Moiré-pattern fluctuations and electron-phason coupling in twisted bilayer graphene

Héctor Ochoa

Moiré-pattern fluctuations and electron-phason coupling in twisted bilayer graphene

Héctor Ochoa

Department of Physics, Columbia University, New York, NY 10027, USA

In twisted bilayer graphene, long-wavelength lattice fluctuations on the scale of the moiré period are dominated by phason modes, i.e., acoustic branches of the incommensurate lattice resulting from coherent superpositions of optical phonons. In the limit of small twist angles, these modes describe the sliding motion of stacking domain walls separating regions of partial commensuration. The resulting soliton network is a soft elastic manifold, whose reduced rigidity arises from the competition between intralayer (elastic) and interlayer (adhesion) forces governing lattice relaxation. Shear deformations of the beating pattern dominate the electron-phason coupling to the leading order in $t_\perp/t$, the ratio between interlayer and intralayer hopping parameters. This coupling lifts the layer degeneracy of the Dirac cones at the corners of the moiré Brillouin zone, which could explain the observed 4-fold (instead of 8-fold) Landau level degeneracy. Electron-phason scattering gives rise to a linear-in-temperature contribution to the resistivity that increases with decreasing twist angle due to the reduction of the stiffness of the soliton network. This contribution alone, however, seems to be insufficient to explain the huge enhancement of the resistivity of the normal state close to the magic angle.

I. INTRODUCTION

Two graphene layers rotated with respect to each other an incommensurate angle θ form a quasi-periodic structure known as moiré pattern.[1] At a *magic angle* $\theta \sim 1°$, quantum interference of electrons in the associated superlattice potential gives rise to narrow low-energy bands,[2] setting the stage for strongly correlated phenomena.[3,4] In addition to what appears to be Mott insulating states at half-filling of the lowest flat bands[3] and the onset of superconductivity under doping[4] or hydrostatic pressure,[5] new many-body insulating states have been reported for different filling factors, some of them with apparent topological character.[6–8]

Qualitative differences in phenomenology from one device to another suggest that we should envision the moiré pattern not as a rigid potential landscape but as a spatially inhomogeneous and (as I will argue here) most likely fluctuating one. Structural inhomogeneities around the magic angle are revealed by variations in the electronic densities of full-filled (single-particle) insulating states, along with evidences from quantum interference measurements of the formation of insulating islands within the superconducting state.[4,5] For smaller twist angles, spatial variations of the moiré period have been directly visualized by transmission electron[9] and scanning tunnel microscopies[10] as well as near-field optical techniques,[11] showing the formation of regions of partial commensuration separated by stacking domain walls.[12,13] These structural differences can induce/favor different symmetry-broken states. The interaction with the encapsulating boron nitride in transport devices can break the sublattice symmetry (usually referred as $\mathcal{C}_2\mathcal{T}$ symmetry in the literature[14,15]), opening a gap in the low-energy bands, which acquire nonzero valley-Chern numbers. This state could serve as a precursor for the formation of a (quantum) anomalous Hall ferromagnet at odd fillings as the bands become spin-valley polarized due to electron correlations.[6,16,17] Mean-field calculations suggest that this symmetry can also be spontaneously broken,[18] which has been invoked to explain some features in the magnetotransport of the most homogeneous samples at one-quarter filling of the hole band.[7] The electronic spectrum is also sensitive to the dielectric environment due to electrostatic effects.[19] In open samples, suitable for local probes, tunneling spectroscopy has unveiled the emergence of charge ordering with broken 3-fold symmetry as the Fermi level varies.[20–22]

Another striking observation is the remarkably large, linear in temperature (T) resistance of the normal state.[23,24] Here there appear to be subtle difference between samples too: While in the devices of Cao *et al.*[23] a linear-T resistivity is only apparent for fractional fillings (more prominently at half-filling), suggesting a connection with the correlated state at lower temperatures, the devices of Polshyn *et al.*[24] show qualitatively the same behavior (albeit non-monotonic) for a broader range of carrier densities, with increasing values of the resistivity as the twist angle decreases. This latter behavior has been attributed to electron-phonon scattering.[25–27] These studies focus on the original acoustic modes of graphene, but as noticed in Ref. 28, relative displacements of the two layers have a stronger impact on the moiré interference pattern and, potentially, on the electronic structure. Here I analyze the case with the account of interlayer interaction forces, which play a fundamental role in the energetics of these modes.

The structure of the article is as follows. I will discuss first in Sec. II the emergence of new collective modes, *phasons*, associated with the broken translational symmetry of the incommensurate lattice, or more accurately, the invariance of the system under relative translations of the two layers. These modes define two acoustic branches in the spectrum of small oscillations, which is determined by the competition of the two relevant length scales in the problem: the moiré period, and a length scale related to the curvature of the adhesion potential that define the characteristic width of stacking textures (solitons)

connecting degenerate adhesion-energy minima. Already for angles $\theta \lesssim 3°$, phonon softening is very pronounced, marking the instability of the system towards the formation of a soliton network. Phason dynamics is then governed by the effective elasticity of these objects. In Sec. III, I will analyze how phasons couple to electrons. Transverse modes (i.e., shear and rotational deformations of the interference pattern) have a stronger impact on the low-energy spectrum and, in particular, can explain the reduced Landau level degeneracy observed in magnetotransport.[3–5] Although the coupling considered here is intrinsically weak (proportional to the ratio between interlayer and intralayer hopping parameters), scattering by large (due to the reduced rigidity of stacking solitons) phason fluctuations limit the electron mobilities, leading to a linear-T resistivity that increases with decreasing twist angle. However, this contribution seems to be insufficient to explain the dramatic increase of the resistivity around the magic angle reported in the experiments. I will finally discuss other possible scenarios in Sec. IV. Details of the calculations are saved for the appendices.

II. PHASONS IN MOIRÉ-PATTERNED BILAYER GRAPHENE

Let me first introduce the geometry of the moiré superlattice. In the absence of lattice relaxation, atomic positions in the bottom and top layers are spanned by primitive vectors $\mathbf{a}_{1,2}$ and $\mathbf{a}'_{1,2} = R(\theta)\,\mathbf{a}_{1,2}$ of a triangular Bravais lattice, where $R(\theta)$ is a SO(2) matrix describing the relative rotation. For concreteness, I am going to consider the situation depicted in Fig. 1, in which carbon atoms sit on top of each other (AA stacking) prior to the relative rotation. This choice is arbitrary but ultimately inconsequential in the limit of small (generically incommensurate) twist angles, for in that case the system explores all possible stacking configurations. For future reference, lateral positions will be referred to the rotation axis; the coordinate system is such that the x and y axes lie along the two-fold symmetry axes highlighted in Fig. 1 a. Positive twist angles will correspond to anticlockwise rotations of the top layer.

The moiré superlattice is defined by a beating pattern resulting from the periodicity of the individual triangular lattices, i.e., the Fourier components of the atomic density $\rho(\mathbf{r})$ are peaked at vectors $\mathbf{G}$ of the incommensurate reciprocal lattice,

$$\rho(\mathbf{r}) \approx \sum_{\{\mathbf{G}\}} \rho_{\mathbf{G}}\, e^{i\mathbf{G}\cdot\mathbf{r}} = \sum_{\{\mathbf{G}\}} |\rho_{\mathbf{G}}|\, e^{-i\phi_{\mathbf{G}}+i\mathbf{G}\cdot\mathbf{r}}, \quad (1)$$

where in the last expression I have separated the Fourier components in modulus and phase (note that $\phi_{-\mathbf{G}} = -\phi_{\mathbf{G}}$, so $\rho(\mathbf{r})$ is real). The set $\{\mathbf{G}\}$ corresponds to the lattice spanned by primitive vectors $\mathbf{G}_{1,2} = \mathbf{b}_{1,2} - \mathbf{b}'_{1,2}$, where $\mathbf{b}_{1,2}$ and $\mathbf{b}'_{1,2} = R(\theta)\,\mathbf{b}_{1,2}$ are primitive vectors of the bottom and top reciprocal lattices, respectively (see
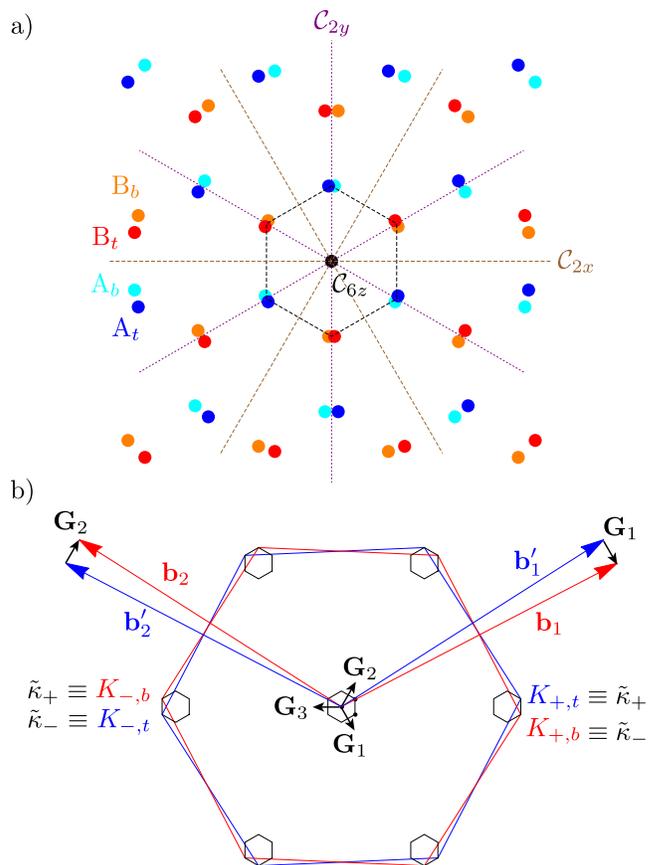

FIG. 1: a) Microscopic lattice (no relaxation) around a beating pattern maximum (local AA stacking) for a twist angle of $\theta = 5°$. Dashed lines highlight the hexagonal (D_6) symmetry of the continuum model, with a 6-fold rotation axis along the common center and 2-fold rotation axis within the plane. b) Superimposed Brillouin zones of top (in blue) and bottom (in red) layers. In commensurate approximants, valleys $K_{\tau,\mu}$ lie at the two inequivalent corners of the moiré Brillouin zone, $\tilde{\kappa}_\eta$; hereafter $\tau = \pm 1$ labels the valley, $\mu = \pm 1\,(t/b)$ labels the top (t) and bottom (b) layers, and $\eta = \tau \times \mu$ labels the corresponding $\tilde{\kappa}_\eta$ point in type-I or sublattice-exchange odd commensurate structures,[15,29,30] which are dense in the limit of small twist angles.

Fig. 1 b). The moiré superlattice is just the dual to $\{\mathbf{G}\}$, spanned by primitive vectors[31]

$$\mathbf{R}_{1,2} = \left[1 - R^{-1}(\theta)\right]^{-1} \mathbf{a}_{1,2}. \quad (2)$$

It is also convenient to introduce the function $\boldsymbol{\Delta}(\mathbf{r}) \equiv \mathbf{r} - R^{-1}(\theta)\,\mathbf{r}$, which measures the distance between a given point $\mathbf{r}$ in the top layer with respect to its original position in the bottom layer before the twist (assuming yet no lattice relaxation). Note that at the maxima (local AA stacking) of the interference pattern $\boldsymbol{\Delta}$ coincides with a lattice vector of the graphene Bravais lattice, $\boldsymbol{\Delta}(n\,\mathbf{R}_1 + m\,\mathbf{R}_2) = n\,\mathbf{a}_1 + m\,\mathbf{a}_2$, with n, m integers. The beating pattern maxima are separated by

a distance $L_M = |\mathbf{R}_{1,2}| = \sqrt{3}a/(2\sin\theta/2)$, defining the moiré superlattice period.

So far, interlayer forces have not appeared in our discussion. Lattice relaxation of structures with a nominal twist angle ($|\rho_{\mathbf{G}}| \neq 0$) could be generically described by a Landau-like free-energy expansion in powers of the beating-pattern density which, as inferred from the previous discussion, gives us an idea of the degree of overlap in the lateral position of the atoms in both layers (and therefore of the interlayer coupling). Minimization of this functional would lead to constraints of the form $\sum_{i=1...n} \phi_{\mathbf{G}_i} = \gamma$, with $\sum_{i=1...n} \mathbf{G}_i = 0$, coming from terms in the nth power of $\rho(\mathbf{r})$. The reader should note that such *relaxed structures* are only metastable states that, in real devices, are likely to be stabilized by the unavoidable tensions generated during the fabrication process. In fact, it is likely that the samples present patches with different twist angles, i.e., different period of the moiré patterning. Here we are concerned only about long-wavelength fluctuations (smooth on the scale of the moiré period L_M) around a local minimum.

The previous constraints defining the configuration of minimum energy leave a certain number N of the phases $\phi_{\mathbf{G}_i}$ unspecified; in other words, variations of N phases generate physically distinct but energetically equivalent quasi-periodic structures and, thus, represent soft modes of the system. In this case, the number of soft modes is $N = 2$, for $\{\mathbf{G}\}$ is a two-dimensional Bravais lattice. Consider, for example, a beating pattern truncated to the first six Fourier harmonics (first star), which set the sixfold symmetry and the single independent length scale of the moiré superlattice: $\rho(\mathbf{r}) \approx \rho_0 + \rho_1 \sum_{i=1,2,3} \cos(\mathbf{G}_i \cdot \mathbf{r} - \phi_{\mathbf{G}_i})$, with $\mathbf{G}_3 \equiv -\mathbf{G}_1 - \mathbf{G}_2$. A generic cubic term in the phenomenological free energy fix one of the three phases. A convenient parametrization is then of the form

$$\phi_{\mathbf{G}_i} = \tilde{\mathbf{u}} \cdot \mathbf{G}_i + \frac{\gamma}{3}, \tag{3}$$

where $\tilde{\mathbf{u}}$ is a two-dimensional vector describing the *phason* modes of the incommensurate lattice.[32] Note that changes in $\tilde{\mathbf{u}}$ translate rigidly the beating pattern, while an arbitrary change in γ (which is not a soft mode) distorts the pattern. In a totally incommensurate or *floating* state, i.e., when lattice relaxation is ignored, $\tilde{\mathbf{u}}$ can be identified straightforwardly with a relative displacement between layers, $\mathbf{u}$. Note that in that case $\mathbf{\Delta}(\mathbf{r}) \to \mathbf{r} - R^{-1}(\theta)(\mathbf{r} - \mathbf{u}) = \mathbf{\Delta}(\mathbf{r}) + R^{-1}(\theta)\,\mathbf{u}$, therefore, the maxima of the beating pattern are translated a distance[33]

$$\tilde{\mathbf{u}} = [1 - R(\theta)]^{-1}\,\mathbf{u} = \left(\frac{1}{2} + \frac{1}{2}\cot\frac{\theta}{2}\,\hat{\mathbf{z}}\times\right)\mathbf{u}. \tag{4}$$

More generically, phason modes correspond to coherent superpositions of optical phonons with momenta separated by a superlattice vector $\mathbf{G}$. When lattice relaxation is taken into account, these coherent superpositions describe more complex atomic rearrangements than a simple relative displacement of the two layers. The coordinate $\tilde{\mathbf{u}}$ can be identified then with the sliding motion of domain walls separating regions of partial commensuration (alternating AB and energetically equivalent BA Bernal stackings, where atoms of different sublattices lie on top of each other).

A. Soliton network

The problem of lattice relaxation can be addressed in the framework of a two-dimensional version of the Frenkel-Kontorova model.[34–36] Neglecting entropic terms in the free energy, the problem reduces to solve the following equation for $\mathbf{u}(\mathbf{r})$, understood now as a field in the continuum describing smooth (in the scale of the interatomic distance $a \approx 1.42$ Å) relative displacements of the layers:

$$\frac{\lambda+\mu}{2}\,\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{u}) + \frac{\mu}{2}\,\nabla^2\mathbf{u} = \frac{\partial}{\partial\mathbf{u}}\mathcal{V}_{\text{ad}}[\mathbf{r},\mathbf{u}(\mathbf{r})]. \tag{5}$$

Here $\mu \approx 3\lambda \approx 9$ eV/Å^2 are the Lamé coefficients of graphene;[37] I am disregarding corrugations provided that the bending energy is inconsequential on length scales longer than $\sqrt{\kappa/(\lambda+2\mu)} \sim a$, where $\kappa \approx 1$ eV is the bending rigidity.

