Randomly twisted bilayer graphene: Cascade transitions

Baruch Horovitz [1] and Pierre Le Doussal [2]

[1]*Department of Physics, Ben Gurion University of the Negev, Beer Sheva 84105, Israel*

[2]*Laboratoire de Physique de l'Ecole Normale Supérieure, CNRS, ENS and PSL Université, Sorbonne Université, Université Paris Cité, 24 rue Lhomond, 75005 Paris, France,*

(Received 7 September 2025; revised 29 April 2026; accepted 19 May 2026; published 5 June 2026)

Twisted bilayer graphene (TBG) is known to have disorder in its twist angle. We show that in terms of a Dirac equation with a random gauge potential $\mathbf{A}(\mathbf{r})$, this disorder becomes huge when the average twist angle is near the magic angle where the Dirac velocity vanishes. The density of states then diverges at the Dirac point as $\rho(E) \sim E^{(2/z)-1}$ with $z \gg 1$, and we deduce that electrons occupy energies very near $E = 0$. We solve a strong disorder problem by deriving a sum rule on the disorder averaged eigenfunctions from which we deduce that each added electron contributes equal intraband Coulomb interaction energy. The various bands in TBG are related by either $\mathbf{A}(\mathbf{r}) \to \mathbf{A}(-\mathbf{r})$ or $\mathbf{A}(\mathbf{r}) \to -\mathbf{A}(\mathbf{r})$, which affects the interband interaction energy. We find, within Hartree-Fock, jumps in the chemical potential at each integer filling, accounting for the cascade transitions.

DOI: 10.1103/c6rt-6qg1

I. INTRODUCTION

Twisted bilayer graphene (TBG) is extensively studied as a unique platform for strongly correlated electrons. It has been shown that near the corners of the moiré Brillouin zone the electrons obey a 2D Dirac equation with a velocity that vanishes [1] at a magic twist angle of $\theta_m \approx 1.1^\circ$. By approaching θ_m the interaction vs kinetic energies is easily tuned. The remarkable cascade transitions [2–5] show a sequence of peaks in the inverse compressibility $d\mu/d\nu$ when the filling factor of electrons in a moiré unit cell is near an integer $|\nu| = 0, 1, 2, 3$, corresponding to several flavors, each having a Dirac point. We note that the shape of the cascade transitions depends markedly on the size of the probed system, a pattern that recurs in symmerically twisted trilayer graphene (TTG) [6,7]. The peaks are broad and shifted from integer ν when probing $\approx$0.5 μm [2] or $\approx 0.1\mu$ in TTG [6], while they are sharper and more accurate at integer ν when probing larger areas by either directly averaging over $\approx$2 μm [3] or by a global measurements over a few micrometers [4,5] and up to $\approx$10 μm in TTG [7]. We propose that this size dependence is related to the disorder in the twist angle. The latter was studied in TBG and shown to be present on scales $\gtrsim$0.5 μm [8]. There has been a considerable theoretical progress in understanding TBG and its cascade transitions [1,2,4,9]; however, none of these theories includes twist angle disorder. We claim that disorder average is essential in any long-scale measurement such as those in global experiments or in transport data. In the present work we solve a strong disorder problem. We first show that the twist angle disorder generates a random gauge potential $\mathbf{A}(\mathbf{r})$ within a 2D Dirac equation. It can be taken as transverse [10], i.e., $A_x = \partial_y V$, $A_y = -\partial_x V$, with potential V and effective magnetic field $\mathsf{B} = \nabla^2 V$. The disorder average, $\overline{\mathsf{B}(\mathbf{q})\mathsf{B}(\mathbf{q}')} = \pi\sigma q^2 (2\pi)^2 \delta(\mathbf{q}+\mathbf{q}')$ in Fourier space, defines the dimensionless disorder parameter σ, which, as we show, becomes huge near the magic angle. It is known that gauge disorder produces an anomalous density of states (DOS) [11] $\rho(E) \sim E^{(2/z)-1}$, which undergoes a freezing transition, and becomes divergent at large σ, with $z \sim \sqrt{\sigma}$ [12]. We infer then that the electrons occupy states very near $E = 0$ for each of the flavors. Furthermore, we prove a sum rule for a disorder average of the eigenfunctions of the Dirac equation, from which we deduce that each added electron carries equal average Coulomb interaction energy. Using known symmetries and explicitly known $E = 0$ solutions, we deduce the average interaction also between the various flavors. Finally, we evaluate the Hartree-Fock interaction energy and show that the chemical potential μ jumps at each integer ν, accounting for the main feature of the cascade transitions. The Appendixes provide details on symmetries and on the reduced Hamiltonian for both TBG and TTG, on density overlaps of the explicitly known $E = 0$ states, and on disorder averages of ρ^2 that lead to the sum rules.

II. TWIST ANGLE DISORDER

We parametrize the deviation of the twist angle from its mean value as $\delta\theta(\mathbf{r})$, and its disorder average by

$$\overline{\delta\theta(\mathbf{r}+\delta\mathbf{r})\delta\theta(\mathbf{r})} = \gamma \left(\frac{a}{L_m}\right)^2 \mathrm{e}^{-|\delta\mathbf{r}|/\ell_0}, \tag{1}$$

where a, L_m are the lattice constants of a single layer and of the bilayer moiré lattice, respectively, with $L_m \approx 13$ nm, and ℓ_0 is a correlation length. Reference [8] provides detailed analysis of the $\delta\theta(\mathbf{r})$ distribution, showing two samples with standard deviations of 0.025° and 0.022°. Since $\frac{a}{L_m} = 2\sin\frac{1}{2}\theta \approx \theta_m$, we infer $\gamma \approx 5 \times 10^{-4}$. Assuming that ℓ_0 is dominated by the maximal $\nabla\theta$ these samples imply $\ell_0 =$ 0.13–0.21 μm, alternatively taking their "typical gradient" of 0.05°/μm implies $\ell_0 \approx 0.5$ μm. The onset of Landau fans at $B \approx 1$ T, corresponding to a cyclotron resonance of 30 nm, indicates a shorter ℓ_0.

 ©2026 American Physical Society

Consider next the resulting strain. A single graphene layer whose unit cell is tilted by an angle $\delta\theta(\mathbf{r})$ has a shear strain [13]

$$\tfrac{1}{2}(\partial_x u_y - \partial_y u_x) = \delta\theta. \tag{2}$$

This relation involves the transverse component of the displacement pattern (u_x, u_y), whose Fourier components are $u_\alpha(\mathbf{q}) = 2i\epsilon_{\alpha\beta}q_\beta\delta\theta(\mathbf{q})/q^2$. The strain implies that the nearest-neighbor C-C spacing $a' = a/\sqrt{3}$ as well as the corresponding transfer integral are modulated; hence, the Dirac equation acquires a gauge vector potential [14–17]

$$\tilde{A}_x(\mathbf{q}) = i\frac{\beta}{2a'}(q_x u_x - q_y u_y) = -\frac{2\beta}{a'}\frac{q_x q_y}{q^2}\delta\theta$$
$$\tilde{A}_y(\mathbf{q}) = -i\frac{\beta}{2a'}(q_y u_x + q_x u_y) = \frac{\beta}{a'}\frac{q_y^2 - q_x^2}{q^2}\delta\theta, \tag{3}$$

where $\beta \approx 3$ (Ref. [14]) leading to a fictitious magnetic field $\mathsf{B}(\mathbf{q}) = iq_x\frac{3q_y^2-q_x^2}{q^2}\frac{\beta}{a'}\delta\theta(\mathbf{q})$. Upon angular averaging for the scale $\ell_0 \gg L_m$ we can identify the disorder parameter as

$$\pi\tilde{\sigma} \equiv \int d^2r \sum_{j=x,y} \overline{\tilde{A}_j(\mathbf{r})\tilde{A}_j(\mathbf{r}')} = \frac{1}{2}\left(\frac{\beta}{a'}\right)^2 \int d^2r \overline{\delta\theta(\mathbf{r})\delta\theta(0)}$$
$$\approx 27\pi\gamma\left(\frac{\ell_0}{L_m}\right)^2. \tag{4}$$

Each layer has actually $\frac{1}{2}\delta\theta$ with $\gamma = 5\times10^{-4}/4$; hence, for the range $\ell_0 = 200$–500 nm we estimate $\tilde{\sigma} = 0.8$–5 per layer. We note that the strain due to lattice relaxation [15,17] is also described by a gauge potential which is, however, periodic on the moiré lattice, while in our case $\mathbf{A}(\mathbf{r})$ varies on the scale $\ell_0 \gg L_m$.

Consider next a one-layer Hamiltonian $v_F\boldsymbol{\sigma}\cdot\mathbf{k}$ with velocity v_F, Pauli matrices $\boldsymbol{\sigma}$, and momentum $\mathbf{k}$ measured from a chosen moiré Brillouin-zone corner (a Dirac point). The perturbative approach [1,15,17] (Appendix A) couples this Dirac point by wave vectors $\mathbf{q}_i$ to the three neighboring corners that correspond to states on the other layer. This interlayer tunneling is u_0 on AA regimes and u_1 on AB regimes (AA or AB refers to the relative arrangement of the two layers). The interlayer coupling contributes an opposite sign $\sim -v_F\boldsymbol{\sigma}\cdot\mathbf{k}$; hence, the effective velocity v vanishes at θ_m. Including $\tilde{\mathbf{A}}(\mathbf{r})$ in this derivation is valid since $\ell_0 \gg L_m$, i.e., the $\mathbf{q}_i$ apply on a scale where $\tilde{\mathbf{A}}(\mathbf{r})$ is locally constant. Since $\tilde{\mathbf{A}}(\mathbf{r}) \sim \delta\theta(\mathbf{r})$ has opposite signs on the two layers, the near cancellation of v_F does not apply to, leading to an effective Hamiltonian:

$$\mathcal{H}_{\rm eff} = v\boldsymbol{\sigma}\cdot[-i\nabla - \mathbf{A}(\mathbf{r})]$$
$$v = v_F\frac{1-3\bar{u}_1^2}{1+3(\bar{u}_0^2+\bar{u}_1^2)},$$
$$\mathbf{A}(\mathbf{r}) = \frac{1+3\bar{u}_1^2}{1-3\bar{u}_1^2}\tilde{\mathbf{A}}(\mathbf{r}), \tag{5}$$

where $\bar{u}_i = u_i\frac{3a}{4\pi v_F\theta}$, $i = 1, 2$. The velocity v coincides with the original [1] result when $u_0 = u_1$. It indicates the presence of a magic angle, and is in fact very close to the exact (numerical) result [1,15,17] for $\mathbf{A} = 0$. The same form of $\mathcal{H}_{\rm eff}$ is found also for symmetric TTG (Appendix A). We note that in deriving $\mathcal{H}_{\rm eff}$ factors $e^{\pm i\theta\sigma_z/4}$ that can lead to non-Abelian disorder [18] are neglected here and elsewhere [15,17,19] since θ is small (Appendix A).

TABLE I. Flavors and eigenstates.

Symmetry relation	Flavor Hamiltonians $\mathcal{H}_{v_m,v}$	Eigenstates $E\to-0$	Eigenstates $E\to+0$
1	$\mathcal{H}_{+,+} = v\boldsymbol{\sigma}\cdot[-i\nabla - \mathbf{A}(\mathbf{r})]$	$\begin{pmatrix} e^{V(\mathbf{r})} \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ e^{-V(\mathbf{r})} \end{pmatrix}$
η_x	$\mathcal{H}_{-,+} = v\boldsymbol{\sigma}\cdot[-i\nabla + \mathbf{A}(\mathbf{r})]$	$\begin{pmatrix} e^{-V(\mathbf{r})} \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ e^{V(\mathbf{r})} \end{pmatrix}$
$\sigma_x R$	$\mathcal{H}_{+,-} = v\boldsymbol{\sigma}^*\cdot[i\nabla - \mathbf{A}(-\mathbf{r})]$	$\begin{pmatrix} 0 \\ e^{V(-\mathbf{r})} \end{pmatrix}$	$\begin{pmatrix} e^{-V(-\mathbf{r})} \\ 0 \end{pmatrix}$
$\sigma_x R\eta_x$	$\mathcal{H}_{-,-} = v\boldsymbol{\sigma}^*\cdot[i\nabla + \mathbf{A}(-\mathbf{r})]$	$\begin{pmatrix} 0 \\ e^{-V(-\mathbf{r})} \end{pmatrix}$	$\begin{pmatrix} e^{V(-\mathbf{r})} \\ 0 \end{pmatrix}$

It is remarkable that at the magic angle $3\bar{u}_1^2 \to 1$ the coefficient of $\tilde{\mathbf{A}}(\mathbf{r})$ remains finite; thus, the disorder effect becomes stronger and would induce a divergent DOS [11] and a freezing transition [12] before reaching the magic angle. An expansion in θ yields $v(\theta+\delta\theta) = v(\theta) + v_F\frac{\delta\theta}{\theta}$. With the given standard deviation we consider $\delta\theta = 0.025^\circ$ as the closest approach to θ_m; the velocity is then reduced to $v \approx v_F\, 0.025/1.1 \approx v_F/40$. At this situation the disorder parameter $\pi\sigma = \int d^2r \sum_{j=x,y}\overline{A_j(\mathbf{r})A_j(\mathbf{r}')}$ is enhanced to $\sigma = 40^2\tilde{\sigma}$, i.e., given the previous estimate, σ can reach huge values of 10^3–10^4.

