V}_d]_{11}(R_{\frac{2\pi}{3}}\mathbf{X}) = [\mathbb{V}_d]_{11}(\mathbf{X})$, where $R_{\frac{2\pi}{3}}$ is the $2\pi/3$ rotation. Recalling the definition of $G(\mathbf{X})$ in Eq. (15) and using that $\mathbf{q}_3 = R_{\frac{2\pi}{3}}\mathbf{q}_1$ while $\mathbf{q}_2 = (R_{\frac{2\pi}{3}})^2\mathbf{q}_1$, we obtain

$$w^d_{\mathrm{AA}} = \frac{1}{3|\Omega_{\mathrm{M}}|}\left(\sum_{n=1}^{3}\int_{\Omega_M}[\mathbb{V}_d]_{11}(\mathbf{X})\mathrm{e}^{\mathrm{i}\mathbf{q}_n\cdot\mathbf{X}}\,\mathrm{d}\mathbf{X}\right) = \frac{1}{|\Omega_{\mathrm{M}}|}\left(\int_{\Omega_M}[\mathbb{V}_d]_{11}(\mathbf{X})\mathrm{e}^{\mathrm{i}\mathbf{q}_1\cdot\mathbf{X}}\,\mathrm{d}\mathbf{X}\right).$$

From the definition of $\mathbb{V}$ and the fact that $\mathbf{q}_1 = J\mathbf{K}$, while $|\Omega_M| = |\Omega|$, we get

$$w^d_{\mathrm{AA}} = \int_{\mathbb{R}}\left(\int_{\Omega}\overline{\left[\left(V + V_{\mathrm{int},d}\left(\cdot + \frac{d}{2}\right)\right)u_1\right]\left(\mathbf{x}, z - \frac{d}{2}\right)}\,\mathrm{d}\mathbf{x}\right)\left(\fint_{\Omega} u_1\left(\mathbf{x}, z + \frac{d}{2}\right)\mathrm{d}\mathbf{x}\right)\mathrm{d}z.$$

A similar calculation leads to

$$w^d_{\mathrm{AB}} = \int_{\mathbb{R}}\left(\int_{\Omega}\overline{\left[\left(V + V_{\mathrm{int},d}\left(\cdot + \frac{d}{2}\right)\right)u_2\right]\left(\mathbf{x}, z - \frac{d}{2}\right)}\,\mathrm{d}\mathbf{x}\right)\left(\fint_{\Omega} u_1\left(\mathbf{x}, z + \frac{d}{2}\right)\mathrm{d}\mathbf{x}\right)\mathrm{d}z.$$

In the Supplemental Materials [23], we prove that $\overline{u_j(x_1, x_2, z)} = -u_j(-x_1, x_2, z)$. Since V and V_{int} are real valued, with $V(x_1, x_2, z) = V(-x_1, x_2, z)$, the parameters w^d_{AA} and w^d_{AB} are real valued. In addition, we also proved that $u_2(x_1, -x_2, z) = u_1(x_1, x_2, z)$. Together with the fact that $V(x_1, -x_2, z) = V(x_1, x_2, z)$, we deduce $w^d_{\mathrm{AA}} = w^d_{\mathrm{AB}}$.

[1] E. Y. Andrei, D. K. Efetov, P. Jarillo-Herrero, A. H. MacDonald, K. F. Mak, T. Senthil, E. Tutuc, A. Yazdani, and A. F. Young, The marvels of moiré materials, Nat. Rev. Mater. **6**, 201 (2021).

[2] S. Carr, S. Fang, and E. Kaxiras, Electronic-structure methods for twisted moiré layers, Nat. Rev. Mater. **5**, 748 (2020).

[3] 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**, 43 (2018).

[4] E. J. Mele, Commensuration and interlayer coherence in twisted bilayer graphene, Phys. Rev. B **81**, 161405(R) (2010).

[5] R. Bistritzer and A. H. MacDonald, Moiré butterflies in twisted bilayer graphene, Phys. Rev. B **84**, 035440 (2011).

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

[7] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Continuum model of the twisted graphene bilayer, Phys. Rev. B **86**, 155449 (2012).