The right-hand side of Eq. (5) describes variations in configurational space (stackings) of the adhesion potential (here with units of energy density) between layers. The notation emphasizes the different periodicities in real and configurational spaces: while the dependence on $\mathbf{r}$ is modulated on the scale of the moiré pattern, i.e., it admits a Fourier expansion in $\{\mathbf{G}\}$, the dependence on stacking configurations changes on the atomic unit cell. A first-star expansion of the latter compatible with sixfold symmetry reduces to

$$\mathcal{V}_{\text{ad}}[\mathbf{u}(\mathbf{r})] = V\sum_{i=1}^{3}\left\{\frac{1}{2} + \cos[\mathbf{b}_i\cdot(\mathbf{u}+\mathbf{\Delta}_0)]\right\}, \tag{6}$$

where $\mathbf{\Delta}_0$ is the displacement of the top layer with respect to the bottom layer in Bernal (AB) stacking, which sets the reference in energies. The phenomenological parameter $V \approx 90$ meV/nm^2 measures the energy difference between Bernal and AA stackings.[38] The adhesion-energy landscape deduced from this model is shown in Fig. 2 a. In this approximation, the modulation in real space follows from Eq. (6) just by noting that when layers are rotated, the separation between unit cells depends on the position $\mathbf{r}$ via the substitution $\mathbf{\Delta}_0 \to \mathbf{\Delta}(\mathbf{r})$, and $\mathbf{b}_i\cdot\mathbf{\Delta}(\mathbf{r}) = \mathbf{G}_i\cdot\mathbf{r}$ from the previous definitions.

In addition to the periodicity of the moiré pattern, there is another length scale encrypted in Eq. (5) related to the curvature of the adhesion potential,

$$\ell = \sqrt{\frac{\mu}{2\frac{\partial^2 \mathcal{V}_{\text{ad}}}{\partial u^2}\big|_{\text{AB}}}} = \frac{a}{\pi}\sqrt{\frac{\mu}{2V}} \approx 3.2 \text{ nm}. \tag{7}$$
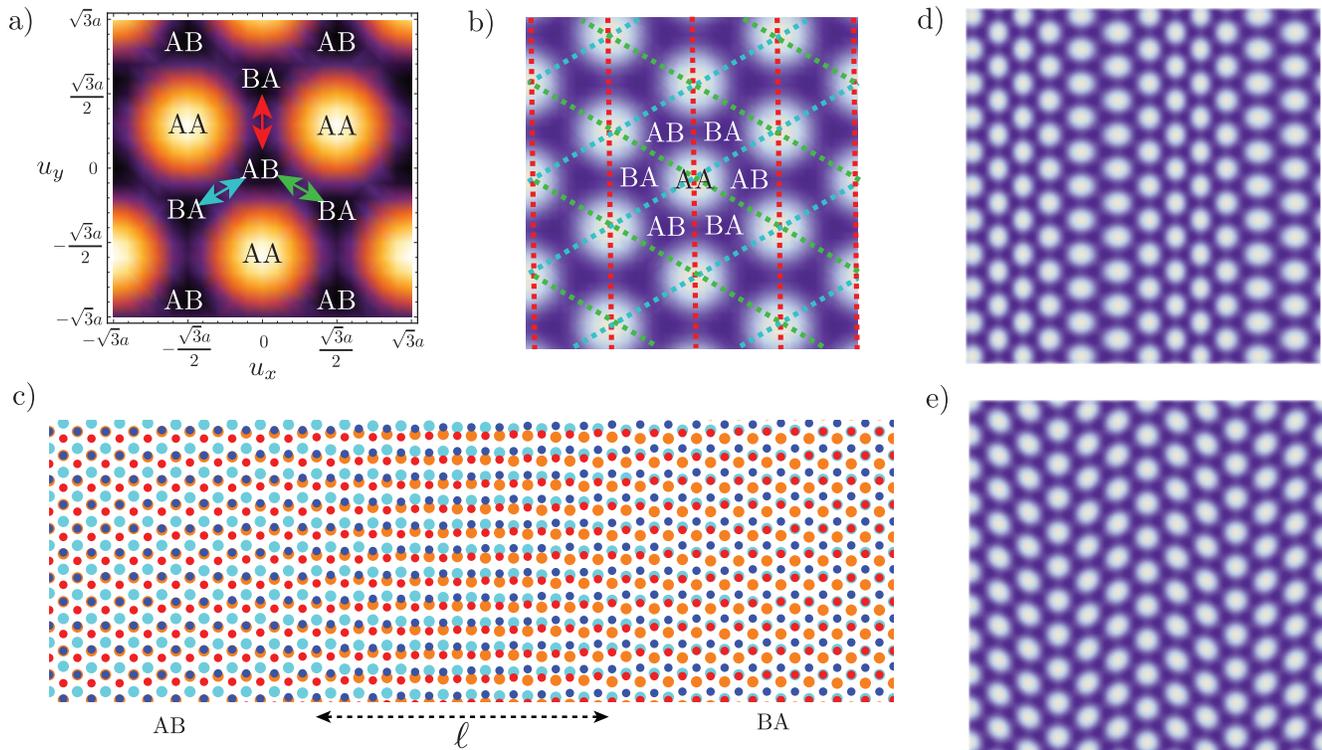
FIG. 2: a) Adhesion-energy landscape deduced from Eq. (6). Arrows represent stacking textures connecting degenerate minima (AB and BA Bernal stackings). b) Generic beating pattern (Eq. 1) formed by two floating layers. When ℓ is ostensibly smaller than L_M, sharper stacking textures are formed, corresponding to domain walls between regions of partial commensuration. The domain walls are represented as dashed lines in the same colors as their representation in configurational space, panel a. c) Sine-Gordon shear soliton connecting AB and BA stacking regions. Colored dots represent atomic positions as in Fig. 1. d) Longitudinal phason distortion of the beating pattern with wavelength $\lambda = 3\sqrt{3}L_M$. e) Transverse phason with the same wavelength.

This length scale characterizes the spatial extension of stacking textures connecting equivalent minima (AB and BA stackings). The competition between their cost in elastic energy and the adhesion energy of large incommensurate areas governs the degree of lattice relaxation. When L_M and ℓ are comparable, lattice relaxation is negligible and the beating pattern is well approximated by two floating layers. However, when L_M is larger than ℓ, it is energetically cheaper for the system to form regions of partial commensuration separated by domain walls of characteristic width ℓ, where the cost in elastic and adhesion energies is concentrated.

The competition between these two length scales is reflected in the the spectrum of small oscillations around a local minimum, see Appendix A. The harmonic expansion of the adhesion potential around a floating state only couples modes with momentum mismatch in the first star, originally separated by a frequency of the order of $\omega_M = 4\pi c/(\sqrt{3}L_M)$, where $c = \sqrt{\mu/\rho_0} \approx 14$ Km/s is the sound velocity of transverse phonons in graphene. When $L_M \sim \ell$, the strength of these harmonics is negligible with respect to ω_M, and the spectrum resembles

the acoustic phonons of single-layer graphene folded into the moiré Brillouin zone. When L_M is ostensibly larger than ℓ, this simple model predicts a strong softening of the two acoustic branches, indicating that the floating state is unstable and lattice relaxation is not longer negligible. Sharper stacking textures are formed,[36] giving rise to more harmonics in the expansion of the adhesion potential that have to be included in a new calculation of the spectrum of oscillations. Already in this regime, the simple identification in Eq. (4) breaks down and the energetics of phasons are no longer described by the elasticity of individual layers.

The first step is to determine the stacking texture. In order to construct approximated long-wavelength density profiles, let me consider first one-dimensional solutions of Eq. (5) of the form $\mathbf{u}(\mathbf{r}) = u(\varrho)\,\hat{\mathbf{u}}$, where the spatial dependence (along unit vector $\hat{\boldsymbol{\varrho}}$) is not necessarily collinear with lattice relaxation. A simple inspection of Eq. (6) (plugging the one-dimensional ansatz and projecting over $\hat{\mathbf{u}}$ and $\hat{\mathbf{z}} \times \hat{\mathbf{u}}$) shows that only *tensile* ($\hat{\mathbf{u}} \parallel \hat{\boldsymbol{\varrho}}$) or *shear* ($\hat{\mathbf{u}} \perp \hat{\boldsymbol{\varrho}}$) solutions exist in this approximation.[39] I am going to focus on the latter, which are less energetic;

this assumption agrees with numerical calculations[36,38] (note also that the softening is more pronounced for the transverse oscillation mode) and is obviously favored by the orientation of the moiré superlattice with respect to the graphene lattice in the limit of small twist angles. As implied by the energy landscape in Figs. 2 a, the relaxation is more likely to occur along an armchair direction. Let me focus first on the case $\hat{\mathbf{u}}_3 \approx \hat{\mathbf{y}}$, $\hat{\boldsymbol{\varrho}}_3 \approx \hat{\mathbf{x}}$ (solitons marked in red in the figure); the other two solutions with the same energy follow from rotating the axes 120°. The notation emphasizes that $\mathbf{G}_i \approx -\frac{2\pi}{L}\hat{\boldsymbol{\varrho}}_i$ for small twist angles, with $L = |\hat{\mathbf{x}} \cdot \mathbf{R}_{1,2}| \approx \sqrt{3}L_M/2$.

In the asymptotic limit $L_M \gg \ell$, we can neglect the modulation on the scale of the moiré (first argument in the adhesion potential) and write $\mu\, u'' = (V\pi/a)\sin(2\pi u/a)$, whose general solution is a train of domain walls separated a certain fixed length. The latter should be determined from energetic considerations,[40] but in our heuristic construction we can just identify this length scale with L. In this limit, the domain-wall profile is well described by a sine-Gordon soliton,

$$u_3(x) \approx \frac{2a}{\pi}\arctan\left(e^{\frac{x-x_3}{\ell}}\right), \qquad (8)$$

as numerical calculations confirm.[36] The corresponding stacking texture is represented in Fig. 2 c. Associated with the domain wall, there is a tension $\sigma = a\sqrt{2\mu V}/\pi$ characterizing its energy cost. This energy does not depend on the soliton center, x_3: while the separation between domain walls is set by the moiré periodicity, rigid displacements of the soliton solution does not cost energy.

The density wave associated with the one-dimensional train of domain walls is of the form

$$\rho_{1D}(\mathbf{r}) = \sum_{n=0}^{\infty}|\rho_n|\cos\left[\frac{2\pi n}{L}(x-x_3)\right] \qquad (9)$$
$$\approx \sum_{n=0}^{\infty}|\rho_n|\cos\left[n\left(\hat{\boldsymbol{\varrho}}_3 \cdot \mathbf{G}_3\right)(x-x_3)\right],$$

where the Fourier components $|\rho_n|$ can be estimated from the approximate profile in Eq. (8). The two-dimensional beating pattern can be approximated then by the superposition of three density waves like Eq. (9) with director

vectors rotated 120°. At *rigid* soliton crossings, the system explores all possible stacking configurations. Adhesion forces will relax the structure, but some areas will remain pinned to saddle points of the potential, including AA stackings, introducing a large free-energy cost. Therefore, the configuration of minimum energy is such that the number of soliton crossings is minimized, and consequently, from the three phases related to the soliton centers, $x_{1,2,3}$, only two are really independent. In the parametrization of Eq. (3), the soliton centers are related to the phason field as

$$x_i \approx \hat{\boldsymbol{\varrho}}_i \cdot \tilde{\mathbf{u}} - \frac{\gamma L}{6\pi}. \qquad (10)$$

Minimizing the number of soliton crossings corresponds to the condition $x_0 + x_1 + x_2 \approx 0$ (modulo $L/3$), or equivalently, γ must be an integer of 2π.[41] The resulting beating pattern can be envisioned as a triangular lattice of AA-stacked regions connected by sine-Gordon domain walls, as illustrated in Fig. 2 b.

B. Generalized elasticity

Provided that ℓ is smooth on the interatomic distance, the free-energy cost of long-wavelength fluctuations of the beating pattern can be described by a phenomenological expansion in terms of derivatives of the phason field, $\partial_i \tilde{u}_j$, constrained only by the symmetries of the continuum model. Since the superlattice vectors imposes a preferential orientation of the beating pattern (i.e., only translations and not rotations are soft[42]), introducing a symmetric strain tensor is not sensible for this problem. Instead, $\partial_i \tilde{u}_j$ can be arranged in irreducible representations of D_6 (see Fig. 1),

$$\partial_x \tilde{u}_x + \partial_y \tilde{u}_y \sim A_1, \qquad (11a)$$
$$\partial_x \tilde{u}_y - \partial_y \tilde{u}_x \sim A_2, \qquad (11b)$$
$$\begin{bmatrix} \partial_x \tilde{u}_x - \partial_y \tilde{u}_y \\ -\partial_x \tilde{u}_y - \partial_y \tilde{u}_x \end{bmatrix} \sim E_2, \qquad (11c)$$

representing compressional, rotational (or tilting), and shear deformations of the soliton network, respectively. A generic harmonic expansion reads then

$$\mathcal{F}[\tilde{\mathbf{u}}(\mathbf{r})] = \frac{1}{2}\int d\mathbf{r}\left[\tilde{\lambda}\left(\boldsymbol{\nabla}\cdot\tilde{\mathbf{u}}\right)^2 + \frac{\tilde{\mu}}{2}\left(\partial_i\tilde{u}_j + \partial_j\tilde{u}_i\right)^2 + \gamma\left(\boldsymbol{\nabla}\times\tilde{\mathbf{u}}\right)^2\right], \qquad (12)$$

where the cost of compressional and shear deformations is expressed in terms of new Lamé coefficients and γ is the tilt modulus accounting for the cost of rotations with respect to the moiré superlattice. The elastic constants can be estimated from the approximated profile in Eq. (8)

as (see Appendix A)

$$\tilde{\lambda}(\theta) \approx \frac{(1+\nu)\,\ell\,V\sin\left(\frac{\theta}{2}\right)}{a\,(\nu-1)}, \qquad (13a)$$

$$\tilde{\mu}(\theta) \approx \frac{(3-\nu)\,\ell\,V\sin\left(\frac{\theta}{2}\right)}{a\,(1-\nu)}, \qquad (13b)$$

$$\gamma(\theta) \approx \frac{4\ell\,V\sin\left(\frac{\theta}{2}\right)}{a\,(1-\nu)}, \qquad (13c)$$

where $\nu \approx 0.2$ is graphene's Poisson ratio.

The soliton network appears to have, within the limitations of the present calculation, a negative Poisson's ratio $(\tilde{\lambda} < 0)$.[43] This is not so surprising for a two-dimensional incommensurate structure,[44] and manifests the tendency of the resulting beating pattern to preserve the six-fold symmetry of the moiré superlattice. Here I should emphasize that a relative compression/expansion of the layers (like the uniaxial hetero-strain considered in Ref. 45) introduces a shear deformation of the beating pattern, and viceversa, relative shear between layers (which is energetically cheaper) introduces a longitudinal distortion. The latter can be interpreted as a modulation of the moiré period or a nonuniform twist angle, which is systematically observed in topographic images acquired by scanning tunneling microscopy.[20–22] The lower energy for longitudinal deformations is ultimately ascribed to the fact that the soliton network connecting AA stacked regions must be understood as a system of strings under tension, whose energy scales linearly (instead of quadratically, like in a system of springs) with length. The tension of the domain walls manifest the metastability of bilayers with a nominal twist angle. These structures can be further stabilized by entropic terms in the free energy, not included in the purely mechanical model discussed here. Two-dimensional soliton configurations carry a lot of entropy,[40,46] which is also associated with the fact that their energy scales linearly with length. Consequently, thermal fluctuations can contribute to renormalize the elastic constants.[44,47] The existence of patches or domains with a nominal twist angle can also be interpreted as the fact that the transition from commensurate to incommensurate structures is first order and dominated by entropy.

Longitudinal and transverse phason modes are represented in Fig. 2 d and e, respectively. Their frequencies deduced from the harmonic expansion in Eq. (12) reads

$$\omega_{\mathbf{q}}^{(L)} = \sqrt{\frac{\tilde{\lambda} + 2\tilde{\mu}}{\tilde{\rho}}}\,|\mathbf{q}| \approx \sqrt{\frac{5 - 3\nu}{4\,(1 - \nu)}}\,c\,|\mathbf{q}|, \qquad (14a)$$

$$\omega_{\mathbf{q}}^{(T)} = \sqrt{\frac{\tilde{\mu} + \gamma}{\tilde{\rho}}}\,|\mathbf{q}| \approx \sqrt{\frac{7 - \nu}{4\,(1 - \nu)}}\,c\,|\mathbf{q}|, \qquad (14b)$$

where I have introduced in the mass density of the soliton network (see Appendix A):

$$\tilde{\rho}\,(\theta) \approx \rho\,\sqrt{\frac{8V}{\pi^2 \mu}}\,\sin\left(\frac{\theta}{2}\right). \qquad (15)$$

The inertia of the soliton system is reduced because the formation of sharper (on the scale of L_M) stacking textures implies that, effectively, a smaller fraction of the atoms within the moiré cell takes part of the sliding motion of one layer with respect to the other. This compensates the reductions of the elastic constants, so the dispersion of the phason modes do not change much with respect to the acoustic phonons of individual graphene

layers, in agreement with numerical calculations.[48] This does not imply, as we have seen, that interlayer adhesion forces are negligible, or that the elasticity of individual layers can describe the energetics of phason modes.

