

# Magic angle semimetals SUPPLEMENTARY INFORMATION

Yixing Fu,<sup>1,\*</sup> Elio J. König,<sup>1,\*</sup> Justin H. Wilson,<sup>2,1,\*</sup> Yang-Zhi Chou,<sup>3,4</sup> and Jedediah H. Pixley<sup>1</sup>

<sup>1</sup>*Department of Physics and Astronomy, Center for Materials Theory, Rutgers University, Piscataway, NJ 08854 USA*

<sup>2</sup>*Institute of Quantum Information and Matter and Department of Physics,  
California Institute of Technology, Pasadena, California 91125 USA*

<sup>3</sup>*Condensed Matter Theory Center and the Joint Quantum Institute,  
Department of Physics, University of Maryland, College Park, MD 20742 USA*

<sup>4</sup>*Department of Physics and Center for Theory of Quantum Matter,  
University of Colorado Boulder, Boulder, Colorado 80309, USA*

(Dated: September 8, 2020)

## Supplementary Note 1. MODELS

In this section we define the models which we analyze.

*Semimetals in an incommensurate scalar potential.* The tight-binding Hamiltonians of models dubbed “perfect” spin orbit coupling (SOC) are given by

$$\hat{T}_{SOC} = \sum_{\mathbf{r},\mu} \left[ \frac{it}{2} c_{\mathbf{r}}^{\dagger} \sigma_{\mu} c_{\mathbf{r}+\hat{\mu}} + \text{h.c.} \right], \quad (1)$$

$$\hat{V}_{SOC} = W \sum_{\mathbf{r},\mu} \cos(Qr_{\mu} + \phi_{\mu}) c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}}, \quad (2)$$

where  $t$  is the hopping matrix element,  $\sigma_{\mu}$  are the Pauli matrices,  $c_{\mathbf{r}}$  are electron annihilation operators,  $Q$  is our quasiperiodic wave vector,  $\phi_{\mu}$  are the offsets of the origin of the potential, and  $W$  is the amplitude of the quasiperiodic potential. In the two-dimensional (2D) case  $\mu = x, y$  and in the three-dimensional (3D) case  $\mu = x, y, z$  and  $\mathbf{r}$  takes values in the set of all lattice points on a square (cubic) lattice. The Hamiltonian for the  $\pi$ -flux model has the same potential term in 2D. The hopping term is modified as follows

$$\hat{T}_{\pi} = -t \sum_{\mathbf{r},\mu=x,y} [c_{\mathbf{r}}^{\dagger} e^{iA_{\mu}(\mathbf{r})} c_{\mathbf{r}+\hat{\mu}} + \text{h.c.}], \quad (3)$$

where we choose the gauge with  $A_x(\mathbf{r}) = \pi/2$  for all sites  $\mathbf{r}$  on the square lattice, and  $A_y(\mathbf{r}) = -(-1)^{r_x} \pi/2$ . For the chosen gauge, periodic boundary conditions require the lattice size in  $x$  direction to be even.

The spinless honeycomb (HC) lattice model is given by a Hamiltonian of the form

$$\hat{T}_{HC} = -t \sum_{\mathbf{r}_A, i} [c_A^{\dagger}(\mathbf{r}_A) c_B(\mathbf{r}_A + \mathbf{d}_i) + \text{h.c.}], \quad (4)$$

$$\hat{V}_{HC} = W \sum_{\mathbf{r}, \delta_{\mu}} \cos(Q\mathbf{r} \cdot \delta_{\mu} + \phi_{\mu}) c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}}. \quad (5)$$

The sum over  $\mathbf{r}_A$  is over one of the two sub-lattices, while  $\mathbf{r}$  is over all points. The index  $i$  labels the three nearest neighbors of  $\mathbf{r}_A$ , and  $\mathbf{d}_i$  is the vector from  $\mathbf{r}_A$  to its nearest neighbor  $i$ . The vectors  $\delta_{\mu}$  are a choice of each particular model and for numerics we choose  $\delta_1 = \mathbf{d}_1 = (2/3)\hat{x}$  and  $\delta_2 = \mathbf{d}_2 = -(1/3)\hat{x} + (1/\sqrt{3})\hat{y}$ . The kinetic part of the Hamiltonian for the one dimensional model with power law dispersion<sup>1</sup> is given in momentum space

$$\hat{T}_{1D} = -t \sum_k \text{sgn}[\cos(k)] |\cos(k)|^{\sigma} c_k^{\dagger} c_k. \quad (6)$$

We assume  $\sigma < 1$ , this expression can be readily Fourier transformed to a tight binding model with long range hopping (LRH). This yields a hopping amplitude

$$t_{ij} \sim -2t[1 - (-1)^{|i-j|}] \sin[\pi(|i-j| - \sigma)/2] \Gamma(1 + \sigma) |i-j|^{-(1+\sigma)} \quad (7)$$

for  $|i - j| \gg 1$  and  $\Gamma(x)$  is the Gamma function. The potential is

$$\hat{V}_{1D} = W \sum_r \cos(Qr + \phi) c_r^\dagger c_r. \quad (8)$$

*Chiral Twisted Bilayer Graphene.* The model we use for twisted bilayer graphene simulates the chiral version of the continuum model<sup>2-4</sup>.

To write down the full model, we have first that

$$\hat{T}_{\text{cTBG}} = -t \sum_{a=1,2} \sum_{\langle ij \rangle} c_{a,\mathbf{r}_i}^\dagger c_{a,\mathbf{r}_j}, \quad (9)$$

where  $\langle ij \rangle$  indicates nearest neighbors on the honeycomb lattice and  $a$  labels the two layers. This model has four Dirac nodes: two per layer. Furthermore, for reporting figures, we take  $t = 2.8 \text{ eV}$  as it is for graphene.

We then couple the two layers with a quasiperiodically modulated tunnelling to simulate the effect of twisting

$$\begin{aligned} \hat{V}_{\text{cTBG}} = W \sum_{\mathbf{r}} \sum_j \left\{ (c_{1,A,\mathbf{r}+\boldsymbol{\eta}_j}^\dagger c_{2,B,\mathbf{r}} + c_{1,B,\mathbf{r}}^\dagger c_{2,A,\mathbf{r}+\boldsymbol{\eta}_j}) \cos[\mathbf{q}_j \cdot (\mathbf{r} + \boldsymbol{\eta}_j/2) + \phi_j] \right. \\ \left. - \frac{1}{3\sqrt{3}} \sum_{n=1}^6 (-1)^n (c_{1,A,\mathbf{r}+\boldsymbol{\eta}_j+\mathbf{a}_n}^\dagger c_{2,B,\mathbf{r}} + c_{1,B,\mathbf{r}}^\dagger c_{2,A,\mathbf{r}+\boldsymbol{\eta}_j+\mathbf{a}_n}) \sin[\mathbf{q}_j \cdot (\mathbf{r} + (\boldsymbol{\eta}_j + \mathbf{a}_j)/2) + \phi_j] \right\} + \text{h.c.}, \quad (10) \end{aligned}$$

where  $\mathbf{r}$  is on the triangular Bravais lattice and  $\boldsymbol{\eta}_j$  describe the vectors to connect nearest neighbors (B sites to A sites) and they are given by

$$\begin{aligned} \boldsymbol{\eta}_1 &= (0, 1), \\ \boldsymbol{\eta}_2 &= (-\sqrt{3}/2, -1/2), \\ \boldsymbol{\eta}_3 &= (\sqrt{3}/2, 1/2). \end{aligned} \quad (11)$$

The vectors  $\mathbf{a}_j$  are the six nearest neighbors on the triangular lattice defined by  $\mathbf{a}_1 = \boldsymbol{\eta}_1 - \boldsymbol{\eta}_2$ ,  $\mathbf{a}_2 = \boldsymbol{\eta}_1 - \boldsymbol{\eta}_3$ ,  $\mathbf{a}_3 = \mathbf{a}_2 - \mathbf{a}_1$ ,  $\mathbf{a}_4 = -\mathbf{a}_1$ ,  $\mathbf{a}_5 = -\mathbf{a}_2$ , and  $\mathbf{a}_6 = \mathbf{a}_1 - \mathbf{a}_2$ . Last,  $\mathbf{q}_j$  are defined by the twist angle

$$\begin{aligned} \mathbf{q}_1 &= k_\theta(0, -1), \\ \mathbf{q}_2 &= k_\theta(\sqrt{3}/2, 1/2), \\ \mathbf{q}_3 &= k_\theta(-\sqrt{3}/2, 1/2), \end{aligned} \quad (12)$$

where  $k_\theta = 2k_D \sin(\theta/2)$  for twist angle  $\theta$  and  $k_D = 4\pi/(3\sqrt{3})$  is the distance from the  $\Gamma$  point to  $\mathbf{K}$  point.

To show how this reproduces the chiral model of twisted bilayer graphene, we can rewrite the above in  $\mathbf{k}$ -space

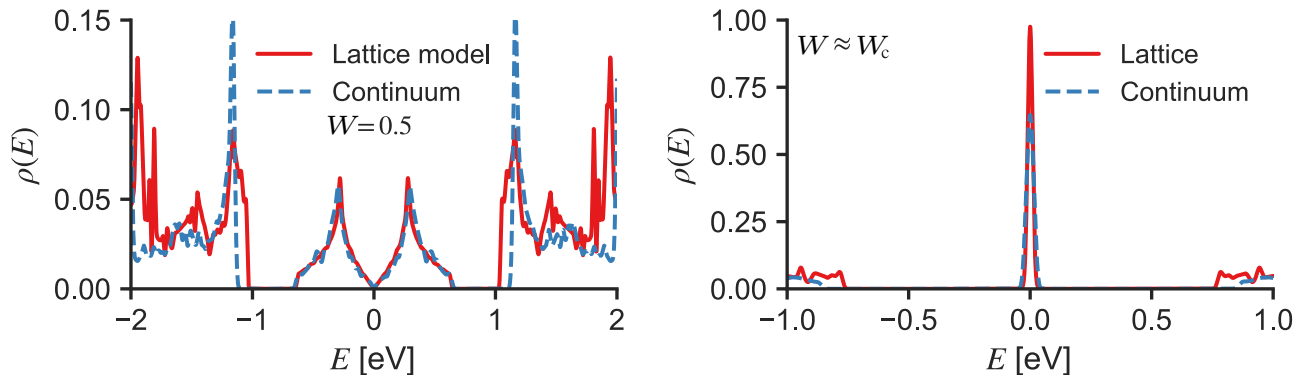
$$\begin{aligned} \hat{V}_{\text{cTBG}} = W \sum_{\mathbf{k}} \left[ c_{1,\mathbf{k}+\frac{\mathbf{q}_1}{2}}^\dagger \sigma_x c_{2,\mathbf{k}-\frac{\mathbf{q}_1}{2}} e^{i\phi_1} + c_{1,\mathbf{k}+\frac{\mathbf{q}_2}{2}}^\dagger (e^{-i\mathbf{k}\cdot\mathbf{a}_1} \sigma_+ + e^{i\mathbf{k}\cdot\mathbf{a}_1} \sigma_-) c_{2,\mathbf{k}-\frac{\mathbf{q}_2}{2}} e^{i\phi_2} \right. \\ \left. + c_{1,\mathbf{k}+\frac{\mathbf{q}_3}{2}}^\dagger (e^{-i\mathbf{k}\cdot\mathbf{a}_2} \sigma_+ + e^{i\mathbf{k}\cdot\mathbf{a}_2} \sigma_-) c_{2,\mathbf{k}-\frac{\mathbf{q}_3}{2}} e^{i\phi_3} \right] f(-\mathbf{k}) + w \sum_{\mathbf{k}} \left[ c_{1,\mathbf{k}-\frac{\mathbf{q}_1}{2}}^\dagger \sigma_x c_{2,\mathbf{k}+\frac{\mathbf{q}_1}{2}} e^{-i\phi_1} \right. \\ \left. + c_{1,\mathbf{k}-\frac{\mathbf{q}_2}{2}}^\dagger (e^{-i\mathbf{k}\cdot\mathbf{a}_1} \sigma_+ + e^{i\mathbf{k}\cdot\mathbf{a}_1} \sigma_-) c_{2,\mathbf{k}+\frac{\mathbf{q}_2}{2}} e^{-i\phi_2} + c_{1,\mathbf{k}-\frac{\mathbf{q}_3}{2}}^\dagger (e^{-i\mathbf{k}\cdot\mathbf{a}_2} \sigma_+ + e^{i\mathbf{k}\cdot\mathbf{a}_2} \sigma_-) c_{2,\mathbf{k}+\frac{\mathbf{q}_3}{2}} e^{-i\phi_3} \right] f(\mathbf{k}), \quad (13) \end{aligned}$$

where  $c_{a,\mathbf{k}} = (c_{aA\mathbf{k}}, c_{aB\mathbf{k}})^T$  with  $a$  labeling the layer, A and B labeling the sublattice, and  $\mathbf{k}$  is the lattice wave vector. The function  $f$  is real-valued and has the form

$$f(\mathbf{k}) = \frac{1}{2} + \frac{1}{6i\sqrt{3}} \sum_{n=1}^6 (-1)^n e^{-i\mathbf{k}\cdot\mathbf{a}_n}. \quad (14)$$

If we then concentrate near the  $\mathbf{K}$  or  $\mathbf{K}'$  points ( $\mathbf{K} = (-\frac{4\pi}{3\sqrt{3}}, 0)$  and  $\mathbf{K}' = (\frac{4\pi}{3\sqrt{3}}, 0)$ ), we have  $f(\mathbf{K}) = 1$  and  $f(\mathbf{K}') = 0$ . Furthermore,  $\mathbf{K} \cdot \mathbf{a}_1 = -2\pi/3$ ,  $\mathbf{K} \cdot \mathbf{a}_2 = 2\pi/3$ ,  $\mathbf{K}' \cdot \mathbf{a}_1 = 2\pi/3$ , and  $\mathbf{K}' \cdot \mathbf{a}_2 = -2\pi/3$ . The result is that

$$\hat{V}_{\text{cTBG}}|_{\text{near } \mathbf{K}} \approx W \sum_{\mathbf{k}} \sum_j c_{1,\mathbf{k}-\frac{\mathbf{q}_j}{2}}^\dagger T_j c_{2,\mathbf{k}+\frac{\mathbf{q}_j}{2}} e^{-i\phi_j}, \quad (15)$$



Supplementary Figure 1: (left) Comparison of density of states  $\rho(E)$  of the lattice and continuum models at  $W = 0.5$  eV at an angle  $\theta \approx 8.958^\circ$  approximated by taking the fraction  $k_\theta = \frac{4\pi}{3} \frac{11}{122}$ , we see excellent agreement in the miniband itself, and expected disagreement for higher energy bands beginning with the gap from miniband. (right) Comparison of density of states near the critical value of  $W$  for the angle  $\theta \approx 8.958^\circ$ . For the lattice model we observe this to happen for  $W_c \approx 0.875$  eV while for the continuum model we know this occurs at  $0.930$  eV. The lattice model and the continuum model do not match at high energy, which results in a modified value of the critical value of  $W$ . We see that the band is quite flat in either case, and just as in the  $W = 0.5$  eV case, the higher bands are at slightly different energies.

with

$$\begin{aligned}
 T_1 &= \sigma_x, \\
 T_2 &= e^{2\pi i/3} \sigma_+ + e^{-2\pi i/3} \sigma_-, \\
 T_3 &= e^{-2\pi i/3} \sigma_+ + e^{2\pi i/3} \sigma_-.
 \end{aligned} \tag{16}$$

This exactly describes the continuum model as described in Refs.<sup>2,4</sup> with AA tunnelling set to zero (the chiral limit). If we also look at  $\mathbf{K}'$  we find that we obtain the same low-energy model up to a unitary transformation.

