2D Materials

CrossMark

Universal principles of moiré band structures

Jan Attig[1] , Jinhong Park , Michael M Scherer , Simon Trebst , Alexander Altland
and Achim Rosch

Institute for Theoretical Physics, University of Cologne, 50937 Cologne, Germany
[1] These two authors contributed equally to this work.

E-mail: jinhong@thp.Uni-Koeln.DE

Abstract

Moiré materials provide a highly tunable environment for the realization of band structures with engineered physical properties. Specifically, moiré structures with Fermi surface flat bands—a synthetic environment for the realization of correlated phases—have moiré unit cells containing thousands of atoms and tantalizingly complex band structures. In this paper we show that *statistical principles* go a long way in explaining universal physical properties of these systems. Our approach builds on three conceptual elements: the presence of quantum chaos caused by the effective irregularity of the atomic configurations on short length scales, Anderson localization in momentum space, and the presence of approximate crystalline symmetries. Which of these principles dominates depends on material parameters such as the extension of the Fermi surface or the strength of the moiré lattice potential. The phenomenological consequences of this competition are predictions for the characteristic group velocity of moiré bands, a primary indicator for their average flatness. In addition to these generic features, we identify structures outside the statistical context, notably almost flat bands close to the extrema of the unperturbed spectra, and the celebrated zero energy 'magic angle' flat bands, where the latter require exceptionally fine tuned material parameters.
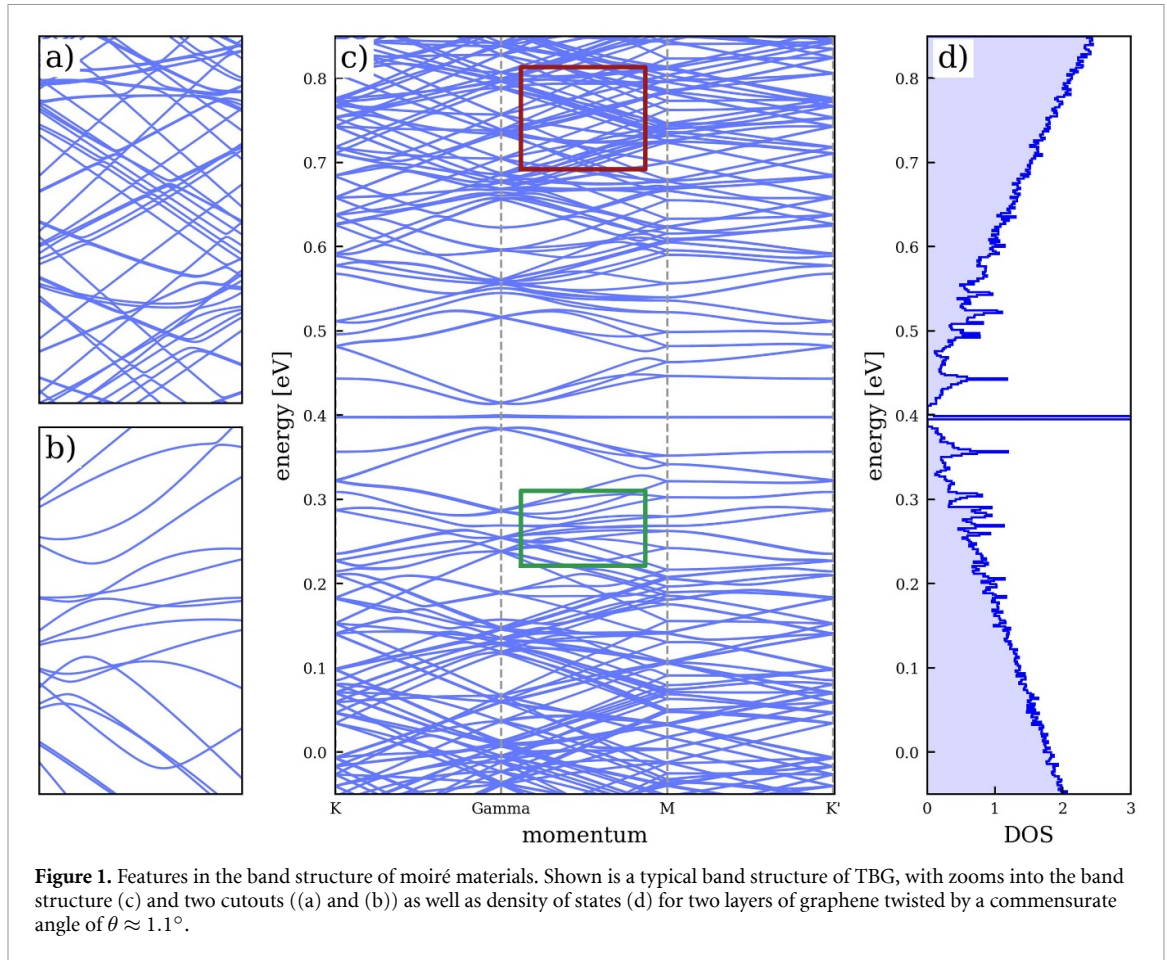
1. Introduction

Sheets of two-dimensional materials stacked at relative twist angles or with a mismatch in lattice constant define a class of quantum matter known as moiré materials. At low twist angles or small lattice constant mismatch, moiré materials can have tens of thousands of atoms in their effective unit cells, and as many energy bands in their Brillouin zones. Controlled variations of twist angles, and/or the (corrugated) van der Waals coupling between layers affords the unique opportunity of band structure engineering [1–8]. The recent realization of almost non-dispersive bands in twisted bilayer graphene (TBG) (and the observation of a wealth of strong correlation effects [9–15] symptomatic for flat band materials) demonstrate the opportunities provided by this type of quantum matter, which besides graphene [16–21] contains hexagonal boron nitride [22, 23], transition metal dichalcogenides [24–26], and others [27, 28] as material platforms.

The tantalizingly complex band structure of moiré materials raises the question for underlying universal principles. For instance, a naked eye inspection of the blow-ups in figure 1 reveals recurrent patterns in the 'spaghetti' of individual energy bands. Most apparent among these are regions with almost linearly dispersive bands ('uncooked spaghetti' in figure 1(a)) interspersed by narrowly avoided crossings, and regions with slack energy bands ('cooked spaghetti' in figure 1(b)) meandering up and down subject to strong band repulsion. The average uniformity of these patterns over wide ranges of momenta and energies suggests that statistical principles are at work. Embedded in these structures we observe anomalous features which clearly are not of statistical nature. Most prominent among these are the celebrated flat bands forming upon fine tuning of twist angles and/or corrugation parameters, see figure 1(c) near the Fermi level of ~ 0.4 eV. In addition to these exceptional flat bands, there are more robust 'super-flat' bands forming next to the band minima of the uncoupled layers, as we will show below.

In this paper, we present a simple semi-phenomenological theory which turns the complexity

Figure 1. Features in the band structure of moiré materials. Shown is a typical band structure of TBG, with zooms into the band structure (c) and two cutouts ((a) and (b)) as well as density of states (d) for two layers of graphene twisted by a commensurate angle of $\theta \approx 1.1°$.

implied by the large number of bands into an advantage and uses it as the basis for a *statistical approach* [29–31]. Our theory builds on three basic principles:

- *Lattice periodicity.* The presence of an effective moiré potential repeating itself over large distance scales defines a lattice structure of periodicity $L \gg 1$ atomic lattice spacings, a. The latter may be looked at in real space or, preferably for our purposes, in momentum space. In that representation, the system is described by a finite lattice of spacing $G_m \sim 1/(La)$ with $\mathcal{O}(L^2)$ sites corresponding to the number of atoms in the moiré unit cell. The moiré potential defines an effective hopping Hamiltonian in this lattice, and the dispersion relation of the unperturbed layers that of an effective on-site potential.
- *Anderson localization.* The aperiodic site-to-site variations of that potential define a source of effective irregularity or quantum *disorder*. For weak moiré potentials, hopping along the effectively one-dimensional equi-potential Fermi surfaces is impeded by the mechanism of Anderson localization (in momentum space). We will see that this manifestation of quantum localization is an efficient promoter of strong energy band dispersion [32].

- *Quantum chaos.* With increasing moiré hopping the Fermi surfaces broaden and eventually turn into quasi two-dimensional structures. The increasing hopping strength driving this development delocalizes wave functions, up to a point that they cover a two-dimensional subset of the lattice almost ergodically. In such regimes, the combination of residual quantum disorder and discrete symmetries characterizing the moiré lattice defines weakly dispersive band structures containing accumulations of almost, but not fully flat bands.

As we will see, the combination of these three elements goes a long way in quantitatively describing the universal features of the moiré band structure. However, it also explains various non-universal features, among them the formation of different types of flat bands or van Hove structures.

We conclude this introductory discussion with a general remark on the lattice structure of the problem. Most moiré systems are *quasi crystals*: the moiré unit cells are not periodically repeated, and translational invariance is absent including at the largest length scales. As we demonstrate in appendix A (see also [3, 33]), this lack of commensurability shows at orders in perturbation theory roughly equal to the linear extension L of the moiré cells. In a series of recent papers [34–37] it has been found that

for flat bands qualitative effects may ensue, notably localization-delocalization transitions driven by the quasi crystalline lack of translational invariance. However, for the 'mesoscopically' large moiré structures, $L = \mathcal{O}(10^2)$, considered in this paper, such effects can safely be ignored and we assume moiré cell periodicity throughout.

2. Setup and general considerations

We consider two-dimensional crystalline systems with lattice constant a—more generally (a_1, a_2)—subject to a perturbation V periodic over distances $a_m = La$, $L \gg 1$ (inset a) of figure 2). This perturbation may be the substrate potential induced by the proximity of a second layer of different chemical composition, or generated by the coupling to a second layer as in TBG. In either case, V defines periodic 'hopping potential' in momentum space, and we discuss them in parallel. More specifically, in the 'extended' Brillouin zone defined by the reciprocal lattice vector of magnitude $G = 2\pi/a$, the potential V defines transition matrix elements between states differing in multiples of the moiré reciprocal lattice vector of magnitude $G_m \equiv 2\pi/a_m = G/L$, and thus defines a momentum space hopping Hamiltonian. Momenta $\boldsymbol{k} \bmod G_m$ are conserved, and on this basis, the spectrum of the system gets organized into $\sim L^2$ bands indexed by $\boldsymbol{k}$. We assume the potential to be strong enough to couple these bands, $V \gtrsim D/L^2$, where D is the total bandwidth and D/L^2 defines the characteristic band spacing. This assumption rests on the efficient coupling of 'sites' in the momentum space moiré lattice over scales $\Delta p \gtrsim G_m$. (In appendix C we discuss in which way this assumption relies on the commensurability properties of the substrate potential and is not entirely innocent.)

In the rest of the paper, we will build on a momentum space lattice picture to obtain information on universal features of the moiré band structure. In particular, we will emphasize connections between the present problem and that of Anderson localization in quasi one-dimensional disordered media. To understand this link, notice that for each realization of $\boldsymbol{k}$ the dispersion of the native two-dimensional material $\epsilon_{\boldsymbol{Q}_n} \equiv \epsilon(\boldsymbol{k} + \boldsymbol{Q}_n)$ is a function on the sites $\boldsymbol{Q}_n \equiv n_1 \boldsymbol{G}_{m,1} + n_2 \boldsymbol{G}_{m,2}$ of the moiré lattice with effectively random site–to–site variations and continuous modulation in $\boldsymbol{k}$.