III. ELECTRON-PHASON COUPLING

In order to evaluate the effect of phason fluctuations on the electronic spectrum and transport properties, I am going to consider the continuum model usually discussed in the literature:[1,2]

$$\hat{\mathcal{H}}^{(\tau)} = \begin{pmatrix} \hat{\mathcal{H}}_D^{(\tau,t)} & \hat{T}^{(\tau)}\,(\mathbf{r}) \\ \left[\hat{T}^{(\tau)}\,(\mathbf{r})\right]^{\dagger} & \hat{\mathcal{H}}_D^{(\tau,b)} \end{pmatrix}. \qquad (16)$$

The Hamiltonian is written in block form, each of them acting on a spinor wave function in a given valley sector (labelled by $\tau = \pm 1$) from top and bottom layers (upper and lower blocks, respectively) describing electronic states around points $K_{\tau,\mu}$ in Fig. 1. The block-diagonal terms are Dirac Hamiltonians of the form

$$\hat{\mathcal{H}}_D^{(\tau,\mu)} = \hbar v_F\,\mathbf{\Sigma}^{(\tau,\mu)} \cdot \left(\hat{\mathbf{k}} - \mathbf{K}_{\tau,\mu}\right), \qquad (17)$$

where $\hat{\mathbf{k}} = -i\boldsymbol{\partial}$ is the crystalline-momentum operator in real-space representation, $\hbar v_F = 3ta/2$, with $t \approx 2.8$ eV being the intralayer hopping parameter, and $\mathbf{\Sigma}^{(\tau,\mu)}$ is a vector of Pauli matrices defined in the spinor (sublattice) space. Since crystalline momentum is expressed in a common frame of reference (defined by $\mathcal{C}_{2x}$ and $\mathcal{C}_{2y}$ axes in Fig. 1) the Pauli matrices have to be properly rotated,

$$\mathbf{\Sigma}^{(\tau,\mu)} = \left(\tau\,e^{\frac{i\mu\theta\hat{\ell}_z}{2}}\hat{\sigma}_x e^{-\frac{i\mu\theta\hat{\ell}_z}{2}}, e^{\frac{i\mu\theta\hat{\ell}_z}{2}}\hat{\sigma}_y e^{-\frac{i\mu\theta\hat{\ell}_z}{2}}\right), \quad (18)$$

where $\hat{\ell}_z = \tau\hat{\sigma}_z/2$ is the generator of spinor rotations (see Appendix B) and $\mu = \pm 1$ for top/bottom layer blocks. Interlayer tunneling processes are described by the off-diagonal blocks, given by

$$\hat{T}^{(\tau)}\,(\mathbf{r}) = t_\perp\left[\hat{T}_0^{(\tau)} + \hat{T}_1^{(\tau)}e^{-i\tau\mathbf{G}_1\cdot(\mathbf{r}-\tilde{\mathbf{u}})} + \hat{T}_2^{(\tau)}e^{i\tau\mathbf{G}_2\cdot(\mathbf{r}-\tilde{\mathbf{u}})}\right], \qquad (19)$$

where $t_\perp \approx 110$ meV parametrizes the strength of the interlayer hopping; matrices $\hat{T}_i^{(\tau)}$ contain the pertinent phases acquired by the wave function when electrons hop between different sublattices:

$$\hat{T}_0^{(\tau)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad (20a)$$

$$\hat{T}_1^{(\tau)} = \begin{pmatrix} 1 & e^{i\tau\frac{2\pi}{3}} \\ e^{-i\tau\frac{2\pi}{3}} & 1 \end{pmatrix}, \qquad (20b)$$

$$\hat{T}_2^{(\tau)} = \begin{pmatrix} 1 & e^{-i\tau\frac{2\pi}{3}} \\ e^{i\tau\frac{2\pi}{3}} & 1 \end{pmatrix}. \qquad (20c)$$

In addition to time-reversal and superlattice-translation symmetries, along with valley conservation imposed by construction, the continuum model is also invariant under D_6 point-group operations[14,15] and infinitesimal *rigid* translations of one layer with respect to the other: a uniform phason field $\tilde{\mathbf{u}}$ parametrizing the center of the beating pattern can be absorbed in a unitary rotation of the wave function (specifically, it can be absorbed as a phase in the Bloch states defined in different copies of the moiré Brillouin zone), so the mini-band spectrum remains invariant. Note also that the model is derived assuming no lattice relaxation (the main steps are reproduced in Appendix B following Ref. 49). Its effect can be incorporated by means of i) strain fields within the Dirac-Hamiltonian blocks

and ii) more harmonics in $\hat{T}(\mathbf{r})$ along with smaller amplitudes for interlayer hoppings between the same sublattice, reflecting the shrinking of AA stacked areas. Although these terms have an important effect on the electronic spectrum,[36,50,51] I am going to neglect them in my estimation of the electron-phason coupling; the observation is that, regardless of the changes in the Hamiltonian, the invariance under relative translations of the two layers imposes the parametric dependence on $\tilde{\mathbf{u}}$ already contained in Eq. (19). Thus, we can obtain the leading contribution in $t_\perp$ by expanding Eq. (19) (and its hermitic conjugate) up to linear order in the smooth (on the scale of the moiré period) nonuniform phason field, just like in the case of a charge-density wave,[52]

$$\delta\hat{T}^{(\tau)}(\mathbf{r}) \approx i\,\tau\,t_\perp\tilde{\mathbf{u}}(\mathbf{r}) \cdot \left[\mathbf{G}_1\,\hat{T}_1^{(\tau)}e^{-i\tau\mathbf{G}_1\cdot\mathbf{r}} - \mathbf{G}_2\,\hat{T}_2^{(\tau)}e^{i\tau\mathbf{G}_2\cdot\mathbf{r}}\right]. \tag{21}$$

Introducing Fourier series for the phason field, decomposing the Fourier components in longitudinal and transverse components, and promoting the latter to boson operators in conventional fashion, provided that $\boldsymbol{\pi} = \tilde{\rho}\,\dot{\tilde{\boldsymbol{u}}}$ is the canonical momentum density conjugate to soliton displacements, I arrive at the following general expression for the electron-phason coupling in second quantization,

$$\hat{H}_{\text{e-ph}} = -\sum_{\tau,\mathbf{k}',\mathbf{k},\mathbf{q}} \tau\,t_\perp \left\{ \sqrt{\frac{\hbar}{2A\tilde{\rho}\omega_{\mathbf{q}}^{(L)}}}\,\hat{A}_{\mathbf{q}}^{(L)} \left[\frac{\mathbf{q}\cdot\mathbf{G}_1}{|\mathbf{q}|}\left(\Psi_{\mathbf{k}'-\tau\mathbf{G}_1}^{(\tau,t)}\right)^\dagger \hat{T}_1^{(\tau)}\Psi_{\mathbf{k}}^{(\tau,b)} - \frac{\mathbf{q}\cdot\mathbf{G}_2}{|\mathbf{q}|}\left(\Psi_{\mathbf{k}'+\tau\mathbf{G}_2}^{(\tau,t)}\right)^\dagger \hat{T}_2^{(\tau)}\Psi_{\mathbf{k}}^{(\tau,b)}\right] \right. \tag{22}$$
$$\left. + \sqrt{\frac{\hbar}{2A\tilde{\rho}\omega_{\mathbf{q}}^{(T)}}}\,\hat{A}_{\mathbf{q}}^{(T)} \left[\frac{(\mathbf{q}\times\mathbf{G}_1)_z}{|\mathbf{q}|}\left(\Psi_{\mathbf{k}'-\tau\mathbf{G}_1}^{(\tau,t)}\right)^\dagger \hat{T}_1^{(\tau)}\Psi_{\mathbf{k}}^{(\tau,b)} - \frac{(\mathbf{q}\times\mathbf{G}_2)_z}{|\mathbf{q}|}\left(\Psi_{\mathbf{k}'+\tau\mathbf{G}_2}^{(\tau,t)}\right)^\dagger \hat{T}_2^{(\tau)}\Psi_{\mathbf{k}}^{(\tau,b)}\right] \right\} \delta_{\mathbf{k}',\mathbf{k}+\mathbf{q}} + \text{h.c.},$$

where A is the area of the system, $\hat{A}_{\mathbf{q}}^{(\nu)} = a_{\mathbf{q}}^\nu + (a_{-\mathbf{q}}^\nu)^\dagger$, and $\Psi_{\mathbf{k}}^{(\tau,\mu)} = (c_{A,\mathbf{k}}^{\tau,\mu}, c_{B,\mathbf{k}}^{\tau,\mu})^T$; operators $(c_{\alpha,\mathbf{k}}^{\tau,\mu})^\dagger/c_{\alpha,\mathbf{k}}^{\tau,\mu}$ create/annihilate Bloch states with crystalline momentum $\mathbf{k}$ around valley τ in sublattice α of layer μ, while $(a_{\mathbf{q}}^\nu)^\dagger/a_{\mathbf{q}}^\nu$ creates/annihilates longitudinal ($\nu = L$) and transverse ($\nu = T$) phasons with momentum $\mathbf{q}$.

A similar coupling was considered in Ref. 53 to estimate the contribution from phonon umklapp scattering in graphene on boron nitride. In fact, Eq. (22) captures both normal and umklapp scattering processes when the reconstruction of the electronic spectrum is taken into account and the resulting mini-bands are represented in a reduced-zone scheme, with crystal momentum $\mathbf{k}$ restricted to the first moiré Brillouin zone. For low carrier concentration, umklapp processes are mediated by phonons with momenta of the order of $\mathbf{G}_i$ and, therefore, their contribution to the resistivity scales as $\varrho \propto e^{-\hbar\omega_{\mathbf{G}_i}/k_BT}/T$. In what follows, I am going to restrict the analysis to quasi-elastic scattering processes within the Fermi circles around the mini-Dirac points of the spectrum.

A. Low-energy Hamiltonian

Before tacking the calculation of the resistivity, let me discuss first the impact of phason fluctuations on the low-energy sector of the electronic spectrum. In a commensurate approximant, valleys $K_{\tau,\mu}$ are folded onto corners $\tilde{\kappa}_\eta$ of the moiré Brillouin zone; Figure 1 illustrates the case for type-I (in the nomenclature of Ref. 15, *sublattice-exchange odd* in Ref. 29) commensurate structures, which are dense in the limit of small angles.[30] A $\mathbf{k}\cdot\mathbf{p}$ expansion around these points reads

$$\hat{\mathcal{H}}^{(\tau,\mu)} = \hbar v_F^*\left(\tau\hat{\sigma}_x, \hat{\sigma}_y\right)\cdot\mathbf{p} + \hat{\mathcal{H}}_{\text{e-ph}}^{(\tau,\mu)}, \tag{23}$$

where $\mathbf{p}$ is the crystalline momentum around $\tilde{\kappa}_\eta$. This Hamiltonian acts on a new spinor basis for each valley sector τ, each entry corresponding to envelope wave functions mostly localized on a given sublattice of layer μ; the hybridization with other sublattices/layers is measured by the parameter $\alpha \equiv t_\perp/(\hbar v_F|\tilde{\kappa}_\eta|)$. The new
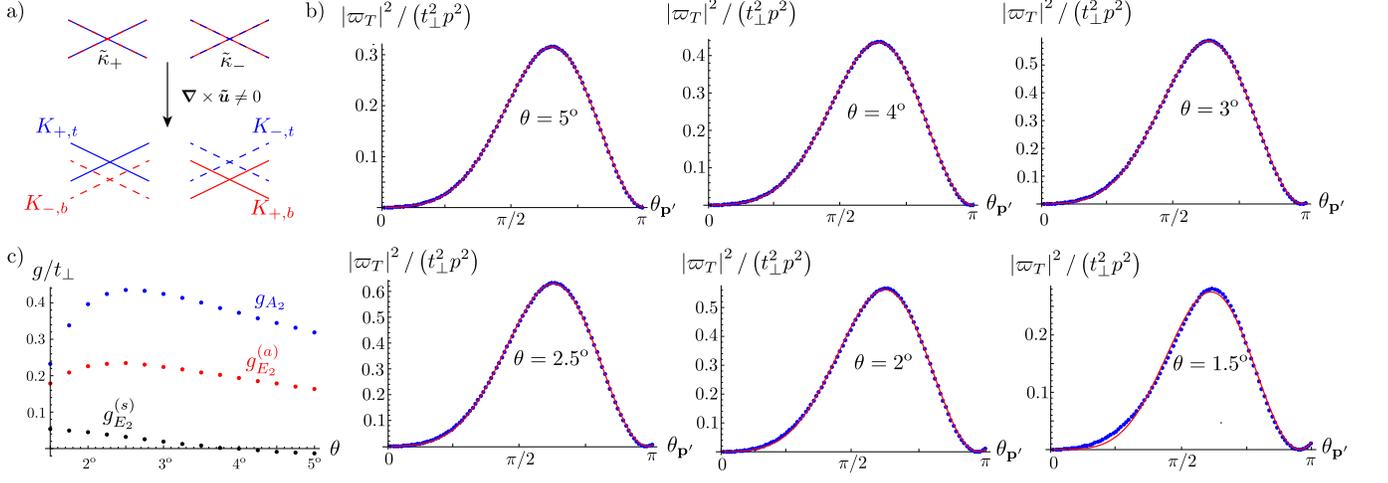
FIG. 3: a) Schematic representation of the splitting of the Dirac mini-bands under a local *tilting* distortion of the moiré pattern (coupling parametrized by g_{A_2}); dashed lines correspond to valley $\tau = +1$, straight lines to valley $\tau = -1$, blue color to top layer $\mu = +1$ (t), and red color to bottom layer $\mu = -1$ (b). b) Electron-phason matrix elements in valley $\tau = +1$ within the conduction band $\zeta = +1$ around κ_+ point as a function of the scattering angle. The incident momentum $\mathbf{p} = p\,\hat{\mathbf{e}}_x$ corresponds to a filling of $n = 10^{11}$ cm^{-2} in the Dirac model, i.e., $pa = 0.0056$. Blue dots correspond to the numerical evaluation in the continuum model for the electronic bands, red curves correspond to the phenomenological fitting in Eq. (26). c) Electron-phason couplings within the low-energy Dirac bands as a function of the twist angle extracted from the fittings.

Fermi velocity can be estimated in perturbation theory[2] as $v_F^*/v_F = (1 - 3\alpha^2)/(1 + 6\alpha^2)$; $\alpha^2 = 1/3$ defines the first magic angle in this approximation.

The second term in Eq. (23) represents the electron-phason coupling projected onto the lowest-energy bands. The most general phenomenological expansion allowed by symmetry reads

$$\hat{\mathcal{H}}_{\text{e-ph}}^{(\tau, t/b)} = g_{A_1} \boldsymbol{\nabla} \cdot \tilde{\mathbf{u}} \pm g_{A_2} \left(\boldsymbol{\nabla} \times \tilde{\mathbf{u}}\right)_z + g_{E_2}^{(s)} \left[\left(\partial_x \tilde{u}_x - \partial_y \tilde{u}_y\right) \hat{\sigma}_x - \tau \left(\partial_x \tilde{u}_y + \partial_y \tilde{u}_x\right) \hat{\sigma}_y\right]$$
$$\pm g_{E_2}^{(a)} \left[\tau \left(\partial_x \tilde{u}_x - \partial_y \tilde{u}_y\right) \hat{\sigma}_y + \left(\partial_x \tilde{u}_y + \partial_y \tilde{u}_x\right) \hat{\sigma}_x\right], \tag{24}$$

where the upper/lower sign applies to top/bottom layer sectors. Note that the electron-phason coupling can only depend on derivatives of the phason field due to the invariance of the electronic spectrum with respect to infinitesimal translations of the moiré pattern. The combinations in Eq. (11) can be paired with electronic operators transforming under the same irreducible representation of D_6 to form invariants under the point group and time reversal operations; details can be found in Appendix B. The coupling constants g_i are phenomenological parameters with units of energy.

The first and third terms in Eq. (24) resemble the *scalar* and *vector* electron-phonon couplings in graphene and are expected to be subleading in layer hybridization, $g_{A_1}, g_{E_2}^{(s)} \sim \mathcal{O}(\alpha^2)$; consequently, the electron-phason coupling is dominated by transverse modes. The other two couplings act with opposite signs on valleys $K_{\tau, \pm}$ coming from different layers. In particular, tilting the

soliton network $(\boldsymbol{\nabla} \times \tilde{\mathbf{u}} \neq 0)$ with respect to the preferential direction imposed by the moiré superlattice lifts the degeneracy of these points, as illustrated in Fig. 3 a; the coupling parametrized by $g_{E_s}^{(a)}$ includes the effect of distortions of the beating pattern that break the 3-fold rotational symmetry, displacing the positions of the Dirac crossings in $\mathbf{k}$-space. These couplings reproduce the effect of relative strain between layers in the band structure[45] and, as noted before, could explain the reduced Landau level degeneracy reported in magneto-transport.

