Coulomb Interaction-Stabilized Isolated Narrow Bands with Chern Numbers $\mathcal{C} > 1$ in Twisted Rhombohedral Trilayer-Bilayer Graphene

Võ Tiến Phong[1, 2, *] and Cyprian Lewandowski[1, 2, †]

[1]*Department of Physics, Florida State University, Tallahassee, FL, 32306, U.S.A.*
[2]*National High Magnetic Field Laboratory, Tallahassee, FL, 32310, U.S.A.*
(Dated: June 26, 2025)

Supplementary Material

CONTENTS

* vophong@magnet.fsu.edu
† clewandowski@magnet.fsu.edu

LIST OF FIGURES

LIST OF TABLES

I. CONTINUUM HAMILTONIAN

We construct twisted rhombohedral multilayer graphene (TRMG_{3+2}) in the following way. We start with a rhombohedral trilayer whose valley-projected Hamiltonian at valley $\nu = \pm$ is given by

$$\mathbb{K}_{3,\nu}(\mathbf{k}) = \begin{pmatrix} \mathbb{K}_{1,\nu}(\mathbf{k}) & \mathbb{U}_\nu(\mathbf{k}) & \mathbb{W} \\ \mathbb{U}_\nu^\dagger(\mathbf{k}) & \mathbb{K}_{1,\nu}(\mathbf{k}) & \mathbb{U}_\nu(\mathbf{k}) \\ \mathbb{W}^\dagger & \mathbb{U}_\nu^\dagger(\mathbf{k}) & \mathbb{K}_{1,\nu}(\mathbf{k}) \end{pmatrix} \tag{S1a}$$

$$\mathbb{K}_{1,\nu}(\mathbf{k}) = \begin{pmatrix} 0 & \hbar v_0 \Pi_\nu(\mathbf{k}) \\ \hbar v_0 \Pi_\nu^\dagger(\mathbf{k}) & 0 \end{pmatrix}, \tag{S1b}$$

$$\mathbb{U}_\nu(\mathbf{k}) = \begin{pmatrix} \hbar v_4 \Pi_\nu(\mathbf{k}) & \hbar v_3 \Pi_\nu^\dagger(\mathbf{k}) \\ \gamma_1 & \hbar v_4 \Pi_\nu(\mathbf{k}) \end{pmatrix}, \tag{S1c}$$

$$\mathbb{W} = \begin{pmatrix} 0 & \gamma_2/2 \\ 0 & 0 \end{pmatrix}, \tag{S1d}$$

where $\mathbf{k}$ is measured from the center of the microscopic Brillouin zone, $\Pi_\nu(\mathbf{k}) = \nu k_x - i k_y - 4\pi/3a$, $a = 2.46$ Å is lattice constant of graphene, $\hbar v_i = -\sqrt{3}\gamma_i a/2$, and γ_i are the hopping parameters. We ignore dimerization energies on eclipsed sites in this work. The aforementioned hopping parameters used in this work are listed in Table I. Next, we stack Bernal bilayer graphene on top of the trilayer in perfect crystallographic alignment. If the orientation of the bilayer is *parallel* to the trilayer, then the bilayer Hamiltonian is

$$\mathbb{K}_{2,\nu}(\mathbf{k}) = \begin{pmatrix} \mathbb{K}_{1,\nu}(\mathbf{k}) & \mathbb{U}_\nu(\mathbf{k}) \\ \mathbb{U}_\nu^\dagger(\mathbf{k}) & \mathbb{K}_{1,\nu}(\mathbf{k}) \end{pmatrix}; \tag{S2}$$

otherwise, if the orientation of the bilayer is *antiparallel* to the trilayer, then the bilayer Hamiltonian is

$$\mathbb{K}_{2,\nu}(\mathbf{k}) = \begin{pmatrix} \mathbb{K}_{1,\nu}(\mathbf{k}) & \mathbb{U}_\nu^\dagger(\mathbf{k}) \\ \mathbb{U}_\nu(\mathbf{k}) & \mathbb{K}_{1,\nu}(\mathbf{k}) \end{pmatrix}. \tag{S3}$$
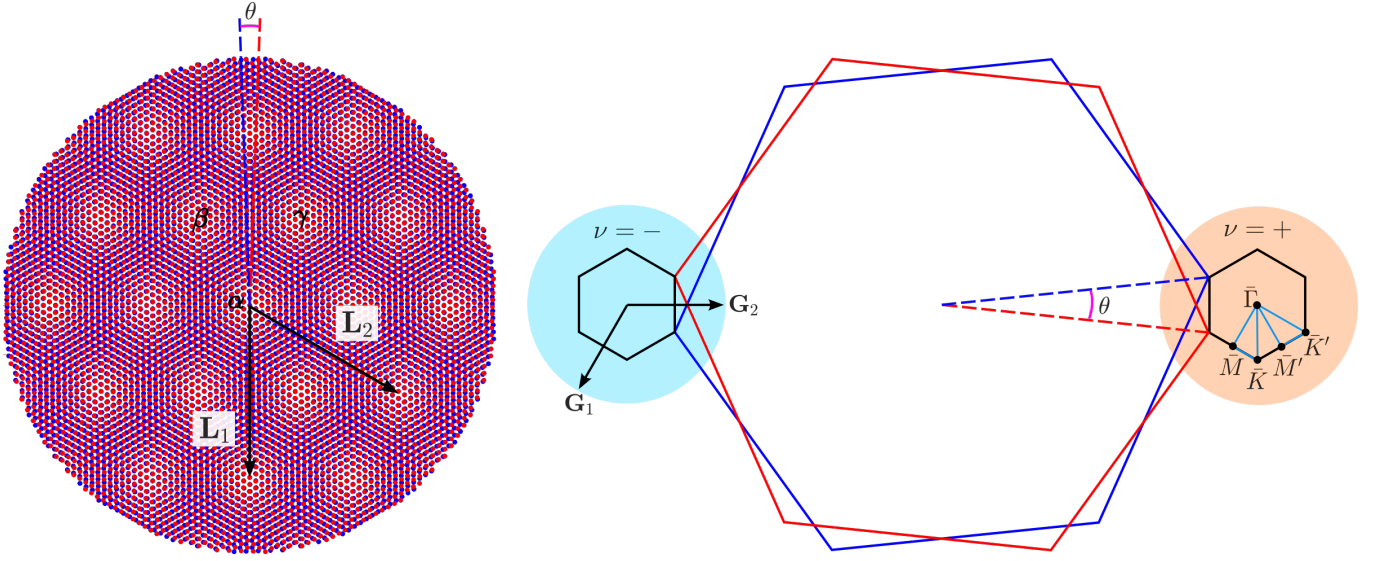
Figure S1. **Lattice representation of twisted rhombohedral trilayer-bilayer graphene.** (Left) Moiré pattern of rhombohedral bilayer graphene, whose atoms are colored red, stacked on top of rhombohedral trilayer graphene, whose atoms are colored blue, with a relative twist angle θ. Regions of high-symmetry stacking order are labeled α, β, γ. (Right) Reconstructed moiré Brillouin zone shown by the black hexagons, where high-symmetry points are marked. The larger blue and red hexagons are the rotated Brillouin zones of the graphene multilayer subsystems before hybridization.

The two inequivalent stacking orders differ by taking the Hermitian adjoint of the interlayer hopping matrix $\mathbb{U}_\nu(\mathbf{k})$ in the bilayer sector. One stacking order is related to the other by a C_{2z} rotation of the bilayer. Next, we rotate the trilayer by $+\theta/2$ and the bilayer by $-\theta/2$ around a common vertical axis. By continuing to measure momentum using a common (non-rotated) coordinate system for both stacks, the momentum in the trilayer gets rotated by $-\theta/2$ while the momentum in the bilayer gets rotated by $+\theta/2$. The rotation in reciprocal space is opposite to the rotation in real space to compensate for the use of a common axis. Thus, the Hamiltonian of a twisted trilayer-bilayer stack is simply

$$\mathbb{K}_{3+2,\nu}(\mathbf{k}) = \begin{pmatrix} \mathbb{K}_{3,\nu}(\mathbb{R}[-\theta/2]\mathbf{k}) & * \\ * & \mathbb{K}_{2,\nu}(\mathbb{R}[+\theta/2]\mathbf{k}) \end{pmatrix}, \tag{S4}$$

where $\mathbb{R}[\theta]$ is the rotation matrix that implements counterclockwise rotation by angle θ.

At the interface between the trilayer and bilayer substacks, a long-wavelength moiré pattern is formed; this pattern modulates the hopping amplitude between regions of different local atomic stacking registry. Denoting the microscopic primitive translation vectors and primitive reciprocal lattice vectors as

$$\mathbf{a}_1 = a\,(1, 0)\,, \quad \mathbf{a}_2 = a\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \quad \mathbf{b}_1 = \frac{4\pi}{\sqrt{3}a}\left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right), \quad \text{and} \quad \mathbf{b}_2 = \frac{4\pi}{\sqrt{3}a}\,(0, 1)\,, \tag{S5}$$

the moiré reciprocal lattice vectors are defined as $\mathbf{G}_i = (\mathbb{R}[-\theta/2] - \mathbb{R}[+\theta/2])\,\mathbf{b}_i$:

$$\mathbf{G}_1 = \frac{4\pi}{\sqrt{3}L}\left(-\frac{1}{2}, -\frac{\sqrt{3}}{2}\right) \quad \text{and} \quad \mathbf{G}_2 = \frac{4\pi}{\sqrt{3}L}\,(1, 0)\,, \tag{S6}$$

where $L = a/2\sin(\theta/2) \approx a/\theta$ is the moiré period. The corresponding moiré lattice vectors $\mathbf{L}_1$ and $\mathbf{L}_2$ are shown in Fig. S1. This period diverges as $\theta \to 0$. In terms of these moiré reciprocal lattice vectors, the tunneling matrix between the trilayer and bilayer substacks is given in real space by

$$\mathbb{T}_\nu(\mathbf{r}) = \begin{pmatrix} \gamma_{\text{AA}} & \gamma_{\text{AB}} \\ \gamma_{\text{AB}} & \gamma_{\text{AA}} \end{pmatrix} + \begin{pmatrix} \gamma_{\text{AA}} & \gamma_{\text{AB}}\omega^{-\nu} \\ \gamma_{\text{AB}}\omega^{\nu} & \gamma_{\text{AA}} \end{pmatrix} e^{-i\nu\mathbf{G}_1\cdot\mathbf{r}} + \begin{pmatrix} \gamma_{\text{AA}} & \gamma_{\text{AB}}\omega^{\nu} \\ \gamma_{\text{AB}}\omega^{-\nu} & \gamma_{\text{AA}} \end{pmatrix} e^{-i\nu(\mathbf{G}_1+\mathbf{G}_2)\cdot\mathbf{r}}, \tag{S7}$$

where $\omega = e^{2\pi i/3}$ and γ_{AA} and γ_{AB} are hopping amplitudes at AA and AB regions. Because of the possibility of lattice relaxation, $|\gamma_{\text{AA}}| \leq |\gamma_{\text{AB}}|$. In twisted bilayer graphene, $\gamma_{\text{AA}}/\gamma_{\text{AB}} = 0.8$ in a typical experimental setup [2]. Presumably, with more layers, it is more difficult to fully relax. We do not scrutinize lattice relaxation in this work. Instead, we leave the ratio

Parameter set name	γ_0	γ_1	γ_2	γ_3	γ_4	γ_{AA}	γ_{AB}	Description
`Param_one`	-3100	380	-15	290	141	100	120	Includes some lattice relaxation
`Param_two`	-3100	380	-15	290	141	120	120	No lattice relaxation
`Param_three`	-2700	380	-15	290	141	80	100	Includes some lattice relaxation, reduced Fermi velocity
`Param_four`	-2700	380	-15	290	141	100	100	No lattice relaxation, reduced Fermi velocity

Table I. **Hopping parameters for rhombohedral graphene multilayers.** All values are quoted in millielectronvolts. γ_0, γ_1, γ_2, γ_3, and γ_4 in `Param_one` and `Param_two` are taken from Ref. [1]. Results presented in the main text come from `Param_one` because we believe it contains values closest to those found in experiments.

γ_{AA}/γ_{AB} as an adjustable parameter in the modeling. The effect of lattice relaxation on the band structure can be an interesting future study. Again, the parameters we use in our study are listed in Table I. For `Param_one` and `Param_two`, we include the parameters from Ref. [1], which were deduced experimentally. They are close to the values from this experiment [3]. Therefore, we believe these two sets to represent parameters somewhat closer to experiments. For γ_{AB}, we take it to be roughly one-third of γ_1 since they are essentially the same term if the twist were not there. In experimental Ref. [2] and theoretical Refs. [4, 5], $\gamma_{AB} = 110$ meV, which is roughly what we use $\gamma_{AB} = 120$ meV. For `Param_three` and `Param_three`, we reduce $\gamma_0 = -2700$ meV and $\gamma_{AB} = 100$ meV to values closer to what density functional theory calculations predict [5–7]. It is worth mentioning parenthetically that we include γ_2, hopping between next-nearest layers, in the rhombohedral trilayer stack, but we do not include such a hopping across the twist interface. This is because we assume that the misalignment due to twist diminishes that already-weak tunneling significantly. By testing parameters from both DFT calculations and experiments, we demonstrate that our principle conclusions are robust against these deviations in parameters. Now, including only nearest-layer hopping across the twist interface, the Hamiltonian in real space now becomes

$$\mathbb{K}_{3+2,\nu}(\mathbf{r}) = \begin{pmatrix} \mathbb{K}_{3,\nu}(-i\mathbb{R}[-\theta/2]\nabla_\mathbf{r}) & \mathbb{T}_\nu(\mathbf{r}) \\ \mathbb{T}_\nu^\dagger(\mathbf{r}) & \mathbb{K}_{2,\nu}(-i\mathbb{R}[+\theta/2]\nabla_\mathbf{r}) \end{pmatrix}. \tag{S8}$$

Finally, we add to the Hamiltonian a vertical interlayer electrostatic potential (corresponding to the external displacement field):

$$\mathbb{D} = \Delta \begin{pmatrix} -2 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & 0 & +2 \end{pmatrix}_\ell \otimes \mathbb{1}_\sigma, \tag{S9}$$

where σ and ℓ denote sublattice and layer degrees of freedom respectively and Δ is the energy difference between adjacent layers. The energy difference between the outermost layers in this five-layer stack is $4|\Delta|$. This vertical displacement is an essential feature of our study since it functions as a tunable knob with which to control band flatness and isolation.

By themselves, rhombohedral graphene multilayers already do not respect C_{2z} symmetry. So, of course, TRMG_{3+2} also does not respect C_{2z} symmetry. Because the top and bottom stacks do not have the same number of layers, any possible symmetry exchanging layers from the trilayer with those of the bilayer, such as C_{2x} and C_{2y}, is violated, even when $\Delta = 0$. There are also no vertical mirror planes in this system. However, C_{3z} and time-reversal $\mathcal{T}$ symmetries are respected when both valleys are considered. All in all, this system has only a few symmetries. In our simulations, we ensure that both $\mathcal{T}$ and $\mathcal{C}_{3z}$ symmetries are numerically respected in every converged solution.

Throughout this work, we shall restrict ourselves to $\theta > 0$ without loss of generality. For $\theta < 0$, we can map it to a value in $\theta > 0$ using the following transformation $\{\theta, \nu, k_x, x\} \mapsto -\{\theta, \nu, k_x, x\}$. This is M_x mirror reflection. One can see that this is a good transformation in the following way. First, by noting that $\theta \mapsto -\theta$ leads to $\mathbf{G}_i \mapsto -\mathbf{G}_i$, we immediately find that $\{\theta, \nu\} \mapsto -\{\theta, \nu\}$ leaves $\nu\mathbf{G}_i$ invariant. However, $\nu \mapsto -\nu$ flips the two matrices involving ν in the hopping matrix of Eq. (S7). To switch them back, we implement $x \mapsto -x$ to exchange $\mathbf{G}_1$ and $\mathbf{G}_1 + \mathbf{G}_2$. This proves that the hopping matrix is invariant under this transformation. The diagonal elements are also invariant since $(\nu k_x - i k_y) e^{\pm i\nu\theta/2}$ is manifestly invariant under this transformation. Thus, every state in valley ν at momentum (k_x, k_y) and twist angle θ has a partner in valley $-\nu$ at momentum $(-k_x, k_y)$ and twist angle $-\theta$. This immediately implies that the bands at $-\theta$ have the same spectral and topological properties as those at θ. So, we do not need to assess $-\theta$ separately.
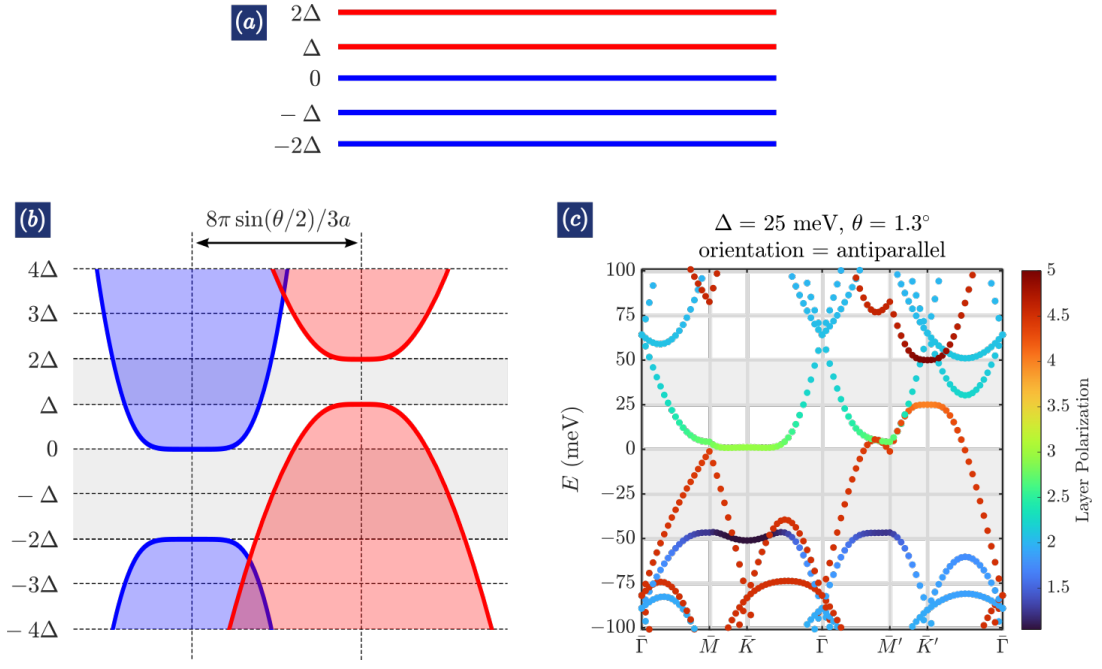
Figure S2. **Band alignment of a twisted stack without hopping across the twist interface.** (a) Real-space distribution of the on-site energies induced by an interlayer displacement field. (b) Schematic representation of the band alignment of the trilayer and bilayer spectra showing the absence of a global energy gap even the local spectrum of each substack is individually gapped. (c) Band structure with trigonal warping terms (γ_i from `Param_one`) but without hopping across the twist interface ($\gamma_{AA} = \gamma_{AB} = 0$). Here, states belonging to one substack as opposed to another are identified by their layer polarization. States belonging to the trilayer have layer polarization between 1 and 3, while those of the bilayer have layer polarization between 4 and 5.

II. NON-INTERACTING PHASE DIAGRAMS

A. Band Alignment Considerations

In this section, we study the phase diagrams calculated in the non-interacting limit. Without interactions, the four isospin flavors are necessarily degenerate; thus, it is sufficient to inspect just the $\{\nu = +, s = \uparrow\}$ flavor. To begin, it is instructive to examine the band alignment of the two substacks in a minimal model that neglects all trigonal warping terms as well as hopping across the twist interface. In this case, the two substacks are completely independent from each other. At this level of simplicity, the discussion is valid for both the antiparallel and parallel configurations. Because of the relative twist, the band structures of the substacks are shifted away from each other by $8\pi \sin(\theta/2)/3a$ in momentum space. Near their respective rotated zone corners, the energy spectra of the two substacks are [8]

$$\varepsilon_3(\mathbf{k}) = -\Delta \pm \sqrt{\Delta^2 + \frac{(\hbar v_0 |\mathbf{k}|)^6}{\gamma_1^4}}, \tag{S10a}$$

$$\varepsilon_2(\mathbf{k}) = \frac{3}{2}\Delta \pm \sqrt{\frac{\Delta^2}{4} + \frac{(\hbar v_0 |\mathbf{k}|)^4}{\gamma_1^2}}. \tag{S10b}$$

The trilayer spectrum is centered at $-\Delta$ with a gap of $2|\Delta|$ and disperses away its extrema as $\pm|\mathbf{k}|^3$, while the bilayer spectrum is centered at $3\Delta/2$ with a gap of $|\Delta|$ and disperses away from its extrema as $\pm|\mathbf{k}|^2$. Importantly, for $\Delta > 0$, the conduction band minimum of the trilayer is located at $E = 0$ while the valence band maximum of the bilayer is located at $E = \Delta$. Therefore, the two spectra overlap, and there are no global gaps that result from local gap openings due to Δ because of the spectral shifts of the two substacks. For $\Delta < 0$, the valence band maximum of the trilayer is at $E = 0$, while the conduction band minimum is at $E = \Delta$. Thus, the same conclusion applies that there are no resulting global gaps. This is illustrated schematically in Fig. S2(b). This basic observation survives even if the trigonal warping terms are switched back on, as shown in Fig. S2(c). In this
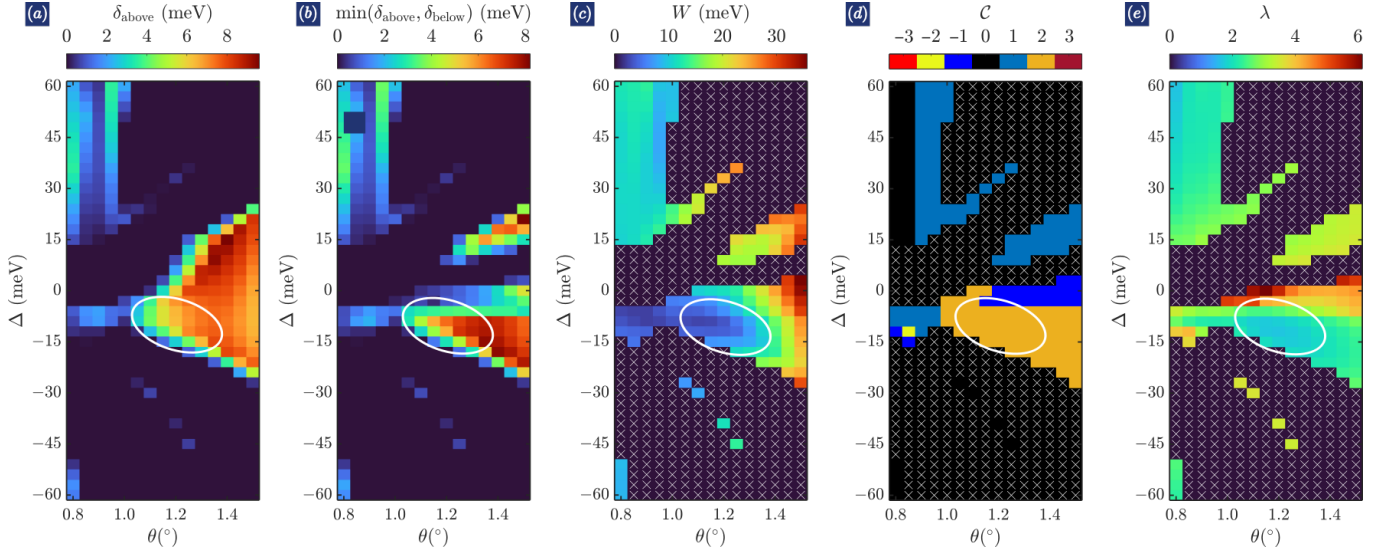
Figure S3. **Non-interacting phase diagram of the antiparallel stacking order with lattice relaxation.** (a) Value of the global gap above the first conduction band. (b) Global gap of the first conduction band defined as the minimum of the gap below and above that band. (c) Bandwidth of the first conduction band in cases where it is isolated, *i.e.* $\min(\delta_{\mathrm{above}}, \delta_{\mathrm{below}}) > 0$. (d)-(e) Chern number and trace condition of isolated first conduction bands. × denotes points where the gap is zero, for which the task of defining an associated bandwidth, Chern number, or trace condition violation either makes no sense or is useless. `Param_one` is used for the results reported here.

band structure, we identify a state $|\Psi_{\mathbf{k}}\rangle$ as belonging to one substack as opposed to another according to its layer polarization

$$\chi_\ell = \sum_{\ell=1}^{5} \ell \langle \Psi_{\mathbf{k}} | \mathbb{P}_\ell | \Psi_{\mathbf{k}} \rangle, \tag{S11}$$

where $\mathbb{P}_\ell$ is the layer projector matrix. If $1 \leq \chi_\ell \leq 3$, the state belongs to the trilayer sector; if $4 \leq \chi_\ell \leq 5$, the state belongs to the bilayer sector. From this simple exercise, we observe that the displacement field plays a minimal role in generating robust gaps in our system of twisted rhombohedral multilayers despite the fact that the displacement field generates significant gaps in each of the substack individually. Consequently, global gaps in our system are produced predominantly by hybridization across the twist interface, demonstrating the importance of the moiré potential in our work.

We now turn on the interlayer hopping across the twist interface. This generically produces anticrossings where the bands approach each other in energy unless otherwise prohibited by a symmetry. Because our system belongs to a small symmetry group, there are not many, if any, protected band crossings. This is the origin of many gaps, sometimes global, in the band structure. We are interested in the formation of *global* gaps at charge neutrality and above the first conduction band. Wherever these two gaps are significantly larger than the bandwidth of the first conduction band, that band can be treated to a good approximation as an isolated band wherein physics driven purely by electron-electron interactions may be qualitatively important. Even if there is no global gap at charge neutrality, having a gap above the first conduction is sometimes enough to stabilize an anomalous quantum Hall effect at filling of one electron per moiré unit cell. Thus, we are also interested in situations where there are only global gaps above the first conduction band. However, in these cases, it is not expected that interaction-driven physics can be ascertained by projecting to just a single isolated conduction band. Nonetheless, by considering a subset of nearby bands, we can still ascertain whether this situation can give rise to a symmetry-broken phase in the presence of Coulomb interactions. Given these considerations, we first inspect the spectral gap formation in the band structures. For our system of interest, twist angle θ and interlayer displacement field Δ are the two experimentally tunable parameters. So, we scan the phase space of θ vs Δ looking for these gaps.

B. Antiparallel Stacking Order

1. With Some Lattice Relaxation

For the antiparallel stacking order simulated using `Param_one`, which includes some lattice relaxation controlled by a non-unity ratio $\gamma_{\mathrm{AA}}/\gamma_{\mathrm{AB}}$, the phase diagram results are shown in Fig. S3. Here, δ_{above} is the global gap between the first conduction band and second conduction band while δ_{below} is the global gap between the first conduction band and the first valence band.

δ_{above} is plotted in Fig. S3(a) where it is shown that for small $\theta < 1°$, there is a region in parameter space at large positive $\Delta > 15$ meV where there exists small gaps $\lesssim 4$ meV above the first conduction band. More importantly, at larger $\theta > 1°$, there is a region surrounding $\Delta = 0$ meV where δ_{above} can be as large as 9 meV. This region grows in area as θ increases. To search for an anomalous Hall phase, δ_{above} is a useful metric to suggest the possibility of symmetry breaking once interactions are accounted for. To find isolated bands, we further need δ_{below} to be non-zero. The gap associated with a band is defined as the lesser of δ_{above} and δ_{below}. This is plotted in Fig. S3(b). As demonstrated here, only a subset of the nonzero region in Fig. S3(a) remains nonzero in Fig. S3(a) since the complement has a vanishing δ_{below}. We find the region of the largest $\min(\delta_{\mathrm{above}}, \delta_{\mathrm{below}})$ occurs around $\Delta = 0$, starts from $\theta \sim 1.0°$, and enlarges as θ increases. An additional indicator of the tendency to stabilize interaction-driven phases is the bandwidth of the band nearest to chemical potential. In Fig. S3(c), we plot the bandwidth W of the first conduction band only for values of (Δ, θ) where $\min(\delta_{\mathrm{above}}, \delta_{\mathrm{below}}) > 0$. Where $\min(\delta_{\mathrm{above}}, \delta_{\mathrm{below}}) = 0$, we simply identify that point in parameter space as having $W = 0$. The smallest bandwidths for isolated bands occur near $\Delta = 0$ for $\theta < 1.3°$. This overlaps partially with the region of largest $\min(\delta_{\mathrm{above}}, \delta_{\mathrm{below}})$. Therefore, we identify this region, marked by a while oval in Fig. S3, as the most conducive to symmetry-broken phases based on non-interacting band structures. Next, we characterize the topology of the isolated bands. The results are in shown Fig. S3(d). We find a large region where the $\mathcal{C} = 2$ as well as regions where the $\mathcal{C} = \pm 1$. The $\mathcal{C} = 2$ region is particularly interesting because it overlaps in phase space with regions of small bandwidths and large band gaps. In addition to the Chern number, we also characterize the isolated bands by the trace condition, which is a dimensionless quantity defined as

$$\lambda = \frac{1}{2\pi} \int d^2\mathbf{k} \left[\mathrm{Tr} g_{\mu\nu}(\mathbf{k}) - |\mathcal{B}_z(\mathbf{k})| \right], \tag{S12}$$

$g_{\mu\nu}(\mathbf{k})$ and $\mathcal{B}_z(\mathbf{k})$ are the quantum metric and Berry curvature as defined in Sec. VIII. This trace condition measures how "close" a band is to the zeroth Landau level in terms its quantum geometry in $\mathbf{k}$-space. A small value of λ means that the trace of the quantum metric nearly saturates the bound $g_{\mu\nu}(\mathbf{k}) \geq |\mathcal{B}_z(\mathbf{k})|$. The trace condition is smallest also in the white oval region of Fig. S3(e). The smallest values are $\lambda \approx 1.7$ in that region. All in all, from a non-interacting phase diagram, we identify the white oval region as most promising for stabilizing interaction-driven symmetry-broken phases.

To gain insight about whether interactions would significantly alter the non-interacting phase diagram, we now investigate the charge distribution of the first conduction band. For concreteness, let us take $\theta = 1.25°$ since this angle contains many of the important qualitative features for the other angles as well. In Fig. S4, we plot the evolution of the band structures as Δ varies for $\theta = 1.25°$. At $\Delta = 0$ meV, there are two bands near charge neutrality isolated from the rest of the band structure: one band is occupied (valence band) and one band is empty (conduction band) at neutrality. These two bands have mixed layer character, meaning that there are states with χ_ℓ in the trilayer substack and those with χ_ℓ in the bilayer substack coexisting in the same band. As Δ is increased in the negative direction, these two bands separate, and the resulting conduction band acquires a Chern number of $\mathcal{C} = 2$. Recall that negative Δ is the regime where in the trilayer substack is shifted higher in energy than the bilayer substack in the absence of hopping across the twist interface. Therefore, for small negative Δ, the first conduction band is essentially polarized to the trilayer substack. This band initially becomes narrower in bandwidth as Δ grows in magnitude before the trend reverses as the the state at $\bar{K}$ is pushed upwards in energy to band invert with the second conduction band at larger Δ, causing the bandwidth to increase and the global gap surrounding the first conduction band to collapse. In the regime where $\mathcal{C} = 2$ and there is global gap surrounding the first conduction band, the Berry curvature peaks near the $\bar{\Gamma}$ point, but it also has non-negligible density spreading across the moire Brillouin zone as well. Going in the opposite direction to increase positive Δ, we also observe the first conduction getting pushed to higher energy. Sometimes, this conduction band is isolated global gaps and has a Chern number of $\mathcal{C} = +1$. However, there is no value of positive Δ for which its bandwidth is narrow. This is because this band primarily polarizes to the bilayer substack for which band flatness over a large region of momentum space emerges at higher values of Δ.

The above analysis reveals that in the $\mathcal{C} = 2$ phase, the average density is mostly in the trilayer substack. However, this is only part of the picture as variations of the charge density are not captured in the calculation of their average. To gain more insight, we now inspect the real-space charge density for $\Delta = -12$ meV and $\theta = 1.25°$ as an illustration. In Fig. S5, we map the charge density for each layer for states at the high-symmetry points $\bar{\Gamma}$, $\bar{K}$, and $\bar{K}'$. We observe significant charge redistribution moving from one $\mathbf{k}$ point to the next. At $\bar{\Gamma}$, most of the charge is concentrated at γ regions on layer 1 with some residual charge located on layers 2 and 5 at α regions. At $\bar{K}$, nearly all of the charge is located on layer 1 at γ regions. There is virtually no charge on the remaining four layers. At $\bar{K}'$, most of the charge is localized in α regions divided up among layers 1,3, and 4. In layer 2, there is some charge density in γ regions. Even if all the layers were treated as one, there would still be significant charge reshuffling laterally as one moves across the mBZ. Therefore, we expect Hartree-Fock (HF) renormalization to change the shape of this band. Moreover, there is also charge redistribution across the layers. Thus, we expect out-of-plane Coulomb interactions to contribute as well. As such, in order to accurately predict symmetry breaking in the presence of Coulomb interactions between charges, we take into consideration whether the charges are of the same layer or different layers, as detailed later.

Equipped with the real-space representation of the wave functions, it is straightforward to calculate the C_3 symmetry eigenvalues $\omega(\mathbf{k}_*)$ of wavefunctions:

$$\hat{C}_3 |\Psi_{\mathbf{k}_*}\rangle = \omega(\mathbf{k}_*) |\Psi_{\mathbf{k}_*}\rangle, \tag{S13}$$
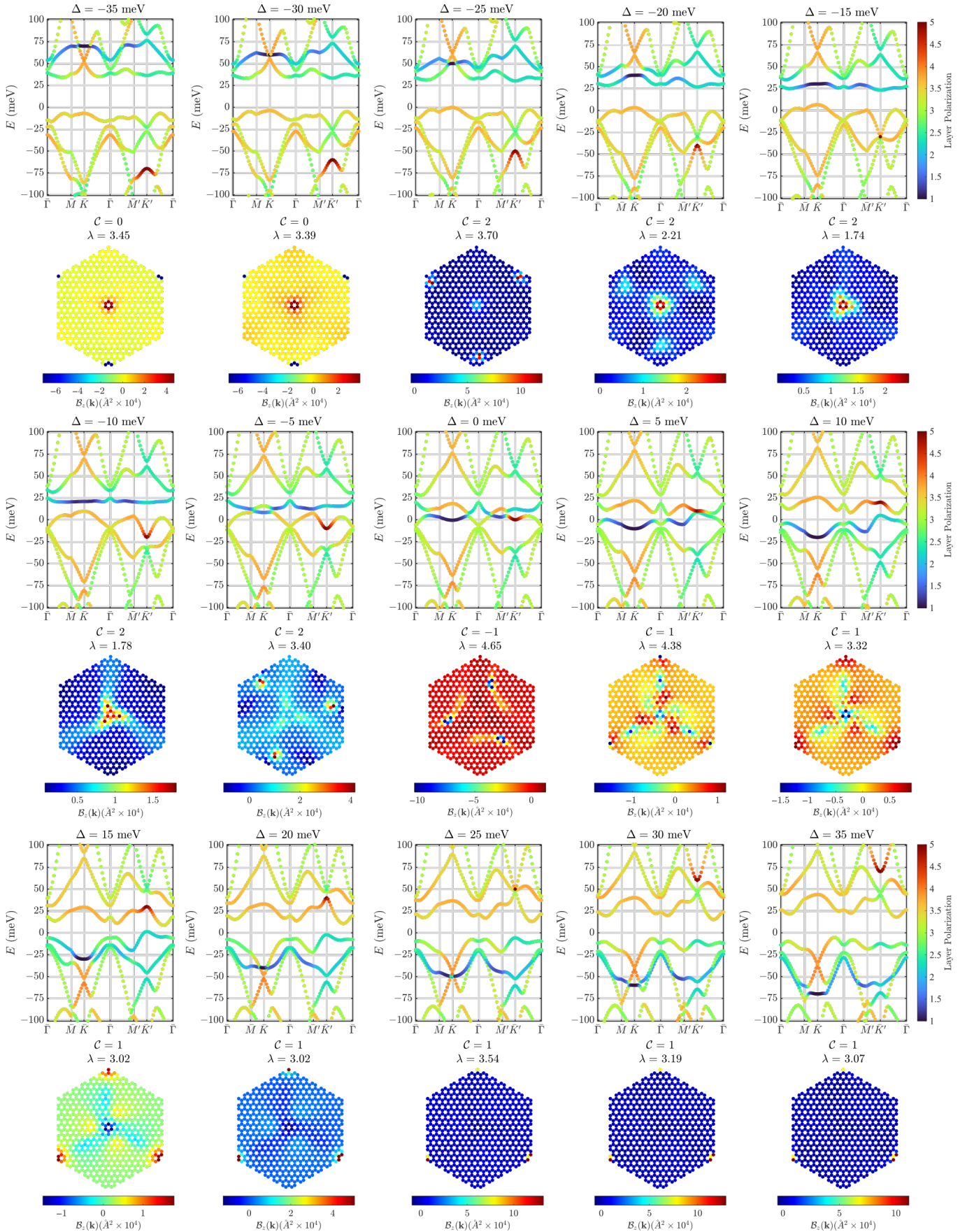
Figure S4. **Evolution of band structures as a function of Δ at $\theta = 1.25°$ for the antiparallel stacking order with lattice relaxation.** For each Bloch state, we color code it by its layer polarization. Below each band structure is shown the corresponding Berry curvature distribution of its first conduction band. The Chern number and trace condition violation for that band are also displayed. `Param_one` is used for the results reported here.
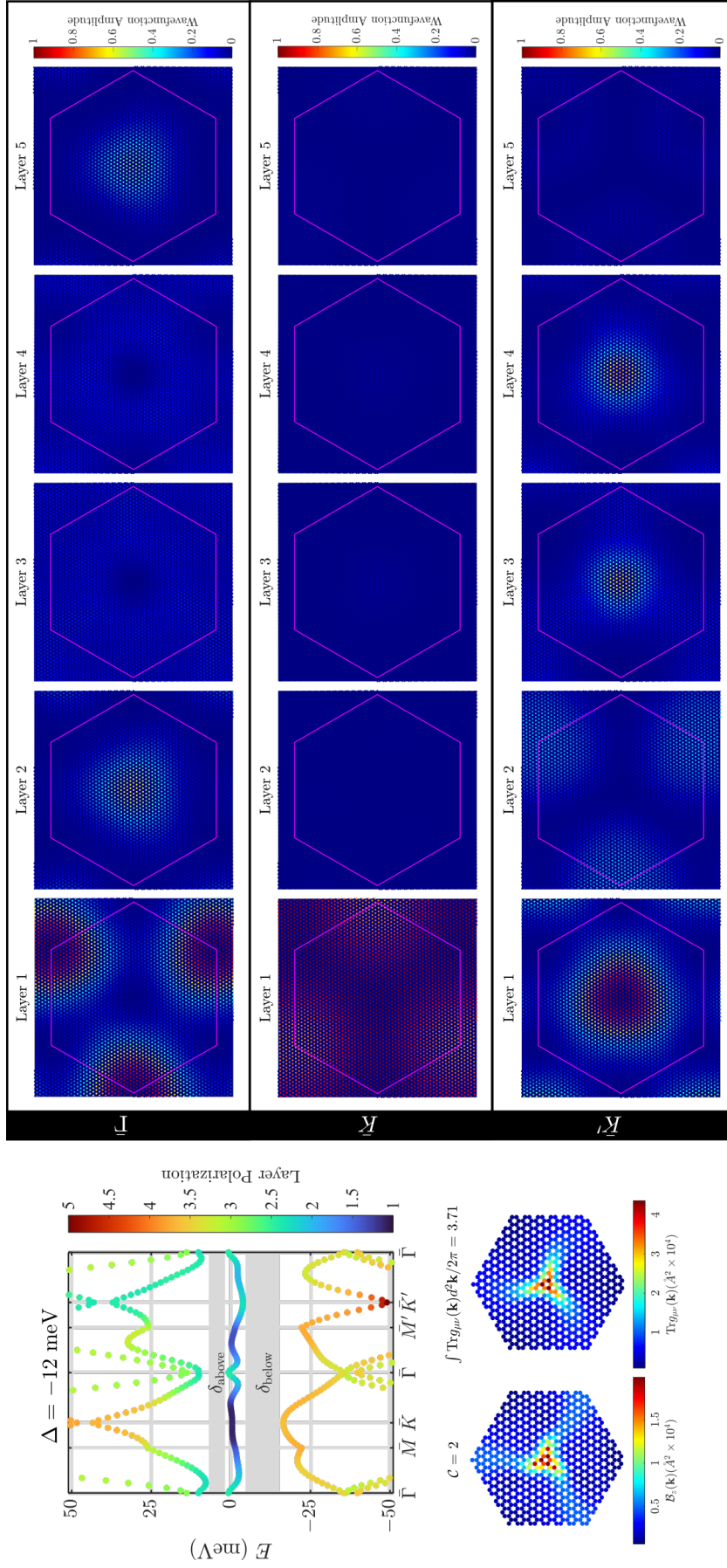
Figure S5. **Charge density of** $C = 2$ **band in the non-interacting limit with some lattice relaxation for the antiparallel configuration.** We show the charge density in each layer for states at the high-symmetry points $\bar{\Gamma}$, $\bar{K}$, and $\bar{K}'$ for the first conduction band that has Chern number $C = 2$. We also compute the C_{3z} symmetry eigenvalues: $\omega(\bar{\Gamma}) = 1$ and $\omega(\bar{K}) = \omega(\bar{K}') = \exp(-2\pi i/3)$. This confirms that the Chern number is indeed two. The quantum metric and Berry curvature are shown below the band structure. Param_one is used for the results reported here.