In momentum space, $\epsilon_{\boldsymbol{Q}_n}$ acts as an *effective potential*. In combination with weak (translationally invariant) hopping $t_\perp$ induced by V, this potential has two principal effects. First, it defines broadened *quasi one-dimensional* Fermi surfaces, centered around the Fermi lines $\epsilon_{\boldsymbol{Q}_n} = \epsilon_F$ of the unperturbed system. More precisely, typical wave functions will probe a region of width $\sim t_\perp/(v_F G_m)$ lattice spacings around these
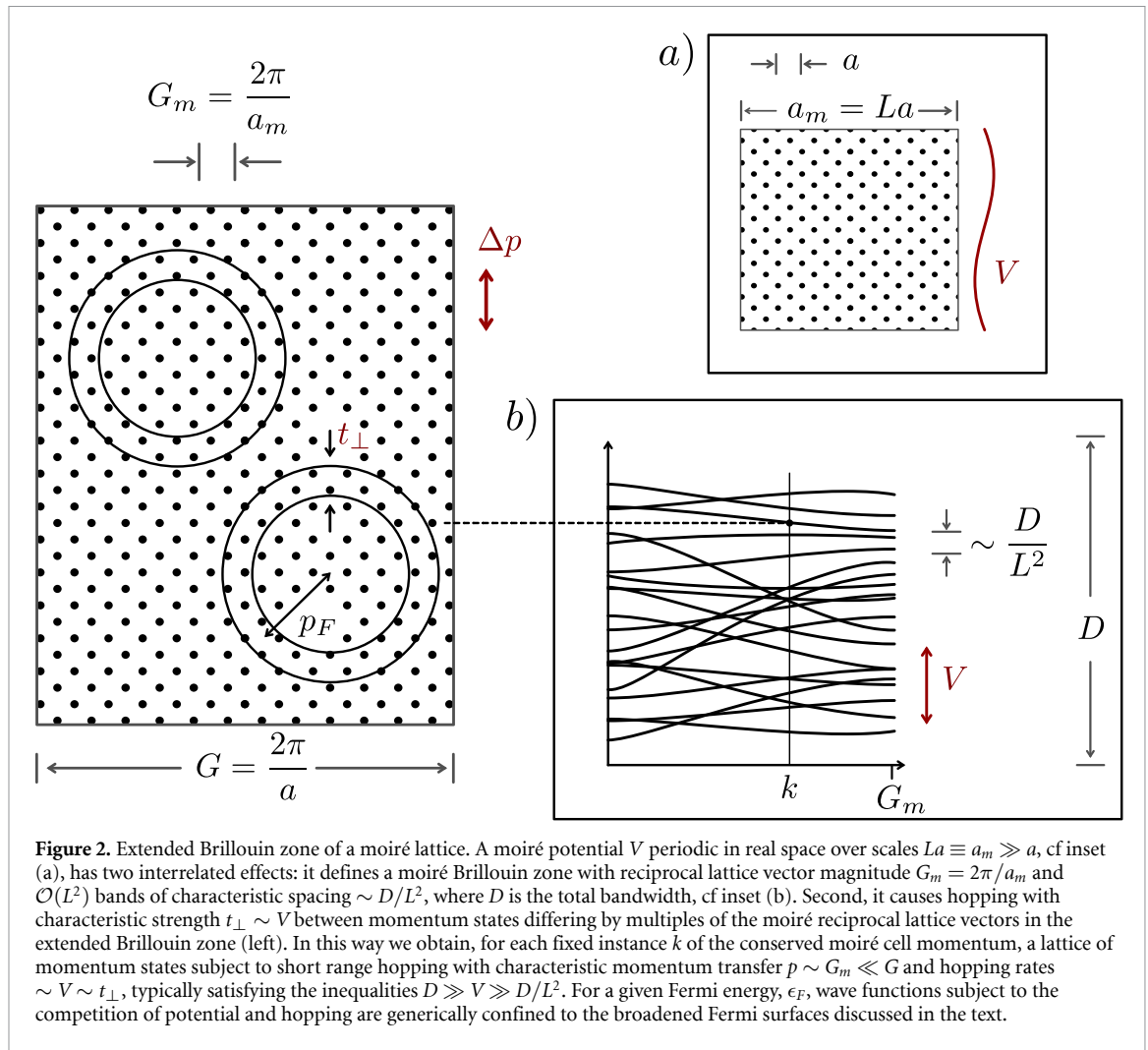
contours, see figure 2. Second, as the reciprocal lattice is discrete, the site-to-site variations of $\epsilon_{\boldsymbol{Q}_n}$, define a source of effective disorder.

Such *quasi* one-dimensional disordered systems are subject to Anderson localization weaker than in strictly one-dimensional materials (potential roadblocks can be efficiently sidestepped) but stronger than in two-dimensions (there is limited phase space for transverse diffusion). We aim to explore how the ensuing physics of wave function confinement due to effective disorder manifests itself in the universal band structure of moiré materials and what exceptions to such a universal framework exists.

To sharpen the question, let us for a moment ignore the localization principle and speculate on the ramifications of the coupling V in the band structure. Under the above assumptions, $V \gtrsim D/L^2$ is a strong and effectively random perturbation parametrically dependent on $\boldsymbol{k}$. On this basis, we should expect efficient level repulsion, i.e. variations of bands over scales $\Delta\epsilon \sim D/L^2$ under parametric variations of extension $\Delta k \sim G_m$—the 'cooked spaghetti' scenario. Characteristic level 'velocities' in this case would be of order $\Delta\epsilon/\Delta k \sim v_F(a/L)^\alpha$, where $v_F \sim D/a$ is a typical group velocity of the underlying two-dimensional material and the exponent $1/2 \leq \alpha \leq 1$ will depend on how strongly neighboring 'cooked spaghetti' will wiggle relatively to each other, see below. Regions of such small dispersion are observed for sufficiently large $t_\perp$, but they are not generic. Far more frequently do we see 'uncooked spaghetti', of steeper velocity $\mathcal{O}(L^0)$. This is proof–by–contradiction that 'quantum disorder' or 'quantum chaos' reasoning by itself does not explain the generic band structure. We aim to demonstrate how the localization principle is the missing element in the story.

While much of our phenomenological reasoning does not rely on model specific assumptions, the concrete calculations below are performed for the case of twisted bilayer structures, more specifically, bilayers of honeycomb lattices as relevant to the case of magic-angle graphene. The details of this model setup are summarized in appendix A.

In this paper, we will look at this system through the lenses of two complementary numerical models. The first is a real space tight-binding model of twisted honeycomb layers with Slater–Koster parameters [38–40], cf appendix B. This model describes the system from first principles and is suited to explore wide portions of its spectrum, including those far detached from its Dirac points. However, this freedom comes at the expense of relatively high computational demands. We use this approach to explore large scale statistical features of the spectrum. The second is a continuum model developed by MacDonald and Bistritzer for bilayer graphene [3] (see appendix F for a review). This model assumes linear dispersion of the uncoupled layers, and hence is limited to the vicinity of the Dirac points. However,

Figure 2. Extended Brillouin zone of a moiré lattice. A moiré potential V periodic in real space over scales $La \equiv a_m \gg a$, cf inset (a), has two interrelated effects: it defines a moiré Brillouin zone with reciprocal lattice vector magnitude $G_m = 2\pi/a_m$ and $\mathcal{O}(L^2)$ bands of characteristic spacing $\sim D/L^2$, where D is the total bandwidth, cf inset (b). Second, it causes hopping with characteristic strength $t_\perp \sim V$ between momentum states differing by multiples of the moiré reciprocal lattice vectors in the extended Brillouin zone (left). In this way we obtain, for each fixed instance k of the conserved moiré cell momentum, a lattice of momentum states subject to short range hopping with characteristic momentum transfer $p \sim G_m \ll G$ and hopping rates $\sim V \sim t_\perp$, typically satisfying the inequalities $D \gg V \gg D/L^2$. For a given Fermi energy, ϵ_F, wave functions subject to the competition of potential and hopping are generically confined to the broadened Fermi surfaces discussed in the text.

due to its computational efficiency it gives us highly resolved insight into the spectra and wave functions near these points.

The rest of the paper is organized as follows: In section 3 we discuss key observations on the band structures of moiré materials obtained within the framework of the real space approach. In section 4 we introduce the real and momentum space description of these structures as a basis for our subsequent discussion of the statistical approach. In the central section 5 we discuss how the band velocity distributions relates to the concept of Anderson localization in momentum space. This is followed by the detailed numerical study of the continuum model in section 6 where we consider spectral and wave function statistics as indicators of chaos and localization, and relate them to the observed characteristics of the velocity statistics. We conclude in section 7, technical details of our analysis are relegated to several appendices.

3. Statistics of velocity distribution

To set the stage for our analysis, we first discuss the band structure of the twisted honeycomb bilayer moiré system, with an emphasis on the distribution

of *band velocities* over wide ranges of the spectrum. Our method of choice in this endeavor is the real space model [38–40] mentioned above. We diagonalize the model to extract the dispersion of the energy bands $\epsilon_n(\boldsymbol{k})$ with band index n and wave vectors $\boldsymbol{k}$ within the first moiré Brillouin zone. The electron group velocity with wave vector $\boldsymbol{k}$ in band n is then given as $\boldsymbol{v}_n(\boldsymbol{k}) = \nabla_{\boldsymbol{k}}\epsilon_n(\boldsymbol{k})$. To obtain the velocity distribution, we collect the absolute values $|\boldsymbol{v}_n(\boldsymbol{k})|$ in all bands for a large set of randomly selected wave vectors. We show their relative occurrence in the right three panels of the upper row of figure 3 for three different choices of the interlayer coupling and corrugation, ranging from decoupled flat layers (cf panels b + f) over the experimental system (cf panels c + g) to strongly coupled and corrugated layers (c.f. panels d + h). The corresponding panels in the lower row show which velocities appear at a given energy.

We first observe that typical electron velocities are of order v_F (cf the vertical dashed lines in panels (b) and (c) in figure 3 for average values), and hence do not suffer the suppression down to scales $\sim 1/L$ which would be expected on the basis of the naive 'band repulsion' picture formulated above. On the other hand, the distributions tend to exhibit enhanced
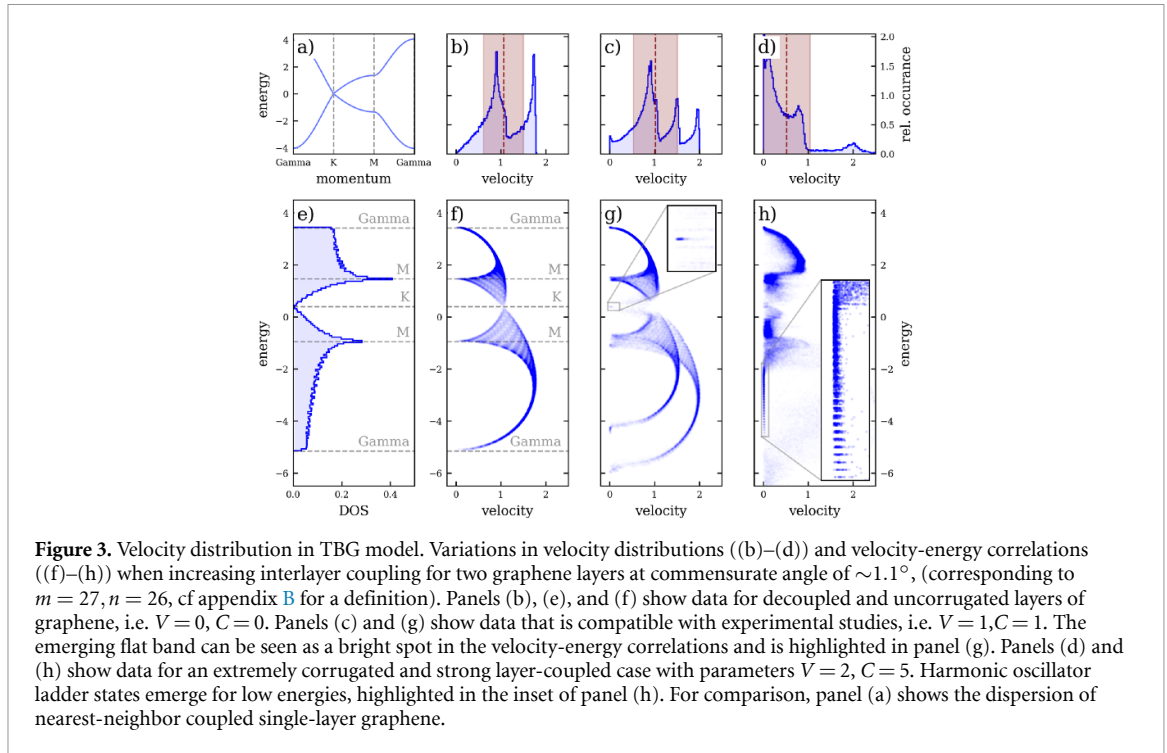
Figure 3. Velocity distribution in TBG model. Variations in velocity distributions ((b)–(d)) and velocity-energy correlations ((f)–(h)) when increasing interlayer coupling for two graphene layers at commensurate angle of $\sim$1.1°, (corresponding to $m = 27, n = 26$, cf appendix B for a definition). Panels (b), (e), and (f) show data for decoupled and uncorrugated layers of graphene, i.e. $V = 0$, $C = 0$. Panels (c) and (g) show data that is compatible with experimental studies, i.e. $V = 1, C = 1$. The emerging flat band can be seen as a bright spot in the velocity-energy correlations and is highlighted in panel (g). Panels (d) and (h) show data for an extremely corrugated and strong layer-coupled case with parameters $V = 2$, $C = 5$. Harmonic oscillator ladder states emerge for low energies, highlighted in the inset of panel (h). For comparison, panel (a) shows the dispersion of nearest-neighbor coupled single-layer graphene.

probabilities for smaller velocities upon increasing corrugation and interlayer coupling, cf panel (d) in figure 3 where the interlayer tunneling is twice as large as experimentally reported and the corrugation is increased by a factor of five. In this extreme parameter regime, the main peak of the velocity distribution is shifted toward small velocities right above zero. The corresponding panel (h) indicates that the small velocities come from different energy regions at and near the van-Hove filling.