Figure 3 b shows the numerical evaluation (blue dots) of the matrix elements of the coupling with transverse phason modes in the second line of Eq. (22):

$$\langle \zeta', \tau, \mathbf{k}'| \hat{H}_{\text{ep}} |\zeta, \tau, \mathbf{k}\rangle = \sqrt{\frac{\hbar}{2A\tilde{\rho}\omega_{\mathbf{q}}^{(T)}}} \, \varpi_{T,\tau}^{\zeta, \zeta'} \left(\mathbf{q}, \mathbf{k}, \mathbf{k}'\right) \delta_{\mathbf{k}', \mathbf{k}+\mathbf{q}}. \tag{25}$$

Here $|\zeta, \tau, \mathbf{k}\rangle$ represents electronic states from valley τ in minin-band ζ with crystalline momentum $\mathbf{k}$ restricted to the first moiré Brillouin zone, which are obtained by diagonalizing the Hamiltonian in Eq. (16); the resulting matrix is truncated to a finite number of momentum values in each layer coming from different copies of the moiré zone, ranging from 7 for the largest twist angle ($\theta = 5^{\circ}$)

to 61 for the smallest ($\theta = 1.5^{\circ}$). The results are normalized by the incident momentum (measured with respect to the corresponding mini-Dirac point) along the x axis corresponding to a filling of $n = 10^{11}$ cm^{-2} in the lowest-energy electron band and fitted (red curves) to the phenomenological expression derived from Eq. (24),

$$\varpi_{T,\tau}^{\mu,\zeta}(\mathbf{q}, \mathbf{p}, \mathbf{p}') = -\mu \, g_{A_2} \, |\mathbf{q}| \cos\left(\frac{\theta_{\mathbf{p}'} - \theta_{\mathbf{p}}}{2}\right) - \mu\tau\zeta \, g_{E_2}^{(a)} \, |\mathbf{q}| \cos\left(2\theta_{\mathbf{q}} + \frac{\theta_{\mathbf{p}'} + \theta_{\mathbf{p}}}{2}\right) + \tau\zeta \, g_{E_2}^{(s)} \, |\mathbf{q}| \sin\left(2\theta_{\mathbf{q}} + \frac{\theta_{\mathbf{p}'} + \theta_{\mathbf{p}}}{2}\right),$$

$$(26)$$

from which I estimate the coupling constants shown in Fig. 3 c for different twist angles. The calculation is restricted to intra-band ($\zeta = \zeta'$), quasi-elastic processes, for which $|\mathbf{q}| = |\mathbf{p}' - \mathbf{p}| = 2|\mathbf{p}| \sin(\frac{\theta_{\mathbf{p}'} - \theta_{\mathbf{p}}}{2})$. Note that in these last expressions momenta are measured with respect to $\tilde{\kappa}_\eta$, with $\eta = \mu \times \tau$ as prescribed by the folding scheme in Fig. (1), and $\theta_{\mathbf{k}}$ parametrizes the direction of momentum $\mathbf{k}$, namely, $\mathbf{k} = |\mathbf{k}|(\cos\theta_{\mathbf{k}}, \sin\theta_{\mathbf{k}})$. In particular, the calculations of Fig. 3 corresponds to $\tau = +1$ and $\mu = +1$; the values of the couplings do not change appreciably when the calculation is performed for different incident momenta or band/valley numbers within the low-energy Dirac cones. The accuracy of the fitting curves is very good down to angles of the order of $\theta = 1.2^{\circ}$.

The layer-symmetric coupling $g_{E_2}^{(s)}$ is only appreciable for the smallest angles, as expected. The layer-asymmetric couplings are non-monotonic with twist angle. This behavior is reproduced by the $\mathbf{k} \cdot \mathbf{p}$ perturbative expansion presented in Appendix B, which predicts a stronger coupling with *tilting* deformations of the beating pattern; second-order perturbation theory gives

$$g_{A_2} \approx \frac{3\,\alpha\,t_\perp}{1 + 6\alpha^2}\left(\frac{v_F^*}{v_F}\right). \quad (27)$$

The coupling grows first as the layer hybridization of the electronic wave function $\alpha/(1 + 6\alpha^2)$ increases, but around $\theta \lesssim 3^{\circ}$ it starts to decrease following the same trend as the Fermi velocity.

B. Phason-limited electronic transport

Electron scattering by strong phason fluctuations is expected to contribute to the T-dependent resistivity of twisted bilayer graphene at small incommensurate angles. I am going to consider a semiclassical treatment in the framework of Boltzmann transport theory, which implicitly assumes that $k_F \ell \gg 1$, where k_F is the momentum of carriers within the Fermi surface and ℓ is the phason-limited mean free path. The electron-phason coupling

enters explicitly in the collision integral via a scattering probability rate, which can be computed from Fermi's golden rule as $\mathcal{W}_i^f = 2\pi\hbar^{-1}|\langle f|\hat{H}_{\mathrm{ep}}|i\rangle|^2 \delta(E_f - E_i)$; the initial state reads $|i\rangle = |\zeta, \tau, \mathbf{k}\rangle \otimes |n_{\mathbf{q}}^{(\nu)}\rangle$, where $|n_{\mathbf{q}}^{(\nu)}\rangle$ represents a state with $n_{\mathbf{q}}^{(\nu)}$ phasons in branch ν. Note that, although the electron-phason coupling projected onto the low-energy bands results from coherent superpositions of electronic states in the two layers, at this point we are neglecting inter-band coherences in the calculation of the resistivity, which could alter the results as the system approaches the neutrality point.[54,55] I am also going to assume that the phason ensemble thermalize much faster than electrons, so $n_{\mathbf{q}}^{(\nu)}$ reduces to a equilibrium Bose-Einstein distribution function. Phason emission/absorption processes scatterer the initial state into $|f\rangle = |\zeta', \tau, \mathbf{k}'\rangle \otimes |n_{\mathbf{q}}^{(\nu)} \pm 1\rangle$ at a rate

$$
\begin{aligned}
\mathcal{W}_{\zeta,\mathbf{k}}^{\zeta',\mathbf{k}'} =&\frac{2\pi}{\hbar} \sum_{\nu=L,T} \sum_{\mathbf{q}} \frac{\hbar \left|\varpi_{\nu,\tau}^{\zeta,\zeta'}(\mathbf{q}, \mathbf{k}, \mathbf{k}')\right|^2}{2A\tilde{\rho}\omega_{\mathbf{q}}^{(\nu)}} \delta_{\mathbf{k}',\mathbf{k}+\mathbf{q}} \times \\
& \left[n_{\mathbf{q}}^{(\nu)} \, \delta\left(\varepsilon_{\zeta',\tau,\mathbf{k}'} - \varepsilon_{\zeta,\tau,\mathbf{k}} - \hbar\omega_{\mathbf{q}}^{(\nu)}\right) \right. \\
& \left. + \left(n_{\mathbf{q}}^{(\nu)} + 1\right) \delta\left(\varepsilon_{\zeta',\tau,\mathbf{k}'} - \varepsilon_{\zeta,\tau,\mathbf{k}} + \hbar\omega_{\mathbf{q}}^{(\nu)}\right) \right],
\end{aligned}
$$

$$(28)$$

In order to obtain analytical formulas, I am going to restrict the analysis to intra-band, quasi-elastic processes within the Dirac cones, dominated by the g_{A_2} coupling with transverse phason modes according to the estimates presented in the previous subsection. The resistivity can be obtained from a variational method[56] applied to the linearized Boltzmann equation describing the evolution of deviations of the electronic distribution function from equilibrium. The calculation is analogous to the case of the phonon-limited resistivity in graphene[57] and I am not going to reproduce the details here. The final result reads

$$\varrho = \frac{\hbar}{e^2} \frac{(2k_F g_{A_2})^2}{\tilde{\rho} \, (v_F^*)^2 \, k_B T} \, I\left(\frac{T}{T_{\mathrm{BG}}}\right), \quad (29)$$
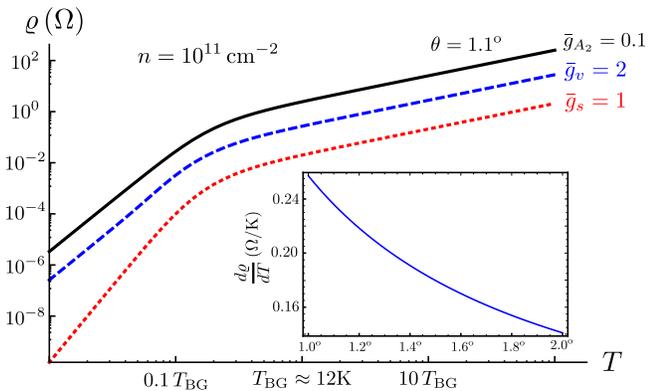
FIG. 4: T-dependent resistivity (in logarithmic scale) deduced from Eq. (29) (black curve, corresponding to twist angle $\theta = 1.1°$). Dashed blue and dotted red curves show the contribution from the scalar ($\bar{g}_s$) and vectorial ($\bar{g}_v$) couplings with in-plane (layer symmetric) phonons.[57] The inset shows the dependence of $d\varrho/dT$ with twist angle in the high-T regime.

where $I(x)$ is a dimensionless function defined by

$$I(x) = \int_0^1 dy \, y^4 \sqrt{1 - y^2} \, \frac{e^{y/x}}{\left(e^{y/x} - 1\right)^2}. \qquad (30)$$

The temperature scale T_{BG}, akin to the Bloch-Grüneisen temperature in the problem of electron-phonon scattering, is related to the maximum momentum transfer ($2k_F$) between electronic states in a quasi-elastic scattering event,

$$k_B T_{\mathrm{BG}} = \hbar \omega_{2k_F}^{(T)}. \qquad (31)$$

At temperatures much lower than T_{BG}, the resistivity is dominated by small-angle scattering events and scales as $\varrho \sim (T/T_{\mathrm{BG}})^4$. In the high-temperature regime, $T \gg T_{\mathrm{BG}}$, the resistivity grows linearly with T as

$$\varrho \approx \frac{\pi \left(g_{A_2}\right)^2 k_B T}{16 e^2 \hbar \left(v_F^*\right)^2 (\tilde{\mu} + \gamma)}. \qquad (32)$$

Given the dispersion relation in Eq. (14b) and the relation between the Fermi momentum and the carrier concentration in the low-energy Dirac model, $k_F = \sqrt{\pi n/2}$, the crossover takes place at temperatures of the order of $T_{\mathrm{BG}} \approx 12\sqrt{n}$ K, with n measured in units of 10^{11} cm^{-2}. Figure 4 (black curve) shows the dependence on temperature for fixed values of the carrier density ($n = 10^{11}$ cm^{-2}) and the dimensionless electron-phason coupling $\bar{g}_{A_2} \equiv a g_{A_2}/\hbar v_F^* = 0.1$.

IV. DISCUSSION

A linear-T resistivity is expected in general if electron transport is limited by scattering off boson fluctuations above a certain temperature scale, the latter defined by phase-space constraints in these scattering events. In particular, for the coupling with either phasons or layer-symmetric phonons (corresponding to acoustic in-phase vibrations of both layers, for which the crossover temperature is of the same order as T_{BG} defined above), the slope of the resistivity with temperature can be written as

$$\frac{d\varrho_i}{dT} \approx R_0 \left(\bar{g}_i\right)^2 \frac{k_B}{16 a^2 \mathcal{K}_i}, \qquad (33)$$

where $R_0 = \pi \hbar/e^2$ is the quantum of resistance, $\mathcal{K}_i$ is the suitable elastic modulus ($\tilde{\mu} + \gamma$ for transverse phasons, $\lambda + 2\mu$ and μ for longitudinal and transverse phonons, respectively), and the dimensionless parameter $\bar{g}_i$ measures the strength of the coupling within the mini-Dirac bands in units of the bandwidth parameter $\hbar v_F^*/a$. In the case of phasons, we have $\bar{g}_{A_2} \sim t_\perp/t$, where the exact fraction depends on the grade of layer hybridization in the wave function and, therefore, is expected to be sensitive to the effect of lattice relaxation on the electronic spectrum, which is not taken into account in the estimates of Sec. III. Regarding layer-symmetric phonons, since the vector potential couples directly to the electron velocity operator, interference of the electronic wave function leads to the same cancelation as in the Fermi velocity, giving $\bar{g}_v \approx \beta \equiv -\partial \ln t/\partial \ln a \sim 2 - 3$. In the case of the scalar coupling, the same cancelation takes place due to the reconstruction of the electronic spectrum, which enters through the electrostatic screening of the deformation potential;[56] in a Thomas-Fermi treatment,[57] the dimensionless coupling reduces to

$$\bar{g}_s \approx \frac{aD}{8 e^2 k_e} \sim 1, \qquad (34)$$

where $D \approx 20 - 30$ eV and $k_e = 1/4\pi\epsilon_0$ are the bare deformation potential[58] and Coulomb constants, respectively.

The conclusions of this analysis are the following: i) The effective couplings (normalized by the bandwidth) within the reconstructed Dirac cones should not depend strongly on the twist angle with the exception of, maybe, the coupling with phason modes due to its sensitivity to layer hybridization. ii) Despite the weaker coupling with phasons, these modes can dominate the resistivity at small twist angles due to the reduced stiffness of the stacking domain-wall system. The inset of Fig. 4 shows the dependence of the slope of the resistivity as a function of the twist angle prescribed by the scaling of the elastic constants in Eqs. (13) (i.e., neglecting changes in the effective couplings). The resistivity increases as the twist angle decreases due to the reduction of the rigidity of the soliton network, but the slope is at least two orders of magnitude smaller than those reported in the experiments.[25,26] This points to a different mechanism, possibly related to strong electron correlations around the magic angle. Recently, González and Stauber have

argued that perfect nesting for fillings close to the emergent van Hove singularity at the band edge gives rise to a marginal Fermi liquid scaling of the quasiparticle lifetime.[59] Partially related to this observation, a recent model for the linear-T resistivity in the normal state of cuprate superconductors invoking umklapp scattering[60] starts from the assumption of a Fermi surface reconstruction to maximize commensurate nesting; this could be induced by a spin-wave instability that, also in the present case,[61,62] can compete with the superconducting order. Other indications are the fact that the low-temperature crossover is systematically smaller than $T_{\rm BG}$, as pointed out in Ref. 26, and the absence of saturation of the resistivity around the Fermi temperature (only when higher-energy bands start to be populated the resistivity drops), which may indicate that electronic quasiparticles are not well defined.

The conclusions presented here are based on Boltzmann transport theory, which does not include mesoscopic effects ascribed to the intrinsically disordered nature of the devices. In particular, in the comparison with the contribution from conventional phonons, I have neglected the effect of disorder on the beating pattern, which could pin the soliton network and open a gap in the phason spectrum. The reduction of the effective rigidity of the soliton network implies that phason fluctuations are enhanced and, consequently, the moiré pattern is also more sensitive to perturbations induced by the substrate. The simplest perturbation to the harmonic theory in Eq. (12) is a *weak* (in the sense of Larkin[63]) disorder potential of the form

$$V_{dis} = - \int d\mathbf{r}\, \mathbf{f}\left(\mathbf{r}\right) \cdot \tilde{\mathbf{u}}\left(\mathbf{r}\right). \qquad (35)$$

Here $\mathbf{f}\left(\mathbf{r}\right)$ represent forces acting independently on the stacking solitons. This potential breaks the translational invariance of the incommensurate lattice and pin the soliton system. Positional order in the moiré superlattice is lost at distances of the order of

$$L_c \approx \frac{6\left(\tilde{\lambda} + 2\tilde{\mu}\right)}{\pi \sigma_f}, \qquad (36)$$

where $\sigma_f^2 \equiv \overline{f^2}$ represents the dispersion in the distribution of forces. In the case of encapsulated samples, for example, we can estimate this parameter as $\sigma_f \sim V_{\rm BN}/L_{\rm BN}$, where $V_{\rm BN} \approx 0.1$ eV/nm^2 characterizes the adhesion energy between graphene and boron nitride,[64] and $L_{\rm BN} \approx 14$ nm is the characteristic size of the moiré due to the lattice mismatch, about 1.8%. According to Eqs. (13) and (36), around the magic angle $\theta \sim 1^\circ$ positional order in the corresponding moiré superlattice is lost at distances $L_c \sim 26$ nm, about twice the size of the moiré period; L_c collapses to L_M at angles $\theta \sim 0.7^\circ$.

Equation (36) must be interpreted as a collective pinning length[65,66] below which the soliton network responds elastically or, more accurately, the harmonic expansion in Eq. (12) holds. What happens beyond that length it is difficult to say due to the evident shortcomings of the model in Eq. (35), which, for example, neglects the quasi-periodicity of the beating pattern and, possibly, of the interaction with boron nitride if the latter is aligned with the sample. The difficulty arises from the fact that realistic disorder potentials vary on length scales much shorter than the moiré period. These considerations will be taken into account in a future study. Nevertheless, it is worth emphasizing that, although these are only tentative estimates, it is precisely the low stiffness of the soliton network defining the beating pattern at small twist angle what makes the system so sensitive to structural disorder, explaining the widespread presence of *twist angle disorder* in the samples.

Finally, another possible mechanism for the softening of the soliton network is the presence of strong nematic fluctuations of electronic origin, as suggested by tunneling spectroscopy.[20–22] This observation is particularly relevant for the nematicity around the charge neutrality point, which seems to be more pronounced at stacking domain walls.[21,22] The phenomenological Hamiltonian presented in Eq. (24) provides a good description of the coupling with phason fluctuations close to neutrality. The symmetry analysis presented in Appendix B can be used to construct phenomenological theories describing the coupling of nematic order parameters with lattice degrees of freedom. The interplay between electrons, phasons, and nematic fluctuations is also of potential relevance to understand the T-dependent resistivity observed in twisted bilayer graphene. On the experimental front, it would be interesting to have a systematic study of the anisotropy in transport as a function of carrier density and in the absence and presence of applied tensions (although this could be affected by other mesoscopic effects). It would be also very useful to compare the resistivity of devices under different hydrostatic pressure (which have been shown to display a similar phenomenology than *magic angle* bilayers[5]) in order to determine the dominant role of phasons in transport, since the application of pressure can pin the soliton network and suppress this contribution to the resistivity.

