Anderson's Theorem for Correlated Insulating States in Twisted Bilayer Graphene

Kryštof Kolář,[1] Gal Shavit[ID],[2] Christophe Mora[ID],[3] Yuval Oreg,[2] and Felix von Oppen[ID][1]

[1]*Dahlem Center for Complex Quantum Systems and Fachbereich Physik, Freie Universität Berlin, 14195 Berlin, Germany*
[2]*Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot, Israel 7610001*
[3]*Université Paris Cité, CNRS, Laboratoire Matériaux et Phénomènes Quantiques, 75013 Paris, France*

The emergence of correlated insulating phases in magic-angle twisted bilayer graphene exhibits strong sample dependence. Here, we derive an Anderson theorem governing the robustness against disorder of the Kramers intervalley coherent (K-IVC) state, a prime candidate for describing the correlated insulators at even fillings of the moiré flat bands. We find that the K-IVC gap is robust against local perturbations, which are odd under $\mathcal{PT}$, where $\mathcal{P}$ and $\mathcal{T}$ denote particle-hole conjugation and time reversal, respectively. In contrast, $\mathcal{PT}$-even perturbations will in general induce subgap states and reduce or even eliminate the gap. We use this result to classify the stability of the K-IVC state against various experimentally relevant perturbations. The existence of an Anderson theorem singles out the K-IVC state from other possible insulating ground states.

Introduction.—Twisted bilayer graphene (TBG) is recently attracting much attention as a highly tunable platform of strongly correlated electrons. Twisting two graphene sheets introduces a moiré lattice, which supports exceptionally flat bands for certain twist angles [1]. At these magic angles, the kinetic energy of the electrons is effectively suppressed and the system is prone to developing interaction-driven correlated phases. Corresponding experiments exhibit signatures of correlated insulators, superconductivity, nematicity, integer and fractional Chern insulators, spontaneous flavor polarization, as well as orbital ferromagnetism [2–17].

A prime candidate for understanding the correlated insulating phases, which occur near even integer fillings of the moiré flat bands are Kramers intervalley coherent (K-IVC) states [18–24]. These states exhibit a pattern of magnetization currents, which triple the graphene unit cell, thereby breaking the lattice translation symmetry as well as time reversal. The associated spontaneous coherence between the two valleys of the TBG band structure gaps out the moiré Dirac points and induces insulating behavior. A recent work reports evidence for the K-IVC state by measuring the magnetic-field dependence of the thermodynamic gap [25]. In general, the appearance and strength of insulating states tend to be device dependent [8,26]. A possible explanation for this sample-specific behavior lies in residual disorder associated with random strain or impurity potentials. Thus, it is important to study and understand their effects.

Previous works [22,27–31] predominantly considered smooth disorder and twist-angle variations, for which the associated long-range domain patterns can be directly mapped in experiment [6,7,9,10,15,32]. Here, we focus on local impurities. We show that for the K-IVC state, one can systematically classify impurities according to their ability to induce subgap excitations (Fig. 1), which diminish or even eliminate the insulating gap. Our discussion is strongly informed by a far-reaching analogy with the familiar problem of classifying impurities in s-wave superconductors [33].

Physically, these analogies can be understood by noting that K-IVC states can be thought of as binding holes in one valley to electrons in the other, akin to excitonic insulators. Evidently, this is similar to binding time-reversed electrons into Cooper pairs. According to Anderson's theorem [33], the ability of impurities to induce subgap excitations in s-wave superconductors is controlled by whether or not they respect time-reversal symmetry. We find that particle-hole symmetry plays a similar role for the K-IVC state as time-reversal symmetry does for superconductors. This is consistent with the picture of bound electron-hole pairs.

In contrast to the K-IVC state, we find that there are no corresponding Anderson theorems for other possible insulating ground states such as the valley polarized and valley Hall states.
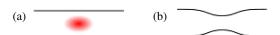

FIG. 1. Examples of impurities in twisted bilayer graphene (schematically shown as black lines), which are (a) even (charged impurity located between the layers; schematically shown in red) and (b) odd (local change in interlayer tunneling) under $\mathcal{PT}$ ($\mathcal{P}$ and $\mathcal{T}$ denote particle-hole and time-reversal conjugation, respectively). $\mathcal{PT}$-even impurities induce subgap states in the K-IVC phase of TBG, while $\mathcal{PT}$-odd impurities do not due to Anderson's theorem for the K-IVC state.