In addition to the generic features of the energy–velocity distribution, we observe a number of anomalies in the form of almost perfectly non-dispersive (zero velocity) bands. Among these, the most prominent is the celebrated flat band of TBG, visible as a bright spot in the inset of panel g). The interpretation of this anomaly in the mindset of the present approach is discussed in section 7. At strong corrugation we observe a different type of flat bands, distinguished by their uniform spacing in energy (inset of panel h). This ladder structure affords a natural interpretation as a tunneling phenomenon in the moiré momentum space lattice structure, as we discuss in appendix D. Here, we will focus on the analysis of the 'generic' regions of the velocity distribution and their explanation in terms of momentum space localization.

4. Momentum space localization

With an eye on the generic regions of moiré band structures, we now proceed to develop the momentum space picture in detail and how it can be used to characterize the universal features of moiré band structures. For the sake of definiteness, we again consider the case of TBG. However, most of the discussion applies to different types of moiré materials with little or no alteration.

We consider a lattice in its momentum space representation, as schematically depicted in the inset of figure 4(a). General wave functions $|\psi_n(\boldsymbol{k})\rangle$ with momentum $\boldsymbol{k}$ in the first moiré Brillouin zone and band index n are given by

$$|\psi_n(\boldsymbol{k})\rangle = \sum_{\boldsymbol{Q}} \sum_{\alpha=U,L} A_n^\alpha(\boldsymbol{k}-\boldsymbol{Q})|\psi_\alpha(\boldsymbol{k}-\boldsymbol{Q}-\boldsymbol{K}^\alpha)\rangle.$$

(1)

Here $\boldsymbol{Q}$ are the moiré reciprocal lattice vectors, α is the layer index ($\alpha = U, L$), $\boldsymbol{K}^\alpha$ is the K point of layer α in the first moiré Brillouin zone, cf figure 2(b), $A_n^\alpha(\boldsymbol{k}-\boldsymbol{Q})$ are expansion coefficients, and $|\psi_\alpha(\boldsymbol{k}+\boldsymbol{Q}-\boldsymbol{K}^\alpha)\rangle$ are two-component vectors whose components describe the wave function amplitudes on the A and B sites of the bipartite honeycomb lattice, respectively. Using $|\psi_n(\boldsymbol{k})\rangle$ as basis vectors, one can evaluate the tunneling matrix elements between layers and derive an effective Hamiltonian, see appendix F.

Hereafter, we concentrate on physics near the K points for simplicity. By folding out of the first moiré Brillouin zone to the extended zone, a lattice spanned by the reciprocal moiré lattice vectors $\boldsymbol{Q}$ (cf gray dots in figure 4) emerges.

For the following discussion it is useful to simplify our setup still further and temporarily ignore the layer and the K-point index. The essential physics is then described by the Hamiltonian
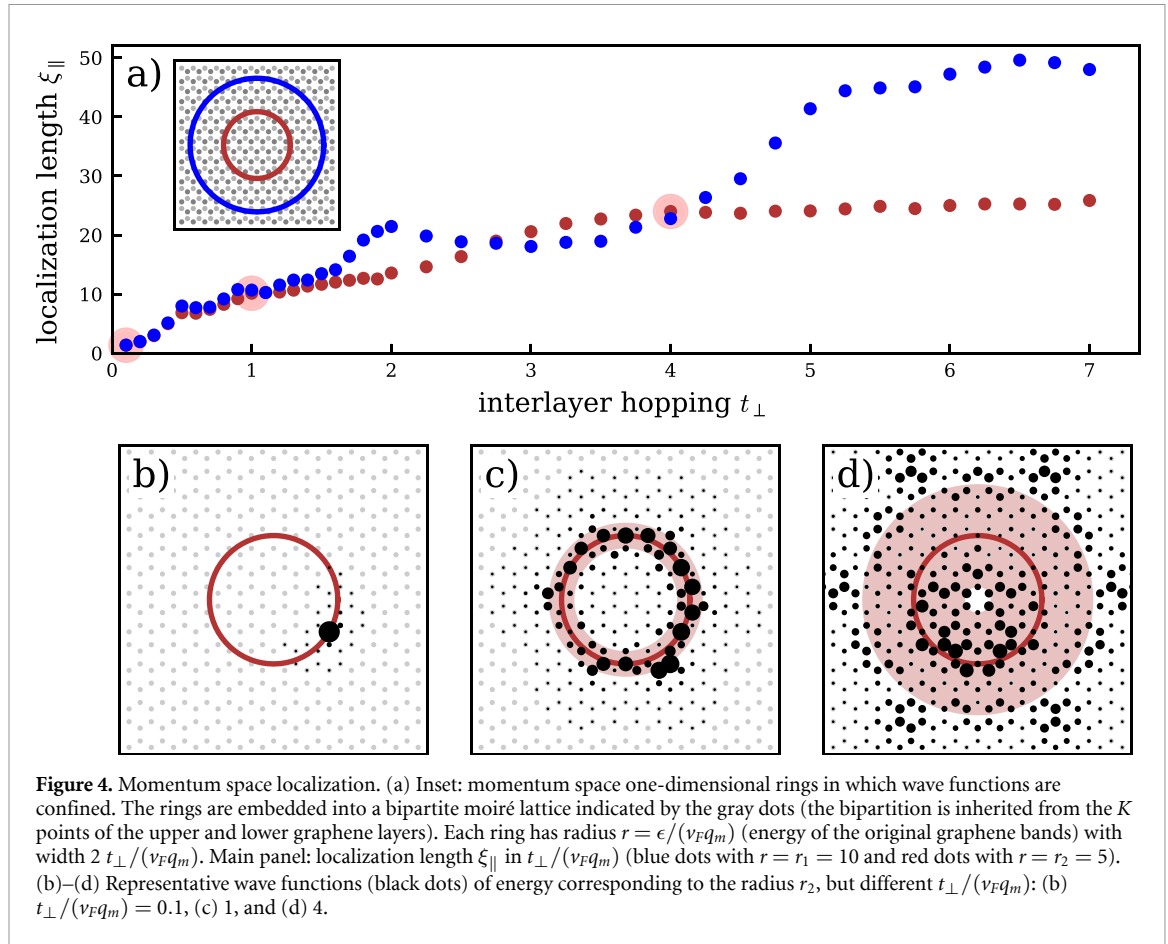
Figure 4. Momentum space localization. (a) Inset: momentum space one-dimensional rings in which wave functions are confined. The rings are embedded into a bipartite moiré lattice indicated by the gray dots (the bipartition is inherited from the K points of the upper and lower graphene layers). Each ring has radius $r = \epsilon/(v_F q_m)$ (energy of the original graphene bands) with width $2\,t_\perp/(v_F q_m)$. Main panel: localization length $\xi_\parallel$ in $t_\perp/(v_F q_m)$ (blue dots with $r = r_1 = 10$ and red dots with $r = r_2 = 5$). (b)–(d) Representative wave functions (black dots) of energy corresponding to the radius r_2, but different $t_\perp/(v_F q_m)$: (b) $t_\perp/(v_F q_m) = 0.1$, (c) 1, and (d) 4.

$$\hat{H}_{\boldsymbol{k}} = -t_\perp \sum_{\langle \boldsymbol{Q}, \boldsymbol{Q}' \rangle} (c^\dagger_{\boldsymbol{Q}} c_{\boldsymbol{Q}'} + \text{h. c.}) + \sum_{\boldsymbol{Q}} \epsilon(\boldsymbol{k} + \boldsymbol{Q}) c^\dagger_{\boldsymbol{Q}} c_{\boldsymbol{Q}},$$

(2)

where $c^\dagger_{\boldsymbol{Q}}$ creates a particle at momentum $\boldsymbol{k} + \boldsymbol{Q}$ and we made the parametric dependence on the conserved momentum $\boldsymbol{k}$ explicit.

As discussed above, the Hamiltonian equation (2) describes the spreading of wave functions along contours of constant energy $\epsilon_{\boldsymbol{Q}} \simeq \epsilon = \text{const}$. A wave function hybridizes over two nearest neighbors of the momentum lattice on these shells provided the energy difference is of order $\Delta \epsilon \sim v_F G_m \lesssim t_\perp$. This hybridization criterion gives the shells a width of $\sim t_\perp / v_F G_m$ in *transverse* direction. In the *longitudinal* direction, the effectively random site energy variations make quasi one-dimensional localization an inevitable consequence.

However, the question remains under what conditions that localization length ξ is smaller or larger than the circumference of the momentum space Fermi ring. Unfortunately, finding parametric estimates for the dependence $\xi(t_\perp, \epsilon, \{\epsilon_{\boldsymbol{Q}}\})$ of the localization length on the relevant system parameters is not easy under the present circumstances. The reason is that for most systems of interest, $t_\perp$ is of the same order as the characteristic energy differences $\delta \epsilon \sim |\epsilon_{\boldsymbol{Q}} - \epsilon_{\boldsymbol{Q}'}|$ between nearest neighbors, i.e. $t_\perp \simeq \delta \epsilon$. We are thus sitting between the two chairs of localization

in strongly and weakly disordered media, respectively. On top of that, the lattice contains stretches of sites approximately aligned with the Fermi surface (see appendix G for further discussion and illustration). Along these, site-to-site energy differences are atypically small, defining local corridors of near ballistic wave function propagation.

5. Three localization regimes

The momentum space setup introduced above sets the stage for the identification of three different regimes of qualitatively different phenomenology. The nature of these is best understood by considering what happens as the interlayer coupling, $t_\perp$, is gradually increased for a system at given Fermi energy ϵ:

I. Deep localization regime

For small site hopping, $t_\perp$, we are in a regime of strong Anderson localization. Individual eigenfunctions, ψ_n are centered around specific momenta $\boldsymbol{k}_n = \langle \psi_n | \hat{\boldsymbol{k}} | \psi_n \rangle$ of magnitude $k_n \sim \epsilon/v_F$. The group velocity of the electron is computed from $\langle \psi_n | \frac{\partial H_{\boldsymbol{k}}}{\partial \boldsymbol{k}} | \psi_n \rangle$ and changes very little as long as the wavefunction is localized in close proximity to $\boldsymbol{k}_n$. At the same time, $|\partial_{\boldsymbol{k}} H_{\boldsymbol{k}}| \sim v_F$ is large, giving the band dispersion the structure of steep almost linear functions of $\boldsymbol{k}$—the regime of uncooked spaghetti. In this regime

all effects arising from the fact that the Fermi surface is closed are exponentially suppressed. We finally notice that the complete momentum space localization implies the absence of correlations between distinct eigenfunctions and their eigenvalues, with the observable consequence of Poissonian spectral statistics. This regime is of direct relevance for bilayer graphene at large twist angle, see below.

II. Quasi one-dimensional (de-)localization
An increasing of $t_\perp$ causes a gradual compromising of the pristine momentum space localization. Wave functions begin to spread around the Fermi ring, and thus become correlated. While analytic computations are not straightforward, our numerical analysis below suggests an approximately linear dependence, $\xi \sim (t_\perp/v_F)$ for the localization length in momentum space for small $t_\perp$. Assuming that in the same regime the radius of the Fermi surface $\sim \epsilon/v_F$ is linear in the Fermi energy, we obtain a crossover scale $t_\perp \sim \epsilon$ for the I/II regime boundary. Inside regime II, delocalization combined with the presence of effective randomness should lead to chaotic (Wigner–Dyson) correlations in the energy spectra. At the same time, individual states remain inhomogeneously distributed, with non-vanishing expectation values of magnitude $k_n \sim \epsilon/v_F$. On this basis, we expect bands with continued steep slope, but showing the 'level repulsion' symptomatic for chaotic spectra. Metaphorically, this is a regime of semi-cooked spaghetti.