Importantly, this model has  $C_3$  symmetry in addition to  $C_2$  symmetry and  $T$  (time reversal). These symmetries are contingent on  $e^{i\phi_1+i\phi_2+i\phi_3} = 1$  (these phases pick an origin of rotation for the symmetry being applied), but are otherwise built into the model.

To test our model against the continuum model, we now compare the density of states. In the continuum model, there is one parameter that controls the physics  $\alpha = W/(vk_\theta)$  where  $W$  is the tunneling strength between layers,  $v = \frac{3}{2}t$  is the Fermi velocity. Therefore, the continuum model has the same physics at  $\theta = 1.05^\circ$  as for larger angles such as  $\theta = 8.958^\circ$  as we are considering. At larger angles, the continuum model as an approximation for twisted bilayer graphene breaks down due to effects such as band curvature, and the same is true of the lattice model in this section, but we can observe density of state comparisons to see how well the lattice model is capturing the continuum model as a low-energy approximation. The results are in Supplementary Figure 1 and we see that the miniband is captured quite accurately.

In order to simulate this system on finite sizes with periodic boundary conditions, we find that we need  $k_\theta = \frac{4\pi}{3} \frac{n}{L}$  for integer  $n$  and system size  $L$  (for  $L^2$  Bravais lattice sites). Then, in order to fully capture an irrational number, such as  $k_\theta = \frac{4\pi}{3} \frac{1}{\varphi^5}$  for the golden ratio  $\varphi$ , we can use the continued fraction expansion defined such that

$$[a_0; a_1, a_2, \dots] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \dots}}, \tag{17}$$

if we truncate this such that for instance,  $[a_0; a_1, a_2, m]$  with  $0 < m \leq a_3$ , we get a rational that approximates the irrational number ( $m = a_3$  being quite a good approximation). Writing  $[a_0; a_1, a_2, m] = \frac{n}{L}$  for integers  $n, L$  gives us a system size  $L$  at which to simulate our system. For example,  $1/\varphi^5 = [0; 11, 11, 11, \dots]$  and we have  $[0; 11, 11] = \frac{11}{122}$  while  $[0; 11, 11, 3] = \frac{34}{377}$  (approximates of our irrational number for system sizes  $L = 122$  and  $L = 377$  respectively).

## Supplementary Note 2. STRUCTURE OF THE SPECTRUM AND SCALING OF THE DENSITY OF STATES

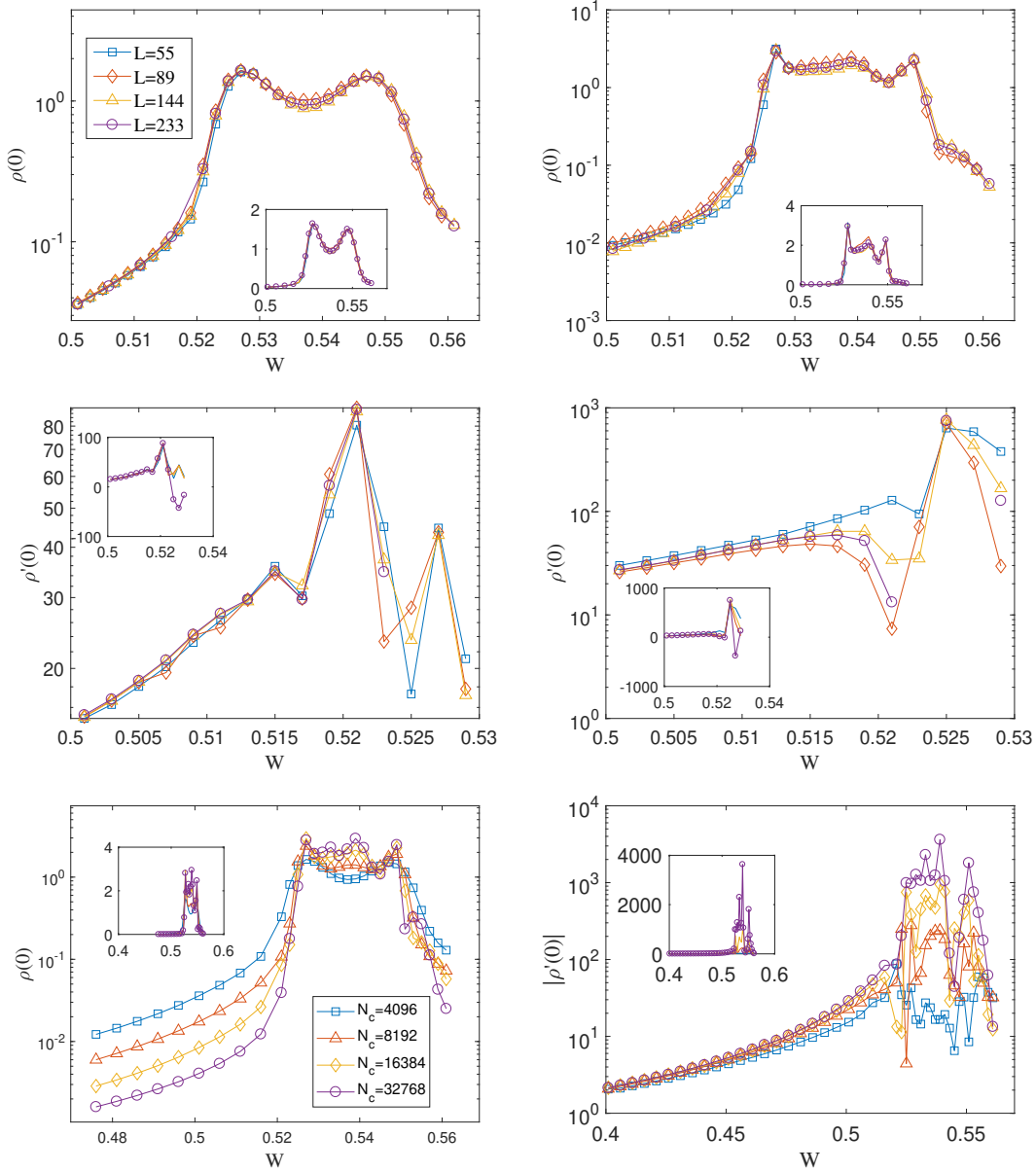
In this section we discuss the numerical method used in the analysis of the spectrum and the finite size effects of the method. We use the kernel polynomial method (KPM) to calculate the density of states  $\rho(E)$ , which expands  $\rho(E)$  in terms of Chebyshev polynomials up to an order  $N_c$ , and we use the Jackson kernel to filter out Gibbs oscillations due to the finite expansion order. To determine the velocity  $v$ , in two-dimensions for example, we fit the low- $|E|$  asymptote  $\rho(E) \approx \rho'(E=0)|E|$  to extract  $\rho'(E=0) \sim 1/v^2$ . Note that in 2D formally  $\rho'(E=0)$  is not just a single derivative due to the  $|E|$  scaling, but we use this notation to unify 2D and 3D; the latter it is simply a second derivative. For details on the KPM technique see Ref. 5. We use twisted boundary conditions and we average over random twists to reduce finite size effects. Now we discuss the effect of finite lattice size  $L$  and finite cutoff  $N_c$  on  $\rho(E=0)$  and  $\rho'(E=0)$ .

As an exemplary case we present results here for the “perfect” SOC and the cTBG models defined in the main text. Results on the other models are similar and we also present results on the 1D model below. Supplementary Figure 2 for the 2D SOC illustrates the dependence on  $L$  and  $N_c$ . For smaller  $N_c$  such as  $N_c = 2^{12} = 4096$ ,  $\rho'(0)$  for all choice of  $L \geq 55$  almost overlap for  $W \leq 0.515$ . For  $N_c = 2^{14} = 16384$ , the  $\rho'(0)$  data converges as a function of  $L$  only for  $L \geq 144$ . Still, the  $L$  convergence is only valid for  $W \leq 0.515$ . This demonstrates that the observed convergence in  $L$  is strongly dependent on  $N_c$  and therefore requires studying the scaling in  $N_c$  for fixed  $L$ . In Supplementary Figure 3(bottom row) we summarize similar features for cTBG showing Dirac nodes before and after the “magic-angle” and within the metallic phase.

When fixing  $L$  and varying  $N_c$ , the semimetal-to-metal transition becomes sharper as  $\rho(E=0)$  rises more abruptly approaching a sharp step as shown in Supplementary Figure 2. This sharpening allows us to pinpoint the location of the transitions accurately, in this case we find  $W_c = 0.525 \pm 0.005$  and  $W'_c = 0.551 \pm 0.005$ . Importantly, the peak of  $\rho(0)$  does not shrink as we vary  $L$  or  $N_c$ , providing strong evidence of the presence of the intermediate metallic phase. In addition, we find that  $\rho'(0)$  does not saturate as we increase the expansion order, indicating within our numerical accuracy that at the transition  $\rho'(0)$  diverges, similar to what was found in 3D<sup>6</sup>. From the above data of  $\rho'(0)$  we determine the scaling exponent  $\beta$  defined by  $\rho'(0) \sim |W_c - W|^{-\beta}$ . We use  $\rho'(0)$  data obtained for  $N_c = 2^{14}$  and  $L = 144$  and we extract  $\beta$  from a log-log fit of  $1/\rho'(0)$  versus  $|W - W_c|$ , see Supplementary Figure 4.

*Dispersion and velocity.* Here we demonstrate the identification of the kinetic velocity as obtained from the twist dispersion with the parameter entering the low-energy asymptote of the DOS. We also compare these numerical results with perturbation theory. We implement twisted boundary conditions by including a factor  $e^{i\theta \cdot \mathbf{r}/L}$  for each real space field located at  $\mathbf{r}$  and twist vector  $\theta$ . Each component of  $\theta$  takes value in  $(0, 2\pi)$  and we compute the energy eigenstates  $E(\theta)$  using exact diagonalization for various values of the twist  $\theta$ . Such a change of boundary condition has no effect on the bulk physics, but effectively moves the origin of the finite size induced momentum grid, so that plotting the spectrum as a function of the twist shows a projection of the dispersion onto  $1/L$ th of the Brillouin zone. Supplementary Figure 5 shows the twist dispersion for various models in one, two, and three dimensions, which clearly demonstrates the dramatic flattening of the bands at the transition. These results were obtained for system size  $L = 233$  in 1D,  $L = 144$  in 2D, and  $L = 21$  in 3D. Using the twist dispersion we can estimate the velocity by fitting the lowest energy band near 0 twist to a straight line. We compare the velocity as calculated from the twist dispersion with the KPM result of the DOS and fourth order perturbation theory in Supplementary Figure 6, which all agree well.

*1D powerlaw hopping model.* The parameter  $\sigma$  defined in equation (6) determines the behavior of the dispersion relation near  $k = 0$ . This can be seen directly from the twist dispersion in Supplementary Figure 5. We present detailed results for the 1D LRH model in Supplementary Figure 7. The DOS depends on  $\sigma$  by  $\rho(E) \sim |E|^{1/\sigma-1}$  (Ref. 7), which is demonstrated in Supplementary Figure 7. In the following we present detailed results for  $\sigma = 1/3$  and leave the full exploration of this 1D model for future work. Focusing on  $\sigma = 1/3$  is numerically advantageous since as we approach the transition the scaling  $\rho(E) \sim |W - W_c|^{-\beta}|E|^2$  allows us to use the second derivative of the DOS  $\rho''(0)$  to estimate  $\beta$  and we can compute  $\rho''(0)$  accurately using the KPM. Notice that the power-law remains constant when varying  $W$  in the semimetal phase, showing the 1D model is also stable to a weak quasiperiodic potential. Upon approaching the transition we find  $\rho''(0)$  displays a clear divergence with no sign of saturation as we increase the expansion order (see Supplementary Figure 7), similar to the 2D model we have discussed above. We find  $W_c = 2.05 \pm 0.03$  and from the power-law scaling  $\rho''(0) \sim |W - W_c|^{-\beta}$  we extract  $\beta = 4.0 \pm 0.8$  for  $\sigma = 1/3$ , see Supplementary Figure 4. Distinct from our results in 2D and 3D the transition in the 1D model is accompanied by real space localization. To demonstrate this we calculate the IPR in real space and momentum space at zero energy. The real space IPR becomes finite and momentum space IPR vanishes near the critical  $W_c$ . In addition, when the momentum space IPR goes to zero the DOS becomes non-zero demonstrating that the generation of DOS is tied with momentum space delocalization, similar to the higher dimensional models.



Supplementary Figure 2: The DOS  $\rho(E=0)$  and its derivative  $\rho'(E=0)$  for 2D perfect SOC with  $Q = 2\pi F_{n-2}/F_n$  at various  $L$  and  $N_c$  near the semimetal-to-metal transition  $W_c = 0.525 \pm 0.005$  and back to the reentrant semimetal  $W'_c = 0.551 \pm 0.005$ . Top and middle left:  $N_c = 4096$  varying  $L$ . Top and middle right:  $N_c = 16384$  varying  $L$ . The key is shared across the top four figures. Bottom:  $L = 233$ , varying  $N_c$ , with a shared key across the two. The insets are the same plots with linear scale.

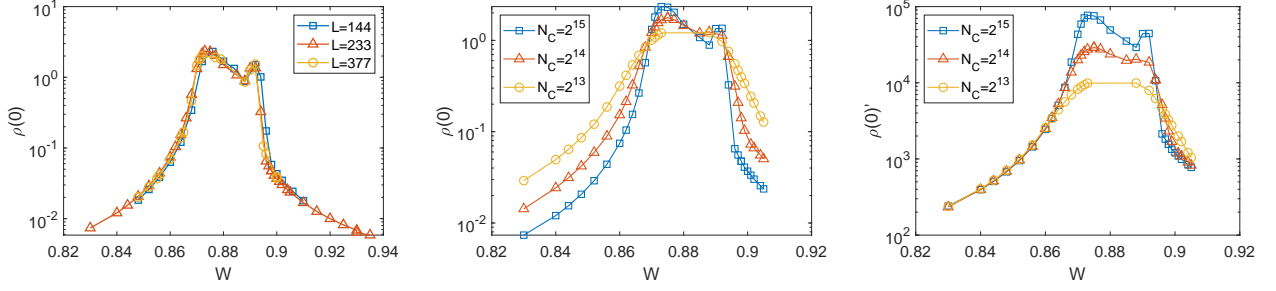
### Supplementary Note 3. ANALYTIC RESULTS

This supplementary note is devoted to the summary of details on analytical arguments presented in the main text.