We proceed to identify the various flavors and their symmetry relations (also in Appendix A) related to the symmetries of the ordered system [19,20]. The operator η_x for exchanging the two layers results in the same Hamiltonian equation (5) except that $\mathbf{A}(\mathbf{r})$ changes sign; hence, $\eta_x = \{\mathbf{A}(\mathbf{r}) \to -\mathbf{A}(\mathbf{r})\}$. This symmetry relates the two moiré valleys, a flavor denoted as $v_m = \pm$. Consider next states built from distinct graphene valleys $v = \pm$ via complex conjugation and space reflection R; hence, $\boldsymbol{\sigma}\to\boldsymbol{\sigma}^*$, $\mathbf{A}(\mathbf{r})\to\mathbf{A}(-\mathbf{r})$. Adding finally the electron spin $s = \pm$ we end up with 3 flavors v_m, v, s and therefore 8 bands, consistent with the $\mathbf{A}(\mathbf{r}) = 0$ case [21]. A significant operation is the particle hole (p-h) symmetry, $C = \eta_x\sigma_x\mathcal{K}$, where $\mathcal{K}$ is complex conjugation, since $C\mathcal{H}_{\rm eff}C = -\mathcal{H}_{\rm eff}$, similar to the $\mathbf{A}(\mathbf{r}) = 0$ case [19,20]. We also note a time-reversal symmetry operator $\sigma_y\eta_y\mathcal{K}$ and a chiral symmetry $\sigma_z\eta_z$ as in similar cases [22]. Thus, our system of 8 Dirac points corresponds to class BDI.

The $E = 0$ explicit eigenfunctions are known [11]; hence, the symmetries determine these also for the other flavors. In Table I we summarize the symmetry relations, the Hamiltonian of each flavor, and the exact $E = 0$ eigenstates in terms of $V(\mathbf{r})$, where we recall that $A_x(\mathbf{r}) = \partial_y V(\mathbf{r})$, $A_y(\mathbf{r}) = -\partial_x V(\mathbf{r})$, and the spin variable s is implicit, unaffecting the Hamiltonians. Each Hamiltonian is 2×2 in its subspace, i.e., $\boldsymbol{\sigma}$ act on the AA and AB states of the moiré unit cell; thus, the whole 3-flavor system is a 16×16 matrix. Table I shows two eigenstates for each of the flavors, corresponding to $E\to\pm0$, related by the p-h operator C. The third column in the table is chosen to represent $E\to-0$ states as generated from each other by the symmetries in the first column, while

the fourth column represents $E \to +0$ states (the third and fourth columns may be interchanged as a whole, unaffecting the following results).

For simplified reading one may proceed directly to Eqs. (12) and (17) for just the $E = 0$ states that can be straighforwardly derived, and then proceed to the Hartree-Fock section. The purpose of the next sections on DOS and sum rules is to prove that Eq. (17) is valid for all eigenstates.

III. DOS

The local DOS is found by projecting the fermion action into the subspace of energy $E + i\epsilon$ (E, ϵ real) and then bosonizing [11,12]. The disorder is averaged with the replica method, generating boson fields $\phi_a(\mathbf{r})$ with $a = 1, ..., m$, eventually $m \to 0$. The noninteracting bosonic action is

$$\mathcal{S}_0 = \int d^2r \frac{1}{8\pi} \sum_{a,b} (K^{-1}\delta_{a,b} + \sigma)\nabla\phi_a \cdot \nabla\phi_b, \tag{6}$$

where the K parameter is a conventional generalization [23], although up to now $K = 1$. The system is solved with $E = 0$, where the bare interaction $\frac{\epsilon}{2\pi\alpha v}\sum_a \int d^2r e^{i\phi_a(\mathbf{r})} + \text{H.c.}$ involves a cutoff α of the fermion-boson transformation [23], $\alpha \approx L_m$. Since the disorder sets in at a longer scale ℓ_0 we first integrate out wave vectors in the range $1/\ell_0 < q < 1/\alpha$ so that $\langle e^{i\phi_a(\mathbf{r})}\rangle_0 = (\frac{\alpha}{\ell_0})^K e^{i\phi_a(\mathbf{r})}$ with $\phi_a(\mathbf{r})$ now varies on scales $\gtrsim \ell_0$. Thus, all states correspond now to wave vectors $q \lesssim 1/\ell_0$ or densities $\lesssim 1/\ell_0^2$. The bare interaction becomes $\sum_a \int \frac{d^2r}{\ell_0^2}\tilde{\epsilon} e^{i\phi_a(\mathbf{r})} + \text{H.c.}$, where $\tilde{\epsilon} = \frac{\epsilon\alpha}{2\pi v}(\frac{\alpha}{\ell_0})^{K-2}$. We note that deviations from a linear dispersion at $q > 1/\ell_0$ could change the definition of $\tilde{\epsilon}$. In the following we use $\tilde{\epsilon}$ as the input for further disorder renormalization on scales $q < 1/\ell_0$.

To treat the strong disorder case properly and obtain the correct scaling dimension of $\tilde{\epsilon}\cos\phi_a(\mathbf{r})$, we include from the start all higher-order terms as generated by renormalization group [24–28]. The interaction is then $\sum_{\{\mathbf{n}\}} Y[\mathbf{n}] e^{i\mathbf{n}\cdot\boldsymbol{\phi}}$, where $\{\mathbf{n}\}$ is a set of vectors of length m and entries $0, \pm 1$, $\boldsymbol{\phi} = \{\phi_1(\mathbf{r}), ..., \phi_m(\mathbf{r})\}$, and $Y[\mathbf{n}] = \Pi_a \tilde{\epsilon}^{n_a^2}$; the bare interaction has $\sum_a n_a^2 = 1$. The bosonic action is then

$$\mathcal{S}_B = \mathcal{S}_0 - \int \frac{d^2r}{\ell_0^2} \sum_{\{\mathbf{n}\}} Y[\mathbf{n}] e^{i\mathbf{n}\cdot\boldsymbol{\phi}}. \tag{7}$$

The DOS, either in terms of the eigenfunctions of the Dirac Hamiltonian $\varphi_n(\mathbf{r})$ and their eigenvalues E_n or in terms of the bosonic variables [12], is

$$\begin{aligned}\overline{\rho_\epsilon(E,\mathbf{r})} &= -\frac{1}{\pi}\text{Im}\overline{\sum_n \frac{1}{E - E_n + i\epsilon}|\varphi_n(\mathbf{r})|^2} \\ &= \frac{1}{\pi\ell_0^2}\frac{\partial}{\partial\epsilon}\left\langle \sum_{\{\mathbf{n}\}} Y[\mathbf{n}] e^{i\mathbf{n}\cdot\boldsymbol{\phi}}\right\rangle,\end{aligned} \tag{8}$$

where $\langle ...\rangle$ averages $\phi_a(\mathbf{r})$ with S_B; the notation $|\varphi_n(\mathbf{r})|^2 \equiv \varphi_n^\dagger(\mathbf{r}) \cdot \varphi_n(\mathbf{r})$ means a spinor product. We evaluate $\rho_\epsilon(\mathbf{r}) = \rho_\epsilon(E = 0, \mathbf{r})$ with a variational method and summarize here the results [12]. The interaction is replaced by $\sum_{a,b} \frac{1}{2}[\sigma_c(K,\tilde{\epsilon})\delta_{ab} + \sigma_0(K,\tilde{\epsilon})]\phi_a(\mathbf{r})\phi_b(\mathbf{r})$ and the variational condition determines $\sigma_0(K,\tilde{\epsilon}) \lesssim \sigma_c(K,\tilde{\epsilon}) \approx \ell_0^{-2}\tilde{\epsilon}^{2/z}$ up to a numerical prefactor. The exponent is $z = K(\sqrt{8\sigma} - 1)$ for strong disorder $\sigma > 2/K^2$, while $z = 2 - K + \sigma K^2$ for weak disorder. We note that in the classical limit ($K = 0$, any σ), where $\cos\phi_a(\mathbf{r})$ is expanded near its minimum, $\sigma_c(K = 0, \tilde{\epsilon}) = \frac{\epsilon}{2\pi v\alpha}$ is correctly reproduced. Furthermore, for $\sigma = 0$ $\sigma_c(K,\tilde{\epsilon})$ is ℓ_0 independent, as it should be.

The average DOS (8) is then

$$\overline{\rho_\epsilon(\mathbf{r})} \approx \frac{1}{2\pi^2 v\alpha}\left(\frac{\alpha}{\ell_0}\right)^K \tilde{\epsilon}^{(2/z)-1}. \tag{9}$$

To find the conventional DOS $\rho(E,\mathbf{r}) = \sum_n \delta(E - E_n)|\varphi_n(\mathbf{r})|^2$, we note the relation

$$\begin{aligned}\rho_\epsilon(\mathbf{r}) &= \frac{1}{\pi}\int dE \frac{\epsilon}{\epsilon^2 + E^2}\rho(E,\mathbf{r}) \\ \Rightarrow \overline{\rho(E,\mathbf{r})} &\approx \frac{\sin(\pi/z)}{2\pi^2 v\alpha}\left(\frac{\alpha}{\ell_0}\right)^K \tilde{E}^{(2/z)-1},\end{aligned} \tag{10}$$

where $\tilde{E} = \frac{E\alpha}{2\pi v}(\frac{\alpha}{\ell_0})^{K-2}$; hence, the DOS diverges at $E = 0$ when $z > 2$. The average density at a chemical potential μ is then

$$\begin{aligned}\overline{\rho(\mathbf{r})} &= \int_0^\mu \overline{\rho(E,\mathbf{r})}dE = \frac{\sin\pi/z}{\pi\ell_0^2}\int_0^{\tilde{\mu}} \tilde{E}^{(2/z)-1}d\tilde{E} \\ &\approx \frac{1}{2\ell_0^2}\tilde{\mu}^{2/z},\end{aligned} \tag{11}$$

where in the last relation $z \gg 1$ and $\tilde{\mu} = \frac{\mu\alpha}{2\pi v}(\frac{\alpha}{\ell_0})^{K-2}$. Hence, μ is flat for the admissible densities $\overline{\rho(\mathbf{r})} \lesssim 1/\ell_0^2$, reflecting the divergent DOS.

IV. SUM RULES

Our aim in this section is to evaluate density overlaps, i.e., disorder average $\overline{|\varphi_i(\mathbf{r})|^2|\varphi_j(\mathbf{r})|^2}$ for the various 2-spinor eigenstates of the Dirac equation (5). We consider first the $E = 0$ states listed in Table I. These are evaluated in Appendix B (including their normalization factors) using replica methods, aiming to find their L dependence. The results in terms of the scaler $V(\mathbf{r})$ are, for $\sigma > \sigma_c = 2$,

$$\overline{\int d^2r e^{4V(\mathbf{r})}/N_+^2} \approx \frac{1}{\ell_0^2}, \tag{12a}$$

$$\overline{\int d^2r/N_+N_-} \approx \left(\frac{L}{\ell_0}\right)^{2-8\sqrt{\sigma/2}}\frac{1}{\ell_0^2} \to 0, \tag{12b}$$

$$\overline{\int d^2r e^{[2V(\mathbf{r})\pm 2V(-\mathbf{r})]}/N_+N_\pm} \approx \frac{1}{L^2}, \tag{12c}$$

where the normalization factors are $N_\pm = \int d^2r e^{\pm 2V(\mathbf{r})}$. These results determine readily the 64 density overlaps between all the states listed in Table I. We note that (12a) is actually the inverse participation ratio, which has been studied by various methods [25,29,30], consistent with the present result. We note further that (12b) shows that the intensity overlap of each state with its p-h conjugate, as well as same energy overlap of states with $v_m = \pm$ and same v, s are all vanishingly small.

We proceed to evaluate density overlaps of all $E \neq 0$ states by proving sum rules, derived by squaring both the fermionic

and bosonic forms of (8) and then averaging. Consider first the fermionic form of (8):

$$\rho_\epsilon^2(\mathbf{r}) = \frac{1}{\pi^2}\sum_{i,j}\frac{\epsilon^2}{(E_i^2+\epsilon^2)(E_j^2+\epsilon^2)}|\varphi_i(\mathbf{r})|^2|\varphi_j(\mathbf{r})|^2$$
$$\approx \frac{1}{\pi^2\epsilon^2}\sum_{i,j}^{|E_i|,|E_j|<\epsilon}|\varphi_i(\mathbf{r})|^2|\varphi_j(\mathbf{r})|^2, \tag{13}$$

so that ϵ can be interpreted as the chemical potential μ. The disorder average $\overline{\rho_\epsilon^2}$ in its bosonic form is in fact a two-layer problem [12] as detailed in Appendix C. We find

$$\overline{\rho_\epsilon^2} = \left(\frac{1}{\pi\ell_0^2}\right)^2\frac{\ell_0^2}{\epsilon^2}\sigma_c(2K,\tilde\epsilon^2) \approx \frac{2\overline{\rho(\mathbf{r})}}{\pi^2\ell_0^2\epsilon^2}, \tag{14}$$

where the exponent $z = K(\sqrt{8\sigma}-1) \to 2z$ so that $\sigma_c(2K,\tilde\epsilon^2) \approx \ell_0^{-2}(\tilde\epsilon^2)^{2/2z} = \ell_0^{-2}\tilde\mu^{2/z} = 2\overline{\rho(\mathbf{r})}$ using Eq. (11). We note that $\overline{\rho_\epsilon^2} = \frac{1}{L^2\ell_0^2\epsilon^2}f[\tilde\epsilon(\frac{L}{\ell_0})^z]$ satisfies a scaling form with $f(0)=1$, $f(x)\sim x^{2/z}$ at $x\gg 1$, interpolating between the single-particle (12a) and the finite-density case (14). A similar scaling for class AIII was proposed [31], though for individual terms of Eq. (13).