III. Strong coupling and dimensional crossover
Naively, one would expect that further increase of $t_\perp$ results in an *ergodic* phase characterized by uniform wave function distribution around the Fermi surface and strong level correlations. However, the actually observed behavior is more nuanced. In fact, the increase of the coupling, starts several developments, the confluence of which determines the observable phenomenology: for coupling strength approaching $t_\perp \sim \epsilon$, wave functions are no longer confined to a ring, they flood the interior of the constant 'potential' energy circle, and significantly extend beyond it. We note that for models with Dirac dispersion this criterion for the II/III boundary is parametrically of the same order than that for the I/II boundary, indicating that the intermediate regime II may not have a parametrically wide support. On the other hand, the three regimes are defined by physically different principles and our analysis below demonstrates the prevalence of regime II over a numerically wide interval for relevant model parameters.

Second, the diminishing influence of the effectively random fluctuations in ϵ_Q implies a higher degree of wave function isotropy. For the same reason, discrete *symmetries* begin to play a role. The moiré momentum space lattice shows crystal discrete symmetries, namely C_3 rotation symmetry and the mirror symmetry M_y ($y \to -y$) combined with

an operation that flips the upper and lower layers. Except at few high symmetry points in the reduced $\boldsymbol{k}$-Brillouin zone, these symmetries are broken by the parametric momentum $\boldsymbol{k}$ in the function $\epsilon_Q = \epsilon_Q(\boldsymbol{k})$. However for strong $t_\perp$ this symmetry breaking gets relatively weaker, especially in the neighborhood of high symmetry points. In principle, one may expect a situation where the Hilbert space is reorganized into irreducible symmetry representation spaces of these, with chaotic correlations inside each symmetry sector due to the residual influence of ϵ_Q and only weak correlations between different sectors. This anticipation suggests spectral statistics intermediate between Wigner–Dyson and Poissonian. At the same time, we expect band velocities parametrically smaller than in regime II. To be a little more concrete, assuming that a characteristic wave function spreads over N lattice sites exploring both positive and negative velocities. In this case, the central limit theorem suggests a velocity expectation value of order $v_F/\sqrt{N}$ with $N \sim L^2$ in the two-dimensional regime while $N \sim L$ in the quasi one-dimensional regime. While this estimate may be too crude, our analysis below confirms hybrid Wigner–Dyson/Poisson hybrid statistics and band velocities drastically reduced compared to those in regime II. Note that the statements given above only hold for closed Fermi surfaces of the underlying single-layer dispersion. For example, if one considers a moiré where single-layer transport is quasi-one dimensional with two disconnected open Fermi surfaces, then the spread of the wave function along a single Fermi surface will still result in a large net velocity.

6. Continuum model

From our previous discussion it is evident that both the structure of single wave functions and correlations in the energy spectrum play a crucial role in understanding the physics of the above regimes I–III. On this basis, we focus on two sets of statistical observables throughout: (i) wave-function statistics, as characterized by inverse participation ratios (IPRs), and (ii) spectral statistics described by the so-called Kullback–Leibler (KL) divergence. The latter is particularly suited to the quantification of spectral statistics in hybrid regimes where neither Poisson nor Wigner–Dyson statistics prevails in pure form. We will analyze these quantities within the framework of a continuum model developed by MacDonald and Bistritzer for bilayer graphene [3] (see appendix F for a review). Focusing on the strong corrugation limit [41], we now discuss how increasing the interlayer coupling $t_\perp$ brings the three regimes discussed above to life.

6.1. Wave function statistics
The degree to which a wave function, defined for a set of lattice sites $\boldsymbol{Q}$, is delocalized is conveniently quantified by the IPR[42], $\sum_{\boldsymbol{Q}} |\psi_{\boldsymbol{Q}}|^4$. In the limiting cases

of fully delocalized and perfectly localized states, this quantity assumes the values $1/L^2$ and 1, respectively. More generally, the IPR probes the inverse square of the localization length. More precisely, we define the localization length $\xi_\parallel(\epsilon)$ along a ring of given energy ϵ, by the IPR projected onto the ring as

$$\xi_\parallel(\epsilon) = 1/\overline{\left(\sum_{S_l} \left(\sum_{(\boldsymbol{Q}+\boldsymbol{K}^\alpha)\in S_l(\epsilon)} \left| A_n^\alpha(\boldsymbol{k}-\boldsymbol{Q}) \right|^2 \right)^2 \right)}.$$

(3)

Here, the first sum runs over a set of momentum vectors S_ℓ given by $\boldsymbol{Q}+\boldsymbol{K}^\alpha$ that fulfill the condition

$$\text{Round}\left[\epsilon/(v_F q_m)\left(\text{Arg}(\boldsymbol{Q}+\boldsymbol{K}^\alpha)+\pi\right)\right]$$
$$= \ell \quad \text{for} \quad \ell \in \mathbb{Z},$$

where $q_m = |\boldsymbol{K}^U - \boldsymbol{K}^L|$ is a momentum difference between momenta in the upper layer, $\boldsymbol{K}^U$, and lower layer, $\boldsymbol{K}^L$. The overbar represents the averaging over momenta $\boldsymbol{k}$ in the first moiré Brillouin zone and band index n with eigenenergy $\epsilon_n(\boldsymbol{k})$ in the vicinity of ϵ.

This localization length $\xi_\parallel$ is plotted as a function of the interlayer hopping $t_\perp$ in the top panel (a) of figure 4 for two different dimensionless ring radii $r \equiv \epsilon/(v_F q_m)$, $r_1 = 10$ (blue dots) and $r_2 = 5$ (red dots). The most obvious structure to notice is an increase of the localization length followed by the eventual saturation at a plateau. For small $t_\perp$, the IPR does not show significant energy dependence, and the two curves approximately coalesce. We interpret this observation as localization driven by the local *incommensurability* of the lattice structure, cf figure 9 in Appendix G for an illustration. Roughly, this means that the extension of wave functions depends on the geometric orientation of the locally straight Fermi surface relative to the momentum lattice. With increasing $t_\perp$, the wave function explores larger regions of the Fermi surface. For the system of smaller energy, the circular geometry of the latter becomes visible at values of the localization length, where the larger surface still looks approximately straight. This is the reason for the deviations between the two curves at some intermediate coupling strength. Eventually saturation of the localization length is expected. Contrary to what one might expect, we observe saturation at a value somewhat different from the ring circumference. The origin of this deviation will be discussed below. We have no convincing explanation for the growing tendency for non-monotonous behavior of the localization length with increasing energy.

It is illuminating to relate this discussion to the structure of actual wave functions. At small interlayer coupling $t_\perp \approx 0.1 \, v_F q_m$, wave functions are fully localized at single (but arbitrary) points in momentum space, as shown in panel (b) of figure 4. For intermediate interlayer coupling $t_\perp \sim v_F q_m$, they start to spread out along the ring and become mutually correlated, panel (c). However, a still inhomogeneous

distribution of wave function weight implies a characteristic band velocity of $\mathcal{O}(\epsilon/k_F)$ in line with the discussion of the intermediate regime II in the previous section. For even larger interlayer couplings $t_\perp > v_F q_m$, the wave functions begin to spread out into the two-dimensional momentum space, cf panel (d). (This excursion into the second dimension may explain the above mentioned numerical discrepancy between the saturation value of $\xi_\parallel$ and $2\pi r$.) At the same time we observe regular features, in the shown example an approximate reflection symmetry at the vertical axis. These structures herald the increasing importance of lattice symmetries and the entrance into regime III.
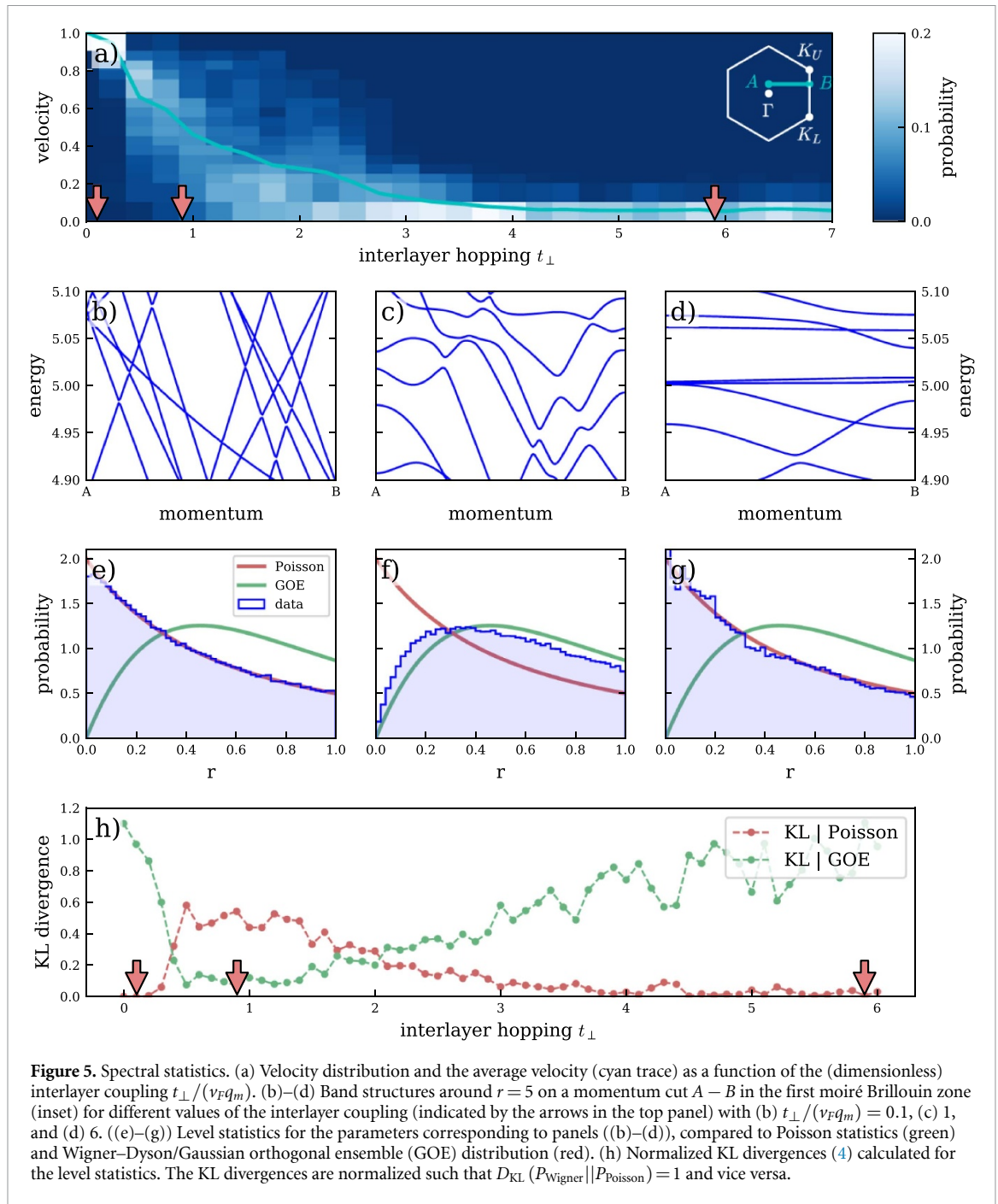
6.2. Spectral statistics

To further characterize the different regimes, we now turn to spectral statistics as a second diagnostic tool. Our results for the spectral correlations characterizing the system are summarized in figure 5.

The top panel (a) provides an overview of how the band velocities are distributed in dependence of of the interlayer coupling $t_\perp$. Starting from the Fermi velocity (set to 1 in this plot), the average velocity slowly decreases until it asymptotically hovers around a small finite value of $v_{\text{avg}} \approx 0.05$ for a numerical value of $t_\perp \gtrsim 4$. While this evolution of the velocity distribution in itself might not point at three different regimes, such a distinction becomes apparent when looking at typical band structures for increasing values of the interlayer coupling as depicted in panels (b–d) in the middle row of figure 5.

Panel (b) on the left shows the scenario of 'uncooked spaghetti', regime I, with different energy levels criss-crossing one another at small interlayer coupling. Here the average band velocities are of the order of v_F and level repulsion is exponentially small. That this is a momentum space localized regime becomes evident in the *level statistics* of the spectrum. To this end, we calculate the distribution of ratios of adjacent level spacings $r_n = \Delta E_n / \Delta E_{n+1}$ [43] (in order to avoid level unfolding). The resulting distribution is then compared to what is expected for these ratios in Poisson or Wigner–Dyson statistics (orthogonal symmetry class, Gaussian orthogonal ensemble (GOE)), i.e. to what is expected for a localized versus delocalized phase, respectively. As demonstrated in the lower row of panels, the data almost perfectly follows the Poisson distribution of a localized phase.