In conclusion, long-wavelength fluctuations of a moiré beating pattern in the limit of small twist angles are dominated by phason modes. Their contribution to resistivity grows linearly with T, with increasing slope as the twist angle decreases due to the reduction of the stiffness of the soliton network. This contribution alone, however, seems to be insufficient to explain the fast growth of the resistivity when the *magic angle* is approached, pointing to a different mechanism that might involve the presence of strong nematic fluctuations or a Fermi surface reconstruction linked to the correlated phenomena at lower temperatures.

Acknowledgments

I would like to thank Francisco Guinea, Abhay Pasupathy, Matthew Yankowitz, and Cory Dean for valuable discussions, and Ricardo Zarzuela for his careful reading of an early version of this manuscript. This work has been supported by the NSF MRSEC program DMR-1420634.

Note: While finishing this manuscript, I came across a recent preprint,[67] which presents a full calculation of the spectrum of oscillation with the account of lattice relaxation, following the same recipe as in Appendix A. This work highlights the role of stacking domain walls, reaching the same conclusions as in Sec. II, in particular, the scaling of the elastic constants of the moiré superlattice in Eqs. (13).

Appendix A: Mechanical model

Lattice relaxation is described by an elastic free energy of the form

$$F = F_{\text{el}} + F_{\text{ad}}. \tag{A1}$$

The first terms accounts for in-plane elastic distortions of graphene layers,

$$F_{\text{el}} = \sum_{\mu=t,b} \int d\mathbf{r} \left[\frac{\lambda}{2} \left(\boldsymbol{\nabla} \mathbf{u}^{(\mu)} \right)^2 + \frac{\mu}{4} \left(\partial_i u_j^{(\mu)} + \partial_j u_i^{(\mu)} \right)^2 \right], \tag{A2}$$

where $\mathbf{u}^{(\mu)}$ describes displacements of unit cells in layer μ with respect to their equilibrium positions in the absence of interlayer forces. It is convenient to introduce the relative displacement, $\mathbf{u} = \mathbf{u}^{(t)} - \mathbf{u}^{(b)}$, and total displacement, $\mathbf{v} = \mathbf{u}^{(t)} + \mathbf{u}^{(b)}$. In the simplest approximation, the adhesion energy, second term in Eq. (A1), is a functional of the former field only,

$$F_{\text{ad}} = \int d\mathbf{r} \, \mathcal{V}_{\text{ad}} \left[\mathbf{r}, \mathbf{u} \left(\mathbf{r} \right) \right], \tag{A3}$$

where $\mathcal{V}_{\text{ad}}$ is the adhesion potential introduced in the main text. Equation (5) comes from minimizing the free energy of variations with respect to $\mathbf{u}$. Dynamical equations are derived from the total Lagrangian $L = K - F$, where the kinetic energy reads

$$K = \frac{\rho}{4} \int d\mathbf{r} \left[\dot{\mathbf{v}}^2 + \dot{\mathbf{u}}^2 \right]. \tag{A4}$$

Here $\rho = 7.6 \times 10^{-7}$ kg/m² is the mass density of individual graphene layers.

The first field $\mathbf{v}$ represents in-phase displacements of both layers and can be identified with the original acoustic phonons.[57] From this point on, I am going to focus on the dynamics of relative displacements, $\mathbf{u}$. Let me consider deviations from a metastable configuration,

$\delta\mathbf{u}(t, \mathbf{r}) = \mathbf{u}(t, \mathbf{r}) - \mathbf{u}^{(0)}(\mathbf{r})$, where $\mathbf{u}^{(0)}(\mathbf{r})$ is a solution of Eq. (5). Plugging this ansatz into the previous equations gives

$$F_{\text{el}} \left[\mathbf{u}^{(0)} \right] + F_{\text{el}} \left[\delta\mathbf{u} \right] + F_{\text{mix}} \left[\delta\mathbf{u}, \mathbf{u}^{(0)} \right] + F_{\text{ad}} \left[\delta\mathbf{u}, \mathbf{u}^{(0)} \right], \tag{A5}$$

where the first and second terms are just Eq. (A2) evaluated with the metastable solution and the corresponding deviation, while the third term mixes both,

$$F_{\text{mix}} \left[\delta\mathbf{u}, \mathbf{u}^{(0)} \right] = \int d\mathbf{r} \left[\frac{\lambda}{2} \partial_i u_i^{(0)} \partial_j \delta u_j \right. \tag{A6}$$
$$\left. + \frac{\mu}{4} \left(\partial_i u_j^{(0)} + \partial_j u_i^{(0)} \right) \left(\partial_i \delta u_j + \partial_j \delta u_i \right) \right].$$

The last term comes from the adhesion energy; expanding up to quadratic order gives

$$F_{\text{ad}} \left[\delta\mathbf{u}, \mathbf{u}^{(0)} \right] \approx \int d\mathbf{r} \left\{ \mathcal{V}_{\text{ad}} \left[\mathbf{r}, \mathbf{u}^{(0)} \right] + \delta u_i \frac{\partial \mathcal{V}_{\text{ad}}}{\partial u_i} |_{\mathbf{u}^{(0)}} \right.$$
$$\left. + \frac{1}{2} \delta u_i \delta u_j \frac{\partial^2 \mathcal{V}_{\text{ad}}}{\partial u_i \partial u_j} |_{\mathbf{u}^{(0)}} \right\}. \tag{A7}$$

Hereafter repeated latin indices are summed up. Integration by parts (dropping boundary terms) leads to

$$F \approx F_0 + U \left[\delta\mathbf{u} \right] + \int d\mathbf{r} \, \delta u_i \left[\frac{\partial \mathcal{V}_{\text{ad}}}{\partial u_i} |_{\mathbf{u}^{(0)}} \right. \tag{A8}$$
$$\left. - \frac{\lambda}{2} \partial_i \partial_j u_j^{(0)} - \frac{\mu}{2} \partial_j \left(\partial_i u_j^{(0)} + \partial_j u_i^{(0)} \right) \right].$$

The last term is 0 just from Eq. (5). The first term represents the free energy of the equilibrium solution,

$$F_0 = \int d\mathbf{r} \left\{ \frac{\lambda}{4} \left(\boldsymbol{\nabla} \mathbf{u}^{(0)} \right)^2 + \frac{\mu}{8} \left(\partial_i u_j^{(0)} + \partial_j u_i^{(0)} \right)^2 \right.$$
$$\left. + \mathcal{V}_{\text{ad}} \left[\mathbf{r}, \mathbf{u}^{(0)} \left(\mathbf{r} \right) \right] \right\}, \tag{A9}$$

while $U[\delta\mathbf{u}]$ describes the spectrum of harmonic oscillations,

$$U \left[\delta\mathbf{u} \right] = \frac{1}{2} \int d\mathbf{r} \left\{ \frac{\lambda}{2} \left(\boldsymbol{\nabla} \delta\mathbf{u} \right)^2 + \frac{\mu}{4} \left(\partial_i \delta u_j + \partial_j \delta u_i \right)^2 \right.$$
$$\left. + \delta u_i \delta u_j \frac{\partial^2 \mathcal{V}_{\text{ad}}}{\partial u_i \partial u_j} |_{\mathbf{u}^{(0)}} \right\}. \tag{A10}$$

The Euler-Lagrange equations describing oscillations around the metastable state reads then

$$-\rho \, \delta\ddot{u}_i + \frac{\lambda + \mu}{2} \partial_i \partial_j \delta u_j + \frac{\mu}{2} \partial_j \partial_j \delta u_i = \delta u_j \frac{\partial^2 \mathcal{V}_{\text{ad}}}{\partial u_i \partial u_j} |_{\mathbf{u}^{(0)}}. \tag{A11}$$

By introducing Fourier series,

$$\delta u_i \left(t, \mathbf{r} \right) = \frac{1}{\sqrt{A}} \sum_{\mathbf{q}} \int \frac{d\omega}{2\pi} u_i \left(\omega, \mathbf{q} \right) e^{i\mathbf{q}\cdot\mathbf{r} - i\omega t}, \tag{A12}$$
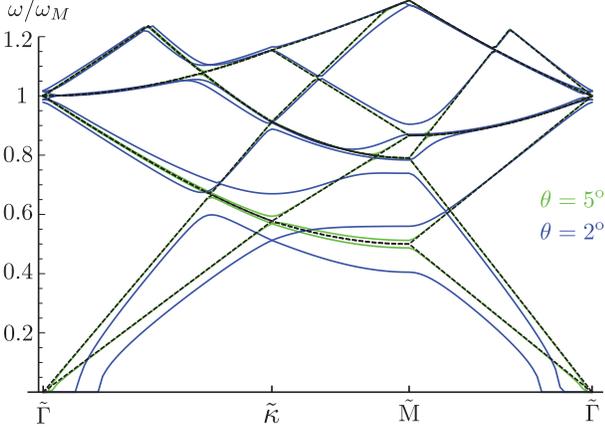
FIG. 5: Spectrum of oscillations of two floating layers. When ℓ and L_M are comparable (for example, curves in green for $\theta = 5°$), the spectrum resembles the acoustic phonon branches of graphene folded onto the moiré Brillouin zone (in dashed black for reference). For smaller angles, the strong softening of the phonon modes around the zone center mark the instability of the system towards the formation of sharper stacking textures.

the problem reduces to solve the following secular equation,

$$\rho \omega^2 u_i (\omega, \mathbf{q}) = \left[\mu |\mathbf{q}|^2 \delta_{ij} + (\lambda + \mu) q_i q_j \right] u_j (\omega, \mathbf{q}) + \sum_{\mathbf{G}} K_{ij} (\mathbf{G}) u_j (\omega, \mathbf{q} - \mathbf{G}), \quad (A13)$$

where I have introduced

$$K_{ij} (\mathbf{G}) = \frac{4}{\sqrt{3} L_M^2} \int_{\text{moiré}} d\mathbf{r} \, e^{-i\mathbf{G} \cdot \mathbf{r}} \frac{\partial^2 \mathcal{V}_{\text{ad}}}{\partial u_i \partial u_j} \Big|_{\mathbf{u}^{(0)}}. \quad (A14)$$

1. Floating layers

For large twist angles such that L_M and ℓ (defined in Eq. 7) are comparable, we can neglect lattice relaxation, $\mathbf{u}^{(0)} \approx 0$, and consider oscillations around two floating layers. The average free energy per moiré cell is just $\bar{F}_0 = 3V/2$. The adhesion potential reduces to

$$\mathcal{V}_{\text{ad}} [\mathbf{r}, 0] = V \sum_{i=1}^{3} \left[\frac{1}{2} + \cos (\mathbf{G}_i \cdot \mathbf{r}) \right]. \quad (A15)$$

In this approximation, all the harmonics $K_{ij}(\mathbf{G})$ are 0 except for the ones in the first star ($\alpha = 1, 2, 3$)

$$K_{ij} (\pm \mathbf{G}_\alpha) = -\frac{V}{2} (\mathbf{b}_\alpha)_i (\mathbf{b}_\alpha)_j. \quad (A16)$$

The spectra in Fig. 5 are the result of truncating the secular equation to include the first 36 harmonics, giving

a 37×37 matrix to diagonalize. The lowest-frequency branches are strongly softened (blue curves) when the matrix elements in Eq. (A16) start to be comparable with $\omega_M = 4\pi c/(\sqrt{3} L_M)$.

2. Soliton network

Let me assume now that the lattice relaxes and the solution is roughly described by the superposition of the three density waves introduced in the main text,

$$\mathbf{u}^{(0)} (\mathbf{r}) = \sum_{\alpha=1,2,3} \mathbf{u}_\alpha (\varrho), \quad (A17)$$

where $\mathbf{u}_\alpha (\varrho) \approx u (\varrho) \hat{\mathbf{u}}_\alpha$ is a soliton train along $\hat{\boldsymbol{\varrho}}_\alpha = \hat{\mathbf{z}} \times \hat{\mathbf{u}}_\alpha$, $u(\varrho)$ is given by Eq. (8), and the unit vectors $\hat{\mathbf{u}}_\alpha$ lie along the three armchair directions:

$$\hat{\mathbf{u}}_1 = \left(-\frac{\sqrt{3}}{2}, -\frac{1}{2} \right), \quad (A18a)$$

$$\hat{\mathbf{u}}_2 = \left(\frac{\sqrt{3}}{2}, -\frac{1}{2} \right), \quad (A18b)$$

$$\hat{\mathbf{u}}_3 = (0, 1). \quad (A18c)$$

To the leading order in ℓ/L_M, we can neglect soliton interactions at the crossings and approximate

$$\bar{F}_0 \approx \frac{1}{\sqrt{3} L_M} \sum_\alpha \int d\varrho \left\{ \frac{\mu}{2} (\partial_\varrho u_\alpha)^2 + 2 V [0, \mathbf{u}_\alpha (\varrho)] \right\} \approx \frac{2\sqrt{3} \, \sigma}{L_M}. \quad (A19)$$

For a sine-Gordon soliton, the tension reduces to

$$\sigma = \int d\varrho \left\{ \frac{\mu}{4} (\partial_\varrho u_\alpha)^2 + V [0, \mathbf{u}_\alpha (\varrho)] \right\} \quad (A20)$$
$$= V \int d\varrho \, \text{sech}^2 \left(\frac{\varrho - x_\alpha}{\ell} \right) = 2\ell V = \frac{a}{\pi} \sqrt{2\mu V}.$$

As argued in the main text, the acoustic branches in this limit can be identified with traveling-wave modes of the soliton network parametrized by collective coordinates $\tilde{\mathbf{u}}$. Specifically, if $\mathbf{u}^{(0)} (\mathbf{r} - \tilde{\mathbf{u}})$ is the soliton-network solution centered at $\tilde{\mathbf{u}}$, then a smoothly distorted profile can be approximated by the functional

$$\mathbf{u} [\tilde{\mathbf{u}} (\mathbf{r})] = \mathbf{u}^{(0)} (\mathbf{r} - \tilde{\mathbf{u}} (\mathbf{r})), \quad (A21)$$

where the collective coordinate has been promoted to a field $\tilde{\mathbf{u}}(\mathbf{r})$. Spatial derivatives can be approximated as

$$\partial_i u_j \approx \partial_i u_j^{(0)} - \partial_k u_j^{(0)} \partial_i \tilde{u}_k, \quad (A22)$$

and therefore we can identify $\partial_i \delta u_j \approx -\partial_k u_j^{(0)} \partial_i \tilde{u}_k$. Plugging this result in the first line of Eq. (A10) gives

$$\mathcal{F} [\tilde{\mathbf{u}} (\mathbf{r})] \approx \frac{1}{2} \int d\mathbf{r} \left\{ \frac{\lambda}{2} \partial_j u_i^{(0)} \partial_l u_k^{(0)} (\partial_i \tilde{u}_j \partial_k \tilde{u}_l) \quad (A23) \right.$$
$$\left. + \frac{\mu}{2} \left[\partial_k u_j^{(0)} \partial_l u_j^{(0)} (\partial_i \tilde{u}_k \partial_i \tilde{u}_l) + \partial_k u_j^{(0)} \partial_l u_i^{(0)} (\partial_i \tilde{u}_k \partial_j \tilde{u}_l) \right] \right\}.$$

We can now estimate from this last result the coefficients in the free-energy expansion of Eq. (12). Since the spatial dependence of the phason field $\tilde{\mathbf{u}}(\mathbf{r})$ must be smooth on the scale of the moiré superlattice, the derivatives of $\mathbf{u}^{(0)}$ inside the integral can be approximated by their average over a moiré unit cell. We end up with

$$\mathcal{F}\left[\tilde{\mathbf{u}}\left(\mathbf{r}\right)\right] \approx \frac{C_{ijkl}}{2} \int d\mathbf{r}\, \partial_i \tilde{u}_j \partial_k \tilde{u}_l, \qquad \text{(A24)}$$

where

$$C_{ijkl} = \frac{\lambda}{\sqrt{3}L_M^2} \int d\mathbf{r}\, \partial_j u_i^{(0)} \partial_l u_k^{(0)} \qquad \text{(A25)}$$
$$+ \frac{\mu}{\sqrt{3}L_M^2} \int d\mathbf{r} \left\{ \partial_j u_k^{(0)} \partial_l u_i^{(0)} + \delta_{ik}\, \partial_j u_m^{(0)} \partial_l u_m^{(0)} \right\}.$$

Note that the symmetry-adapted expansion in the main text adopts the same form, where the tensor of elastic coefficients reads

$$C_{ijkl} = (\gamma + \tilde{\mu})\, \delta_{ik}\delta_{jl} + \tilde{\lambda}\, \delta_{ij}\delta_{kl} + (\tilde{\mu} - \gamma)\, \delta_{jk}\delta_{il}. \quad \text{(A26)}$$

In order to evaluate the elastic constants, we can proceed as before and plug the superposition of the three soliton waves into Eq. (A25). Ignoring again the subleading contribution from soliton crossings, I find

$$C_{ijkl} = I \sum_\alpha \left[\mu\, \delta_{ik} (\hat{\mathbf{u}}_\alpha)_m (\hat{\mathbf{u}}_\alpha)_m (\hat{\boldsymbol{\varrho}}_\alpha)_j (\hat{\boldsymbol{\varrho}}_\alpha)_l \right.$$
$$\left. + (\lambda + \mu)(\hat{\mathbf{u}}_\alpha)_i (\hat{\mathbf{u}}_\alpha)_k (\hat{\boldsymbol{\varrho}}_\alpha)_j (\hat{\boldsymbol{\varrho}}_\alpha)_l \right], \quad \text{(A27)}$$

where the prefactor reads

$$I = \frac{1}{\sqrt{3}L_M} \int d\varrho\, (\partial_\varrho u)^2 = \frac{4V\ell}{\sqrt{3}\mu L_M}. \qquad \text{(A28)}$$

The tensors between parenthesis are just

$$\delta_{ik} \sum_\alpha (\hat{\mathbf{u}}_\alpha)_m (\hat{\mathbf{u}}_\alpha)_m (\hat{\boldsymbol{\varrho}}_\alpha)_j (\hat{\boldsymbol{\varrho}}_\alpha)_l = \frac{3}{2}\delta_{ik}\delta_{jl} \qquad \text{(A29a)}$$

$$\sum_\alpha (\hat{\mathbf{u}}_\alpha)_i (\hat{\mathbf{u}}_\alpha)_k (\hat{\boldsymbol{\varrho}}_\alpha)_j (\hat{\boldsymbol{\varrho}}_\alpha)_l = \frac{9}{8}\delta_{ik}\delta_{jl} - \frac{3}{8}\delta_{ij}\delta_{kl} - \frac{3}{8}\delta_{jk}\delta_{il}. \qquad \text{(A29b)}$$

Comparing these expressions with Eq. (A26), I arrive at the final formulas in Eqs. (13) of the main text.