Figure S6. **Non-interacting phase diagram of the antiparallel stacking order without lattice relaxation.** (a) Value of the global gap above the first conduction band. (b) Global gap of the first conduction band defined as the minimum of the gap below and above that band. (c) Bandwidth of the first conduction band in cases where it is isolated, *i.e.* $\min(\delta_{\text{above}}, \delta_{\text{below}}) > 0$. (d)-(e) Chern number and trace condition of isolated first conduction bands. $\times$ denotes points where the gap is zero, for which the task of defining an associated bandwidth, Chern number, or trace condition violation either makes no sense or is useless. `Param_two` is used for the results reported here. Like the results reported in Fig. S3, the phase diagram here features a $\mathcal{C} = 2$ phase. However, unlike in the case with lattice relaxation, the band gaps here are much smaller.

where $\hat{C}_3$ is a three-fold rotation operator and $\mathbf{k}_*$ is a high-symmetry momentum that brings a wavefunction to itself (up to a phase) under C_3 rotation. Using the C_3 eigenvalues, we can verify the Chern number (modulo three) using the formula [9]

$$\omega(\bar{K})\omega(\bar{K}')\omega(\bar{\Gamma}) = \exp\left(-\frac{2\pi i \mathcal{C}}{3}\right). \tag{S14}$$

This serves as a check of the Chern number calculated by numerically integrating the local Berry curvature. As an example, for the states in Fig. S5, $\omega(\bar{\Gamma}) = 1$ and $\omega(\bar{K}) = \omega(\bar{K}') = \exp(-2\pi i/3)$, confirming indeed that $\mathcal{C} = 2$.

2. Without Lattice Relaxation

We now analyze the non-interacting phase diagram simulated using parameters in `Param_two`. This set of parameters includes no lattice relaxation across the twist interface since $\gamma_{\text{AA}} = \gamma_{\text{AB}} = 120$ meV. The phase diagram is shown in Fig. S6. Qualitatively, this phase diagram is generally the same as the phase diagram in Fig. S3 in the following ways. The two feature a $\mathcal{C} = 2$ phase starting around $\theta > 1.1°$. There are also $\mathcal{C} = \pm 1$ phases, but they appear at different regions in the two phase diagrams. On the other hand, the two phase diagrams differ in important ways as well. First, the Chern-nontrivial regions in parameter space are smaller in Fig. S6 compared to Fig. S3. Second, the band gaps are much smaller without lattice relaxation. Lattice relaxation tends to enhance the direct gap at $\bar{\Gamma}$, and therefore is better at spectrally isolating the Chern-nontrivial phases ($\mathcal{C} = 2$ phase in particular). Without lattice relaxation, the largest gaps we obtain in the $\mathcal{C} = 2$ phase are about 2 meV. Values like this are too small to well-isolate a band into which one can project the Coulomb interactions. Therefore, on the basis on non-interacting band structures, one might dismiss these systems without lattice relaxation as poor candidate for anomalous integer and fractional Chern insulating phases. However, as we shall show, non-interacting band structures are not necessarily a good predictor of the emergence of Chern phases stabilized by interactions. Therefore, we shall revisit this phase diagram after accounting for HF corrections in a later section. In addition to the fact that the first conduction band is poorly isolated from the other bands in the region of interest in parameter space, that band also has a relatively larger trace condition violation compared to the results with lattice relaxation. For the $\mathcal{C} = 2$ phase, we find that the smallest violation is around $\lambda \approx 2.7$ (with lattice relaxation, this violation has smallest value around $\lambda \approx 1.7$). All in all, it does not appear from a non-interacting analysis that the bands calculated without lattice relaxation would be good candidates for anomalous Chern phases.

To gain more insights into the small band gaps, we next inspect the band structures at a fixed $\theta = 1.25°$. The results are shown in Fig. S7. Starting at $\Delta = 0$ meV, we already see that the bands are not well separated from each other near charge neutrality, unlike in the situation with corrugation. In particular, the $\bar{\Gamma}$ point is barely gapped. As we increase Δ in the negative direction,

the two bands at charge neutrality spectrally disentangle. However, the first conduction band is never well separated from the higher-energy continuum, with its $\bar{\Gamma}$ point very close to the next energy band. This is in contrast to the situation with lattice relaxation where there is a region of negative Δ where this band is reasonably well isolated from the rest of the band structure. However, it is still true in the present case without lattice relaxation that this first conduction band is quite flat and is localized, on average, mostly in the trilayer substack. In some ways, this is reminiscent of the situation in pentalayer graphene aligned with hexagonal boron nitride where the narrow bottom of the conduction band is not separated from the rest of the energy spectrum. In the positive Δ region, the first band conduction is seldom flat or isolated but it does feature a Chern number of one. In this way, the positive Δ region here is qualitatively the same as the positive Δ region in the case with lattice relaxation. We have also inspected the charge density distribution across the layers. In Fig. S8, we show an example from $\theta = 1.25°$ and $\Delta = -10$ meV. The general patterns look very much like those in Fig. S5. In particular, the symmetry eigenvalues at the high-symmetry points remain unchanged and they multiply to give the correct Chern number.

3. With Reduced Fermi Velocity γ_0

To close our considerations of the non-interacting phase diagrams for the antiparallel configuration, we show in Figs. S9 and S10 the phase diagrams with a reduced $\gamma_0 = 2700$ meV with and without lattice relaxation respectively. Here, we use parameters from `Param_three` and `Param_four` respectively. Overall, there Chern-two phase remains robust on all cases. Compared to having lattice relaxation, the phase diagrams without lattice relaxation feature much smaller gaps and higher trace condition violation. This is true for $\gamma_0 = 3100$ meV and $\gamma_0 = 2700$ meV. The Chern-two phase of interest remains around $\theta = 1.25°$ and $\Delta = -10$ meV with reduced γ_0. This analysis inspires confidence that qualitative conclusions about the Chern-two phase are robust against changes in modeling parameters since using both parameters experimentally determined and calculated from density functional theory methods, we obtain basically the same results.

C. Parallel Stacking Order

1. With Some Lattice Relaxation

We now examine the phase diagram for the parallel configuration. To start, we consider the results simulated with `Param_one` shown in Fig. S11. This phase diagram is richer than the antiparallel configuration's phase diagrams in the sense that it features isolated Chern bands with $|\mathcal{C}| = 1, 2, 3$. Referring to Fig. S11, we find that the $\mathcal{C} = 3$ and $\mathcal{C} = 2$ phases exist exclusively in the $\Delta < 0$ region. The latter develops prominently for $\theta > 1°$, while the latter exists around $\theta \approx 1°$. However, the $\mathcal{C} = 2$ phase features simultaneously small bandwidths and band gaps, on the order of 2 meV or so. On the other hand, the $\mathcal{C} = 3$ phase exists with larger bandwidths and band gaps. Part of the reason is simply the fact that the $\mathcal{C} = 3$ generically exists at larger angles where the bands disperse more due to a larger mBZ (because of a smaller real-space unit cell). Since the $\mathcal{C} = 3$ phase is qualitatively different compared to the antiparallel configuration, we focus on this phase more carefully. The evolution of the band structures at $\theta = 1.25°$ as Δ varies ins shown in Fig. S12. Based on these spectra alone, there does not appear to be much qualitative difference between the parallel configuration and the antiparallel configuration at the same values in parameter space. We obverse the following features that are also present in the corresponding antiparallel stacking order. As Δ is increased in the negative direction, the first conduction separates from the two-band manifold at charge neutrality and migrates upward in energy. Before it merges into the continuum, that band is spectrally isolated and has a Chern number of $\mathcal{C} = 3$. The density of that band is, on average, localized on the trilayer substack. The Berry curvature peaks around the $\bar{\Gamma}$ point in the mBZ, but it has appreciable spread throughout the mBZ. It is worth emphasizing that the trace condition violation in this case is much smaller than in the antiparallel configuration. The best value is around 0.7, which is about half the trace condition violation of the antiparallel configuration. On the opposite side of Δ, the first conduction band is mostly localized on the bilayer substack and is never really well-isolated from the rest of the band structure or narrow in bandwidth.

Next, we inspect the real-space charge density of $\theta = 1.25°$ and $\Delta = -10$ meV as an example. We map the charge density for each layer for the three C_{3z}-symmetric wavefunctions, shown in Fig. S13. For the $\bar{\Gamma}$ and $\bar{K}$ states, the charge density is basically the same as that in Fig. S5. For the $\bar{\Gamma}$ state, the charge is localized mostly at γ regions on layer 1 with a smaller concentration of charge on layer 2 at α regions. For the $\bar{K}$ state, almost all of the charge is located on layer 1, surrounding γ regions. Both of these states transform in the same way as in the antiparallel configuration. In particular, $\omega(\bar{\Gamma}) = 1$ and $\omega(\bar{K}) = \exp(-2\pi i/3)$. On the contrary, the charge density for the $\bar{K}'$ state is completely different to that of the antiparallel case. In the present parallel case, the charge density is localized at γ regions of layer 1, with residual charge density located at β regions of layer 2 and α regions on layer 3 and layer 4. The symmetry eigenvalue of this state is $\omega(\bar{K}') = \exp(+2\pi i/3)$. This change in the charge distribution at the $\bar{K}'$ point is responsible for changing the Chern number from two to three.
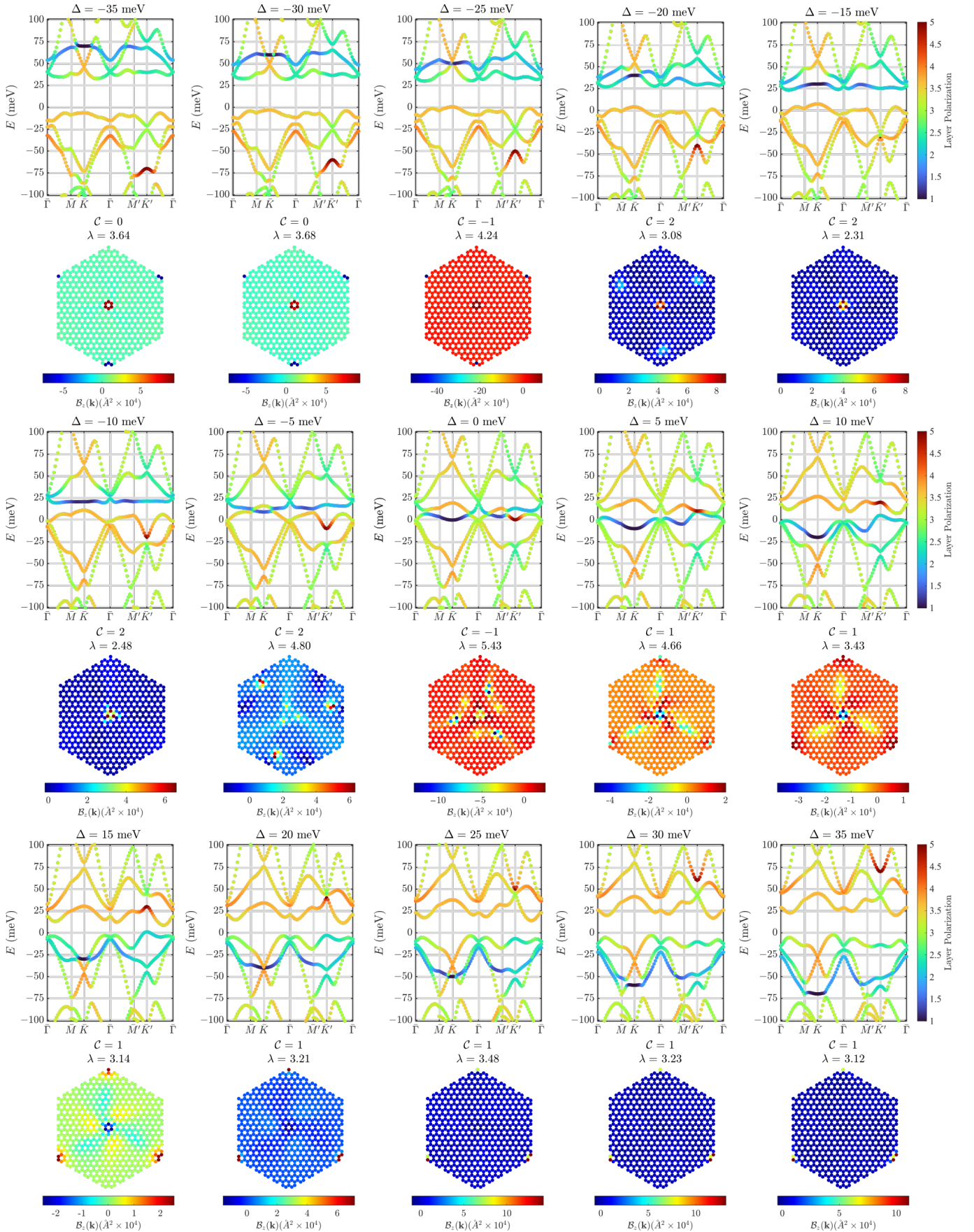
Figure S7. **Evolution of band structures as a function of** Δ **at** $\theta = 1.25°$ **for the antiparallel stacking order without lattice relaxation.** For each Bloch state, we color code it by its layer polarization. Below each band structure is shown the corresponding Berry curvature distribution of its first conduction band. The Chern number and trace condition violation for that band are also displayed. `Param_two` is used for the results reported here.
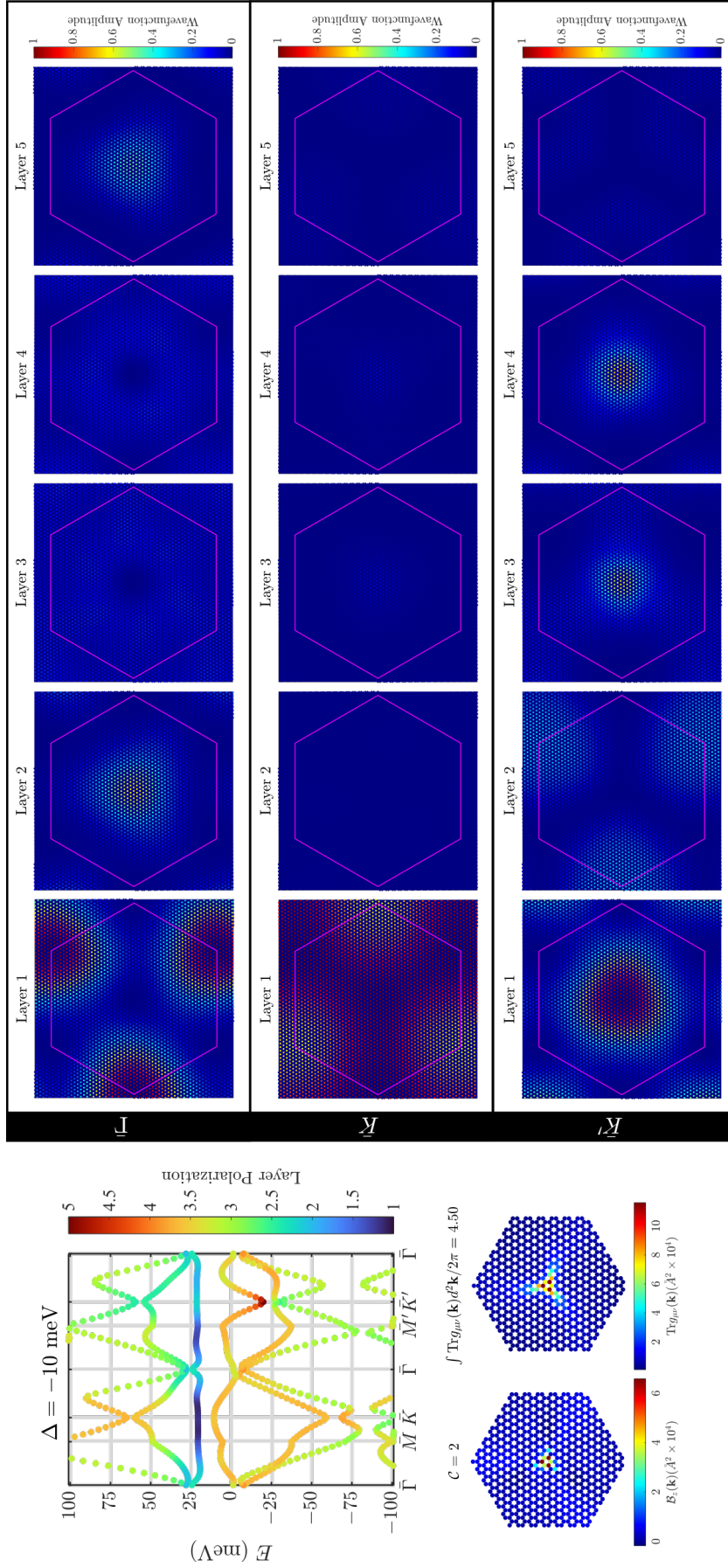
Figure S8. **Charge density of** $C = 2$ **band in the non-interacting limit without lattice relaxation for the antiparallel configuration.** From the band structure, we observe that the first conduction band is not well isolated from its surrounding high-energy bands. We show the charge density in each layer for states at the high-symmetry points $\bar{\Gamma}$, $\bar{K}$, and $\bar{K}'$ for the first conduction band that has Chern number $C = 2$. We also compute the C_{3z} symmetry eigenvalues: $\omega(\bar{\Gamma}) = 1$ and $\omega(\bar{K}) = \omega(\bar{K}') = \exp(-2\pi i/3)$. This confirms that the Chern number is indeed two. The quantum metric and Berry curvature are shown below the band structure. Param_two is used for the results reported here.
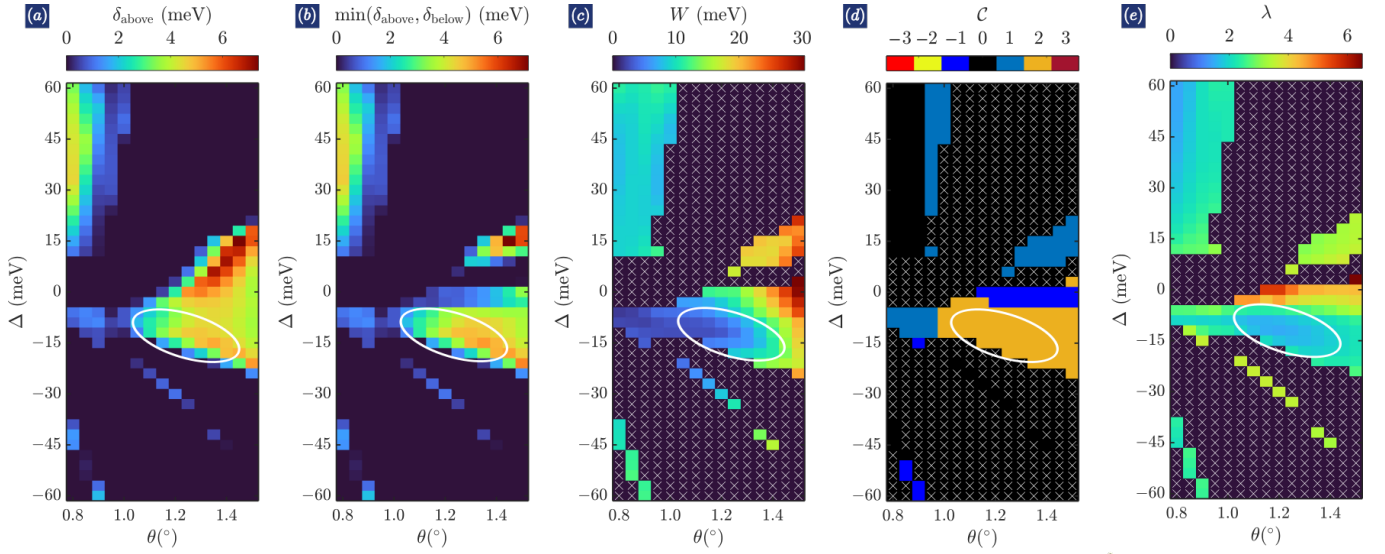
Figure S9. **Non-interacting phase diagram of the antiparallel stacking order with reduced γ_0 and with lattice relaxation.** (a) Value of the global gap above the first conduction band. (b) Global gap of the first conduction band defined as the minimum of the gap below and above that band. (c) Bandwidth of the first conduction band in cases where it is isolated, *i.e.* $\min(\delta_{\mathrm{above}}, \delta_{\mathrm{below}}) > 0$. (d)-(e) Chern number and trace condition of isolated first conduction bands. $\times$ denotes points where the gap is zero, for which the task of defining an associated bandwidth, Chern number, or trace condition violation either makes no sense or is useless. `Param_three` is used for the results reported here. Like the results reported in Fig. S3, the phase diagram here features a $\mathcal{C} = 2$ phase. However, unlike in the case with a larger γ_0 and with lattice relaxation, the band gaps here are smaller.

Figure S10. **Non-interacting phase diagram of the antiparallel stacking order with reduced γ_0 and without lattice relaxation.** (a) Value of the global gap above the first conduction band. (b) Global gap of the first conduction band defined as the minimum of the gap below and above that band. (c) Bandwidth of the first conduction band in cases where it is isolated, *i.e.* $\min(\delta_{\mathrm{above}}, \delta_{\mathrm{below}}) > 0$. (d)-(e) Chern number and trace condition of isolated first conduction bands. $\times$ denotes points where the gap is zero, for which the task of defining an associated bandwidth, Chern number, or trace condition violation either makes no sense or is useless. `Param_four` is used for the results reported here. Like the results reported in Fig. S9, the phase diagram here features a $\mathcal{C} = 2$ phase. However, unlike in the case with a larger γ_0 and with lattice relaxation, the band gaps here are smaller. Also, the region of $\mathcal{C} = 2$ is much smaller. This trend mirrors the trend seen in Fig. S6 with a larger γ_0.
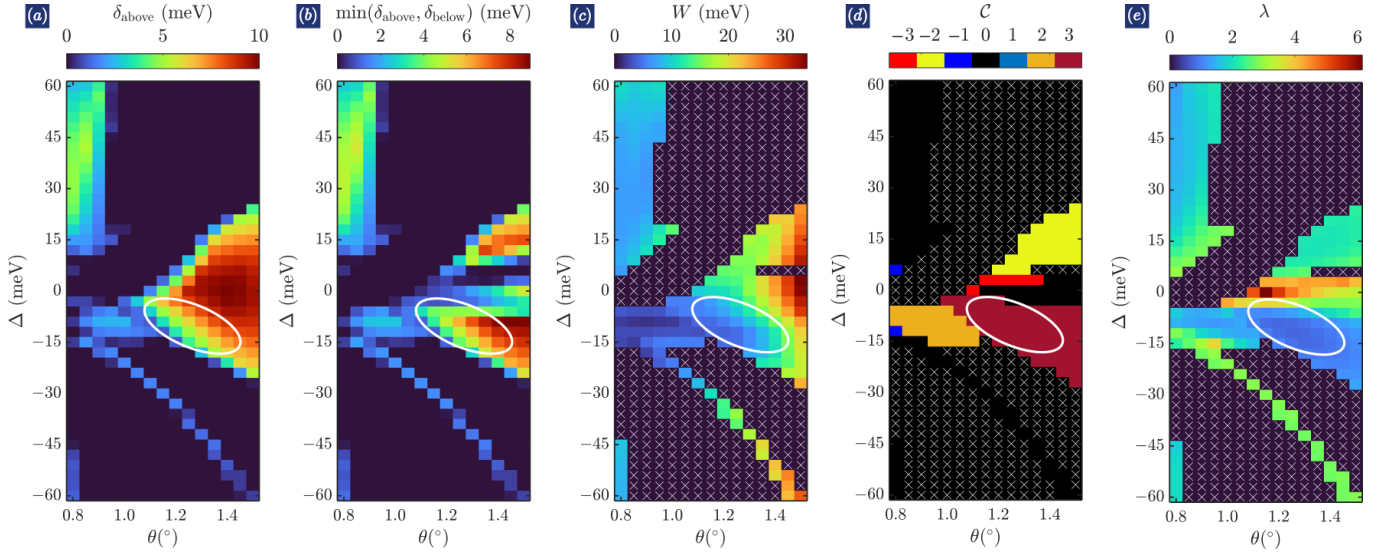
Figure S11. **Non-interacting phase diagram of the parallel stacking order with lattice relaxation.** (a) Value of the global gap above the first conduction band. (b) Global gap of the first conduction band defined as the minimum of the gap below and above that band. (c) Bandwidth of the first conduction band in cases where it is isolated, *i.e.* $\min(\delta_{\text{above}}, \delta_{\text{below}}) > 0$. (d)-(e) Chern number and trace condition of isolated first conduction bands. There is a large region with $\mathcal{C} = 3$ and small trace condition violation. $\times$ denotes points where the gap is zero, for which the task of defining an associated bandwidth, Chern number, or trace condition violation either makes no sense or is useless. `Param_one` is used for the results reported here.

2. Without Lattice Relaxation

Without lattice relaxation, $\gamma_{\text{AA}} = \gamma_{\text{AB}} = 120$ meV, the phase diagram is shown in Fig. S14. The first obvious difference here compared to Fig. S11 is the significant reduction in the $\mathcal{C} = 3$ region. While it is still there, it is now mostly succeeded by a $\mathcal{C} = 2$ phase as Δ is increased in the negative direction at a fixed θ. The band gaps in general are also smaller in Fig. S14 compared to Fig. S11. Also, the trace condition violation is larger without lattice relaxation when compared to values obtained in the case with lattice relaxation. It is worth mentioning that the $\mathcal{C} = 2$ bands also exist with lattice relaxation too, as can be seen in Fig. S12, but they typically appear without a fully developed global gap above. So they do not show up in the phase diagram, which only displays a Chern phase if it has a clear gap above and below. All of these trends are also observed in the antiparallel configuration when comparing results with and without lattice relaxation. Inspecting the band structures from Fig. S15, we observe the main difference here compared to Fig. S12 is the fact that the first conduction band is barely isolated from the other bands for the values of Δ shown. Take $\Delta \approx -10$ meV as an example, we observe that the $\bar{\Gamma}$ point is almost touching the next higher-energy band. Because of this, the Berry curvature is almost singular at the $\bar{\Gamma}$ point. Again, this is also observed in the antiparallel configuration. So it appears that the general effects of reducing lattice relaxation are to diminish the regions of nontrivial Chern phases south of $\Delta = 0$ meV, decrease the bandwidths of the remaining topologically nontrivial regions, and exaggerates the corresponding trace condition violation. Nonetheless, it is worth emphasizing again that although the $\mathcal{C} = 3$ is significantly diminished in phase space in the absence of corrugation effects, it is still present. Once interactions are switched on at the mean-field level, it may be stabilized.

3. With Reduced Fermi Velocity γ_0

Next, we present the phase diagrams for the cases with a reduced Fermi velocity γ_0. The results with relaxation are reported in Fig. S16, while the results without relaxation are reported in Fig. S17. With relaxation, the phase diagram is similar to the results in Fig. S11. There is a large region with $\mathcal{C} = 3$ and quite small λ. The smallest trace condition violation is around 0.75. However, the reduced γ_0 leads to relatively smaller global band gaps in this nontrivial phase. This is observed in the antiparallel configuration too where the reduction in the Fermi velocity generally leads to a decrease in the global band gaps. Moving onto the results without relaxation, we see here that the $\mathcal{C} = 3$ phase almost completely disappears, leaving only a robust $\mathcal{C} = 2$ phase that has a decently small trace condition violation $\lambda \approx 1.4$. Recalling Fig. S14, we see that the effect of removing lattice relaxation is to decrease the region in phase space where the first conduction band is an isolated, narrow, and $\mathcal{C} = 3$ band. That reduction is more dramatic here with $\gamma_0 = 2700$ meV than in the case where $\gamma_0 = 3100$ meV. In fact, the reduction in the $\gamma_0 = 2700$ meV case is so significant that we cannot conclude with confidence that the $\mathcal{C} = 3$ exists here at all. The phase area
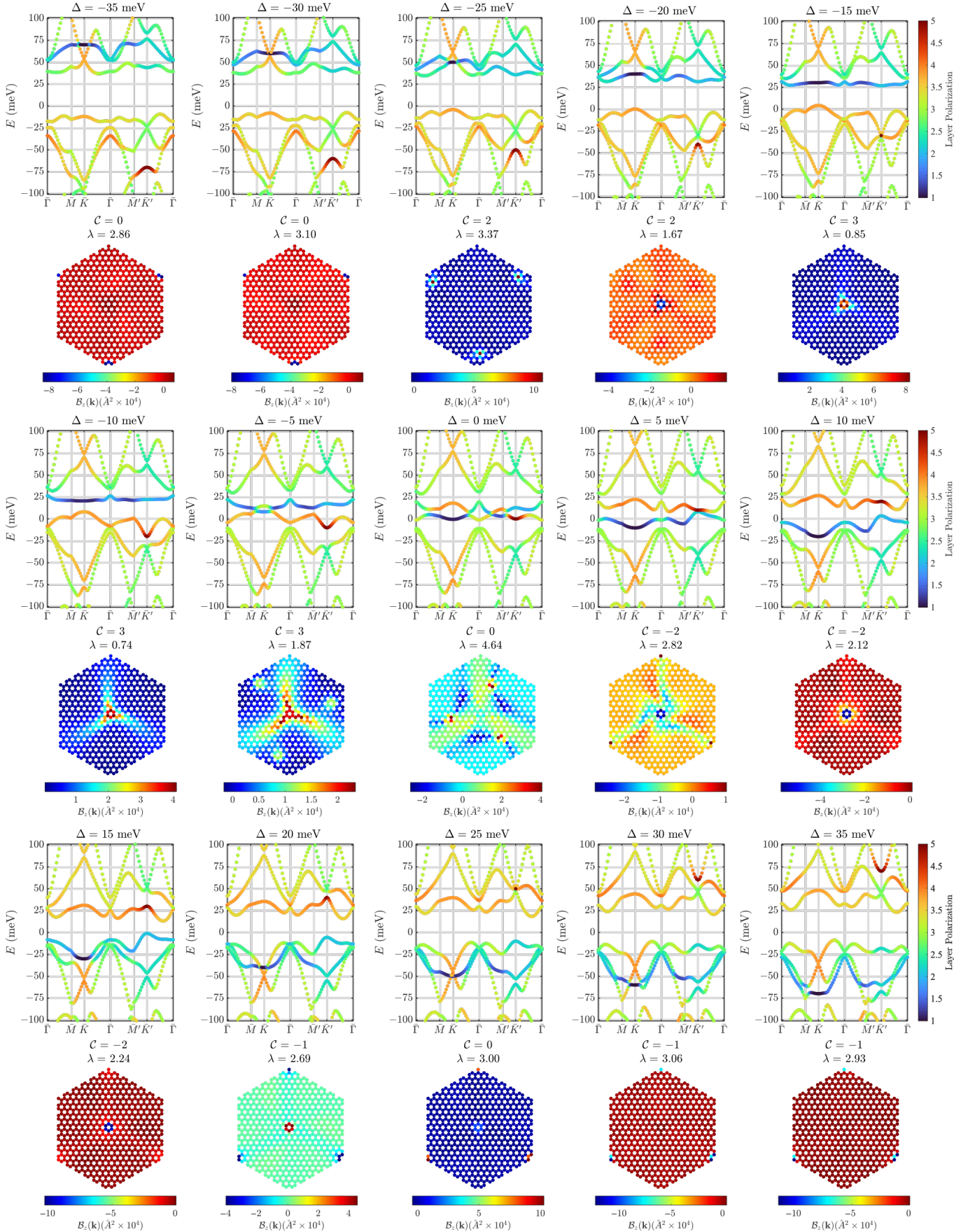
Figure S12. **Evolution of band structures as a function of Δ at $\theta = 1.25°$ for the parallel stacking order.** For each Bloch state, we color code it by its layer polarization. Below each band structure is shown the corresponding Berry curvature distribution of its first conduction band. The Chern number and trace condition violation for that band are also displayed. `Param_one` is used for the results reported here.
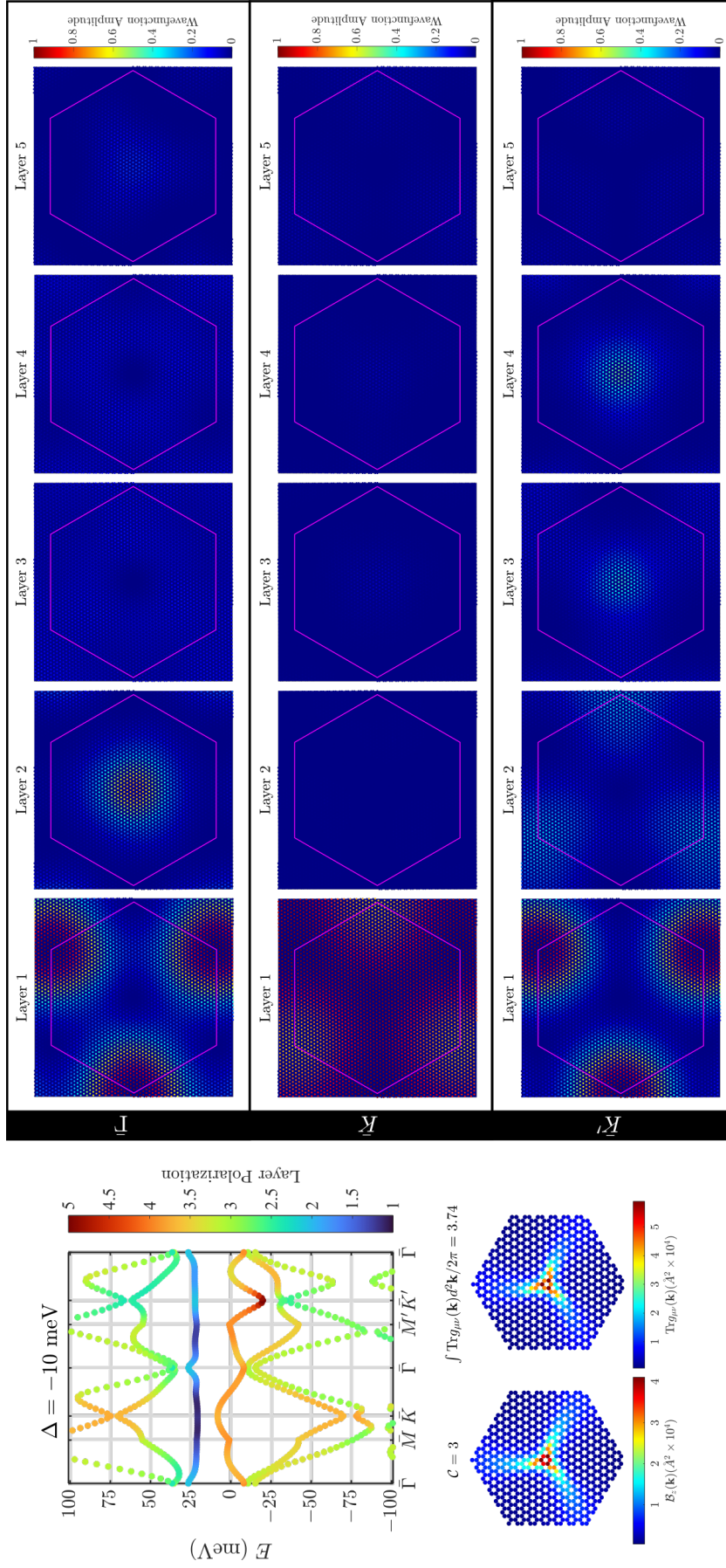
Figure S13. **Charge density of $C = 3$ band in the non-interacting limit with some lattice relaxation for the parallel configuration.** We show the charge density in each layer for states at the high-symmetry points $\bar{\Gamma}$, $\bar{K}$, and $\bar{K}'$ for the first conduction band that has Chern number $C = 3$. We also compute the C_{3z} symmetry eigenvalues: $\omega(\bar{\Gamma}) = 1$, $\omega(\bar{K}) = \exp(-2\pi i/3)$, and $\omega(\bar{K}) = \exp(+2\pi i/3)$. This confirms that the Chern number is indeed three. The quantum metric and Berry curvature are shown below the band structure. Param_one is used for the results reported here.
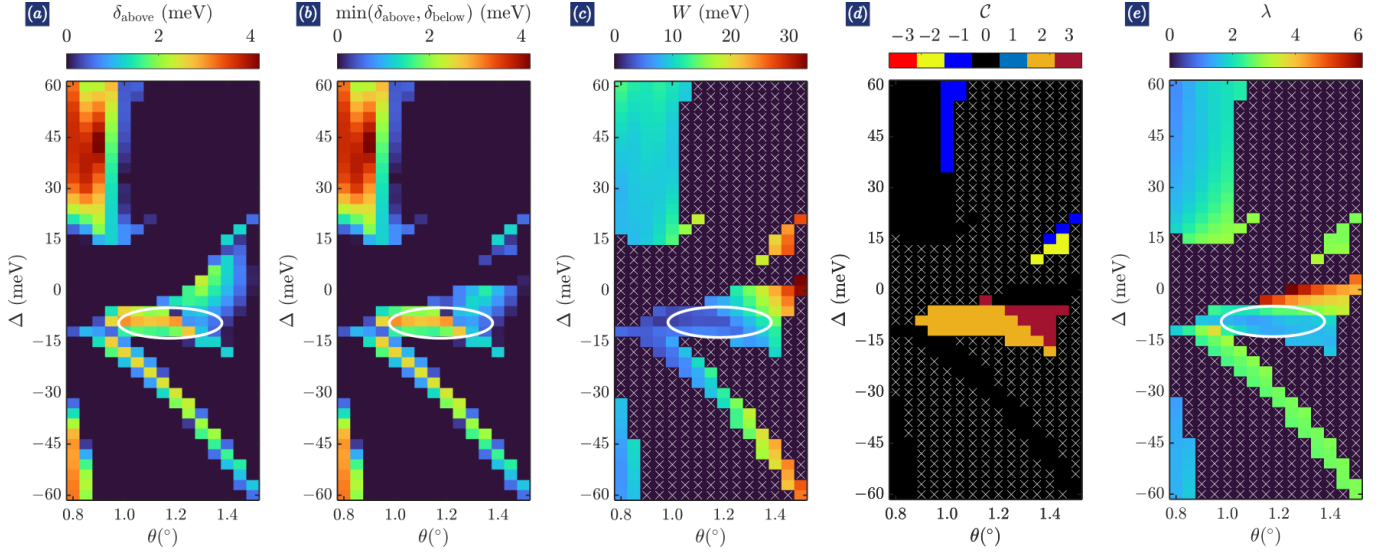
Figure S14. **Non-interacting phase diagram of the parallel stacking order without lattice relaxation.** (a) Value of the global gap above the first conduction band. (b) Global gap of the first conduction band defined as the minimum of the gap below and above that band. (c) Bandwidth of the first conduction band in cases where it is isolated, $i.e.$ $\min\left(\delta_{\text{above}}, \delta_{\text{below}}\right) > 0$. (d)-(e) Chern number and trace condition of isolated first conduction bands. There is a small region with $\mathcal{C} = 3$ and relatively small trace condition violation. $\times$ denotes points where the gap is zero, for which the task of defining an associated bandwidth, Chern number, or trace condition violation either makes no sense or is useless. `Param_two` is used for the results reported here.

is so small that it could just be a numerical artifact due to fluctuation of the Chern number near a band inversion.