Further increasing the interlayer coupling one enters the regime II of 'semi-cooked spaghetti' with a significant amount of level repulsion and still steep average dispersion. This is the regime which we identified as the momentum-space localized regime above where the wavefunction is restricted to a ring in momentum space, i.e. delocalized in one dimension but localized with regard to the perpendicular (radial) direction. In terms of level statistics, this delocalization becomes evident in the observation of

Figure 5. Spectral statistics. (a) Velocity distribution and the average velocity (cyan trace) as a function of the (dimensionless) interlayer coupling $t_\perp/(v_F q_m)$. (b)–(d) Band structures around $r = 5$ on a momentum cut $A - B$ in the first moiré Brillouin zone (inset) for different values of the interlayer coupling (indicated by the arrows in the top panel) with (b) $t_\perp/(v_F q_m) = 0.1$, (c) 1, and (d) 6. ((e)–(g)) Level statistics for the parameters corresponding to panels ((b)–(d)), compared to Poisson statistics (green) and Wigner–Dyson/Gaussian orthogonal ensemble (GOE) distribution (red). (h) Normalized KL divergences (4) calculated for the level statistics. The KL divergences are normalized such that $D_{KL}(P_{Wigner}||P_{Poisson}) = 1$ and vice versa.

Wigner–Dyson statistics, more precisely a GOE distribution of energy level ratios as shown in panel (f) in the lower row of figure 5. The spectrum at fixed momentum is described by the orthogonal symmetry class due to a combination of time-reversal symmetry and inversion which maps the momentum $\boldsymbol{k}$ to itself. Note that even a weak external magnetic field will reduce the symmetry to the unitary symmetry class. It is an interesting open question whether this leads to a singular response in the band structure.

Finally, for yet stronger coupling we enter the 'cooked spaghetti' regime III, with weakly dispersive bands. Naked eye inspection of the dispersion reveals the presence of level repulsion, but also level crossings. We tentatively interpret this observation in terms of the almost decoupled symmetry multiplets discussed in the end of section 4. Notice that the weak dispersiveness (k-dependence) reflects the relatively weaker influence of the effective disorder: it is no longer strong enough to localize, giving the wave functions a symmetric a roughly isotropic distribution along the fermi surface. Under these circumstances, velocity expectations values are small and depend only weakly on the parameter k entering the lattice potential. The observed Poisson statistics indicates the absence of correlations between different symmetry sectors of the Hilbert space. However, we have not managed to identify subspaces of cleanly

realized lattice symmetries, indicating that the Hilbert space decomposition is only approximated.

In more quantitative terms [44], the identity of the different regimes can be resolved by monitoring the proximity of the spectral distribution to either Poisson or Wigner–Dyson statistics via an entropic measure. To this end, we consider the KL divergence [45]

$$D_{KL}(P||Q) = \sum_s p_s \log\left(\frac{p_s}{q_s}\right), \qquad (4)$$

as a logarithmic measure for the difference between two distributions *P* and *Q*. Specifically, the lowest panel in figure 5(h) shows the KL divergences between the observed distribution and the Poisson/Wigner–Dyson distribution as a function of the interlayer coupling. The crossover from regime I to regime II can clearly be seen in the crossover from Poisson to Wigner–Dyson type distributions. However, we also notice that neither of the limiting distributions is generically realized in pure form. This is particularly true for regime III, where we observe strong fluctuations of the KL divergences. This sensitivity to small variations of the interlayer coupling reflects the fact that spectral statistics, in particular chaotic regimes following Wigner–Dyson statistics, are particularly sensitive to the presence of discrete symmetries. The formation of decoupled symmetry multiplets in regime III therefore leads to the return of Poisson statistics for large interlayer couplings.

6.3. Beyond statistics

We already mentioned that the spectra of moiré systems feature structures outside the statistical approach. Indicated by van Hove singularities, visible as spikes in the spectral density shown in figure 1, these include different types of anomalously flat bands.

We first note that close to some of their minima the unperturbed bands are approximately parabolic, $\sim p^2/(2M)$ (in symbolic notation neglecting the two-dimensionality of the problem). The coupling between the layers adds an effective periodic potential $\sim t_\perp \cos(G_m r)$, defining an quantum mechanical washboard Hamiltonian. Referring for a more detailed discussion to appendix D, the potential term for relevant model parameters is effectively strong. Under these circumstances we obtain a spectrum comprising bands exponentially small in the ratio of $t_\perp$ and the recoil energy $E_R = G_m^2/(2M)$, $t_\perp/E_R \gg 1$, centered around the equidistant levels of the parabolic potential minima. Examples of such bands are visible in panel (h) of figure 3.

Then there is of course the celebrated flat band of *magic angle* graphene. In the reading of this paper, this band is the result of an 'magically' high level of fine tuning resulting in a band with exceptionally low,

but not vanishing *k*-dependence. Besides fine tuning, a factor supporting the flatness of this band is its positioning at zero energy. The average particle-hole symmetry of the spectrum visible in figure 1 implies a tendency for 'locking' at this value. At the same time, this band sits at a 'Fermi circle' of vanishing radius, which makes momentum space localization a non-issue. At any rate, the universal principles addressed in this paper have nothing of relevance to say on the engineering of this type of magically flat bands.

7. Discussion

In this paper, we applied statistical concepts to the description of moiré band structures. Our starting point was the observation that the incommensurability of the inter-layer coupling relative to the uncoupled system's dispersion is a source of effective disorder. On this basis, we defined three regimes governed by different principles: Regime I where weak interlayer coupling leads to momentum space localization and approximately linearly dispersive statistically independent energy bands, regime II at intermediate coupling where states become mutually correlated and partially delocalized leading to a nonlinearly dispersive spectrum, and regime III where strong interlayer coupling makes the symmetries of the moiré lattice a relevant feature and the energy bands become nearly, but not perfectly flat.

Embedded in the tangle of the statistical spectrum there exist various non-generic features, notably bands of exceptional flatness localized near band extrema. We also interpreted the celebrated magic angle flat band as a genuinely anomalous structure which owes its existence to multi-parameter fine tuning rather than to a universal principle.

On the basis of the above discussion the best bet for encountering an accumulation of exceptionally flat bands is regime III. However, that regime requires $t_\perp \gtrsim \epsilon_F$, i.e. huge interlayer coupling, or Fermi energies scaled to the vicinity of an effective Dirac point (as is the case for the magic angle flat bands.)

Some open questions that merit further exploration include the role of screening and its band structure effects. One might also entertain the question whether one can engineer longer-range hopping in momentum space in order to drive the system into the delocalized, chaotic regime already for small interlayer coupling and thereby effectively flatten all bands. This could possibly be achieved by a manipulation and rearrangement of atoms (while keeping the underlying moiré periodicity unchanged) to generate a short-range potential in real space, which in turn induces a long-range hopping in momentum space.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgment

Appendix A. Magic-angle graphene

We here define the model of TBG which serves as a basis for the numerical studies of this paper. Graphene has a lattice constant a of the order of angstrom and a total bandwidth $D \sim \mathcal{O}(10\,\text{eV})$ [40, 48]. At the magic angle of $1.1°$, the stacked material has a moiré lattice constant $a_m \approx 50a$ [3, 49]. A moiré potential V is caused by the interlayer coupling $t_\perp \sim \mathcal{O}(0.1\ \text{eV}) \sim D/L$, which is subject to spatial variations on the large scale of the moiré unit cell. At such small angles, incommensuration effects coming from $LG_m \neq G$ are suppressed by a factor $(t_\perp/D)^L$, i.e. no quasicrystal physics occurs. The variation of the interlayer coupling may be amplified through the corrugation of the layers in the out-of-plane direction, e.g. the interlayer distance is slightly different in the AA-stacked regions and in the AB-stacked regions, see appendix B. We take this into account in our modeling, below. Note that the bandwidths of the much discussed flat bands at the Fermi level in magic-angle graphene are reported in a range of 20–40 meV [12, 15]. This is significantly larger than theoretical predictions of $\lesssim 10$ meV [50], including corrugation but not in-plane lattice relaxation effects. Putting the experimentally reported bandwidth into context with G_m, the typical velocities v in these bands are only moderately suppressed when compared to the typical velocity v_F of single-layer graphene, i.e. $v \sim 0.1, \ldots, 0.2 \cdot v_F$.

Appendix B. Real-space lattice model

To perform band structure calculations for TBG, we start from a real-space lattice model by considering a honeycomb lattice with lattice vectors

$$\boldsymbol{a}_1 = a\left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right), \quad \boldsymbol{a}_2 = a\left(\frac{\sqrt{3}}{2}, +\frac{1}{2}\right) \quad \text{(B1)}$$

and lattice constant a.

Twisting two graphene layers with respect to each other by a finite angle θ gives rise to a large-scale interference pattern—the moiré pattern. At commensurate twist angles, moiré *unit cells* are formed which are strictly periodic. These commensurate twist angles

are achieved by pairs of integer numbers m and n which define the twisting of the Bravais lattice site $R_u = m\boldsymbol{a}_1 + n\boldsymbol{a}_2$ in the upper layer on top of site $R_l = n\boldsymbol{a}_1 + m\boldsymbol{a}_2$. The twist angle $\theta_{m,n}$ for a given pair (m,n) is then defined as

$$\cos(\theta_{m,n}) = \frac{1}{2}\frac{m^2 + n^2 + 4mn}{m^2 + n^2 + mn}. \quad \text{(B2)}$$

The corresponding moiré Bravais lattice vectors are

$$\boldsymbol{a}_{1,M} = m\boldsymbol{a}_1 + n\boldsymbol{a}_2, \quad \boldsymbol{a}_{2,M} = R(60°)\boldsymbol{a}_{1,M}. \quad \text{(B3)}$$

We further use a generic tight-binding Hamiltonian with all-to-all hopping amplitudes $t(\boldsymbol{r}_{i,j})$, reading

$$\mathcal{H} = \sum_{i,j} t(\boldsymbol{r}_{i,j}) c_j^\dagger c_i. \quad \text{(B4)}$$

The distance-dependent hopping amplitudes are chosen to be in the Slater–Koster form [38–40], i.e.

$$t(\boldsymbol{r}) = V_{pp\pi}(r)\left(1 - \left(\frac{\boldsymbol{r}\cdot\boldsymbol{e}_z}{r}\right)^2\right) + V_{pp\sigma}(r)\left(\frac{\boldsymbol{r}\cdot\boldsymbol{e}_z}{r}\right)^2, \quad \text{(B5)}$$

$$V_{pp\pi}(r) = V_{pp\pi}^0 e^{-(r-a_0)/\delta_0}, \quad \text{(B6)}$$

$$V_{pp\sigma}(r) = V_{pp\sigma}^0 e^{-(r-d_0)/\delta_0}. \quad \text{(B7)}$$

Here, we have introduced the intralayer nearest-neighbor distance $a_0 = a/\sqrt{3}$, the mean interlayer distance d_0 (see next paragraph), the decay distance of orbital overlap δ_0, as well as the two overlap integrals $V_{pp\pi}^0$ and $V_{pp\sigma}^0$. In our calculations we base our parameters on the experimentally determined values of $a_0 = 0.142$ nm, $\delta_0 = 0.319$ nm as well as $V_{pp\pi}^0 = -2.7$ eV and $V_{pp\sigma}^0 = 0.48$ eV.

For the modeling of the real-space structure of TBG in a setting close to experiments, we further take into account corrugation effects which buckle the lattice on scales of the moiré cell, i.e. we introduce a periodic variation of the interlayer distance. This effect stems from the interactions between the atoms of the two layers which are stacked either in an AA or AB fashion, depending on the relative position within the moiré cell. The layer distance therefore varies between d_{AA} and d_{AB} with the periodicity of the moiré cell as

$$d(\boldsymbol{R}) = d_0 + 2d_1 \sum_{i=1}^{3} \cos\left(2\pi\frac{\boldsymbol{R}\cdot\boldsymbol{C}_i}{|\boldsymbol{C}_i|^2}\right). \quad \text{(B8)}$$

Here, d_0 and d_1 are based on the AA and AB distances as

$$d_0 = \frac{1}{3}(d_{\text{AA}} + 2d_{\text{AB}}), \quad \text{(B9)}$$

$$d_1 = \frac{1}{9}(d_{\text{AA}} - d_{\text{AB}}). \quad \text{(B10)}$$

The corrugation spanning vectors C_i which are enclosing an angle of $60°$ span the moiré pattern and read

$$C_1 = \frac{1}{2}(a_{1,M} + a_{2,M}), \tag{B11}$$

$$C_2 = R(60°)C_1, \tag{B12}$$

$$C_3 = R(120°)C_1. \tag{B13}$$

The AA and AB distances have been experimentally determined to be $d_{AA} = 0.360\,\text{nm}$ and $d_{AB} = 0.335\,\text{nm}$ [50–52].