Anderson's theorem for s-wave superconductors.—To clearly bring out the analogies, as well as differences, between TBG and superconductors, we begin our discussion with a review of Anderson's theorem for *s*-wave superconductors [33–36], using a formulation which turns out to be adaptable to TBG. Starting from the second-quantized BCS mean-field Hamiltonian $\mathcal{H} = \frac{1}{2} \int d\mathbf{r} \Psi^{\dagger}(\mathbf{r}) H \Psi(\mathbf{r})$, we write the Bogoliubov–de Gennes (BdG) Hamiltonian

$$H = \begin{pmatrix} H_e & \Delta \\ \Delta & H_h \end{pmatrix}, \tag{1}$$

in a four-component Nambu formalism, using the basis $\Psi = [\psi_{\uparrow}, \psi_{\downarrow}, \psi_{\downarrow}^{\dagger}, -\psi_{\uparrow}^{\dagger}]$. The normal-state Hamiltonians $H_{e/h}$ for electrons (*e*) and holes (*h*) as well as the pairing Δ are matrices in spin space. For *s*-wave pairing, Δ is proportional to the unit matrix. In the following, we assume that Δ is spatially uniform and chosen to be real.

The BdG Hamiltonian is constrained by antisymmetry under particle-hole conjugation, $\mathcal{P}H\mathcal{P}^{-1} = -H$. As a consequence, the normal-state Hamiltonians of electrons and holes in Eq. (1) are related by time reversal $\mathcal{T}$,

$$H_h = -\mathcal{T}H_e\mathcal{T}^{-1}. \tag{2}$$

This can be deduced by defining Pauli matrices τ_{α} and s_{α} in particle-hole and spin space, respectively. Then, particle-hole conjugation is implemented by $\mathcal{P} = -i\tau_y\mathcal{T} = \tau_y s_y \mathcal{K}$ and squares to unity, $\mathcal{P}^2 = 1$, while time reversal takes the form $\mathcal{T} = is_y\mathcal{K}$ with $\mathcal{T}^2 = -1$ ($\mathcal{K}$ implements complex conjugation).

We separate the normal-state Hamiltonian $H_e = H_0 + U$ into a spatially homogeneous part H_0 and a (local) perturbation U. While we assume $H_0 = \mathcal{T}H_0\mathcal{T}^{-1}$ to be time-reversal symmetric, a general perturbation $U = U_+ + U_-$ can have components $U_{\pm} = \pm\mathcal{T}U_{\pm}\mathcal{T}^{-1}$, which are even (+) or odd (−) under time reversal. Combining these symmetry properties under time reversal with Eq. (2), the BdG Hamiltonian in Eq. (1) can be written compactly as

$$H = H_0\tau_z + \Delta\tau_x + U_+\tau_z + U_-\tau_0. \tag{3}$$

Importantly, one observes that time-reversal-symmetric perturbations anticommute with the order-parameter term, $\{\Delta\tau_x, U_+\tau_z\} = 0$, while the time-reversal breaking term, U_-, commutes.

It can now be seen quite generally that time-reversal-even perturbations do not reduce the BdG gap (Anderson's theorem). Given that antisymmetry under particle-hole conjugation $\mathcal{P}$ enforces the eigenenergies to be symmetric about zero energy, the spectrum can be deduced from the square of H,

$$H^2 = (H_0 + U_+)^2 + \Delta^2, \tag{4}$$

implying that the magnitude of the eigenvalues of H is bounded from below by Δ. This argument uses the assumption that the gap remains uniform in the presence of the perturbation, but holds regardless of the particular spatial structure of the impurity potential.

Conversely, perturbations, which are odd under time reversal, generally reduce the gap. A uniform Zeeman field described by $U_- = \mathbf{B} \cdot \mathbf{s}$ reduces the gap to $\Delta - |\mathbf{B}|$, provided the normal-state Hamiltonian is spin-rotation invariant. Local magnetic impurities with $U_- = J\mathbf{S} \cdot \mathbf{s}\delta(\mathbf{r})$ are well known to induce Yu-Shiba-Rusinov states at subgap energies [37–40].