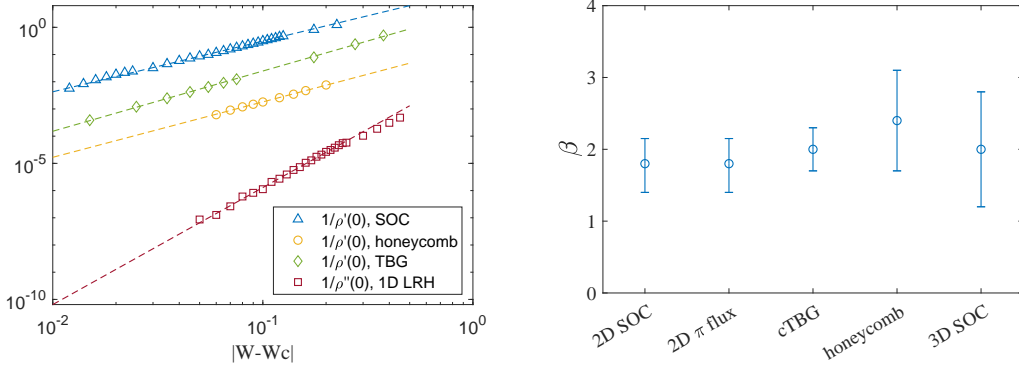
*Perturbative calculation of velocity renormalization.* We present the perturbative calculation of velocity renormalization using the language of retarded Green's functions,

$$\hat{G}_0(E) = [E + i\eta - \hat{T}]^{-1}, \quad \hat{G}(E) = [E + i\eta - \hat{T} - \hat{V}]^{-1}, \quad (18a)$$

and are interested in diagonal components  $G_{\mathbf{k},\mathbf{k}'}$  with  $\mathbf{k} = \mathbf{k}'$ , only (e.g. for the DOS we only need  $\rho(E) = -(1/\pi)\text{Im}\sum_{\mathbf{k}}\text{Tr}G_{\mathbf{k},\mathbf{k}}(E)$ ). We define the self energy at momentum  $\mathbf{k}$  by all diagrams which are  $G_0(\mathbf{k}, E)$  irreducible



Supplementary Figure 3: The DOS (left and middle)  $\rho(E=0)$  and its derivative (right)  $\rho'(E=0)$  for cTBG at  $k_\theta = 2\pi F_{n-5}/F_n$  for different system sizes (left) and for different KPM expansion orders for  $L = F_n = 233$  at various  $N_c$  (middle and right) near the semimetal-to-metal transition  $W_c = 0.8725 \pm 0.001$  and back to the reentrant semimetal  $W'_c = 0.892 \pm 0.002$ .



Supplementary Figure 4: Estimate of the scaling exponent  $\beta$ . Left: Extracting  $\beta$  from fitting (dashed lines)  $1/\rho'(0)$  in 2D and  $1/\rho''(0)$  in 1D for  $\sigma = 1/3$  versus  $|W - W_c|$  on a log-log plot displaying a clear power law. For clarity we have shifted the data of the honeycomb (HC) model so that it doesn't overlap with the cTBG model. Right:  $\beta$  estimate for the models we investigated in this paper, note that the estimate of  $\beta$  for the 3D Weyl model quotes the result from Ref.<sup>6</sup> and for the 1D long range hopping model we find  $\beta = 4 \pm 0.8$  for  $\sigma = 1/3$ .

and write

$$G_{\mathbf{k},\mathbf{k}}(E) = [G_0(\mathbf{k}, E)^{-1} - \Sigma(\mathbf{k}, E)]^{-1}. \quad (18b)$$

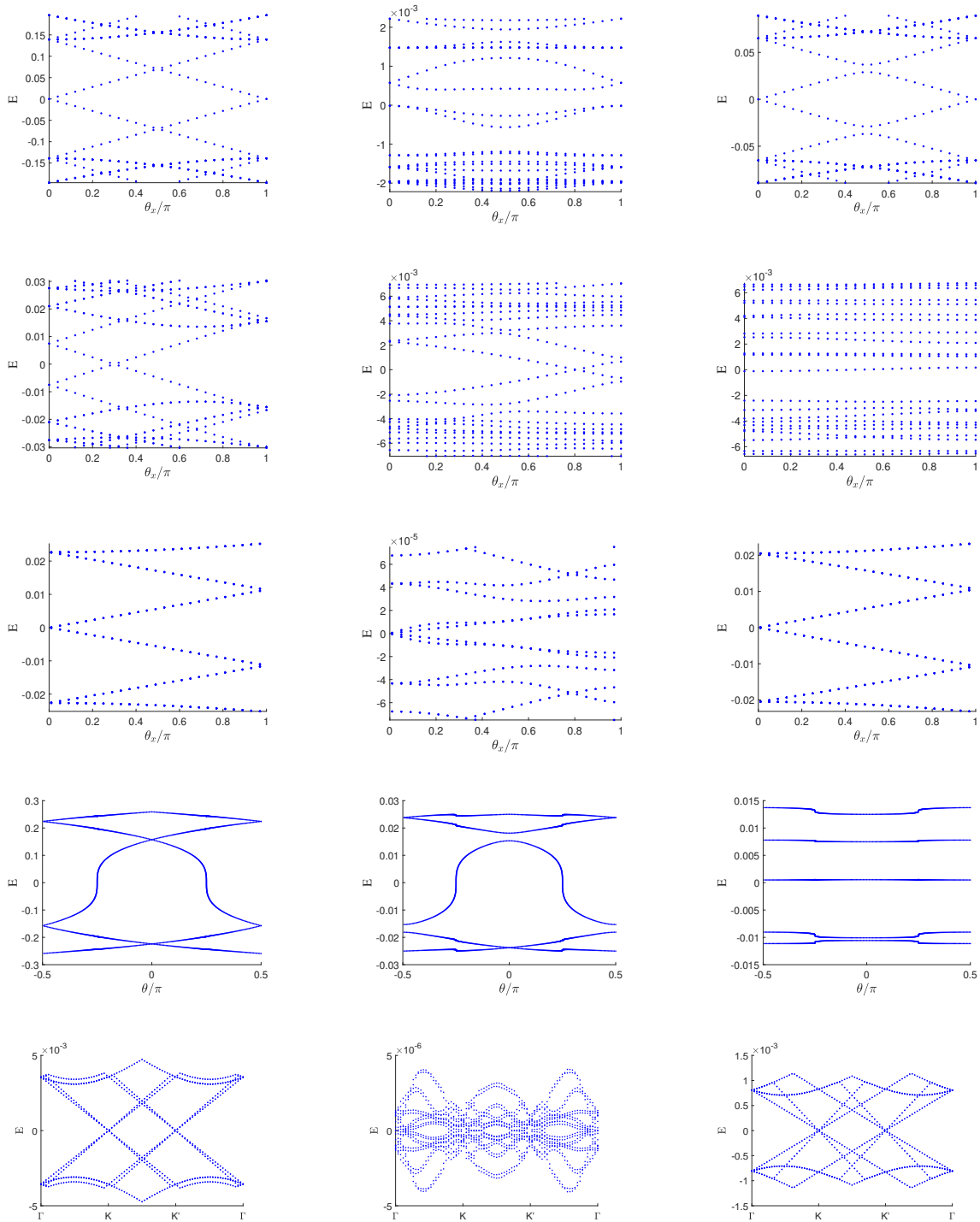
We expand about a given node  $\mathbf{K}_i$  of the dispersion  $T(\mathbf{K}_i + \mathbf{p}) \simeq T(\mathbf{K}_i) + h^{(\mathbf{K}_i)}(\mathbf{p})$  to leading order in  $p \ll 1/a$ . For models which satisfy the symmetry constraints exposed in the main text (see also Supplementary Note Supplementary Note 3)  $\Sigma(\mathbf{K}_i + \mathbf{p}, E) = E\Sigma_E + h(\mathbf{p})\Sigma_p$  to leading order in  $E, p$ . Henceforth, we choose the energy offset such that  $T(\mathbf{K}) = 0$ . Then,

$$G_{\mathbf{k},\mathbf{k}}(E) = Z[E - (v/v_0)h(\mathbf{k})]^{-1} \quad \text{with } Z^{-1} = 1 - \Sigma_E \text{ and } v/v_0 = (1 + \Sigma_p)Z. \quad (19)$$

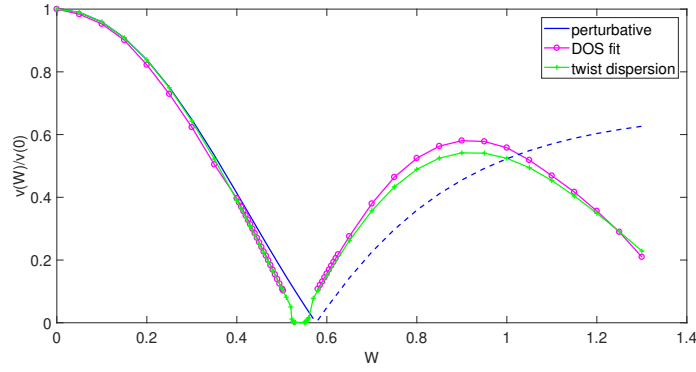
In this section we evaluate the self energy to leading and, for some models, next to leading order in powers of  $W$  and summarize them in Supplementary Table 1. A discussion of infinite order perturbation theory can be found at the end of this note.

To illustrate the procedure we analyze the model of 2D perfect SOC for which the states at small  $\mathbf{k}$  with Hamiltonian  $H(\mathbf{k}) = t(\sin(k_x)\sigma_x + \sin(k_y)\sigma_y) \simeq t\mathbf{k} \cdot \boldsymbol{\sigma}$  are connected to the states at  $\mathbf{k} \pm Q\hat{e}_{x,y}$  and therefore to leading order perturbation theory

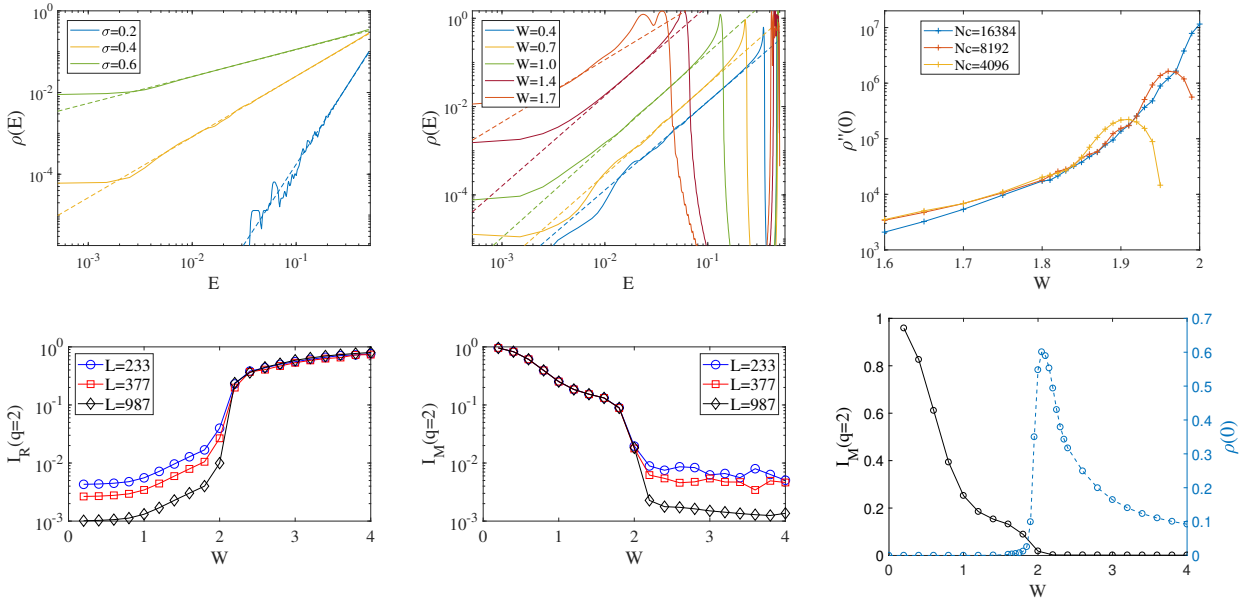
$$\Sigma^{(2)}(\mathbf{k}) = (W/2)^2 \sum_{\pm} \frac{1}{E - H(\mathbf{k} \pm Q\hat{e}_x)} + x \leftrightarrow y \simeq -E4\alpha^2 - t\mathbf{k} \cdot \boldsymbol{\sigma}(2\alpha^2(1 - \cos(Q))) \quad (20)$$



Supplementary Figure 5: Twist dispersion for various models displaying the characteristic magic angle feature of flat bands. All results shown except for the second and last rows (honeycomb model and cTBG respectively) are on the twist trajectory  $\theta_y = 0$ , and for 3D  $\theta_z = 0$ . Whereas for the honeycomb model and cTBG,  $\theta_y = -\theta_x$  (going through high symmetry points as displayed in the last row). First row: 3D SOC model investigated in Ref. 6 with  $Q = 2\pi F_{n-2}/L$ ,  $L = 21$ ,  $W = 0.1, 0.384, 0.5$  corresponding to before the transition, in the metallic phase and after the transition. Second row: 2D honeycomb model at  $Q = 2\pi F_{n-3}/L$ ,  $L = 144$ ,  $W = 0.8, 1.0, 1.08$  representing the states well before the transition, right before the transition, and a flattened band. Third row: 2D SOC model at  $Q = 2\pi F_{n-2}/L$ ,  $L = 144$ ,  $W = 0.35, 0.54, 0.8$  depicting states before the transition, in the metallic phase and after the transition. We have numerically checked that the twist dispersion of the  $\pi$ -flux model coincides with the 2D SOC, see the arguments exposed in Supplementary Figure 8, below. Fourth row: 1D power law hopping model at  $Q = 2\pi F_{n-3}/L$ ,  $\sigma = 1/3$ ,  $L = 1597$ ,  $W = 0.5, 1.7, 2.051$  of states well before the transition, the formation of a miniband, and a flattened band. Last row: 2D cTBG model at  $k_\theta = 4\pi/3 \times 21/L$ ,  $L = 233$ ,  $W = 0.8, 0.876, 0.9$  for states before the magic-angle, at the magic-angle, and after the magic-angle respectively.

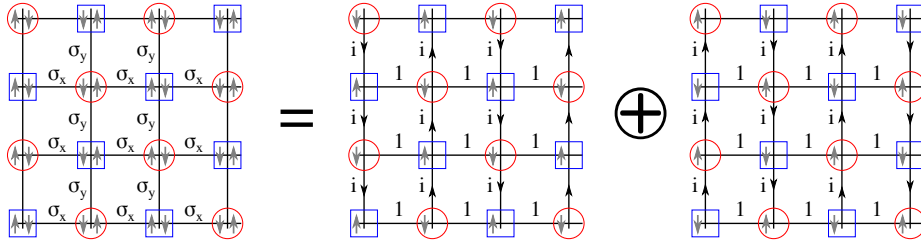


Supplementary Figure 6: Velocity for the 2D perfect SOC model at  $Q = 2\pi F_{n-2}/F_n$ , as obtained from fitting the DOS near  $E = 0$  (with  $L = 144$ ,  $N_c = 16384$  KPM result), fitting the twist dispersion near 0 twist (with  $L = 144$  exact diagonalization), and the perturbative calculation, Eq. (4) of the main text. The dashed line is the sign-reversed perturbative result for post-transition  $W$ .