D. Summary of Observations

In the preceding sections, we have shown the phase diagrams for the antiparallel and parallel configurations using different sets of parameters that include or not include the effects of corrugation as well as having a larger or smaller γ_0. While there are some qualitatively important differences between the cases, there are clearly also some general conclusions that are insensitive to microscopic details. In this section, we summarize those robust conclusions. To aid in that effort, we recognize that the ample amount of data presented calls for the introduction of a figure of merit that can quantitatively capture the region in phase space most desirable. For the antiparallel configuration, let us focus only on the $\mathcal{C} = 2$ phase, while for the parallel configuration, let us focus only on the $\mathcal{C} = 2$ and $\mathcal{C} = 3$ phases. Then, within those regions, we define a dimensionless figure of merit as

$$\mathfrak{f.o.m.} = \frac{\min\left(\delta_{\text{above}}, \delta_{\text{below}}\right)}{W\lambda}. \tag{S15}$$

The larger this figure of merit, naively, the likelier the band is at stabilizing a Chern phase in the presence of interactions.

Referring to Table II for guidance, we see the following trends for the antiparallel configuration:

1. The $\mathcal{C} = 2$ phase that exists in the negative Δ region around $\theta \approx 1.2°$ is robust as it can be seen with and without relaxation and with a larger or smaller γ_0 value.

2. Disregarding lattice relaxation increases the bandwidths of the first conduction band, decrease the gap above that band while increases the gap below it. Not including lattice relaxation also decreases the phase space area of this $\mathcal{C} = 2$ phase.

3. The absence of lattice relaxation increases the trace condition violation.

4. Reducing γ_0 decreases δ_{above} and increases δ_{below}. It does not seem to affect the W much. λ is slightly decreased with reduced γ_0.

For the parallel configuration, the $\mathcal{C} = 2$ phase has the following trends, many of which are opposite to the trends of the $\mathcal{C} = 2$ phase in the antiparallel configuration:

1. The $\mathcal{C} = 2$ phase that exists in the negative Δ region around $\theta \approx 1.0°$ is robust as it can be seen with and without relaxation and with a larger or smaller γ_0 value.
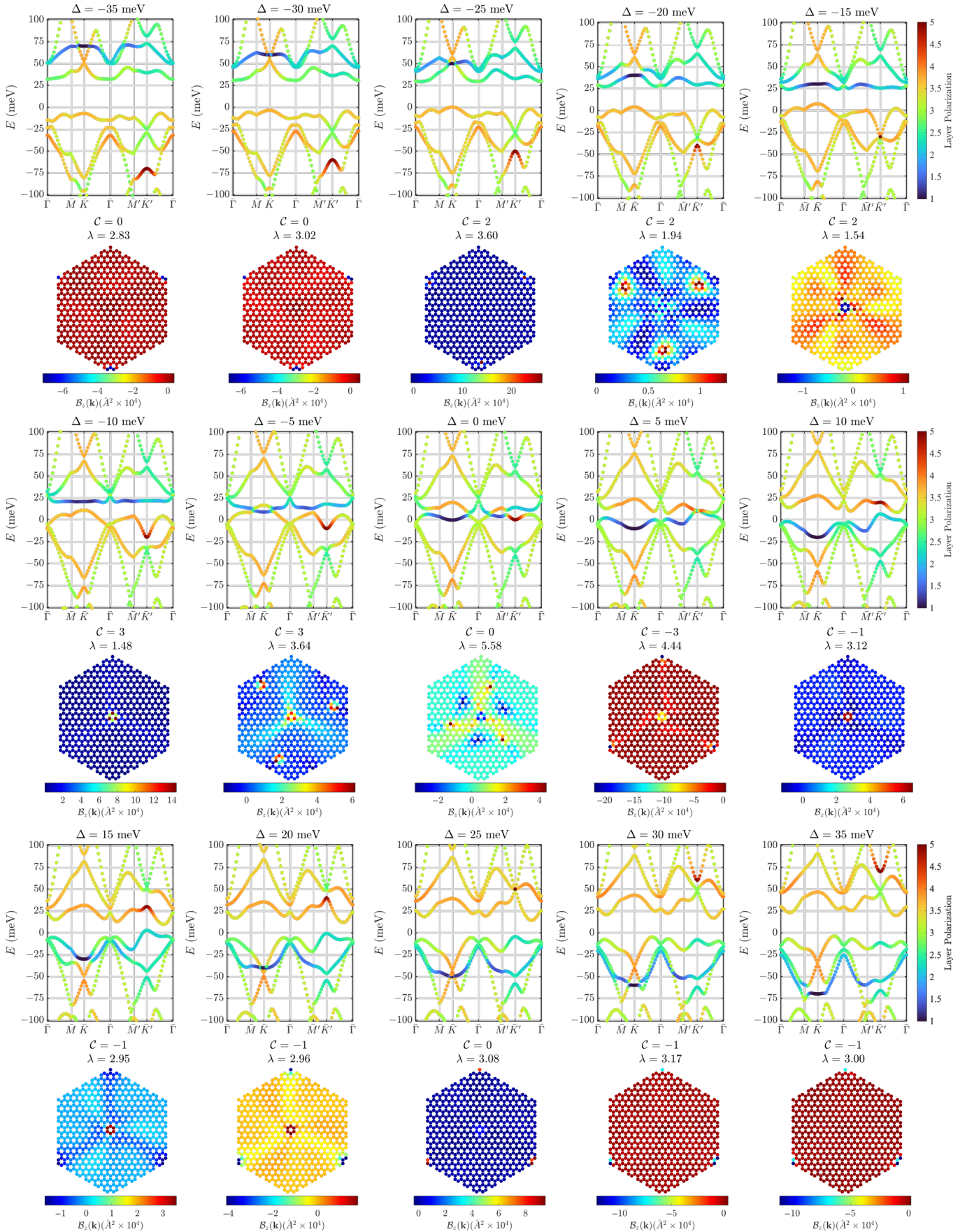
Figure S15. **Evolution of band structures as a function of Δ at $\theta = 1.30°$ for the parallel stacking order.** For each Bloch state, we color code it by its layer polarization. Below each band structure is shown the corresponding Berry curvature distribution of its first conduction band. The Chern number and trace condition violation for that band are also displayed. `Param_two` is used for the results reported here.
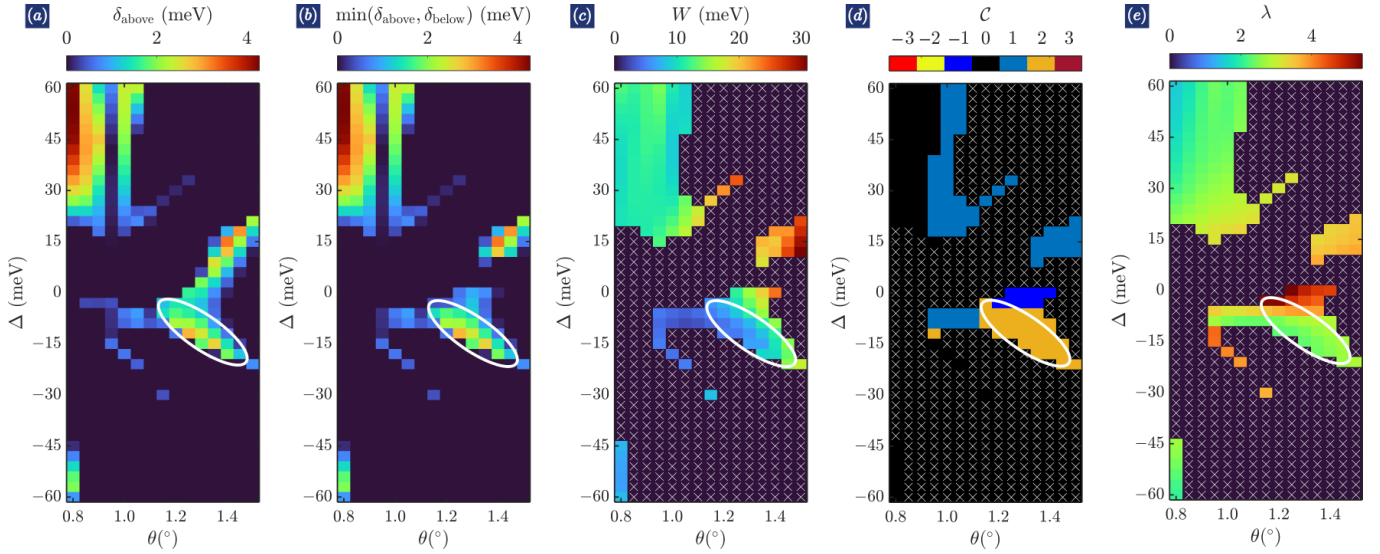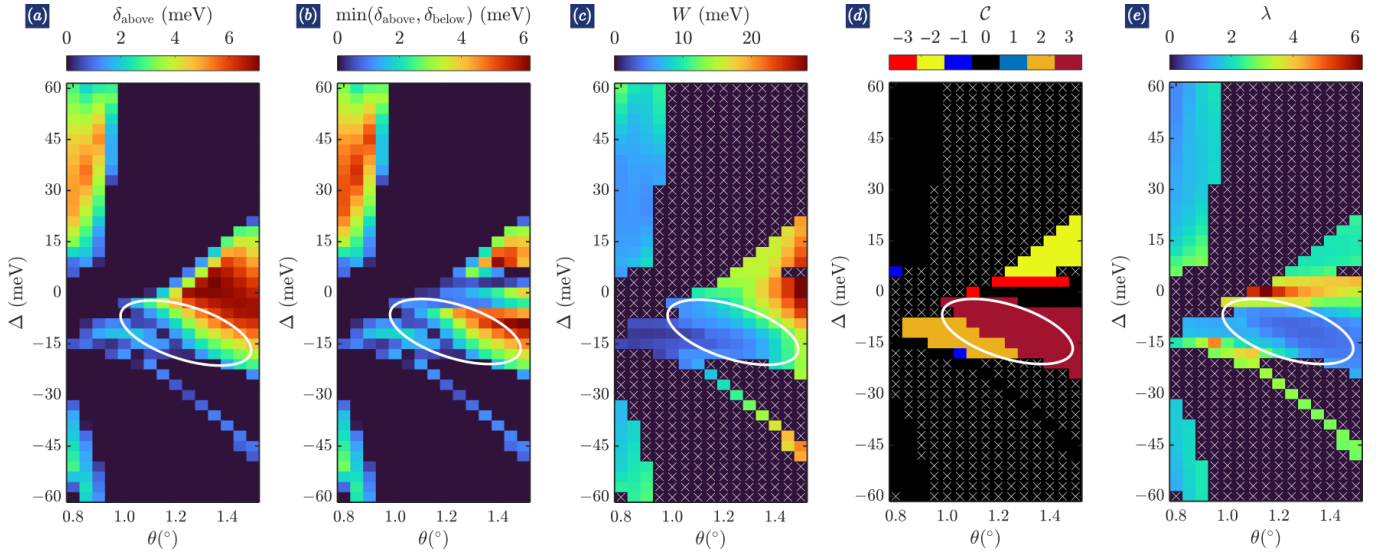
Figure S16. **Non-interacting phase diagram of the parallel stacking order with reduced γ_0 and with lattice relaxation.** (a) Value of the global gap above the first conduction band. (b) Global gap of the first conduction band defined as the minimum of the gap below and above that band. (c) Bandwidth of the first conduction band in cases where it is isolated, *i.e.* $\min(\delta_{\text{above}}, \delta_{\text{below}}) > 0$. (d)-(e) Chern number and trace condition of isolated first conduction bands. There is a large region with $\mathcal{C} = 3$ and relatively small trace condition violation. There is also a region with $\mathcal{C} = 2$ and small trace condition violation. $\times$ denotes points where the gap is zero, for which the task of defining an associated bandwidth, Chern number, or trace condition violation either makes no sense or is useless. `Param_three` is used for the results reported here.
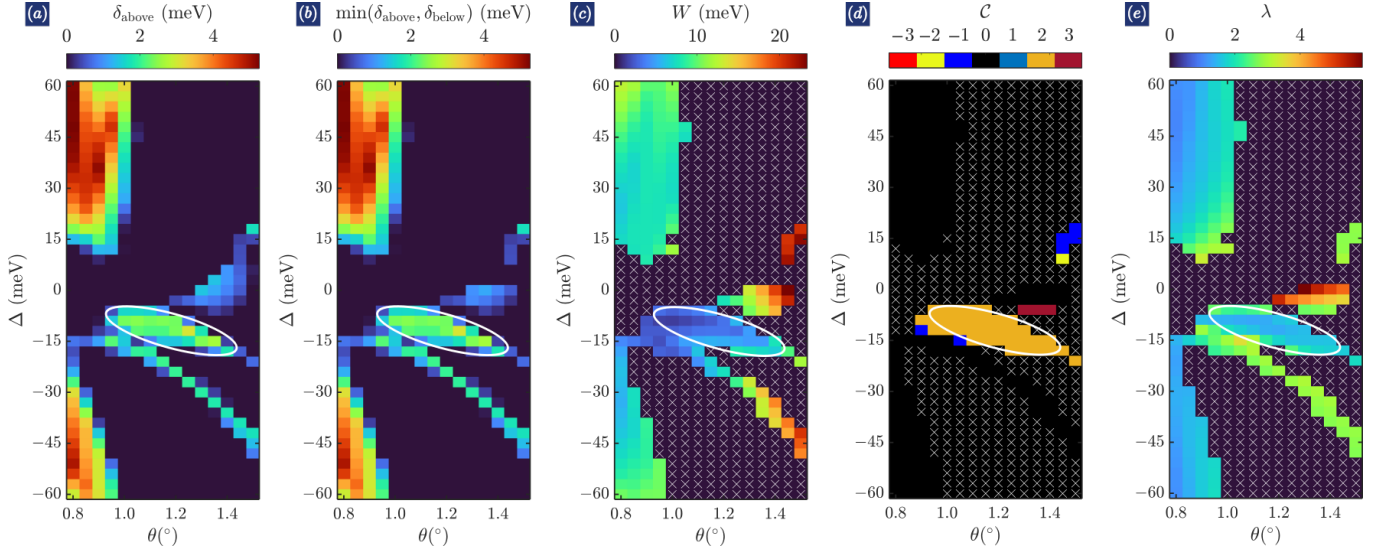
Figure S17. **Non-interacting phase diagram of the parallel stacking order with reduced γ_0 and without lattice relaxation.** (a) Value of the global gap above the first conduction band. (b) Global gap of the first conduction band defined as the minimum of the gap below and above that band. (c) Bandwidth of the first conduction band in cases where it is isolated, *i.e.* $\min(\delta_{\text{above}}, \delta_{\text{below}}) > 0$. (d)-(e) Chern number and trace condition of isolated first conduction bands. The $\mathcal{C} = 3$ is almost completely diminished. However, there is a region with $\mathcal{C} = 2$ and small trace condition violation. $\times$ denotes points where the gap is zero, for which the task of defining an associated bandwidth, Chern number, or trace condition violation either makes no sense or is useless. `Param_four` is used for the results reported here.

2. Removing lattice relaxation slightly increases the phase space area of this phase, increases δ_{above}, decreases δ_{below}, and increases W.

3. The absence of lattice relaxation increases the trace condition violation.

For the parallel configuration, the $\mathcal{C} = 3$ phase has the following trends

1. The $\mathcal{C} = 3$ phase is robust at $\gamma_0 = 3100$ meV with and without lattice relaxation, but is fragile at $\gamma_0 = 2700$ meV since it

Configuration	$\mathcal{C}$	# of points	θ (degrees)	Δ (meV)	δ_{above} (meV)	δ_{below} (meV)	W (meV)	λ
Antiparallel with relaxation	2	58	1.15°	−9	5.63	7.78	2.75	1.71
Antiparallel without relaxation	2	27	1.25°	−12	2.87	11.51	4.68	2.32
Antiparallel with reduced γ_0 and with relaxation	2	56	1.20°	−12	4.21	14.07	2.99	1.50
Antiparallel with reduced γ_0 and without relaxation	2	26	1.30°	−15	2.36	19.16	4.45	2.16
Parallel with relaxation	2	21	0.90°	−9	2.33	8.58	1.47	1.27
Parallel without relaxation	2	26	1.00°	−9	3.05	5.17	2.23	1.50
Parallel with reduced γ_0 and with relaxation	2	21	0.95°	−12	1.81	16.20	1.63	1.52
Parallel with reduced γ_0 and without relaxation	2	34	1.00°	−9	2.36	5.95	1.38	1.39
Parallel with relaxation	3	51	1.25°	−12	4.89	17.24	4.39	0.74
Parallel without relaxation	3	12	1.30°	−9	0.88	7.39	8.45	1.58
Parallel with reduced γ_0 and with relaxation	3	48	1.30°	−12	3.44	18.63	6.00	0.76
Parallel with reduced γ_0 and without relaxation	3	3	1.35°	−6	0.34	2.78	12.90	2.69

Table II. **Summary of observations from non-interacting band structures.** The first column lists the configuration and parameters used. The second column denotes the non-trivial Chern phase of interest. The third column provides the number of data points belonging to such phase and serves as a measure for the area of that phase within the phase diagram. For the data point within that area which minimizes the figure of merit in Eq. S15, we display the angle and displacement field where that occurs in columns four and five, the gaps above and below the first conduction band in columns six and seven, the bandwidth of that band in column eight, and the corresponding trace condition in column nine.
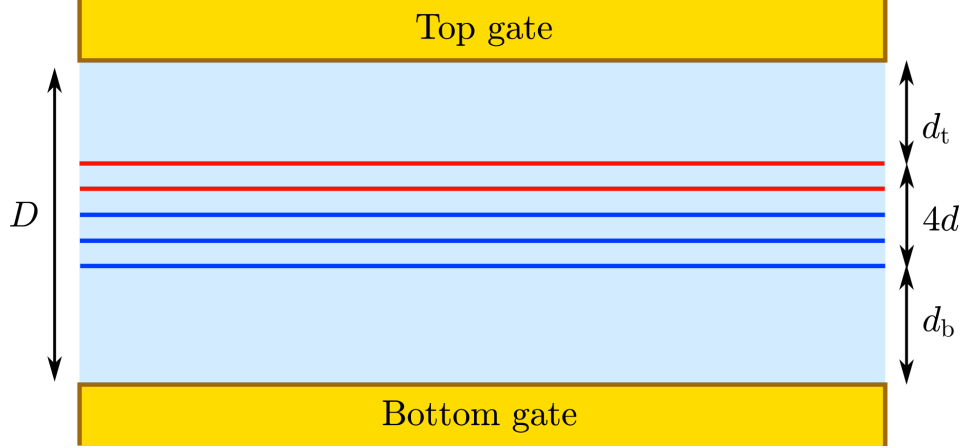
Figure S18. **Schematic representation of a dual-gate experiment.** The blue (red) lines represent layers of the trilayer stack (bilayer stack).

appears prominently with relaxation and essentially disappears without relaxation.

2. Turing off lattice relaxation decreases both δ_{above} and δ_{below} and increases W.

3. The trace condition violation more than doubles when lattice relaxation is removed.

All in all, the $\mathcal{C} = 2$ is quite stable in the non-interacting limit. The $\mathcal{C} = 3$ is more sensitive to changing γ_0 and including or not including lattice relaxation. We end this section with a cautionary note that the preceding analysis using non-interacting band structures only provides some guidance to the likelihood of stabilizing non-trivial Chern phases. However, this analysis is not definitive, and because of that, in the succeeding sections, we carefully examine the effects of Coulomb repulsion on stabilizing non-trivial Chern phases using the self-consistent mean-field.

III. LAYER-DEPENDENT COULOMB POTENTIAL ENERGY

A. Green's Function Derivation

A distinguishing feature of our work is the employment of a layer-dependent electrostatic potential. In this section, we derive that Coulomb potential energy between two charges trapped inside a dual-gate setup. To that end, we solve a related problem of a δ-function charge in between two infinite metallic plates located at the $z = 0$ and $z = D$ planes. The electrostatic potential
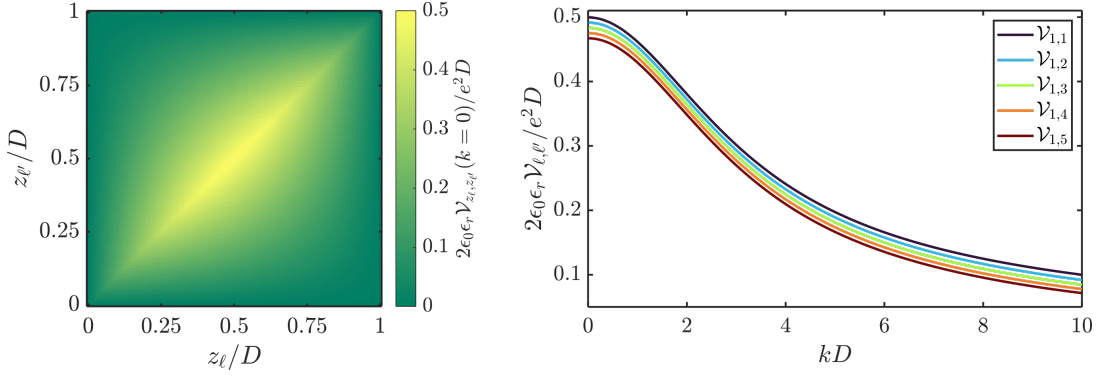
Figure S19. **Layer-dependent Coulomb potential energy.** (Left) Coulomb potential energy at $k = 0$ for different locations of the two layers, where $0 < z_\ell < D$ and $0 < z_{\ell'} < D$. (Right) Momentum dependence of the Coulomb potential energy. Here, we use $d/D = 3.35/400$. When $\ell = \ell'$, the Coulomb potential is essentially $e^2 \tanh(kD/2)/2\epsilon_0\epsilon_r k$ since the gates are much farther away from the active material than the thickness of said material. However, for $\ell \neq \ell'$, the deviation from the layer-independent potential is not negligible. In the present case, that deviation is about 7%. Hence, it is important to consider layer dependence to even capture qualitative physics.

$\varphi(\mathbf{r}, z, z')$ can be integrated from Poisson's equation

$$\nabla^2 \varphi(\mathbf{r}, z, z') = \frac{e}{\epsilon_0 \epsilon_r} \delta^{(2)}(\mathbf{r} - \mathbf{r}')\delta(z - z'), \tag{S16}$$

where $\epsilon_0 \epsilon_r$ is the relative permittivity, $\mathbf{r} = (x, y)$ is the two-dimensional vector, and $(\mathbf{r}', z')$ is the location of the charge $(0 < z' < D)$. For simplicity, let us absorb $e/\epsilon_0\epsilon_r$ into φ for now and resurrect it only when necessary. The metallic boundary conditions are $\varphi(\mathbf{r}, z = 0, z') = \varphi(\mathbf{r}, z = D, z') = 0$. We note that the equation $\nabla^2 \varphi(\mathbf{r}, z, z') = \delta^{(2)}(\mathbf{r} - \mathbf{r}')\delta(z - z')$ is just the familiar Green's function equation for the Laplace operator. We assume that solutions take the form

$$\varphi(\mathbf{r}, z, z') = \sum_{n=1}^{\infty} \varphi_n(\mathbf{r}, z') \sin\left(\frac{\pi n}{D} z\right). \tag{S17}$$

This ansatz clearly satisfies the boundary conditions. It is simply an expansion using a Fourier sine series with coefficients $\varphi_n(\mathbf{r}, z')$ to be determined. Substituting this ansatz into Poisson's equation, we obtain

$$\frac{\partial^2 \varphi_n(\mathbf{r}, z')}{\partial x^2} + \frac{\partial^2 \varphi_n(\mathbf{r}, z')}{\partial y^2} - \frac{\pi^2 n^2}{D^2} \varphi_n(\mathbf{r}, z') = \frac{2}{D} \sin\left(\frac{\pi n}{D} z'\right) \delta^{(2)}(\mathbf{r} - \mathbf{r}'). \tag{S18}$$

We now expand $\varphi_n(\mathbf{r}, z')$ as a Fourier transform: $\varphi_n(\mathbf{r}, z') = \int \frac{d^2\mathbf{k}}{(2\pi)^2} e^{i\mathbf{k}\cdot\mathbf{r}} \varphi_n(\mathbf{k}, z')$ and $\varphi_n(\mathbf{k}, z') = \int d^2\mathbf{r} e^{-i\mathbf{k}\cdot\mathbf{r}} \varphi_n(\mathbf{r}, z')$. Substituting this ansatz into Poisson's equation, we obtain an *algebraic* equation

$$\left[-k^2 - \frac{\pi^2 n^2}{D^2}\right] \varphi_n(\mathbf{k}, z') = \frac{2}{D} \sin\left(\frac{\pi n}{D} z'\right) e^{-i\mathbf{k}\cdot\mathbf{r}'}, \tag{S19}$$

where $k = |\mathbf{k}|$. Thus, the coefficients are

$$\varphi_n(\mathbf{k}, z') = -\left[k^2 + \frac{\pi^2 n^2}{D^2}\right]^{-1} \frac{2}{D} \sin\left(\frac{\pi n}{D} z'\right) e^{-i\mathbf{k}\cdot\mathbf{r}'}. \tag{S20}$$

The full Green's function is therefore

$$\varphi(\mathbf{r}, z, z') = \int \frac{d^2\mathbf{k}}{(2\pi)^2} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \varphi(\mathbf{k}, z, z'),$$

$$\varphi(\mathbf{k}, z, z') = -\frac{1}{D} \sum_{n=-\infty}^{\infty} \left[k^2 + \frac{\pi^2 n^2}{D^2}\right]^{-1} \sin\left(\frac{\pi n}{D} z'\right) \sin\left(\frac{\pi n}{D} z\right). \tag{S21}$$

To do the sum, we can use Poisson's summation formula: $\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n)$, where $\hat{f}(\omega)$ (the hat here denotes Fourier transform and not operator) is the Fourier transform of $f(t)$ [1]. In our case, we have

$$
\begin{aligned}
f(t) &= \left[k^2 + \frac{\pi^2 t^2}{D^2}\right]^{-1} \sin\left(\frac{\pi t}{D}z'\right)\sin\left(\frac{\pi t}{D}z\right) = \frac{1}{2}\left[k^2 + \frac{\pi^2 t^2}{D^2}\right]^{-1}\left[\cos\left(\frac{\pi(z-z')}{D}t\right) - \cos\left(\frac{\pi(z+z')}{D}t\right)\right] \\
&= \frac{1}{4}\left[k^2 + \frac{\pi^2 t^2}{D^2}\right]^{-1}\left[e^{\pi i z_- t/D} + e^{-\pi i z_- t/D} - e^{\pi i z_+ t/D} - e^{-\pi i z_+ t/D}\right],
\end{aligned}
$$

(S22)

where $z_\pm = z \pm z'$. The Fourier transform of this function is

$$
\hat{f}(\omega) = \frac{D}{4k}\left(\exp\left[-2kD\left|\omega - \frac{z_-}{2D}\right|\right] + \exp\left[-2kD\left|\omega + \frac{z_-}{2D}\right|\right] - \exp\left[-2kD\left|\omega - \frac{z_+}{2D}\right|\right] - \exp\left[-2kD\left|\omega + \frac{z_+}{2D}\right|\right]\right).
$$

(S23)

This is calculated as follows. We consider a function of the following form, with various constants to be determined later,

$$
\begin{aligned}
f(t) &= \frac{e^{i\beta t}}{k^2 + \alpha^2 t^2} = b\int_{\mathbb{R}} e^{2\pi i\omega t}e^{-a|\omega-\omega'|}d\omega = b\left[\int_{\omega'}^{\infty} e^{2\pi i\omega t}e^{-a(\omega-\omega')}d\omega + \int_{-\infty}^{\omega'} e^{2\pi i\omega t}e^{+a(\omega-\omega')}d\omega\right] \\
&= be^{2\pi i\omega' t}\left[\int_0^{\infty} e^{2\pi i\omega t}e^{-a\omega}d\omega + \int_{-\infty}^0 e^{2\pi i\omega t}e^{+a\omega}d\omega\right] = be^{2\pi i\omega' t}\left[\frac{2a}{a^2 + (2\pi t)^2}\right].
\end{aligned}
$$

(S24)

From this, we find $\omega' = \beta/2\pi$, $a = 2\pi k/\alpha$, and $b = \pi/\alpha k$. Consequently, the Fourier transform of $f(t)$ is

$$
\hat{f}(\omega) = \frac{\pi}{\alpha k}\exp\left[-\frac{2\pi k}{\alpha}\left|\omega - \frac{\beta}{2\pi}\right|\right].
$$

(S25)

Now, matching α and β with D and $z_\pm$, we obtain the desired result. It is clear that $\hat{f}(\omega)$ is even in both $z_\pm$. Therefore, we can write $z_\pm$ as $|z_\pm|$ (actually, $z_+ > 0$ so the absolute value on z_+ is not necessary). Furthermore, we also note that $|z_\pm|/2D < 1$. This allows us to compute the sums exactly. For example,

$$
\begin{aligned}
\sum_{n=-\infty}^{\infty} \exp\left[-2kD\left|n - \frac{|z_-|}{2D}\right|\right] &= \sum_{n=1}^{\infty}\exp\left[-2kD\left(n - \frac{|z_-|}{2D}\right)\right] + \sum_{n=-\infty}^{0}\exp\left[2kD\left(n - \frac{|z_-|}{2D}\right)\right] \\
&= \operatorname{csch}(kD)\cosh(k(D - |z_-|)).
\end{aligned}
$$

(S26)

Putting the pieces together, we find

$$
\sum_{n=-\infty}^{\infty} \hat{f}(n) = \frac{D\operatorname{csch}(kD)}{2k}\left[\cosh(k\left[D - |z_-|\right]) - \cosh(k\left[D - |z_+|\right])\right].
$$

(S27)

Rewriting the full Green's function in terms of this resummed result, we find a very simple formula, with electric charge e and permittivity $\epsilon_0\epsilon_r$ restored:

$$
\varphi(\mathbf{k}, z, z') = -\frac{e\operatorname{csch}(kD)}{2\epsilon_0\epsilon_r k}\left[\cosh(k\left[D - |z_-|\right]) - \cosh(k\left[D - |z_+|\right])\right].
$$

(S28)

Clearly, $\varphi(\mathbf{k}, z=0, z') = \varphi(\mathbf{k}, z=D, z') = 0$ as desired. The Coulomb potential energy is just $\mathcal{V}(\mathbf{k}, z, z') = -e\varphi(\mathbf{k}, z, z')$ [10, 11]:

$$
\mathcal{V}(\mathbf{k}, z, z') = \frac{e^2\operatorname{csch}(kD)}{2\epsilon_0\epsilon_r k}\left[\cosh(k\left[D - |z - z'|\right]) - \cosh(k\left[D - |z + z'|\right])\right].
$$

(S29)

This is the Coulomb potential energy of two charges, one located at z and one located at z', inside two metallic plates a distance D apart. As a consistency check, the symmetric-gate setup where $z = z' = D/2$ has the potential energy $\mathcal{V}(\mathbf{k}, z = z' = D/2) =$

---

[1] Here is a quick proof of the formula, which can be found in any introductory text on Fourier analysis. Let $f(t)$ be a function whose Fourier transform $\hat{f}(\omega) = \int_{\mathbb{R}} e^{-2\pi i\omega t}f(t)dt$ exists. Define a periodic extension of $f(t)$ with period $1: P(t) = \sum_{n\in\mathbb{Z}} f(t+n)$. Because this function is periodic, it has a Fourier series whose coefficients are given by $a_n = \int_0^1 P(t)e^{-2\pi int}dt = \int_0^1 \sum_{m\in\mathbb{Z}} f(t+m)e^{-2\pi int}dt = \int_0^1 \sum_{m\in\mathbb{Z}} f(t+m)e^{-2\pi in(t+m)}dt = \int_{\mathbb{R}} f(t)e^{-2\pi int}dt = \hat{f}(n)$. Consequently, we have $P(t) = \sum_{n\in\mathbb{Z}} f(t+n) = \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{2\pi int}$. Evaluating this at $t = 0$ gives the desired result.

$\frac{e^2}{2\epsilon_0\epsilon_r k}\tanh(kD/2)$. As a final check, we confirm that the Green's function we obtained actually solves Poisson's equation with a δ-function source: $\nabla^2\varphi(\mathbf{r}-\mathbf{r}',z,z')=0$ if $z\neq z'$ since $\partial_z^2\varphi(\mathbf{k},z,z')=+k^2\varphi(\mathbf{k},z,z')$. Exactly at $z=z'$, we have a jump discontinuity when $\mathbf{r}=\mathbf{r}': \lim_{z\to(z')^+}\partial_z\varphi(\mathbf{r}-\mathbf{r}',z,z')-\lim_{z\to(z')^-}\partial_z\varphi(\mathbf{r}-\mathbf{r}',z,z')=\frac{e}{\epsilon_0\epsilon_r}\delta^{(2)}(\mathbf{r}-\mathbf{r}')$. We end this section by considering the $k\to 0$ limit. We do this two different ways and show that they agree. In the first, we simply take the limit of $\mathcal{V}(\mathbf{k},z,z')$

$$\lim_{k\to 0}\mathcal{V}(\mathbf{k},z,z')=\frac{e^2}{2\epsilon_0\epsilon_r}\frac{2D|z+z'|-2D|z-z'|+|z-z'|^2-|z+z'|^2}{2D}=\frac{e^2}{\epsilon_0\epsilon_r}\left[\min(z,z')-\frac{zz'}{D}\right]. \qquad (S30)$$

Apparently, this must equal to the following for $0<z'<D$ and $0\leq z\leq D$:

$$\frac{2D}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}\sin\left(\frac{\pi n}{D}z'\right)\sin\left(\frac{\pi n}{D}z\right)=\frac{D}{2\pi^2}\left[\mathrm{Li}_2\left(e^{\frac{i\pi}{D}(z-z')}\right)-\mathrm{Li}_2\left(e^{\frac{i\pi}{D}(z+z')}\right)+\mathrm{Li}_2\left(e^{-\frac{i\pi}{D}(z-z')}\right)-\mathrm{Li}_2\left(e^{-\frac{i\pi}{D}(z+z')}\right)\right], \qquad (S31)$$

where $\mathrm{Li}_2(z)$ is the dilogarithm. It is not obvious that Eqs. (S30) and (S31) are equal to each other, but one can use the identity $\mathrm{Li}_2(z)+\mathrm{Li}_2(z^{-1})=-\pi^2/6-[\log(-z)]^2/2$ to show that they do indeed agree [12]. We have also checked numerically that they agree. It is worth pointing out that if the metallic gates are held at non-zero potentials $\varphi_0\neq 0$, as long as both gates have the same potential, then all we need to do is add this constant to the electrostatic potential $\varphi(\mathbf{r},z,z')\mapsto\varphi(\mathbf{r},z,z')+\varphi_0$. This has no effect on the Coulomb potential energy since such an overall shift in the electrostatic potential is accounted for in the electron occupation. For metallic gates held at unequal potentials, without loss of generality, let us assume $\varphi(\mathbf{r},z=0,z')=0$ and $\varphi(\mathbf{r},z=D,z')=\varphi_0$. This changes the potential to $\varphi(\mathbf{r}-\mathbf{r}',z,z')\mapsto\varphi(\mathbf{r}-\mathbf{r}',z,z')+\varphi_0 z/D$. The term $\varphi_0 z/D$ is independent of $\mathbf{r}$ and $\mathbf{r}'$. It is most conveniently represented in the Hamiltonian as a layer-dependent electrostatic potential energy. Therefore, the Coulomb potential energy in Eq. (S29) is the most general form for our purpose. For our system of interest, we use the setup shown in Fig. S18. Accordingly, Eq. (S29) should be modified such that

$$\begin{aligned} D &= d_{\mathrm{b}}+d_{\mathrm{t}}+4d, \\ z_\ell &= d_{\mathrm{b}}+(\ell-1)d, \end{aligned} \qquad (S32)$$

where d_{b} (d_{t}) is the distance from the bottom (top) to the nearest graphene layer and $d=3.35$ Å is the separation distance between adjacent graphene layers. Using this notation, we write $\mathcal{V}_{\ell,\ell'}(\mathbf{k})=\mathcal{V}(\mathbf{k},z_\ell,z_{\ell'})$. To simplify, we assume that $d_{\mathrm{t}}=d_{\mathrm{b}}$ so that we only need to specify the total gate-to-gate distance D. In this work, we primarily use $D=400$ Å (the distance from the middle of the five-layer stack to one of the gates is 200 Å). For these values, the Coulomb potential energy is shown in Fig. S19. We see that when $\ell\neq\ell'$, the Coulomb potential is slightly suppressed (about 7% at maximum) for all momenta. When $\ell=\ell'$, we find that the Coulomb potential is essentially equal to the simple $(e^2/2\epsilon_0\epsilon_r)\tanh(kD/2)/k$ potential.