When referring to specific combinations of model parameters, we always use relative values with respect to the experimental parameters, e.g. a corrugation of 0 refers to a layer distance of the AB regions everywhere in the lattice. In total, we modify the layer distance, the strength of corrugation, as well as the orbital overlap. To that end, we introduce the interlayer coupling parameter V which denotes the interlayer Slater–Koster parameter $V_{pp\sigma}^0$ in units of its experimental value. Further, we parametrize the interlayer distance D_l in units of the equilibrium distance as well as the level of corrugation by a parameter C which denotes the corrugated part of the interlayer distance $d_{AA} - d_{AB}$ in units of the experimentally determined value.

Appendix C. One-dimensional moiré system

In this section, we study a one-dimensional moiré system to explore the effective range of the coupling between lattice sites in momentum space. We show that this coupling decays exponentially on the scale $\Delta p \gtrsim G_m$, where G_m is the shortest moiré reciprocal lattice vector.

To that end, we consider a one-dimensional moiré system consisting of two chains with different lattice constants a and b. Each layer feels a 'substrate potential' induced by the other layer, cf figure 6(a). To obtain a large-scale moiré interference pattern, $|a - b|$ is assumed to be much smaller than a and b. The length of the moiré unit cell is given by $a_m = ab/|a - b|$ and the length of the unit of the moiré reciprocal lattice vector is $G_m = 2\pi/a_m$.

The potential V in the upper chain as induced by the lower chain reads

$$V(R^U) = \sum_{R^L} U(R^U - R^L), \tag{C1}$$

where $U(R^U - R^L)$ is the microscopic interlayer coupling which only depends on the distance $|R^U - R^L|$. Through the potential V an electron with initial momentum p can be scattered into a state with momentum p'. This process is described by the matrix element

$$\langle p'|V|p \rangle = \frac{1}{N} \sum_{R^U} V(R^U) e^{i(p-p')R^U}$$
$$= \frac{1}{N} \sum_{R^U, R^L} U(R^U - R^L) e^{i(p-p')R^U}, \tag{C2}$$

where N is the number of sites in the upper chain. To study the behavior of this matrix element for long wavelengths, we introduce a smoothing function $f(x)$ [38]. The smoothing property of f allows us to treat the moiré potential as a continuous and smooth modulation and estimate the induced hopping range. To realize the smoothing, $f(x)$ is chosen to decay on a scale much larger than the atomic scale a, but much smaller than the moiré scale a_m and it is normalized as $\int_{a_m} dx f(x) = a_m$. Inserting this normalization condition into equation (C2) yields

$$\langle p'|V|p \rangle = \frac{1}{Na_m} \int_x \sum_{R^U, R^L} f(x - R^U) U(R^U - R^L) e^{i(p-p')R^U}$$
$$\sim \frac{1}{Na_m} \sum_{R^U, R^L} \int_x e^{i(p-p')x} f(x - R^U) U(R^U - R^L)$$
$$= f_{p-p'} \sum_{G^U, G^L, q} U(q) \delta_{p-p'+q, G^U} \delta_{q, G^L}. \tag{C3}$$

Here, we used $e^{i(p-p')R^U} \sim e^{i(p-p')x}$ when going from the first to the second line as the smoothing function acts as a delta function on the moiré scale and we introduced the reciprocal lattice vectors of the upper-/lower chains $G^{U/L}$. The two Kronecker deltas in the last line of equation (C3) lead to the relations

$$\Delta p = p - p' = G^U - G^L = 2\pi\left(\frac{m_1}{a} - \frac{m_2}{b}\right) \equiv Q,$$
$$q = G^L = 2\pi\frac{m_2}{b}, \tag{C4}$$

with $m_1, m_2 \in \mathbb{Z}$ and the moiré reciprocal lattice vectors Q. To obtain a non-negligible contribution to $\langle p'|V|p \rangle$, the smoothing function $f(\Delta p)$ needs to be sizeable, which is the case for small momentum transfer Δp smaller than the decay length ($\ll 1/a$ or $1/b$) of $f(\Delta p)$. This is realized for the case $m_1 = m_2$, which we discuss in the following.

Since $U(x)$ decays on the atomic scale a, the potential $U(q)$ decays on the scale $\sim 1/a$. Therefore, the dominant contribution comes from the $q = 0$ part, but it gives rise to the condition $p = p'$, and thus results in a trivial energy shift by a constant term. The lowest non-trivial contribution comes from the $m_1 = m_2 = \pm 1$. According to equation (C4), we then obtain $\Delta p = \pm 2\pi(1/a - 1/b)$ and $q = \pm 2\pi/b$ leading to the matrix element $\langle p'|V|p \rangle = U(2\pi/a)$. More generally, choosing $m = m_1 = m_2$, i.e. $\Delta p = 2\pi m(1/a - 1/b)$ and $q = 2\pi m/b$, yields the matrix element

$$\langle p'|V|p \rangle = U\left(\frac{2\pi m}{b}\right). \tag{C5}$$

Figure 6. One-dimensional moiré system. (a) Two layers with different lattice constant a and b. $|a - b|$ is assumed to be much smaller than a and b, or equivalently a, b are much smaller than the size of moiré unit cell, a_m. The lattice sites in the upper (lower) layer are denoted as red (blue) dots. (b) and (c) commensurate lattices with the condition $\ell_1 a = \ell_2 b$ with integers ℓ_1, ℓ_2. (b) $\ell_1 = 15$ and $\ell_2 = 14$. (c) $\ell_1 = 36$ and $\ell_2 = 33$. There is a three-fold approximate symmetry within the exact periodicity. After washing out the fine structures of the lattices by a bootstrap (e.g. disorder), the periodicity of the approximate symmetry governs features of the coupling of lattice sites in the momentum space.

Hence, together with the assumption that the microscopic interlayer coupling U decays exponentially on the atomic scale a, the coupling of the lattice sites in momentum space decays exponentially over scale $\Delta p \gtrsim G_m$.

We note that the above argument does not depend on whether the lattice configuration is commensurate or incommensurate since we never explicitly impose the condition of $LG_m = G^U$ or $LG_m = G^L$. It is instructive, however, to briefly revisit the reasoning for the commensurate case, where $l_1 a = l_2 b$ with integers ℓ_1 and ℓ_2. When $|\ell_1 - \ell_2| = 1$, cf figure 6(b), the lowest non-trivial condition $p - p' = 2\pi(1/a - 1/b) = \pm 2\pi/(\ell_1 a) = \pm G_m$ is fulfilled; the range of the coupling of the lattice sites in momentum space decays over the scale G_m. Even for $|\ell_1 - \ell_2| \neq 1$, we always find the approximate symmetry that satisfies $\ell_1' a \simeq \ell_2' b$ with $|\ell_1' - \ell_2'| = 1$, cf figure 6(c), only if $|a - b| \ll a, b$. This periodicity of the approximate symmetry governs features of the coupling of lattice sites after washing out the fine structures of the lattices by a possible bootstrap, e.g. disorder. Therefore, the statement that the coupling of lattice sites in momentum space decays over the scale $\Delta p \gtrsim G_m$ holds generally for moiré systems.

Appendix D. Flat band ladders

An exceptional feature of our real-space TBG model is the occurrence of a sequence of equally-spaced flat bands. This sequence emerges near the minimum of the graphene bands of the uncoupled system, cf the inset in panel (h) of figure 3 in the main text. The underlying physical mechanism at play here can be readily understood in terms of a harmonic-oscillator level spacing. To that end, consider electrons near the band minimum of the uncoupled layers tunneling in the periodic moiré potential, which—along one chosen direction—is dominated by the modulation term $\sim t_\perp \cos(G_m r)$ [53]. Near a band minimum, the kinetic energy of the electrons can be approximated using a quadratic approximation $\sim k^2/(2m)$, where

m is given by the curvature at that minimum. The potential is large as compared to the recoil energy, since the parameters in the presented model imply

$$t_\perp \gtrsim 1/L \gg G_m^2/(2m) \sim 1/L^2.$$

This allows us to extract the corresponding level spacing of the harmonic oscillator eigenenergies yielding

$$\sqrt{t_\perp G_m^2/m} \sim 1/L^{3/2} \gg 1/L^2.$$

In consequence, we obtain a sequence of equally spaced bands with an exponentially small bandwidth

$$\Delta W \sim \exp(-\sqrt{t_\perp m} G_m) \sim \exp(-\sqrt{N}).$$

Definite numerical evidence of this scenario is provided in figure 7, where we show the wavefunction solutions of the eigenstates in this sequence as plotted in the cascade of panels on the right—a beautiful illustration of two-dimensional harmonic oscillator states.

Appendix E. Spectral statistics for TBG

We complement the analysis of spectral statistics for the continuum model in the main text (section 6.2) with numerical data for the real-space TBG model here. While such an analysis for the real-space model does not give the same quantitative clarity as for the continuum model, we observe very similar qualitative behavior as a function of increasing interlayer coupling. A summary is provided in figure 8. Here, the upper two rows show the velocity statistics as a function of increasing interlayer coupling (different columns), cf figure 3 of the main text. Therein, we present a histogram of velocities in the first row and energy-resolved 'waterfall' plots in the second row. The lower three rows show the spectral statistics. Here, the two rows at the bottom are snapshots of the band structure in the two energy windows

Figure 7. Formation of ladder states near band minimia. The 'waterfall' plot (akin to figure 3 in the main text) on the left highlights the formation, at finite interlayer coupling, of a sequence of equally-spaced flat bands emerging near the minimum of the graphene band of the uncoupled system. The cascade of plots on the right illustrates the wavefunctions for the first five energy levels extracted from exact diagonalization of our real-space TBG model. The parameters are identical to those in figure 3 panels (d) and (h), i.e. strong interlayer coupling $V = 2$ and strong corrugation $C = 5$.

Figure 8. Spectral statistics for real-space TBG model as a function of interlayer coupling at an angle corresponding to $m = 15, n = 12$, cf (B2), i.e. $\theta \approx 2.28°$. We have choosen an interlayer distance of 55% of the experimental distance and a corrugation that is five times stronger than the experimental value, i.e. $D_l = 0.55$ and $C = 5$, cf section B. Columns are shown from left to right with increasing interlayer coupling parameter V which denotes the interlayer Slater–Koster parameter $V_{pp\sigma}^0$ in units of the experimental value $V_{pp\sigma}^0 = 0.48$ eV. The upper two rows show velocity statistics whereas the lower three rows illustrate the spectral statistics. The middle row shows histograms of r values which are compared to Poisson or Wigner Dyson statistics indicate localization versus chaotic regimes. The underlying histograms of values for the ratios of adjacent level spacings r are computed by averaging over several momenta within the first moiré Brillouin zone as well as all energies. The spectral statistics of the real-space model are in overall agreement with what is discussed for the momentum space model in section 6.2 in the main text.

as indicated on the very left. A quantitative analysis of the level spacings in these band structure plots is provided in the middle row. Here, we present the level-spacing distribution as a histogram that is compared to the Poisson and Wigner–Dyson distributions, which are characteristic for localized and delocalized chaotic regimes, respectively.

Looking at the evolution for increasing interlayer coupling parameter V, we observe a similar progression as discussed for the continuum model in the

main text, cf section 6.2: starting from a Poisson-like distribution the level statistics evolves to a broad histogram as resembling a Wigner–Dyson-like distribution. We note, however, that the formation of flat band ladders, cf section D, gives rise to a δ-function like peak at vanishing ratio of adjacent level spacing, i.e. at $r = 0$. This phenomenon somewhat obscures the evolution of the level spacing distribution. For large interlayer coupling, the distribution moves back to a monotonously decaying Poisson-like

distribution, akin to the reentrance behavior discussed for the continuum model in section 6.2.

Appendix F. Momentum-space continuum model for TBG

In this section, we review the continuum model for TBG as developed in [3] as valid for small twist angles θ. We employ this model for the calculations in section 6 as a simple tool to understand the complex band structures of the TBG based on statistical principles.