Equating the two forms (13), (14) yields a sum rule on the disorder average of all occupied states of the Dirac equation,

$$\sum_{i,j}^{|E_i|,|E_j|<\mu}\int d^2r\overline{|\varphi_i[\mathbf{A}(\mathbf{r}),\mathbf{r})]|^2|\varphi_j[\mathbf{A}(\mathbf{r}),\mathbf{r}]|^2} \approx \frac{N}{\pi\ell_0^2}, \tag{15}$$

where $\overline{\rho(\mathbf{r})}L^2 = N$ is the total number of electrons in the partial filled band with $E\geqslant 0$; $\mathbf{A}(\mathbf{r})$ is specified here to allow for other overlaps below. We note that with the single $E=0$ state the result with $N=1$ reproduces Eq. (12a); the latter was derived without using the fermion-boson transformation. The change $N\to N+1$ corresponds to an addition of one overlap in (15) with either $E_i=\pm E_j=\mu$. The sign choice is prompted by the known $E=0$ overlap and the shared symmetry of either $E>0$ or $E<0$ states. Since at equal $E\to 0$ the $E_i=E_j$ overlap is dominant while $E_i=-E_j$ is vanishingly small [Eq. (12b)], we deduce that all the terms in (15) are equal to $\delta_{i,j}/\pi\ell_0^2$.

Consider next the density overlaps between the states of $\mathcal{H}_{++}$ and $\mathcal{H}_{-+}$, generated by $\overline{\rho_\epsilon[\mathbf{A}(\mathbf{r})]\rho_\epsilon[-\mathbf{A}(\mathbf{r})]}$. This case also leads to a similar sum rule [interchanging $\phi_a^\pm$ in Eq. (C1)]:

$$\sum_{i,j}^{|E_i|,|E_j|<\mu}\int d^2r\overline{|\varphi_i[\mathbf{A}(\mathbf{r}),\mathbf{r}]|^2|\varphi_j[-\mathbf{A}(\mathbf{r}),\mathbf{r}]|^2} \approx \frac{N}{\pi\ell_0^2}. \tag{16}$$

In this case the$E\to+0$ state of $\mathcal{H}_{++}$ in Table I strongly overlaps with the $E\to-0$ state of $\mathcal{H}_{-+}$, while equal energy states have a vanishingly small overlap. Thus, we deduce in this case that the change $N\to N+1$ generates an $E_i=-E_j$ overlap and that all terms in (16) are equal to $\delta_{i,-j}/\pi\ell_0^2$, where $\delta_{i,-j}$ means $E_i=-E_j$.

Next consider the overlaps of $\mathcal{H}_{++}$ with $\mathcal{H}_{+-}$ states. The latter involve $\mathbf{r}\to-\mathbf{r}$; hence, the overlap is that of states localized at different positions and we expect $\int d^2r\overline{|\varphi_i[\mathbf{A}(\mathbf{r}),\mathbf{r}]|^2|\varphi_j[\mathbf{A}(-\mathbf{r}),-\mathbf{r}]|^2} \approx \delta_{i,j}/L^2$ consistent with Eq. (12c). In this case, however, we do not have an exact sum rule; yet this can be motivated by considering the overlap of densities $e^{2V(\mathbf{r})}$ and $e^{2V(\mathbf{r}-\mathbf{r}_0)}$ with $\mathbf{r}_0\gg\ell_0$, representing states localized at different positions. This overlap is evaluated in Eq. (B20) and yields the same result as Eq. (12c). The same reasoning and result apply to overlaps of $\mathcal{H}_{++}$ with $\mathcal{H}_{--}$ states.

Defining interaction parameters by multiplying the various overlaps by $U\ell_0^2$ (to be used in the next section), we finally obtain

$$U_1 = U\ell_0^2\int d^2r\overline{|\varphi_i[\mathbf{A}(\mathbf{r}),\mathbf{r})]|^2|\varphi_j[\mathbf{A}(\mathbf{r}),\mathbf{r}]|^2} \approx U\delta_{i,j}$$
$$U_1 = U\ell_0^2\int d^2r\overline{|\varphi_i[\mathbf{A}(\mathbf{r}),\mathbf{r})]|^2|\varphi_j[-\mathbf{A}(\mathbf{r}),\mathbf{r}]|^2} \approx U\delta_{i,-j}$$
$$U_2 = U\ell_0^2\int d^2r\overline{|\varphi_i[\mathbf{A}(\mathbf{r}),\mathbf{r})]|^2|\varphi_j[\mathbf{A}(-\mathbf{r}),-\mathbf{r}]|^2}$$
$$\approx U\frac{\ell_0^2}{L^2}\delta_{i,j}. \tag{17}$$

These results are fully consistent with the various overlaps in Eq. (12) of the $E=0$ states. For the latter states we can find the matrix elements also for small σ; see Eq. (B21).

V. HARTREE-FOCK

The Coulomb interaction $V_c(\mathbf{r})$ has a screening length [2] of $d/\epsilon_B\approx 10$ nm, where d is the distance to the gate; the ϵ_B is the dielectric constant of the spacer h-BN. Since the range of the eigenstates ℓ_0 is much longer we can use a local interaction defined by $U\equiv\int V_c(\mathbf{r})d^2r/\ell_0^2$. The average interaction is for occupied states in the flavors α (including spin) with i are states within each band, i.e., the eigenstates are now 16 long spinors $\varphi_{i\alpha}(\mathbf{r})$:

$$\langle\mathcal{H}\rangle_{\rm int} = \frac{1}{2}U\ell_0^2\int d^2r\sum_{ij,\alpha\beta}[\varphi_{i\alpha}^\dagger(\mathbf{r})\varphi_{i\alpha}(\mathbf{r})\varphi_{j\beta}^\dagger(\mathbf{r})\varphi_{j\beta}(\mathbf{r})$$
$$-\varphi_{j\beta}^\dagger(\mathbf{r})\varphi_{i\alpha}(\mathbf{r})\varphi_{i\alpha}^\dagger(\mathbf{r})\varphi_{j\beta}(\mathbf{r})]. \tag{18}$$

The first term is the direct (Hartree) interaction and the second one is the exchange (Fock) interaction. The spinor product implies that the exchange vanishes for $\alpha\neq\beta$ (orthogonal spinors). For $\alpha=\beta$ we note that within that subspace the eigenstates have the form $\tilde\varphi_i(\mathbf{r})b_i$, where b_i are 2-spinors and $\tilde\varphi_i(\mathbf{r})$ are scalars and the latter represent the AA and AB moiré lattice sites that are close on the scale ℓ_0; hence, they have a common $\mathbf{r}$ dependence. The disorder average involves, from the first Eq. (17), $\overline{|\tilde\varphi_i(\mathbf{r})|^2|\tilde\varphi_j(\mathbf{r})|^2}\sim\delta_{i,j}$ for both exchange and direct terms, which therefore precisely cancel. The disorder average Hartree-Fock energy is then

$$\overline{\langle\mathcal{H}_{\rm int}\rangle} = \frac{1}{2}U\ell_0^2\int d^2r\sum_{ij,\alpha\neq\beta}\overline{|\varphi_{i\alpha}|^2|\varphi_{j\beta}|^2}. \tag{19}$$

The initial interpretation [2] that has accounted for the main features of the cascade transitions has assumed that there are 4 effective bands, each allowing filling of $-1<\nu<1$. We find (consistent with [21]), however, that there are 8 noninteracting flavors each allowing filling of $-\frac{1}{2}<\nu<\frac{1}{2}$.

TABLE II. HF flavor filling and chemical potentials.

ν	States	μ
–4	$\uparrow++, \downarrow-+ \; E<0$	0
−3	$\uparrow+-, \downarrow-- \; E<0$	$2U_2$
−2	$\downarrow++, \uparrow-+ \; E<0$	U_1+2U_2
−1	$\downarrow+-, \uparrow-- \; E<0$	U_1+4U_2
0	$\uparrow++, \downarrow-+ \; E>0$	$2U_1+4U_2$
1	$\uparrow+-, \downarrow-- \; E>0$	$2U_1+6U_2$
2	$\downarrow++, \uparrow-+ \; E>0$	$3U_1+6U_2$
3	$\downarrow+-, \uparrow-- \; E>0$	$3U_1+8U_2$

This apparent contradiction is reconciled by our key result that the states with opposite v_m have vanishingly small interactions for all equal $v, s, \mathrm{sign}(E)$, as in Eq. (12b). In contrast, e.g., spin-↑ and spin-↓ states with the same $v_m, v, \mathrm{sign}(E)$ interact strongly as in Eq. (12a). Our system is then equivalent to the Hartree-Fock result (19) with 4 bands each having $-1 < \nu < 1$. We note that each of the effective bands is separated into 2 by $\mathrm{sign}(E)$ of their kinetic energy. This separation is not due to the kinetic energy itself, which is negligible, but due to the significantly different interaction terms as listed in (17).

Finally, we address the cascade transitions. The DOS divergence of each flavor implies that the occupied states are very near $E = 0$ and therefore the kinetic energies are negligible. Hence, we need just the direct term with $\alpha \neq \beta$ in (19) and use the interactions in Eq. (17). There are a number of choices for filling up the various flavor states, all leading to equivalent sequences of chemical potentials. In Table II we have chosen states with equal ↑, ↓ populations since spin polarization (in the absence of magnetic fields) is not observed [5,32]. We find that the stronger jumps are those into even fillings while the weaker jumps depend on the probed size L.

VI. CONCLUSIONS

Our first observation is the huge disorder parameter near the magic angle. We claim that averaging on twist angle disorder is essential in measurements on scales $L \gg \ell_0$, where $\ell_0 = 0.2$–$0.5\,\mu$m is the disorder correlation length in TBG [8], while less precisely known for TTG. Our second finding is a solution of a strong disorder problem with results on $E \neq 0$ eigenstates of the Dirac equation with random gauge in terms of sum rules and their individual terms, Eqs. (15)–(17). Finally, we apply a Hartree-Fock average to show jumps of the chemical potential at integer ν, as in the cascade transitions [2–7].

We recall that the shape of the cascade transitions is correlated with the measurement scale L as discussed in the Introduction. In particular, the peaks are broad and shifted from integer ν when probing ≈0.5 μm [2] or ≈0.1μ in TTG [6], while they are sharper and more accurately at integer ν when probing larger areas by either directly averaging over ≈2 μm [3] or by global measurements over a few micrometers [4,5] and up to ≈10 μm in TTG [7]. We propose then that the disorder average at $L \gg \ell_0$ leads to sharper transitions as in our Hartree-Fock result, accounting for the main feature of the data.

We finally note that the jump values in Table II are independent of σ, i.e., the cascade transitions should not change with the distance to the magic angle, as long as $\sigma > 2$. Upon increasing $|\theta - \theta_m|$ the disorder may reach $\sigma = 2$; at this elusive freezing transition we expect the cascade transitions to change considerably [see Eqs. (B21)].

In conclusion, we have shown that twist angle disorder is crucial for interpreting data on long scales $L \gg \ell_0$ near the magic angle. This interpretation reveals insights on eigenstates of the Dirac equation in the presence of a random gauge potential.

ACKNOWLEDGMENTS

We thank S. Ilani, E. Zeldov, E. Y. Andrei, A. Stern, F. Guinea, T. Giamarchi, D. Bernard, and E. Akkermans for highly stimulating discussions. P.L.D. thanks Ben-Gurion University Physics Department for hospitality. B.H. thanks the Laboratoire de Physique de l'Ecole Normale Supérieure for hospitality.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: SYMMETRIES AND THE REDUCED HAMILTONIAN

In this Appendix we derive the symmetries of the effective Hamiltonian in the presence of a random vector potential and derive the reduced Hamiltonian near the magic angle for both TBG and TTG.

A given Dirac point in graphene, i.e., a given valley, is split by the relative rotation θ into 2 Dirac points that are connected by the wave vectors [1]

$$\mathbf{q}_1 = k_\theta(0,\,-1),\;\; \mathbf{q}_2 = \tfrac{1}{2}k_\theta(\sqrt{3},\,1),\;\; \mathbf{q}_3 = \tfrac{1}{2}k_\theta(-\sqrt{3},\,1), \tag{A1}$$

where $k_\theta = \frac{4\pi}{3L_m}$. The effective Hamiltonian near these Dirac points is

$$\mathcal{H}^{(1)} = \begin{pmatrix} v_F\boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla}_{\mathbf{r}} - \tilde{\mathbf{A}}(\mathbf{r})] & e^{i(\theta/4)\sigma_z}\hat{T}(\mathbf{r})e^{i(\theta/4)\sigma_z} \\ e^{-i(\theta/4)\sigma_z}\hat{T}^\dagger(\mathbf{r})e^{-i(\theta/4)\sigma_z} & v_F\boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla}_{\mathbf{r}} + \tilde{\mathbf{A}}(\mathbf{r})] \end{pmatrix}$$

$$\hat{T}(\mathbf{r}) = \sum_{i=1}^{3} e^{i\mathbf{q}_i\cdot\mathbf{r}}T_i,$$

$$T_j = u_0\mathbb{1} + u_1\cos[(j-1)\phi]\sigma_x + u_1\sin[(j-1)\phi]\sigma_y,$$

$$j = 1, 2, 3, \tag{A2}$$

where $\phi = 2\pi/3$, v_F is the bare Dirac velocity of a single layer, $\sigma_{x,y,z}$ are Pauli matrices, and the interlayer couplings at either AA or AB overlaps are u_0 and u_1, respectively. This effective Hamiltonian with $\tilde{\mathbf{A}} = 0$ was given in Ref. [1] and was extended [15,17] to cases with a vector potential induced by lattice relaxation. In our case $\tilde{\mathbf{A}}(\mathbf{r})$ is random; it has opposite signs on the two layers since $\tilde{\mathbf{A}}(\mathbf{r}) \sim \pm\delta\theta$.