B. Screening in the Classical Limit

In this section, we consider the $\mathbf{k}=\mathbf{0}$ component of the layer-dependent Coulomb potential, which corresponds to uniform charge densities. We show that the layer-dependent Coulomb potential agrees exactly with classical considerations, boosting confidence in the correctness of our formulas. For this purpose, we consider a parallel-plate capacitor model with two gates and two layers of the active material (the generalization to many layers is already included in the formulas of the previous section and is unnecessary for the classical consideration here). We consider the geometry shown in Fig. S20, which is the same as the geometry in Fig. S18 in the limit with only two active layers. The charge densities on the bottom and top gates are σ_{b} and σ_{t} and on the first and second layers are σ_1 and σ_2. The associated displacement fields in the regions below and above the active material are

$$\mathcal{D}_{\mathrm{b}}=+\frac{\sigma_{\mathrm{b}}}{\epsilon_0}, \qquad (S33a)$$

$$\mathcal{D}_{\mathrm{t}}=-\frac{\sigma_{\mathrm{t}}}{\epsilon_0}. \qquad (S33b)$$

The field in between the two layers of the active material can be found using Gauss's law:

$$\tilde{\mathcal{D}}-\mathcal{D}_{\mathrm{b}}=\frac{\sigma_1}{\epsilon_0}, \qquad (S34a)$$

$$\mathcal{D}_{\mathrm{t}}-\tilde{\mathcal{D}}=\frac{\sigma_2}{\epsilon_0}. \qquad (S34b)$$
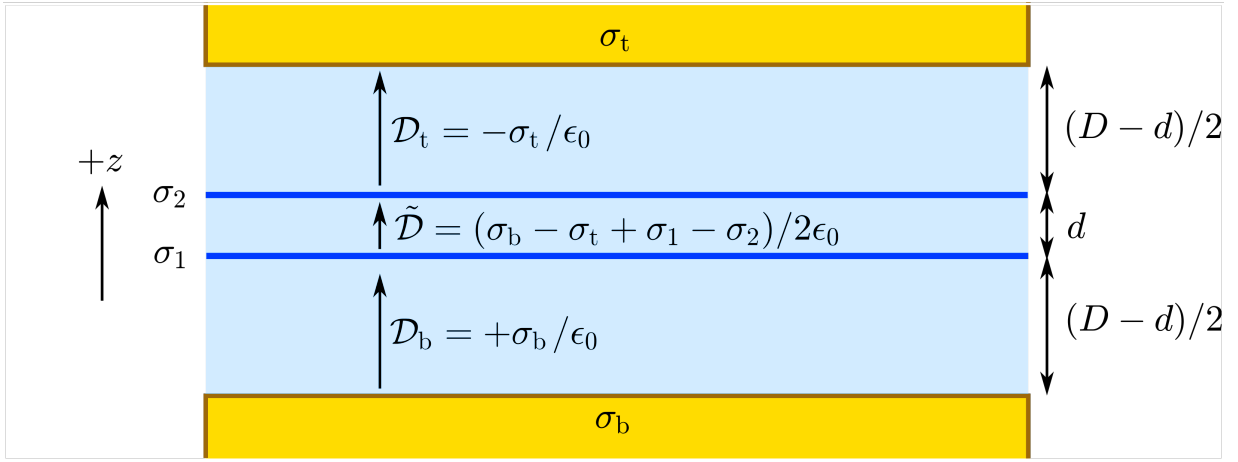
Figure S20. **Parallel-plate capacitor model.** Two active layers are encapsulated between a top and bottom metallic gate. The charge densities σ and associated displacement fields $\mathcal{D}$ are indicated.

Solving for $\tilde{\mathcal{D}}$, we find

$$\tilde{\mathcal{D}} = \frac{\mathcal{D}_b + \mathcal{D}_t}{2} + \frac{\sigma_1 - \sigma_2}{2\epsilon_0} = \mathcal{D}_{ave} + \frac{\sigma_1 - \sigma_2}{2\epsilon_0}, \tag{S35}$$

where $\mathcal{D}_{ave} = (\mathcal{D}_b + \mathcal{D}_t)/2 = (\sigma_b - \sigma_t)/2\epsilon_0$ is the average displacement field often quoted in experimental works as *the* displacement field experimentally controlled by varying gate biases. Using classical electrostatics, we find that the areal energy density is

$$\mathcal{U} = \frac{\epsilon_0}{2\epsilon_r} \int |\mathcal{D}(z)|^2 dz = \frac{1}{2\epsilon_0\epsilon_r} \frac{D-d}{2} \left(\sigma_b^2 + \sigma_t^2\right) + \frac{d}{2\epsilon_0\epsilon_r} \frac{(\sigma_b - \sigma_t + \sigma_1 - \sigma_2)^2}{4}. \tag{S36}$$

It is convenient to group terms based on the total density $\sigma = \sigma_t + \sigma_b = -\sigma_1 - \sigma_2$ and the density fluctuations $\sigma_t - \sigma_b = \delta\sigma_{gate}$ and $\sigma_2 - \sigma_1 = \delta\sigma_{sample}$. We have enforced charge neutrality in the system. We can then rewrite the potential energy density as

$$\mathcal{U} = \frac{1}{8\epsilon_0\epsilon_r} (D-d) \left(\sigma^2 + \delta\sigma_{gate}^2\right) + \frac{d}{8\epsilon_0\epsilon_r} (\delta\sigma_{gate} + \delta\sigma_{sample})^2. \tag{S37}$$

We emphasize that this is *not* just the energy density of interaction between the charges on the active material but is the *total* energy density of putting together this charge configuration, including the energy it costs to deposit charges on the gates and their interaction with the charges on the active material. For now, we assume that the gates are held at the same potential. We show that the screened Coulomb potential allows us to calculate this total energy density in the Hartree approximation. To do that, we need to determine the potential difference between the top and bottom gates in the presence of charges as indicated and then set that potential difference equal to zero to determine the relationship between the charges. We find difference in the electrostatic potential between the top and bottom gates to be

$$\Delta\varphi = \varphi_t - \varphi_b = -\frac{\tilde{\mathcal{D}}}{\epsilon_r}d - \frac{\mathcal{D}_t + \mathcal{D}_b}{\epsilon_r}\frac{D-d}{2} = \frac{\delta\sigma_{sample}}{2\epsilon_0\epsilon_r}d + \frac{\delta\sigma_{gate}}{2\epsilon_0\epsilon_r}D = 0 \rightarrow d\delta\sigma_{sample} = -D\delta\sigma_{gate}. \tag{S38}$$

Using this restriction, Eq. (S36) can be rewritten as

$$\mathcal{U} = \frac{1}{8\epsilon_0\epsilon_r} (D-d)\sigma^2 + \frac{1}{8\epsilon_0\epsilon_r} \frac{d(D-d)}{D} \delta\sigma_{sample}^2. \tag{S39}$$

In terms of the screened Coulomb potential, the energy density of system is given by

$$\mathcal{U} = \frac{1}{2e^2} \sum_{\ell,\ell'} \mathcal{V}_{\ell,\ell'}(\mathbf{k}=\mathbf{0})\sigma_\ell\sigma_{\ell'}, \tag{S40}$$

where we have chosen zero potential at the middle of the stack. Recalling that

$$\mathcal{V}_{1,1}(\mathbf{0}) = \mathcal{V}_{2,2}(\mathbf{0}) = \frac{e^2}{4\epsilon_0\epsilon_r D} \left[D^2 - d^2\right], \quad \mathcal{V}_{1,2}(\mathbf{0}) = \mathcal{V}_{2,1}(\mathbf{0}) = \frac{e^2}{4\epsilon_0\epsilon_r D} \left[D-d\right]^2, \tag{S41}$$

we find

$$\mathcal{U} = \frac{1}{8\epsilon_0\epsilon_r D}\left[D^2 - d^2\right]\left(\sigma_1^2 + \sigma_2^2\right) + \frac{1}{4\epsilon_0\epsilon_r D}\left[D - d\right]^2\sigma_1\sigma_2,$$
$$= \frac{1}{8\epsilon_0\epsilon_r}(D - d)\sigma^2 + \frac{1}{8\epsilon_0\epsilon_r}\frac{d(D - d)}{D}\delta\sigma_{\text{sample}}^2, \tag{S42}$$

which agrees exactly with the other method, showing consistency between the two.

Next, we consider the situation of unequal potentials on the metallic gates. This case is slightly more complicated due to inhomogeneous boundary conditions. The electrostatic potential energy in Eq. (S36) is correct irrespective of whether the two metallic gates have the same potential. This is because that formula calculates the energy directly from the energy density of electric fields. Consequently, it does not suffer from gauge ambiguity hidden in formulas involving potentials that require fixing some reference point. For us, potentials are measured relative to ground. To check consistency with the potential approach, we calculate the energy density using the following expression

$$\mathcal{U} = \frac{1}{2V_{\text{sys}}}\int\rho(\mathbf{r}, z)\varphi(\mathbf{r}, z)d^2\mathbf{r}dz + \frac{1}{2}\varphi_t\sigma_t + \frac{1}{2}\varphi_b\sigma_b, \tag{S43}$$

where V_{sys} is the area of the system, the charge density distribution inside the gates is given by

$$\rho(\mathbf{r}, z) = \rho(z) = \sigma_1\delta(z - [D - d]/2) + \sigma_2\delta(z - [D + d]/2) \tag{S44}$$

and the electrostatic potential can be found by convolving the Green's function in Sec. III A with the charge density in Eq. (S44)

$$\varphi(\mathbf{r}, z) = -\frac{1}{e}\int\rho(z')\varphi(\mathbf{r} - \mathbf{r}', z, z')d^2\mathbf{r}'dz' + \frac{\epsilon_0\epsilon_r}{e}\int\left[\varphi_t\frac{\partial\varphi(\mathbf{r} - \mathbf{r}', z, z')}{\partial z'}\bigg|_{z'=D} - \varphi_b\frac{\partial\varphi(\mathbf{r} - \mathbf{r}', z, z')}{\partial z'}\bigg|_{z'=0}\right]d^2\mathbf{r}' \quad \text{(S45)}$$

subject to Dirichlet boundary conditions $\varphi(\mathbf{r}, z = D) = \varphi_t$ and $\varphi(\mathbf{r}, z = 0) = \varphi_b$. In Eq. (S36), we have separated two contributions to the energy: (1) the integral contribution contains effects coming from the charges in between the gates, and (2) the latter two contributions contain effects coming from the charges at the gates. From this division, it is simple to observe that even if there were absolutely no charges in the gates, i.e. $\rho(\mathbf{r}, z) = 0$, the electrostatic potential energy would still be non-zero due purely to charges on the gates setting up an electric field. It is worth emphasizing that the electrostatic potential $\varphi(\mathbf{r}, z)$ contains information about the entire setup, i.e. all charges everywhere. So the distinction drawn here is only a rough division for bookkeeping purposes. It is absolutely *not* correct to say that the integral term in Eq. (S36) contains only contributions coming from interactions among charges on the active material.

Because the charge density is uniform laterally, translational invariance dictates that the potential does not depend on the in-plane coordinates. Therefore, integration over $\mathbf{r}$ and $\mathbf{r}'$ allows us to replace the real-space potential with the $\mathbf{k} = \mathbf{0}$ component of the Fourier-transformed potential. Henceforth, all potentials are in reciprocal space at $\mathbf{k} = 0$. Performing this integration, Eq. (S43) can be simplified to

$$\mathcal{U} = \frac{1}{2}\sigma_1\varphi(z = D/2 - d/2) + \frac{1}{2}\sigma_2\varphi(z = D/2 + d/2) + \frac{1}{2}\varphi_t\sigma_t + \frac{1}{2}\varphi_b\sigma_b, \tag{S46}$$

where the potential at layer 1 is

$$\varphi(z = D/2 - d/2) = \frac{1}{\epsilon_0\epsilon_r}\left[\frac{\sigma_1}{D}\left(\frac{D^2}{4} - \frac{d^2}{4}\right) + \frac{\sigma_2}{D}\left(\frac{D}{2} - \frac{d}{2}\right)^2\right] + \frac{\varphi_b + \varphi_t}{2} - \frac{d}{D}\frac{\varphi_t - \varphi_b}{2}, \tag{S47}$$

and potential at layer 2 is

$$\varphi(z = D/2 + d/2) = \frac{1}{\epsilon_0\epsilon_r}\left[\frac{\sigma_1}{D}\left(\frac{D}{2} - \frac{d}{2}\right)^2 + \frac{\sigma_2}{D}\left(\frac{D^2}{4} - \frac{d^2}{4}\right)\right] + \frac{\varphi_b + \varphi_t}{2} + \frac{d}{D}\frac{\varphi_t - \varphi_b}{2}. \tag{S48}$$

We note that these functions have the correct limits as $d \to D$. The terms in square brackets in Eqs. (S47) and (S48) are the same terms originating from the homogeneous Green's function as they are independent of the gate potentials; the terms involving $\varphi_t + \varphi_b$ are constant shifts in potential needed to fix the boundary conditions; the terms involving $\varphi_t - \varphi_b$ are of the form $\Delta\varphi z/D$, linear functions of position. Indeed, Eqs. (S47) and (S48) are quite simple that one could have easily guessed them without needing to compute the Green's function at all, but the purpose of the exercise here is to check that the Green's function

produces the correct result. Using the fact the that $\varphi_b + \varphi_t = \sigma(D - d)/2\epsilon_0\epsilon_r$, we find

$$
\begin{aligned}
\mathcal{U} &= \frac{1}{8\epsilon_0\epsilon_r D} \left[D^2 - d^2\right] \left(\sigma_1^2 + \sigma_2^2\right) + \frac{1}{4\epsilon_0\epsilon_r D} \left[D - d\right]^2 \sigma_1\sigma_2 - \frac{1}{8\epsilon_0\epsilon_r}\sigma^2(D - d) + \frac{d}{4D}\delta\sigma_{\text{sample}}\Delta\varphi + \frac{1}{2}\varphi_t\sigma_t + \frac{1}{2}\varphi_b\sigma_b \\
&= \frac{1}{8\epsilon_0\epsilon_r D} \left[D^2 - d^2\right] \left(\sigma_1^2 + \sigma_2^2\right) + \frac{1}{4\epsilon_0\epsilon_r D} \left[D - d\right]^2 \sigma_1\sigma_2 + \frac{1}{4D}\left[d\delta\sigma_{\text{sample}} + D\delta\sigma_{\text{gate}}\right]\Delta\varphi \\
&= \frac{1}{8\epsilon_0\epsilon_r D} \left[D^2 - d^2\right] \left(\sigma_1^2 + \sigma_2^2\right) + \frac{1}{4\epsilon_0\epsilon_r D} \left[D - d\right]^2 \sigma_1\sigma_2 + \frac{\epsilon_0\epsilon_r}{2D}(\Delta\varphi)^2 \\
&= \frac{1}{8\epsilon_0\epsilon_r} \left(D - d\right)\sigma^2 + \frac{1}{8\epsilon_0\epsilon_r D}\left[dD - d^2\right]\delta\sigma_{\text{sample}}^2 + \frac{\epsilon_0\epsilon_r}{2D}(\Delta\varphi)^2.
\end{aligned}
$$

(S49)

Now, using Eq. (S38), we can eliminate $\Delta\varphi$ to find

$$
\begin{aligned}
\mathcal{U} &= \frac{1}{8\epsilon_0\epsilon_r} \left(D - d\right)\sigma^2 + \frac{1}{8\epsilon_0\epsilon_r}\left[2d\sigma_{\text{sample}}\sigma_{\text{gate}} + D\sigma_{\text{gate}}^2 + d\sigma_{\text{sample}}^2\right], \\
&= \frac{1}{8\epsilon_0\epsilon_r} \left(D - d\right)\left(\sigma^2 + \delta\sigma_{\text{gate}}^2\right) + \frac{d}{8\epsilon_0\epsilon_r}(\delta\sigma_{\text{gate}} + \delta\sigma_{\text{sample}})^2,
\end{aligned}
$$

(S50)

which agrees exactly with Eq. (S37), again showing that our approach is correct even at unequal potential values at the gates.

One might worry that in order to minimize the total electrostatic energy that one needs to not only rearrange the charges on the active material but also redistribute the charges at the gates as well. We end this section by remarking that this is not necessary as long as the gate potentials are fixed. This can be seen by noting that the potential energy at fixed potential difference between the gates can be written as

$$
\mathcal{U} = \frac{1}{2e^2} \sum_{\ell,\ell'} \mathcal{V}_{\ell,\ell'}(\mathbf{k} = \mathbf{0})\sigma_\ell\sigma_{\ell'} + \frac{\epsilon_0\epsilon_r}{2D}(\Delta\varphi)^2,
$$

(S51)

where the last term is a fixed constant once the gate potentials are specified. This difference in the gate potentials is related to the Δ parameter in the single-particle Hamiltonian. In Eq. (S51), importantly, there is no information about the gate charge densities. Therefore, minimization of the Coulomb potential energy can be done by varying only the electron densities in the active material. This justifies the implementation of the gate-screened Coulomb potential in the HF algorithm, where no information about the gate charge densities appears.

IV. SELF-CONSISTENT MEAN-FIELD THEORY

To study electron-electron interactions, it is most convenient to adopt second-quantized notation. First, we rewrite the kinetic energy in second-quantized notation

$$
\hat{\mathcal{K}} = \sum_{\mathbf{k}\in\text{mBZ},\mathbf{G},\mathbf{G}'} \sum_{\mu,\mu'} \hat{c}_{\mu,\mathbf{k}+\mathbf{G}}^\dagger \mathbb{K}_{\mathbf{G},\mathbf{G}'}^{\mu,\mu'}(\mathbf{k})\hat{c}_{\mu',\mathbf{k}+\mathbf{G}'},
$$

(S52)

where $\mathbb{K}_{\mathbf{G},\mathbf{G}'}^{\mu,\mu'}(\mathbf{k})$ is same matrix defined in Sec. I with a slight change in notation to emphasize the dependence on reciprocal lattice vectors, $\hat{c}_{\mu,\mathbf{k}+\mathbf{G}}^\dagger$ creates a plane-wave state $|\psi_{\mu,\mathbf{k}+\mathbf{G}}\rangle$ with μ denoting all the internal indices, which include $\{s(\text{spin}), \nu(\text{valley}), \sigma(\text{sublattice}), \ell(\text{layer})\}$. Throughout, we use lower-case momenta to denote momenta confined within a moiré-Brillouin zone (mBZ) and capitalized momenta to denote moiré reciprocal lattice vectors. We can diagonalize the kinetic energy by finding eigenstates $|\Psi_{n,\xi,\mathbf{k}}\rangle = e^{i\mathbf{k}\cdot\hat{\mathbf{r}}} |u_{n,\xi,\mathbf{k}}\rangle = \sum_{\mathbf{G},\sigma,\ell} \phi_{n,\xi,\sigma,\ell,\mathbf{k}+\mathbf{G}} |\psi_{\xi,\sigma,\ell,\mathbf{k}+\mathbf{G}}\rangle$ with energy $E_{n,\xi}(\mathbf{k})$, where n is band index, $\xi \in \{s, \nu\}$ denotes only the non-spatial indices, and $\phi_{n,\xi,\sigma,\ell,\mathbf{k}+\mathbf{G}}$ are numerical coefficients obtained from diagonalizing $\mathbb{K}$. The operators that create $|\Psi_{n,\xi,\mathbf{k}}\rangle$ are denoted $\hat{\Psi}_{n,\xi,\mathbf{k}}^\dagger$. Normalization is enforced by $\sum_{\mathbf{G},\sigma,\ell} |\phi_{n,\xi,\sigma,\ell,\mathbf{k}+\mathbf{G}}|^2 = 1$. In this band basis, the kinetic energy operator is diagonal

$$
\hat{\mathcal{K}} = \sum_{\mathbf{k}\in\text{mBZ}} \sum_{n,\xi} E_{n,\xi}(\mathbf{k})\hat{\Psi}_{n,\xi,\mathbf{k}}^\dagger \hat{\Psi}_{n,\xi,\mathbf{k}}.
$$

(S53)

Moving onto the four-fermion Coulomb interaction, we can also write it in the band basis as

$$\frac{1}{2}\sum_{\ell,\ell',\alpha,\alpha'}\int d^2\mathbf{r}d^2\mathbf{r}'\mathcal{V}_{\ell,\ell'}(\mathbf{r}-\mathbf{r}') : \hat{\rho}_{\alpha,\ell}(\mathbf{r})\hat{\rho}_{\alpha',\ell'}(\mathbf{r}') :$$

$$= \frac{1}{2}\sum_{\substack{\mathbf{k},\mathbf{p},\mathbf{q}\in\mathrm{mBZ}\\n_1,n_2,n_3,n_4\\\xi_1,\xi_2}}\mathbb{V}_{\xi_1,\xi_2}^{n_1,n_2,n_3,n_4}(\mathbf{k},\mathbf{p},\mathbf{q})\hat{\Psi}_{n_1,\xi_1,\mathbf{k}+\mathbf{q}}^{\dagger}\hat{\Psi}_{n_2,\xi_2,\mathbf{p}-\mathbf{q}}^{\dagger}\hat{\Psi}_{n_4,\xi_2,\mathbf{p}}\hat{\Psi}_{n_3,\xi_1,\mathbf{k}}, \tag{S54}$$

where $\alpha = \{s,\nu,\sigma\}$ to single out the layer index ℓ and the matrix elements are computed using the one-particle wave functions:

$$\mathbb{V}_{\xi_1,\xi_2}^{n_1,n_2,n_3,n_4}(\mathbf{k},\mathbf{p},\mathbf{q}) = \frac{1}{V_{\mathrm{sys}}}\sum_{\mathbf{Q},\ell_1,\ell_2}\left[\sum_{\mathbf{G},\sigma_1}\phi_{n_1,\xi_1,\sigma_1,\ell_1,\mathbf{k}+\mathbf{G}+\mathbf{q}+\mathbf{Q}}^{\dagger}\phi_{n_3,\xi_1,\sigma_1,\ell_1,\mathbf{k}+\mathbf{G}}\right]\mathcal{V}_{\ell_1,\ell_2}(\mathbf{q}+\mathbf{Q})\times$$

$$\times\left[\sum_{\mathbf{P},\sigma_2}\phi_{n_2,\xi_2,\sigma_2,\ell_2,\mathbf{p}+\mathbf{P}-\mathbf{q}-\mathbf{Q}}^{\dagger}\phi_{n_4,\xi_2,\sigma_2,\ell_2,\mathbf{p}+\mathbf{P}}\right]. \tag{S55}$$

The layer-dependent screened Coulomb potential energy $\mathcal{V}_{\ell_1,\ell_2}(\mathbf{k})$ is given in Eq. (S29) of Sec. III. We can simplify by defining layer-resolved form factors

$$\mathbb{A}_{\xi,\ell}^{n,n'}(\mathbf{k},\mathbf{q}+\mathbf{Q}) = \sum_{\mathbf{G},\sigma}\phi_{n,\xi,\sigma,\ell,\mathbf{k}+\mathbf{G}+\mathbf{q}+\mathbf{Q}}^{\dagger}\phi_{n',\xi,\sigma,\ell,\mathbf{k}+\mathbf{G}} = \langle u_{n,\xi,\mathbf{k}+\mathbf{q}}|\, e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}}\mathbb{P}_{\ell}\,|u_{n',\xi,\mathbf{k}}\rangle, \tag{S56}$$

where $\mathbb{P}_{\ell}$ is the projector onto the ℓ layer. These form factors obey $\left[\mathbb{A}_{\xi,\ell}^{n,n'}(\mathbf{k},\mathbf{q}+\mathbf{Q})\right]^{\dagger} = \mathbb{A}_{\xi,\ell}^{n',n}(\mathbf{k}+\mathbf{q},-\mathbf{q}-\mathbf{Q})$. With these form factors, we can write the matrix elements as

$$\mathbb{V}_{\xi_1,\xi_2}^{n_1,n_2,n_3,n_4}(\mathbf{k},\mathbf{p},\mathbf{q}) = \frac{1}{V_{\mathrm{sys}}}\sum_{\mathbf{Q},\ell_1,\ell_2}\mathbb{A}_{\xi_1,\ell_1}^{n_1,n_3}(\mathbf{k},\mathbf{q}+\mathbf{Q})\mathcal{V}_{\ell_1,\ell_2}(\mathbf{q}+\mathbf{Q})\mathbb{A}_{\xi_2,\ell_2}^{n_2,n_4}(\mathbf{p},-\mathbf{q}-\mathbf{Q})$$

$$= \frac{1}{V_{\mathrm{sys}}}\sum_{\mathbf{Q},\ell_1,\ell_2}\mathbb{A}_{\xi_1,\ell_1}^{n_1,n_3}(\mathbf{k},\mathbf{q}+\mathbf{Q})\mathcal{V}_{\ell_1,\ell_2}(\mathbf{q}+\mathbf{Q})\left[\mathbb{A}_{\xi_2,\ell_2}^{n_4,n_2}(\mathbf{p}-\mathbf{q},\mathbf{q}+\mathbf{Q})\right]^{\dagger}. \tag{S57}$$

We note that the form factors $\mathbb{A}_{\xi,\ell}^{n,n'}(\mathbf{k},\mathbf{q})$ are periodic in the first momentum $\mathbf{k}$ but *not* in the second momentum $\mathbf{q}$ assuming that the energy eigenfunctions are calculated with enough $\mathbf{G}$ vectors to ensure convergence; meanwhile, the band-projected Coulomb matrix elements $\mathbb{V}_{\xi_1,\xi_2}^{n_1,n_2,n_3,n_4}(\mathbf{k},\mathbf{p},\mathbf{q})$ are periodic in all three momenta, again assuming that enough $\mathbf{Q}$ vectors are included in the sum. Let us now implement the HF decoupling assuming that moiré translational symmetry is *not* broken:

$$\frac{1}{2}\sum_{\substack{\mathbf{k},\mathbf{p},\mathbf{q}\in\mathrm{mBZ}\\n_1,n_2,n_3,n_4\\\xi_1,\xi_2}}\mathbb{V}_{\xi_1,\xi_2}^{n_1,n_2,n_3,n_4}(\mathbf{k},\mathbf{p},\mathbf{q})\hat{\Psi}_{n_1,\xi_1,\mathbf{k}+\mathbf{q}}^{\dagger}\hat{\Psi}_{n_2,\xi_2,\mathbf{p}-\mathbf{q}}^{\dagger}\hat{\Psi}_{n_4,\xi_2,\mathbf{p}}\hat{\Psi}_{n_3,\xi_1,\mathbf{k}}$$

$$\approx \sum_{\substack{\mathbf{k},\mathbf{p}\in\mathrm{mBZ}\\n_1,n_2,n_3,n_4\\\xi_1,\xi_2}}\mathbb{V}_{\xi_1,\xi_2}^{n_1,n_2,n_3,n_4}(\mathbf{k},\mathbf{p},0)\langle\hat{\Psi}_{n_1,\xi_1,\mathbf{k}}^{\dagger}\hat{\Psi}_{n_3,\xi_1,\mathbf{k}}\rangle\hat{\Psi}_{n_2,\xi_2,\mathbf{p}}^{\dagger}\hat{\Psi}_{n_4,\xi_2,\mathbf{p}}$$

$$- \sum_{\substack{\mathbf{k},\mathbf{p}\in\mathrm{mBZ}\\n_1,n_2,n_3,n_4\\\xi_1,\xi_2}}\mathbb{V}_{\xi_1,\xi_2}^{n_1,n_2,n_3,n_4}(\mathbf{k},\mathbf{p},\mathbf{p}-\mathbf{k})\langle\hat{\Psi}_{n_1,\xi_1,\mathbf{p}}^{\dagger}\hat{\Psi}_{n_4,\xi_2,\mathbf{p}}\rangle\hat{\Psi}_{n_2,\xi_2,\mathbf{k}}^{\dagger}\hat{\Psi}_{n_3,\xi_1,\mathbf{k}} + \dots \tag{S58}$$

Defining the density matrix $\mathbb{D}_{\xi,\xi'}^{n,n'}(\mathbf{p}) = \langle\hat{\Psi}_{n,\xi,\mathbf{p}}^{\dagger}\hat{\Psi}_{n',\xi',\mathbf{p}}\rangle$, we can now write the full Coulomb potential energy in the HF approximation as

$$\hat{\mathcal{V}} \approx \sum_{\substack{\mathbf{k},\mathbf{p}\in\mathrm{mBZ}\\n_1,n,n_3,n'\\\xi_1,\xi}}\mathbb{V}_{\xi_1,\xi}^{n_1,n,n_3,n'}(\mathbf{p},\mathbf{k},0)\mathbb{D}_{\xi_1,\xi_1}^{n_1,n_3}(\mathbf{p})\hat{\Psi}_{n,\xi,\mathbf{k}}^{\dagger}\hat{\Psi}_{n',\xi,\mathbf{k}} - \sum_{\substack{\mathbf{k},\mathbf{p}\in\mathrm{mBZ}\\n_1,n,n',n_4\\\xi',\xi}}\mathbb{V}_{\xi',\xi}^{n_1,n,n',n_4}(\mathbf{k},\mathbf{p},\mathbf{p}-\mathbf{k})\mathbb{D}_{\xi',\xi}^{n_1,n_4}(\mathbf{p})\hat{\Psi}_{n,\xi,\mathbf{k}}^{\dagger}\hat{\Psi}_{n',\xi',\mathbf{k}} + \dots,$$

$$\tag{S59}$$

where the ellipsis stands for (ignored) constants that can only shift the overall spectrum. There is much redundancy in the Coulomb matrix elements. Such redundancy can be exploited to speed up numerical implementation. To see this, we write the matrix elements explicitly:

$$\mathbb{V}^{n_1,n_2,n_3,n_4}_{\xi_1,\xi_2}(\mathbf{k},\mathbf{p},\mathbf{p}-\mathbf{k}) = \frac{1}{V_{\text{sys}}} \sum_{\mathbf{Q},\ell_1,\ell_2} \langle u_{n_1,\xi_1,\mathbf{p}}| e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}} \mathbb{P}_{\ell_1} |u_{n_3,\xi_1,\mathbf{k}}\rangle \, \mathcal{V}_{\ell_1,\ell_2}(\mathbf{p}-\mathbf{k}+\mathbf{Q}) \left[\langle u_{n_4,\xi_2,\mathbf{p}}| e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}} \mathbb{P}_{\ell_2} |u_{n_2,\xi_2,\mathbf{k}}\rangle\right]^\dagger,$$

$$\mathbb{V}^{n_1,n_2,n_3,n_4}_{\xi_1,\xi_2}(\mathbf{k},\mathbf{p},\mathbf{0}) = \frac{1}{V_{\text{sys}}} \sum_{\mathbf{Q},\ell_1,\ell_2} \langle u_{n_1,\xi_1,\mathbf{k}}| e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}} \mathbb{P}_{\ell_1} |u_{n_3,\xi_1,\mathbf{k}}\rangle \, \mathcal{V}_{\ell_1,\ell_2}(\mathbf{Q}) \left[\langle u_{n_4,\xi_2,\mathbf{p}}| e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}} \mathbb{P}_{\ell_2} |u_{n_2,\xi_2,\mathbf{p}}\rangle\right]^\dagger.$$

$$\text{(S60)}$$

From the first line of Eq. (S60), we find $\mathbb{V}^{n_1,n_2,n_3,n_4}_{\xi_1,\xi_2}(\mathbf{k},\mathbf{p},\mathbf{p}-\mathbf{k}) = \left[\mathbb{V}^{n_3,n_4,n_1,n_2}_{\xi_1,\xi_2}(\mathbf{p},\mathbf{k},\mathbf{k}-\mathbf{p})\right]^\dagger = \left[\mathbb{V}^{n_4,n_3,n_2,n_1}_{\xi_2,\xi_1}(\mathbf{k},\mathbf{p},\mathbf{p}-\mathbf{k})\right]^\dagger$.
In the first equality, band indices ($n_1 \leftrightarrow n_3, n_2 \leftrightarrow n_4$) and momenta ($\mathbf{k} \leftrightarrow \mathbf{p}$) are exchanged, while valley-spin indices are fixed. In the second equality, band indices ($n_1 \leftrightarrow n_4, n_2 \leftrightarrow n_3$) and valley-spin indices ($\xi_1 \leftrightarrow \xi_2$) are exchanged, while momenta are fixed. From the second line of Eq. (S60), we find $\mathbb{V}^{n_1,n_2,n_3,n_4}_{\xi_1,\xi_2}(\mathbf{k},\mathbf{p},\mathbf{0}) = \left[\mathbb{V}^{n_3,n_4,n_1,n_2}_{\xi_1,\xi_2}(\mathbf{k},\mathbf{p},\mathbf{0})\right]^\dagger = \left[\mathbb{V}^{n_4,n_3,n_2,n_1}_{\xi_2,\xi_1}(\mathbf{p},\mathbf{k},\mathbf{0})\right]^\dagger$.
In the first equality, band indices ($n_1 \leftrightarrow n_3, n_2 \leftrightarrow n_4$) are exchanged, while valley-spin indices and momenta are fixed. In the second equality, band indices ($n_1 \leftrightarrow n_4, n_2 \leftrightarrow n_3$), valley-spin indices ($\xi_1 \leftrightarrow \xi_2$), and momenta ($\mathbf{k} \leftrightarrow \mathbf{p}$) are all exchanged. These properties are used in the derivation of Eq. (S58) to eliminate factors of $\frac{1}{2}$. Collecting terms, we can write the mean-field Hamiltonian as

$$\hat{\mathcal{H}} = \sum_{\substack{\mathbf{k}\in\text{mBZ} \\ n,n',\xi,\xi'}} \hat{\Psi}^\dagger_{n,\xi,\mathbf{k}} \left[\mathbb{K}^{n,n'}_{\xi,\xi'}(\mathbf{k}) + \mathbb{H}^{n,n'}_{\xi,\xi'}(\mathbf{k}) + \mathbb{F}^{n,n'}_{\xi,\xi'}(\mathbf{k})\right] \hat{\Psi}_{n',\xi',\mathbf{k}},$$

$$\text{(S61)}$$

where the kinetic energy, Hartree potential, and Fock potential matrices are

$$\mathbb{K}^{n,n'}_{\xi,\xi'}(\mathbf{k}) = \delta_{\xi,\xi'}\delta_{n,n'} E_{n,\xi}(\mathbf{k}), \tag{S62a}$$

$$\mathbb{H}^{n,n'}_{\xi,\xi'}(\mathbf{k}) = +\delta_{\xi,\xi'} \sum_{\substack{\mathbf{p}\in\text{mBZ} \\ n_1,n_3,\xi_1}} \mathbb{V}^{n_1,n,n_3,n'}_{\xi_1,\xi}(\mathbf{p},\mathbf{k},\mathbf{0})\mathbb{D}^{n_1,n_3}_{\xi_1,\xi_1}(\mathbf{p}), \tag{S62b}$$

$$\mathbb{F}^{n,n'}_{\xi,\xi'}(\mathbf{k}) = -\sum_{\substack{\mathbf{p}\in\text{mBZ} \\ n_1,n_4}} \mathbb{V}^{n_1,n,n',n_4}_{\xi',\xi}(\mathbf{k},\mathbf{p},\mathbf{p}-\mathbf{k})\mathbb{D}^{n_1,n_4}_{\xi',\xi}(\mathbf{p}). \tag{S62c}$$

The Hartree and Fock terms depend on the density matrix; so they must be solved self-consistently. On the other hand, the kinetic piece is a one-particle term that does not depend on density; it does not need to be solved self-consistently.

A. Background Subtraction

Some amount of electron-electron interactions should have already been accounted for in *ab-initio* calculations to obtain the non-interacting band structure at charge neutrality in zero displacement field. To avoid double counting the effects of electron-electron interactions, we have to implement a background subtraction in HF renormalization. There is not yet a unique way to perform such a background subtraction; there are several possible candidates assumed in the literature. In this work, we consider a simple background subtraction that assumes the background charge density is that of a uniform half-filled lattice of isolated atoms. The reference density for such a background is $\mathbb{D}_{\text{ref}}(\mathbf{p}) = \frac{1}{2}\delta_{n,n'}\delta_{\xi,\xi'}$. Using this, we replace the density matrix in Eq. (S62) by $\delta\mathbb{D} = \mathbb{D} - \mathbb{D}_{\text{ref}}$, that is, only charge fluctuations away from the neutral background contribute to the HF Hamiltonian.

B. Convergence Criterion

In our iterative algorithm, no special updating method of the density matrix is used to accelerate convergence. Instead, at each iterative step, we simply recalculate the density matrix by filling the appropriate number of lowest energy eigenstates of the HF Hamiltonian obtained from the previous iteration. We find that even without using a special updating method, every calculation converged in a reasonable amount of time. In our simulations, convergence at step n of the calculation is defined as satisfying a criterion whereby the density matrix in the next step is not changing by much. We take this criterion to be $\|\mathbb{D}_n - \mathbb{D}_{n+1}\| < 10^{-8}$, where $\mathbb{D}_n$ is the density matrix at step n and the matrix norm is taken to be the absolute value of the maximal matrix element.