Supplementary Figure 7: Results for the long-range hopping model with  $Q/2\pi = F_{n-3}/F_n$ . Top left: Density of states  $\rho(E)$  as a function of  $E$  on a log-log scale calculated using KPM at  $L = 1597$ , with an  $N_c = 4096$  with  $W = 0$ . By varying  $\sigma$  the power-law decreases following  $\rho(E) \sim |E|^{1/\sigma-1}$ . Top middle: Fixed  $\sigma = 1/3$  and varying  $W$ , when  $W$  is small, the power law is well preserved. As the model approaches the transition the power law regime is pushed to lower and lower energy due to the formation of minibands. The solid line represent actual data, and dashed lines are fits to the expected form  $\rho(E) \sim |E|^{1/\sigma-1} = E^2$ . To accurately extract the scaling of the prefactor of the DOS, i.e.  $\rho(E) \sim |W - W_c|^{-\beta} |E|^{1/\sigma-1}$ , we turn to the second derivative of  $\rho(E)$ . Top right: The second derivative  $\rho''(0)$  as a function of  $W$  for various  $N_C$  with  $L = 28657$  displaying a clear divergence as the transition is approached, signifying the DOS is becoming non-analytic. As shown in Supplementary Figure 4, we find  $\rho''(0)$  diverges with  $|W - W_c|$  in a power law fashion. To avoid the issues involved with fitting, that we find are most severe in this 1D model we use  $\rho''(0)$  computed directly with the KPM. Bottom left and middle: Real space and momentum space inverse participation ratio (IPR) of energy eigenstate closest to  $E = 0$ , denoted as  $\mathcal{I}_R(q=2)$  and  $\mathcal{I}_M(q=2)$  respectively, calculated for various  $L$  and averaged over 200 realizations. The  $L$  dependence clearly shows that when the real space IPR is delocalized  $\mathcal{I}_R(q=2) \sim 1/L$  the momentum space IPR is localized and vice-versa. This demonstrates the transition in 1D goes from a semimetal to an Anderson insulator. Bottom right: The momentum space IPR  $\mathcal{I}_M(q=2)$  (left vertical axis) and the zero energy DOS  $\rho(0)$  (right vertical axis) for  $L = 987$  as a function of  $W$ , which demonstrates that the zero energy DOS becomes non-zero when the wavefunctions delocalize in momentum space.





Supplementary Figure 8: Graphic demonstration that the model of perfect SOC in 2D is a direct sum of two decoupled  $\pi$  flux models. The model of perfect SOC, on the left of the equality sign, is characterized by direction dependent hopping matrices. Using blue squares and red circles to depict the bipartition, hopping only connects  $|\square, \uparrow\rangle$  with  $|\circ, \downarrow\rangle$ , and *separately*  $|\square, \downarrow\rangle$  with  $|\circ, \uparrow\rangle$ . The hopping in  $y$ -direction is imaginary and directed (this results from the asymmetry of  $\sigma_y$ ) and, in conclusion, leads to the inclusion of a flux  $\pi$  per plaquette. The onsite potential does not violate the described block-diagonalization.

model	$Z^{-1}$	$v/v_0$
2D SOC	$1 + \left[ \frac{W}{t \sin(Q)} \right]^2$	$Z \left( 1 - \frac{1 - \cos(Q)}{2} \left[ \frac{W}{t \sin(Q)} \right]^2 \right)$
3D SOC	$1 + \frac{3}{2} \left[ \frac{W}{t \sin(Q)} \right]^2$	$Z \left( 1 - \frac{2 - \cos(Q)}{2} \left[ \frac{W}{t \sin(Q)} \right]^2 \right)$
1D LRH	$1 + 2 \left[ \frac{W}{t  \sin(Q) ^\sigma} \right]^2$	$Z$
2D cTBG	$1 + \frac{W^2}{t^2} \frac{[\cos(\frac{3k_\theta}{2}) + 6 + 2 \cos(\frac{3k_\theta}{4})]}{12 \sin^2(\frac{3k_\theta}{4})}$	$Z \left( 1 - \frac{W^2}{t^2 \sin^2(\frac{3k_\theta}{4})} \frac{[2 \cos(\frac{3k_\theta}{2}) - 16 \cos(\frac{3k_\theta}{4}) + 4 \cos(\frac{9k_\theta}{4}) + 5 \cos(3k_\theta) - 49]}{72} \right)$
HC, $\delta_\mu = \mathbf{d}_\mu$	$1 + \frac{W^2}{t^2} \frac{3}{8 \sin^2(3Q/4)}$	$Z \left( 1 - \frac{W^2}{t^2} \frac{1/4}{1 + 2 \cos(Q/2)} \right)$
HC, $\delta_\mu = O_{\pi/2} \mathbf{d}_\mu$	$1 + \frac{W^2}{t^2} \frac{3[2 + \cos(\sqrt{3}Q/2)]}{8 \sin^2(3\sqrt{3}Q/4)}$	$Z \left( 1 + \frac{W^2}{t^2} \frac{5 + 4 \cos(\sqrt{3}Q/2)}{4[1 + \cos(\sqrt{3}Q/2)]^2} \right)$
HC, $\delta_\mu = \mathbf{d}_{1,2}$	$1 + \frac{W^2}{t^2} \frac{1}{4 \sin^2(3Q/4)}$	$Z \left( 1 - \frac{W^2}{t^2} \frac{\begin{pmatrix} 3[\cos(Q/2) - 2 \cos(Q)] & -\sqrt{3}[\cos(Q/2) + 2 \cos(Q)] \\ -\sqrt{3}[\cos(Q/2) + 2 \cos(Q)] & 5 \cos(Q/2) - 2 \cos(Q) \end{pmatrix}}{48 \sin^2(3Q/4)^2} \right)$

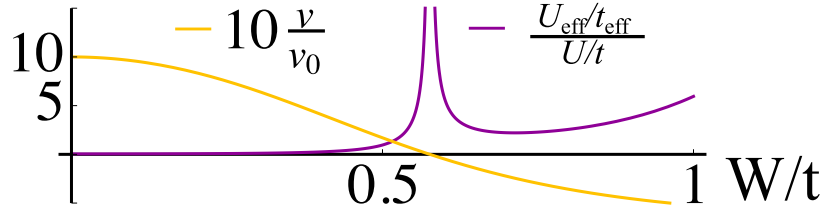
Supplementary Table 1: Perturbative corrections to quasiparticle weight  $Z$  and velocity  $v/v_0$  for a variety of magic angle models. Note that for the honeycomb model and  $\delta_\mu = \mathbf{d}_{1,2}$ , the symmetry protection of nodes is lost. It implies a relocation of  $K$ -point node  $\delta \mathbf{k} = W^2(1, -\sqrt{3})^T / [12t^2(1 + 2 \cos(Q/2))]$  and a distorted velocity matrix. For cTBG, the momentum dependent self-energy has the form  $\Sigma_{\mathbf{p}} h(p_x, p_y) + \tilde{\Sigma}_{\mathbf{p}} h(p_y, -p_x)$  and we extract  $v/v_0 = Z(1 + \Sigma_{\mathbf{p}})$ .

For the next to leading order, all states at Manhattan distance  $2Q$  from the origin are integrated out and we obtain

$$\begin{aligned} \Sigma^{(4)}(\mathbf{k}) \simeq & -\frac{E}{16} \left( \frac{W}{t} \right)^4 (4 \cos(Q) + 10 \cos(2Q) + 11) \csc^4(Q) \sec^2(Q) \\ & + \left( \frac{W}{t} \right)^4 \frac{t \mathbf{k} \cdot \boldsymbol{\sigma}}{16} (4 - 5 \cos(Q) + 6 \cos(2Q)) \csc(Q)^4 \sec(Q) \end{aligned} \quad (21)$$

This is the origin of Eq. (4) of the main text. It turns out that the results obtained for the 2D model of perfect SOC directly apply to the  $\pi$  flux model. This is best graphically shown, see Supplementary Figure 8: the model of 2D perfect SOC is a direct sum of two  $\pi$ -flux models which in the absence of interactions completely decouple. By consequence, all single particle results obtained for model of 2D perfect SOC also hold for the  $\pi$ -flux model.

*Renormalization of interactions.* We present an analytical estimate of the renormalization of the interaction upon



Supplementary Figure 9: Divergence of contact interaction according to Eq. (27) for the model of 2D SOC. Here, the fourth order perturbative self energy was employed and we used  $\gamma = 1/5$ .

projection onto certain minibands and approaching the transition for the model of 2D SOC. Let the bare ( $W = 0$ ) model in the continuum be written as ( $\mathbf{K}_i$  are various Dirac/Weyl nodes, with linear  $\mathbf{k} \cdot \mathbf{p}$  Hamiltonian  $h^{(\mathbf{K}_i)}(\mathbf{p})$ )

$$S = \sum_{\mathbf{K}_i} \int^{1/a} (dp) \int d\tau \bar{c}^{(\mathbf{K}_i)}(\mathbf{p}) [\partial_\tau + h^{(\mathbf{K}_i)}(\mathbf{p})] c^{(\mathbf{K}_i)}(\mathbf{p})$$

$$+ \sum_{\mathbf{K}_{1,2,3,4}} \int^{1/a} \prod_{j=1}^4 (dp_j) \int d\tau (2\pi)^d g_{\{\mathbf{K}_i\}} \delta_{\mathbf{K}_1 + \mathbf{K}_3 - \mathbf{K}_2 - \mathbf{K}_4} \delta(\mathbf{p}_1 + \mathbf{p}_3 - \mathbf{p}_2 - \mathbf{p}_4) \frac{[\bar{c}^{(\mathbf{K}_1)}(\mathbf{p}_1) c^{(\mathbf{K}_2)}(\mathbf{p}_2)] [\bar{c}^{(\mathbf{K}_3)}(\mathbf{p}_3) c^{(\mathbf{K}_4)}(\mathbf{p}_4)]}{|\mathbf{p}_1 - \mathbf{p}_2|^\sigma}. \quad (22)$$

The spectrum of  $h^{(\mathbf{K}_i)}(\mathbf{p})$  has the form  $v_0 |\mathbf{p}|$  with bare value  $v_0 \sim ta$  and, for contact interaction ( $\sigma = 0$ ),  $g_{\{\mathbf{K}_i\}} \sim Ua^d$ , while for Coulomb interaction ( $\sigma = d - 1$ )  $g_{\{\mathbf{K}_i\}} \propto \delta_{\mathbf{K}_1, \mathbf{K}_2}$ . Perturbation theory indicates a dimensionless parameter

$$\frac{g_{\{\mathbf{K}_i\}}}{v_0/a^{d-1}} \sim \frac{U}{t} \text{ for } \sigma = 0 \text{ (contact interaction),} \quad (23)$$

$$\alpha = \frac{g_{\{\mathbf{K}_i\}}}{v_0} \text{ for } \sigma = d - 1 \text{ (Coulomb interaction).} \quad (24)$$

We now consider the effect of integrating out high energy states and projecting onto a miniband with effective Brillouin zone size  $1/a'$ . This leads to

$$S = \sum_{\mathbf{K}_i} \int^{1/a'} (dp') \int d\tau \bar{c}_{<}^{(\mathbf{K}_i)}(\mathbf{p}') Z^{-1} [\partial_\tau + \frac{v}{v_0} h^{(\mathbf{K}_i)}(\mathbf{p}')] c_{<}^{(\mathbf{K}_i)}(\mathbf{p}')$$

$$+ \sum_{\mathbf{K}_{1,2,3,4}} \int^{1/a'} \prod_{j=1}^4 (dp'_j) \int d\tau (2\pi)^d g_{\{\mathbf{K}_i\}} \delta_{\mathbf{K}_1 + \mathbf{K}_3 - \mathbf{K}_2 - \mathbf{K}_4} \delta(\mathbf{p}'_1 + \mathbf{p}'_3 - \mathbf{p}'_2 - \mathbf{p}'_4) \frac{[\bar{c}_{<}^{(\mathbf{K}_1)}(\mathbf{p}'_1) c_{<}^{(\mathbf{K}_2)}(\mathbf{p}'_2)] [\bar{c}_{<}^{(\mathbf{K}_3)}(\mathbf{p}'_3) c_{<}^{(\mathbf{K}_4)}(\mathbf{p}'_4)]}{|\mathbf{p}'_1 - \mathbf{p}'_2|^\sigma}. \quad (25)$$

The renormalizations  $Z$  and  $v/v_0$  originate from scalar and matrix components of the self-energy and were calculated perturbatively above. We now first rescale  $p' = \frac{a}{a'} p$  with  $p \in (0, 1/a)$  and then define  $c_{<}^{(\mathbf{K})}(a\mathbf{p}/a') Z^{-1/2} (a/a')^{d/2} = c^{(\mathbf{K})}(\mathbf{p})$ . Under this rescaling, we restore the form of Eq. (22), including its UV cut-off  $1/a$ , but obtain the rescaling  $v_0 \rightarrow va/a'$ ,  $g \rightarrow g(a/a')^{d-\sigma} Z^2$ . From this we obtain the final formula for renormalization of the dimensionless coupling constant

$$\frac{U_{\text{eff}}/t_{\text{eff}}}{U/t} = \gamma \left(\frac{a}{a'}\right)^{d-1} \frac{Z^2}{v/v_0} \quad (\text{contact interaction}), \quad (26)$$

$$\frac{\alpha_{\text{ren.}}}{\alpha_{\text{bare}}} = \frac{Z^2}{v/v_0} \quad (\text{Coulomb}). \quad (27)$$

Here,  $\gamma$  is an unknown constant of order unity which depends on details of the cut-off of the linearized theory. Importantly, the integration reduces the bare contact interaction by a factor  $(a/a')^{d-1}$ , except in the closest vicinity of the magic angle where the vanishing velocity overtakes the reduction, see Supplementary Figure 9.

*Relationship to number theory.* In addition, we show the relationship of the sequence of relevant perturbative processes with certain well known sequences from number theory. Starting from the scattering process of order  $l_1 = 1$  we want to determine the sequence  $\{l_n\}_{n=1}^\infty$  for which the  $l_n$ th order momentum transfer carves out smaller minibands than the  $l_{n-1}$ th order. In formulae, this implies for the 2D model of perfect SOC of the main text

and arbitrary incommensuration wavevector  $Q$  the condition  $\sin^2(l_n Q) < \sin^2(l_{n-1} Q)$ . We now concentrate on the specific case  $Q = 2\pi/\phi^2 = \pi(3 - \sqrt{5})$ . For this situation, the defining condition on the sequence of  $l_n$  is  $\sin^2(\pi l_n \sqrt{5}) < \sin^2(\pi l_{n-1} \sqrt{5})$ . The sequence  $\{l_n\}_{n=1}^{\infty}$  for which  $l_n \sqrt{5}$  successively approaches integers is the sequence of denominators of the leading rational approximants, i.e. the sequence of denominators of continued fraction convergents of  $\sqrt{5}$  (OEIS ID A001076). This sequence is also half the value of the even fibonaccis  $l_n = F_{3n}/2$ .

This sequence also connects to the formation of minibands as found with the finite size  $Q_n = 2\pi F_{n-2}/F_n$ . Intuitively, when  $F_n$  is even ( $n$  is a multiple of 3), then at order  $F_n/2$ , the Dirac nodes gap out, but then for  $F_{n+1}$  and  $F_{n+2}$  this perturbative gap must have moved to small but finite energy, forming the miniband. This motivates using  $Q_{3n+1}$  and  $Q_{3n+2}$  to study the effective model of successive minibands, as we do in Supplementary Note Supplementary Note 6.

Therefore, we offer a proof that connects this sequence to the perturbative minibands which requires the following facts about Fibonaccis

- $F_{3m}$  is even while  $F_{3m+1}$  and  $F_{3m+2}$  is odd.
- Catalan's Identity:  $(-1)^{n-r} F_r^2 = F_n^2 - F_{n+r} F_{n-r}$ .
- $(-1)^n = F_n F_{n-1} - F_{n-2} F_{n+1}$ .

To determine the order of perturbation theory where a gap is opened, we need to find an integer  $g_n$  such that  $g_n Q_n = 2\pi \times q/F_n$  such that  $q$  is an integer closest to  $F_n/2$  modulo  $F_n$ . This is accomplished by the following theorem

**Theorem.** *Let  $n = 3m + r$  for  $r = 0, 1, 2$ , then the integer  $g_n = F_{3m}/2$  is the smallest integer such that  $g_n F_{n-2} \equiv \frac{F_n + \delta_n}{2} \pmod{F_n}$  with integer  $|\delta_{3m}| \leq 1$ . In particular,  $\delta_{3m} = 0$ ,  $\delta_{3m+1} = (-1)^m$ , and  $\delta_{3m+2} = (-1)^{m+1}$ .*

*Proof.* We break this up into cases. Since  $F_{n-2}$  does not divide  $F_n$ , we merely need to find a  $g_n$  that satisfies the relevant cases in order to find the unique  $g_n$ , as long as  $g_n < F_n$ .