The Hamiltonian near the second valley is

$$\mathcal{H}^{(2)} = \begin{pmatrix} v_F\boldsymbol{\sigma}^* \cdot [i\boldsymbol{\nabla}_{\mathbf{r}} - \tilde{\mathbf{A}}(-\mathbf{r})] & \mathrm{e}^{-i(\theta/4)\sigma_z}\hat{T}^*(\mathbf{r})\mathrm{e}^{-i(\theta/4)\sigma_z} \\ \mathrm{e}^{i(\theta/4)\sigma_z}\hat{T}^{*\dagger}(\mathbf{r})\mathrm{e}^{i(\theta/4)\sigma_z} & v_F\boldsymbol{\sigma}^* \cdot [i\boldsymbol{\nabla}_{\mathbf{r}} + \tilde{\mathbf{A}}(-\mathbf{r})] \end{pmatrix}. \tag{A3}$$

This is obtained by rotating (A2) by C_{2z}, which involves the operator $R = \{\mathbf{r} \to -\mathbf{r}\}$ and interchange carbon sites $A \leftrightarrow B$, i.e., σ_x; hence, $\mathcal{H}^{(2)} = R\sigma_x\mathcal{H}^{(1)}\sigma_x R$ using $\sigma_x T(-\mathbf{r})\sigma_x = T^*(\mathbf{r})$.

We note that the factors $\mathrm{e}^{\pm i\theta\sigma_z/4}$ in Eqs. (A2) and (A3) are commonly neglected [15,17,19] since θ is small. However, these factors lead to non-Abelian disorder that can be relevant [18]. These factors modify the u_0 interlayer coupling to $u_0\mathrm{e}^{i\theta\sigma_z/2}$ so that Eqs. (A2) and (A3) contain to leading order to $\mathcal{H}' = \mp\frac{1}{2}u_0\theta\sigma_z\eta_y\sum_j \mathrm{e}^{i\mathbf{q}_j\cdot\mathbf{r}}$, where $\eta_{x,y,z}$ are Pauli matrices in the subspace of the two layers. Since $\theta(\mathbf{r}) = \overline{\theta} + \delta\theta(\mathbf{r})$ $\mathcal{H}'$ represents non-Abelian disorder. In comparison, the Abelian disorder $v_F\boldsymbol{\sigma}\cdot\tilde{\mathbf{A}}(\mathbf{r})\eta_z$ contains $\approx v_F\delta\theta\beta/a'$ [Eq. (3)]; hence, the ratio of coefficients is $\frac{u_0/2}{v_F\beta/a'} \approx 10^{-2}$, assuming $u_0 \approx u_1 \approx \frac{4\pi v_F\theta}{3a\sqrt{3}}$ near the magic angle $\theta = 1.1^\circ \approx 0.02$ rad. In the following we therefore neglect these $\mathrm{e}^{\pm i\theta\sigma_z/4}$ factors.

We consider next the symmetries of the Hamiltonian, extending previous studies [19,20] to our case with a random $\tilde{\mathbf{A}}(\mathbf{r})$. Note first that the Hamiltonian of both layers, with Pauli matrices $\boldsymbol{\tau}$ in the valley subspace, possess the symmetry operation $\sigma_x\tau_x R$:

$$\sigma_x\tau_x R\mathcal{H}\sigma_x\tau_x R = R\begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix}\begin{pmatrix} \mathcal{H}^{(1)} & 0 \\ 0 & \mathcal{H}^{(2)} \end{pmatrix}\begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix}R = \mathcal{H}. \tag{A4}$$

This is an extension of the single-valley symmetry [20] $\sigma_x\mathcal{K}R$ for the case $\tilde{\mathbf{A}}(\mathbf{r}) = 0$, where $\mathcal{K}$ represents complex conjugation.

Since we neglect the small θ in Eqs. (A2) and (A3) the system has a particle-hole (p-h) symmetry operator $\mathcal{C} = \eta_y\sigma_x\mathcal{K}$,

$$\begin{aligned}\mathcal{C}\mathcal{H}^{(1)}\mathcal{C} = \mathcal{K}&\begin{pmatrix} 0 & -i\sigma_x \\ i\sigma_x & 0 \end{pmatrix} \\ &\times\begin{pmatrix} \boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla} - \tilde{\mathbf{A}}(\mathbf{r})] & \hat{T}(\mathbf{r}) \\ \hat{T}^\dagger(\mathbf{r}) & \boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla} + \tilde{\mathbf{A}}(\mathbf{r})] \end{pmatrix} \\ &\times\begin{pmatrix} 0 & -i\sigma_x \\ i\sigma_x & 0 \end{pmatrix}\mathcal{K} = -\mathcal{H}^{(1)},\end{aligned} \tag{A5}$$

and similarly with $\mathcal{H}^{(2)}$. This is actually the same operator as in the $\tilde{A}(\mathbf{r}) = 0$ case [20]. Thus, for each eigenstate $|\psi\rangle$ with eigenvalue E there is an eigenstate $\mathcal{C}|\psi\rangle$ with eigenvalue $-E$. In view of the symmetry (A4) the p-h symmetry can also be represented by $\tau_x\eta_y R\mathcal{K}$, which mixes the two valleys. We note that a time-reversal operator $\tau_x\mathcal{K}$, which is a valid symmetry for $\tilde{\mathbf{A}}(\mathbf{r}) = 0$, is broken when $\tilde{\mathbf{A}}(\mathbf{r}) \neq 0$. The symmetries in the reduced Hamiltonian near the magic angle are also considered in Sec. II.

We proceed to find the reduced Hamiltonian near the magic angle. We follow the perturbation expansion [1,15,17] in u_0, u_1 connecting a state in layer I to 3 states in layer II. We then find an effective Hamiltonian near the Dirac point of layer I. The one-layer Hamiltonian $v_F\boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla} - \tilde{\mathbf{A}}(\mathbf{r})]$ with a state $\psi_0(\mathbf{r})$ is connected by T_j to the 3 states $\psi_j(\mathbf{r})\mathrm{e}^{-i\mathbf{q}_j\cdot\mathbf{r}}$, $j = 1, 2, 3$. Ignoring the small θ rotation (exact for $u_0 = 0$), the 4×4 Hamiltonian equation (A2) becomes an 8×8 one,

$$\tilde{\mathcal{H}} = v_F\begin{pmatrix} \boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla} - \tilde{\mathbf{A}}(\mathbf{r})] & T_1/v_F & T_2/v_F & T_3/v_F \\ T_1/v_F & \boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla} + \tilde{\mathbf{A}}(\mathbf{r}) - \mathbf{q}_1] & 0 & 0 \\ T_2/v_F & 0 & \boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla} + \tilde{\mathbf{A}}(\mathbf{r}) - \mathbf{q}_2] & 0 \\ T_3/v_F & 0 & 0 & \boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla} + \tilde{\mathbf{A}}(\mathbf{r}) - \mathbf{q}_3] \end{pmatrix}. \tag{A6}$$

Note that $\tilde{A}(\mathbf{r})$ is slowly varying on the scale L_m so that it is kept on the diagonal. This is in contrast to the vector potentials from lattice relaxation [15,17] that carry momenta $\mathbf{q}_i - \mathbf{q}_j$ $(i \neq j)$ and therefore appear in this matrix in off-diagonal positions; the latter case results in a shift of the magic angle [15,17].

We next repeat the known limit $i\boldsymbol{\nabla} = \tilde{\mathbf{A}}(\mathbf{r}) = 0$ showing that it has an eigenstate $\Psi_0 = (\psi_0, \psi_1, \psi_2, \psi_3)$ with zero eigenvalue. Defining matrices h_j, $j = 1, 2, 3$,

$$\begin{aligned} h_j &= -v_F\boldsymbol{\sigma}\cdot\mathbf{q}_j = iv_Fk_\theta\begin{pmatrix} 0 & -\mathrm{e}^{-i(j-1)\phi} \\ \mathrm{e}^{i(j-1)\phi} & 0 \end{pmatrix} \\ h_j^{-1} &= h_j/(v_Fk_\theta)^2,\ T_jh_j^{-1}T_j = \left(u_0^2 - u_1^2\right)h_j/(v_Fk_\theta)^2. \end{aligned} \tag{A7}$$

Note that $\sum_j T_jh_j^{-1}T_j = 0$ for any u_0, u_1. The eigenvalue equation of (A6) yields

$$\begin{aligned} &\text{lines 2,3,4} \quad T_j\psi_0 + h_j\psi_j = 0 \ \Rightarrow \psi_j = -h_j^{-1}T_j\psi_0 \\ &\text{line 1} \ \sum_j T_j\psi_j = \left(\sum_j T_jh_i^{-1}T_j\right)\psi_0 = 0, \end{aligned} \tag{A8}$$

which identifies ψ_j and proves that it has vanishing eigenvalue. Assuming $\langle\psi_0|\psi_0\rangle = 1$, we find the normalization $\langle\Psi_0|\Psi_0\rangle = 1 + 3(\bar{u}_0^2 + \bar{u}_1^2)$ using $(h_j^{-1})^2 = \mathbb{1}/(v_Fk_\theta)^2$ and defining $\bar{u}_i = u_i/v_Fk_\theta$, $i = 1, 2$.

Consider next the effective Hamiltonian projected to the ψ_0 space, to first order in $-i\boldsymbol{\nabla}$, $\tilde{\mathbf{A}}(\mathbf{r})$, defined by

$$\begin{aligned}\langle\psi_0|\mathcal{H}_{\mathrm{eff}}|\psi_0\rangle \sim \langle\Psi_0 + O(k, \tilde{\mathbf{A}})|v_F\boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla}\mathbb{1} - \tilde{\mathbf{A}}(\mathbf{r})D]|\Psi_0 \\ + O(k, \tilde{\mathbf{A}})\rangle,\end{aligned} \tag{A9}$$

where D is a diagonal matrix with elements $(1,-1,-1,-1)$. Including the normalization, we obtain

$$\begin{aligned}\langle\psi_0|\mathcal{H}_{\rm eff}|\psi_0\rangle &= \frac{1}{1+3(\bar{u}_0^2+\bar{u}_1^2)}\langle\Psi_0|v_F\boldsymbol{\sigma}\cdot[\mathbf{k}\mathbb{1}-\tilde{\mathbf{A}}(\mathbf{r})D]|\Psi_0\rangle\\ &= \frac{v_F}{1+3(\bar{u}_0^2+\bar{u}_1^2)}\langle\psi_0|\boldsymbol{\sigma}\cdot[\mathbf{k}-\tilde{\mathbf{A}}(\mathbf{r})]+\sum_j T_j\big(h_j^{-1}\big)^\dagger\boldsymbol{\sigma}\cdot[\mathbf{k}+\tilde{\mathbf{A}}(\mathbf{r})]h_j^{-1}T_j|\psi_0\rangle\\ &= \frac{v_F}{1+3(\bar{u}_0^2+\bar{u}_1^2)}\langle\psi_0|\boldsymbol{\sigma}\cdot[\mathbf{k}-\tilde{\mathbf{A}}(\mathbf{r})]-3\bar{u}_1^2\boldsymbol{\sigma}\cdot[\mathbf{k}+\tilde{\mathbf{A}}(\mathbf{r})]|\psi_0\rangle. \end{aligned} \tag{A10}$$

This then identifies the reduced Hamiltonian, Eq. (5).

We note that in addition to $\tilde{\mathbf{A}}(\mathbf{r})\approx\frac{\beta}{a'}\delta\theta$ [Eq. (3)] the Dirac Hamiltonian equation (5) contains a $\delta\theta$ dependence from $v(\theta+\delta\theta)=v(\theta)+v_F\frac{\delta\theta}{\theta}$, hence a Hamiltonian term $\approx -iv_F\frac{\delta\theta(\mathbf{r})}{\theta}\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}$, possibly relating to interference between different moiré periodicities [33]. However, the ratio of this term to that of $\tilde{\mathbf{A}}(\mathbf{r})$ is small, $\frac{a'}{\ell_0\theta\beta}\ll 1$, and is therefore neglected, taking $|\boldsymbol{\nabla}|\sim 1/\ell_0$.

We next interchange the two layers with an operator η_x, i.e.,

$$\eta_x\mathcal{H}^{(1)}\eta_x = \begin{pmatrix}\boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla}_{\mathbf{r}}+\tilde{\mathbf{A}}(\mathbf{r})] & e^{-i(\theta/4)\sigma_z}\hat{T}^\dagger(\mathbf{r})e^{-i(\theta/4)\sigma_z}\\ e^{i(\theta/4)\sigma_z}\hat{T}(\mathbf{r})e^{i(\theta/4)\sigma_z} & \boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla}_{\mathbf{r}}-\tilde{\mathbf{A}}(\mathbf{r})]\end{pmatrix}, \tag{A11}$$

and build the perturbation expansion with a state on layer II that is coupled to 3 states in layer I. This corresponds to expanding near the Dirac point of layer II representing a distinct Moiré valley. Neglecting θ, the only change in the perturbative solution is $\mathbf{q}_j\to-\mathbf{q}_j$, which does not change the result. The effective Hamiltonian is then the same as in (A10), i.e., Eq. (5), except significantly that $\tilde{\mathbf{A}}(\mathbf{r})\to-\tilde{\mathbf{A}}(\mathbf{r})$. Hence, we represent η_x explicitly by $\tilde{\mathbf{A}}(\mathbf{r})\to-\tilde{\mathbf{A}}(\mathbf{r})$.