C. Initial Ansatz for Density Matrix

In this work, we only consider filling factor of one, leaving other possibilities to future studies. Said differently, we electron dope the system with one particle per moiré unit cell. We use two different ansatzes for the density matrix to search for different ground states using the HF algorithm. In the first, we are after an isospin-polarized ground state, and therefore, the initial ansatz is a density matrix that is half-filled equally in all four flavors plus an extra $N_{\mathbf{k}}^2$ electrons placed into the remaining available lowest-energy states in only one of the four flavors. In the second, we seek the symmetric ground state, and therefore, the initial ansatz is a density matrix that is filled equally in all four flavors. To determine which ground state is favored, we compare their total energy and choose the one with the lower energy. To ensure consistency, we confirm that C_{3z} is numerically preserved for all converged solutions. Ideally, to determine the "true" ground state, one has to, the very at least, perform many repeated searches with randomized ansatz for the initial density matrix and then compare their energies to choose the best solution. However, because we prefer to work with relatively large grids and to include a large number of bands to ensure convergence, doing randomized searches is currently computationally prohibitive for us. With that said, since our work is focused on the possibility of inducing isolated narrow Chern bands at filling factor of one, our choice of initial density matrices is physically motivated. Because the energy differences between the symmetric ground states and the $\mathcal{T}$-broken ground states for wide areas of parameter space are relatively large (on the order of millielectronvolts), we expect these $\mathcal{T}$-broken ground phases to be robust. Thus, the qualitative conclusion of the *existence* of isolated narrow Chern bands should survive in more refined mean-field searches even if the phase boundaries may shift.

V. SELF-CONSISTENT MEAN-FIELD PHASE DIAGRAMS

All of our simulations are done with a grid of $N_{\mathbf{k}} \times N_{\mathbf{k}} = 18 \times 18$. All results are simulated with at least 6 bands per isospin. However, many of the data points presented in the main text are obtained with 8 bands per isospin. We have checked that even a $N_{\mathbf{k}} \times N_{\mathbf{k}} = 12 \times 12$ grid gives qualitatively similar results to those with a $N_{\mathbf{k}} \times N_{\mathbf{k}} = 18 \times 18$ grid. Therefore, we believe that our calculations have converged. Of course, definitive conclusion about convergence is difficult to make since it would require, in addition to increasing mesh sizes, some sort of analytic argument about the thermodynamic limit [13]. Increasing beyond $N_{\mathbf{k}} \times N_{\mathbf{k}} = 18 \times 18$ is possible, but is costly and is beyond our available computational resources. There is a wide range of values for the dielectric constant used in theoretical works for twisted systems. There does not seem to be any consensus on what that value should be, even when given the experimental setup. Here, for the majority of the simulations, we use $\epsilon_r = 6$ and $\epsilon_r = 12$ for the dielectric constant. While the smaller value of $\epsilon_r = 6$ is closer to the experimental value of the dielectric constant of hexagonal boron nitride, the larger value of $\epsilon_r = 12$ is probably more realistic because it may account for some effects of internal screening. All in all, dielectric screening does indeed alter the phase boundaries of our system, but does not affect to a significant degree the existence of various phases of interest. For consistency, all results have been checked to respect C_{3z} symmetry and for cases of the symmetric state, have also been checked to additionally respect $\mathcal{T}$ symmetry.

The phase diagrams show a variety of phases. We characterize them based on the following criteria:

1. **$\mathcal{T}$-preserving metal**: In our HF calculations, we compare the energies of the converged solutions obtained from an initial ansatzes which populate equally all four isospins and those of the converged solutions calculated from initial ansatzes which populates only one isospin. Whenever the energy of the symmetric solution is lower (or very close to the symmetry-broken solution), we assign that state a $\mathcal{T}$-preserving metal. This state retains all symmetries of the non-interacting model, and is necessarily metallic because filling one electron per moiré unit cell with a fourfold degeneracy means that each band is only one-fourth filled.

2. **$\mathcal{T}$-breaking metal**: This state is obtained from the symmetry-broken ansatz. Therefore, it breaks time-reversal symmetry by unequally filling the four otherwise degenerate bands. We diagnose it as metallic by calculating the gap to excited states. When that gap is zero, the state is metallic.

3. **Correlated trivial insulator**: This state emerges from the symmetry-broken ansatz. Coulomb interactions stabilize the separation of the first conduction band in a single isospin flavor from the higher-energy bands. When the gap to excited states finite, i.e. $\delta_{\text{above}} > 0$ in the notation of Sec. II, we have an insulating state. If $\sum_{n \in \text{occupied}} \mathcal{C}_n = 0$, where n is band index, we call that state a trivial insulator.

4. **Anomalous Chern insulator**: This state differs from the correlated trivial insulator only in that $\sum_{n \in \text{occupied}} \mathcal{C}_n \neq 0$. However, the band immediately below the correlated gap is *not* isolated. So we cannot assign that band a Chern number. In that regard, we do not anticipate this state to be a favorable precursor to a fractional Chern state upon hole doping. Nonetheless, when the chemical potential is parked in the gap above that band, a non-zero Hall conductivity can be measured.

5. **Anomalous Chern insulator with isolated Chern band**: This state is a refined characterization of the anomalous Chern insulator state with $\delta_{\text{below}} > 0$. In other words, this state not only has quantized Hall conductivity but it also features an isolated narrow band with finite Chern number upon doping of which a fractional Chern insulator (FCI) state may emerge. As a consequence, this state can be labeled by two quantum numbers: the Chern number of the band which is isolated, $\mathcal{C}$, and the Chern number associated with the Hall conductivity of the gap right above that band, $\sum_{n \in \text{occupied}} \mathcal{C}_n$. These two quantum numbers usually agree for the majority of phase space. However, they do not necessarily have to the same, and indeed, there are points in phase space where they do not agree. Since our interest is in the isolated Chern band, for brevity and clarity, we show only $\mathcal{C}$ in our phase diagrams.

A. Antiparallel Stacking Order

In this section, we study the phase diagrams of the antiparallel stacking order using the four parameter sets listed in Table I.

1. With Some Lattice Relaxation

We first study the phase diagrams simulated using `Param_one`, which uses a large Fermi velocity $\gamma_0 = -3100$ meV and unequal interlayer hoppings $\{\gamma_{\text{AA}}, \gamma_{\text{AB}}\} = \{100, 120\}$ meV to include some effects of lattice relaxation. The phase diagram for $\epsilon_r = 12$ is shown in Fig. S21(a). The phase diagram is unsurprisingly complex, featuring all of the phases aforementioned, including metallic and insulating phases, some of which preserve $\mathcal{T}$-symmetry and some of which break $\mathcal{T}$-symmetry spontaneously. In S21(b,c,d), we show a representative band structure, polarized by isospin, for an anomalous Chern insulator without a gap below, a $\mathcal{T}$-preserving metal, and a $\mathcal{T}$-breaking metal respectively. These three phases are not the focus of our study; so we do not characterize these phases further. Moving onto the Chern insulating phases where there is an isolated Chern band right below Fermi level E_F, we find that for the antiparallel configuration reported here, there are three different phases with $\mathcal{C} = \pm 1$ and $\mathcal{C} = 2$. For these phases, we show the band structures both polarized by isospin and by layer index in S21(e,f,g). It is clear that all of these phases have average electron densities primarily localized on the trilayer substack. We shall characterize the lateral distribution of the electron density later. For each of the Chern phases, we also show the associated Berry curvature and quantum metric distribution in $\mathbf{k}$-space for the representative values of (θ, Δ). Compared to the $|\mathcal{C}| = 1$ phase, the $\mathcal{C} = 2$ phase has smaller λ, which is favorable for the emergence of fractionalized phases. For this phase, we observe in Fig. S21(e) that the Chern band, chosen for the optimal f.o.m., is relatively narrow with Berry curvature localized around the $\bar{\Gamma}$ point in momentum space. For this state, we have $\{\theta, \Delta, \delta_{\text{above}}, \delta_{\text{below}}, W, \lambda\} = \{1.20°, -18 \text{ meV}, 11.13 \text{ meV}, 14.62 \text{ meV}, 7.25 \text{ meV}, 1.32\}$.

Next, we focus on just the Chern phases with an isolated band. We further characterize these phases in Fig. S22. Here, $\times$ decates (θ, Δ) where other less interesting phases stabilize. For values of (θ, Δ) where we have a spectrally-isolated Chern band, we show the band gap of that band in Fig. S22(a), the bandwidth in Fig. S22(b), the Chern number in Fig. S22(c), the trace condition violation in Fig. S22(d), and the band-average layer polarization in Fig. S22(e). The band average layer polarization is defined as $\bar{\chi}_\ell = \frac{1}{N_{\mathbf{k}}^2} \sum_{\mathbf{k}} \chi_\ell(\mathbf{k})$, where $\chi_\ell(\mathbf{k})$ is the layer polarization for state $\mathbf{k}$ in the band of interest. This is an average measure for the charge localization of an entire band, not just a single Bloch state, and thus provides a further characterization of the charge distribution. From Fig. S22, we see that as a general trend, the $\mathcal{C} = 2$ phase features smaller bandwidths, larger band gaps, and smaller trace condition violations compared to the $\mathcal{C} = \pm 1$ phases. Therefore, this phase is not only interesting because it contains a higher Chern number, but it is also interesting because it contains favorable spectral and quantum geometric properties for further strong electron-electron effects at partial filling.

To characterize the lateral charge distribution, we choose to focus only on the $\mathcal{C} = 2$ phase. Using the value of (θ, Δ) that yields optimal f.o.m., we plot the charge density in real space summed over all states in the band of interest for all five layers, resolved down to the sublattices, in Fig. S23. First, we show the fractional share of the charge for each sublattice on each layer in the bar graph. As is clear, most of the charges are localized on the A sublattice on layer 1. However, we see that there are non-zero, albeit much smaller, charge densities on all the other layers. This is reflected in the real-space charge distribution. On layer 1, we see a weak modulation of the charge density on the moiré scale showing slightly higher concentrations at the α and β regions. The pattern on layer 2 is similar to the pattern on layer 1, though layer 2 has significantly less charge. On layers 3 and 4, the charge is concentrated at the α regions. On the basis of the charge density summed over states in the active band, we can conclude that this is a Hall crystal because it features a charge density that not only occupies layers unequally but also fluctuates laterally on the moiré scale. However, we should note that the fluctuation is only mild. Also, we have not summed all states in the other valence bands, which if done, may change the charge density profile.

We now move onto the phase diagram simulated using `Param_one` with $\epsilon_r = 6$. The results are shown in Figs. S24,S25, and S26 analogous to Figs. S21,S22, and S23 for $\epsilon_r = 12$. In the present case, because $\epsilon_r = 6$ is smaller, Coulomb interactions are twice as strong. Not surprisingly then, we observe in the phase diagram of Fig. S24 that symmetry-broken phases dominate phase space. The regions of Chern phases are enlarged compared to the simulations with $\epsilon_r = 12$. However, all qualitative features remain the same as before, suggesting that our conclusions are robust in a range of electron-electron interactions.

Here, the optimal f.o.m. occurs $\{\theta, \Delta, \delta_{\mathrm{above}}, \delta_{\mathrm{below}}, W, \lambda\} = \{1.25°, -24\,\mathrm{meV}, 26.94\,\mathrm{meV}, 17.37\,\mathrm{meV}, 12.36\,\mathrm{meV}, 1.24\}$. The optimal trace condition violation here is slightly better than for $\epsilon_r = 12$. The charge density distribution shown in Fig. S26 is nearly indistinguishable from the charge distribution shown in Fig. S23 (though they are not identical as the reader can attest upon careful inspection).

2. Without Lattice Relaxation

In this section, we analyze the phase diagrams simulated with `Param_two` where $\{\gamma_{\mathrm{AA}}, \gamma_{\mathrm{AB}}\} = \{120, 120\}$ meV. This batch of results, shown in Fig. S27, represents a twisted structure without any lattice relaxation. It is an idealization that we do *not* expect an experimental platform to respect. It is nonetheless good to analyze this case to assess the stability of our results. In brief, all relevant trends observed with lattice relaxation are also observed without, again demonstrating robustness of our results. In more details, the top panel of Fig. S27 shows results for $\epsilon_r = 12$ while the bottom panel shows results for $\epsilon_r = 6$. In both cases, we observe a prominent $\mathcal{C} = 2$ phase at negative Δ over a range of angles. Increasing Coulomb interactions enlarges the phase-space area of this phase. For $\epsilon_r = 12$, optimal f.o.m. occurs at $\{\theta, \Delta, \delta_{\mathrm{above}}, \delta_{\mathrm{below}}, W, \lambda\} = \{1.30°, -21\,\mathrm{meV}, 10.14\,\mathrm{meV}, 18.33\,\mathrm{meV}, 8.37\,\mathrm{meV}, 1.67\}$. For $\epsilon_r = 6$, optimal f.o.m. occurs at $\{\theta, \Delta, \delta_{\mathrm{above}}, \delta_{\mathrm{below}}, W, \lambda\} = \{1.30°, -27\,\mathrm{meV}, 20.68\,\mathrm{meV}, 19.40\,\mathrm{meV}, 12.02\,\mathrm{meV}, 1.44\}$. The band structures for these optimal states are shown in Fig. S27(b,g). We notice that the trace condition violations here are generally larger than the results with lattice relaxation. Here, we also observe that the $\mathcal{C} = 2$ phase exhibits enhanced Berry curvature at the $\bar{\Gamma}$ point in momentum space. Also, optimal angles are pushed higher without lattice relaxation compared to with lattice relaxation. Therefore, it appears that lattice relaxation is favorable to the stabilization of the $\mathcal{C} = 2$ phase we seek.

3. With Reduced Fermi Velocity γ_0

Here, we study the phase diagrams generated with parameters from `Param_three` and `Param_four`. The Fermi velocity has been reduced to $\gamma_0 = -2700$ meV. The results with lattice relaxation $\{\gamma_{\mathrm{AA}}, \gamma_{\mathrm{AB}}\} = \{80, 100\}$ meV are shown in the top panel of Fig. S28, while the results without lattice relaxation $\{\gamma_{\mathrm{AA}}, \gamma_{\mathrm{AB}}\} = \{100, 100\}$ meV are shown in the bottom panel of Fig. S28. For those sets of simulations, we set $\epsilon_r = 12$. We observe $\mathcal{C} = 2$ phases in both scenarios. We see that lattice relaxation pushes the range of angles where this phase is observed to higher values. With lattice relaxation, optimal f.o.m. occurs at $\{\theta, \Delta, \delta_{\mathrm{above}}, \delta_{\mathrm{below}}, W, \lambda\} = \{1.25°, -18\,\mathrm{meV}, 12.09\,\mathrm{meV}, 15.76\,\mathrm{meV}, 7.13\,\mathrm{meV}, 1.14\}$. Without lattice relaxation, optimal f.o.m. occurs at $\{\theta, \Delta, \delta_{\mathrm{above}}, \delta_{\mathrm{below}}, W, \lambda\} = \{1.35°, -21\,\mathrm{meV}, 11.10\,\mathrm{meV}, 20.64\,\mathrm{meV}, 8.43\,\mathrm{meV}, 1.35\}$. We again observe concentration of the Berry curvature at the zone center for the $\mathcal{C} = 2$ states, consistent with previous results using different sets of parameters.

B. Parallel Stacking Order

In this section, we study the phase diagrams of the parallel stacking order using the four parameter sets listed in Table I. This stacking order is richer than the antiparallel stacking order because it features $\mathcal{C} = 2$ and $\mathcal{C} = 3$ phases. As we shall show, it is also more sensitive to parameter choice than the previous analysis.

1. With Some Lattice Relaxation

Using parameters from `Param_one`, we simulate the results shown in Fig. S29 with $\epsilon_r = 12$. Similar to the antiparallel case, we observe incredible complexity in the evolution of different phases in Fig. S29(a). This is because the band structure for this system is quite involved. In Fig. S29(b,c,d), we show an anomalous Chern insulator without an isolator band, a $\mathcal{T}$-preserving metal, and a $\mathcal{T}$-breaking metal respectively. Unlike in the antiparallel configuration, we observe two prominent higher-Chern phases: $\mathcal{C} = 2$ and $\mathcal{C} = 3$. The $\mathcal{C} = 2$ phase occurs at lower angles than the $\mathcal{C} = 3$ phase. Both phases roughly occupy the same areas in phase space. The optimal $\mathcal{C} = 2$ state occurs at $\{\theta, \Delta, \delta_{\mathrm{above}}, \delta_{\mathrm{below}}, W, \lambda\} = \{1.10°, -21\,\mathrm{meV}, 7.85\,\mathrm{meV}, 27.04\,\mathrm{meV}, 3.75\,\mathrm{meV}, 0.87\}$. The band structure for this state is shown in Fig. S29(e). We draw particular attention to the very narrow band right below Fermi level. The Berry curvature for this state is more delocalized throughout the mBZ compared to the other nontrivial states that we have shown. The trace condition violation here is $\lambda = 0.87$. We also see that the Bloch states for the active band are localized mostly in the trilayer substack. For the $\mathcal{C} = 3$ phase, the optimal state occurs at $\{\theta, \Delta, \delta_{\mathrm{above}}, \delta_{\mathrm{below}}, W, \lambda\} = \{1.30°, -21\,\mathrm{meV}, 9.54\,\mathrm{meV}, 25.55\,\mathrm{meV}, 11.04\,\mathrm{meV}, 0.55\}$. The band structure for this state is shown in Fig. S29(f). The band of interest has some dispersion after HF renormalization. The Berry curvature is centered around the $\bar{\Gamma}$ point. The trace condition violation is especially small here, only around $\lambda = 0.55$.

Focusing on the higher-Chern-nontrivial phases, we show further spectral and topological characterization in Fig. S30. Here, $\times$ decates (θ, Δ) where other less interesting phases stabilize. For values of (θ, Δ) where we have a spectrally-isolated Chern band, we show the band gap of that band in Fig. S30(a), the bandwidth in Fig. S30(b), the Chern number in Fig. S30(c), the trace condition violation in Fig. S30(d), and the band-average layer polarization in Fig. S30(e). Some trends are evident, the $\mathcal{C} = 2$ phase has smaller band gaps and bandwidths than the $\mathcal{C} = 3$ phase. However, the $\mathcal{C} = 2$ phase has large λ than the the $\mathcal{C} = 3$ phase. Both phases have average $\bar{\chi}_\ell \lesssim 2$, including that they both harbor majority charge density in the trilayer substack as opposed to the bilayer substack.

Moving onto the lateral charge distribution, we map the charge density on each of the five layers for the two optimal states aforementioned. In Fig. S31, we show the charge density for the $\mathcal{C} = 2$ state at $\theta = 1.10°$ and $\Delta = -21$ meV. From the bar graph, we see that most of the charge resides on layer 1 on the A sublattice. There are smaller distributions on the other layers. Of note, on layer 5, the charge density is localized also exclusively on the B sublattice. Laterally, the charge distribution on layer 1 is fairly uniform, with some inhomogeneity due to slightly higher concentrations at the α and γ regions. On layer 1, the charge density is depleted at the α regions, while it is peaked at the α regions on layers 3 and 4. On layer 5, there is small charge density is localized at the β regions. In Fig. S32, we show the charge density for the $\mathcal{C} = 3$ state at $\theta = 1.30°$ and $\Delta = -21$ meV. The charge density here is quite similar to that shown in Fig. S31. On layer 1, we see that charges coalesce slightly more at the γ regions than other regions. On layer 2, the charge is fairly uniform. On layers 3 and 4, the charges localize at the α regions. On layer 5, there is very little charge, with inhomogeneity due to slight localization at the γ regions. Indeed, in both phases, the charge is *not* uniform across the layers and within each layer, featuring clear inhomogeneities.

Next, we consider the phase diagram simulated with $\epsilon_r = 6$ for the parallel configuration using `Param_one`. The results are shown in Fig. S33. In Fig. S33(a), we notice that because Coulomb interactions are stronger, the spontaneously-formed phases outcompete the symmetric phase. Remarkably, we find four prominent Chern phases with $\mathcal{C} = -1$, $\mathcal{C} = -2$, $\mathcal{C} = 2$, and $\mathcal{C} = 3$. The first two occur at positive Δ while the last two occur at negative Δ. The band structures for the $\mathcal{C} = -1$ and $\mathcal{C} = -2$ phases are shown in Fig. S33(e,f) respectively. For the $\mathcal{C} = -1$ phase, the trace condition violation is higher than the other phases; it is $\lambda = 2.21$ for the present case. For the $\mathcal{C} = -2$ phase, the active band is quite dispersive with gaps to other bands that are not as large. These two phases are localized primarily, but not exclusively, on the bilayer substack. It is also worth noting that these two phases do not exist when $\epsilon_r = 12$. So they are not stable to changing Coulomb interactions in the range of parameters of interest. For the $\mathcal{C} = 2$ and $\mathcal{C} = 3$ phases, we show the band structures with optimal f.o.m. in Fig. S33(c,d). For the $\mathcal{C} = 2$, we observe a narrow isolated band with large gaps both above and below said band. The Berry curvature distribution and quantum metric do peak around the $\bar{\Gamma}$ point but also have non-negligible spread throughout the mBZ. The optimal f.o.m. here is $\{\theta, \Delta, \delta_{\text{above}}, \delta_{\text{below}}, W, \lambda\} = \{1.15°, -30 \text{ meV}, 16.33 \text{ meV}, 40.88 \text{ meV}, 6.51 \text{ meV}, 0.74\}$. For the $\mathcal{C} = 3$ phase, the band is, again, dispersive, but is is well isolated from the rest of the band structure. The optimal f.o.m. here is $\{\theta, \Delta, \delta_{\text{above}}, \delta_{\text{below}}, W, \lambda\} = \{1.40°, -27 \text{ meV}, 21.22 \text{ meV}, 30.13 \text{ meV}, 21.44 \text{ meV}, 0.48\}$. The Berry curvature here is also peaked at the $\bar{\Gamma}$ point. Further characterization of the Chern-nontrivial phases is shown in Fig. S34. From there, we observe clearly at the Chern phases at positive Δ have comparatively larger trace condition violations and are localized primarily on the bilayer substack as diagonosed by $\bar{\chi}_\ell$. For the Chern phases at negative Δ, we observe that the trace condition violation is generally smaller for the $\mathcal{C} = 3$ phase compared to the $\mathcal{C} = 2$ phase. The ratio of the bandwidth to band gap is smaller for the $\mathcal{C} = 2$ phase than the $\mathcal{C} = 3$ phase. The real-space charge densities for these two phases are shown in Figs. S35 and S36. They are very similar to Figs. S31 and S32. Therefore, the discussion about the latter figures apply to the former as well.

2. Without Lattice Relaxation

Here, we analyze the phase diagrams for the parallel configuration simulated using `Param_two` which sets $\{\gamma_{\text{AA}}, \gamma_{\text{AB}}\} = \{120, 120\}$ meV. This parameter set represents the absence of lattice relaxation. We first focus on $\epsilon_r = 12$. In Fig. S37, we show the results for $\epsilon_r = 12$ and, in Fig. S38, we show the results for $\epsilon_r = 6$. From Fig. S37(a), we observe that the $\mathcal{C} = 3$ phase has been reduced significantly in size compared to the results with lattice relaxation. However, the $\mathcal{C} = 2$ phase remains prominent without lattice relaxation. As before, the $\mathcal{C} = 2$ phase features a relatively narrow and isolated band with favorable quantum geometry. In Fig. S37(e), we show a representative band structure from the optimal state with f.o.m. here is $\{\theta, \Delta, \delta_{\text{above}}, \delta_{\text{below}}, W, \lambda\} = \{1.10°, -18 \text{ meV}, 8.59 \text{ meV}, 14.32 \text{ meV}, 3.92 \text{ meV}, 0.89\}$. The $\mathcal{C} = 3$ phase here does not look appealing here because it does not have as small a trace condition violation compared to the results with lattice relaxation. The band of interest is dispersive and is not well isolated from the rest of the band structure. Even the optimal state for this phase, shown in S37(f), does not appear optimal from the point of view of being conducive to further electron-electron interaction effects at partial doping. The optimal state here is $\{\theta, \Delta, \delta_{\text{above}}, \delta_{\text{below}}, W, \lambda\} = \{1.30°, -15 \text{ meV}, 4.29 \text{ meV}, 5.37 \text{ meV}, 17.40 \text{ meV}, 1.14\}$. Therefore, we observe that the lattice relaxation is important for the existence of the $\mathcal{C} = 3$ phase at $\epsilon_r = 12$. The $\mathcal{C} = 2$ phase remains robust without lattice relaxation at $\epsilon_r = 12$.

We now turn our attention to the results in Fig. S38 simulated with $\epsilon_r = 6$. Here, Coulomb interactions are stronger so symmetry-breaking phases dominate. We again observe Chern phases at both positive and negative Δ but for reasons described in the previous section with lattice relaxation, we only focus on the phases at negative Δ. At this value of ϵ_r,

we recover both the $\mathcal{C} = 2$ and $\mathcal{C} = 3$ to their former prominence. The $\mathcal{C} = 2$ phase again has quite narrow bands with favorable quantum geometry, as shown in S38(e). The optimal state here occurs at $\{\theta, \Delta, \delta_{\text{above}}, \delta_{\text{below}}, W, \lambda\} = \{1.30°, -30 \text{ meV}, 19.04 \text{ meV}, 33.48 \text{ meV}, 6.13 \text{ meV}, 0.99\}$. The $\mathcal{C} = 3$ phase is still dispersive but it is isolated from the rest of the energy spectrum, as shown in S38(f). The Berry curvature and quantum metric are also centered at $\bar{\Gamma}$. The optimal state here occurs at $\{\theta, \Delta, \delta_{\text{above}}, \delta_{\text{below}}, W, \lambda\} = \{1.35°, -24 \text{ meV}, 11.25 \text{ meV}, 12.63 \text{ meV}, 25.64 \text{ meV}, 0.78\}$. The existence of both the $\mathcal{C} = 2$ and $\mathcal{C} = 3$ phases at $\epsilon_r = 6$ suggest that these phases are stable against the inclusion of lattice relaxation or lack thereof. However, one needs to tune the strength of Coulomb interactions in order to stabilize them. Lattice relaxation is definitely helpful to that effort.

3. With Reduced Fermi Velocity γ_0

Finally, we examine the phase diagrams simulated with `Param_three` and `Param_four` as shown in Figs. S39 and S40 respectively. These two parameter sets have reduced $\gamma_0 = -2700 \text{ meV}$. The former includes some lattice relaxation while the latter does not. For both simulations, we use $\epsilon_r = 12$. We observe in both cases that the $\mathcal{C} = 2$ and $\mathcal{C} = 3$ phases exist. However, with lattice relaxation, the $\mathcal{C} = 3$ is much more prominent. In both cases, the qualitative features observed before remain: the $\mathcal{C} = 2$ state has narrow, well-isolated bands with favorable quantum geometry, while the $\mathcal{C} = 3$ state also has favorable quantum geometry, but its bands are more dispersive. With lattice relaxation, the optimal $\mathcal{C} = 2$ state occurs at $\{\theta, \Delta, \delta_{\text{above}}, \delta_{\text{below}}, W, \lambda\} = \{1.15°, -24 \text{ meV}, 6.50 \text{ meV}, 35.02 \text{ meV}, 5.10 \text{ meV}, 1.01\}$, while the optimal $\mathcal{C} = 3$ state occurs at $\{\theta, \Delta, \delta_{\text{above}}, \delta_{\text{below}}, W, \lambda\} = \{1.35°, -18 \text{ meV}, 11.22 \text{ meV}, 15.27 \text{ meV}, 16.32 \text{ meV}, 0.54\}$. With lattice relaxation, the optimal $\mathcal{C} = 2$ state occurs at $\{\theta, \Delta, \delta_{\text{above}}, \delta_{\text{below}}, W, \lambda\} = \{1.30°, -24 \text{ meV}, 9.68 \text{ meV}, 28.45 \text{ meV}, 4.65 \text{ meV}, 1.01\}$, while the optimal $\mathcal{C} = 3$ state occurs at $\{\theta, \Delta, \delta_{\text{above}}, \delta_{\text{below}}, W, \lambda\} = \{1.40°, -18 \text{ meV}, 4.45 \text{ meV}, 10.56 \text{ meV}, 19.84 \text{ meV}, 1.01\}$.

C. Summary of Observations

We close this section with a few remarks summarizing the general trends observed from the extensive simulation results reported. Summary information about the optimal $\mathfrak{f.o.m.}$ states is tabulated in Tables III, IV, and V. We shall use these tables for guidance. All of the relevant higher-Chern phases are observed at negative Δ where the charge density of the band of interest is localized primarily on layer 1 of the trilayer substack. For the antiparallel configuration, we observe only a $\mathcal{C} = 2$ phase, while for the parallel configuration, we observe a $\mathcal{C} = 2$ and a $\mathcal{C} = 3$ phase. The $\mathcal{C} = 2$ phase in the antiparallel configuration appears most robust to changes to parameters. The phases in the parallel configuration are more sensitive to parameter changes, as detailed below.

For the $\mathcal{C} = 2$ phase of the antiparallel configuration, we find the following:

1. The optimal angles are around $\theta \sim 1.2 - 1.3°$. The range of optimal displacement fields is $\Delta \in [-27, -18]$ meV.

2. For $\gamma = -3100$ meV, increasing Coulomb interactions increases the phase-space area of this phase, decreases the optimal λ, and increases the bandwidths and band gaps.

3. Lattice relaxation appears to be beneficial to this phase as it decreases the optimal λ while simultaneously lowers the bandwidths. The dependence of band gaps on lattice relaxation appears more complicated: it increases δ_{above} but decreases δ_{below}.

4. With reduced $\gamma_0 = -2700$ meV, this phase remains robust. Here, we also observe that adding lattice relaxation decreases λ and decreases W, consistent with the other simulations.

For the $\mathcal{C} = 2$ phase of the parallel configuration, we observe the following:

1. The optimal angles are around $\theta \sim 1.1 - 1.3°$. The range of optimal displacement fields is $\Delta \in [-30, -18]$ meV.

2. The bandwidths are remarkably narrower in this phase compared to the other nontrivial Chern phases in both configurations.

3. The trace condition violations for this phase are generally smaller than the $\mathcal{C} = 2$ phase of the antiparallel configuration but are generally larger than the $\mathcal{C} = 3$ phase.

4. Strengthening Coulomb interactions decreases λ when lattice relaxation is kept but increases λ when lattice relaxation is ignored. However, increasing Coulomb interactions increases band gaps and bandwidths whether or not lattice relaxation is accounted for. Therefore, increasing Coulomb interactions is helpful to this phase. However, the phase areas have a complicated relationship on the strength of Coulomb interactions.

Parameter set	ϵ_r	# of points	θ (degrees)	Δ (meV)	δ_{above} (meV)	δ_{below} (meV)	W (meV)	λ
`Param_one`	12	26	1.20°	−18	11.26	14.62	7.25	1.32
`Param_one`	6	39	1.25°	−24	26.94	17.37	12.36	1.24
`Param_two`	12	23	1.30°	−21	10.14	18.33	8.38	1.67
`Param_two`	6	38	1.30°	−27	20.68	19.40	12.02	1.44
`Param_three`	12	33	1.25°	−18	12.09	15.76	7.13	1.14
`Param_four`	12	34	1.35°	−21	11.10	20.64	8.43	1.35

Table III. **Summary of observations of the mean-field $\mathcal{C} = 2$ phase in the antiparallel configuration.** The first column lists the parameter set used. The second column denotes the dielectric constant ϵ_r. The third column provides the number of data points belonging to such phase and serves as a measure for the area of that phase within the phase diagram. For the data point within that area which minimizes the figure of merit in Eq. S15, we display the angle and displacement field where that occurs in columns four and five, the gaps above and below the first conduction band in columns six and seven, the bandwidth of that band in column eight, and the corresponding trace condition in column nine.

Parameter set	ϵ_r	# of points	θ (degrees)	Δ (meV)	δ_{above} (meV)	δ_{below} (meV)	W (meV)	λ
`Param_one`	12	22	1.10°	−21	7.85	27.04	3.75	0.87
`Param_one`	6	16	1.15°	−30	16.33	40.88	6.51	0.74
`Param_two`	12	27	1.10°	−18	8.60	14.32	3.92	0.89
`Param_two`	6	26	1.30°	−30	19.04	33.48	6.13	0.99
`Param_three`	12	10	1.15°	−24	6.50	35.02	5.10	1.01
`Param_four`	12	24	1.30°	−24	9.68	28.45	4.65	1.01

Table IV. **Summary of observations of the mean-field $\mathcal{C} = 2$ phase in the parallel configuration.** The first column lists the parameter set used. The second column denotes the dielectric constant ϵ_r. The third column provides the number of data points belonging to such phase and serves as a measure for the area of that phase within the phase diagram. For the data point within that area which minimizes the figure of merit in Eq. S15, we display the angle and displacement field where that occurs in columns four and five, the gaps above and below the first conduction band in columns six and seven, the bandwidth of that band in column eight, and the corresponding trace condition in column nine.

5. For $\gamma = -3100$ meV, lattice relaxation does not appear to affect the phase much. While including lattice does decrease optimal λ, it also decrease the phase areas.

6. With reduced $\gamma_0 = -2700$ meV, this phase area is reduced noticeably with relaxation, but the phase area remains relatively similar to the other parameter choices when relaxation is included.

For the $\mathcal{C} = 3$ phase of the parallel configuration, we note the following:

1. The optimal angles are around $\theta \sim 1.3 - 1.4°$. The range of optimal displacement fields is $\Delta \in [-27, -18]$ meV.

2. As a general trend, this phase has the lowest λ compared to the other two phases.

3. This phase contains active bands that are appreciably more dispersive than the bands in the other nontrivial phases.

4. For $\gamma_0 = -3100$ meV, lattice relaxation is favorable to this phase as it decreases λ, decreases the bandwidths, and increases the band gaps.

5. Increasing Coulomb interactions stabilizes this phase, reducing λ and increasing band gaps as well as bandwidths.

6. With reduced $\gamma_0 = -2700$ meV, this phase area is reduced noticeably without relaxation. This phase follows the opposite trend compared to the $\mathcal{C} = 2$ phase in the parallel configuration.

We emphasize also that ranges of (θ, Δ) where nontrivial Chern phases exist after HF renormalization do not often overlap with the ranges of (θ, Δ) where Chern bands appear in the non-interacting limit, suggesting the importance of Coulomb interactions.

VI. NUMERICAL STABILITY ANALYSIS

We briefly address the question of convergence of our results in this section. For concreteness, we focus only on our principal results presented in the main text. That is, we show that the optimal states for both configurations simulated with `Param_one` and $\epsilon_r = 12$ are stable with changes in the number of bands included and the mesh size. In Fig. S41, we show the optimal $\mathcal{C} = 2$ state for the antiparallel configuration. In Figs. S42 and S43, we show the optimal $\mathcal{C} = 2$ and $\mathcal{C} = 3$ states for the parallel

Parameter set	ϵ_r	# of points	θ (degrees)	Δ (meV)	δ_{above} (meV)	δ_{below} (meV)	W (meV)	λ
Param_one	12	33	$1.30°$	-21	9.54	25.55	11.04	0.55
Param_one	6	38	$1.40°$	-27	21.22	30.13	21.44	0.48
Param_two	12	9	$1.30°$	-15	4.29	5.37	17.40	1.14
Param_two	6	17	$1.35°$	-24	11.25	12.63	25.64	0.78
Param_three	12	37	$1.35°$	-18	11.22	15.27	16.32	0.54
Param_four	12	14	$1.40°$	-18	4.45	10.56	19.84	1.01

Table V. **Summary of observations of the mean-field $\mathcal{C} = 3$ phase in the parallel configuration.** The first column lists the parameter set used. The second column denotes the dielectric constant ϵ_r. The third column provides the number of data points belonging to such phase and serves as a measure for the area of that phase within the phase diagram. For the data point within that area which minimizes the figure of merit in Eq. S15, we display the angle and displacement field where that occurs in columns four and five, the gaps above and below the first conduction band in columns six and seven, the bandwidth of that band in column eight, and the corresponding trace condition in column nine.

configuration respectively. In all three cases, we see that already at $N_{\mathbf{k}} \times N_{\mathbf{k}} = 6 \times 6$, we have excellent agreement in the energetics. However, the trace condition violations do deviate from the "true" values at such a smaller mesh. For a larger mesh of $N_{\mathbf{k}} \times N_{\mathbf{k}} = 12 \times 12$, we observe in all three cases that the violations are within 6% of the "true" values. This analysis adds convincing evidence that our principal results concerning the nontrivial topological states have likely converged numerically.

VII. PHASE DIAGRAM COMPARISON WITH LAYER-*IN*DEPENDENT COULOMB POTENTIAL

In many previous studies on related twisted and non-twisted systems, the Coulomb potential is often regarded as independent of the layer index, i.e. all of the layers are treated as residing on the same plane. This is probably a good approximation when the number of layers is small (say two or three layers) or when the distance to the gates is much greater than the thickness of the stack, $D/2 \gg N_\ell d$, where N_ℓ is the number of layers. In this work, neither of these assumptions is true: (1) we have a reasonably thick stack (5 layers) and (2) the gates are not set significantly far away since experimentally they can be rather close to the active material. We have therefore considered the full layer dependence of the Coulomb potential in this work. In this section, we compare the results of considering the full layer dependence to those obtained by neglecting that layer index.