**Case 1:**  $r = 0$ . For this case we can prove the above by noting that if  $g_{3m} = F_{3m}/2$ , and  $F_{n-2}$  is necessarily odd, then (We use  $N$  to represent an arbitrary, unimportant, integer.)

$$g_n F_{n-2} = \frac{1}{2} F_{3m} (2N + 1) \equiv \frac{1}{2} F_{3m} \pmod{F_n}, \quad (28)$$

with  $\delta_{3m} = 0$ .

The value  $g_n$  is smallest since  $\delta_n$  must vary by 2 in order for the equation  $\frac{F_n + \delta_n}{2}$  to remain integer valued, and the condition  $|\delta_n| < 1$  prevents that.

**Case 2:**  $r = 1$ . For this case we observe using the previous fact about Fibonaccis that

$$F_{3m} F_{3m-1} = F_{3m+1} F_{3m-2} + (-1)^{3m}, \quad (29)$$

and therefore since  $F_{3m-2}$  is odd

$$\frac{1}{2} F_{3m} F_{3m-1} = (N + \frac{1}{2}) F_{3m+1} + \frac{1}{2} (-1)^{3m} \equiv \frac{F_{3m+1} + (-1)^{3m}}{2} \pmod{F_n}, \quad (30)$$

and therefore  $g_{3m+1} = F_{3m}/2$  with  $\delta_{3m+1} = (-1)^m$ .

In order to show that  $g_n$  is smallest, only  $\delta_{3m+1} = -(-1)^m$  could be a problem, but we notice that  $\delta_{3m+1} = -(-1)^m$  is satisfied (uniquely) for  $g_n = F_{3m+1} - F_{3m}/2 > F_{3m}/2$ .

**Case 3:**  $r = 2$ . For this case we take Catalan's identity with  $r = 2$  to get

$$F_{3m}^2 = F_{3m+2} F_{3m-2} - (-1)^{3m}, \quad (31)$$

and therefore since  $F_{3m-2}$  is odd

$$\frac{1}{2} F_{3m} F_{3m} = (N + \frac{1}{2}) F_{3m+2} - \frac{1}{2} (-1)^{3m} \equiv \frac{F_{3m+2} - (-1)^{3m}}{2} \pmod{F_n}, \quad (32)$$

so again  $g_{3m+2} = F_{3m}/2$  with  $\delta_{3m+2} = (-1)^{m+1}$ .

In order to show that  $g_n$  is smallest, only  $\delta_{3m+1} = (-1)^m$  could be a problem, but we notice that  $\delta_{3m+1} = (-1)^m$  is satisfied (uniquely) for  $g_n = F_{3m+2} - F_{3m}/2 > F_{3m}/2$ .  $\square$

Therefore, the order of perturbation theory that opens up a gap nearest to  $E = 0$  for  $Q = 2\pi F_{3m+r-2}/F_{3m+r}$  for  $r = 0, 1, 2$  is  $F_{3m}/2$ .

*Generality of the magic-angle phenomenon - symmetry protection.* In this part of the supplementary note we discuss the generality of our findings by highlighting the general condition for the appearance of the magic angle phenomenon, namely the stability of the semimetal at weak coupling.

We concentrate on nodes in the kinetic term  $\hat{T}$  which are protected by a symmetry group  $G_T$ . For example, this analysis applies to each model we have considered in 2D as well as Dirac semimetals in 3D. Note that in general  $G_T$  is a subgroup of all symmetry operations of the kinetic term. Let  $U_{S_T}$  be the representation of  $S_T \in G_T$  in the (e.g. spinorial) Hilbert space, then the symmetry of the Hamiltonian implies  $T(\mathbf{k}) = U_{S_T}^\dagger T(S_T \mathbf{k}) U_{S_T}$ . We concentrate on high symmetry points where  $S_T \mathbf{K} = \mathbf{K}$ ,  $\forall S_T \in G_T$ . Then, a non-trivial representation implies degeneracy in view of  $[T(\mathbf{K}), U_{S_T}] = 0 \forall S_T \in G_T$  (formally, two non-commuting  $U_{S_T}$  are needed). We further assume a group  $G_V$  of spatial (point group) symmetries of the quasiperiodic background  $\hat{V}$ , such that

$$V(\mathbf{x}) = \sum_{S_V \in G_V} \tilde{U}_{S_V} W \tilde{U}_{S_V}^\dagger e^{iQ\mathbf{x} \cdot S_V \hat{e}_0} + h.c.. \quad (33)$$

Here now,  $\tilde{U}_{S_V}$  is the representation of  $S_V \in G_V$  and  $\hat{e}_0$  is an arbitrary vector in  $\mathbb{R}^d$ .

In this section, we consider Eq. (18) formally to all orders in perturbation theory. The semimetallic behavior persists if a)  $\Sigma(\mathbf{k})$  is hermitian and b)  $T(\mathbf{k}) + \Sigma(\mathbf{k})$  has the same symmetry protected touching point as  $T(\mathbf{k})$ , i.e. if  $\Sigma(\mathbf{k})$  respects the symmetries ensuring the semimetal. In view of the incommensuration, perfect resonance is formally absent to any order in perturbation theory and therefore, the decay rate  $1/\tau \sim \sum_{\mathbf{k}'} |\mathbf{T}_{\mathbf{k}, \mathbf{k}'}|^2 \delta(E_{\mathbf{k}} - E_{\mathbf{k}'})$  (more generally: the anti-hermitian part of the self-energy) vanishes ( $\mathbf{T}_{\mathbf{k}, \mathbf{k}'}$  denotes the T-matrix). Thus a) is fulfilled and  $1/\tau \neq 0$  signals the breakdown of perturbation theory (spontaneous unitarity breaking). We can then show to all orders in perturbation theory that the semimetal is stable provided  $G_T$  is a subgroup of  $G_V$ .

We proceed to the proof of  $\Sigma(\mathbf{k}) = U_{S_T}^\dagger \Sigma(S_T \mathbf{k}) U_{S_T}$  under the outlined assumptions. To get a feeling, we first consider second order perturbation theory.

$$\Sigma^{(2)}(\mathbf{k}) = \sum_{S_V \in G_V} \tilde{U}_{S_V} W \tilde{U}_{S_V}^\dagger [E^+ - T(\mathbf{k} + Q S_V \hat{e}_0)]^{-1} \tilde{U}_{S_V} W^\dagger \tilde{U}_{S_V}^\dagger. \quad (34)$$

We compare to

$$\begin{aligned} U_{S_T}^\dagger \Sigma^{(2)}(S_T \mathbf{k}) U_{S_T} &= \sum_{S_V \in G_V} U_{S_T}^\dagger \tilde{U}_{S_V} W \tilde{U}_{S_V}^\dagger U_{S_T} [E^+ - U_{S_T}^\dagger T(S_T \mathbf{k} + Q S_V \hat{e}_0) U_{S_T}]^{-1} U_{S_T}^\dagger \tilde{U}_{S_V} W \tilde{U}_{S_V}^\dagger U_{S_T} \\ &= \sum_{S_V \in G_V} U_{S_T}^\dagger \tilde{U}_{S_V} W \tilde{U}_{S_V}^\dagger U_{S_T} [E^+ - T(\mathbf{k} + Q S_T^{-1} S_V \hat{e}_0)]^{-1} U_{S_T}^\dagger \tilde{U}_{S_V} W^\dagger \tilde{U}_{S_V}^\dagger U_{S_T} \end{aligned} \quad (35)$$

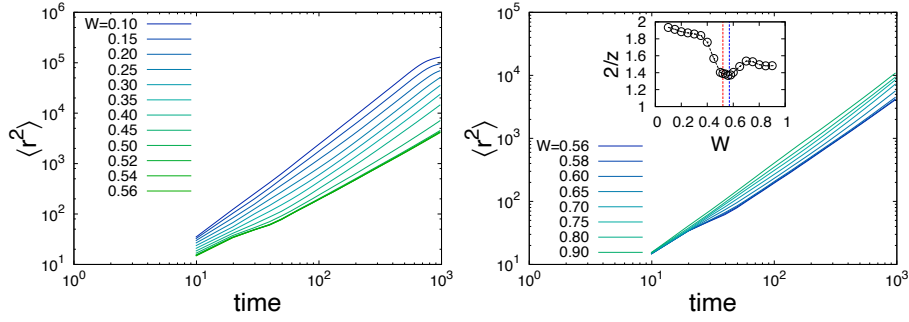
This expression is invariant provided the action of  $S_T$  onto  $G_V$  is a bijection of  $G_V$  onto itself  $\forall S_T \in G_T$ , i.e.  $S_T S_V \in G_V \forall S_V \in G_V$  and  $S_T G_V = G_V$  as this allows to uniquely relate the summation index. Taking  $S_V = \mathbf{1}$  implies that  $S_T \in G_V$  and hence  $G_T$  is a subgroup of  $G_V$ . By consequence, the representation in the Hilbert space fulfills  $\tilde{U}_{S_T^{-1} S_V} = U_{S_T}^\dagger \tilde{U}_{S_V}$  and  $\Sigma^{(2)}(\mathbf{k})$  is invariant under the symmetries protecting the semimetal. We now continue with the next order  $\Sigma^{(4)}$ , from there the generality of the statement becomes apparent,

$$\begin{aligned} \Sigma^{(4)}(\mathbf{k}) &= \sum_{S_V \in G_V} \sum_{\substack{S'_V \in G_V \\ S'_V \hat{e}_0 \neq -S_V \hat{e}_0}} \tilde{U}_{S_V} W \tilde{U}_{S'_V}^\dagger [E^+ - T(\mathbf{k} + Q S_V \hat{e}_0)]^{-1} \\ &\quad \tilde{U}_{S'_V} W \tilde{U}_{S'_V}^\dagger [E^+ - T(\mathbf{k} + Q S_V \hat{e}_0 + Q S'_V \hat{e}_0)]^{-1} \\ &\quad \tilde{U}_{S'_V} W^\dagger \tilde{U}_{S'_V}^\dagger [E^+ - T(\mathbf{k} + Q S_V \hat{e}_0)]^{-1} \tilde{U}_{S_V} W^\dagger \tilde{U}_{S_V}^\dagger. \end{aligned} \quad (36)$$

The exclusion  $S'_V \hat{e}_0 + S_V \hat{e}_0 \neq 0$  ensures the irreducibility with respect to  $G_{\mathbf{k}}$ . Again we can apply an  $S_T$  transformation and exploit the two conditions exposed above to relabel both  $S_V$  and  $S'_V$ . This implies the invariance of  $\Sigma_{(4)}$ . This procedure can be used to arbitrary order in perturbation theory.

#### Supplementary Note 4. WAVEPACKET DYNAMICS

*Non-interacting Wavepacket Expansion as a Probe of the Single Particle Quantum Phase Transition.* Firstly we demonstrate that wavepacket dynamics can be used as a tool to observe the single particle quantum phase transition in the models that possess a reentrant semimetal in two dimensions. Note that in three dimensions due to the diffusive



Supplementary Figure 10: Wave packet dynamics displaying the spread of the wavepacket  $\langle r(t)^2 \rangle$  as a function of time computed from the KPM on a system size of  $L = 987$  with a KPM expansion order  $N_C = 2^{12}$ . (Left) Wavepacket spreading for  $W$  in the semimetal and passing into the metallic phase at  $W \approx 0.54$ . (Right) Similar results but for  $W$  passing from the metallic phase to the reentrant semimetal at  $W \approx 0.57$ . The data clearly shows a non-monotonic wavepacket spreading for increasing quasiperiodic potential strength which gives rise to a non-monotonic behavior in the dynamic exponent  $z$  shown in the inset, which we extract via fits to the long time dynamics (the dashed lines mark the entry into the metallic phase).

states at finite energy that dominate the dynamics this probe is not useful<sup>6</sup>. We initialize a wavepacket to the origin  $|\psi_0\rangle$  and use the KPM to time evolve the state to obtain  $|\psi(t)\rangle = e^{-iHt}|\psi_0\rangle$  and from this we compute the spread of the wavepacket as a function of time from  $\langle r(t)^2 \rangle \equiv \langle \psi(t)|r^2|\psi(t)\rangle$ . From the long time dynamics we extract the dynamical exponent from the scaling  $\langle r(t)^2 \rangle \sim t^{2/z}$ . We focus on the perfect SOC Hamiltonian from the main text. The KPM expands the time evolution operator in terms of Chebyshev polynomials up to an order  $N_C$ , which dictates the final time that can be reached in the numerical calculations. Here we focus on a large linear system size of  $L = 987$  and use a KPM expansion order  $N_C = 2^{12}$ , which allows us to time evolve the wavepacket until it spreads out across the sample. As shown in Supplementary Figure 10 we find rather unusual wavepacket dynamics which is a signature of a sequence of semimetal-metal-semimetal transitions. As a function of increasing  $W$  we find that the speed at which the wavepacket spreads out monotonically decreases until we reach the metallic phase where the dependence on  $W$  is rather weak. Then, upon reentering the semimetal phase at larger  $W$  the wavepacket spreading speeds back up. This is clearly demonstrated in the dynamical exponent  $z$  showing non-monotonic behavior in the inset. Interestingly, our estimate of  $z$  is not diffusive, consistent with the marginal nature of 2D. We expect that this wavepacket signature can be used to detect the transition in cold atom experiments.

*Wavepacket expansion in the hydrodynamic regime of strong interactions.* Now we turn to the dynamics of a cold atomic cloud in magic-angle emulators. We concentrate on the model of perfect SOC in 2D, and incorporate velocity and wave function renormalization perturbatively, as exposed in Supplementary Note Supplementary Note 3. Complementary to the previous section, we here consider strong interactions leading to fast equilibration, justifying an effective hydrodynamic approach.