Considering next the other graphene valley with $\mathcal{H}^{(2)}$, the perturbation proceeds with $h_j\to-h_j^*=-h_j(-\phi)$, so the effective Hamiltonian is (A10) except $\tilde{\mathbf{A}}(\mathbf{r})\to\tilde{\mathbf{A}}(-\mathbf{r})$ and $\mathcal{H}_{\rm eff}\to\mathcal{H}_{\rm eff}^*$.

The system is thus decoupled into three subspaces labeled with s, v^m, v, where $s=\pm$ for spin, $v^m=\pm$ for moiré valley, and $v=\pm$ for graphene valley. Each of the resulting 8 bands has a single Dirac cone (positive and negative energies) with 2 spinor components defined (in real space) on AA and AB sites, respectively.

Trilayer twisted graphene. We finally consider mirror-symmetric TTG, in particular the ABA configuration, as in the relevant experiments [6,7]. We also note that the ABA configuration is energetically favored, having a stronger interlayer coupling [34].

We derive the reduced Hamiltonian, following the previous TBG derivation and its TTG extension [35]. We allow initially for arbitrary variations in the local twist angle of each layer, leading to 3 random vector potentials $\mathbf{A}_j(\mathbf{r})$, $j=1,2,3$. The previous interlayer matrices T_j for *AA* overlap at the middle of the moiré unit cell need to be replaced by $T_j\to e^{-i\mathbf{q}_j\cdot\mathbf{r}_0}T_j$ for *AB* overlap, where $\mathbf{r}_0=\frac{4\pi}{3\sqrt{3}k_\theta}(-1,0)$; eventually, this phase factor has no effect. The Hamiltonian (A6) becomes, for the ABA configuration,

$$\tilde{\mathcal{H}}=\begin{pmatrix} v_F\boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla}-\mathbf{A}_1(\mathbf{r})] & T_1 & T_2 & T_3 & 0\\ T_1^\dagger & v_F\boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla}-\mathbf{A}_2(\mathbf{r})-\mathbf{q}_1] & 0 & 0 & T_1^\dagger\\ T_2^\dagger & 0 & v_F\boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla}-\mathbf{A}_2(\mathbf{r})-\mathbf{q}_2] & 0 & T_2^\dagger\\ T_3^\dagger & 0 & 0 & v_F\boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla}-\mathbf{A}_2(\mathbf{r})-\mathbf{q}_3] & T_3^\dagger\\ 0 & T_1 & T_2 & T_3 & v_F\boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla}-\mathbf{A}_3(\mathbf{r})]\end{pmatrix}. \tag{A12}$$

At $i\boldsymbol{\nabla}=\mathbf{A}_j(\mathbf{r})=0$ the eigenstates $\Psi_0=(\alpha,\beta_1,\beta_2,\beta_3,\gamma)$ for vanishing eigenvalue satisfy

$$\begin{aligned}&\text{lines 2,3,4}\quad T_j^\dagger\alpha+h_j\beta_j+T_j^\dagger\gamma=0\Rightarrow\beta_j=-h_j^{-1}T_j^\dagger(\alpha+\gamma)\\ &\text{lines 1,5}\quad\sum_j T_j\beta_j=0=-\sum_j T_jh_j^{-1}T_j^\dagger(\alpha+\gamma),\end{aligned} \tag{A13}$$

and the mirror symmetry allows $\alpha=\pm\gamma$. The normalization for the case $\alpha=\gamma$ is

$$\langle\Psi_0|\Psi_0\rangle=2+\langle\alpha+\gamma|\sum_j T_j\big(h_j^{-1}\big)^\dagger h_j^{-1}T_j^\dagger|\alpha+\gamma\rangle=2+12\big(\bar{u}_0^2+\bar{u}_1^2\big)\equiv C.$$

The effective Hamiltonian, projected on $|\alpha\rangle$, to first order in $i\boldsymbol{\nabla}$, $\mathbf{A}_j(\mathbf{r})$ is

$$\begin{aligned}\langle\alpha|\mathcal{H}_{\rm eff}|\alpha\rangle &= \langle\Psi_0|v_F\boldsymbol{\sigma}\cdot\begin{pmatrix}[-i\boldsymbol{\nabla}-\mathbf{A}_1(\mathbf{r}) & 0 & 0 & 0 & 0\\ 0 & -i\boldsymbol{\nabla}-\mathbf{A}_2(\mathbf{r}) & 0 & 0 & 0\\ 0 & 0 & -i\boldsymbol{\nabla}-\mathbf{A}_2(\mathbf{r}) & 0 & 0\\ 0 & 0 & 0 & -i\boldsymbol{\nabla}-\mathbf{A}_2(\mathbf{r}) & 0\\ 0 & 0 & 0 & 0 & -i\boldsymbol{\nabla}-\mathbf{A}_3(\mathbf{r})\end{pmatrix}|\Psi_0\rangle\\ &= \frac{v_F}{C}\langle\alpha|\boldsymbol{\sigma}\cdot[-2i\boldsymbol{\nabla}-\mathbf{A}_1(\mathbf{r})-\mathbf{A}_3(\mathbf{r})]+\sum_j\langle\alpha+\gamma|T_j\big(h_j^{-1}\big)^\dagger\boldsymbol{\sigma}[-i\boldsymbol{\nabla}-\mathbf{A}_2(\mathbf{r})]h_j^{-1}T_j^\dagger|\alpha+\gamma\rangle\\ &= \frac{v_F}{C}\langle\alpha|\boldsymbol{\sigma}\cdot[-2i\boldsymbol{\nabla}-\mathbf{A}_1(\mathbf{r})-\mathbf{A}_3(\mathbf{r})]-12\bar{u}_1^2\boldsymbol{\sigma}[-i\boldsymbol{\nabla}-\mathbf{A}_2(\mathbf{r})]|\alpha\rangle\\ \Rightarrow\ \mathcal{H}_{\rm eff} &= \frac{v_F}{C}\boldsymbol{\sigma}\cdot\big[-2i\boldsymbol{\nabla}\big(1-6\bar{u}_1^2\big)-\big(\mathbf{A}_1(\mathbf{r})+\mathbf{A}_3(\mathbf{r})-12\bar{u}_1^2\mathbf{A}_2(\mathbf{r})\big)\big].\end{aligned}\tag{A14}$$

The Dirac velocity is then

$$v = v_F\frac{1-6\bar{u}_1^2}{1+6\big(\bar{u}_0^2+\bar{u}_1^2\big)}.\tag{A15}$$

The critical value where v vanishes, relative to that of TBG, is $\bar{u}_1^c(\mathrm{TTG}) = \bar{u}_1^c(\mathrm{TBG})/\sqrt{2}$; since $\bar{u}_1 \sim 1/\theta$ the magic angle satisfies $\theta_m(\mathrm{TTG}) = \sqrt{2}\theta_m(\mathrm{TBG})$. The disorder terms can be written as

$$\frac{v}{4}\Big\{\frac{1+6\bar{u}_1^2}{1-6\bar{u}_1^2}[\mathbf{A}_1(\mathbf{r})+\mathbf{A}_3(\mathbf{r})-2\mathbf{A}_2(\mathbf{r})] + \mathbf{A}_1(\mathbf{r})+\mathbf{A}_3(\mathbf{r})+2\mathbf{A}_2(\mathbf{r})\Big\}.$$

The first term diverges at the critical point, hence dominant, while the second term is finite and is therefore neglected. The effective disorder vector potential is then

$$\begin{aligned}\mathbf{A}(\mathbf{r}) &= \frac{1+6\bar{u}_1^2}{4\big(1-6\bar{u}_1^2\big)}[\mathbf{A}_1(\mathbf{r})+\mathbf{A}_3(\mathbf{r})-2\mathbf{A}_2(\mathbf{r})]\\ \mathcal{H}_{\rm eff} &= v\boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla}-\mathbf{A}(\mathbf{r})],\end{aligned}\tag{A16}$$

exactly the same form as Eq. (A5). In terms of the local twist angles $\delta\theta_i(\mathbf{r})$ we have $\mathbf{A}(\mathbf{r}) \sim \delta\theta_1(\mathbf{r}) + \delta\theta_3(\mathbf{r}) - 2\delta\theta_3(\mathbf{r}) = \delta\theta_{12}(\mathbf{r}) + \delta\theta_{32}(\mathbf{r})$. We note that for the solution $\alpha = -\gamma$ Eq. (A13) implies $\beta_j = 0$; hence, the interlayer coupling has no effect and the Dirac Hamiltonian is just $v_F\boldsymbol{\sigma}\cdot[-i\boldsymbol{\nabla} - \mathbf{A}_i(\mathbf{r})]$, $i = 1, 3$, i.e., the velocity v_F is maintained and the disorder has a minor effect.

APPENDIX B: DENSITY OVERLAPS FOR $E = 0$ STATES

We wish to evaluate the density overlaps between the various $E = 0$ eigenstates listed in Table I. By density overlap we mean the disorder average $\overline{|\varphi_i(\mathbf{r})|^2|\varphi_j(\mathbf{r})|^2}$ for various 2-spinor eigenstates of the Dirac equation with a random gauge potential.

We start with replica symmetric method that applies to weak disorder [11]. Consider, where eventually $n \to 0$, and using $\langle V^2(\mathbf{r})\rangle = \frac{1}{2}\sigma\ln\frac{L}{\ell_0}$, $\langle V(\mathbf{r})V(0)\rangle = \frac{1}{2}\sigma\ln\frac{L}{r}$,

$$\begin{aligned}P_{q,q'} &= \overline{\frac{\int \mathrm{e}^{2q'V(\mathbf{r})}d^2r}{[\int \mathrm{e}^{2V(\mathbf{r})}d^2r]^q}} = \overline{\int \mathrm{e}^{2q'V(\mathbf{r})}d^2r\,\Pi_{j=1}^{n-q}\int d^2r_j\mathrm{e}^{2V(\mathbf{r}_j)}} = \int d^2r d^2r_1\ldots d^2r_{n-q}\mathrm{e}^{2\overline{[q'V(\mathbf{r})+\sum_{j=1}^{n-q}V(\mathbf{r}_j)]^2}}\\ &= \mathrm{e}^{\sigma(q'^2+n-q)\ln\frac{L}{\ell_0}}\int d^2r d^2r_1\ldots d^2r_{n-q}\mathrm{e}^{2\sigma q'\sum_{j=1}^{n-q}\ln\frac{L}{|\mathbf{r}-\mathbf{r}_j|}+2\sigma\sum_{i<j}^{n-q}\ln\frac{L}{|\mathbf{r}_i-\mathbf{r}_j|}} \approx \left(\frac{L}{\ell_0}\right)^{\sigma(q'^2+n-q)}L^{2+2(n-q)} \overset{n\to0}{=} \left(\frac{L}{\ell_0}\right)^{\sigma(q'^2-q)}L^{2-2q}.\end{aligned}\tag{B1}$$

Since $n \to 0$ at the end, we always omit factors $Z(1)^n$ in denominators, where $Z(1) = \int \mathrm{e}^{2V(\mathbf{r})}d^2r$. The last step assumes that all differences $|\mathbf{r}-\mathbf{r}_i|$, $|\mathbf{r}_i-\mathbf{r}_j| \sim L$ so the the dimension of the integral is $2+2(n-q)$, implying that σ is not too large. For the needed intensity overlaps, $P_{2,2} \sim L^{2\sigma-2}$ and $P_{2,0} \sim L^{-2-2\sigma}$.

To treat large σ we follow the packet partition of replicas leading to replica symmetry breaking (RSB), as discussed in Ref. [24] Appendix D and in Ref. [26] Eqs. 77–81. For strong σ we expect strong attraction between replica variables; thus, we allow the option of forming packets where p replica variables $\mathbf{r}_i$ are close to each other. Eventually p is determined by a minimum condition. Consider first the packet division of P_{q0}. In (B1) divide the $n-q$ replicas into $(n-q)/p$ packets having p replicas each. Each packet has $\mathbf{r}_{ki}$, $k = 1, \ldots, p$ close to each other while the interpacket distance is $\sim L$. Thus the number of

integrations $\sim L^2$ is $1+(n-q)/p$ (including 1 from the nominator of P_{q0}),

$$P_{q,0} = \mathrm{e}^{2\sigma(n-q)\ln\frac{L}{\ell_0}} L^{2+2\frac{n-q}{p}} \left\{ \int \Pi_{i=1}^{(n-q)/p} \frac{d^2r_{1i}...d^2r_{pi}}{L^2} \mathrm{e}^{2\sigma\sum_{k,k'=1}^{p}\sum_{i<j}^{(n-q)/p}\ln\frac{L}{|\mathbf{r}_{ki}-\mathbf{r}_{k'j}|}} \right\} \mathrm{e}^{2\sigma\frac{1}{2}p(p-1)\frac{n-q}{p}\ln\frac{L}{\ell_0}}$$
$$\approx L^{2+2\frac{n-q}{p}} \left(\frac{L}{\ell_0}\right)^{\sigma(n-q)+\sigma(p-1)(n-q)} \ell_0^{2(p-1)\frac{n-q}{p}} \overset{n\to 0}{=} \ell_0^{2-2q} \left(\frac{L}{\ell_0}\right)^{2-2q(\frac{1}{p}+\frac{1}{2}\sigma p)}, \tag{B2}$$

where the $\{\}$ integral is by assumption convergent for each packet so it is $\approx \ell_0^{2(p-1)}$. Minimizing $\frac{1}{p}+\frac{1}{2}\sigma p$ leads to $p = \min[\frac{1}{\sqrt{\sigma/2}}, 1]$, hence a critical $\sigma_c = 2$. For $\sigma < 2$ there is no RSB ($p = 1$) and the result coincides with (B1). For $\sigma > 2$ RSB is present ($p < 1$), hence

$$P_{q,0} \approx \ell_0^{2-2q} \left(\frac{L}{\ell_0}\right)^{2-4q\sqrt{\sigma/2}} \qquad \sigma > 2. \tag{B3}$$