Twisted bilayer graphene.—We begin our discussion of TBG by introducing its band Hamiltonian $h(\mathbf{k})$, after projection to the eight flat bands. It is conveniently written in the Chern basis spanned by the spin, valley (Pauli matrices τ_{α}), and band (Pauli matrices σ_{α}) degrees of freedom [18],

$$h(\mathbf{k}) = h_0(\mathbf{k})\tau_z + h_x(\mathbf{k})\sigma_x + h_y(\mathbf{k})\sigma_y\tau_z. \tag{5}$$

Time-reversal symmetry enforces $h_0(\mathbf{k}) = -h_0(-\mathbf{k})$ and $h_{x,y}(\mathbf{k}) = h_{x,y}(-\mathbf{k})$. Because of the negligible spin-orbit coupling of graphene, the Hamiltonian is a unit matrix in spin space. The label σ_z is associated with the Chern number $C = \sigma_z\tau_z$ and (partial) sublattice polarization. While the Pauli matrices τ_{α} refer to different degrees of freedom in our descriptions of superconductors and TBG, we shall see that they actually play rather analogous roles.

In addition to spatial and spin rotation symmetries as well as charge conservation, the Hamiltonian $h(\mathbf{k})$ conserves valley charge, obeys spinless time-reversal symmetry, and has an (approximate) particle-hole antisymmetry [18,41–43]. The latter three symmetries are central to our discussion. The conservation of valley charge is associated with invariance under $U(1)$ valley rotations $U_V = e^{i\theta\tau_z}$, spinless time reversal is implemented by $\mathcal{T} = \tau_x\mathcal{K}$ with $\mathcal{T}^2 = 1$, and particle-hole conjugation takes the form $\mathcal{P} = i\sigma_y\tau_z\mathcal{K}$ with $\mathcal{P}^2 = -1$. The particle-hole antisymmetry of TBG emerges when neglecting the small relative twist of the Dirac Hamiltonians of the two layers in the Bistrizer-MacDonald model [1].

Within the mean-field approximation for the K-IVC state [18], the band Hamiltonian $h(\mathbf{k})$ is complemented by the order parameter $h_{\mathrm{IVC}} = \Delta\sigma_y(\tau_x\cos\theta + \tau_y\sin\theta)$, where θ denotes an arbitrary phase. In view of the associated magnetization currents, the K-IVC state breaks time reversal spontaneously, $\mathcal{T}h_{\mathrm{IVC}}\mathcal{T}^{-1} = -h_{\mathrm{IVC}}$. However, it preserves a modified time-reversal symmetry [18]

$$\mathcal{T}' = i\tau_y\mathcal{K}, \tag{6}$$

which concatenates $\mathcal{T}$ with a valley rotation, $\mathcal{T}' = \tau_z \mathcal{T}$. Both $\mathcal{T}$ and valley rotations are symmetries of the single-particle Hamiltonian, so that the mean-field Hamiltonian $H(\mathbf{k}) = h(\mathbf{k}) + h_{\text{IVC}}$ conserves the Kramers time reversal $\mathcal{T}'$ (with $\mathcal{T}'^2 = -1$) as a whole.

The Hamiltonian $H(\mathbf{k})$ can be thought of as the analog of the BdG Hamiltonian for the K-IVC state. We will now make the analogies yet more explicit by a change of basis $H \to \mathcal{U} H \mathcal{U}^\dagger$ with

$$\mathcal{U} = \begin{pmatrix} 1 & 0 \\ 0 & i\sigma_y \end{pmatrix}. \qquad (7)$$

In the new basis, which we refer to as the particle-hole basis, the Chern number becomes $C = \sigma_z$. Transforming the Hamiltonian in this manner, we find

$$H(\mathbf{k}) = H_0(\mathbf{k})\tau_z + \Delta(\tau_x \cos\tilde{\theta} + \tau_y \sin\tilde{\theta}) \qquad (8)$$

($\tilde{\theta} = \theta + (\pi/2)$). Here, we make the dependence on the valley Pauli matrices τ_α explicit, while H_0 and Δ are still matrices in sublattice space. We find $H_0(\mathbf{k}) = h_0(\mathbf{k}) + h_x(\mathbf{k})\sigma_x + h_y(\mathbf{k})\sigma_y$ for the single-particle Hamiltonian of the K valley, while Δ is simply proportional to the unit matrix. The transformation (7) to the particle-hole basis also changes the explicit forms of the time-reversal and charge-conjugation operations, $\mathcal{T} \to \mathcal{U}\mathcal{T}\mathcal{U}^\dagger$ and $\mathcal{P} \to \mathcal{U}\mathcal{P}\mathcal{U}^\dagger$, which yields $\mathcal{P} = i\sigma_y \mathcal{K}$ and $\mathcal{T}' = -\tau_x \mathcal{P}$.