From the practical point of view, neglecting the layer index can lead to a substantial reduction in required computational resources for HF calculations. This is due to the fact that the form factors no longer need to be resolved to individual layers, significantly lowering memory and CPU requirements for the evaluations of the Coulomb matrix elements, which, importantly, are the most intensive routine in our the HF algorithm. Without layer index, Eqs. (S56) and (S57) simplify:

$$\mathbb{V}_{\xi_1,\xi_2}^{n_1,n_2,n_3,n_4}(\mathbf{k},\mathbf{p},\mathbf{q}) = \frac{1}{V_{\text{sys}}} \sum_{\mathbf{Q}} \mathbb{A}_{\xi_1}^{n_1,n_3}(\mathbf{k},\mathbf{q}+\mathbf{Q})\mathcal{V}(\mathbf{q}+\mathbf{Q}) \left[\mathbb{A}_{\xi_2}^{n_4,n_2}(\mathbf{p}-\mathbf{q},\mathbf{q}+\mathbf{Q})\right]^\dagger,$$

$$\mathbb{A}_{\xi}^{n,n'}(\mathbf{k},\mathbf{q}+\mathbf{Q}) = \sum_{\mathbf{G},\sigma,\ell} \phi_{n,\xi,\sigma,\ell,\mathbf{k}+\mathbf{G}+\mathbf{q}+\mathbf{Q}}^\dagger \phi_{n',\xi,\sigma,\ell,\mathbf{k}+\mathbf{G}} = \langle u_{n,\xi,\mathbf{k}+\mathbf{q}}| e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}} |u_{n',\xi,\mathbf{k}}\rangle, \qquad (S63)$$

$$\mathcal{V}(\mathbf{q}) = \frac{e^2}{2\epsilon_0\epsilon_r|\mathbf{q}|} \tanh\left(D|\mathbf{q}|/2\right).$$

As illustrated, there is no longer a need to calculate $\mathbb{A}_{\xi,\ell}^{n,n'}$ since the only quantities which are needed to calculate $\mathbb{V}_{\xi_1,\xi_2}^{n_1,n_2,n_3,n_4}$ are $\mathbb{A}_{\xi}^{n,n'} = \sum_\ell \mathbb{A}_{\xi,\ell}^{n,n'}$. This simple change might have important qualitative implications for the phase diagram, which we now examine. The phase diagrams simulated using Param_one with $\epsilon_r = 12$ and $\epsilon_r = 6$ for both antiparallel and parallel configurations are shown in Fig. S44. Qualitatively, the phase diagrams in Fig. S44 are similar to the phase diagrams obtained using the layer-dependent potential. For instance, there is a robust $\mathcal{C} = 2$ phase in the antiparallel configuration for both values of the dielectric constant. There is also a robust $\mathcal{C} = 2$ phase and a robust $\mathcal{C} = 3$ phase for the parallel configuration, again, for both values of the dielectric constant. To further understand the difference between using the layer-dependent potential and layer-independent potential, we tabulate properties of the optimal states in Table VI. In all cases, we observe that the optimal states are pushed to lower values of Δ when layer dependence is included. This is perfectly sensible because layer screening reduces the externally-applied displacement field. For the antiparallel configuration, the optimal states appear quite similar with and without layer dependence. For the parallel configuration, λ is reduced without layer dependence. Thus, it is important to include layer dependence in order to accurately capture quantum geometry as closeness to ideal quantum geometry is an important aspect of our conclusions. This alone suggests the necessity of using the layer-dependent Coulomb potential energy for realistic modeling. We leave further analysis of the difference between the two potential energy functions to future studies.

Configuration	Layer dependence?	ϵ_r	$\mathcal{C}$	# of points	θ (degrees)	Δ (meV)	δ_{above} (meV)	δ_{below} (meV)	W (meV)	λ
Antiparallel	No	12	2	23	1.20°	−12	11.64	12.53	7.45	1.32
Antiparallel	Yes	12	2	26	1.20°	−18	11.13	14.62	7.25	1.32
Antiparallel	No	6	2	48	1.25°	−15	27.37	16.14	14.04	1.27
Antiparallel	Yes	6	2	39	1.25°	−24	26.94	17.37	12.36	1.24
Parallel	No	12	2	25	1.10°	−15	8.83	25.49	4.39	0.75
Parallel	Yes	12	2	22	1.10°	−21	7.85	27.04	3.75	0.87
Parallel	No	6	2	29	1.15°	−18	18.29	35.32	7.69	0.59
Parallel	Yes	6	2	16	1.15°	−30	16.33	40.88	6.51	0.74
Parallel	No	12	3	27	1.25°	−12	11.17	14.56	12.56	0.45
Parallel	Yes	12	3	33	1.30°	−21	9.54	25.55	11.04	0.55
Parallel	No	6	3	36	1.35°	−18	21.77	27.96	18.96	0.36
Parallel	Yes	6	3	38	1.40°	−27	21.22	30.13	21.44	0.48

Table VI. **Comparison of the optimal states obtained with layer-dependent and layer-independent Coulomb potential energy.** Column one indicates the configuration. Column two indicates whether a layer-independent or layer-dependent Coulomb potential energy is used. Column three denotes the dielectric constant. Columns four and five indicate the Chern phase and its number of points in phase space. For the data point within that area which minimizes the figure of merit in Eq. S15, we display the angle and displacement field where that occurs in columns six and seven, the gaps above and below the first conduction band in columns eight and nine, the bandwidth of that band in column ten, and the corresponding trace condition in column eleven.

VIII. NUMERICAL EVALUATION OF QUANTUM GEOMETRIC TENSOR

The gauge-invariant geometric tensor for Bloch states $|u_{\mathbf{k}}\rangle$ from a single isolated band is [14, 15]

$$\mathcal{Q}_{\mu\nu}(\mathbf{k}) = \langle \partial_\mu u_{\mathbf{k}} | \partial_\nu u_{\mathbf{k}} \rangle - \langle \partial_\mu u_{\mathbf{k}} | u_{\mathbf{k}} \rangle \langle u_{\mathbf{k}} | \partial_\nu u_{\mathbf{k}} \rangle , \tag{S64}$$

where all partial derivatives are taken with respect to $\mathbf{k}$. Recalling that $\langle u_{\mathbf{k}} | \partial_\nu u_{\mathbf{k}} \rangle$ is purely imaginary, we notice that the second factor in the geometric tensor is purely real. We therefore define the Berry curvature and quantum metric as

$$\mathcal{F}_{\mu\nu}(\mathbf{k}) = -i \left(\langle \partial_\mu u_{\mathbf{k}} | \partial_\nu u_{\mathbf{k}} \rangle - \langle \partial_\nu u_{\mathbf{k}} | \partial_\mu u_{\mathbf{k}} \rangle \right) = 2\text{Im}\mathcal{Q}_{\mu\nu}(\mathbf{k}), \tag{S65a}$$

$$g_{\mu\nu}(\mathbf{k}) = \frac{\langle \partial_\mu u_{\mathbf{k}} | \partial_\nu u_{\mathbf{k}} \rangle + \langle \partial_\nu u_{\mathbf{k}} | \partial_\mu u_{\mathbf{k}} \rangle}{2} - \langle \partial_\mu u_{\mathbf{k}} | u_{\mathbf{k}} \rangle \langle u_{\mathbf{k}} | \partial_\nu u_{\mathbf{k}} \rangle = \text{Re}\mathcal{Q}_{\mu\nu}(\mathbf{k}). \tag{S65b}$$

To facilitate numerical calculation, let us consider finite differences of the form

$$\langle u_{\mathbf{k}} | u_{\mathbf{k}+d\mathbf{k}} \rangle \approx 1 + \langle u_{\mathbf{k}} | \partial_\mu u_{\mathbf{k}} \rangle dk^\mu + \frac{1}{2} \langle u_{\mathbf{k}} | \partial_\mu \partial_\nu u_{\mathbf{k}} \rangle dk^\mu dk^\nu + ..., \tag{S66}$$

where we have expanded $|u_{\mathbf{k}+d\mathbf{k}}\rangle$ to second order in $d\mathbf{k}$ and the contribution of the normalization, if any, is neglected since it is beyond second order. Taking the absolute value and only keeping up to second order, we find

$$|\langle u_{\mathbf{k}} | u_{\mathbf{k}+d\mathbf{k}} \rangle| \approx \sqrt{\left(1 + \frac{1}{2}\text{Re} \langle u_{\mathbf{k}} | \partial_\mu \partial_\nu u_{\mathbf{k}} \rangle dk^\mu dk^\nu \right)^2 + \left(-i \langle u_{\mathbf{k}} | \partial_\mu u_{\mathbf{k}} \rangle dk^\mu + \frac{1}{2}\text{Im} \langle u_{\mathbf{k}} | \partial_\mu \partial_\nu u_{\mathbf{k}} \rangle dk^\mu dk^\nu \right)^2}$$

$$\approx 1 + \frac{1}{2} \left[\text{Re} \langle u_{\mathbf{k}} | \partial_\mu \partial_\nu u_{\mathbf{k}} \rangle + \langle \partial_\mu u_{\mathbf{k}} | u_{\mathbf{k}} \rangle \langle u_{\mathbf{k}} | \partial_\nu u_{\mathbf{k}} \rangle \right] dk^\mu dk^\nu + ... \tag{S67}$$

We can simplify noting that $\text{Re} \langle u_{\mathbf{k}} | \partial_\mu \partial_\nu u_{\mathbf{k}} \rangle = -\text{Re} \langle \partial_\mu u_{\mathbf{k}} | \partial_\nu u_{\mathbf{k}} \rangle$ because $\langle u_{\mathbf{k}} | \partial_\mu \partial_\nu u_{\mathbf{k}} \rangle + \langle \partial_\mu u_{\mathbf{k}} | \partial_\nu u_{\mathbf{k}} \rangle$ is purely imaginary [16]:

$$|\langle u_{\mathbf{k}} | u_{\mathbf{k}+d\mathbf{k}} \rangle| \approx 1 - \frac{1}{2} g_{\mu\nu} dk^\mu dk^\nu + ... \rightarrow g_{\mu\nu} dk^\mu dk^\nu \approx 2 - 2 |\langle u_{\mathbf{k}} | u_{\mathbf{k}+d\mathbf{k}} \rangle| \approx -2\text{Re}\ln \left[\langle u_{\mathbf{k}} | u_{\mathbf{k}+d\mathbf{k}} \rangle \right] \tag{S68}$$

For purposes of numerical evaluation, let us identify Eq. (S68) with the quantum metric half-way between $\mathbf{k}$ and $\mathbf{k} + d\mathbf{k}$: $g_{\mu\nu}(\mathbf{k} + d\mathbf{k}/2)$. On a grid, we have wave functions at discrete points $\mathbf{k} = n_1 \mathbf{g}_1 + n_2 \mathbf{g}_2$, where $\mathbf{g}_i = \mathbf{G}_i / N_{\mathbf{k}}$ and $N_{\mathbf{k}}^2$ is the number of points sampled in the moiré Brillouin zone. We compute the quantum metric and Berry curvature using two different *counterclockwise* triangular loops: (1) $\mathbf{k} \rightarrow \mathbf{k} + \mathbf{g}_1 \rightarrow \mathbf{k} + \mathbf{g}_1 + \mathbf{g}_2 \rightarrow \mathbf{k}$ and (2) $\mathbf{k} \rightarrow \mathbf{k} + \mathbf{g}_1 + \mathbf{g}_2 \rightarrow \mathbf{k} + \mathbf{g}_2 \rightarrow \mathbf{k}$. For the first

triangular loop, the quantum metric is approximately

$$
\begin{aligned}
&- 2\mathrm{Re}\ln\left[\langle u_{\mathbf{k}}|u_{\mathbf{k}+\mathbf{g}_1}\rangle \langle u_{\mathbf{k}+\mathbf{g}_1}|u_{\mathbf{k}+\mathbf{g}_1+\mathbf{g}_2}\rangle \langle u_{\mathbf{k}+\mathbf{g}_1+\mathbf{g}_2}|u_{\mathbf{k}}\rangle\right] \\
&\approx g_{xx}(\mathbf{k}+\mathbf{g}_1/2)\frac{\kappa^2}{4} + 2g_{xy}(\mathbf{k}+\mathbf{g}_1/2)\frac{\sqrt{3}\kappa^2}{4} + g_{yy}(\mathbf{k}+\mathbf{g}_1/2)\frac{3\kappa^2}{4} + g_{xx}(\mathbf{k}+\mathbf{g}_1+\mathbf{g}_2/2)\kappa^2 \\
&+ g_{xx}(\mathbf{k}+\mathbf{g}_1/2+\mathbf{g}_2/2)\frac{\kappa^2}{4} - 2g_{xy}(\mathbf{k}+\mathbf{g}_1/2+\mathbf{g}_2/2)\frac{\sqrt{3}\kappa^2}{4} + g_{yy}(\mathbf{k}+\mathbf{g}_1/2+\mathbf{g}_2/2)\frac{3\kappa^2}{4} \\
&\approx \left[g_{xx}(\mathbf{k}+(2\mathbf{g}_1+\mathbf{g}_2)/3) + g_{yy}(\mathbf{k}+(2\mathbf{g}_1+\mathbf{g}_2)/3)\right]\frac{3\kappa^2}{2}
\end{aligned}
\tag{S69}
$$

For the second triangular loop, the quantum metric is approximately

$$
\begin{aligned}
&- 2\mathrm{Re}\ln\left[\langle u_{\mathbf{k}}|u_{\mathbf{k}+\mathbf{g}_1+\mathbf{g}_2}\rangle \langle u_{\mathbf{k}+\mathbf{g}_1+\mathbf{g}_2}|u_{\mathbf{k}+\mathbf{g}_2}\rangle \langle u_{\mathbf{k}+\mathbf{g}_2}|u_{\mathbf{k}}\rangle\right] \\
&\approx g_{xx}(\mathbf{k}+\mathbf{g}_1/2+\mathbf{g}_2/2)\frac{\kappa^2}{4} - 2g_{xy}(\mathbf{k}+\mathbf{g}_1/2+\mathbf{g}_2/2)\frac{\sqrt{3}\kappa^2}{4} + g_{yy}(\mathbf{k}+\mathbf{g}_1/2+\mathbf{g}_2/2)\frac{3\kappa^2}{4} \\
&+ g_{xx}(\mathbf{k}+\mathbf{g}_1/2+\mathbf{g}_2)\frac{\kappa^2}{4} + 2g_{xy}(\mathbf{k}+\mathbf{g}_1/2+\mathbf{g}_2)\frac{\sqrt{3}\kappa^2}{4} + g_{yy}(\mathbf{k}+\mathbf{g}_1/2+\mathbf{g}_2)\frac{3\kappa^2}{4} + g_{xx}(\mathbf{k}+\mathbf{g}_2/2)\kappa^2 \\
&\approx \left[g_{xx}(\mathbf{k}+(\mathbf{g}_1+2\mathbf{g}_2)/3) + g_{yy}(\mathbf{k}+(\mathbf{g}_1+2\mathbf{g}_2)/3)\right]\frac{3\kappa^2}{2}
\end{aligned}
\tag{S70}
$$

Here, $\kappa = |\mathbf{g}_i|$. In the above derivations, we have assumed that the off-diagonal elements of $g_{\mu\nu}$ approximately cancel since the loops close. Eqs. (S69) and (S70) are essentially the same expressions in Ref. [16]. There is a slightly different way (but similar in spirit) to numerically calculate the trace of the quantum metric presented in Ref. [17]. For the Berry curvature, we can rewrite $\mathcal{B}_z(\mathbf{k}) = \mathcal{F}_{xy}(\mathbf{k}) = \hat{z}\cdot[\nabla \times (-i\langle u_{\mathbf{k}}|\nabla_{\mathbf{k}}|u_{\mathbf{k}}\rangle)]$. So, the integral of the Berry curvature over a triangular area can be turned into a loop integral around the boundary of that area via Stokes' theorem (for the single-band case we are currently assuming). Thus, we have

$$
\begin{aligned}
\mathcal{B}_z(\mathbf{k}+(2\mathbf{g}_1+\mathbf{g}_2)/3)\frac{\sqrt{3}\kappa^2}{4} &= \mathrm{Im}\ln\left[\langle u_{\mathbf{k}}|u_{\mathbf{k}+\mathbf{g}_1}\rangle \langle u_{\mathbf{k}+\mathbf{g}_1}|u_{\mathbf{k}+\mathbf{g}_1+\mathbf{g}_2}\rangle \langle u_{\mathbf{k}+\mathbf{g}_1+\mathbf{g}_2}|u_{\mathbf{k}}\rangle\right], \\
\mathcal{B}_z(\mathbf{k}+(\mathbf{g}_1+2\mathbf{g}_2)/3)\frac{\sqrt{3}\kappa^2}{4} &= \mathrm{Im}\ln\left[\langle u_{\mathbf{k}}|u_{\mathbf{k}+\mathbf{g}_1+\mathbf{g}_2}\rangle \langle u_{\mathbf{k}+\mathbf{g}_1+\mathbf{g}_2}|u_{\mathbf{k}+\mathbf{g}_2}\rangle \langle u_{\mathbf{k}+\mathbf{g}_2}|u_{\mathbf{k}}\rangle\right].
\end{aligned}
\tag{S71}
$$

We can compactify the notation by defining $\triangle(\mathbf{k}_*)$ as the gauge-invariant argument of the logarithm around a counterclockwise loop centered at $\mathbf{k}_*$. Then, we have

$$
\mathrm{Tr}\, g_{\mu\nu}(\mathbf{k}_*) = -\frac{4}{3\kappa^2}\mathrm{Re}\ln \triangle(\mathbf{k}_*),
\tag{S72a}
$$

$$
\mathcal{B}_z(\mathbf{k}_*) = \frac{4}{\sqrt{3}\kappa^2}\mathrm{Im}\ln \triangle(\mathbf{k}_*).
\tag{S72b}
$$

This shows that both the quantum metric and Berry curvature by just taking real and imaginary parts of $\triangle(\mathbf{k}_*)$ (and then multiplying by non-universal factors that depend on the grid geometry). It is worth emphasizing in closing that the expressions derived Eq. (S72) are manifestly gauge-invariant. Also, it is clear from Eq. (S72) that the trace of the quantum metric is positive semidefinite. This is because $\mathrm{Re}\ln \triangle(\mathbf{k}_*) = \ln|\triangle(\mathbf{k}_*)| \leq 0$ since $0 \leq |\triangle(\mathbf{k}_*)| \leq 1$. For the Berry curvature, we choose the branch of the logarithm such that $-\pi \leq \mathrm{Im}\ln \triangle(\mathbf{k}_*) < \pi$.

The primary interest of this study is to explore the topology of the first conduction band in cases where it is well-isolated spectrally from the rest of the energy spectrum. For this purpose, the above prescription to calculate the Berry curvature and corresponding Chern number of the first conduction band is sufficient. However, there is a subtlety that may be relevant experimentally. The Chern number of the first conduction band is not necessarily equal to the Hall conductivity (in units of e^2/h) of the gap right above it since the latter is equal to the sum of all the Chern numbers (in units of e^2/h) of the occupied bands. In an experiment, what is measured is the Hall conductance (or equivalently, Hall resistance) associated with a gap, and not directly the Chern number associated with a band. So, it is meaningful to calculate such a Hall conductance. When computing the Chern numbers of all the other bands to obain the Hall conductivity of a gap, a numerical instability is occasionally observed when the remote bands are close together in energy. To mitigate this effect, we instead calculate the non-Abelian Berry curvature whose trace over bands once integrated over the mBZ gives the *composite* Chern number. This calculation has the advantage that it is well defined even if there are degeneracies in the occupied manifold as long as the occupied and unoccupied states are separated from each other by a sufficiently large energy gap, which we always assume for robust insulating states. In this multiband formalism, the non-Abelian Berry connection is [18]

$$
\mathcal{A}^{n,m}(\mathbf{k}) = -i\langle u_{n,\mathbf{k}}|\nabla_{\mathbf{k}}|u_{m,\mathbf{k}}\rangle,
\tag{S73}
$$

where m, n are band indices of the occupied bands. This is now a matrix with the number of rows equaling the number of occupied bands. The non-Abelian Berry curvature matrix is defined using a covariant derivative $\mathfrak{D}_\mu = \partial_\mu - \mathcal{A}_\mu$:

$$\mathcal{B}_z(\mathbf{k}) = -\left[\mathfrak{D}_x, \mathfrak{D}_y\right] = \partial_x \mathcal{A}_y - \partial_y \mathcal{A}_x - \left[\mathcal{A}_x, \mathcal{A}_y\right]. \tag{S74}$$

The non-Abelian Berry curvature differs from the Abelian Berry curvature in the nontrivial commutator $\left[\mathcal{A}_x, \mathcal{A}_y\right]$, which clearly vanishes in the Abelian case. The composite Chern number is

$$\mathcal{C}_{\text{tot}} = \frac{1}{2\pi} \int d^2\mathbf{k} \text{Tr}\left[\mathcal{B}_z(\mathbf{k})\right], \tag{S75}$$

where the trace is over all occupied bands. To compute this quantity numerically, we use the method described in Ref. [19]. We can use the same formulas as before with the replacement:

$$|u_\mathbf{k}\rangle \mapsto |U_\mathbf{k}\rangle = \left(|u_{1,\mathbf{k}}\rangle \ |u_{2,\mathbf{k}}\rangle \ |u_{3,\mathbf{k}}\rangle \ ... \ |u_{n,\mathbf{k}}\rangle\right),$$
$$\langle u_\mathbf{k}| \mapsto \langle U_\mathbf{k}| = \left(\langle u_{1,\mathbf{k}}| \ \langle u_{2,\mathbf{k}}| \ \langle u_{3,\mathbf{k}}| \ ... \ \langle u_{n,\mathbf{k}}|\right)^T. \tag{S76}$$

Then, we have

$$\mathcal{B}_z(\mathbf{k} + (2\mathbf{g}_1 + \mathbf{g}_2)/3)\frac{\sqrt{3}\kappa^2}{4} = \text{Im}\ln\left[\langle U_\mathbf{k}|U_{\mathbf{k}+\mathbf{g}_1}\rangle \langle U_{\mathbf{k}+\mathbf{g}_1}|U_{\mathbf{k}+\mathbf{g}_1+\mathbf{g}_2}\rangle \langle U_{\mathbf{k}+\mathbf{g}_1+\mathbf{g}_2}|U_\mathbf{k}\rangle\right], \tag{S77a}$$

$$\mathcal{B}_z(\mathbf{k} + (\mathbf{g}_1 + 2\mathbf{g}_2)/3)\frac{\sqrt{3}\kappa^2}{4} = \text{Im}\ln\left[\langle U_\mathbf{k}|U_{\mathbf{k}+\mathbf{g}_1+\mathbf{g}_2}\rangle \langle U_{\mathbf{k}+\mathbf{g}_1+\mathbf{g}_2}|U_{\mathbf{k}+\mathbf{g}_2}\rangle \langle U_{\mathbf{k}+\mathbf{g}_2}|U_\mathbf{k}\rangle\right]. \tag{S77b}$$

Under suitable conditions, tracing over the Berry curvature can be replaced by taking determinants inside the logarithm. We end by emphasizing again that this method is only reliable when the occupied bands are reasonably separated spectrally from the unoccupied bands. Otherwise, this multi-band method suffers the same problems as the single-band method does.

IX.   DESCRIPTIONS OF ACCOMPANYING .MP4 MOVIES

This manuscript is accompanied by several .mp4 files that animate the evolution of band structures using different sets of parameters. In this section, we provide descriptions for all of these .mp4 files.

1. `noninteracting_band_structure_antiparallel_with_relaxation.mp4`: This movie shows the evolution of the non-interacting band structure in the *antiparallel configuration* for $\nu = +$ as $\Delta \in [-60, 60]$ meV and $\theta \in [0.8°, 1.5°]$ are varied. Each energy state is color-coded by its layer polarization as defined in Eq. (S11). The Chern number of the first conduction band is indicated in the title. We have used parameters in `Param_one` which *includes some lattice relaxation* due to the difference in γ_{AA} and γ_{AB}.

2. `noninteracting_band_structure_antiparallel_no_relaxation.mp4`: This movie shows the evolution of the non-interacting band structure in the *antiparallel configuration* for $\nu = +$ as $\Delta \in [-60, 60]$ meV and $\theta \in [0.8°, 1.5°]$ are varied. Each energy state is color-coded by its layer polarization as defined in Eq. (S11). The Chern number of the first conduction band is indicated in the title. We have used parameters in `Param_two` which *does not include any lattice relaxation* since $\gamma_{\text{AA}} = \gamma_{\text{AB}}$.

3. `noninteracting_band_structure_parallel_with_relaxation.mp4`: This movie shows the evolution of the non-interacting band structure in the *parallel configuration* for $\nu = +$ as $\Delta \in [-60, 60]$ meV and $\theta \in [0.8°, 1.5°]$ are varied. Each energy state is color-coded by its layer polarization as defined in Eq. (S11). The Chern number of the first conduction band is indicated in the title. We have used parameters in `Param_one` which *includes some lattice relaxation* due to the difference in γ_{AA} and γ_{AB}.

4. `noninteracting_band_structure_parallel_no_relaxation.mp4`: This movie shows the evolution of the non-interacting band structure in the *parallel configuration* for $\nu = +$ as $\Delta \in [-60, 60]$ meV and $\theta \in [0.8°, 1.5°]$ are varied. Each energy state is color-coded by its layer polarization as defined in Eq. (S11). The Chern number of the first conduction band is indicated in the title. We have used parameters in `Param_two` which *does not include any lattice relaxation* since $\gamma_{\text{AA}} = \gamma_{\text{AB}}$.

5. `noninteracting_band_structure_antiparallel_reduced_vF_with_relaxation.mp4`: This movie shows the evolution of the non-interacting band structure in the *antiparallel configuration* for $\nu = +$ as $\Delta \in [-60, 60]$ meV and $\theta \in [0.8°, 1.5°]$ are varied. Each energy state is color-coded by its layer polarization as defined in Eq. (S11). The Chern number of the first conduction band is indicated in the title. We have used parameters in `Param_three` which *includes some lattice relaxation* due to the difference in γ_{AA} and γ_{AB}.

6. `noninteracting_band_structure_antiparallel_reduced_vF_no_relaxation.mp4`: This movie shows the evolution of the non-interacting band structure in the *antiparallel configuration* for $\nu = +$ as $\Delta \in [-60, 60]$ meV and $\theta \in [0.8°, 1.5°]$ are varied. Each energy state is color-coded by its layer polarization as defined in Eq. (S11). The Chern number of the first conduction band is indicated in the title. We have used parameters in `Param_four` which *does not include any lattice relaxation* since $\gamma_{AA} = \gamma_{AB}$.

7. `noninteracting_band_structure_parallel_reduced_vF_with_relaxation.mp4`: This movie shows the evolution of the non-interacting band structure in the *parallel configuration* for $\nu = +$ as $\Delta \in [-60, 60]$ meV and $\theta \in [0.8°, 1.5°]$ are varied. Each energy state is color-coded by its layer polarization as defined in Eq. (S11). The Chern number of the first conduction band is indicated in the title. We have used parameters in `Param_three` which *includes some lattice relaxation* due to the difference in γ_{AA} and γ_{AB}.

8. `noninteracting_band_structure_parallel_reduced_vF_no_relaxation.mp4`: This movie shows the evolution of the non-interacting band structure in the *parallel configuration* for $\nu = +$ as $\Delta \in [-60, 60]$ meV and $\theta \in [0.8°, 1.5°]$ are varied. Each energy state is color-coded by its layer polarization as defined in Eq. (S11). The Chern number of the first conduction band is indicated in the title. We have used parameters in `Param_four` which *does not include any lattice relaxation* since $\gamma_{AA} = \gamma_{AB}$.

In all of these band structures, the presented Chern numbers are calculated numerically by summing over the local Berry curvature computed with Eq. (S72) using a finite grid of $N_{\mathbf{k}} \times N_{\mathbf{k}} = 18 \times 18$ points. This method is quite reliable except in cases where the *direct* gaps between bands become vanishingly small, such as when a band inversion occurs. To mitigate this, one has to increase the number of mesh points. However, that is quite expensive, especially for HF calculations. Thus, Chern numbers at a phase boundary should be taken with some skepticism. With that said, the Chern numbers for states deep inside a phase is very reliable. None of our conclusion relies on data points at a phase boundary. So this small numerical instability has no material effect on our study.

---

[1] H. Zhou, T. Xie, A. Ghazaryan, T. Holder, J. R. Ehrets, E. M. Spanton, T. Taniguchi, K. Watanabe, E. Berg, M. Serbyn, *et al.*, Half-and quarter-metals in rhombohedral trilayer graphene, Nature **598**, 429 (2021).

[2] S. Lisi, X. Lu, T. Benschop, T. A. de Jong, P. Stepanov, J. R. Duran, F. Margot, I. Cucchi, E. Cappelli, A. Hunter, *et al.*, Observation of flat bands in twisted bilayer graphene, Nat. Phys. **17**, 189 (2021).

[3] Y. Shi, S. Xu, Y. Yang, S. Slizovskiy, S. V. Morozov, S.-K. Son, S. Ozdemir, C. Mullan, J. Barrier, J. Yin, *et al.*, Electronic phase separation in multilayer rhombohedral graphite, Nature **584**, 210 (2020).

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

[5] N. Leconte, S. Javvaji, J. An, A. Samudrala, and J. Jung, Relaxation effects in twisted bilayer graphene: A multiscale approach, Phys. Rev. B **106**, 115410 (2022).

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

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

[8] M. Koshino and E. McCann, Trigonal warping and berry's phase $n\pi$ in abc-stacked multilayer graphene, Phys. Rev. B **80**, 165409 (2009).

[9] C. Fang, M. J. Gilbert, and B. A. Bernevig, Bulk topological invariants in noninteracting point group symmetric insulators, Phys. Rev. B **86**, 115112 (2012).

[10] C. Huang, T. M. R. Wolf, W. Qin, N. Wei, I. V. Blinov, and A. H. MacDonald, Spin and orbital metallic magnetism in rhombohedral trilayer graphene, Phys. Rev. B **107**, L121405 (2023).

[11] K. c. v. Kolář, Y. Zhang, S. Nadj-Perge, F. von Oppen, and C. Lewandowski, Electrostatic fate of n-layer moiré graphene, Phys. Rev. B **108**, 195148 (2023).

[12] D. Zagier, The dilogarithm function, in *Frontiers in Number Theory, Physics, and Geometry II: On Conformal Field Theories, Discrete Groups and Renormalization*, edited by P. Cartier, P. Moussa, B. Julia, and P. Vanhove (Springer, 2007) pp. 3–65.

[13] K. Huang, X. Li, S. Das Sarma, and F. Zhang, Self-consistent theory of fractional quantum anomalous hall states in rhombohedral graphene, Phys. Rev. B **110**, 115146 (2024).

[14] A. K. Pati, Relation between "phases" and "distance" in quantum evolution, Phys. Lett. A **159**, 105 (1991).

[15] R. Cheng, Quantum geometric tensor (fubini-study metric) in simple quantum system: A pedagogical introduction, arXiv preprint arXiv:1012.1337 https://doi.org/10.48550/arXiv.1012.1337 (2010).

[16] B. Mera and T. Ozawa, Engineering geometrically flat chern bands with fubini-study kähler structure, Phys. Rev. B **104**, 115160 (2021).

[17] N. Marzari and D. Vanderbilt, Maximally localized generalized wannier functions for composite energy bands, Phys. Rev. B **56**, 12847 (1997).

[18] F. Wilczek and A. Zee, Appearance of gauge structure in simple dynamical systems, Phys. Rev. Lett. **52**, 2111 (1984).

[19] T. Fukui, Y. Hatsugai, and H. Suzuki, Chern numbers in discretized brillouin zone: Efficient method of computing (spin) hall conductances, J. Phys. Soc. Jpn. **74**, 1674 (2005).

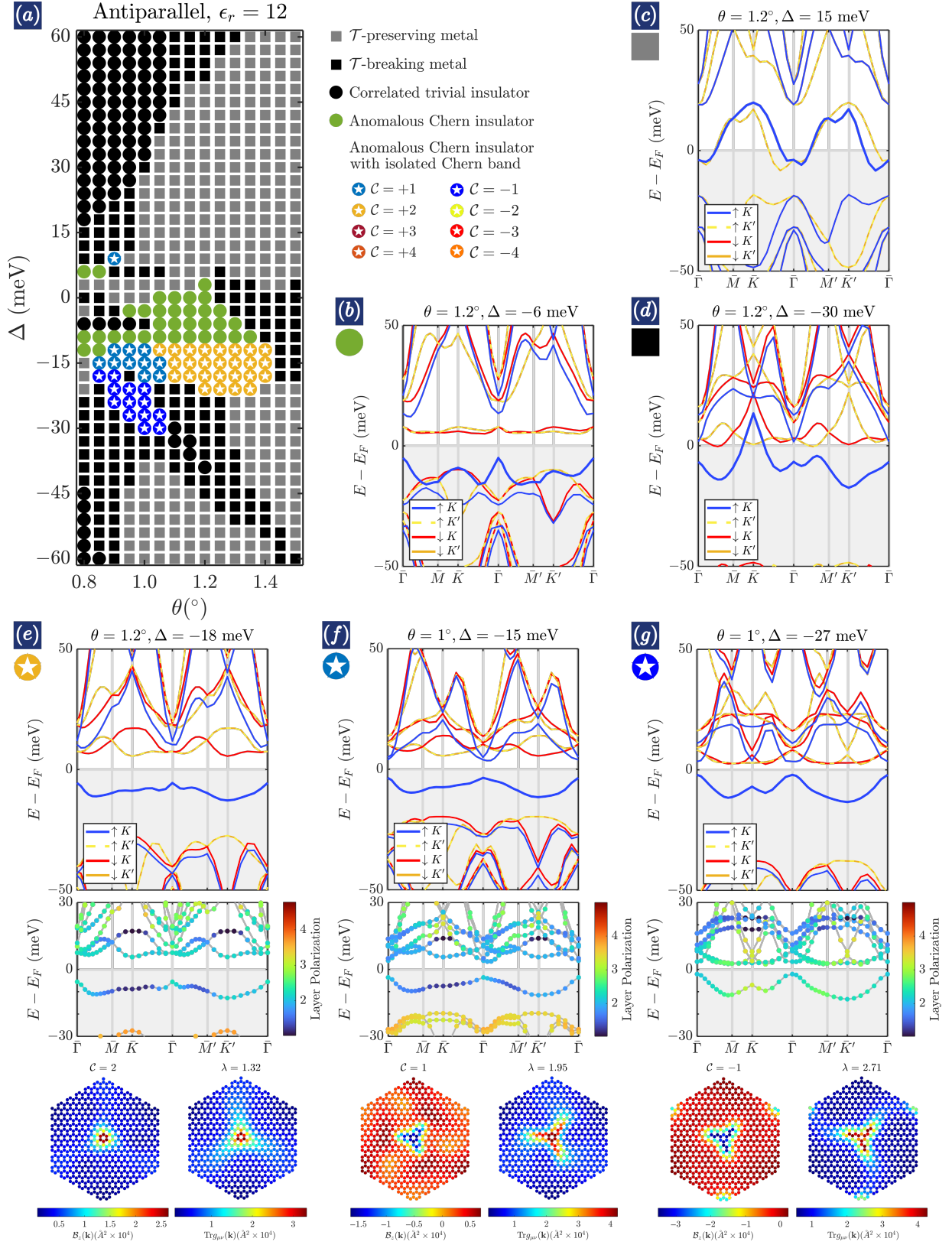Figure S21. **Mean-field phase diagram for the antiparallel configuration with lattice relaxation and** $\epsilon_r = 12$. (a) Phase diagram as a function of θ and Δ. The results here are simulated using parameters from `Param_one`. Representative band structures of the different phases are shown in (b)-(g): an anomalous Chern insulator with no gap below the band closest to chemical potential (b), a $\mathcal{T}$-preserving metal (c), a $\mathcal{T}$-breaking metal (d), and an anomalous Chern insulator with isolated Chern band with $\mathcal{C} = +2$ (e), $\mathcal{C} = +1$ (f), or $\mathcal{C} = -1$ (g). For (b)-(d), we only show the isospin polarization of the energy bands. For (e)-(g), we additionally color code the layer polarization of the Bloch states; furthermore, we show the corresponding Berry curvature and quantum metric of the band closest to chemical potential.
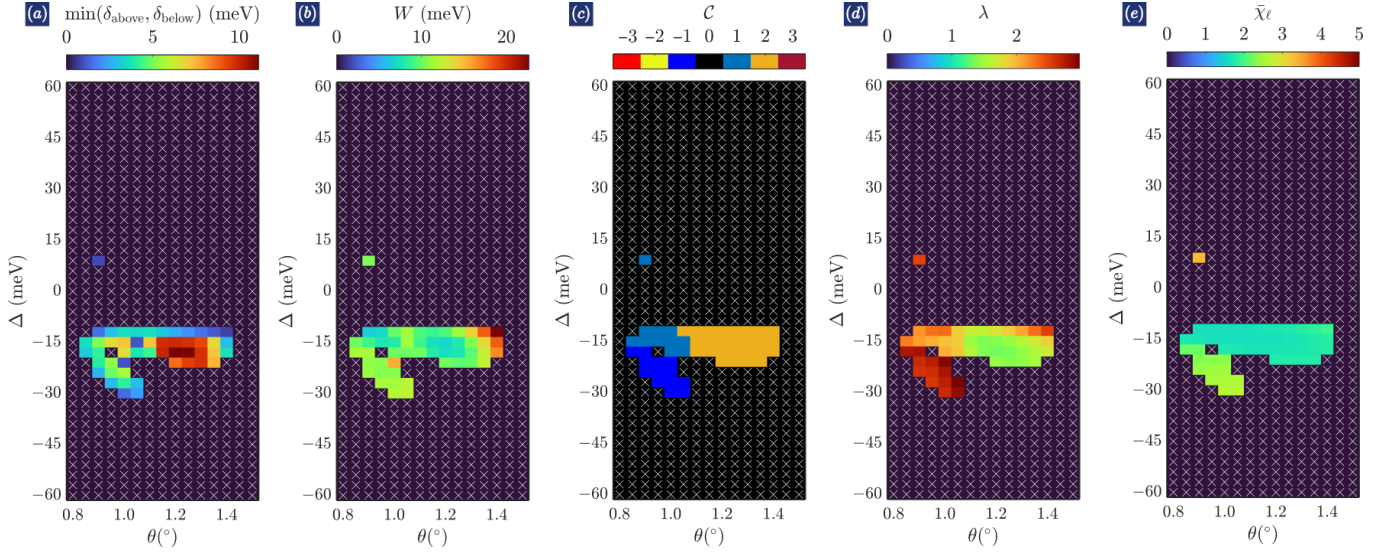
Figure S22. **Further spectral and topological characterization of anomalous Chern phases in Fig. S21.** For anomalous Chern phases with an isolated Chern band, we show following properties of that isolated band: spectral gap (a), bandwidth (b), Chern number (c), trace condition violation (d), and average layer polarization (e).
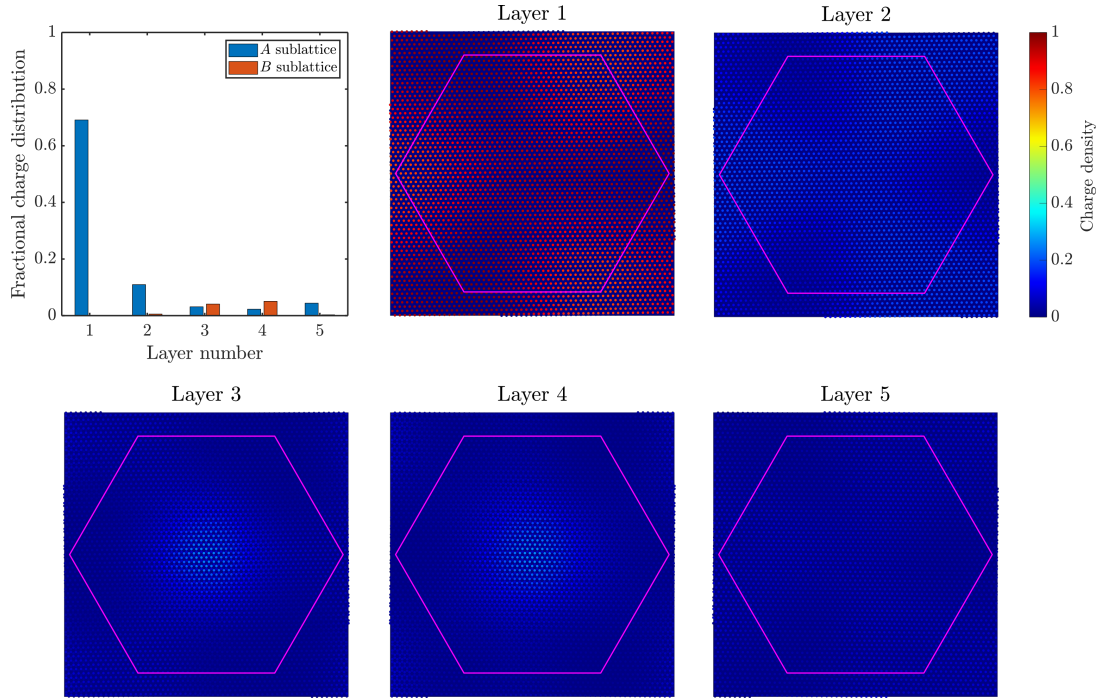