In the following we actually need a slightly different average,

$$P'_{2,0} = \overline{\frac{\int d^2r}{\int d^2r\, \mathrm{e}^{2V(\mathbf{r})} \int d^2r\, \mathrm{e}^{-2V(\mathbf{r})}}} = \int d^2r \prod_{i=1}^{n-1} d^2r_i d^2r'_i \overline{\mathrm{e}^{2\sum_{i=1}^{n-1}[V(\mathbf{r}_i)-V(\mathbf{r}'_i)]}}. \tag{B4}$$

Since the groups $\mathbf{r}_i$ and $\mathbf{r}'_i$ come with opposite signs of V they interact with $-\sigma < 0$, which does not favor packet formation. Hence, we treat the two groups separately. Divide each group of $n-1$ replicas into $i = 1, 2, ..., (n-1)/p$ packets having p replicas each. Each packet has either $\mathbf{r}_{ki}$, $k = 1, ..., p$ or $\mathbf{r}'_{ki}$, $k = 1, ..., p$ close to each other while the interpacket distance is $\sim L$. Thus, the number of integrations $\sim L^2$ is $1+2(n-1)/p$:

$$P'_{20} = \mathrm{e}^{2\sigma(n-1)\ln\frac{L}{\ell_0}} L^{2+4(n-1)/p} \left\{ \int \Pi_{i=1}^{(n-1)/p} \frac{d^2r_{1i}...d^2r_{pi}}{L^2} \int \Pi_{i=1}^{(n-1)/p} \frac{d^2r'_{1i}...d^2r'_{pi}}{L^2} \right.$$
$$\left. \times\, \mathrm{e}^{[2\sigma\sum_{k,k'=1}^{p}\sum_{i<j}^{(n-1)/p}\ln\frac{L}{|\mathbf{r}_{ki}-\mathbf{r}_{k'j}|}]+[\text{all } \mathbf{r}_{ki}-\mathbf{r}_{k'j}\to\mathbf{r}'_{ki}-\mathbf{r}'_{k'j}]-2[\text{all } \mathbf{r}_{ki}-\mathbf{r}_{k'j}\to\mathbf{r}_{ki}-\mathbf{r}'_{k'j}]} \right\} \mathrm{e}^{2\sigma\ln\frac{L}{\ell_0}\frac{1}{2}p(p-1)2(n-1)/p}$$
$$\approx L^{2+4(n-1)/p} \left(\frac{L}{\ell_0}\right)^{2\sigma(n-1)+2\sigma(p-1)(n-1)} \ell_0^{2(p-1)2(n-1)/p} \overset{n\to 0}{=} \ell_0^{-2} \left(\frac{L}{\ell_0}\right)^{2-4(\frac{1}{p}+\frac{1}{2}\sigma p)}, \tag{B5}$$

where $\{...\}$ is L independent. The result is identical to $P_{2,0}$.

For $q' \neq 0$ the $\mathbf{r}$, $\mathbf{r}_i$ integrations are not symmetric; therefore, allow for one packet of size p_0 that interacts with $q'V(\mathbf{r})$, i.e., its variables are $\mathbf{r}$, $\mathbf{r}_{k0}$, $k = 1, \ldots, p_0$. The other variables are divided into packets of size p, i.e., $i = 1, ..., \frac{n-q-p_0}{p}$ packets. Each packet has a c.m. that spans the area L^2, while the spacing between packet members is $\approx \ell_0$. Each intrapacket interaction is $\sim \ln\frac{L}{\ell_0}$, while all interpacket terms are $\sim \ln\frac{L}{|\mathbf{r}_{ki}-\mathbf{r}_{k'j}|}$, $i \neq j$. Hence,

$$P_{q,q'} = \mathrm{e}^{\sigma(q'^2+n-q)\ln\frac{L}{\ell_0}} L^{2+2\frac{n-q-p_0}{p}} \left\{ \int \frac{d^2r}{L^2} \prod_{k=1}^{p_0} \int d^2r_{k0} \prod_{i=1}^{n-q-p_0} \frac{1}{L^2} \prod_{k=1}^{p} \int d^2r_{ki} \right.$$
$$\left. \times\, \mathrm{e}^{2\sigma q'\sum_{k=1}^{p}\sum_{j=1}\ln\frac{L}{\mathbf{r}-\mathbf{r}_{kj}|}+2\sigma\sum_{k=1}^{p_0}\sum_{k'=1}^{p}\sum_j\ln\frac{L}{|\mathbf{r}_{k0}-\mathbf{r}_{k'j}|}+2\sigma\sum_{k,k'=1}^{p}\sum_{i<j}\ln\frac{L}{|\mathbf{r}_{ki}-\mathbf{r}_{k'j}|}} \right\}$$
$$\times\, \mathrm{e}^{2\sigma\ln\frac{L}{\ell_0}[q'p_0+\frac{1}{2}p_0(p_0-1)+\frac{1}{2}p(p-1)\frac{n-q-p_0}{p}]}. \tag{B6}$$

The $\{...\}$ factor is again L independent; hence,

$$\frac{\ln P_{q,q'}}{\ln L} = \sigma(q'^2+n-q) + 2\left(1+\frac{n-q-p_0}{p}\right) + 2\sigma q'p_0 + \sigma p_0(p_0-1) + \sigma p(p-1)\frac{n-q-p_0}{p}$$
$$\overset{n\to 0}{=} \sigma q'^2 + 2 - 2\left(\frac{1}{p}+\tfrac{1}{2}\sigma p\right)(q+p_0) + 2\sigma q'p_0 + \sigma p_0^2. \tag{B7}$$

Optimizing with respect to p leads as before to $p = \min(1, 1/\sqrt{\sigma/2})$, hence a critical $\sigma_c = 2$. Optimizing with respect to p_0 yields three regimes:

$$p_0 + q' = \begin{cases} q', & \sigma < \frac{2}{2q'-1} \\ \frac{1}{2}\left(1+\frac{2}{\sigma}\right), & \frac{2}{2q'-1} < \sigma < 2 \\ 1/\sqrt{\sigma/2}, & \sigma > 2. \end{cases} \tag{B8}$$

In the first two regimes we assume $q' \geqslant 1$. In the first, small $\sigma < \frac{2}{2q'-1}$ regime one has $p = 1,\ p_0 = 0$, which leads to $P_{q,q'} \sim L^{\sigma(q'^2-q)+2-2q}$ in agreement with (1), while in the third, $\sigma > 2$ regime one finds

$$P_{q,q'} \approx \left(\frac{L}{\ell_0}\right)^{-4\sqrt{\sigma/2}(q-q')} \ell_0^{2-2q}, \quad \sigma > 2,\ q' > 0. \tag{B9}$$

Note that at $q' = 0$ this result differs from the previous $P_{q,0}$ in (2) by a factor L^2; in this case the variable p_0 is redundant since there is no $q'V(\mathbf{r})$ to induce it.

Finally, we find an intermediate regime for $\frac{2}{2q'-1} < \sigma < 1$. This regime was obtained previously in the case $q' = q$ by a completely different replica method, and called "prefreezing" [30] (in that case, it represents the disorder-averaged inverse participation ratio). Here we find in that regime

$$P_{q,q'} \approx \left(\frac{L}{\ell_0}\right)^{-(2+\sigma)(q-q')-\frac{(2-\sigma)^2}{4\sigma}} \ell_0^{2-2q}, \quad \frac{2}{2q'-1} < \sigma < 2,\ q' > 1. \tag{B10}$$

Although it is gratifying to obtain here the prefreezing regime by a distinct replica method, we will not need it for our applications, which focus on large σ.

Remark. It is important to note that the $P_{q,q'}$ computed here are *disorder averages* of the ratios $\overline{Z(q')/Z(1)^q}$, where we denote $Z(q) = \int \mathrm{e}^{2qV(\mathbf{r})} d^2r$, also called annealed quantities. One can also compute the *typical values* of these ratios, as was done for $q' = q$ in Ref. [26] Eq. 138, as well as in Ref. [29] Eq. 4.5 within Derrida's random energy model. These studies have shown that at large L

$$\frac{\log Z(q)}{\log L} \to \begin{cases} 2 + \sigma q^2, & \sigma < 2/q^2 \\ 4q\sqrt{\sigma/2}, & \sigma > 2/q^2. \end{cases} \tag{B11}$$

Since $\log(Z(q')/Z(1)^q) = \log Z(q') - q \log Z(1)$ it leads to three regimes:

$$\frac{\log(Z(q')/Z(1)^q)}{\log L} \to \begin{cases} \sigma((q')^2 - q) + 2 - 2q, & \sigma < 2/(q')^2 \\ 4q'\sqrt{\sigma/2} - q(2+\sigma), & 2/(q')^2 < \sigma < 2 \\ 4\sqrt{\sigma/2}(q' - q), & \sigma > 2. \end{cases} \tag{B12}$$

Comparing with the above we see that the typical and average differ only in the intermediate disorder regime, i.e., the nature of the prefreezing is slightly different in both cases.

Consider next the density overlap,

$$Q_{1,-1} = \overline{\frac{\int d^2r\, \mathrm{e}^{2V(\mathbf{r})-2V(-\mathbf{r})}}{\int d^2r\, \mathrm{e}^{2V(\mathbf{r})} \int d^2r\, \mathrm{e}^{-2V(-\mathbf{r})}}} = \int d^2r \prod_{i=1}^{n-1} d^2r_i d^2r'_i\, \overline{\mathrm{e}^{2V(\mathbf{r})-2V(-\mathbf{r})+2\sum_{i=1}^{n-1}[V(\mathbf{r}_i)-V(-\mathbf{r}'_i)]}}. \tag{B13}$$

Since the groups $\mathbf{r}, \mathbf{r}_i$ and $-\mathbf{r}, -\mathbf{r}'_i$ come with opposite signs of V they interact with $-\sigma < 0$, which does not favor packet formation. Hence, we treat the two groups separately. Consider $\mathbf{r}_{k1},\ k = 1, ..., p_1$ as one packet with $p_1 - 1$ replicas, which includes $\mathbf{r} = \mathbf{r}_{11}$ for convenience, and then packets $\mathbf{r}_{ki},\ k = 1, ..., p$ each with p replicas and $i = 2, 3, ..., n' = \frac{n-p_1}{p} + 1$, which altogether represent $n - 1$ replicas. Similarly, consider all $\mathbf{r}_{ki} \to -\mathbf{r}'_{ki}$ with $\mathbf{r}'_{11} = -\mathbf{r}$. Note that the pair $\mathbf{r}, -\mathbf{r}$ cannot be in the same packet since then $r \lesssim \ell_0$ and the c.m. of the packet cannot span the whole space; hence, the packets $\mathbf{r}_{k1},\ \mathbf{r}'_{k1}$ are distinct. Thus,

$$\begin{aligned} Q_{1,-1} = \mathrm{e}^{\sigma \ln \frac{L}{\ell_0}(2+2n-2)} L^{2+2(n'-1)2} \Bigg\{ \int \frac{d^2r}{L^2} \prod_{k=2}^{p_1} \int d^2r_{k1} \prod_{i=2}^{n'} \frac{1}{L^2} \prod_{k=1}^{p} d^2r_{ki} \prod_{k=2}^{p_1} \int d^2r'_{k1} \prod_{i=2}^{n'} \frac{1}{L^2} \prod_{k=1}^{p} d^2r'_{ki} \\ \times \mathrm{e}^{[2\sigma \sum_{k=1}^{p_1} \sum_{k'=1}^{p} \sum_{i=2}^{n'} \ln \frac{L}{|\mathbf{r}_{k1} - \mathbf{r}_{k'i}|} + 2\sigma \sum_{k,k'=1}^{p} \sum_{i<j}^{n'} \ln \frac{L}{|\mathbf{r}_{ki} - \mathbf{r}_{k'j}|}] + [\text{all } \mathbf{r}_{ki} - \mathbf{r}_{k'j} \to \mathbf{r}'_{ki} - \mathbf{r}'_{k'j}] - [\text{all } \mathbf{r}_{ki} - \mathbf{r}_{k'j} \to \mathbf{r}_{ki} + \mathbf{r}'_{k'j}]} \Bigg\} \\ \times \mathrm{e}^{2\sigma \ln \frac{L}{\ell_0} [\frac{1}{2} p_1(p_1-1) + \frac{1}{2} p(p-1)(n'-1)]2}, \end{aligned} \tag{B14}$$

where $\{...\}$ is L independent. Hence,

$$\frac{\ln Q_{1,-1}}{\ln L} = 2 + 4\frac{n-p_1}{p} + 2\sigma n + 2\sigma \left[p_1(p_1-1) + p(p-1)\frac{n-p_1}{p} \right] \overset{n\to 0}{=} 2 - 4\frac{p_1}{p} + 2\sigma p_1(p_1 - p).$$

Let us start with the case $\sigma > 2$. Minimizing over p yields $p = 1/\sqrt{\sigma/2}$ and over p_1 is the same, $p_1 = 1/\sqrt{\sigma/2}$. While there is RSB in the packet size p there is no change in $Q_{1,-1}$, reflecting a σ-independent exponent,

$$Q_{1,-1} \approx \left(\frac{L}{\ell_0}\right)^{-2} \ell_0^{-2} \quad \sigma > 0. \tag{B15}$$

Indeed, optimizing now for $\sigma < 2$ one finds $p = 1$ and naively $p_1 = (\frac{1}{2} + 1/\sigma)$. However, the optimum cannot be reached since one must have $p_1 \leqslant 1$; hence, one must set $p = p_1 = 1$ leading to the same -2 exponent.