The effective mass density of the soliton network can be derived in the same manner by approximating the time derivatives in Eq. (A4) as $\dot{u}_i \approx -\partial_j u_i^{(0)} \dot{\tilde{u}}_j$. The mass tensor of a soliton reads then

$$M_{ij} = \frac{\rho}{2} \int d^2\mathbf{r}\, \partial_i u_k^{(0)} \partial_j u_k^{(0)}, \qquad \text{(A30)}$$

where the integral is extended over a moiré unit cell. Proceeding just as before, the final result reads $M_{ij} = M\delta_{ij}$, where M is the inertia of a stacking domain wall,

$$M = \frac{3a^2 \rho L_M}{2\pi^2 \ell}. \qquad \text{(A31)}$$

Dividing this quantity by the area of the moiré unit cell $A_M = \sqrt{3}L_M^2/2$ gives the mass density $\tilde{\rho}$ in Eq. (15). The dispersion relations in Eqs. (14) follow from the corresponding Euler-Lagrange equations, where the Fourier components of the phason field can be decomposed in longitudinal and transverse components as usual,

$$\tilde{\mathbf{u}}\left(\mathbf{q}\right) = \frac{i\,\mathbf{q}}{|\mathbf{q}|}\tilde{u}_L\left(\mathbf{q}\right) + \frac{i\,\hat{\mathbf{z}} \times \mathbf{q}}{|\mathbf{q}|}\tilde{u}_T\left(\mathbf{q}\right). \qquad \text{(A32)}$$

These can be promoted to phason creation/annihilation operators in conventional fashion via the identification (here $\nu = L, T$ labels the branch)

$$\tilde{u}_\nu\left(\mathbf{q}\right) \longrightarrow \sqrt{\frac{\hbar}{2\,\tilde{\rho}\,\omega_{\mathbf{q}}^{(\nu)}}} \left[a_{\mathbf{q}}^{(\nu)} + \left(a_{-\mathbf{q}}^{(\nu)}\right)^\dagger \right], \qquad \text{(A33)}$$

where phason operators satisfy the boson algebra $[a_{\mathbf{q}_1}, (a_{\mathbf{q}_2})^\dagger] = \delta_{\mathbf{q}_1, \mathbf{q}_2}$, following from the conjugacy relations $\{\tilde{u}_i(\mathbf{r}), \pi_j(\mathbf{r}')\} = \delta_{ij}\,\delta(\mathbf{r} - \mathbf{r}')$, where $\boldsymbol{\pi} = \tilde{\rho}\,\dot{\tilde{\mathbf{u}}}$.

Appendix B: Electronic model

Let me start by writing a generic tight-binding Hamiltonian for two graphene layers floating on top of each other, $\hat{H}_{\mathrm{tb}} = \hat{H}_b + \hat{H}_t + \hat{T} + \hat{T}^\dagger$. The first two terms represent intralayer hopping processes; up to first nearest neighbors, we have

$$\hat{H}_\mu = -t \sum_i \left\{ \left| \mathbf{R}_{i,A}^{(\mu)}, \right\rangle \left\langle \mathbf{R}_{i,B}^{(\mu)} \right| + \left| \mathbf{R}_{i,A}^{(\mu)} \right\rangle \left\langle \mathbf{R}_{i,B}^{(\mu)} - \mathbf{a}_1^{(\mu)} \right| \right.$$
$$\left. + \left| \mathbf{R}_{i,A}^{(\mu)} \right\rangle \left\langle \mathbf{R}_{i,B}^{(\mu)} - \mathbf{a}_2^{(\mu)} \right| \right\} + \mathrm{h.c.} \qquad \text{(B1)}$$

The Hamiltonian is written in a monoelectronic basis of Wannier π-orbitals localized on sites A/B of unit cell i of layer μ. We can introduce Bloch states as

$$\left| \mathbf{R}_{i,\alpha}^{(\mu)} \right\rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}_\mu \in \mathrm{BZ}_\mu} e^{-i\mathbf{k}_\mu \cdot \mathbf{R}_{i,\alpha}^{(\mu)}} \left| \mathbf{k}_\mu, \alpha \right\rangle, \qquad (\mathrm{B2})$$

with crystalline momenta $\mathbf{k}_\mu$ restricted to the first Brillouin zone of layer μ (N is the number of unit cells on each layer that I assume the same). The previous Hamiltonian reduces to

$$\hat{H}_\mu = \sum_{\mathbf{k}_\mu \in \mathrm{BZ}_\mu} \sum_{\alpha,\beta=A,B} \left[\hat{H}_{\mathbf{k}_\mu} \right]_{\alpha,\beta} \left| \mathbf{k}_\mu, \alpha \right\rangle \left\langle \mathbf{k}_\mu, \beta \right|, \quad (\mathrm{B3})$$

where the matrix in sublattice space reads

$$\hat{H}_{\mathbf{k}_\mu} = -t \sum_{i=1}^{3} \begin{pmatrix} 0 & e^{i\mathbf{k}_\mu \cdot \boldsymbol{\delta}_i^{(\mu)}} \\ e^{-i\mathbf{k}_\mu \cdot \boldsymbol{\delta}_i^{(\mu)}} & 0 \end{pmatrix}. \qquad (\mathrm{B4})$$

The sum is extended over the three vectors connecting A with nearest B sites in layer μ; in my notation, $\boldsymbol{\delta}_{1,2}^{(\mu)} =$

$\boldsymbol{\delta}_3^{(\mu)} - \mathbf{a}_{1,2}^{(\mu)}$, and $\boldsymbol{\delta}_3^{(\mu)} = \mathbf{R}_{i,B}^{(\mu)} - \mathbf{R}_{i,A}^{(\mu)}$ connects the two sites within the unit cell.

The terms $\hat{T}$ ($\hat{T}^\dagger$) describe interlayer tunneling processes,

$$\hat{T} = \sum_{i,j} \sum_{\alpha,\beta} T_{\alpha\beta}^{ij} \left| \mathbf{R}_{i,\alpha}^{(t)} \right\rangle \left\langle \mathbf{R}_{j,\beta}^{(b)} \right|. \qquad (\mathrm{B5})$$

In a two-center, Slater-Koster-like approximation, $T_{\alpha\beta}^{ij}$ depends only on the relative distance between Wannier centers, so it must admit a Fourier expansion of the form

$$T_{\alpha\beta}^{ij} = A_c \int \frac{d\mathbf{q}}{(2\pi)^2} e^{i\mathbf{q} \cdot \left(\mathbf{R}_{i,\alpha}^{(t)} - \mathbf{R}_{j,\beta}^{(b)} \right)} T(\mathbf{q}), \qquad (\mathrm{B6})$$

where A_c is graphene's unit-cell area. Introducing Bloch states and plugging this last expression into Eq. (B5), we can rewrite the latter as

$$\hat{T} = \sum_{\{\mathbf{b}^{(b)}\}} \sum_{\{\mathbf{b}^{(t)}\}} \sum_{\mathbf{k}_t, \mathbf{k}_b} \sum_{\alpha,\beta} e^{i\mathbf{b}^{(t)} \cdot \boldsymbol{\delta}_\alpha^{(t)} - i\mathbf{b}^{(b)} \cdot \boldsymbol{\delta}_\beta^{(b)}} T\left(\mathbf{k}_b + \mathbf{b}^{(b)} \right) \delta_{\mathbf{k}_b + \mathbf{b}^{(b)}, \mathbf{k}_t + \mathbf{b}^{(t)}} \left| \mathbf{k}_t, \alpha \right\rangle \left\langle \mathbf{k}_b, \beta \right|, \qquad (\mathrm{B7})$$

where the first two sums are on reciprocal vectors of the top and bottom lattices. In deriving this expression, I have made use of the identities

$$\sum_i e^{i(\mathbf{q}-\mathbf{k}_t) \cdot \mathbf{R}_{i,\alpha}^{(t)}} = N\, e^{i\mathbf{b}^{(t)} \cdot \boldsymbol{\delta}_\alpha^{(t)}} \delta_{\mathbf{q}, \mathbf{k}_t + \mathbf{b}^{(t)}}, \qquad (\mathrm{B8a})$$

$$\sum_j e^{-i(\mathbf{q}-\mathbf{k}_b) \cdot \mathbf{R}_{j,\beta}^{(b)}} = N\, e^{-i\mathbf{b}^{(b)} \cdot \boldsymbol{\delta}_\beta^{(b)}} \delta_{\mathbf{q}, \mathbf{k}_b + \mathbf{b}^{(b)}}, \qquad (\mathrm{B8b})$$

where $\boldsymbol{\delta}_\alpha^{(\mu)}$ is the position of site α within a reference unit cell in layer μ, such that $\boldsymbol{\delta}_B^{(\mu)} - \boldsymbol{\delta}_A^{(\mu)} = \boldsymbol{\delta}_3^{(\mu)}$.

Equation (B7) describes scattering events satisfying the general umklapp condition

$$\mathbf{k}_t + n_1' \mathbf{b}_1' + n_2' \mathbf{b}_2' = \mathbf{k}_b + n_1 \mathbf{b}_1 + n_2 \mathbf{b}_2, \qquad (\mathrm{B9})$$

where n_i, n_i' are integers. Here I have introduced the notation of the main text, namely, $\mathbf{b}_i$ and $\mathbf{b}_i' = R(\theta)\mathbf{b}_i$ are the primitive vectors of the reciprocal lattice of the bottom and top layers, respectively; similarly, I write $\boldsymbol{\delta}_\alpha^{(b)} \equiv \boldsymbol{\delta}_\alpha$, $\boldsymbol{\delta}_\alpha^{(t)} = R(\theta)\boldsymbol{\delta}_\alpha + \mathbf{u}$ where, in addition to the relative rotation along a common hexagon center, I consider a relative displacement $\mathbf{u}$ of the top with respect to the bottom layer. Commensurate approximants to the electronic structure simplifies the condition in Eq. (B9) such that $n_i = n_i'$, i.e., $\mathbf{k}_t = \mathbf{k}_b + \mathbf{G}$, where $\mathbf{G}$ is a vector of the moiré superlattice; equation (B7) simplifies then to

$$\hat{T} \approx \sum_{\{\mathbf{b}\}} \sum_{\mathbf{k}_b, \mathbf{k}_t} \sum_{\alpha,\beta} e^{i\mathbf{b} \cdot (\boldsymbol{\delta}_\alpha - \boldsymbol{\delta}_\beta) + i\mathbf{b} \cdot \left(R^T \mathbf{u} \right)} T(\mathbf{k}_b + \mathbf{b}) \, \delta_{\mathbf{k}_t, \mathbf{k}_b + (1-R)\mathbf{b}} \left| \mathbf{k}_t, \alpha \right\rangle \left\langle \mathbf{k}_b, \beta \right|. \qquad (\mathrm{B10})$$

The sum in $\{\mathbf{b}\}$ is also restricted in practice, provided that $T(\mathbf{q})$ is a rapidly decaying function of momentum; in particular, $T(\mathbf{q})$ is strongly suppressed for values $|\mathbf{q}| > 1/d$, where d is the separation between layers.[1,2]

For small twist angles, the electronic spectrum will be dominated by the low-energy Dirac bands of decoupled graphene layers lying around the two inequivalent corners (valleys) of the respective Brillouin zones labelled by

$K_{\tau,\mu}$ in Fig. 1; here $\tau = \pm 1$ labels the valleys, which is assumed to be a good quantum number. In a continuum description,[1,2] band dispersion is simplified by expanding the phases in Eq. (B4) for small momenta around $\mathbf{K}_{\tau,\mu}$,

$$\sum_{i=1}^{3} e^{i(\mathbf{K}_{\tau,\mu}+\mathbf{p}_\mu)\cdot\boldsymbol{\delta}_i^{(\mu)}} \approx -\tau\, p_\mu^{(x)} + i p_\mu^{(y)}, \tag{B11}$$

where the x, y components are adapted to the high-symmetry (zig-zag and armchair, respectively) axes of individual graphene layers. As mentioned in the main text, it is convenient to expresses crystalline momenta in a common frame of reference (defined in this case by the high-symmetry axes highlighted in Fig. 1), so the spinor basis in the corresponding sublattice space of each layer has to be rotated, leading to the block-diagonal terms in Eq. (16).

For the tunneling terms, we can, in the same spirit, neglect small deviations from $\mathbf{K}_{\tau,\mu}$ in the argument of $T(\mathbf{q})$. If we include only one harmonic, $t_\perp \equiv T\left(|\mathbf{K}_{\tau,\mu}|\right)$, the sum in $\{\mathbf{b}\}$ has to be restricted to $\mathbf{b} = \mathbf{0}, -\tau\,\mathbf{b}_1, \tau\,\mathbf{b}_2$, corresponding to the three equivalent positions of a given valley within a single-layer Brillouin zone. Equation (B10) simplifies to

$$\hat{T} \approx t_\perp \sum_{\mathbf{p},\tau} \sum_{\alpha,\beta} \left\{ \left[\hat{T}_0^{(\tau)}\right]_{\alpha,\beta} |\mathbf{k}_t = \mathbf{k}_b, \alpha\rangle \langle \mathbf{k}_b = \mathbf{K}_{\tau,b} + \mathbf{p}, \beta| + \left[\hat{T}_1^{(\tau)}\right]_{\alpha,\beta} e^{i\tau\mathbf{G}_1\cdot\tilde{\mathbf{u}}} |\mathbf{k}_t = \mathbf{k}_b - \tau\,\mathbf{G}_1, \alpha\rangle \langle \mathbf{k}_b = \mathbf{K}_{\tau,b} + \mathbf{p}, \beta| \right.$$
$$\left. + \left[\hat{T}_2^{(\tau)}\right]_{\alpha,\beta} e^{-i\tau\mathbf{G}_2\cdot\tilde{\mathbf{u}}} |\mathbf{k}_t = \mathbf{k}_b + \tau\,\mathbf{G}_2, \alpha\rangle \langle \mathbf{k}_b = \mathbf{K}_{\tau,b} + \mathbf{p}, \beta| \right\}, \tag{B12}$$

where I have used the relation in Eq. (4) along with the identity in footnote 31, so that $\mathbf{b}\cdot(R^T\mathbf{u}) = -\mathbf{G}\cdot\tilde{\mathbf{u}}$. Equation (B12) transformed back to real space corresponds to Eq. (19), with matrices $\hat{T}_i$ given in Eqs. (20).