We describe the dynamics of the system using the Boltzmann equation for the distribution function  $f_n$

$$\dot{f}_n + \dot{\mathbf{r}}_n \nabla_{\mathbf{r}} f_n + \dot{\mathbf{p}} \nabla_{\mathbf{p}} f_n = \text{St}_n[f]. \quad (37)$$

Here, we introduced the multi index  $n = (\mathbf{K}, \mathbf{p}, \xi)$  to label the states (node  $\mathbf{K}$ , momentum relative to the node  $\mathbf{p}$ , band  $\xi = \pm 1$ ). Semiclassical fermions have energy  $\epsilon_n = \xi[v|\mathbf{p}| - v_3|\mathbf{p}|^3 - v_{C_4}(p_x^4 + p_y^4)/|\mathbf{p}|]$  and velocity  $\dot{\mathbf{r}}_n = \partial\epsilon_n/\partial\mathbf{p}$ . The momentum satisfies  $\dot{\mathbf{p}}_n = \mathbf{F}(\mathbf{r})$  for force  $\mathbf{F}$ . The collision kernel (outgoing states are denoted by a bar) is

$$\text{St}_n[f] = - \sum_{m, \bar{n}, \bar{m}} \{W_{nm \rightarrow \bar{n}\bar{m}} f_n f_m (1 - f_{\bar{n}})(1 - f_{\bar{m}}) - W_{\bar{n}\bar{m} \rightarrow nm} f_{\bar{n}} f_{\bar{m}} (1 - f_n)(1 - f_m)\}. \quad (38)$$

The Boltzmann treatment is valid when either thermal length or Fermi wave length exceed the mean free path. We determine the transition probability in the Fermi golden rule approximation (see e.g. Ref. 8)

$$W_{nm \rightarrow \bar{n}\bar{m}} = 2\pi U^2 |\langle n|\bar{n}\rangle|^2 |\langle m|\bar{m}\rangle|^2 (2\pi)^3 \delta(\epsilon_n + \epsilon_m - \epsilon_{\bar{n}} - \epsilon_{\bar{m}}) \delta(\mathbf{p}_n + \mathbf{p}_m - \mathbf{p}_{\bar{n}} - \mathbf{p}_{\bar{m}}) \tilde{\delta}_{\mathbf{K}_n + \mathbf{K}_m, \mathbf{K}_{\bar{n}} + \mathbf{K}_{\bar{m}}}. \quad (39)$$

The tilde on the Kronecker delta denotes equality modulo  $2\pi$ . The collision integral conserves particle number and energy  $\sum_n \text{St}_n[f] = 0 = \sum_n \epsilon_n \text{St}_n[f]$ . While it does not conserve the total crystalline momentum  $\mathbf{p}^{\text{tot}} = (\mathbf{K} + \mathbf{p})$  due to umklapp scattering, the momenta  $\mathbf{p}$  relative to the nodes are conserved  $\sum_n \mathbf{p}_n \text{St}_n[f] = 0$ . Therefore, number density  $\rho_N = \sum_n f_n$  and energy density  $\rho_E = \sum_n \epsilon_n f_n$  are conserved and the energy current  $\mathbf{j}_\epsilon = \sum_n \epsilon_n \dot{\mathbf{r}}_n f_n$

( $\epsilon_n \dot{\mathbf{r}}_n = v^2 \mathbf{p}_n + \mathcal{O}(v_3, v_{C_4})$ ) decays slowly due to nonlinearities in the spectrum. We estimate its decay rate as  $1/\tau = \Gamma_0 f(\alpha) Z(\alpha)^2 / \bar{v}(\alpha)^2$  where  $\Gamma_0 = U_0^2/t$  and  $f(\alpha)$  is a dimensionless regular function of  $\alpha$  which is proportional to  $v_3, v_{C_4}$ .

Following a formalism of hydrodynamics in graphene<sup>9</sup> we introduce smoothly varying Lagrange multipliers (chemical potential, temperature, hydrodynamic velocity) for each of the conserved quantities. The Fermi energy (measured from the Dirac node) is assumed to be small as compared to the mini-bandwidth. In particular the hydrodynamic velocity  $\mathbf{u}$  enters as  $\mathbf{j} = \mathbf{u} \rho_N$ , and for  $\mu \ll T$ ,  $\rho_N \sim \mu T/v^2 + \mathcal{O}(u^2/v^2)$ ,  $\rho_E \sim T^3/v^2 + \mathcal{O}(u^2/v^2)$ . The hydrodynamic equations reduce to

$$\dot{\rho}_N + \nabla_r(\rho_N \mathbf{u}) = 0 \quad (40a)$$

$$\dot{\rho}_E + \nabla_r \left( \frac{3v^2}{2v^2 + u^2} \mathbf{u} \rho_E \right) = \mathbf{F} \cdot \mathbf{u} \rho_N \quad (40b)$$

$$\mathbf{u} \partial_t P + v^2 \nabla P + \mathcal{H}(\partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u}) = \rho_N (v^2 \mathbf{F} - \mathbf{u}(\mathbf{F} \cdot \mathbf{u})) - \frac{\mathcal{H} \mathbf{u}}{\tau} \quad (40c)$$

The quantity  $\mathcal{H} = \frac{3v^2}{2v^2 + u^2} \rho_E$  is the enthalpy and  $P = \frac{v^2 - u^2}{2v^2 + u^2} \rho_E$  is the pressure. The first two formulae are continuity equations, while the last expression is the Euler equation for relativistic fluids. We have neglected all dissipative (viscous) corrections. We incorporate the magic angle effect by using  $v = v(\alpha)$  and the force  $\mathbf{F} = -UZ(\alpha)^2 \nabla \rho_N$  accounts for onsite Hubbard interaction energy in the Hartree approximation.

We numerically solve Eqs. (40), see Supplementary Figure 11 which demonstrates the manifestation of the magic angle effect in various quantities: (i) There is an  $\alpha$  dependent crossover from a short ballistic evolution to diffusive motion at time  $\tau$ . (ii) Near the magic angle, the overdamped dynamics [vanishing  $\tau(\alpha)$ ] prevents the ballistic wave front propagation of energy density (compare c) vs. d)). (iii) Similarly, it prevents the formation of waves in the density profile and (iv) the negative expansion velocity of average position  $\langle r^2 \rangle$  at intermediate time. (v) The behavior as a function of  $\alpha$  (or  $W$ ) is nonmonotonic, compare Supplementary Figure 11 e),f) and Supplementary Figure 10.

*Discussion of experimental observations.* We propose the observation of wave packet dynamics that can be measured in experiment using absorption imaging techniques. We have discussed two complementary limits of wavepacket dynamics: (i) the non-interacting limit, Supplementary Figure 10 (ii) an interacting Fermi gas, Supplementary Figure 11.

The experimental time evolution is limited, e.g. due to particle loss, to about 50 hopping times. This is sufficient to resolve the onset of the crossover from ballistic to (super-)diffusive motion in the non-interacting case.

In the strong coupling regime, and for the chosen parameters in Supplementary Figure 11 ( $\rho_E = v(\alpha = 0)/\xi^2$ ) the energy density at  $\alpha = 0$  implies a typical temperature scale of the order  $T \sim ta/\xi$  ( $a$  is the lattice constant). For  $a/\xi \sim 1/4$  this corresponds to  $T/T_F \sim 1/4$  which is well in the experimentally realistic regime. The observation of the consequences of the magic angle effect at fixed energy density requires a time evolution at least up to  $10\xi/v(\alpha = 0) = 10(\xi/a)/t$ . Thus, for typical cloud sizes  $\xi \sim 4a$  (i) the typical temperature in Supplementary Figure 11 is much smaller than the bandwidth (ii) magic angle effects are well resolvable in present day cold atomic experiments with coherence up to 50 hopping times.

## Supplementary Note 5. MULTIFRACTAL ANALYSIS AND WAVEFUNCTIONS

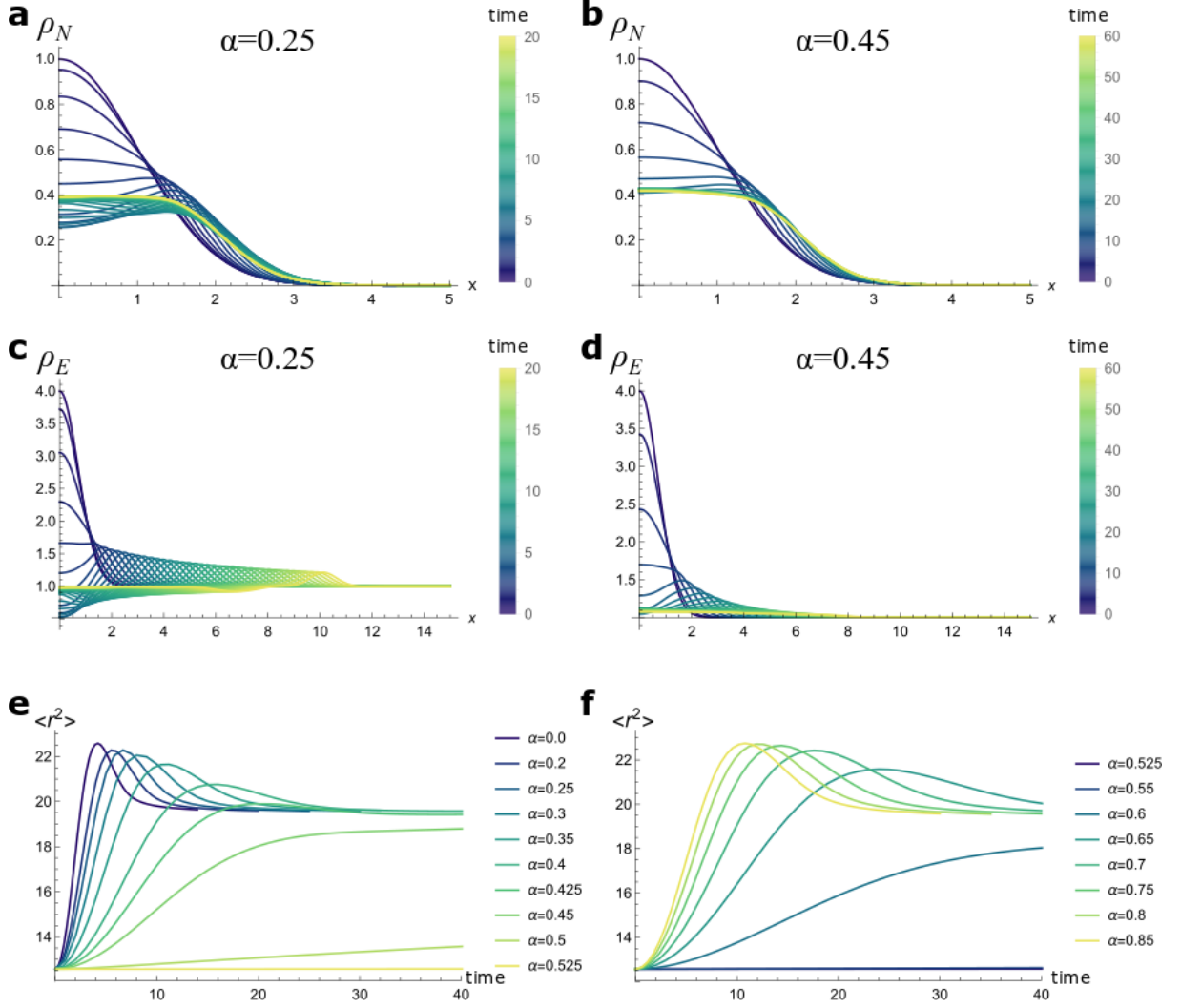
In this supplementary note, we explain the procedure of extracting the multifractal exponents and the underlying physical interpretations.

*Scaling exponent  $\tau_M$ .* We now discuss how we extract  $\tau_M(q)$  from inverse participation ratio (IPR). For a given wavefunction in momentum space, the finest grid is  $2\pi/L \times 2\pi/L$ . We introduce an integer binning factor  $B$  which controls the resolution of the momentum space wavefunction. The IPR constructed in this way has been dubbed the momentum space IPR<sup>6</sup> and given by

$$\mathcal{I}_M(q, L; B) = \sum_{k'} \left| \tilde{\phi}_E(k'; B) \right|^{2q} \sim \left( \frac{Ba}{L} \right)^{\tau_M(q)}, \quad (41)$$

where  $k'$  denotes the wavevector of the binned wavefunction grid and  $\tilde{\phi}_E(k'; B)$  is the binned wavefunction. The scaling behavior only holds when  $1 \ll B \ll L/a$ . To extract  $\tau_M(q)$ , we choose two consecutive values of binning factors,  $B_1$  and  $B_2$ , and then perform numerical derivative as follows:

$$\tau_M(q; B_1, B_2) = \frac{\ln \mathcal{I}_M(q, L, B_1) - \ln \mathcal{I}_M(q, L, B_2)}{\ln B_1 - \ln B_2}. \quad (42)$$



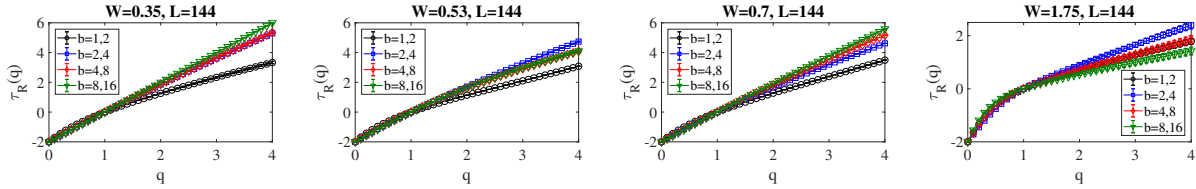
Supplementary Figure 11: Simulation of hydrodynamic evolution, Eq. (40). Initial conditions are  $\rho_E(0) = \frac{v(\alpha=0)}{\xi^3} \left(1 + 3e^{-r^2/\xi^2}\right)$ ,  $\rho(\mathbf{r}, t=0) = \frac{e^{-r^2/[2\xi^2]}}{\xi^2}$ , and  $\mathbf{u}(0) = 0$ . In the plots, all lengths are in units of  $\xi$ , all times in units of  $\xi/v(\alpha=0)$ , in these units  $U(\alpha=0) = 0.1, 1/\tau(\alpha=0) = 0.03$ . Evolution for various times (shown as color) of the particle density a) far from and b) near the magic angle  $\alpha_c = 0.536$ ; c) d) similarly for the evolution of the energy density. e), f): Evolution of  $\langle r^2 \rangle = \int d^2x r^2 \rho_N(\mathbf{r}, t)$ , where the value of  $\alpha$  is shown as color.

At last, we average  $\tau_M(q; B_1, B_2)$  over phases in the quasiperiodic potential. For  $L = 144$ , we construct  $\tau_M$  with 100 realization. We only use 10 realization for  $L = 610$ . For real space scaling exponents ( $\tau_R$ ), we adopt similar procedures to extract their values. For the case of cTBG, we use 100 realizations of  $L = 377$ , but to handle binning we truncate the wavefunction at the edge of the BZ (where the wave function is effectively zero) to produce a  $376 \times 376$  wave function we can then perform the binning scheme on.

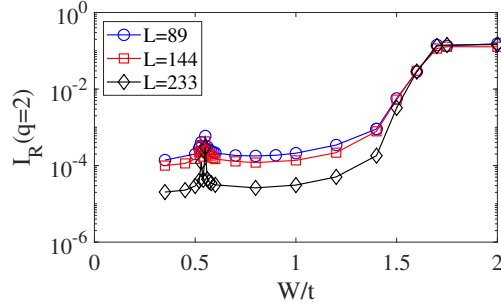
*Real space multifractality.* As a complementary analysis, we study the multifractal spectrum of the real space wavefunctions at zero energy. We construct the corresponding real space scaling exponent spectrum  $\tau_R(q)$  which encodes plane wave, multifractal, and localized wavefunctions.

For  $W/t < 0.525$ , the real space wavefunction is a plane wave, characterized by the Fourier modes at Dirac points  $(k_x, k_y) = (0, 0), (0, \pi), (\pi, 0),$  and  $(\pi, \pi)$ . The multifractal spectrum  $\tau_R(q)$  is characterized by a straight line  $\tau_R(q) = 2(q - 1)$ . Plane wave states can also be found for  $0.55 < W < 1.5$ , the inverted semimetallic regime. For  $0.525 < W/t < 0.55$ , the real space wavefunctions show multifractal behavior characterized by a nonlinear  $\tau_R(q)$  spectrum. For  $W/t \gg 1.5$ , the real space wavefunctions become localized,  $\tau_R(q) = 0$  for  $q > 0$ .

We numerically compute the  $\tau_R(q)$  spectrum for various values of  $W$  in different phases. In Supplementary Figure 12, we compute  $W/t = 0.35, 0.53, 0.7, 1.75$  for  $L = 144$  and average over 100 realizations.  $W/t = 0.35$  and  $W/t = 0.7$



Supplementary Figure 12:  $\tau_R(q)$  with different values of  $W$  with  $L = 144$  and 100 realizations. For  $W = 0.35$  and  $W = 0.7$ , the multifractal spectra approach to plane wave,  $\tau_R(q) = 2(q - 1)$ , as increasing  $b$ . For  $W = 0.53$ , the spectrum is multifractal described by a nonlinear function. For  $W = 1.75$ ,  $\tau_R$  approaches to a localized spectrum,  $\tau_R(q) = 0$  for  $q > 0$ , as increasing  $b$ . The existing data still suffer from a strong finite size effect and do not clearly show a localized spectrum. We therefore turn to the finite size dependence of the real space wavefunction with  $W = 1.75$ , which indeed shows localization as the real space IPR is  $L$  independent in this regime.