Remark. One can check that the typical value of the ratio decays even faster than the average, i.e., it can be argued to equal to $\log Z(\sqrt{2}) - 2\log Z(1)$, which leads to typical exponent -2 for $\sigma < 1$, $4(\sqrt{\sigma} - 1 - \sigma/2)$ for $1 < \sigma < 2$, and $4\sqrt{\sigma/2}(\sqrt{2} - 2)$ for $\sigma > 2$.

Consider next

$$Q_{1,1} = \overline{\frac{\int d^2r\, e^{2V(\mathbf{r})+2V(-\mathbf{r})}}{\int d^2r\, e^{2V(\mathbf{r})} \int d^2r\, e^{2V(-\mathbf{r})}}} = \int d^2r \prod_{i=1}^{2n-2} d^2r_i\, \overline{e^{2V(\mathbf{r})+2V(-\mathbf{r})+2\sum_{i=1}^{2n-2} V(\mathbf{r}_i)}}, \tag{B16}$$

where in the denominator change variable so that $V(-\mathbf{r}) \to V(\mathbf{r})$. Again, the pair $\mathbf{r}, -\mathbf{r}$ cannot be in the same packet since then $r \lesssim \ell_0$ and the c.m. of the packet cannot span the whole space. Consider then one packet $\mathbf{r}_{k1}$, $k = 1, ..., p_1$ with $p_1 - 1$ replicas and $\mathbf{r}_{11} = \mathbf{r}$; a second packet $\mathbf{r}_{k2}$, $k = 1, ..., p_2$ with $p_2 - 1$ replicas and $\mathbf{r}_{12} = -\mathbf{r}$; and then packets $\mathbf{r}_{kj}$, $k = 1, ..., p$ with $j = 3, 4, ..., n' = \frac{2n-p_1-p_2}{p} + 2$, so that the total number of replicas is $2n - 2$.

$$\begin{aligned} Q_{1,1} = e^{\sigma \ln \frac{L}{\ell_0}(2+2n-2)} L^{2+2(n'-2)} \Bigg\{ \int \frac{d^2r}{L^2} \prod_{k=2}^{p_1} \int d^2r_{k1} \prod_{k=2}^{p_2} \int d^2r_{k2} \prod_{j=3}^{n'} \frac{1}{L^2} \prod_{k=1}^{p} \int d^2r_{kj} \\ \times e^{2\sigma[\sum_{k=1}^{p_1}\sum_{k'=1}^{p_2} \ln \frac{L}{|\mathbf{r}_{1k}-\mathbf{r}_{2k'}|} + \sum_{k=1}^{p_1}\sum_{k'=1}^{p}\sum_{j=3}^{n'} \ln \frac{L}{|\mathbf{r}_{1k}-\mathbf{r}_{jk'}|} + \sum_{k=1}^{p_2}\sum_{k'=1}^{p}\sum_{j=3}^{n'} \ln \frac{L}{|\mathbf{r}_{2k}-\mathbf{r}_{jk'}|} + \sum_{k,k'=1}^{p}\sum_{3\leqslant i<j}^{n'} \ln \frac{L}{|\mathbf{r}_{ik}-\mathbf{r}_{jk'}|}]} \Bigg\} \\ \times e^{2\sigma \ln \frac{L}{\ell_0}[\frac{1}{2}p_1(p_1-1) + \frac{1}{2}p_2(p_2-1) + \frac{1}{2}p(p-1)\frac{2n-p_1-p_2}{p}]}, \end{aligned} \tag{B17}$$

where $\{...\}$ is L independent. Hence,

$$\begin{aligned} \frac{\ln Q_{1,1}}{\ln L} &= \sigma[2n + p_1(p_1-1) + p_2(p_2-1) + (p-1)(2n-p_1-p_2)] + 2 + 2\frac{2n-p_1-p_2}{p} \\ &\overset{n\to 0}{=} \sigma\left[p_1^2 + p_2^2 - p(p_1+p_2)\right] + 2 - 2\frac{p_1+p_2}{p}. \end{aligned} \tag{B18}$$

Minimizing with respect to p, p_1, p_2 yields $p = p_1 = p_2 = 1$ for $\sigma < 2$ while $p = p_1 = p_2 = 1/\sqrt{\sigma/2}$ for $\sigma > 2$. Thus, also $Q_{1,1}$ is independent of σ,

$$Q_{1,1} \approx \left(\frac{L}{\ell_0}\right)^{-2} \ell_0^{-2} \qquad \sigma > 0. \tag{B19}$$

Consider next the density overlap of two states displaced by $r_0 \gg \ell_0$, supporting the form of U_2 in Eq. (17),

$$Q(r_0) = \overline{\frac{\int d^2r e^{2V(\mathbf{r})+2V(\mathbf{r}-\mathbf{r}_0)}}{[\int d^2r e^{2V(\mathbf{r})}]^2}} = \int d^2r \prod_{i=1}^{2n-2} d^2r_i\, \overline{e^{2V(\mathbf{r})+2V(\mathbf{r}-\mathbf{r}_0)+2\sum_{i=1}^{2n-2} V(\mathbf{r}_i)}}. \tag{B20}$$

The evaluation is similar to that of $Q_{1,1}$ with $-\mathbf{r}$ replaced by $\mathbf{r} - \mathbf{r}_0$. Since $\mathbf{r}_0 \gg \ell_0$ the packets containing either $\mathbf{r}$ or $\mathbf{r} - \mathbf{r}_0$ are distinct. Hence, the derivation is the same as for $Q_{1,1}$ with the result $Q(\mathbf{r}_0) \approx 1/L^2$.

We note here that our sum rules in Sec. IV are proven for large σ. In fact, one can check the consistency of identifying ϵ as a chemical potential in Eq. (13) by substituting the result (17) in the first line of Eq. (13) and using $\overline{\rho(E, \mathbf{r})}$. This reproduces precisely Eq. (14) except for a factor $(1 - \frac{1}{z})$; hence, the derivation is valid for $z \gg 1$.

We summarize here the various Coulomb matrix elements of Eq. (17) for the $E = 0$ states and for all σ. These matrix elements serve as indicators for the cascade transitions also at $\sigma < 2$.

$$\begin{aligned} U_1 = U\ell_0^2 P_{2,2} &\approx U \quad \sigma > 2 \\ &\approx U\left(\frac{L}{\ell_0}\right)^{-\frac{(2-\sigma)^2}{4\sigma}} \quad \frac{2}{3} < \sigma < 2 \\ &\approx U\left(\frac{L}{\ell_0}\right)^{2\sigma-2} \quad \sigma < \frac{2}{3} \\ U_2 = U\ell_0^2 Q_{1,\pm 1} &\approx \left(\frac{L}{\ell_0}\right)^{-2} \quad \text{all } \sigma \\ U_3 = U\ell_0^2 P_{2,0} &\approx \left(\frac{L}{\ell_0}\right)^{2-8\sqrt{\sigma/2}} \quad \sigma > 2 \\ &\approx \left(\frac{L}{\ell_0}\right)^{-2\sigma-2} \quad \sigma < 2. \end{aligned} \tag{B21}$$

Hence, all matrix elements and the resulting cascade transitions are σ independent at $\sigma > 2$, in particular U_3 is vanishingly small. In contrast, at $\sigma < 2$ the matrix element U_1 decreases while U_3 becomes comparable to U_1, U_2 so we expect that the form and strength of the cascade peaks change considerably. Thus the freezing transition, which is a minor effect in the DOS, becomes prominent when observing the cascade transitions.

APPENDIX C: DISORDER AVERAGE $\overline{\rho_\epsilon^2}$

We evaluate here the disorder average $\overline{\rho_\epsilon^2}$, which is used in Eq. (14) to evaluate the density overlaps of all $E \neq 0$ states. This is achieved by considering a two-layer problem, following Ref. [28], Sec. IV B and Ref. [12], Sec. IV B. The average DOS itself, summarized below Eq. (8), is a similar one-layer problem. The layers $\uparrow$, $\downarrow$ are identical; for later use they have distinct fugacities $\epsilon_\uparrow$ $\epsilon_\downarrow$, so the interaction and the Gaussian terms, and the variational Hamiltonian are [eventually we replace $\epsilon_\uparrow$, $\epsilon_\downarrow \to \tilde{\epsilon}$ as defined below Eq. (6)],

$$\mathcal{H}_{\rm int} = \frac{1}{\pi\alpha}[\epsilon_\uparrow \cos\phi_\uparrow + \epsilon_\downarrow \cos\phi_\downarrow] \to \sum_{\{\mathbf{n}_\uparrow,\mathbf{n}_\downarrow \neq 0\}} \epsilon_\uparrow^{\sum_a n_{\uparrow a}^2} \epsilon_\downarrow^{\sum_a n_{\downarrow a}^2} \mathrm{e}^{i\mathbf{n}_\uparrow\cdot\boldsymbol{\phi}_\uparrow + i\mathbf{n}_\downarrow\cdot\boldsymbol{\phi}_\downarrow}$$

$$\mathcal{H}' = \int d^2r \left[\frac{1}{8\pi} \sum_{a,b} (K^{-1}\delta_{a,b} + 2\sigma) \nabla\phi_a^+ \cdot \nabla\phi_b^+ + \frac{1}{8\pi K} \nabla\phi_a^- \cdot \nabla\phi_b^- \right]$$

$$\mathcal{H}_0 = \mathcal{H}' + \tfrac{1}{2} \int d^2r \sum_{a,b,\pm} (\sigma_c^\pm \delta_{ab} + \sigma_0^\pm) \phi_a^\pm \phi_b^\pm, \tag{C1}$$

where $\phi^\pm = (\phi^\uparrow \pm \phi^\downarrow)/\sqrt{2}$ and the full Hamiltonian is $\mathcal{H} = \mathcal{H}' - \sum_{\mathbf{r}} \mathcal{H}_{\rm int}$. Note that the disorder couples only to ϕ_a^+ with strength 2σ. The Green functions for $\mathcal{H}_0$ yield, defining $u_\pm, A_\pm$ as in Ref. [12]

$$\int_q G_\pm = -4u_\pm \delta_{ab} - 2A_\pm, \quad u_\pm = \frac{K}{4} \ln\left(4\pi K \ell_0^2 \sigma_c^\pm\right)$$

$$A_+ = -\sigma K^2 \ln\left(4\pi K \ell_0^2 \sigma_c^+\right) + K\frac{\sigma_0^+}{2\sigma_c^+} - K^2\sigma, \quad A_- = K\frac{\sigma_0^-}{2\sigma_c^-}, \tag{C2}$$

as in Eq. (58) of Ref. [12] and $1/\ell_0$ serves as a momentum cutoff. The interaction average becomes

$$\begin{aligned}
\langle \mathcal{H}_{\rm int} \rangle_0 &= \sum_{\{\mathbf{n}_\uparrow,\mathbf{n}_\downarrow \neq 0\}} \epsilon_\uparrow^{\sum_a n_{\uparrow a}^2} \epsilon_\downarrow^{\sum_a n_{\downarrow a}^2} \langle \mathrm{e}^{i\mathbf{n}_\uparrow\cdot(\boldsymbol{\phi}_+ + \boldsymbol{\phi}_-)/\sqrt{2} + i\mathbf{n}_\downarrow\cdot(\boldsymbol{\phi}_+ - \boldsymbol{\phi}_-)/\sqrt{2}} \rangle_0 \\
&= \sum_{\{\mathbf{n}_\uparrow,\mathbf{n}_\downarrow \neq 0\}} \epsilon_\uparrow^{\sum_a n_{\uparrow a}^2} \epsilon_\downarrow^{\sum_a n_{\downarrow a}^2} \mathrm{e}^{u_+ \sum_a (n_{\uparrow a}+n_{\downarrow a})^2 + \frac{1}{2}A_+ (\sum_a n_{\uparrow a} + n_{\downarrow a})^2 + u_- \sum_a (n_{\uparrow a} - n_{\downarrow a})^2 + \frac{1}{2}A_- (\sum_a n_{\uparrow a} - n_{\downarrow a})^2} \\
&= \sum_{\{n_{\alpha\beta}\}} \frac{m!}{n_{00}! n_{+0}! \ldots n_{++}!} \langle \epsilon_\uparrow^{\sum_a n_{\uparrow a}^2} \epsilon_\downarrow^{\sum_a n_{\downarrow a}^2} \mathrm{e}^{u_+ \sum_a (n_{\uparrow a}+n_{\downarrow a})^2 + u_- \sum_a (n_{\uparrow a} - n_{\downarrow a})^2 + \omega_+ \sum_a (n_{\uparrow a}+n_{\downarrow a}) + \omega_- \sum_a (n_{\uparrow a} - n_{\downarrow a})} \rangle_\omega \\
\langle \ldots \rangle_\omega &\equiv \iint \mathrm{e}^{-\frac{\omega_+^2}{2A_+} - \frac{\omega_-^2}{2A_-}} \frac{d\omega_+ d\omega_-}{2\pi\sqrt{A_+ A_-}},
\end{aligned} \tag{C3}$$