1. Symmetry-adapted electronic operators

In order to construct phenomenological couplings within the low-energy sector of the spectrum, it is convenient to introduce a basis of electronic operators adapted to the irreducible representations (irreps) of the point-group symmetry D_6 of the continuum model. The internal Hilbert space of the continuum model is spanned by spin, sublattice, layer, and valley degrees of freedom. I am going to ignore the spin since relativistic corrections are weak. Note that neither sublattice nor layer are good quantum numbers, since inter-layer hopping terms mix them, but we can still refer to them to label the transformation properties of the Bloch wave function around $\tilde{\kappa}_\eta$ points. Valley $\tau = \pm 1$ and layer $\mu = \pm 1$ indices label four Dirac crossings reminiscent of the Dirac points of two decoupled graphene layers, while sublattice indices span the associated subspace of Dirac doublets at each crossing. Electronic operators can be written in a basis of Pauli matrices acting on each subspace: $\hat{\sigma}_i$ for sublattice, $\hat{\tau}_i$ for valley, and $\hat{\mu}_i$ for layer. Since we are interested in the coupling to long-wavelength phason fluctuations on the scale of the moiré period, I am going to restrict the analysis to diagonal operators in valley and layer (or $\tilde{\kappa}_\eta$-point) numbers with the pertinent identifications $\hat{\tau}_z \to \tau$, $\hat{\mu}_z \to \mu$. I find convenient, however, to keep the matrix notation here, because the transformation rules of electronic operators are more easily identified from the Pauli-matrix algebra.

For example, elementary rotations along the 6-fold principal axis are implemented by unitary operators

$$\mathcal{C}_{2z} : \hat{\tau}_x\hat{\sigma}_x, \tag{B13a}$$
$$\mathcal{C}_{3z} : e^{\frac{i2\pi}{3}\hat{\ell}_z}, \tag{B13b}$$

where $\hat{\ell}_z = \frac{1}{2}\hat{\tau}_z\hat{\sigma}_z$. Two-fold rotations along x and y axes are given by

$$\mathcal{C}_{2x} : \hat{\mu}_x\hat{\sigma}_x, \tag{B14a}$$
$$\mathcal{C}_{2y} : \hat{\mu}_x\hat{\tau}_x. \tag{B14b}$$

The rest of point-group operations follow from matrix multiplication. Finally, time-reversal symmetry is implemented by the anti-unitary operator

$$\mathcal{T} : \hat{\tau}_x\mathcal{K}, \tag{B15}$$

where $\mathcal{K}$ denotes complex conjugation.

All the possible combinations of valley- and layer-diagonal operators can be classified according to the irreps of D_6 and time-reversal symmetry. The result is summarized in Tab. I, along with the characters of D_6. It is worth emphasizing that there are two ways to lift the Dirac degeneracies without breaking time-reversal symmetry: by breaking both $\mathcal{C}_{2z}$ and $\mathcal{C}_{2y}$ symmetries ($\hat{\mu}_z\hat{\sigma}_z \sim B_1$), i.e., by removing the physical equivalence of both sublattices and layers (a staggered potential of opposite sign in each layer), or by breaking $\mathcal{C}_{2z}$ and $\mathcal{C}_{2x}$ symmetries ($\hat{\sigma}_z \sim B_2$), i.e., by removing only the physical equivalence between sublattices (the same staggered potential in both layers). Only the latter leads to bands with nonzero valley-Chern number. The competition between these two mass terms give rise to the phase diagram discussed in Ref. 16.

The phenomenological electron-phason Hamiltonian in Eq. (24) consists of all the possible invariants formed

Electronic operators ($\mathcal{T}$)	Irrep	$\mathcal{E}$	$\mathcal{C}_{2z}$	$2\,\mathcal{C}_{3z}$	$2\,\mathcal{C}_{6z}$	$3\,\mathcal{C}_{2x}$	$3\,\mathcal{C}_{2y}$
$\hat{\mu}_z\hat{\tau}_z\hat{\sigma}_z\,(-)$	A_1	1	1	1	1	1	1
$\hat{\mu}_z\,(+),\ \hat{\tau}_z\hat{\sigma}_z\,(-)$	A_2	1	1	1	1	-1	-1
$\hat{\tau}_z\,(-),\ \hat{\mu}_z\hat{\sigma}_z\,(+)$	B_1	1	-1	1	- 1	1	-1
$\hat{\sigma}_z\,(+),\ \hat{\mu}_z\hat{\tau}_z\,(-)$	B_2	1	-1	1	-1	-1	1
$\begin{pmatrix}\hat{\tau}_z\hat{\sigma}_x\\ \hat{\sigma}_y\end{pmatrix}(-),\ \begin{pmatrix}-\hat{\mu}_z\hat{\sigma}_y\\ \hat{\mu}_z\hat{\tau}_z\hat{\sigma}_x\end{pmatrix}(-)$	E_1	2	-2	-1	1	0	0
$\begin{pmatrix}\hat{\sigma}_x\\ \hat{\tau}_z\hat{\sigma}_y\end{pmatrix}(+),\ \begin{pmatrix}-\hat{\mu}_z\hat{\tau}_z\hat{\sigma}_y\\ \hat{\mu}_z\hat{\sigma}_x\end{pmatrix}(+)$	E_2	2	2	-1	-1	0	0

TABLE I: Classification of electronic valley- and layer-diagonal operators according to irreducible representations (irreps) of D_6 and parity (± 1, even/odd) under timer-reversal symmetry.

from combinations of the phason-field derivatives in Eqs. (11) with the corresponding electronic operators transforming under the same irrep. Here I should note that the phason field transforms as $\tilde{\mathbf{u}} = (\tilde{u}_x, \tilde{u}_y) \sim E_1$, while the relative displacement between layers transforms as $(-u_y, u_x) \sim E_1$ since, recall, $\mathcal{C}_{2x,y}$ rotations exchange the layers. The relation between these two transformation rules can be understood from the twist (second, dominant term) in Eq. (4) or, with the account of lattice relaxation, the fact that solitons involve relative shear between layers.

2. Perturbative calculation of the electron-phason matrix element

The Hamiltonian of the continuum model can be diagonalized in a basis of Bloch states by restricting the values of crystalline momentum to the first Brillouin zone of the moiré reciprocal lattice and introducing new (band) quantum numbers associated with different copies separated by momenta in $\{\mathbf{G}\}$. In this process, we can absorb a uniform phason field as a phase in the new electronic basis, so the spectrum remains invariant under translations of the moiré pattern. It is useful to consider the positions of the valleys folded into the first moiré Brillouin zone,

$$\mathbf{K}_{\pm,t} = \pm\frac{2\mathbf{G}_2 + \mathbf{G}_1}{3}\ (\equiv \tilde{\kappa}_\pm)\,, \tag{B16a}$$

$$\mathbf{K}_{\pm,b} = \pm\frac{2\mathbf{G}_1 + \mathbf{G}_2}{3}\ (\equiv \tilde{\kappa}_\mp)\,. \tag{B16b}$$

This folding scheme applies to type-I[15] or sublattice-exchange odd[29] commensurate approximants, but once the first sum in Eq. (B10) is restricted to one harmonic, this choice is inconsequential.[30]

Next, we consider the dispersion of low-energy electronic states with momentum $\mathbf{p}$ around these points, in the previous notation, $\mathbf{k}_t = \mathbf{K}_{\tau,t} + \mathbf{p}$ (defining the top layer sector, $\mu = +1$), and $\mathbf{k}_b = \mathbf{K}_{\tau,b} + \mathbf{p}$ (corresponding to the bottom layer sector, $\mu = -1$). For each moiré Brillouin zone in a given layer I am going to truncate the number of copies in the opposite layer to three, so each (decoupled) sector labelled by valley and layer numbers is described by a 8×8 matrix Hamiltonian of the form

$$\hat{\mathcal{H}}_{8\times 8}^{(\tau,\mu)} = \hat{\mathcal{H}}_{\mathbf{p}}^{(\tau,\mu)} + \hat{\mathcal{U}}^{(\tau,\mu)}. \tag{B17}$$

The first term reads just (in block form)

$$\hat{\mathcal{H}}_{\mathbf{p}}^{(\tau,+1)} = \begin{pmatrix} \hbar v_F\,\mathbf{\Sigma}^{(\tau,+1)}\cdot\mathbf{p} & 0 & 0 & 0 \\ 0 & \hbar v_F\,\mathbf{\Sigma}^{(\tau,-1)}\cdot\mathbf{p} & 0 & 0 \\ 0 & 0 & \hbar v_F\,\mathbf{\Sigma}^{(\tau,-1)}\cdot\mathbf{p} & 0 \\ 0 & 0 & 0 & \hbar v_F\,\mathbf{\Sigma}^{(\tau,-1)}\cdot\mathbf{p} \end{pmatrix} \tag{B18}$$

for the top-layer sector, and similarly for the bottom layer,

$$\hat{\mathcal{H}}_{\mathbf{p}}^{(\tau,-1)} = \begin{pmatrix} \hbar v_F\,\mathbf{\Sigma}^{(\tau,+1)}\cdot\mathbf{p} & 0 & 0 & 0 \\ 0 & \hbar v_F\,\mathbf{\Sigma}^{(\tau,+1)}\cdot\mathbf{p} & 0 & 0 \\ 0 & 0 & \hbar v_F\,\mathbf{\Sigma}^{(\tau,+1)}\cdot\mathbf{p} & 0 \\ 0 & 0 & 0 & \hbar v_F\,\mathbf{\Sigma}^{(\tau,-1)}\cdot\mathbf{p} \end{pmatrix}. \tag{B19}$$

From this point on, I am going to neglect the rotation of the spinor basis, which restores the electron-hole symmetry of the spectrum (the error scales with θ^2); while these matrices read the same now, note that they are expressed in a

different basis. The second term in Eq. (B17) reads

$$\hat{\mathcal{U}}^{(\tau,+1)} = \begin{pmatrix} 0 & t_\perp \hat{T}_0^{(\tau)} & t_\perp \hat{T}_1^{(\tau)} & t_\perp \hat{T}_2^{(\tau)} \\ t_\perp \hat{T}_0^{(\tau)} & \hbar v_F \, \boldsymbol{\Sigma}^{(\tau)} \cdot (\mathbf{K}_{\tau,t} - \mathbf{K}_{\tau,b}) & 0 & 0 \\ t_\perp \hat{T}_1^{(\tau)} & 0 & \hbar v_F \, \boldsymbol{\Sigma}^{(\tau)} \cdot (\mathbf{K}_{\tau,t} - \mathbf{K}_{\tau,b} + \tau\mathbf{G}_1) & 0 \\ t_\perp \hat{T}_2^{(\tau)} & 0 & 0 & \hbar v_F \, \boldsymbol{\Sigma}^{(\tau)} \cdot (\mathbf{K}_{\tau,t} - \mathbf{K}_{\tau,b} - \tau\mathbf{G}_2) \end{pmatrix}$$
(B20)

for the top-layer sector, and

$$\hat{\mathcal{U}}^{(\tau,-1)} = \begin{pmatrix} \hbar v_F \, \boldsymbol{\Sigma}^{(\tau)} \cdot (\mathbf{K}_{\tau,b} - \mathbf{K}_{\tau,t}) & 0 & 0 & t_\perp \hat{T}_0^{(\tau)} \\ 0 & \hbar v_F \, \boldsymbol{\Sigma}^{(\tau)} \cdot (\mathbf{K}_{\tau,b} - \mathbf{K}_{\tau,t} - \tau\mathbf{G}_1) & 0 & t_\perp \hat{T}_1^{(\tau)} \\ 0 & 0 & \hbar v_F \, \boldsymbol{\Sigma}^{(\tau)} \cdot (\mathbf{K}_{\tau,b} - \mathbf{K}_{\tau,t} + \tau\mathbf{G}_2) & t_\perp \hat{T}_2^{(\tau)} \\ t_\perp \hat{T}_0^{(\tau)} & t_\perp \hat{T}_1^{(\tau)} & t_\perp \hat{T}_2^{(\tau)} & 0 \end{pmatrix}$$
(B21)

for the bottom-layer sector. The low-energy subspace is defined by the zero-energy eigenstates of $\hat{\mathcal{U}}^{(\tau,\mu)}$, specifically,

$$|\psi_{A,\tau,+1}\rangle = \frac{1}{\sqrt{1+6\alpha^2}} \begin{pmatrix} 1 \\ 0 \\ i\tau\alpha \\ -i\tau\alpha \\ i\tau\alpha \\ -i\tau\alpha\, e^{-i\tau\frac{2\pi}{3}} \\ i\tau\alpha \\ -i\tau\alpha\, e^{i\tau\frac{2\pi}{3}} \end{pmatrix}, \quad |\psi_{B,\tau,+1}\rangle = \frac{1}{\sqrt{1+6\alpha^2}} \begin{pmatrix} 0 \\ 1 \\ i\tau\alpha \\ -i\tau\alpha \\ i\tau\alpha\, e^{i\tau\frac{2\pi}{3}} \\ -i\tau\alpha \\ i\tau\alpha\, e^{-i\tau\frac{2\pi}{3}} \\ -i\tau\alpha \end{pmatrix},$$
(B22a)

$$|\psi_{A,\tau,-1}\rangle = \frac{1}{\sqrt{1+6\alpha^2}} \begin{pmatrix} -i\tau\alpha \\ i\tau\alpha \\ -i\tau\alpha \\ i\tau\alpha\, e^{-i\tau\frac{2\pi}{3}} \\ -i\tau\alpha \\ i\tau\alpha\, e^{i\tau\frac{2\pi}{3}} \\ 1 \\ 0 \end{pmatrix}, \quad |\psi_{B,\tau,-1}\rangle = \frac{1}{\sqrt{1+6\alpha^2}} \begin{pmatrix} -i\tau\alpha \\ i\tau\alpha \\ -i\tau\alpha\, e^{i\tau\frac{2\pi}{3}} \\ i\tau\alpha \\ -i\tau\alpha\, e^{-i\tau\frac{2\pi}{3}} \\ i\tau\alpha \\ 0 \\ 1 \end{pmatrix}.$$
(B22b)

These states form a basis for the D_6 irreps introduced in the previous subsection. The rest of eigenstates are separated by energies $E_i = \pm\hbar v_F |\tilde{\kappa}|, \pm\hbar v_F\sqrt{1+6\alpha^2}\,|\tilde{\kappa}|$ (the last ones are 2-fold degenerate in this approximation).

Next, I am going to produce a $\mathbf{k}\cdot\mathbf{p}$ expansion of the Hamiltonian by projecting-out the high-energy states in a Löwdin perturbative scheme.[68] Let me introduce first the following projection operators:

$$\hat{P}_0^{(\tau,\mu)} \equiv \sum_{\alpha=A,B} |\psi_{\alpha,\tau,\mu}\rangle \langle\psi_{\alpha,\tau,\mu}|,$$
(B23a)

$$\hat{1} - \hat{P}_0^{(\tau,\mu)} = \sum_i \hat{P}_i^{(\tau,\mu)},$$
(B23b)

where i labels the high-energy eigenstates. Up to second order in perturbation theory, where any term in the Hamiltonian of Eq. (B17) (generically written as $\hat{V}^{(\tau,\mu)}$) is treated as a perturbation to $\hat{\mathcal{U}}^{(\tau,\mu)}$, we have

$$\hat{\mathcal{H}}^{(\tau,\mu)} \approx \hat{P}_0^{(\tau,\mu)} \hat{V}^{(\tau,\mu)} \hat{P}_0^{(\tau,\mu)} - \sum_i \hat{P}_0^{(\tau,\mu)} \hat{V}^{(\tau,\mu)} \frac{\hat{P}_i^{(\tau,\mu)}}{E_i} \hat{V}^{(\tau,\mu)} \hat{P}_0^{(\tau,\mu)}.$$
(B24)

First-order perturbation theory in $\hat{\mathcal{H}}_{\mathbf{p}}^{(\tau,\mu)}$ leads to the first term in Eq. (23) with a new Fermi velocity v_F^* reduced by a factor $(1-3\alpha^2)/(1+6\alpha^2)$, the result first obtained by Bistritzer and MacDonald in Ref. 2. For the electron-phason

coupling, we should consider first its expression in the 8×8 Hilbert space; from Eq. (21), we have

$$\hat{V}_{\text{e-ph}}^{(\tau,+1)} = \delta\hat{T}^{(\tau,+1)} + \text{h.c.} = \tau\, t_\perp \begin{pmatrix} 0 & 0 & i\,\mathbf{G}_1\cdot\tilde{\mathbf{u}}\,\hat{T}_1^{(\tau)} & -i\,\mathbf{G}_2\cdot\tilde{\mathbf{u}}\,\hat{T}_2^{(\tau)} \\ 0 & 0 & 0 & 0 \\ -i\,\mathbf{G}_1\cdot\tilde{\mathbf{u}}\,\hat{T}_1^{(\tau)} & 0 & 0 & 0 \\ i\,\mathbf{G}_2\cdot\tilde{\mathbf{u}}\,\hat{T}_2^{(\tau)} & 0 & 0 & 0 \end{pmatrix} \quad \text{(B25)}$$

for the top-layer sector, and

$$\hat{V}_{\text{e-ph}}^{(\tau,-1)} = \delta\hat{T}^{(\tau,-1)} + \text{h.c.} = \tau\, t_\perp \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i\,\mathbf{G}_1\cdot\tilde{\mathbf{u}}\,\hat{T}_1^{(\tau)} \\ 0 & 0 & 0 & -i\,\mathbf{G}_2\cdot\tilde{\mathbf{u}}\,\hat{T}_2^{(\tau)} \\ 0 & -i\,\mathbf{G}_1\cdot\tilde{\mathbf{u}}\,\hat{T}_1^{(\tau)} & i\,\mathbf{G}_2\cdot\tilde{\mathbf{u}}\,\hat{T}_2^{(\tau)} & 0 \end{pmatrix} \quad \text{(B26)}$$

for the bottom layer. First-order perturbation theory gives 0, as expected, since the invariance under translations of the soliton network implies that there must be momentum transfer between electronic states.[69] Treating both $\hat{V}_{\text{e-ph}}^{(\tau,\mu)}$ and $\hat{\mathcal{H}}_{\mathbf{p}}^{(\tau,\mu)}$ up to second order in perturbation theory gives

$$\hat{\mathcal{H}}_{\text{e-ph}}^{(\tau,\mu)}(\mathbf{k},\mathbf{k}') \approx -\sum_i \hat{P}_0^{(\tau,\mu)}\hat{V}_{\text{e-ph}}^{(\tau,\mu)}\frac{\hat{P}_i^{(\tau,\mu)}}{E_i}\hat{\mathcal{H}}_{\mathbf{p}}^{(\tau,\mu)}\hat{P}_0^{(\tau,\mu)} + \text{h.c.}$$

$$\approx -\mu\tau\frac{6\,\alpha\,t_\perp\left(1-3\alpha^2\right)}{\left(1+6\alpha^2\right)^2}\left(\mathbf{p}\cdot\tilde{\mathbf{u}}\right)\hat{\sigma}_z, \quad \text{(B27)}$$

where $\mathbf{p}$ must be interpreted now as the average momentum during a electron-phason scattering event, $\mathbf{p} \equiv (\mathbf{k}'+\mathbf{k})/2$, with $\mathbf{k}$, $\mathbf{k}'$ labelling the initial and final states, respectively.