Supplementary Figure 13: The inverse participation ratio of real space wavefunctions with different system sizes. Blue circles represent the data of  $L = 89$  with 100 realization; red squares represent data of  $L = 144$  with 100 realization; black diamonds represent data of  $L = 233$  with 50 realization. The collapse of data with different sizes ( $W/t \geq 1.75$ ) indicate localized regime.

show plane wave spectrum as increasing the real space binning factor  $b$ .  $W/t = 0.53$  demonstrate a weakly nonlinear multifractal spectrum with a fitting function  $\tau_R(q) \approx 2(q - 1) - 0.16q(q - 1)$  for  $|q| < 1$ . A much larger system size is needed for quantitatively determining the spectrum which is beyond the scope of this work. For  $W/t = 1.75$ , the spectrum gradually approaches to localized like behavior as increasing  $b$ . In our  $L = 144$  data, we do not find a clear localized spectrum, due to finite size effects. We therefore turn to the real space IPR data as a function of system size in Supplementary Figure 13. This indeed shows strong evidence for localization as the IPR is  $L$  independent for  $W/t \geq 1.75$ .

*Interpretation of unfreezing transition.* As we have demonstrated in the maintext, the momentum space wavefunctions serve as a proxy for the semimetal-metal transition. The multifractal analysis in momentum space here is distinct from the conventional notion of wavefunctions in real space<sup>10</sup>. We discuss the interpretation of the multifractal spectrum in depth here.

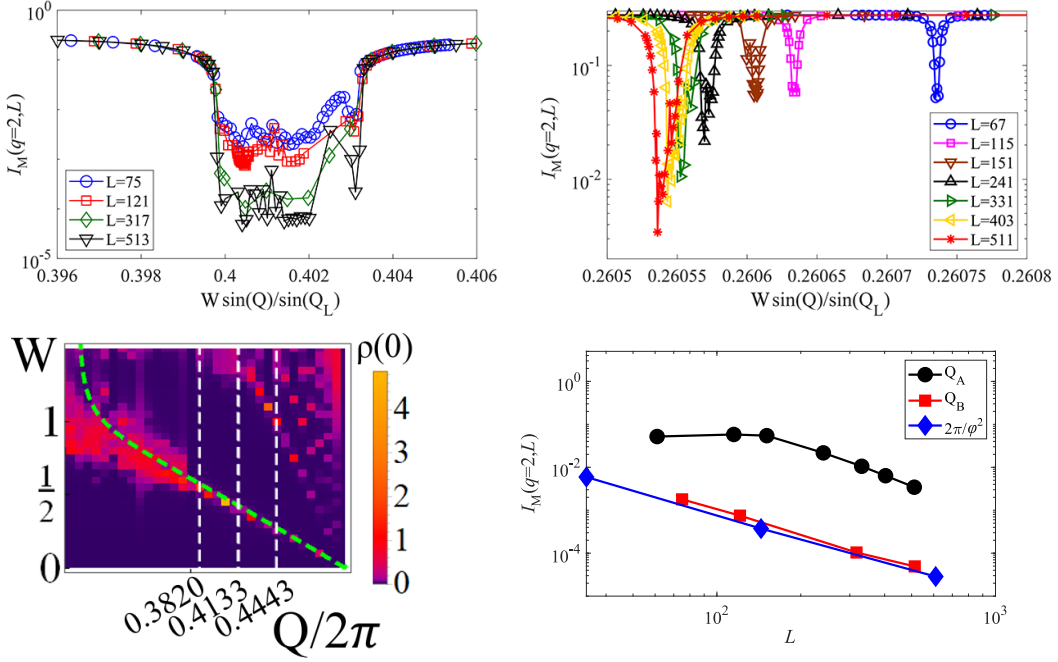
For  $W = 0$ , the zero energy wavefunction is composed of the Fourier modes at the Dirac points. The zero energy states are linear combinations of these four plane waves. The probability distributions (integrating over the spin degrees of freedom) of the momentum space wavefunction generically contains four peaks, which we call “ballistic peaks”. In the multifractal  $\tau_M(q)$  spectrum, these ballistic peaks correspond to a frozen spectrum,  $\tau_M(q) = 0$  for  $q > 1$ .

The frozen feature in the momentum space wavefunction suggests that the ballistic peaks are sharply defined, with the finest localization length  $\pi/L$ . The  $\tau_M(q)$ s extracted from the numerics weakly depend on the choice of binning factors  $B_1$  and  $B_2$ . As  $W$  increases, satellite peaks with weaker amplitude arise. When  $W \approx W_c$ , the  $\tau_M(q)$ s still show freezing behavior for  $B = 1, 2$  but become generically non-zero for larger binning factors. Such features suggests that the distance between a ballistic peak and the nearest satellite peaks is around  $2 * 2\pi/L$  to  $4 * 2\pi/L$ . The ballistic peaks start to hybridize with the satellite peaks when  $W \geq W_c$ . This corresponds to an unfreezing transition in momentum space, which describes a zero-measure set to an extensive set of Fourier modes. The  $W_c$  determined here coincides well with the semimetal-metal transition extracted from the density of states calculations.

For larger values of  $W$ , an inverse semimetal transition, metal-to-semimetal, takes place. The same multifractal analysis in momentum space also applies.

*Minimal system size for momentum space delocalization.* In this part we discuss the minimal system size required





Supplementary Figure 14: Comparisons of IPRs in momentum space that show scaling with system size indicative of delocalization in momentum space, (Top left)  $Q = Q_B = 2\pi(51 - \sqrt{5})/181 \approx 2\pi \times 0.4133$  (Top right)  $Q = Q_A = 2\pi(1/2 - 1/\varphi^6) \approx 2\pi \times 0.4443$ . (Bottom left) The phase diagram, as in Fig 1 **b** of the main text, with cuts along  $Q = 2\pi/\varphi^2 \approx 2\pi \times 0.3820$ ,  $Q = Q_B = 2\pi \times 0.4133$ ,  $Q = Q_A = 2\pi \times 0.4443$ . (Bottom right) Summary of system size dependence of the minimal IPR along these three cuts, demonstrating the existence of a crossover scale between delocalized and localized regimes for  $Q = Q_A$ .

to observe the universal magic angle effect, i.e. momentum space delocalization at the magic angle.

At small twist angle, the separation of length scales between original lattice spacing and moiré supercell becomes large and suppresses the size of the metallic sliver in the phase diagram, Fig. 1**b, c** of the main text. This renders the observation in numerical (i.e. finite size) simulations very difficult. At smallest twist angles,  $\theta \ll 10^\circ$  for TBG and  $(\pi - Q) \ll 1$  for 2D SOC, wavefunction delocalization is not observable at small system sizes.

However, for the specific case of the 2D SOC model and a given  $Q \approx \pi$ , we here demonstrate the appearance of a crossover length scale, beyond which the magic angle effect becomes observable. In Supplementary Figure 14 we present data for  $Q = 2\pi(1/2 - 1/\varphi^6) \approx 2\pi \times 0.4443$ , for which wave function delocalization (i.e. system size dependent IPR) occurs for  $L \gtrsim 160$  and we compare this to other  $Q$  further from  $\pi$ . Note that for  $Q \approx 2\pi \times 0.4443$  the position of the magic angle non-trivially depends on the finite size (i.e. on the concrete numerical rational approximant).

We conclude the discussion of the origin of this critical system size with a physical interpretation. Within our finite size numerics, the incommensurate perturbation repeats itself on the length scale of the entire system. Within our picture, the commensurate structure provides a cut-off for the single particle transition that rounds out the non-analyticity and the wavefunctions do not delocalize in momentum space. Based on this and our numerical observation, we conjecture that for any irrational  $Q$  and sufficiently large system size, wave function delocalization may be observed.

**Supplementary Note 6. WANNIER STATES AND BUILDING HUBBARD MODELS**

In order to analytically build the Hubbard models analyzed in the main text, we use the Hamiltonian as perviously defined

$$\begin{aligned}
\hat{H} &= \hat{T} + \hat{V} + \hat{U}, \\
\hat{T} &= \sum_{\mathbf{r}, \mu} \left[ \frac{it}{2} c_{\mathbf{r}}^{\dagger} \sigma_{\mu} c_{\mathbf{r}+\hat{\mu}} + \text{h.c.} \right], \\
\hat{V} &= W \sum_{\mathbf{r}, \mu} \cos(Q_n r_{\mu} + \phi_{\mu}) c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}}, \\
\hat{U} &= U \sum_{\mathbf{r}} c_{\mathbf{r}, \uparrow}^{\dagger} c_{\mathbf{r}, \downarrow}^{\dagger} c_{\mathbf{r}, \downarrow} c_{\mathbf{r}, \uparrow}.
\end{aligned} \tag{43}$$

We take  $Q_n = 2\pi F_{n-2}/F_n$  and the system size to be  $L = mF_n$  such that  $F_n$  defines a supercell of size  $\ell = F_n$ . The model has a particle-hole symmetry when  $\phi_{\mu} = \pi/2$ , which we will concentrate on in order to isolate bands near  $E = 0$ .

The first task is to isolate bands. With  $L = m\ell$ , our Brillouin Zone will have  $m^2$  sampled points. Due to the imposed discrete translational symmetry, one can write the single particle Hamiltonian  $\hat{H}_{\text{sp}} = \hat{T} + \hat{V}$  as

$$\hat{H}_{\text{sp}} = \bigoplus_{n_x, n_y=0}^{m-1} \hat{H}_{\mathbf{n}}, \tag{44}$$

for  $\mathbf{n} = (n_x, n_y)$  and, using Bloch's theorem, the individual Hamiltonians are (in first quantized notation)

$$\hat{H}_{\mathbf{n}} = \sum_{\mathbf{r}, \mu} \left[ \frac{it}{2} e^{-i2\pi n_{\mu}/L} |\mathbf{n}, \mathbf{r}\rangle \langle \mathbf{n}, \mathbf{r} + \hat{\mu}| \otimes \sigma_{\mu} + \text{h.c.} \right] + W \sum_{\mathbf{r}, \mu} \cos(Q_n r_{\mu} + \phi_{\mu}) |\mathbf{n}, \mathbf{r}\rangle \langle \mathbf{n}, \mathbf{r}|, \tag{45}$$

on a system size of size  $\ell$ . In the thermodynamic limit  $m \rightarrow \infty$ , and we can plot dispersions to see if a gap has opened. Once diagonalized, we have a set of states  $\{E_{\mathbf{n}, j}\}$ , and looking near  $E = 0$ , we can form projectors onto energy states within a miniband

$$\hat{P} = \sum_{E_{\mathbf{n}, j} \in \text{Miniband}} |\mathbf{n}, j\rangle \langle \mathbf{n}, j|. \tag{46}$$

This projector should have an integer multiple of  $m^2$  states within it, and each  $(n_x, n_y)$  pair should contribute the same number of states. This is a check to determine if we have a ‘‘good band.’’

Further, we can get an idea of where the Wannier centers should be by looking at the integrated local density of states

$$\rho_{\text{Band}}(\mathbf{r}) = \langle \mathbf{r} | \hat{P} | \mathbf{r} \rangle. \tag{47}$$

Considering systems with periodic boundary conditions, we consider the position operators  $\{\cos(2\pi\hat{x}/L), \sin(2\pi\hat{x}/L), \cos(2\pi\hat{y}/L), \sin(2\pi\hat{y}/L)\}$ , and we construct the projected operators

$$\{\hat{P} \cos(2\pi\hat{x}/L) \hat{P}, \hat{P} \sin(2\pi\hat{x}/L) \hat{P}, \hat{P} \cos(2\pi\hat{y}/L) \hat{P}, \hat{P} \sin(2\pi\hat{y}/L) \hat{P}\}. \tag{48}$$

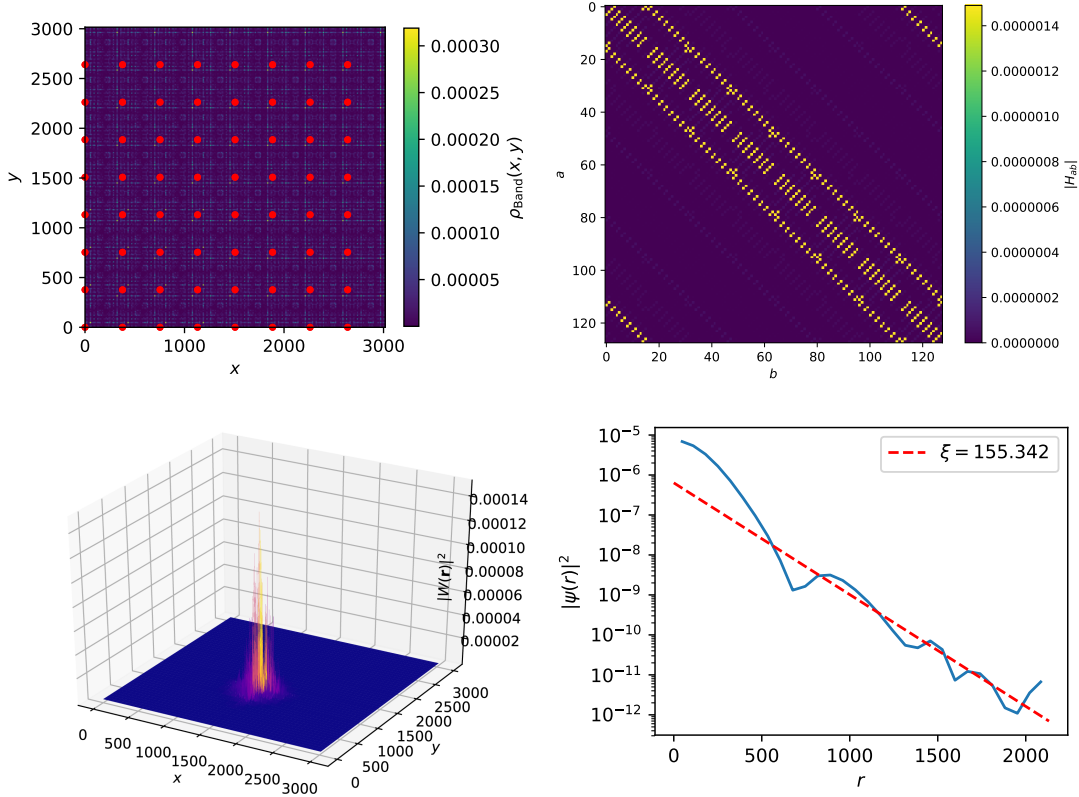
These operators no longer commute with one another and cannot be simultaneously diagonalized. The solution is *approximate joint diagonalization* (AJD) achieved by minimizing a cost function: for a set of operators  $\mathcal{O}$ , we need to find a unitary matrix  $U$  such that

$$\sigma(U) = \sum_{A \in \mathcal{O}} \sum_{i \neq j} |(U^{\dagger} A U)_{ij}|^2 \tag{49}$$

is minimized. An efficient algorithm for this was developed for signal processing<sup>11</sup> and enumerated for maximally localized Wannier states in Ref. 12. Our algorithm to achieve this, written in Julia, can be found in Ref. 13.