Figure S23. **Charge density at** $\theta = 1.20°$ **and** $\Delta = -18$ **meV in Fig. S21.** In the bar graph, we plot the fractional charge distribution on each layer and sublattice summed over all $\mathbf{k}$ in the narrow band in Fig. S21(e), demonstrating a majority population on layer 1. In the density plots, we show the real-space distribution of the summed charge density for that same band. Most of the charge density is concentrated on layer 1. There is slightly more charge concentrated at the α and β regions compared to at other regions. If one carefully inspects these density plots, one can discern the sublattice polarization on each layer.
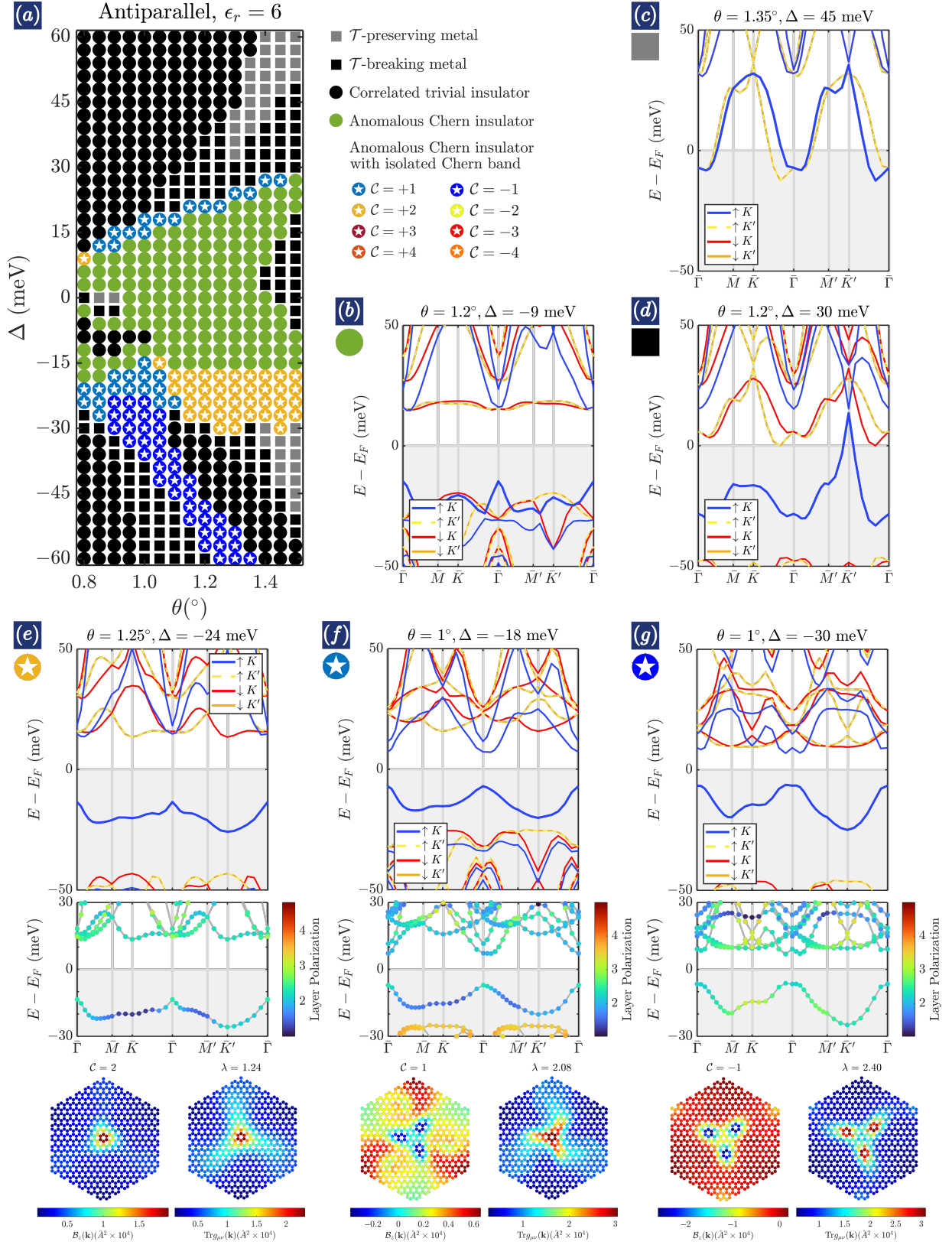
Figure S24. **Mean-field phase diagram for the antiparallel configuration with lattice relaxation and** $\epsilon_r = 6$. (a) Phase diagram as a function of θ and Δ. The results here are simulated using parameters from `Param_one`. Representative band structures of the different phases are shown in (b)-(g): an anomalous Chern insulator with no gap below the band closest to chemical potential (b), a $\mathcal{T}$-preserving metal (c), a $\mathcal{T}$-breaking metal (d), and an anomalous Chern insulator with isolated Chern band with $\mathcal{C} = +2$ (e), $\mathcal{C} = +1$ (f), or $\mathcal{C} = -1$ (g). For (b)-(d), we only show the isospin polarization of the energy bands. For (e)-(g), we additionally color code the layer polarization of the Bloch states; furthermore, we show the corresponding Berry curvature and quantum metric of the band closest to chemical potential.
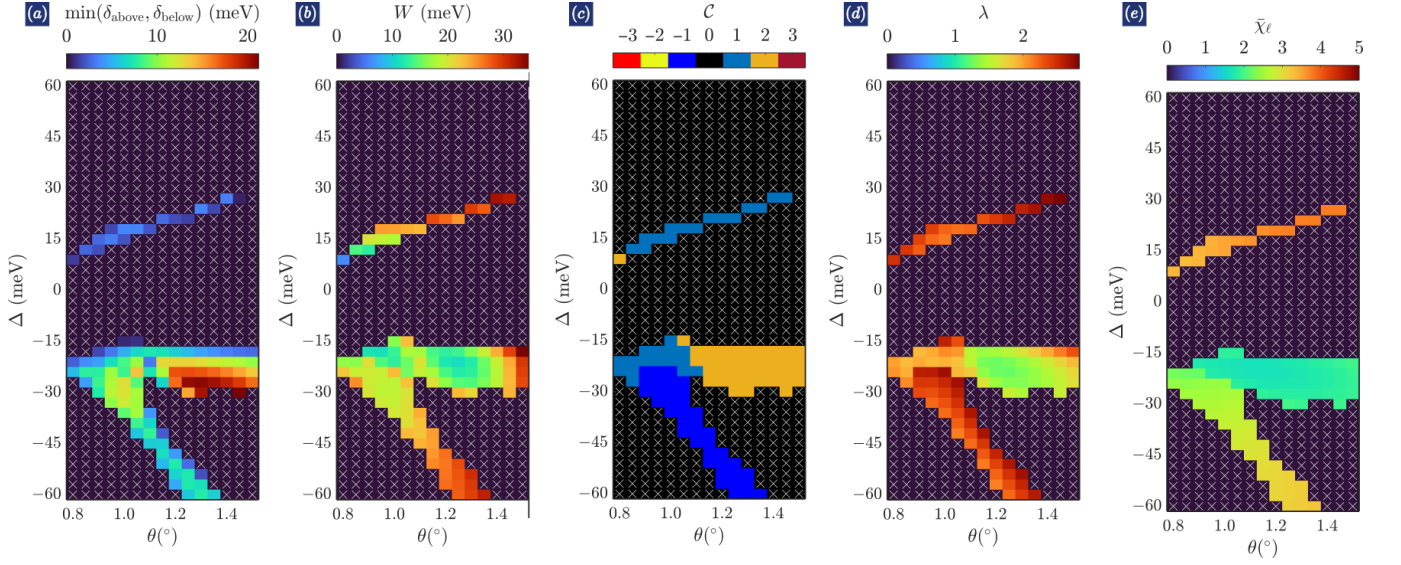
Figure S25. **Further spectral and topological characterization of anomalous Chern phases in Fig. S24.** For anomalous Chern phases with an isolated Chern band, we show following properties of that isolated band: spectral gap (a), bandwidth (b), Chern number (c), trace condition violation (d), and average layer polarization (e).
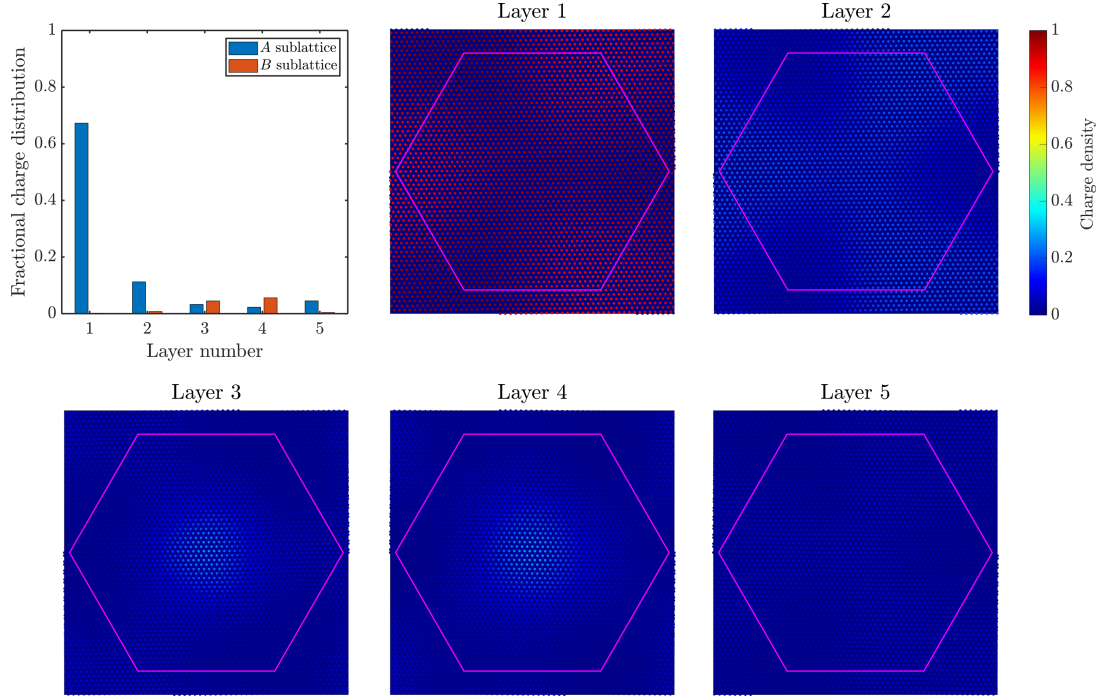

Figure S26. **Charge density at $\theta = 1.25°$ and $\Delta = -24$ meV in Fig. S24.** In the bar graph, we plot the fractional charge distribution on each layer and sublattice summed over all $\mathbf{k}$ in the narrow band in Fig. S24(e), demonstrating a majority population on layer 1. In the density plots, we show the real-space distribution of the summed charge density for that same band. Most of the charge density is on layer 1. α and β regions host slightly more charge density than other regions. We note that this figure is nearly identical to Fig. S23, demonstrating that at least for the $\mathcal{C} = 2$ phase of interest, results simulated with $\epsilon_r = 6$ and $\epsilon_r = 12$ share the same charge density profile.
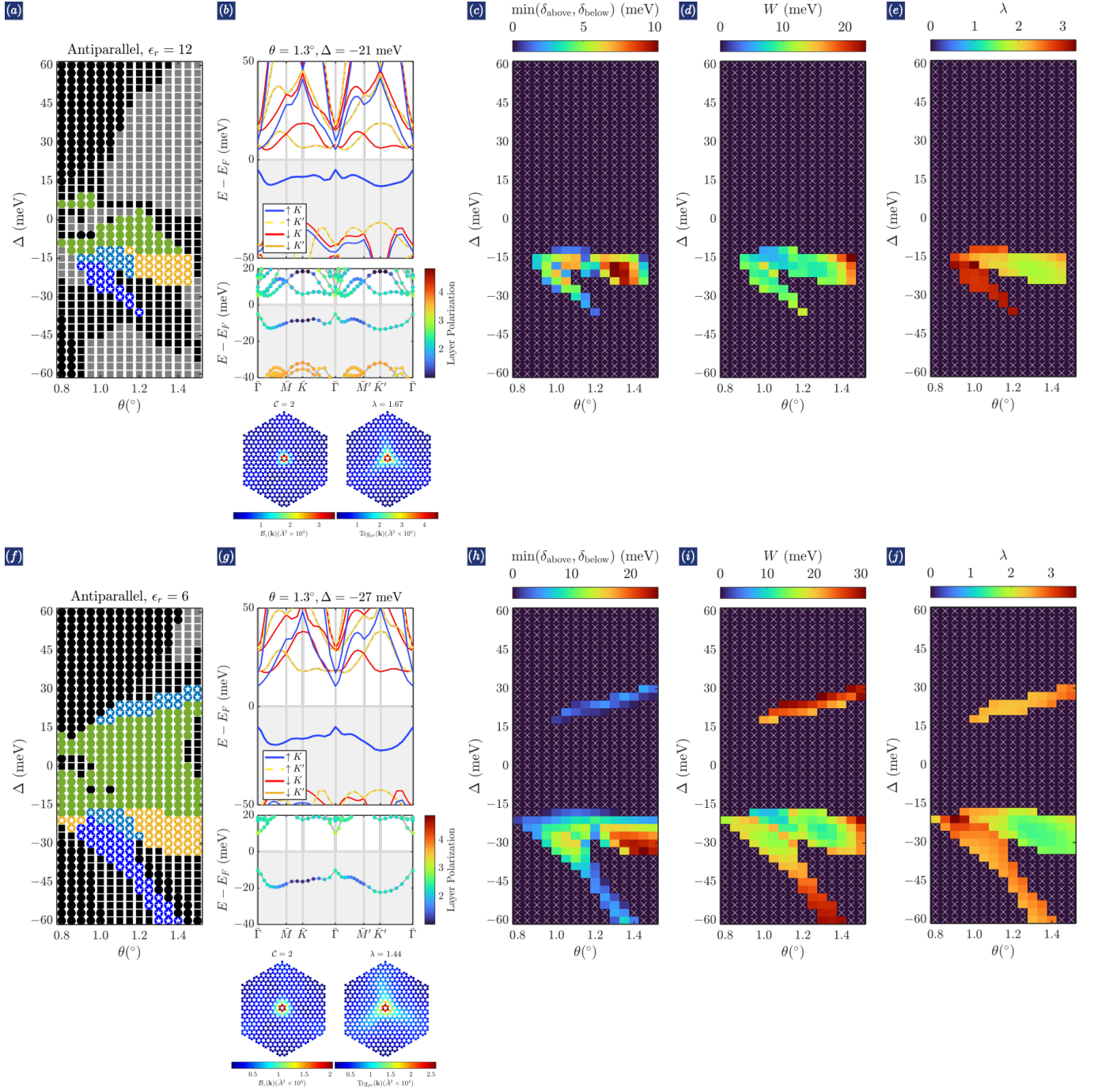
Figure S27. **Mean-field phase diagram for the antiparallel configuration without lattice relaxation.** We show the phase diagram and its further characterization for $\epsilon_r = 12$ in the top panel and for $\epsilon_r = 6$ in the bottom panel. The results here are simulated using parameters from Param_two. In each panel, we show the phase diagram in (θ, Δ) space, the band structures and associated Berry curvature and quantum metric distribution for an optimal f.o.m. phase point, and further characterization of the $\mathcal{C} = 2$ phase including the band gap, bandwidth, and trace condition violation.
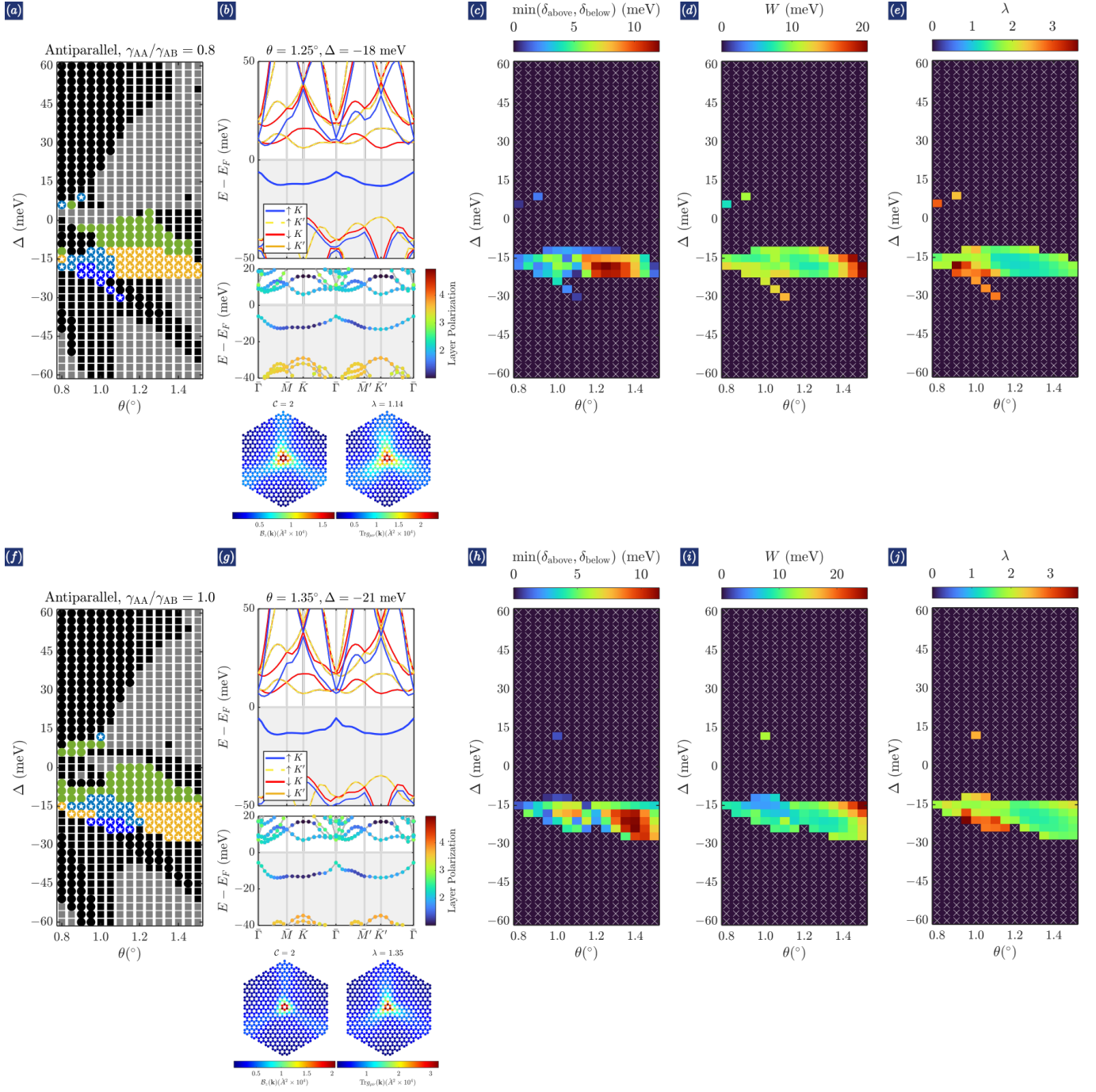
Figure S28. **Mean-field phase diagram for the antiparallel configuration with reduced** γ_0**.** We show the phase diagram and its further characterization with relaxation (`Param_three`) in the top panel and without relaxation (`Param_four`) in the bottom panel. In each panel, we show the phase diagram in (θ, Δ) space, the band structures and associated Berry curvature and quantum metric distribution for an optimal f.o.m. phase point, and further characterization of the $\mathcal{C} = 2$ phase including the band gap, bandwidth, and trace condition violation. Here, $\epsilon_r = 12$.
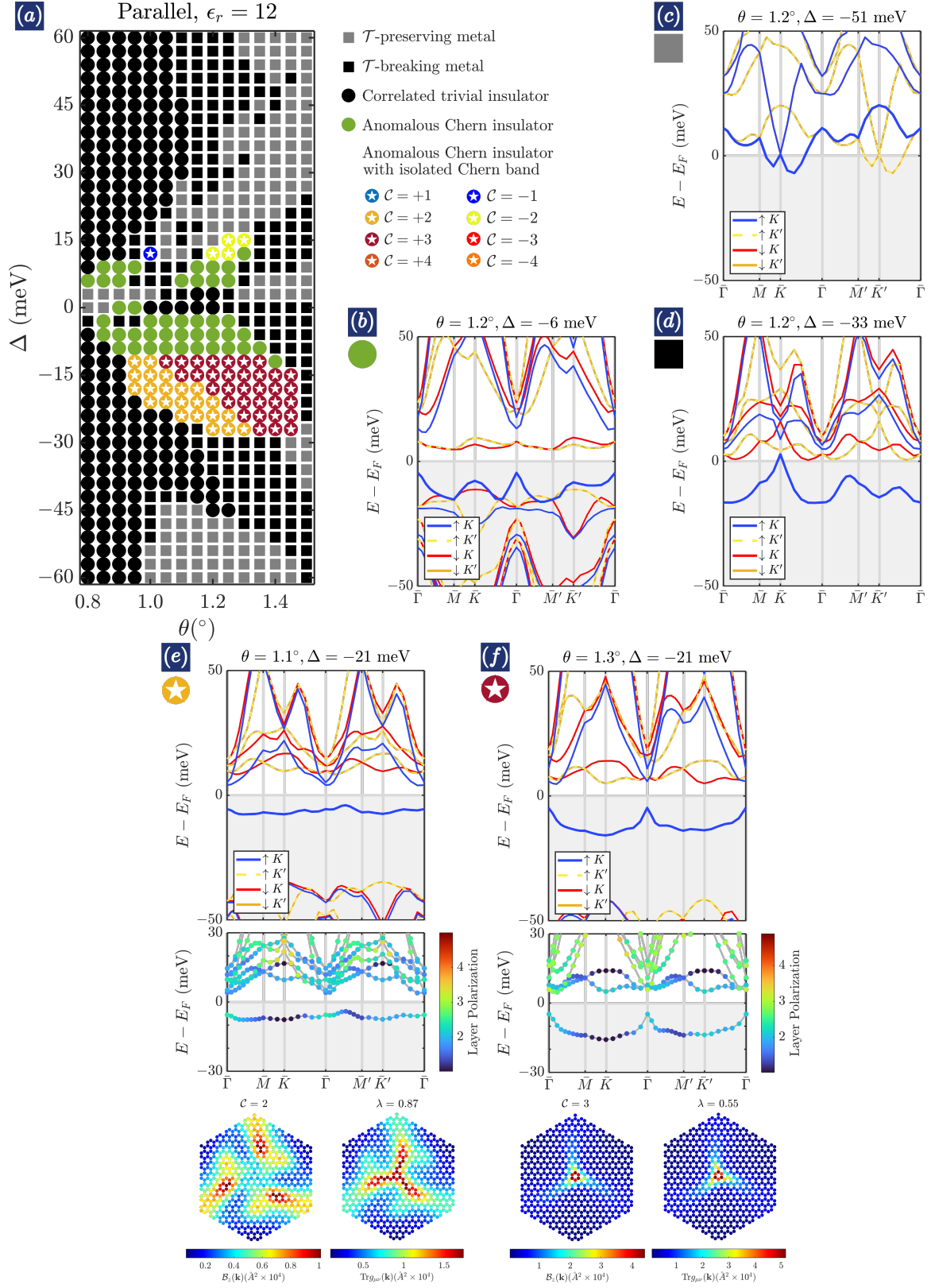
Figure S29. **Mean-field phase diagram for the parallel configuration with lattice relaxation and** $\epsilon_r = 12$. (a) Phase diagram as a function of θ and Δ. The results here are simulated using parameters from `Param_one`. Representative band structures of the different phases are shown in (b)-(f): an anomalous Chern insulator with no gap below the band closest to chemical potential (b), a $\mathcal{T}$-preserving metal (c), a $\mathcal{T}$-breaking metal (d), and an anomalous Chern insulator with isolated Chern band with $\mathcal{C} = +2$ (e) or $\mathcal{C} = +3$ (f). For (b)-(d), we only show the isospin polarization of the energy bands. For (e)-(f), we additionally color code the layer polarization of the Bloch states; furthermore, we show the corresponding Berry curvature and quantum metric of the band closest to chemical potential.
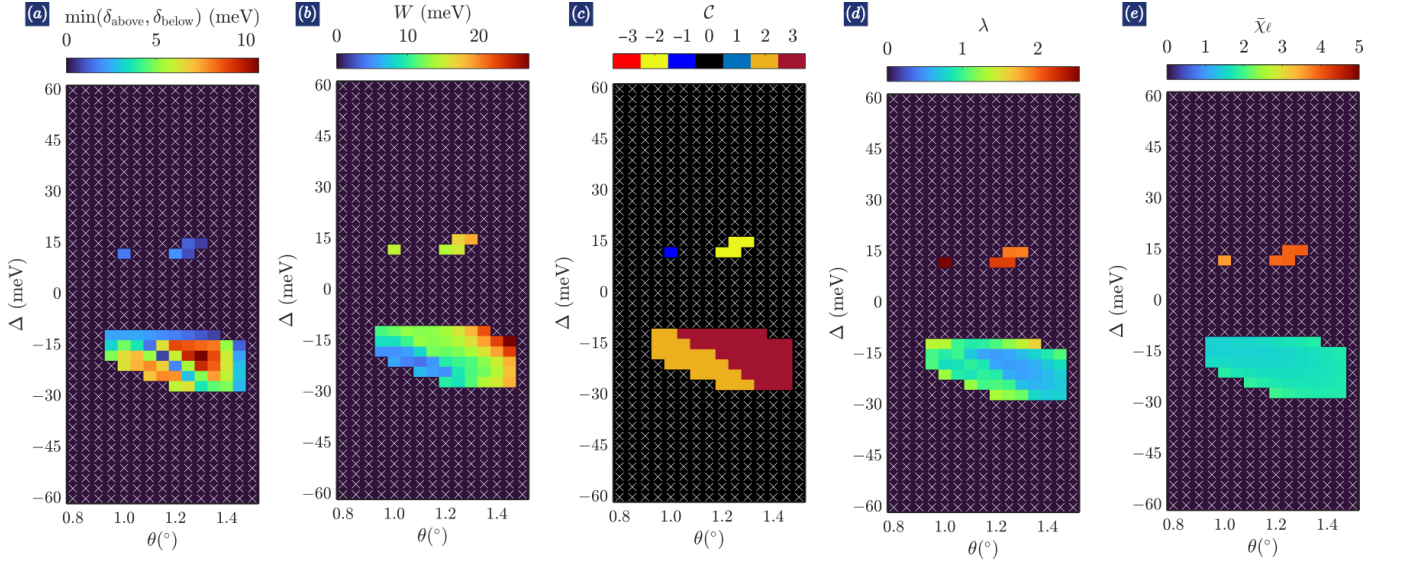
Figure S30. **Further spectral and topological characterization of anomalous Chern phases in Fig. S29.** For anomalous Chern phases with an isolated Chern band, we show following properties of that isolated band: spectral gap (a), bandwidth (b), Chern number (c), trace condition violation (d), and average layer polarization (e).
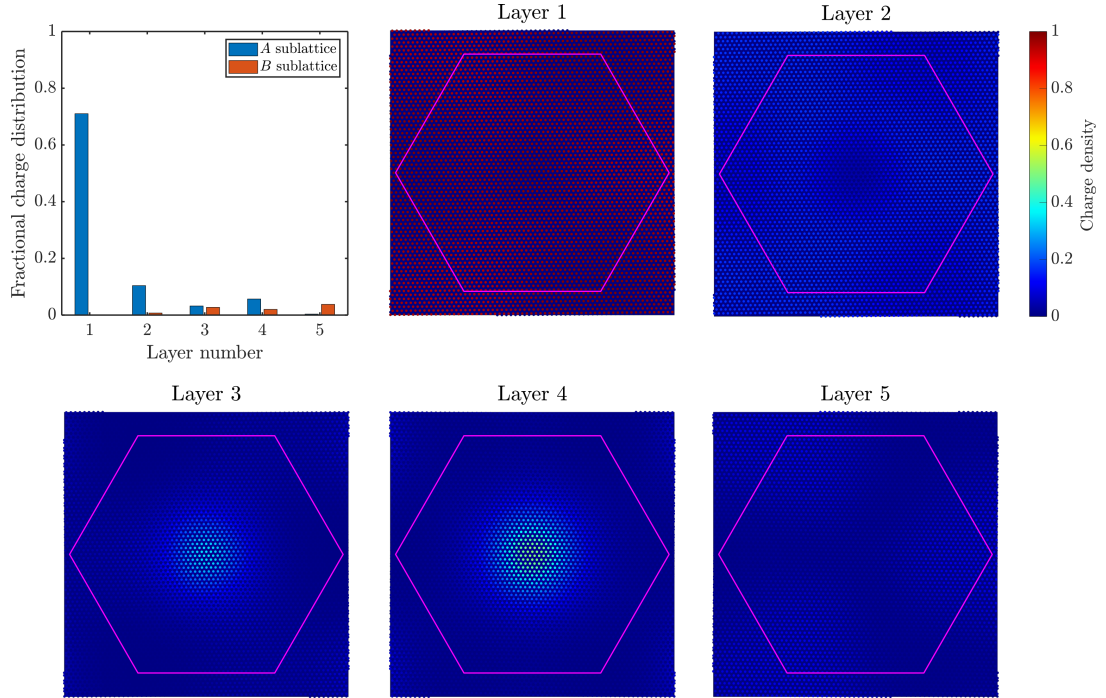

Figure S31. **Charge density at $\theta = 1.10°$ and $\Delta = -21$ meV in Fig. S29.** In the bar graph, we plot the fractional charge distribution on each layer and sublattice summed over all **k** in the narrow band in Fig. S29(e), demonstrating a majority population on layer 1. In the density plots, we show the real-space distribution of the summed charge density for that same band. Most of the charge density is on layer 1, concentrated more at α and γ regions than other regions.
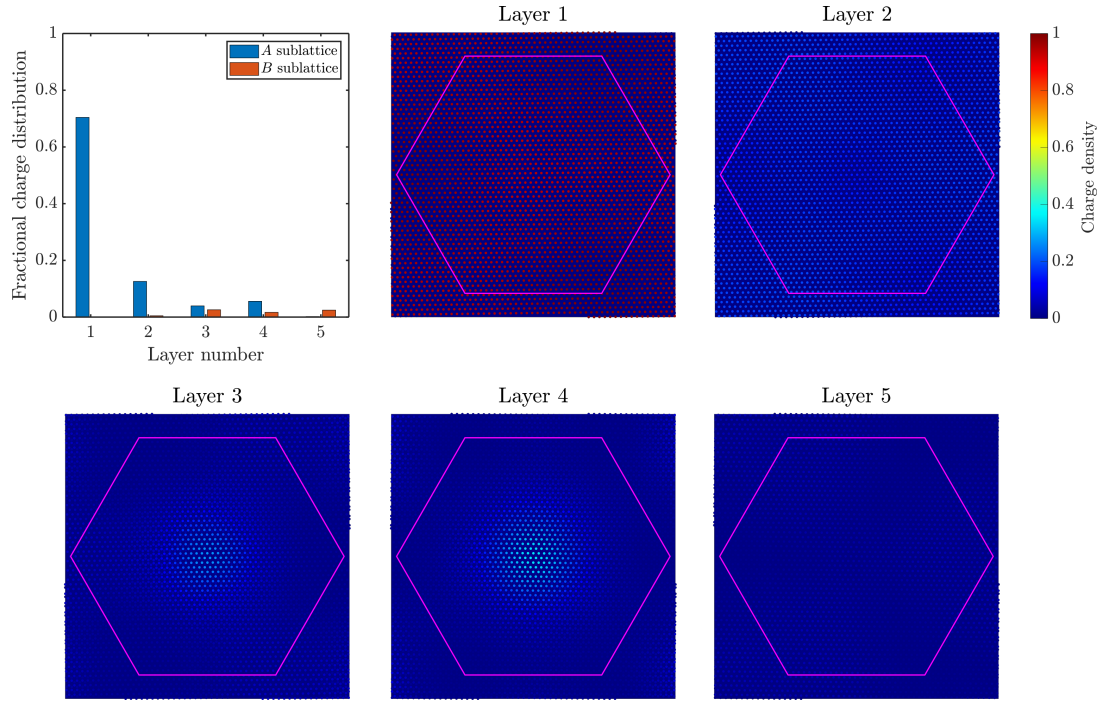
Figure S32. **Charge density at** $\theta = 1.30°$ **and** $\Delta = -21$ **meV in Fig. S29.** In the bar graph, we plot the fractional charge distribution on each layer and sublattice summed over all **k** in the narrow band in Fig. S29(f), demonstrating a majority population on layer 1. In the density plots, we show the real-space distribution of the summed charge density for that same band. Most of the charge density is on layer 1, concentrated more at γ regions than other regions.
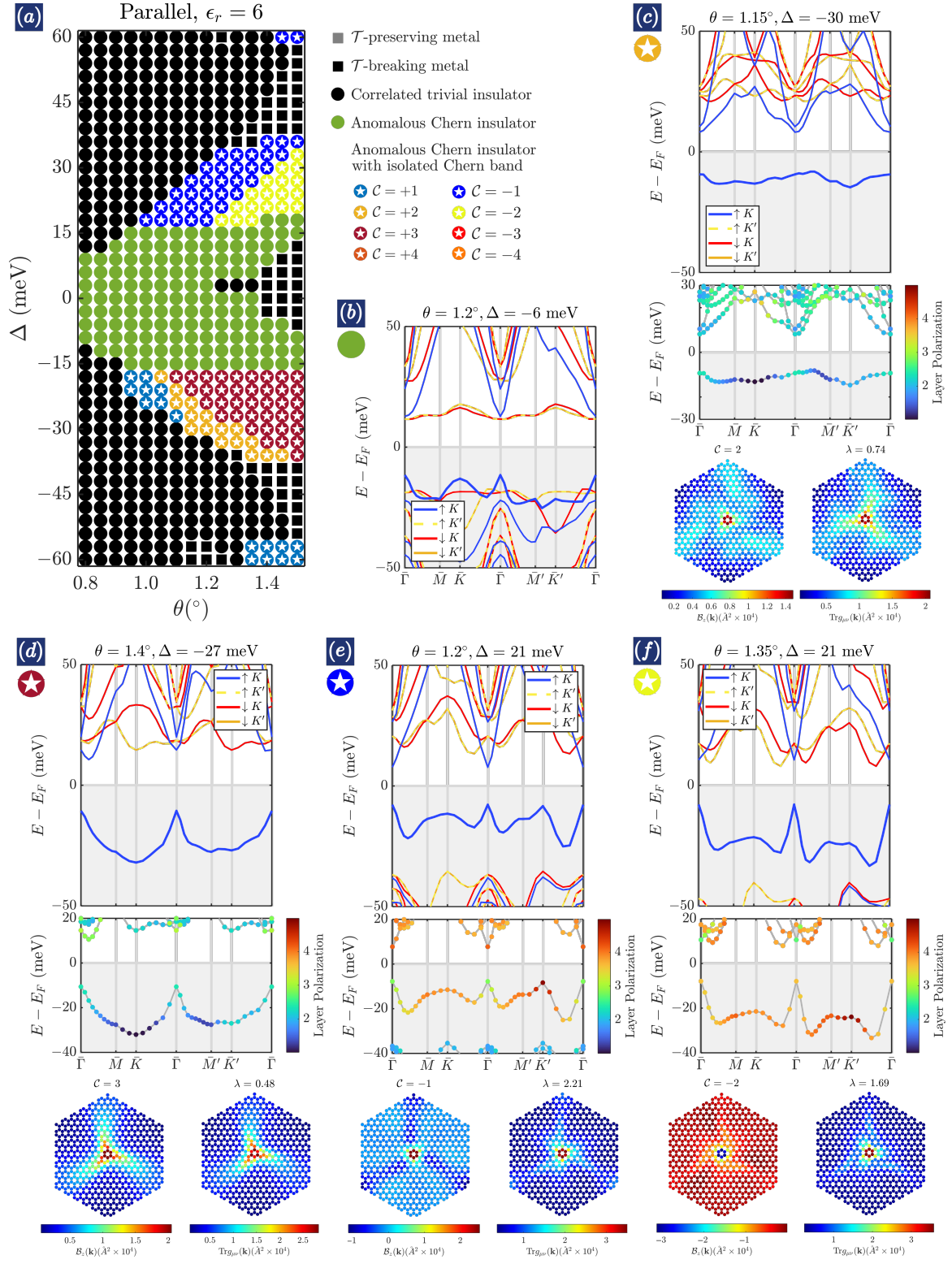
Figure S33. **Mean-field phase diagram for the parallel configuration with lattice relaxation and** $\epsilon_r = 6$. (a) Phase diagram as a function of θ and Δ. The results here are simulated using parameters from `Param_one`. Representative band structures of the different phases are shown in (b)-(f): an anomalous Chern insulator with no gap below the band closest to chemical potential (b) and an anomalous Chern insulator with isolated Chern band with $\mathcal{C} = +2$ (c), $\mathcal{C} = +3$ (d), $\mathcal{C} = -1$ (e), and $\mathcal{C} = -2$ (f). For (c)-(f), we color code the layer polarization of the Bloch states; furthermore, we show the corresponding Berry curvature and quantum metric of the band closest to chemical potential.
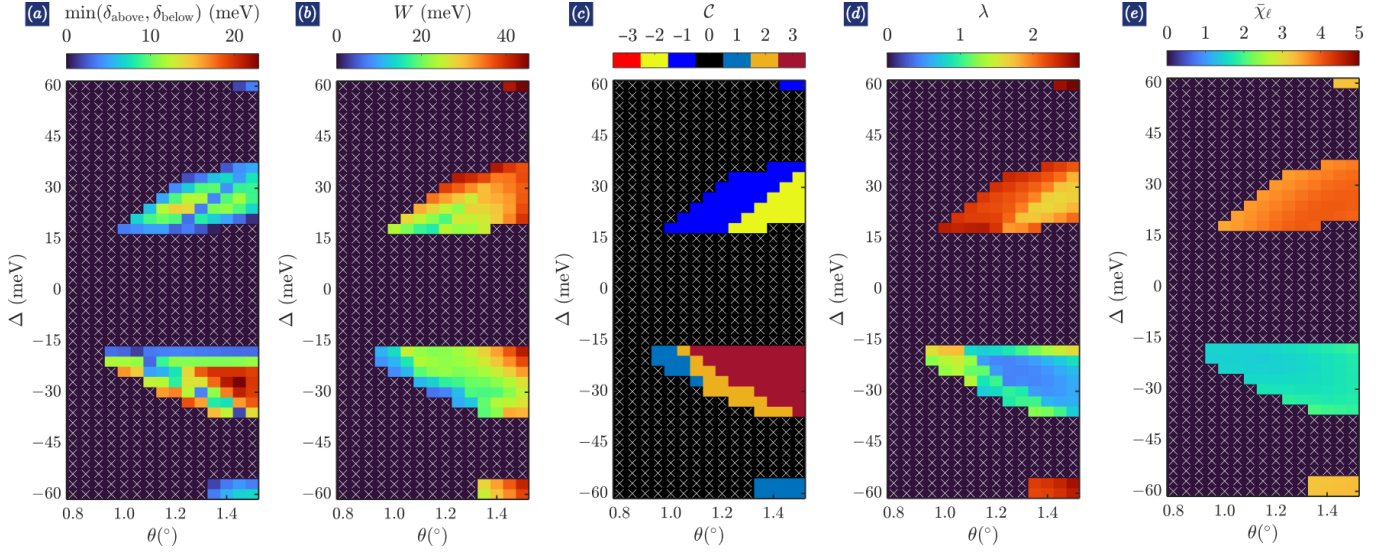
Figure S34. **Further spectral and topological characterization of anomalous Chern phases in Fig. S33.** For anomalous Chern phases with an isolated Chern band, we show following properties of that isolated band: spectral gap (a), bandwidth (b), Chern number (c), trace condition violation (d), and average layer polarization (e).
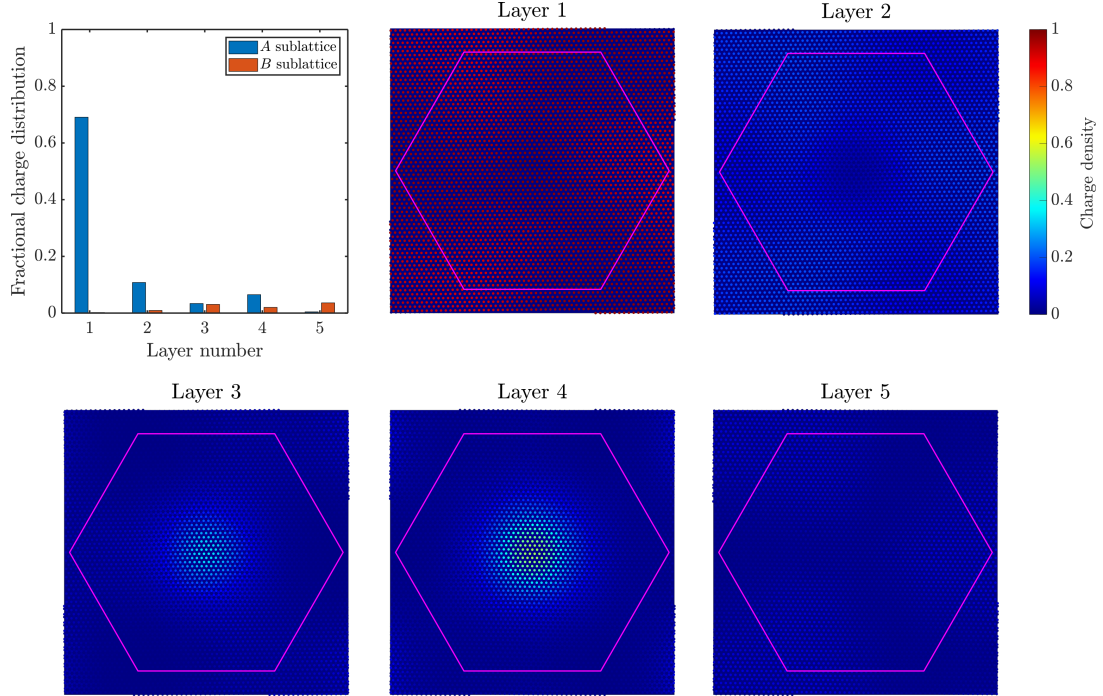