Equation (B27) may look odd at first glance, for it is difficult to recognize the symmetry-allowed couplings in the phenomenological expansion of Eq. (24) from this expression. Note, however, that Eq. (B27) is compatible with time-reversal and D_6 point-group symmetries and, in fact, the combination of valley, sublattice and layer indices is such that when projected over a band state, i.e., and eigenstate of the first term in Eq. (23),

$$|\zeta,\tau,\mu,\mathbf{k}\rangle = \frac{e^{-\frac{i\tau\theta_{\mathbf{k}}}{2}}}{\sqrt{2}}|\psi_{A,\tau,\mu}\rangle + \tau\zeta\frac{e^{\frac{i\tau\theta_{\mathbf{k}}}{2}}}{\sqrt{2}}|\psi_{B,\tau,\mu}\rangle, \quad \text{(B28)}$$

where $\zeta = \pm 1$ labels the electron/hole bands, Eq. (B27) produces the expected matrix element of g_{A_2} coupling. Note first that for intra-band processes we have

$$\langle\zeta,\tau,\mu,\mathbf{k}'|\hat{\mathcal{H}}_{\text{e-ph}}^{(\tau,\mu)}|\zeta,\tau,\mu,\mathbf{k}\rangle \approx \quad \text{(B29)}$$

$$-i\mu\frac{3\,\alpha\,t_\perp\left(1-3\alpha^2\right)}{\left(1+6\alpha^2\right)^2}\sin\left(\frac{\theta_{\mathbf{k}'}-\theta_{\mathbf{k}}}{2}\right)\left(\mathbf{k}+\mathbf{k}'\right)\cdot\tilde{\mathbf{u}}.$$

When the phason field is decomposed in longitudinal and transverse components, the former give rise to scalar products of the form $i\left(\mathbf{k}+\mathbf{k}'\right)\cdot\hat{\mathbf{q}}$, while for transverse modes we have $i\left(\mathbf{k}+\mathbf{k}'\right)\cdot\left(\hat{\mathbf{z}}\times\hat{\mathbf{q}}\right)$; here $\hat{\mathbf{q}}$ is a unit vector along the transferred momentum, $\mathbf{q} = \mathbf{k}'-\mathbf{k}$. For
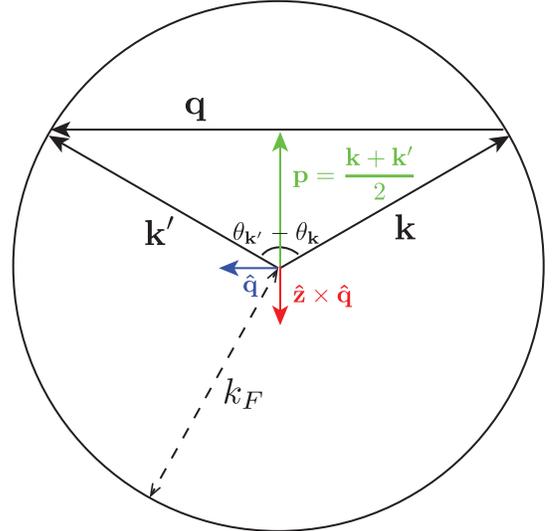

FIG. 6: Kinematics of quasi-elastic scattering events within a Fermi circle centered at one of the corners of the moiré Brillouin zone.

quasi-elastic processes ($|\mathbf{k}'| = |\mathbf{k}|$) illustrated in Fig. 6, we have

$$\left(\mathbf{k}+\mathbf{k}'\right)\cdot\hat{\mathbf{q}} = 0, \quad \text{(B30a)}$$

$$\left(\mathbf{k}+\mathbf{k}'\right)\cdot\left(\hat{\mathbf{z}}\times\hat{\mathbf{q}}\right) = -2\,|\mathbf{k}|\cos\left(\frac{\theta_{\mathbf{k}'}-\theta_{\mathbf{k}}}{2}\right). \quad \text{(B30b)}$$

Collecting all the pieces, I arrive at the following expression for the matrix element of the coupling with transverse phasons,

$$\varpi_{T,\tau}^{\mu}(\mathbf{q},\mathbf{k},\mathbf{k'}) = -\mu \, \frac{6 \, \alpha \, t_\perp}{1+6\alpha^2} \left(\frac{v_F^*}{v_F}\right) |\mathbf{k}| \sin\left(\frac{\theta_{\mathbf{k'}}-\theta_{\mathbf{k}}}{2}\right) \cos\left(\frac{\theta_{\mathbf{k'}}-\theta_{\mathbf{k}}}{2}\right), \tag{B31}$$

where I have regrouped some factors in v_F^*/v_F. This is indeed the matrix element of the coupling with $(\boldsymbol{\nabla}\times\tilde{\mathbf{u}})_z$ since for quasi-elastic processes $|\mathbf{q}| = 2|\mathbf{k}|\sin(\frac{\theta_{\mathbf{k'}}-\theta_{\mathbf{k}}}{2})$.

We end up then with the identification of the phenomenological parameter in Eq. (27).

[1] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Phys. Rev. Lett. **99**, 256802 (2007).
[2] R. Bistritzer and A. H. MacDonald, Proc. Natl. Acad. Sci. **108**, 12233 (2011).
[3] 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, Nature **556**, 80 (2018).
[4] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Nature **556**, 43 (2018).
[5] M. Yankowitz, S. Chen, H. Polshyn, K. Watanabe, T. Taniguchi, D. Graf, A. F. Young, and C. R. Dean, Science **363**, 1059 (2019).
[6] A. L. Sharpe, E. J. Fox, A. W. Barnard, J. Finney, K. Watanabe, T. Taniguchi, M. A. Kastner, and D. Goldhaber-Gordon, Science **365**, 605 (2019).
[7] 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, D. K. Efetov, arXiv:1903.06513.
[8] M. Serlin, C. L. Tschirhart, H. Polshyn, Y. Zhang, J. Zhu, K. Watanabe, T. Taniguchi, L. Balents, and A. F. Young, arXiv:1907.00261.
[9] H. Yoo, R. Engelke, S. Carr, S. Fang, K. Zhang, P. Cazeaux, S. H. Sung, R. Hovden, A. W. Tsen, T. Taniguchi, K. Watanabe, G.-C. Yi, M. Kim, M. Luskin, and E. B. Tadmor, and E. Kaxiras, and P. Kim, Nat. Mater. **18**, 448 (2019).
[10] S. Huang, K. Kim, D. K. Efimkin, T. Lovorn, T. Taniguchi, K. Watanabe, A. H. MacDonald, E. Tutuc, and B. J. LeRoy, Phys. Rev. Lett. **121**, 037702 (2018).
[11] S. S. Sunku, G. X. Ni, B. Y. Jiang, H. Yoo, A. Sternbach, A. S. McLeod, T. Stauber, L. Xiong, T. Taniguchi, K. Watanabe, P. Kim, M. M. Fogler, and D. N. Basov, Science **362**, 1153 (2018).
[12] J. S. Aldena, A. W. Tsena, P. Y. Huanga, R. Hovdena, L. Brownb, J. Parkb, D. A. Mullera, and P. L. McEuen, Proc. Natl. Acad. Sci. **110**, 11256 (2013).
[13] F. Zhang, A. H. MacDonald, and E. J. Mele, Proc. Natl. Acad. Sci. **110**, 10546 (2013).
[14] H. C. Po, L. Zou, A. Vishwanath, and T. Senthil, Phys. Rev. X **8**, 031089 (2018).
[15] L. Zou, H. C. Po, A. Vishwanath, and T. Senthil, Phys. Rev. B **98**, 085435 (2018).
[16] N. Bultinck, S. Chatterjee, and M. P. Zaletel, arXiv:1901.08110.
[17] Y.-H. Zhang, D. Mao, and T. Senthil, arXiv:1901.08209 .
[18] M. Xie and A. H. MacDonald, arXiv:1812.04213.
[19] F. Guinea and N. R. Walet, Proc. Natl. Acad. Sci. **115**, 13174 (2018).
[20] A. Kerelsky, L. McGilly, D. M. Kennes, L. Xian, M. Yankowitz, S. Chen, K. Watanabe, T. Taniguchi, J. Hone, C. Dean, A. Rubio, and A. N. Pasupathy, Nature **572**, 95 (2019).
[21] 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, arXiv:1901.02997.
[22] Y. Jiang, J. Mao, X. Lai, K. Watanabe, T. Taniguchi, K. Haule, and E. Y. Andrei, arXiv:1904.10153.
[23] Y. Cao, D. Chowdhury, D. Rodan-Legrain, O. Rubies-Bigordà, K. Watanabe, T. Taniguchi, T. Senthil, and P. Jarillo-Herrero, arXiv:1901.03710.
[24] H. Polshyn, M. Yankowitz, S. Chen, Y. Zhang, K. Watanabe, T. Taniguchi, C. R. Dean, and A. F. Young, arXiv:1902.00763.
[25] N. Ray, M. Fleischmann, D. Weckbecker, S. Sharma, O. Pankratov, and S. Shallcross, Phys. Rev. B **94**, 245403 (2016).
[26] F. Wu, E. Hwang, and S. Das Sarma, Phys. Rev. B **99**, 165112 (2019).
[27] I. Yudhistira, N. Chakraborty, G. Sharma, D. Y. H. Ho, E. Laksono, O. P. Sushkov, G. Vignale, and S. Adam, Phys. Rev. B **99**, 140302(R) (2019).
[28] B. Lian, Z. Wang, and B. Andrei Bernevig, Phys. rev. Lett. **122**, 257002 (2019).
[29] E. J. Mele, Phys. Rev. B **81**, 161405(R) (2010).
[30] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Phys. Rev. B **86**, 155449 (2012).
[31] Note that $\mathbf{G}_i \cdot \mathbf{R}_j = 2\pi\delta_{ij}$ by definition; equation (2) follows from the identity $(1-R)^T \cdot (1-R^{-1})^{-1} = 1$.
[32] D. Levibe, T. C. Lubensky, S. Ostlund, S. Ramaswamy, P. J. Steinhardt, and J. Toner, Phys. Rev. Lett. **54**, 1520 (1985).
[33] The condition of beating pattern maxima is now $\boldsymbol{\Delta}(n\,\mathbf{R}_1 + m\,\mathbf{R}_2 + \tilde{\mathbf{u}}) = n\,\mathbf{a}_1 + m\,\mathbf{a}_2 - R^{-1}\mathbf{u}$; inverting this relation leads to Eq. (4).
[34] A. M. Popov, I. V. Lebedeva, A. A. Knizhnik, Y. E. Lozovik, and B. V. Potapkin, Phys. Rev. B **84**, 045404 (2011).
[35] X. Gong and E. J. Mele, Phys. Rev. B **89**, 121415(R) (2014).
[36] N. N. T. Nam and M. Koshino, Phys. Rev. B **96**, 075311 (2017).
[37] C. Lee, X. Wei, J. W. Kysar, and J. Hone, Science **321**, 385 (2008).
[38] S. Carr, D. Massatt, S. B. Torrisi, P. Cazeaux, M. Luskin, and E. Kaxiras, Phys. Rev. B **98**, 224102 (2018).
[39] These one-dimensional solutions are only approximations to the domain walls between AB and BA-stacked areas,

which contain both tensile and shear components in general. In the limit $L_M \ll \ell$, the latter dominates, as reflected by the strongest softening of the transverse phonon in the calculation of Fig. 5.

[40] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics*, (Cambridge University Press, Cambridge, 2000).

[41] If we lock the coordinate system to a beating pattern maxima then $\gamma/3 = 0$ modulo 2π.

[42] As we showed before, beating pattern translations result from relative translations of one layer with respect to the other. In an incommensurate structure, this is a soft mode. Beating pattern rotations with respect to the preferential direction imposed by the superlattice potential results from expanding/compressing one layer with respect to the other, which always costs energy.

[43] The bulk modulus $\tilde{\lambda} + \tilde{\mu} \approx \sqrt{3}V\ell/L_M$ remains positive and corresponds, roughly, to the variation of the free-energy density of the soliton network with the moiré period.

[44] S. N. Coppersmith, D. S. Fisher, B. I. Halperin, P. A. Lee, and W. F. Brinkman, Phys. Rev. Lett. **46**, 549 (1981); Phys. Rev. B **25**, 349 (1982).

[45] Z. Bi, N. F. Q. Yuan, and L. Fu, Phys. Rev. B **100**, 035448 (2019).

[46] J. Villain, Surf. Sci. **97**, 219 (1980).

[47] V. L. Pokrovsky and A. L. Talapov, Phys. Rev. Lett. **42**, 65; Zh. Eksp. Teor. Fiz. **78**, 269 (1979) [Sov. Phys. JETP **51**, 134].

[48] A. I. Cocemasov, D. L. Nika, and A. A. Balandin, Phys. Rev. B **88**, 035428 (2013).

[49] M. Koshino, New J. Phys. **17**, 015014 (2015).

[50] M. Koshino, N. F. Q. Yuan, T. Koretsune, M. Ochi, K. Kuroki, and L. Fu, Phys. Rev. X **8**, 031087 (2018).

[51] F. Guinea and N. R. Walet, Phys. Rev. B **99**, 205134 (2019).

[52] M. F. Bishop and A. W. Overhauser, Phys. Rev. B **23**, 3638 (1981).

[53] J. R. Wallbank, R. K. Kumar, M. Holwill, Z. Wang, G. H. Auton, J. Birkbeck, A. Mishchenko, L. A. Ponomarenko, K. Watanabe, T. Taniguchi, K. S. Novoselov, I. L. Aleiner, A. K. Geim, and V. I. Fal'ko, Nat. Phys. **15**, 32 (2019). Umklapp scattering by phonons is discussed in the supplementary material appended to the electronic version of this article.

[54] D. Culcer and R. Winkler, Phys. Rev. B **79**, 165422 (2019).

[55] M. Trushin, J. Kailasvuori, J. Schliemann, and A. H. MacDonald, Phys. Rev. B **82**, 155308 (2010).

[56] J. M. Ziman, *Electrons and Phonons: The Theory of Transport Phenomena in Solids*, (Oxford Univesity Press, London, 1960).

[57] H. Ochoa, E. V. Castro, M. I. Katsnelson, and F. Guinea, Phys. Rev. B **83**, 235416 (2011).

[58] H. Suzuura and T. Ando, Phys. Rev. B **65**, 235412 (2002).

[59] J. González and T. Stauber, arXiv:1903.01376.

[60] T. M. Rice, N. J. Robinson, and A. M. Tsvelik, Phys. Rev. B **96**, 220502(R) (2017).

[61] H. Isobe, N. F. Q. Yuan, and L. Fu, Phys. Rev. X **8**, 041041 (2018).

[62] J. González and T. Stauber, Phys. Rev. Lett. **122**, 026801 (2019).

[63] A. I. Larkin, Zh. Eksp. Teor. Fiz. **58**, 1466 (1970) [Sov. Phys. JETP **31**, 784].

[64] B. Sachs, T. O. Wehling, M. I. Katsnelson, and A. I. Lichtenstein, Phys. Rev. B **84**, 195414 (2011).

[65] H. Fukuyarna and P. A. Lee, Phys. Rev. B **17**, 535 (1978).

[66] A. I. Larkin and Yu. N. Ovchinnikov, J. Low Temp. Phys. **34**, 409 (1979).

[67] M. Koshino and Y.-W. Son, arXiv:1905.09660.

[68] P.-O. Löwdin, J. Chem. Phys. **19**, 1396 (1951).

[69] For this reason, the result in first-order perturbation theory is robust when the hopping amplitudes between different sublattices are changed.