Equation (8) is closely analogous to the BdG Hamiltonian of s-wave superconductors, with particle-hole space replaced by the valley degree of freedom and spin space replaced by sublattice space. In Eq. (8), the band Hamiltonian $H_0(\mathbf{k})$ is analogous to the normal-state Hamiltonian. It multiplies τ_z as a consequence of the chiral antisymmetry $\mathcal{PT} = i\tau_y$ of the TBG Hamiltonian. Moreover, the term describing K-IVC order is analogous to the pairing term in the BdG Hamiltonian, being off diagonal in valley and proportional to the unit matrix in sublattice space.

Beyond the structural similarities of the Hamiltonians, there is also a correspondence of symmetries. Interestingly, the roles of time reversal and particle-hole conjugation are essentially reversed. For superconductors, time reversal acts diagonally in particle-hole space, while particle-hole conjugation is off diagonal. In contrast, in TBG it is time reversal that maps between the two valleys, while particle-hole conjugation acts separately within each valley.

We also note that gauge transformations for superconductors are structurally analogous to U(1) valley rotations in TBG. For superconductors, the BdG Hamiltonian becomes explicitly symmetric under the conventional time reversal operator $\mathcal{T} = is_y \mathcal{K}$ when choosing a gauge, in which Δ is real. In TBG, we can similarly exploit the valley rotation symmetry to choose $\tilde{\theta} = \pi/2$, so that

$H(\mathbf{k}) = H_0(\mathbf{k})\tau_z + \Delta\tau_y$. With this choice, the mean-field K-IVC order is also odd under particle-hole conjugation, so that $\mathcal{P}H(\mathbf{k})\mathcal{P}^{-1} = -H(-\mathbf{k})$. In the following, we make this choice for definiteness. However, just as Anderson's theorem for s-wave superconductors is not specific to a particular gauge, Anderson's theorem for TBG is not limited to this choice.

Impurities and K-IVC states.—Armed with this far-reaching correspondence between the BdG Hamiltonian of s-wave superconductors and the K-IVC state of TBG, we now turn to discussing the effects of impurities on the K-IVC state. We consider impurity potentials which are sufficiently smooth on the scale of the atomic lattice, so that they preserve the $U(1)$ valley symmetry. Intervalley scattering can then be neglected and the impurity potential is diagonal in valley space. With this assumption, the low-energy Hamiltonian in the presence of an impurity potential becomes

$$H = H_0\tau_z + \Delta\tau_y + U_-\tau_z + U_+\tau_0. \qquad (9)$$

Just as for superconductors, the impurity potentials $U_\pm$ (which are matrices in sublattice and spin space) are distinguished by their symmetry properties. For superconductors, antisymmetry under particle-hole conjugation is built into the BdG formalism. For this reason, it was sufficient to classify perturbations according to their behavior under time reversal. In contrast, for TBG, both Kramers time reversal and particle-hole conjugation are physical symmetries of the Hamiltonian. Consequently, we now classify perturbations according to their transformation properties under the combined chiral symmetry operation $\mathcal{PT} = i\tau_y$, namely, $(\mathcal{PT})U_+\tau_0(\mathcal{PT})^{-1} = U_+\tau_0$ and $(\mathcal{PT})U_-\tau_z(\mathcal{PT})^{-1} = -U_-\tau_z$. (Notice that due to valley rotation symmetry, the impurity terms transform in the same way under $\mathcal{PT}$ and $\mathcal{PT}'$.) By comparing with the discussion of Eq. (3), we can now formulate an Anderson's theorem for TBG, our central result: The gap of K-IVC states is robust against valley-preserving perturbations, which are odd under the combined chiral symmetry operation $\mathcal{PT}$. In fact, perturbations which are odd under $\mathcal{PT}$ anticommute with the K-IVC order $\Delta\tau_y$ and cannot reduce the gap. In contrast, similar to time-reversal-breaking impurities in superconductors, perturbations which are even under $\mathcal{PT}$ can induce subgap states in TBG. For perturbations that are local on the moiré scale, this follows as for time-reversal-breaking impurities in s-wave superconductors. A finite density of $\mathcal{PT}$-even impurities can thus suppress or even destroy the K-IVC gap.