$n_{\alpha\beta}$ with α, $\beta = \pm 1, 0$, count the number of pairs that have α entry in the $\uparrow$ layer *and* β entry in the $\downarrow$ layer, $\sum_{\alpha\beta} n_{\alpha\beta} = m$. All the replica sums in the exponent can be written in terms of $n_{\alpha\beta}$. The sum can be written as

$$\begin{aligned}
&\sum_{\{n_{\alpha\beta}\}} \frac{m!}{n_{00}! n_{+0}! \ldots n_{++}!} \langle (\epsilon_\uparrow \mathrm{e}^{u_+ + u_- + \omega_+ + \omega_-})^{n_{+0}} (\epsilon_\downarrow \mathrm{e}^{u_+ + u_- + \omega_+ - \omega_-})^{n_{0+}} (\epsilon_\downarrow \mathrm{e}^{u_+ + u_- - \omega_+ + \omega_-})^{n_{0-}} (\epsilon_\uparrow \mathrm{e}^{u_+ + u_- - \omega_+ - \omega_-})^{n_{-0}} \\
&\times (\epsilon_\uparrow \epsilon_\downarrow \mathrm{e}^{4u_+ + 2\omega_+})^{n_{++}} (\epsilon_\uparrow \epsilon_\downarrow \mathrm{e}^{4u_- + 2\omega_-})^{n_{+-}} (\epsilon_\uparrow \epsilon_\downarrow \mathrm{e}^{4u_+ - 2\omega_+})^{n_{--}} (\epsilon_\uparrow \epsilon_\downarrow \mathrm{e}^{4u_- - 2\omega_-})^{n_{-+}} \rangle_\omega.
\end{aligned} \tag{C4}$$

To prove this, check the exponents of $\epsilon_\uparrow, \epsilon_\downarrow$ and the coefficients of either of $u_\pm$, $\omega_\pm$:

$$\epsilon_\uparrow : n_{+0} + n_{-0} + n_{++} + n_{+-} + n_{--} + n_{-+} = \sum_a n_{\uparrow a}^2$$

$$\epsilon_\downarrow : n_{0+} + n_{0-} + n_{++} + n_{+-} + n_{--} + n_{-+} = \sum_a n_{\downarrow a}^2$$

$$u_+ : n_{+0} + n_{0+} + n_{0-} + n_{-0} + 4n_{++} + 4n_{--} = \sum_a (n_{\uparrow a} + n_{\downarrow a})^2$$

$$u_- : n_{+0} + n_{0+} + n_{0-} + n_{-0} + 4n_{+-} + 4n_{-+} = \sum_a (n_{\uparrow a} - n_{\downarrow a})^2$$
$$\omega_+ : n_{+0} + n_{0+} - n_{0-} - n_{-0} + 2n_{++} - 2n_{--} = \sum_a (n_{\uparrow a} + n_{\downarrow a})$$
$$\omega_- : n_{+0} - n_{0+} + n_{0-} - n_{-0} + 2n_{+-} - 2n_{-+} = \sum_a (n_{\uparrow a} - n_{\downarrow a}). \tag{C5}$$

This confirms (C3) so that the summation is a ninomial expansion,

$$\begin{aligned} Z = 1 &+ \epsilon_\uparrow e^{u_+ + u_- + \omega_+ + \omega_-} + \epsilon_\downarrow e^{u_+ + u_- + \omega_+ - \omega_-} + \epsilon_\downarrow e^{u_+ + u_- - \omega_+ + \omega_-} + \epsilon_\uparrow e^{u_+ + u_- - \omega_+ - \omega_-} \\ &+ \epsilon_\uparrow \epsilon_\downarrow e^{4u_+ + 2\omega_+} + \epsilon_\uparrow \epsilon_\downarrow e^{4u_- + 2\omega_-} + \epsilon_\uparrow \epsilon_\downarrow e^{4u_+ - 2\omega_+} + \epsilon_\uparrow \epsilon_\downarrow e^{4u_- - 2\omega_-} \\ \frac{1}{m}\langle \mathcal{H}_{\rm int}\rangle_0 = \frac{1}{m}\langle Z^m\rangle_\omega &\overset{m\to 0}{=} \langle \ln Z\rangle_\omega. \end{aligned} \tag{C6}$$

As in Ref. [12] we expect the solution where $\sigma_c^- = 0$ (i.e., it is irrelevant). Hence, $u_- \to -\infty$ and

$$Z = 1 + \epsilon_\uparrow \epsilon_\downarrow e^{4u_+ + 2\omega_+} + \epsilon_\uparrow \epsilon_\downarrow e^{4u_+ - 2\omega_+}. \tag{C7}$$

Denoting $\tilde\epsilon = \sqrt{\epsilon_\uparrow \epsilon_\downarrow}$, by a similar calculation as the one leading to (15)–(17) in Ref. [12] the self-consistent equation is now

$$\sigma_c^+ \sim \mathrm{Prob}(2u_+ + \omega_+ + \log\tilde\epsilon > 0) \sim \exp\left(\frac{\left(K + 2\frac{\log\tilde\epsilon}{\log\sigma_c^+}\right)^2}{8\sigma K^2} \log\sigma_c^+ \right), \tag{C8}$$

which means that σ_c^+ is dominated by the tail of the Gaussian distribution of ω_+. This leads to $\sigma_c^+ \sim \tilde\epsilon^{\,2/z}$, with the previous $z = K(\sqrt{8\sigma} - 1)$. A stronger argument, valid more generally is to note (following Ref. [12]) that the solution is the same as in the one-layer case with $K \to 2K$ and $\epsilon_\downarrow \epsilon_\uparrow \to \tilde\epsilon^2$; hence, $\sigma_c^+ \sim (\tilde\epsilon^2)^{2/2z} = \tilde\epsilon^{\,2/z}$ with the previous $z = K(\sqrt{8\sigma} - 1)$.

Our aim is to find the disorder average of $\rho_{\epsilon_\uparrow}\rho_{\epsilon_\downarrow}$. Using the interaction average [Eq. (C6)] and the bosonic form of Eq. (7),

$$\begin{aligned} \overline{\rho_\epsilon^2} &= \left(\frac{1}{\pi\ell_0^2}\right)^2 \left(\frac{\tilde\epsilon}{\epsilon}\right)^2 \frac{\partial^2}{\partial\epsilon_\uparrow \partial\epsilon_\downarrow} \langle \ln[1 + \epsilon_\uparrow\epsilon_\downarrow e^{4u_+ + 2\omega_+} + \epsilon_\uparrow\epsilon_\downarrow e^{4u_+ - 2\omega_+}]\rangle_\omega \\ &= \left(\frac{1}{\pi\ell_0^2}\right)^2 \left(\frac{\tilde\epsilon}{\epsilon}\right)^2 \left\langle \frac{e^{4u_+ + 2\omega_+} + e^{4u_+ - 2\omega_+}}{[1 + \tilde\epsilon^2(e^{4u_+ + 2\omega_+} + e^{4u_+ - 2\omega_+})]^2} \right\rangle_\omega = \left(\frac{1}{\pi\ell_0^2}\right)^2 \frac{\ell_0^2 \sigma_c(2K, \tilde\epsilon^2)}{\epsilon^2}. \end{aligned} \tag{C9}$$

In the last equality we identify the mass parameter σ_c of the DOS evaluation as given in Eq. (15) of Ref. [12] and note that e^{2u} is negligible in that expression. The result (C9) is reproduced in Eq. (14), which eventually leads to the various density overlaps as summarized in Eq. (17).

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

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

[3] A. T. Pierce, Y. Xie, J. M. Park, E. Khalaf, S. H. Lee, Y. Cao, D. E. Parker, P. R. Forrester, S. Chen, K. Watanabe, T. Taniguchi, A. Vishwanath, P. Jarillo-Herrero, and A. Yacoby, Unconventional sequence of correlated Chern insulators in magic-angle twisted bilayer graphene, Nat. Phys. **17**, 1210 (2021).

[4] J. M. Park, Y. Cao, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, Flavour Hund's coupling, Chern gaps and charge diffusivity in moiré graphene, Nature (London) **592**, 43 (2021).

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

[6] A. T. Pierce, Y. Xie, J. M. Park, Z. Cai, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, and A. Yacoby, Tunable interplay between light and heavy electrons in twisted trilayer graphene, Nat. Phys. **21**, 1237 (2025); Y. Xie, *et al.*, Strong interactions and isospin symmetry breaking in a supermoiré lattice, arXiv:2404.01372.

[7] X. Liu, N. J. Zhang, K. Watanabe, T. Taniguchi, and J. I. A. Li, Isospin order in superconducting magic-angle twisted trilayer graphene, Nat. Phys. **18**, 522 (2022).

[8] A. Uri, S. Grover, Y. Cao, J. A. Crosse, K. Bagani, D. Rodan-Legrain, Y. Myasoedov, K. Watanabe, T. Taniguchi, P. Moon, M. Koshino, P. Jarillo-Herrero, and E. Zeldov, Chern mosaic and Berry-curvature magnetism in magic-angle graphene, Nature (London) **581**, 47 (2020).

[9] G. Rai *et al.*, Dynamical correlations and order in magic-angle twisted bilayer graphene, Phys. Rev. X **14**, 031045 (2024).
[10] Its longitudinal part can be gauged away as a random phase, immaterial in the following.
[11] A. W. W. Ludwig, Matthew P. A. Fisher, R. Shankar, and G. Grinstein, Integer quantum Hall transition: An alternative approach and exact results, Phys. Rev. B **50**, 7526 (1994).
[12] B. Horovitz and P. Le Doussal, Freezing transitions and the density of states of two-dimensional random Dirac Hamiltonians, Phys. Rev. B **65**, 125323 (2002).
[13] D. R. Nelson and B. I. Halperin, Dislocation-mediated melting in two dimensions, Phys. Rev. B **19**, 2457 (1979).
[14] F. Guinea, B. Horovitz, and P. Le Doussal, Gauge field induced by ripples in graphene, Phys. Rev. B **77**, 205421 (2008).
[15] A. Ceferino and F. Guinea, Pseudomagnetic fields in fully relaxed twisted bilayer and trilayer graphene, 2D Mater. **11**, 035015 (2024).
[16] N. N. T. Nam and M. Koshino, Lattice relaxation and energy band modulation in twisted bilayer graphene, Phys. Rev. B **96**, 075311 (2017).
[17] Mohammed M. Al Ezzi, G. N. Pallewela, C. De Beule, E. J. Mele, and S. Adam, Analytical model for atomic relaxation in twisted moiré materials, Phys. Rev. Lett. **133**, 266201 (2024).
[18] M. S. Foster, H.-Y. Xie, and Y.-Z. Chou, Topological protection, disorder, and interactions: Survival at the surface of three-dimensional topological superconductors, Phys. Rev. B **89**, 155140 (2014).
[19] B A. Bernevig, Z.-D. Song, N. Regnault, and B. Lian, Twisted bilayer graphene. III. Interacting Hamiltonian and exact symmetries, Phys. Rev. B **103**, 205413 (2021).
[20] Y. Sheffer, R. Queiroz, and A. Stern, Symmetries as the guiding principle for flattening bands of Dirac fermions, Phys. Rev. X **13**, 021012 (2023).
[21] M. Christos, S. Sachdev, and S. Scehurer, Superconductivity, correlated insulators, and Wess–Zumino–Witten terms in twisted bilayer graphene, Proc. Natl. Acad. Sci. USA **117**, 29543 (2020).
[22] B. Sbierski, J. F. Karcher, and M. S. Foster, Spectrum-wide quantum criticality at the surface of Class AIII topological phases: An "energy stack" of integer quantum Hall plateau transitions, Phys. Rev. X **10**, 021025 (2020).
[23] T. Giamarchi, *Quantum Physics in One Dimension* (Clarendon, Oxford, 2003).
[24] S. Scheidl, Glassy vortex state in a two-dimensional disordered XY model, Phys. Rev. B **55**, 457 (1997).
[25] D. Carpentier and P. Le Doussal, Disordered XY models and Coulomb gases: Renormalization via traveling waves, Phys. Rev. Lett. **81**, 2558 (1998); Topological transitions and freezing in XY models and Coulomb gases with quenched disorder: Renormalization via traveling waves, Nucl. Phys. B **588**, 565 (2000).
[26] D. Carpentier and P. Le Doussal, Glass transition of a particle in a random potential, front selection in nonlinear renormalization group, and entropic phenomena in Liouville and sinh-Gordon models, Phys. Rev. E **63**, 026110 (2001).
[27] B. Horovitz and P. Le Doussal, Disorder induced transitions in layered Coulomb gases and superconductors, Phys. Rev. Lett. **84**, 5395 (2000).
[28] B. Horovitz and P. Le Doussal, Disorder-induced transitions in layered Coulomb gases and application to flux lattices in superconductors, Phys. Rev. B **71**, 134202 (2005).
[29] H. E. Castillo, C. de C. Chamon, E. Fradkin, P. M. Goldbart, and C. Mudry, Exact calculation of multifractal exponents of the critical wave function of Dirac fermions in a random magnetic field, Phys. Rev. B **56**, 10668 (1997).
[30] Y. V. Fyodorov, Pre-freezing of multifractal exponents in random energy models with a logarithmically correlated potential, J. Stat. Mech. (2009) P07022.
[31] Y.-Z. Chou and M. S. Foster, Chalker scaling, level repulsion, and conformal invariance in critically delocalized quantum matter: Disordered topological superconductors and artificial graphene, Phys. Rev. B **89**, 165136 (2014).
[32] A. Rozen, J. M. Park, U. Zondiner, Y. Cao, D. Rodan-Legrain, T. Taniguchi, K. Watanabe, Y. Oreg, A. Stern, E. Berg, P. Jarillo-Herrero, and S. Ilani, Entropic evidence for a Pomeranchuk effect in magic-angle graphene, Nature (London) **592**, 214 (2021).
[33] X. Zhang, Z. Zhu, J. H. Wilson, and M. S. Foster, Twisted trilayer graphene, quasiperiodic superconductor, arXiv:2512.22340.
[34] S. Carr, C. Li, Z. Zhu, E. Kaxiras, S. Sachdev, and A. Kruchkov, Ultraheavy and ultrarelativistic Dirac quasiparticles in sandwiched graphenes, Nano Lett. **20**, 3030 (2020).
[35] X. Li, F. Wu, and A. H. MacDonald, Electronic structure of single-twist trilayer graphene, arXiv:1907.12338.