[8] A. B. Watson, T. Kong, A. H. MacDonald, and M. Luskin, Bistritzer-MacDonald dynamics in twisted bilayer graphene, J. Math. Phys. **64**, 031502 (2023).

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

[10] S. Becker, M. Embree, J. Wittsten, and M. Zworski, Spectral characterization of magic angles in twisted bilayer graphene, Phys. Rev. B **103**, 165113 (2021).

[11] A. B. Watson and M. Luskin, Existence of the first magic angle for the chiral model of bilayer graphene, J. Math. Phys. **62**, 091502 (2021).

[12] 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).

[13] G. Trambly de Laissardière, D. Mayou, and L. Magaud, Numerical studies of confined states in rotated bilayers of graphene, Phys. Rev. B **86**, 125413 (2012).

[14] J. Jung and A. H. MacDonald, Accurate tight-binding models for the π bands of bilayer graphene, Phys. Rev. B **89**, 035405 (2014).

[15] S. Carr, D. Massatt, S. Fang, P. Cazeaux, M. Luskin, and E. Kaxiras, Twistronics: Manipulating the electronic properties of two-dimensional layered structures through their twist angle, Phys. Rev. B **95**, 075420 (2017).

[16] S. Carr, S. Fang, P. Jarillo-Herrero, and E. Kaxiras, Pressure dependence of the magic twist angle in graphene superlattices, Phys. Rev. B **98**, 085144 (2018).

[17] H. A. Le and V. N. Do, Electronic structure and optical properties of twisted bilayer graphene calculated via time evolution of states in real space, Phys. Rev. B **97**, 125136 (2018).

[18] P. Moon and M. Koshino, Optical absorption in twisted bilayer graphene, Phys. Rev. B **87**, 205404 (2013).

[19] S. Carr, S. Fang, Z. Zhu, and E. Kaxiras, Exact continuum model for low-energy electronic states of twisted bilayer graphene, Phys. Rev. Res. **1**, 013001 (2019).

[20] 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 (2019).

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

[22] B. A. Bernevig, Z.-D. Song, N. Regnault, and B. Lian, Twisted bilayer graphene. i. matrix elements, approximations, perturbation theory, and a $k \cdot p$ two-band model, Phys. Rev. B **103**, 205411 (2021).

[23] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.107.155403 for a thorough analysis of the symmetries on the moiré-scale reduced model introduced in this

paper, as well as for technical details about the additional terms originating from the nonlocal part of the pseudopotential.

[24] É. Cancès, S. Lahbabi, and M. Lewin, Mean-field models for disordered crystals, J. Math. Pures Appl. **100**, 241 (2013).

[25] G. A. Tritsaris, S. N. Shirodkar, E. Kaxiras, P. Cazeaux, M. Luskin, P. Plecháč, and E. Cancès, Perturbation theory for weakly coupled two-dimensional layers, J. Mater. Res. **31**, 959 (2016).

[26] C. L. Fefferman and M. I. Weinstein, Wave packets in honeycomb structures and two-dimensional dirac equations, Commun. Math. Phys. **326**, 251 (2014).

[27] J. Bezanson, A. Edelman, S. Karpinski, and V. Shah, Julia: A fresh approach to numerical computing, SIAM Rev. **59**, 65 (2017).

[28] M. F. Herbst, A. Levitt, and E. Cancès, DFTK: A julian approach for simulating electrons in solids, Proc. JuliaCon Conf. **3**, 69 (2021).

[29] N. N. T. Nam and M. Koshino, Lattice relaxation and energy band modulation in twisted bilayer graphene, Phys. Rev. B **96**, 075311 (2017).

[30] S. Carr, D. Massatt, S. B. Torrisi, P. Cazeaux, M. Luskin, and E. Kaxiras, Relaxation and domain formation in incommensurate two-dimensional heterostructures, Phys. Rev. B **98**, 224102 (2018).

[31] P. Cazeaux, M. Luskin, and D. Massatt, Energy minimization of two dimensional incommensurate heterostructures, Arch. Ration. Mech. Anal. **235**, 1289 (2020).