The starting point is a monolayer of graphene with primitive lattice vectors as given in equation (B1). Explicitly, the basis of the honeycomb lattice sites reads $\boldsymbol{\tau}_A = 0, \boldsymbol{\tau}_B = a\hat{\boldsymbol{x}}/\sqrt{3}$ and the primitive reciprocal lattice vectors are $\boldsymbol{b}_1 = (2\pi/a)(\hat{\boldsymbol{x}}/\sqrt{3} + \hat{\boldsymbol{y}})$ and $\boldsymbol{b}_2 = (2\pi/a)(\hat{\boldsymbol{x}}/\sqrt{3} - \hat{\boldsymbol{y}})$, where $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$ are euclidean unit vectors in the x and y direction, respectively. The inequivalent $\boldsymbol{K}$ and $\boldsymbol{K}'$ points read

$$\boldsymbol{K} = \frac{4\pi}{3a}\hat{\boldsymbol{y}}, \quad \boldsymbol{K}' = -\boldsymbol{K} = -\frac{4\pi}{3a}\hat{\boldsymbol{y}}. \quad \text{(F1)}$$

Accordingly, the equivalent $\boldsymbol{K}$ points in the first Brillouin zone are given by translations by primitive reciprocal lattice vectors, i.e. $\boldsymbol{K}_1 = \boldsymbol{K}, \boldsymbol{K}_2 = \boldsymbol{K} - \boldsymbol{b}_1$, and $\boldsymbol{K}_3 = \boldsymbol{K} - \boldsymbol{b}_1 + \boldsymbol{b}_2$.

We now consider a twisted graphene bilayer with total relative twist angle θ. For symmetry reasons, the upper layer is twisted by $+\theta/2$ and the lower layer is twisted by $-\theta/2$ with respect to a perfectly aligned AA stacking of the bilayer. The Bloch wave function with momentum $\boldsymbol{k}$ ($\boldsymbol{k}'$) residing on sublattice β (β') in the upper (lower) layer is written as

$$|\Psi^U_{\boldsymbol{k}\beta}\rangle = \frac{1}{\sqrt{N_U}} \sum_{\boldsymbol{R}^U} e^{i\boldsymbol{k}\cdot\boldsymbol{R}^U}|\boldsymbol{R}^U + \boldsymbol{\tau}^U_\beta\rangle,$$

$$|\Psi^L_{\boldsymbol{k}'\beta'}\rangle = \frac{1}{\sqrt{N_L}} \sum_{\boldsymbol{R}^L} e^{i\boldsymbol{k}'\cdot\boldsymbol{R}^L}|\boldsymbol{R}^L + \boldsymbol{\tau}^L_{\beta'}\rangle. \quad \text{(F2)}$$

Here $N_{U(L)}$ is the number of the unit cells in the upper (lower) layer and $\boldsymbol{\tau}^{U(L)}_\beta \equiv e^{\pm i\theta\sigma_z/2}\boldsymbol{\tau}_\beta e^{\mp i\theta\sigma_z/2}$ is the rotated basis.

The bare graphene (Dirac) Hamiltonian of layer $U(L)$ near the $K^{U(L)}$ point then reads

$$H^K_{U/L}(\boldsymbol{p}) = -\frac{\sqrt{3}at}{2}\left(\boldsymbol{\sigma}_{\pm\theta/2}\cdot\boldsymbol{p}\right), \quad \text{(F3)}$$

where $\boldsymbol{\sigma}_\theta = e^{i\theta\sigma_z/2}\boldsymbol{\sigma}e^{-i\theta\sigma_z/2}$. Given momentum $\boldsymbol{p}$ in the first Brillouin zone, the bare graphene Hamiltonian with momentum $\boldsymbol{p} - \boldsymbol{Q}_n$ can be unfolded into the extended Brillouin zone scheme as

$$H^K_{U/L}(\boldsymbol{p} - \boldsymbol{Q}_n) = -\frac{\sqrt{3}at}{2}\left(\boldsymbol{\sigma}_{\pm\theta/2}\cdot(\boldsymbol{p} - \boldsymbol{Q}_n)\right). \quad \text{(F4)}$$

Here $\boldsymbol{Q}_n$ are the moiré reciprocal lattice vectors, associated with the periodicity of the moiré unit cell.

The interlayer hopping matrix describing a process where an electron with momentum $\boldsymbol{k}$ residing on sublattice β in the upper layer hops to a state with momentum $\boldsymbol{k}'$ on sublattice β' in the lower layer reads

$$\langle\Psi^L_{\boldsymbol{k}'\beta'}|H_T|\Psi^U_{\boldsymbol{k}\beta}\rangle = \sum_{\boldsymbol{R}^U,\boldsymbol{R}^L} \frac{t(\boldsymbol{R}^L + \boldsymbol{\tau}^L_{\beta'} - \boldsymbol{R}^U - \boldsymbol{\tau}^U_\beta)}{\sqrt{N_U N_L}}$$
$$\times e^{i\boldsymbol{k}\cdot\boldsymbol{R}^U - i\boldsymbol{k}'\cdot\boldsymbol{R}^L}$$
$$= \sum_{\boldsymbol{R}^U,\boldsymbol{R}^L,\boldsymbol{q}} \frac{t^{\beta'\beta}_{\boldsymbol{q}}}{N_U N_L} e^{i(\boldsymbol{q}-\boldsymbol{k}')\cdot\boldsymbol{R}^L - i(\boldsymbol{q}-\boldsymbol{k})\cdot\boldsymbol{R}^U + i\boldsymbol{q}\cdot(\boldsymbol{\tau}^L_{\beta'} - \boldsymbol{\tau}^U_\beta)}$$
$$= \sum_{\boldsymbol{G}^L,\boldsymbol{G}^U}\sum_{\boldsymbol{q}} t^{\beta'\beta}_{\boldsymbol{q}}\delta_{\boldsymbol{q}-\boldsymbol{k},\boldsymbol{G}^U}\delta_{\boldsymbol{q}-\boldsymbol{k}',\boldsymbol{G}^L} e^{i\boldsymbol{q}\cdot(\boldsymbol{\tau}^L_{\beta'} - \boldsymbol{\tau}^U_\beta)}. \quad \text{(F5)}$$

Here, $\boldsymbol{G}^{U(L)}$ is the reciprocal lattice vector of the upper (lower) graphene layer and we have assumed that $t(\boldsymbol{R}^L + \boldsymbol{\tau}^L_{\beta'} - \boldsymbol{R}^U - \boldsymbol{\tau}^U_\beta)$ only depends on the distance $|\boldsymbol{R}^L + \boldsymbol{\tau}^L_{\beta'} - \boldsymbol{R}^U - \boldsymbol{\tau}^U_\beta|$. The two delta functions in equation (F5) lead to the moiré condition

$$\boldsymbol{k} - \boldsymbol{k}' = \boldsymbol{G}^U - \boldsymbol{G}^L = \boldsymbol{Q}_n. \quad \text{(F6)}$$

Under the assumption that $t_{\boldsymbol{q}}$ decays rapidly around $|\boldsymbol{q}| \sim |\boldsymbol{K}|$, $\boldsymbol{q}$ can be restricted to momenta around the K points in the first Brillouin zone ($\boldsymbol{K}_1, \boldsymbol{K}_2$ and $\boldsymbol{K}_3$); $\boldsymbol{q}$ can be decomposed as $\boldsymbol{q} = \boldsymbol{K}^L_i + \boldsymbol{p}' = \boldsymbol{K}^U_i + \boldsymbol{p}$ for $i = 1, 2, 3$ with small momentum $\boldsymbol{p}$ and $\boldsymbol{p}'$, leading to $\boldsymbol{p} - \boldsymbol{p}' = \boldsymbol{K}^L_i - \boldsymbol{K}^U_i \equiv \boldsymbol{q}_i$. Within those approximations, equation (F5) becomes

$$\langle\Psi^L_{\boldsymbol{p}'\alpha'}|H_T|\Psi^U_{\boldsymbol{p}\alpha}\rangle = \sum_{i=1,2,3} t^{\beta'\beta}_{\boldsymbol{K}_i} e^{i\boldsymbol{K}_i\cdot(\boldsymbol{\tau}^L_{\beta'} - \boldsymbol{\tau}^U_\beta)}\delta_{\boldsymbol{p}-\boldsymbol{p}',\boldsymbol{q}_i}. \quad \text{(F7)}$$

We assume that $t^{\beta'\beta}_{\boldsymbol{K}_i}$ are momentum independent ($t^{AA}_{\boldsymbol{K}_i} = t^{AA}_{\boldsymbol{K}_i} = t^{AA}$ and $t^{AB}_{\boldsymbol{K}_i} = t^{AB}_{\boldsymbol{K}_i} = t^{AB}$), and further neglect the small angle dependence of $\boldsymbol{\tau}^L_\alpha, \boldsymbol{\tau}^U_\alpha \simeq \boldsymbol{\tau}_\alpha$. Then, the hopping term becomes a 2×2 matrix, reading

$$\langle\Psi^L_{\boldsymbol{p}'}|H_T|\Psi^U_{\boldsymbol{p}}\rangle = \sum_{i=1,2,3} \delta_{\boldsymbol{p}-\boldsymbol{p}',\boldsymbol{q}_i}\left(t^{AB}\left[\sigma_x\cos\left(\frac{2\pi}{3}(i-1)\right)\right.\right.$$
$$\left.\left. + \sigma_y\sin\left(\frac{2\pi}{3}(i-1)\right)\right] + t^{AA}\mathbb{1}\right). \quad \text{(F8)}$$

This completes the description of TBG in the continuum model as employed in section 6.

Appendix G. Localization by incommensurability

In this section, we address the localization driven by the local incommensurability as mentioned in section 6. To this end, we consider an approximate Hamiltonian, describing a physical mechanism that takes place in a small fraction of a one-dimensional energy circle of the original graphene dispersion with

Figure 9. Localization by incommensurability. (a) The circle of equal energy crossing momentum $\boldsymbol{p}_0 = p_0(\cos\phi, \sin\phi)$. In the limit of $p_0 \to \infty$ this can be regarded as a straight line (red line) with angle $(\pi - \phi)$. (b) Localization length $\xi_\parallel$ along the line as a function of the angle ϕ with pronounced peaks at commensurate angles. Here, the notion '(in)commensurate angle' means that the distances between the lattice points and the line are (in)commensurate. Different colors represent different $t_\perp$, i.e. $t_\perp = v_F q_m$ (blue dots) and $t_\perp = 0.1 \, v_F q_m$ (red dots). The pronounced peaks of $\xi_\parallel$ at commensurate angles show that incommensurability is a dominant source of localization. (c) $\xi_\parallel$ as a function of $t_\perp$ with fixed $\phi = \pi/12$.

Figure 10. Corrugation effect. (a) Diagrams to show how lattice sites in the momentum space are connected with the nearest neighbor sites by the interlayer hopping with a different level of the corrugation. The corrugation is effectively parameterized with φ as $t^{AA} = t_\perp \cos\varphi$ and $t^{AB} = t_\perp \sin\varphi$. With the eigenstates $|\boldsymbol{k} + \boldsymbol{Q}_n, \alpha\rangle$ of the Hamiltonian in the absence of the interlayer coupling, the strength of the matrix element $|\langle \boldsymbol{k} + \boldsymbol{Q}'_n, U|H_T|\boldsymbol{k} + \boldsymbol{Q}_n, L\rangle|$ is represented as the thickness of the connection between sites. (b) localization length $\xi_\parallel$ as a function of φ with fixed $t_\perp = 0.5 \, v_F q_m$.

large radius, cf the outer-most ring in the inset of figure 4(a).

We start by taking a momentum $\boldsymbol{p}_0 = p_0(\cos\phi, \sin\phi)$, cf figure 9(a), which is large with respect to $\boldsymbol{K}$, i.e. the K point of the untilted graphene band. Then we consider the circle of equal energies with radius $v_F p_0$ passing through momentum $\boldsymbol{p}_0$. In the limit of $p_0 \to \infty$, the fraction of the circle in the vicinity of momenta $\boldsymbol{p}_0$ becomes a straight line with angle $(\pi - \phi)$, cf the red line in figure 9(a). Then, the curvature of the circle is negligible.