Figure S35. **Charge density at $\theta = 1.15°$ and $\Delta = -30$ meV in Fig. S33.** In the bar graph, we plot the fractional charge distribution on each layer and sublattice summed over all $\mathbf{k}$ in the narrow band in Fig. S33(c), demonstrating a majority population on layer 1. In the density plots, we show the real-space distribution of the summed charge density for that same band. Most of the charge density is on layer 1, concentrated more at α and γ regions than other regions.
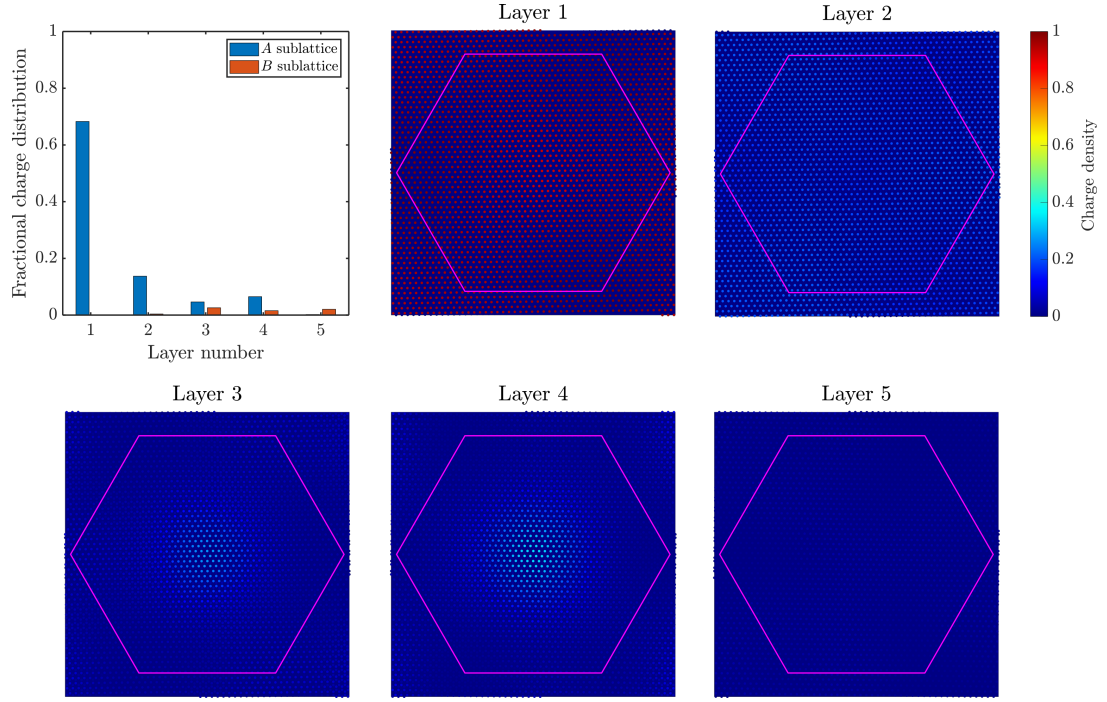
Figure S36. **Charge density at** $\theta = 1.40°$ **and** $\Delta = -27$ **meV in Fig. S33.** In the bar graph, we plot the fractional charge distribution on each layer and sublattice summed over all **k** in the narrow band in Fig. S33(d), demonstrating a majority population on layer 1. In the density plots, we show the real-space distribution of the summed charge density for that same band. Most of the charge density is on layer 1, concentrated more at γ regions than other regions.
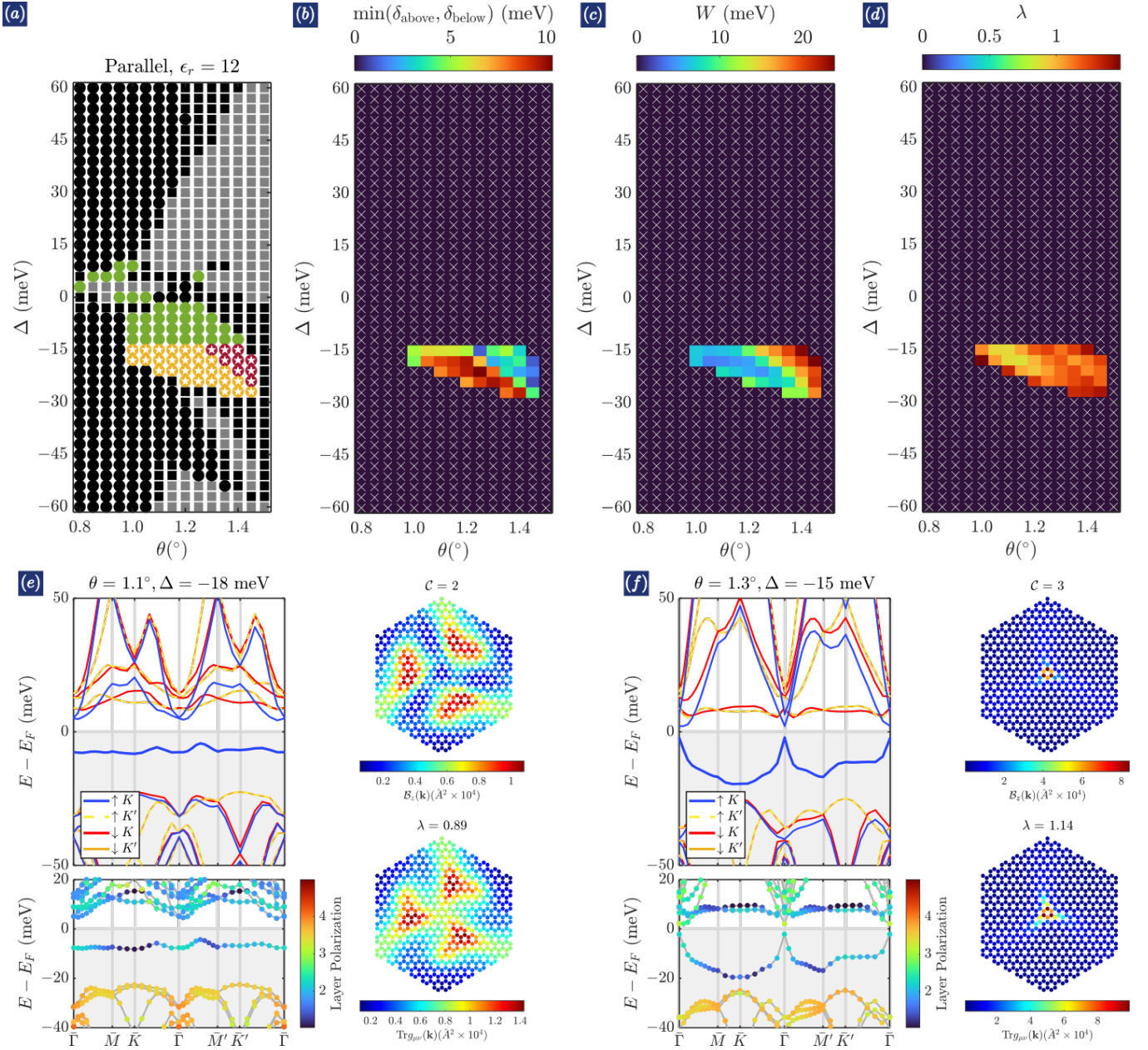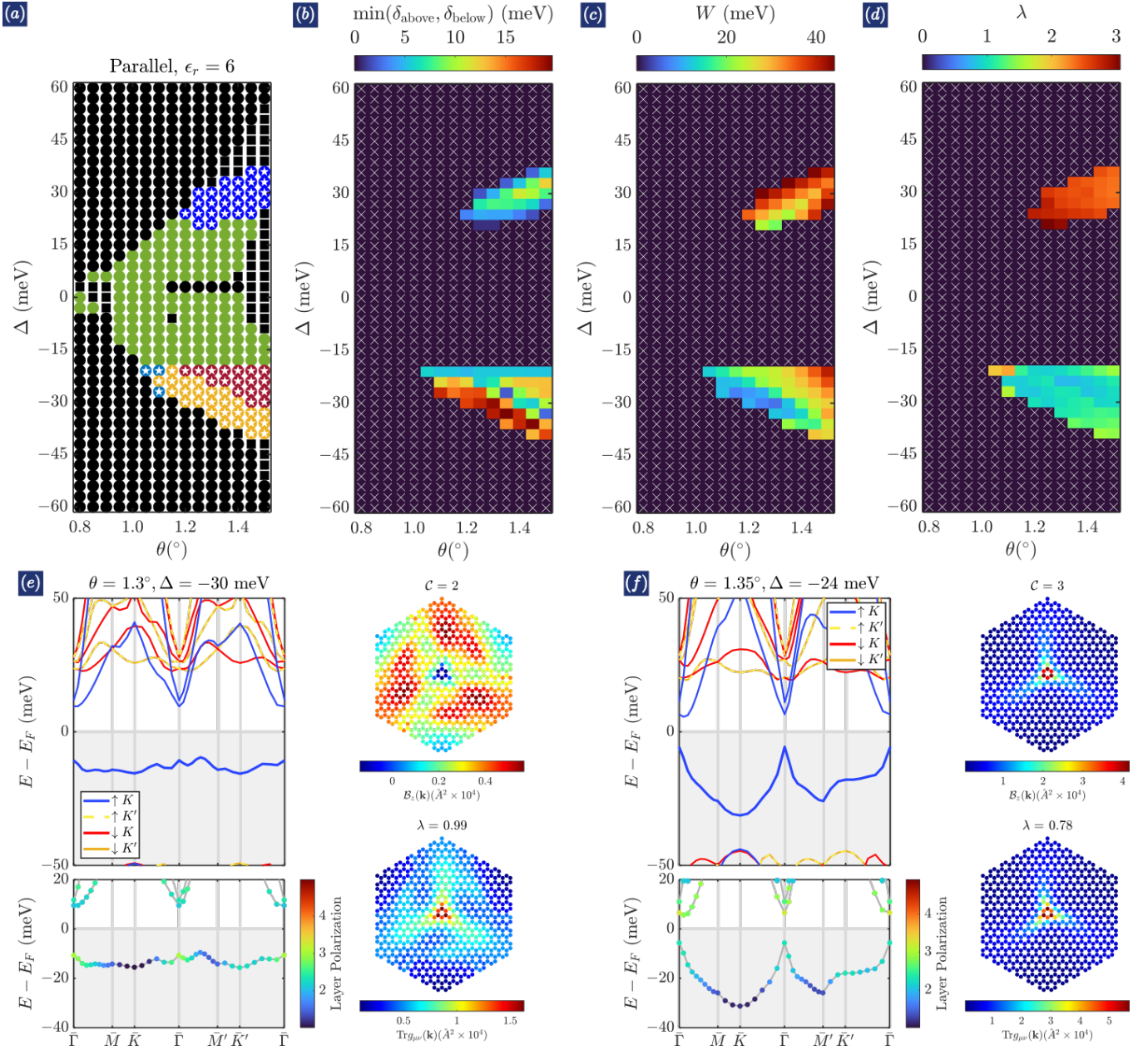
Figure S37. **Mean-field phase diagram for the parallel configuration without lattice relaxation amd** $\epsilon_r = 12$**.** The results here are simulated using Param_two. (a) Phase diagram and (b,c,d) its further characterization: band gap, bandwidth, and trace condition violation. (e) Band structure and associated Berry curvature and quantum metric distribution of an f.o.m. $\mathcal{C} = 2$ state. (f) Band structure and associated Berry curvature and quantum metric distribution of an f.o.m. $\mathcal{C} = 3$ state.

Figure S38. **Mean-field phase diagram for the parallel configuration without lattice relaxation amd** $\epsilon_r = 6$**.** The results here are simulated using `Param_two`. (a) Phase diagram and (b,c,d) its further characterization: band gap, bandwidth, and trace condition violation. (e) Band structure and associated Berry curvature and quantum metric distribution of an f.o.m. $\mathcal{C} = 2$ state. (f) Band structure and associated Berry curvature and quantum metric distribution of an f.o.m. $\mathcal{C} = 3$ state.
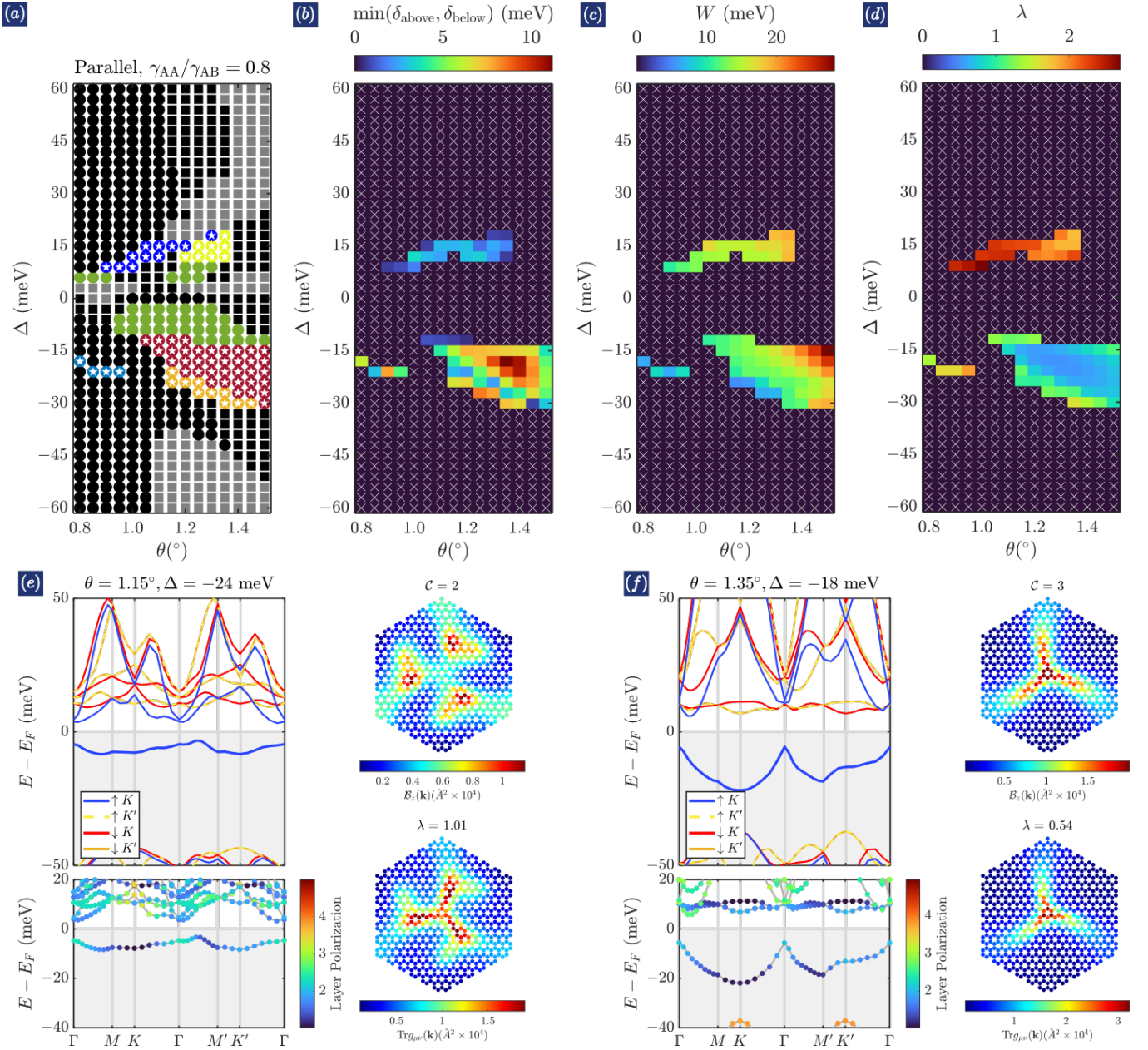
Figure S39. **Mean-field phase diagram for the parallel configuration with reduced γ_0 and $\{\gamma_{AA}, \gamma_{AB}\} = \{80, 100\}$ meV.** The results here are simulated using Param_three. (a) Phase diagram and (b,c,d) its further characterization: band gap, bandwidth, and trace condition violation. (e) Band structure and associated Berry curvature and quantum metric distribution of an f.o.m. $\mathcal{C} = 2$ state. (f) Band structure and associated Berry curvature and quantum metric distribution of an f.o.m. $\mathcal{C} = 3$ state. Here, $\epsilon_r = 12$.
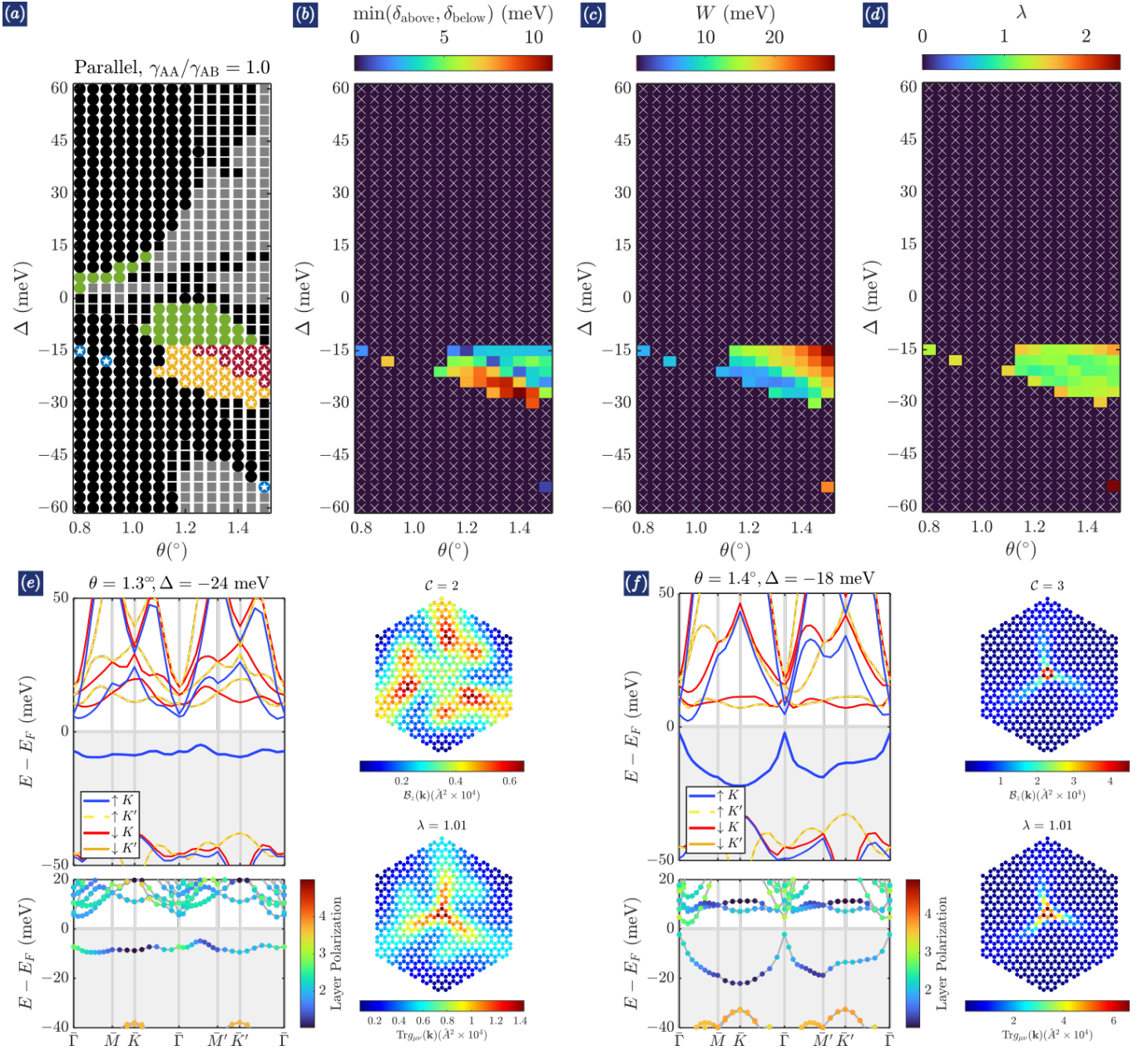
Figure S40. **Mean-field phase diagram for the parallel configuration with reduced** γ_0 **and** $\{\gamma_{AA}, \gamma_{AB}\} = \{100, 100\}$ **meV.** The results here are simulated using Param_four. (a) Phase diagram and (b,c,d) its further characterization: band gap, bandwidth, and trace condition violation. (e) Band structure and associated Berry curvature and quantum metric distribution of an f.o.m. $\mathcal{C} = 2$ state. (f) Band structure and associated Berry curvature and quantum metric distribution of an f.o.m. $\mathcal{C} = 3$ state. Here, $\epsilon_r = 12$.
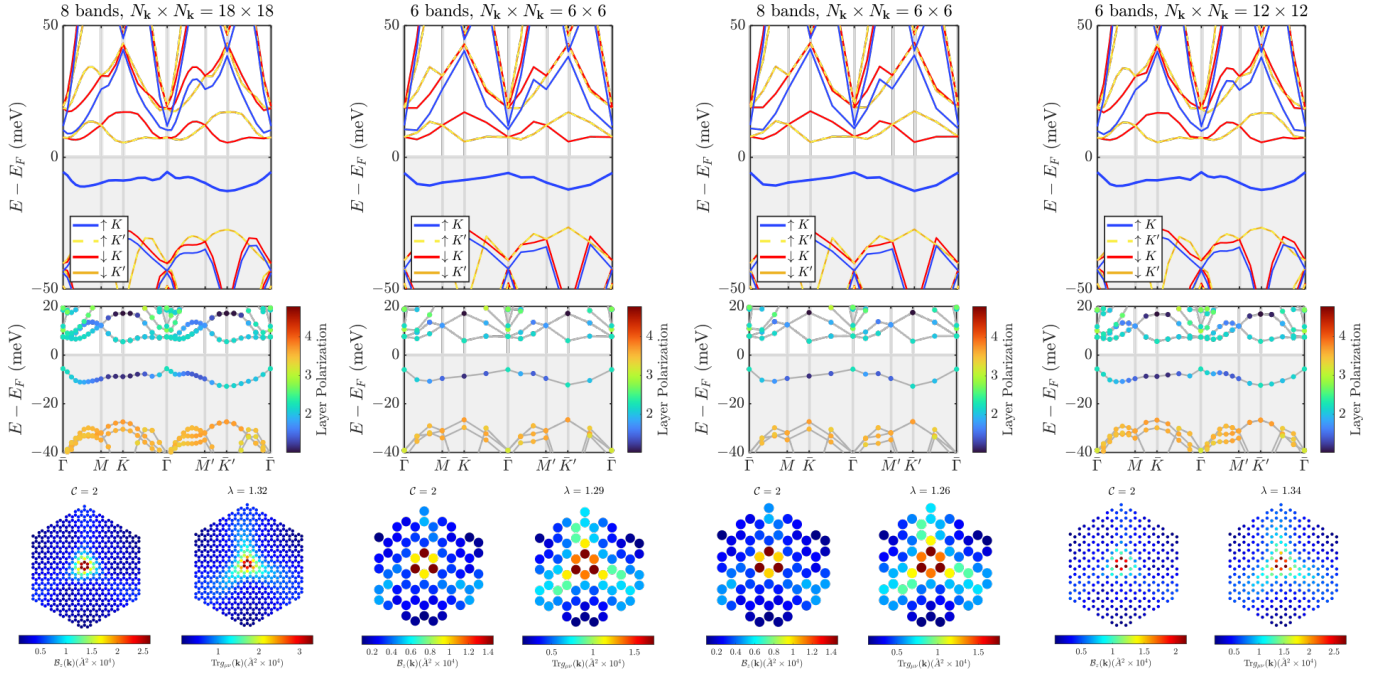
Figure S41. **Convergence of** $\mathcal{C} = 2$ **phase simulation for the antiparallel configuration.** The state is at $(\theta, \Delta) = (1.20°, -18 \text{ meV})$ with $\epsilon_r = 12$. The different band structures are simulated including different numbers of bands and using different mesh sizes, as indicated in the titles. The corresponding Berry curvature and quantum metric are shown below each band structure.
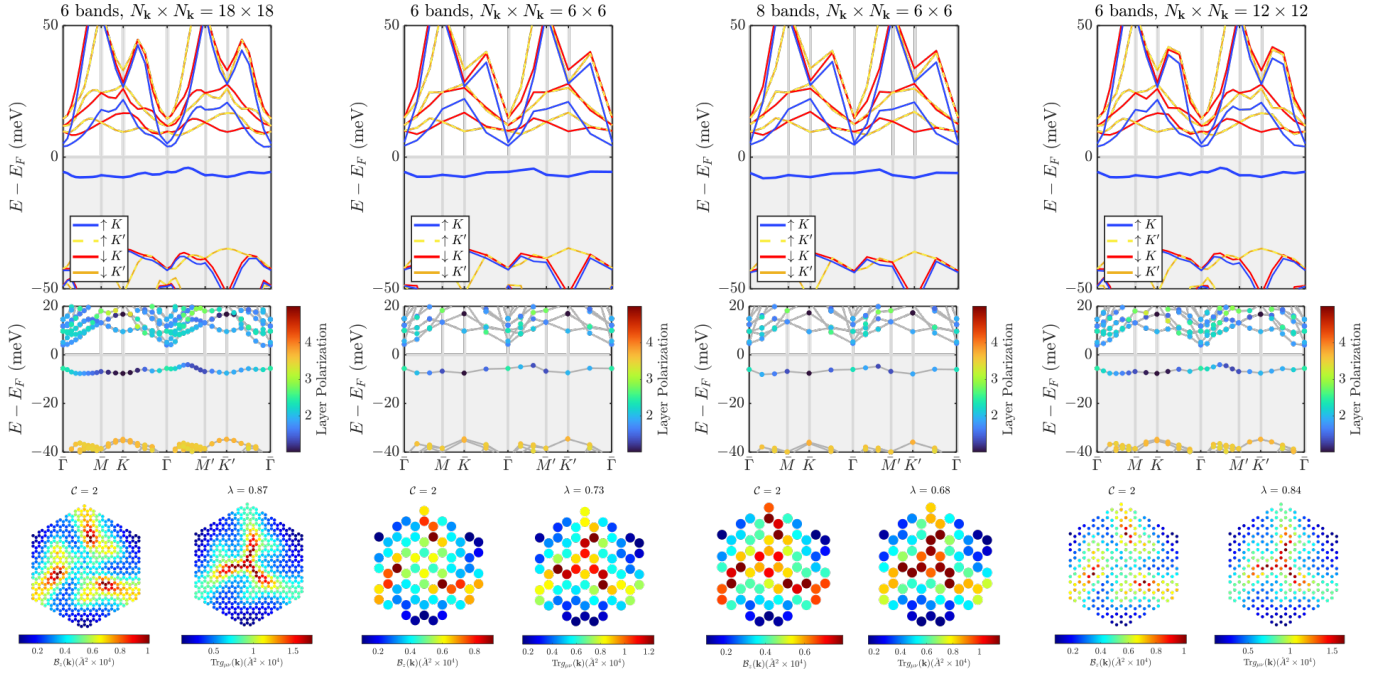

Figure S42. **Convergence of** $\mathcal{C} = 2$ **phase simulation for the parallel configuration.** The state is at $(\theta, \Delta) = (1.10°, -21 \text{ meV})$ with $\epsilon_r = 12$. The different band structures are simulated including different numbers of bands and using different mesh sizes, as indicated in the titles. The corresponding Berry curvature and quantum metric are shown below each band structure.
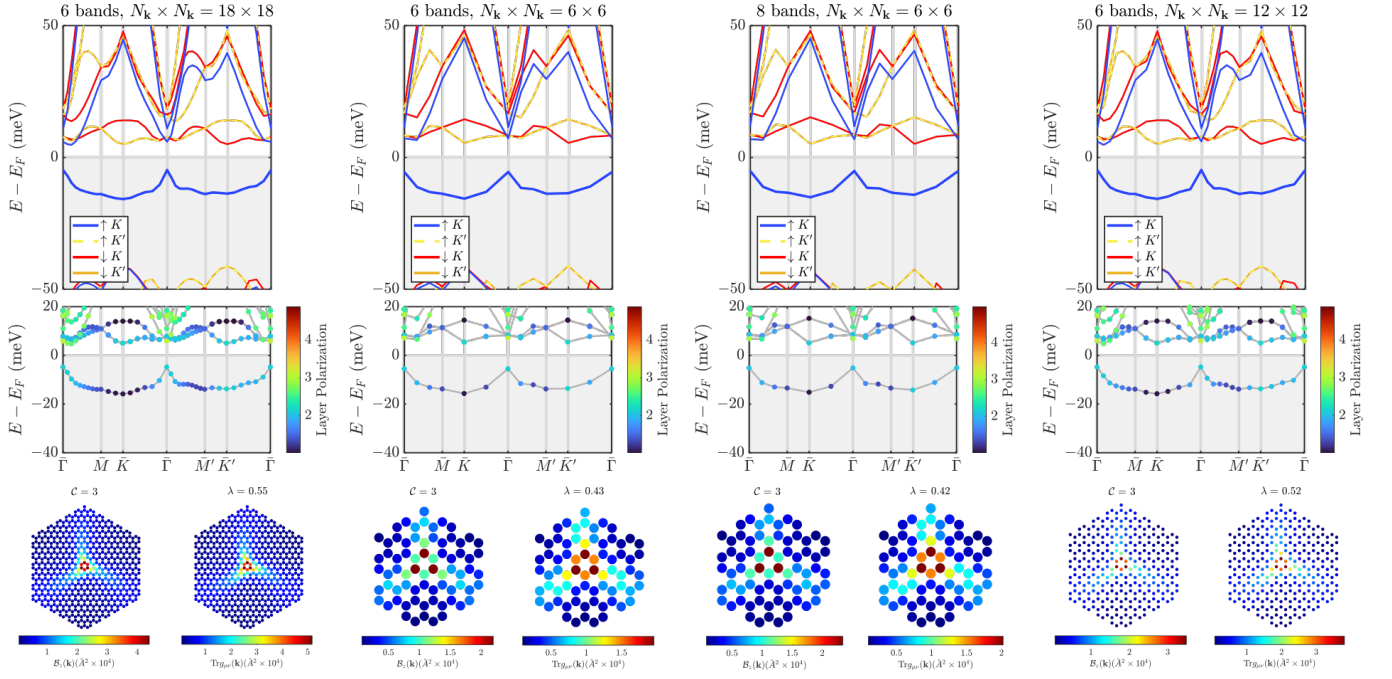
Figure S43. **Convergence of** $\mathcal{C} = 3$ **phase simulation for the parallel configuration.** The state is at $(\theta, \Delta) = (1.30°, -21 \text{ meV})$ with $\epsilon_r = 12$. The different band structures are simulated including different numbers of bands and using different mesh sizes, as indicated in the titles. The corresponding Berry curvature and quantum metric are shown below each band structure.
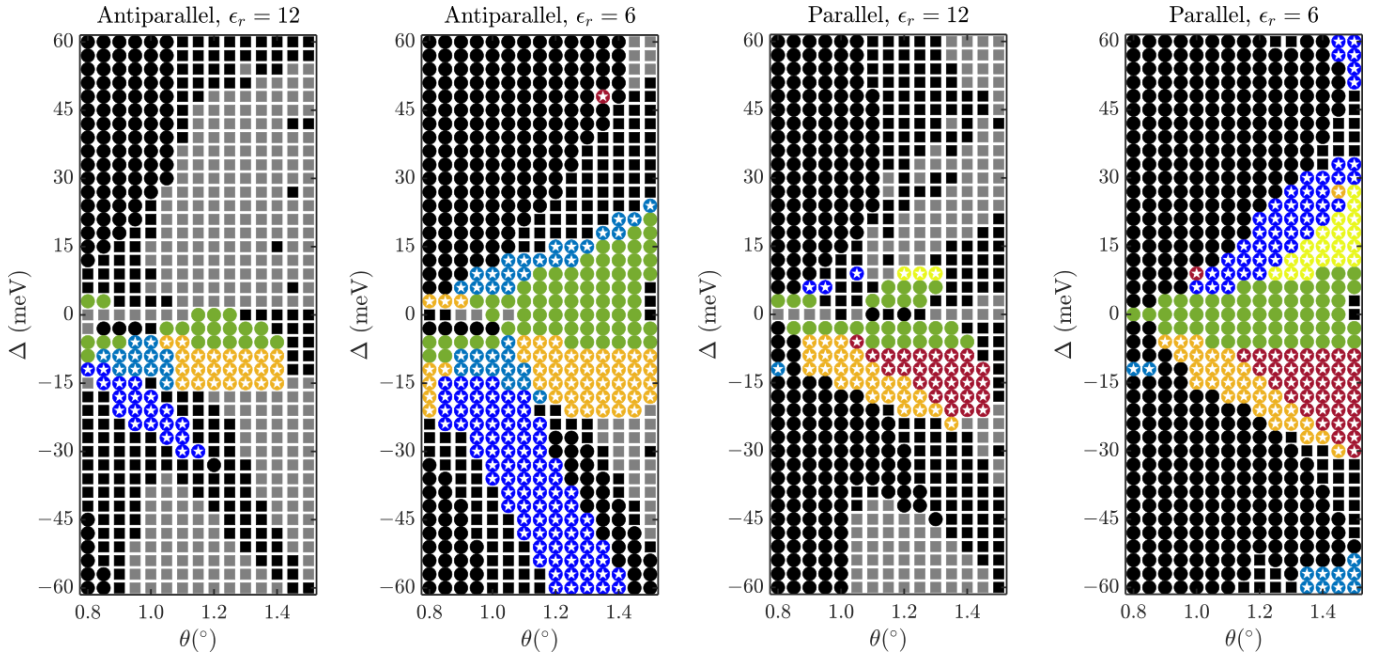

Figure S44. **Mean-field phase diagrams for the antiparallel and parallel configurations simulated with the layer-independent Coulomb potential.** The results here are simulated with `Param_one` and with $\epsilon_r = 12$ or $\epsilon_r = 6$.