So far, our discussion relied on a close structural analogy between the BdG Hamiltonian of s-wave superconductors and the low-energy Hamiltonian of TBG with K-IVC order. More fundamentally, the appearance of an Anderson's theorem in both theories is rooted in the fact that up to $U(1)$ rotations which leave the normal-state Hamiltonian

invariant, their order parameters are proportional to a natural antisymmetry of the model, namely $\mathcal{PT}$. For both superconductors and the K-IVC state of TBG, we have $\Delta\tau_y = -i\Delta\mathcal{PT}$. Up to a gauge transformation (superconductors) or a $U(1)$ valley rotation (TBG), this is equivalent to the general order-parameter term $\Delta(\tau_x\cos\theta + \tau_y\sin\theta)$. This form of the order parameter has two important consequences. First, the order-parameter and single-particle terms in the Hamiltonian anticommute, so that the single-particle energies and the order parameter add in quadrature in the mean-field excitation spectrum. Second, this property persists in the presence of disorder, as long as the latter is odd under $\mathcal{PT}$, which is Anderson's theorem.

One should remember that the derivation of Anderson's theorem relies on several assumptions. In particular, one assumes that the order parameter remains spatially uniform and is momentum independent. Similar to anisotropic superconductors, the order parameter of TBG exhibits some momentum dependence [18]. In the presence of momentum dependence, there will be no systematic anticommutation behavior between the order-parameter term and the impurity potential. Then, Anderson's theorem no longer applies in the strict sense, and implies only enhanced, but not full protection of the gap.

Classifying physical perturbations.—So far, we have phrased our discussion in rather general terms, largely relying on symmetry properties of the TBG flat bands. We now classify perturbations according to their symmetry and tabulate the presence or absence of Anderson's theorem in Table I. For a given behavior under time reversal, it is sufficient to consider their transformation properties under $\mathcal{P}$, which acts separately within each valley.

Usually, we do not know the form of perturbations in the flat-band (i.e., Chern or particle hole) bases, but rather in the microscopic graphene basis. Within the

TABLE I. Table of time-reversal symmetric perturbations (left column) and their effect on the K-IVC gap. The central column gives the matrix structure in the microscopic graphene basis of TBG, where μ_α, σ_α, and τ_α are Pauli matrices in layer, sublattice, and valley space, respectively. The right column indicates the validity of Anderson's theorem. The K-IVC gap is protected against perturbations, for which Anderson's theorem is valid. Notice that in this table, we only consider strain-induced vector potentials (pot.).

Perturbation	Matrix structure	Anderson
Layer-even pot.	1	✗
Layer-odd pot.	μ_z	✓
Layer-even sublattice pot.	σ_z	✓
Layer-odd sublattice pot.	$\sigma_z\mu_z$	✗
Layer-even vector pot.	$\sigma_x, \sigma_y\tau_z$	✗
Layer-odd vector pot.	$\sigma_x\mu_z, \sigma_y\tau_z\mu_z$	✓
Tunneling disorder	see text	✓

Bistrizer-MacDonald model [1], the Hamiltonian H_0 is valley diagonal and takes the form

$$H = \begin{pmatrix} v_D\boldsymbol{\sigma}\cdot\left(\frac{1}{i}\boldsymbol{\nabla} + \mathbf{A}_t\right) + \phi_t & T(\mathbf{r}) \\ T^\dagger(\mathbf{r}) & v_D\boldsymbol{\sigma}\cdot\left(\frac{1}{i}\boldsymbol{\nabla} + \mathbf{A}_b\right) + \phi_b \end{pmatrix} \tag{10}$$