The columns of the resulting unitary matrix  $U$  are the maximally localized Wannier states. The resulting single particle part of tight-binding Hamiltonian is found numerically to be given approximately by

$$\hat{H}_{\text{Wannier}} = U^{\dagger} \hat{P} \hat{H}_{\text{sp}} \hat{P} U \approx \sum_{\mathbf{R}, \mu} \left[ \frac{it_{\text{eff}}}{2} |W_{\mathbf{R}, \tau}\rangle \langle W_{\mathbf{R}+\hat{\mu}, \tau'}| [\sigma_{\mu}]_{\tau, \tau'} + \text{h.c.} \right], \tag{50}$$



Supplementary Figure 15: (Upper left) The integrated density of states  $\rho_{\text{Band}}(x, y)$  along with Wannier centers (red dots) for  $\ell = 377$  and  $m = 8$  and  $W = 0.52445t$ . (Upper right) The magnitude of the elements of the Wannier Hamiltonian  $H_{\text{Wannier}}$ . (Lower left) 3D representation of a single Wannier function  $|W(\mathbf{r})|^2$  and (Lower right) The exponential localization of the Wannier function  $|W(\mathbf{r})|^2 \sim e^{-r/\xi}$ .

for Wannier states  $|W_{\mathbf{R},\tau}\rangle$  labeled by the emergent lattice positions  $\mathbf{R}$  and internal index  $\tau = 1, 2$ . The interaction term also gets renormalized (again, returning to second quantized notation)

$$\hat{H}_{\text{Wannier-int}} = U_{\text{eff}} \sum_{\mathbf{R}} d_{\mathbf{R},1}^\dagger d_{\mathbf{R},2}^\dagger d_{\mathbf{R},2} d_{\mathbf{R},1} \quad (51)$$

where  $d_{\mathbf{R},\tau}$  annihilates the states  $|W_{\mathbf{R},\tau}\rangle$ , and

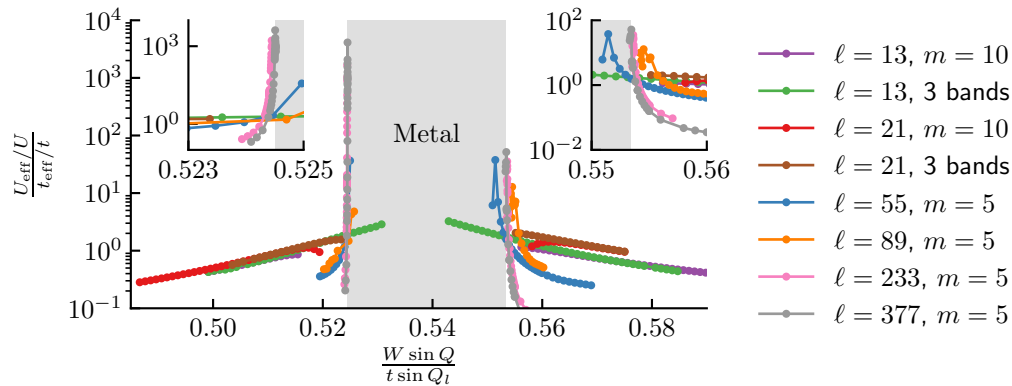
$$U_{\text{eff}} = U \sum_{\mathbf{r}} \{ |W_{\mathbf{R},1}^\uparrow(\mathbf{r})|^2 |W_{\mathbf{R},2}^\downarrow(\mathbf{r})|^2 + |W_{\mathbf{R},2}^\uparrow(\mathbf{r})|^2 |W_{\mathbf{R},1}^\downarrow(\mathbf{r})|^2 - 2\Re[W_{\mathbf{R},1}^\uparrow(\mathbf{r})^* W_{\mathbf{R},2}^\downarrow(\mathbf{r})^* W_{\mathbf{R},1}^\downarrow(\mathbf{r}) W_{\mathbf{R},2}^\uparrow(\mathbf{r})] \}, \quad (52)$$

where  $\langle \mathbf{r}, \sigma | W_{\mathbf{R},\tau} \rangle = W_{\mathbf{R},\tau}^\sigma(\mathbf{r})$  and  $U_{\text{eff}}$  is found to be (numerically) independent of  $\mathbf{R}$ .

To perform the calculations, we look close to the transition  $W_c$  calculated previously and find where the smallest gap opens in the dispersion. In the range where that gap opens and closes, we perform Lanczos to find the Bloch states and then we perform AJD to find the Wannier states. After making sure the states are “good” states (exponentially localized and make the appropriate emergent lattice), we can find the Hamiltonian as described above.

*Computed Wannier functions.* As described in Supplementary Note Supplementary Note 3, we can probe minibands by using successive pairs of odd Fibonacci (even Fibonacci gap out the Dirac nodes, not allowing us to make the effective Hubbard model described above). We perform the calculation for  $\ell = F_n = 13, 21$  (second miniband),  $\ell = F_n = 55, 89$  (third miniband), and  $\ell = 233, 377$  (fourth miniband). In the main text, we present data for  $\ell = 13$ , and here we show data for the point of highest  $U_{\text{eff}}/t_{\text{eff}} \approx 4115U/t$ .

After tracking the gap opening and closing when  $\ell = 377$ , we construct the Wannier states (at  $W = 0.52445t$  and  $m = 8$ ), and the results are in Supplementary Figure 15. We see Wannier centers on top of the  $\rho_{\text{Band}}(\mathbf{r})$  in Supplementary Figure 15 upper-left, along with a visual representation of the resulting  $H_{\text{Wannier}}$  on the upper-right (clearly showing a banded Hamiltonian). The Wannier state is visually seen in Supplementary Figure 15 lower-left and after we bin the data, we find the states are exponentially localized in Supplementary Figure 15 lower-right.



Supplementary Figure 16: The interaction of the center band approaching the transition. Each successive pair of odd  $\ell$  represent the opening of a new miniband. The data labeled “3 bands” comes from the multiorbital models, following the Wannier states continuously connected to the single-orbital model. Note that  $Q = 2\pi/\varphi^2$  for golden ratio  $\varphi$  and  $Q_\ell = 2\pi F_{n-2}/\ell$  with  $\ell = F_n$ .

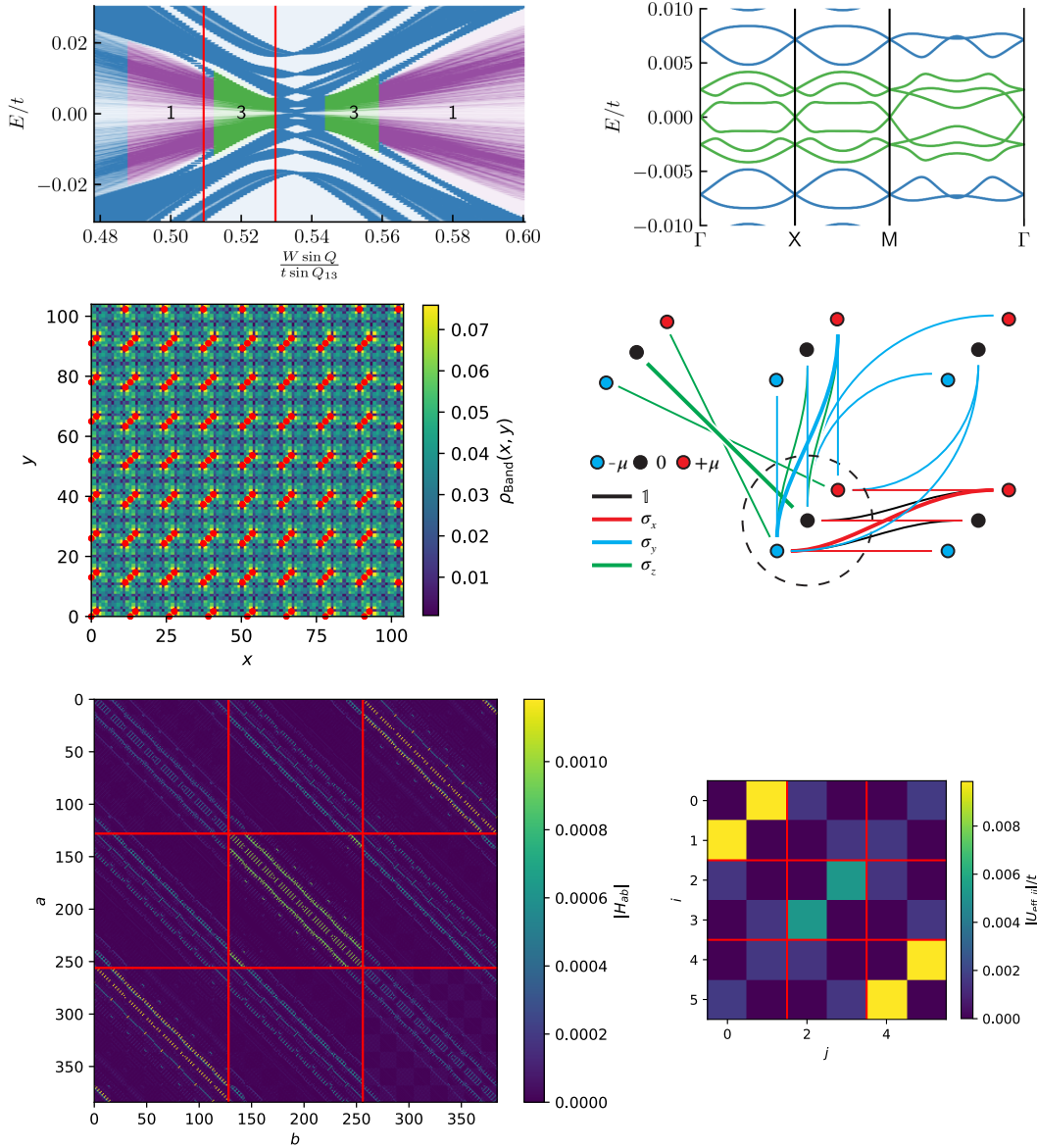
Doing this multiple times leads to an estimate of the effective interactions on either side of the metal transition in Supplementary Figure 16. We should note that for  $\ell = 13, 21$ , part of the data is pulled from a multi-band Hubbard model by tracking  $U_{\text{eff}}/t_{\text{eff}}$  on the Wannier functions with maximal overlap prior to the multi-band Hubbard model. Also, the Wannier states for  $\ell = 89$  could not fully be fully converged, so the data for  $U_{\text{eff}}/t_{\text{eff}}$  is appropriately noisy (and an underestimate).

*Finite energy bands and multi-orbital Hubbard model.* It is interesting to note that at finite energy, flat bands develop and in some cases intersect the band near  $E = 0$  which is the focus of this text. As a clear example, note for  $\ell = 13$ , see that a flat, Dirac bands appear around the region labeled 1 in Supplementary Figure 17(upper left) and intersect the band in the region labeled 3. These flat bands have greatly increased  $U_{\text{eff}}/t_{\text{eff}}$  away from the transition. When the multiorbital Hubbard model appears, the dispersion changes drastically, see Supplementary Figure 17(upper right) and has Wannier centers appearing along diagonals within a supercell as seen in Supplementary Figure 17(middle left). With the computed data, an entire translationally invariant Hubbard model can be constructed and the result leads to hoppings as indicated in Supplementary Figure 17(middle right), from the banded Hamiltonian Supplementary Figure 17(lower left). Lastly, the multiorbital Hubbard  $U$  is a matrix in this case and is visually represented in Supplementary Figure 17(lower right).

---

\* YXF, EJK, and JHW contributed equally.

- [1] Gärttner, M., Syzranov, S. V., Rey, A. M., Gurarie, V. & Radzihovsky, L. Disorder-driven transition in a chain with power-law hopping. *Phys. Rev. B* **92**, 041406 (2015). URL <https://link.aps.org/doi/10.1103/PhysRevB.92.041406>.
- [2] Bistritzer, R. & MacDonald, A. H. Moiré bands in twisted double-layer graphene. *Proceedings of the National Academy of Sciences* **108**, 12233–12237 (2011).
- [3] dos Santos, J. L., Peres, N. & Neto, A. C. Continuum model of the twisted graphene bilayer. *Physical Review B* **86**, 155449 (2012).
- [4] Tarnopolsky, G., Kruchkov, A. J. & Vishwanath, A. Origin of Magic Angles in Twisted Bilayer Graphene. *Physical Review Letters* **122**, 106405 (2019).
- [5] Weiße, A., Wellein, G., Alvermann, A. & Fehske, H. The kernel polynomial method. *Rev. Mod. Phys.* **78**, 275–306 (2006). URL <https://link.aps.org/doi/10.1103/RevModPhys.78.275>.
- [6] Pixley, J. H., Wilson, J. H., Huse, D. A. & Gopalakrishnan, S. Weyl semimetal to metal phase transitions driven by quasiperiodic potentials. *Phys. Rev. Lett.* **120**, 207604 (2018). URL <https://link.aps.org/doi/10.1103/PhysRevLett.120.207604>.
- [7] Syzranov, S. V. & Radzihovsky, L. High-dimensional disorder-driven phenomena in weyl semimetals, semiconductors, and related systems. *Annual Review of Condensed Matter Physics* **9**, 35–58 (2018).
- [8] Kamenev, A. *Field theory of non-equilibrium systems* (Cambridge University Press, 2011).
- [9] Narozhny, B. N., Gornyi, I. V., Mirlin, A. D. & Schmalian, J. Hydrodynamic approach to electronic transport in graphene. *Annalen der Physik* **529**, 1700043 (2017).
- [10] Evers, F. & Mirlin, A. D. Anderson transitions. *Reviews of Modern Physics* **80**, 1355 (2008).
- [11] Cardoso, J. & Souloumiac, A. Jacobi Angles for Simultaneous Diagonalization. *SIAM Journal on Matrix Analysis and*



Supplementary Figure 17: (Upper left) For  $\ell = 13$ , we see a region (labeled 1) where one Dirac band exists around  $E = 0$  (the figures in the main text are constructed here). There is also an extremely flat band at finite energy that intersects the band in the region labeled 3. (This data is taken for  $m \rightarrow \infty$  by random sampling the Brillouin zone and  $\phi_\mu = \pi/2$ .) (Upper right) The dispersion in region 3 at  $W = 0.52t$ , the second red line in the upper left figure. (Middle left) The integrated density of states  $\rho_{\text{Band}}(x, y)$  for the three band model. We can see a line of three states tiled in an  $8 \times 8$  grid. (Middle right) A visual representation of the single-particle part of the multi-orbital Hubbard model. The Wannier states are separated by their average energy  $-\mu, 0, +\mu$ , and they have hopping matrix elements to neighbors and next-to-nearest neighbors on the diagonal. Each Wannier site is double degenerate and the hopping matrix elements are Pauli matrices as color coded. (Bottom left) The Wannier Hamiltonian separated into the 3 orbits (the red lines); we see that it is banded and dominated by the hoppings shown in the middle-right figure. (Bottom right) The Hubbard  $U$  matrix for the multi-orbital model. The highest values of the Hubbard  $U$  matrix comes from states continuously connected to flat bands in the upper-left figure that intersect the middle band in region 3.

- Applications* **17**, 161–164 (1996). URL <http://epubs.siam.org/doi/abs/10.1137/S0895479893259546>.
- [12] Gygi, F., Fattebert, J.-L. & Schwegler, E. Computation of Maximally Localized Wannier Functions using a simultaneous diagonalization algorithm. *Computer Physics Communications* **155**, 1–6 (2003). URL <http://www.sciencedirect.com/science/article/pii/S0010465503003151>.
- [13] Approximate joint diagonalization. <https://github.com/jhwilson/AJD.jl>. [Online; accessed 12-Sept-2018].