The angle ϕ determines the level of incommensurability of the geometrical distances between the line and the sites of the momentum-space lattice. For example, for $\phi = \pi/3$, the distances are commensurate. To describe the physics near $\boldsymbol{p}_0$, we expand the Hamiltonian in equation (F4) in $\delta\boldsymbol{p} = \boldsymbol{p} - \boldsymbol{p}_0$ as $H_l^K(\boldsymbol{p}) \simeq H_l^K(\boldsymbol{p}_0) + \delta H_l^K(\boldsymbol{p})$, with $\delta H_l^K(\boldsymbol{p})$ reading

$$\delta H_l^K(\boldsymbol{p}) = v_F \langle \psi_{\boldsymbol{p}_0, l}^+ | (\delta\boldsymbol{p} + \boldsymbol{Q}_n - \boldsymbol{K}^l) \cdot \boldsymbol{\sigma}_{\pm\theta/2} | \psi_{\boldsymbol{p}_0, l}^+ \rangle$$
$$= \frac{v_F}{p_0} (\delta\boldsymbol{p} + \boldsymbol{Q}_n - \boldsymbol{K}^l) \cdot \boldsymbol{p}_0, \qquad (G1)$$

where $l \in \{U, L\}$ and $|\psi_{\boldsymbol{p}_0, U/L}^+\rangle = (\exp(-i(\phi \mp \frac{\theta}{2})), 1)^T$ is an eigenstate of $H_{U/L}^K(\boldsymbol{p}_0)$ with

positive energy $v_F p_0$. Likewise, the effective interlayer coupling is given by

$$\delta H_T(\boldsymbol{p}_0) = \langle \psi_{\boldsymbol{p}_0, U}^+ | H_T | \psi_{\boldsymbol{p}_0, L}^+ \rangle + \text{H.c.}. \qquad (G2)$$

Using the effective Hamiltonian $(\delta H_{U/L}^K(\boldsymbol{p}_0) + \delta H_T(\boldsymbol{p}_0))$, we compute the localization length $\xi_\parallel$ along the straight line, where $\xi_\parallel$ is defined as the inverse of the projected IPR on the line. We show $\xi_\parallel$ as a function of the angle ϕ in figure 9(b).

Interestingly, for commensurate angles ϕ—as, e.g. found at multiples of $\pi/6$—the localization length exhibits a sharp peak, clearly showing that localization occurs due to incommensuration effects of the underlying moiré lattice. Whenever a wave function crosses the region of an incommensurate angle it gets tied up there, leading to localization. Figure 9(c) shows that the localization length tentatively increases linearly with increasing $t_\perp$, however, with an additional oscillation. The oscillation effect occurs as a function of angle and we did not resolve its origin, here.

Appendix H. Corrugation effect

Finally, we show that in the weak corrugation case, an additional localization mechanism occurs along

the energy circle. Here we address the effect of the corrugation within the momentum space continuum model. This effect is due to interference between the interlayer couplings in the regions of AA stacking and AB stacking. We note that in the main text, we use a strong corrugation case, i.e. $t^{AA} = 0$ in equation (F8), to avoid this effect as it is not of primary interest for the present work.

The corrugation can effectively be taken into account in the interlayer coupling terms of the continuum model, cf equation (F8). Using an angular parameter φ, it can be continuously tuned from weak to strong as $t^{AA} = t_\perp \cos\varphi$ and $t^{AB} = t_\perp \sin\varphi$.

To see how neighboring lattice points are connected by the interlayer coupling, we employ a perturbative approach. First, we calculate the eigenvectors of the Hamiltonian in the absence of the interlayer coupling $|\boldsymbol{k} + \boldsymbol{Q}_n, \alpha = U, L\rangle$. We then compute matrix element $\langle \boldsymbol{k} + \boldsymbol{Q}'_n, U|H_T|\boldsymbol{k} + \boldsymbol{Q}_n, L\rangle$. The absolute value of the matrix element is shown in figure 10(a) and is represented by the thickness of the connections between lattice points.

Without corrugation ($\varphi = \pi/4$), the connections in the angular direction are suppressed due to the formation of dimer states in comparison with the connections in the radial direction. On the other hand, with strong corrugation, i.e. $\varphi = 0$ and $\varphi = \pi/2$, the connection along a contour of equal energy is relatively strong. While three nearest neighbor connections are equivalent in the $\varphi = 0$ ($t^{AB} = 0$) case, the connection in the $\varphi = \pi/2$ ($t^{AA} = 0$) dominantly occurs along the angular direction. Strong dependence on the corrugation suggests an interference effect between the AA and AB interlayer coupling.

The above localization mechanism is further supported by a calculation of the localization length along a ring, figure 10(b). For decreasing corrugation, i.e. closer to $\varphi = \pi/4$, the localization length decreases. This clearly shows that there is an additional localization mechanism in the weak corrugation case.

ORCID iDs

Jan Attig ● https://orcid.org/0000-0002-3762-6010
Jinhong Park ● https://orcid.org/0000-0003-3570-6356
Michael M Scherer ● https://orcid.org/0000-0003-0766-9949
Simon Trebst ● https://orcid.org/0000-0002-1479-9736
Alexander Altland ● https://orcid.org/0000-0002-2991-4805
Achim Rosch ● https://orcid.org/0000-0002-6586-5721

References

[1] Hunt B *et al* 2013 Massive Dirac fermions and Hofstadter butterfly in a van der Waals heterostructure *Science* **340** 1427

[2] MacDonald A H and Bistritzer R 2011 Graphene moiré mystery solved? *Nature* **474** 453

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

[4] Yankowitz M *et al* 2018 Dynamic band-structure tuning of graphene moiré superlattices with pressure *Nature* **557** 404

[5] Rode J C, Smirnov D, Belke C, Schmidt H and Haug R J 2017 Twisted bilayer graphene: interlayer configuration and magneto transport signatures *Ann. Phys., Lpz.* **529** 1700025

[6] Yankowitz M, Xue J, Cormode D, Sanchez-Yamagishi J D, Watanabe K, Taniguchi T, Jarillo-Herrero P, Jacquod P and LeRoy B J 2012 Emergence of superlattice Dirac points in graphene on hexagonal boron nitride *Nat. Phys.* **8** 382

[7] Balents L, Dean C R, Efetov D K and Young A F 2020 Superconductivity and strong correlations in moiré flat bands *Nat. Phys.* **16** 725

[8] Can O, Tummuru T, Day R P, Elfimov I, Damascelli A and Franz M 2021 High-temperature topological superconductivity in twisted double-layer copper oxides *Nat. Phys.* **17** 519–24

[9] Cao Y *et al* 2018 Correlated insulator behaviour at half-filling in magic-angle graphene superlattices *Nature* **556** 80

[10] Cao Y, Fatemi V, Fang S, Watanabe K, Taniguchi T, Kaxiras E and Jarillo-Herrero P 2018 Unconventional superconductivity in magic-angle graphene superlattices *Nature* **556** 43

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

[12] Choi Y *et al* 2019 Electronic correlations in twisted bilayer graphene near the magic angle *Nat. Phys.* **15** 1174

[13] Lu X *et al* 2019 Superconductors, orbital magnets and correlated states in magic-angle bilayer graphene *Nature* **574** 653

[14] Kerelsky A *et al* 2019 Maximized electron interactions at the magic angle in twisted bilayer graphene *Nature* **572** 95

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

[16] Chen G *et al* 2019 Signatures of tunable superconductivity in a trilayer graphene moiré superlattice *Nature* **572** 215

[17] Chen G *et al* 2019 Evidence of a gate-tunable Mott insulator in a trilayer graphene moiré superlattice *Nat. Phys.* **15** 237

[18] Liu X *et al* 2020 Tunable spin-polarized correlated states in twisted double bilayer graphene *Nature* **583** 221

[19] Shen C *et al* 2020 Correlated states in twisted double bilayer graphene *Nat. Phys.* **16** 520

[20] Burg G W, Zhu J, Taniguchi T, Watanabe K, MacDonald A H and Tutuc E 2019 Correlated insulating states in twisted double bilayer graphene *Phys. Rev. Lett.* **123** 197702

[21] He M, Li Y, Cai J, Liu Y, Watanabe K, Taniguchi T, Xu X and Yankowitz M 2021 Symmetry breaking in twisted double bilayer graphene *Nat. Phys.* **17** 26

[22] Xian L, Kennes D M, Tancogne-Dejean N, Altarelli M and Rubio A 2019 Multiflat bands and strong correlations in twisted bilayer boron nitride: doping-induced correlated insulator and superconductor *Nano Lett.* **19** 4934

[23] Ni G X *et al* 2019 Soliton superlattices in twisted hexagonal boron nitride *Nat. Commun.* **10** 4360

[24] Regan E C *et al* 2020 Mott and generalized Wigner crystal states in WSe2/WS2 moiré superlattices *Nature* **579** 359

[25] Wang L *et al* 2020 Correlated electronic phases in twisted bilayer transition metal dichalcogenides *Nat. Mater.* **19** 861

[26] Tang Y *et al* 2020 Simulation of Hubbard model physics in WSe_2/WS_2 moiré superlattices *Nature* **579** 353

[27] Xian L, Claassen M, Kiese D, Scherer M M, Trebst S, Kennes D M and Rubio A 2020 Realization of nearly dispersionless bands with strong orbital anisotropy from destructive interference in twisted bilayer MoS_2 (arXiv:2004.02964)

[28] Kennes D M, Claassen M, Xian L, Georges A, Millis A J, Hone J, Dean C R, Basov D, Pasupathy A and Rubio A 2020 Moiré heterostructures: a condensed matter quantum simulator (arXiv:2011.12638)

[29] Wigner E P 1955 Characteristic vectors of bordered matrices with infinite dimensions *Ann. Math.* **62** 548

[30] Guhr T, Müller-Groeling A and Weidenmüller H A 1998 Random-matrix theories in quantum physics: common concepts *Phys. Rep.* **299** 189

[31] Mehta M L 2004 *Random Matrices* (Amsterdam: Elsevier)

[32] Quantum phase transitions of the the magic-angle flat bands induced by delocalization in momentum space are discussed in [34–37]. Here, we investigate the effect of momentum localization on generic bands away from the magic-angle flat bands

[33] Tarnopolsky G, Kruchkov A J and Vishwanath A 2019 Origin of magic angles in twisted bilayer graphene *Phys. Rev. Lett.* **122** 106405

[34] Fu Y, König E J, Wilson J H, Chou Y-Z and Pixley J H 2020 Magic-angle semimetals *npj Quantum Mater.* **5** 1

[35] Chou Y-Z, Fu Y, Wilson J H, König E J and Pixley J H 2020 Magic-angle semimetals with chiral symmetry *Phys. Rev. B* **101** 235121

[36] Pixley J H, Wilson J H, Huse D A and Gopalakrishnan S 2018 Weyl semimetal to metal phase transitions driven by quasiperiodic potentials *Phys. Rev. Lett.* **120** 207604

[37] Gonçalves M, Olyaei H Z, Amorim B, Mondaini R, Ribeiro P and Castro E V 2020 Incommensurability-induced sub-ballistic narrow-band-states in twisted bilayer graphene (arXiv:2008.07542)

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

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

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

[41] We focus on a strong corrugation limit where the interlayer coupling in the AA region is zero since in the weak corrugation case, an additional interference effect from the interplay between the couplings in the AA and AB region can lead to another form of localization, which is not of primary interest in the present work. This interference driven localization effect is discussed in appendix H

[42] Evers F and Mirlin A D 2008 Anderson transitions *Rev. Mod. Phys.* **80** 1355

[43] Oganesyan V and Huse D A 2007 Localization of interacting fermions at high temperature *Phys. Rev. B* **75** 155111

[44] Using the Kullback–Leibler divergence as a quantitative measure for describing the level statistics has been introduced in two recent studies [46, 47] where it has also been shown [46] that visual inspections of the level statistics may trick one into false conclusions

[45] Mézard M and Montanari A 2009 *Information, Physics and Computation* (Oxford: Oxford University Press)

[46] Monteiro F, Micklitz T, Tezuka M and Altland A 2021 Minimal model of many-body localization *Phys. Rev. Res.* **3** 013023

[47] Berke C, Varvelis E, Trebst S, Altland A and DiVincenzo D P 2020 Transmon platform for quantum computing challenged by chaotic fluctuations (arXiv:2012.05923)

[48] Castro Neto A H, Guinea F, Peres N M R, Novoselov K S and Geim A K 2009 The electronic properties of graphene *Rev. Mod. Phys.* **81** 109

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

[50] Koshino M, Yuan N F Q, Koretsune T, Ochi M, Kuroki K and Fu L 2018 Maximally localized Wannier orbitals and the extended Hubbard model for twisted bilayer graphene *Phys. Rev. X* **8** 031087

[51] Lee J-K, Lee S-C, Ahn J-P, Kim S-C, Wilson J I B and John P 2008 The growth of aa graphite on (111) diamond *J. Chem. Phys.* **129** 234709

[52] Uchida K, Furuya S, Iwata J-I and Oshiyama A 2014 Atomic corrugation and electron localization due to Moiré patterns in twisted bilayer graphenes *Phys. Rev. B* **90** 155451

[53] We note that the full argument for the two-dimensional case, based on the consideration of both basis vectors of the reciprocal lattice, $G_m, 1$ and $G_m, 2$, works analogously