for the K valley. Here, the diagonal and off-diagonal blocks are intra- and interlayer terms, respectively, and the σ_α refer explicitly to the graphene sublattice. Potential disorder introduces layer- and sublattice-dependent potentials $\phi_{t/b}(\mathbf{r})$. Modulations in the interlayer distance cause variations of the interlayer tunneling terms $T(\mathbf{r})$. Strain introduces vector potentials $\mathbf{A}_{t/b}(\mathbf{r})$ and modifies $T(\mathbf{r})$ [44–46]. In terms of the strain-induced displacements $\mathbf{u}_l(\mathbf{r})$ relative to the uniformly twisted bilayer, the components of the vector potential take the form [45,46] $(A_l)_\mu = \mathbf{K}\cdot\partial_\mu\mathbf{u}_l + ([\beta\sqrt{3}]/[2a])(u_{l,xx} - u_{l,yy}, -2u_{l,xy})$. Here, β characterizes the sensitivity of the hopping amplitude to strain-induced displacements and $u_{l,ij}$ is the strain tensor of layer l. Time-reversal symmetry implies that strain-induced vector potentials are odd in valley space, while the scalar potentials are even.

In the microscopic graphene basis of the Bistritzer-MacDonald Hamiltonian in Eq. (10), particle-hole conjugation takes the form [18]

$$\mathcal{P} = i\sigma_x\mu_y\mathcal{K}, \tag{11}$$

where the μ_α are Pauli matrices in layer space ($\mathcal{PT} = i\tau_x\sigma_x\mu_y$). The validity of Anderson's theorem for various perturbations is now readily established and tabulated in Table I for time-reversal symmetric perturbations (also including their structure in the microscopic basis). A sublattice-symmetric potential will commute with $\mathcal{P}$, if it is layer symmetric, and anticommute with $\mathcal{P}$, if it is odd under layer exchange. According to our considerations, we find that layer-symmetric potentials induce subgap states within the K-IVC gap, but layer-odd potentials leave the K-IVC gap intact. These conclusions are reversed for sublattice-odd potentials. Tunneling disorder corresponds to a local variation in the strength of the interlayer tunneling amplitudes and thus in the parameters entering $T(\mathbf{r})$. Consequently, tunneling disorder inherits the $\mathcal{P}$ transformation properties of H_0 and Anderson's theorem applies. Finally, a layer-even vector potential (homostrain) is even under $\mathcal{P}$, while a layer-odd vector potential (heterostrain) is odd. We therefore find that Anderson's theorem applies to local heterostrain only. We note that due to a self-consistency effect, uniform heterostrain decreases the gap [45]. We neglect self-consistency effects since for local perturbations, spatial variations of the order parameter are suppressed by its stiffness on the scale of the correlation length.

Other insulating ground states.—While it has been argued that the K-IVC is the most favorable ground state at even integer fillings of the flat bands of TBG [18,20,21,47], other competing symmetry-broken states can also be considered [48–51]. We find that the K-IVC is distinct from other possible ground states due to the existence of Anderson's theorem for $\mathcal{PT}$-antisymmetric disorder. It is thus conceivable that such kinds of disorder stabilize the K-IVC state relative to competing states.

First, consider an alternative, time-reversal-preserving intervalley coherent state, termed T-IVC. In the particle-hole basis, this state has the order parameter $\Delta\sigma_z(\tau_x\cos\theta + \tau_y\sin\theta)$. The T-IVC gap anticommutes with only one of the three terms of the flat-band Hamiltonian, precluding the derivation of an Anderson's theorem.

The valley-polarized state with order parameter $\Delta\tau_z$ leads to the mean-field Hamiltonian

$$H_{\rm vp} = H_0\tau_z + \Delta\tau_z + U_-\tau_z + U_+\tau_0 \qquad (12)$$

in the particle-hole basis. The order-parameter term commutes with the band Hamiltonian, so that a gap emerges only when Δ shifts the flat bands of the two valleys sufficiently far apart in energy. The impurity problem can be considered separately for the two valleys and regardless of impurity type, there is no robustness due to an Anderson's theorem.

Finally, we consider the valley Hall state with mean field Hamiltonian

$$H_{\rm vh} = H_0\tau_z + \Delta\sigma_z\tau_z + U_-\tau_z + U_+\tau_0 \qquad (13)$$

in the particle-hole basis. The order-parameter term anticommutes with the flat-band Hamiltonian $H_0\tau_z$ [see Eq. (8)] only in the chiral limit, where $h_0(\mathbf{k}) = 0$ [52]. In this idealized (but experimentally remote) limit, the gap is robust against perturbations, which are purely off-diagonal in sublattice space, e.g., strain disorder.

Conclusions.—We have shown that an Anderson-type theorem protects the gap of the K-IVC state, a prime candidate for the observed robust correlated insulators of TBG, from certain types of disorder. Similar to *s*-wave superconductivity, which is robust against time-reversal preserving disorder, we find that the K-IVC gap is robust against disorder, which is odd under $\mathcal{PT}$. The robustness against some types of disorder distinguishes the K-IVC state from other candidate ground states for the correlated insulators and arises due to the special nature of the order parameter, which has the same matrix structure, up to a $U(1)$ valley rotation, as the $\mathcal{PT}$ chiral antisymmetry.

This special structure of the order-parameter term also underlies the close analogy of KIVC states with *s*-wave superconductors. In both cases, this structure guarantees that the band Hamiltonian as well as $\mathcal{PT}$-odd disorder potentials anticommute with a uniform and momentum-independent order-parameter term. For superconductors, $\mathcal{P}$ is an inherent antisymmetry, so that Anderson's theorem applies to $\mathcal{T}$-even perturbations. For TBG, both $\mathcal{P}$ and $\mathcal{T}$ are physical symmetries, so that Anderson's theorem applies to all perturbations which are odd under $\mathcal{PT}$, regardless of their behavior under $\mathcal{P}$ and $\mathcal{T}$ individually.

Our theoretical considerations can be tested by introducing impurity potentials with different behavior under $\mathcal{PT}$. Local $\mathcal{PT}$-even perturbations will in general induce subgap states, which can be probed directly using scanning tunneling microscopy. One expects that a finite density of impurities can reduce or even eliminate the K-IVC gap. We leave a detailed study of this last situation to a separate publication [53].

[1] R. Bistritzer and A. H. MacDonald, Proc. Natl. Acad. Sci. U.S.A. **108**, 12233 (2011).

[2] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, R. C. Ashoori, and P. Jarillo-Herrero, Nature (London) **556**, 80 (2018).

[3] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Nature (London) **556**, 43 (2018).

[4] X. Lu, P. Stepanov, W. Yang, M. Xie, M. A. Aamir, I. Das, C. Urgell, K. Watanabe, T. Taniguchi, G. Zhang, A. Bachtold, A. H. MacDonald, and D. K. Efetov, Nature (London) **574**, 653 (2019).

[5] M. Yankowitz, S. Chen, H. Polshyn, Y. Zhang, K. Watanabe, T. Taniguchi, D. Graf, A. F. Young, and C. R. Dean, Science **363**, 1059 (2019).

[6] Y. Choi, J. Kemmer, Y. Peng, A. Thomson, H. Arora, R. Polski, Y. Zhang, H. Ren, J. Alicea, G. Refael, F. von Oppen, K. Watanabe, T. Taniguchi, and S. Nadj-Perge, Nat. Phys. **15**, 1174 (2019).

[7] A. Kerelsky, L. J. McGilly, D. M. Kennes, L. Xian, M. Yankowitz, S. Chen, K. Watanabe, T. Taniguchi, J. Hone, C. Dean, A. Rubio, and A. N. Pasupathy, Nature (London) **572**, 95 (2019).

[8] Y. Saito, J. Ge, K. Watanabe, T. Taniguchi, and A. F. Young, Nat. Phys. **16**, 926 (2020).

[9] U. Zondiner, A. Rozen, D. Rodan-Legrain, Y. Cao, R. Queiroz, T. Taniguchi, K. Watanabe, Y. Oreg, F. von Oppen, A. Stern, E. Berg, P. Jarillo-Herrero, and S. Ilani, Nature (London) **582**, 203 (2020).

[10] D. Wong, K. P. Nuckolls, M. Oh, B. Lian, Y. Xie, S. Jeon, K. Watanabe, T. Taniguchi, B. A. Bernevig, and A. Yazdani, Nature (London) **582,** 198 (2020).

[11] X. Liu, Z. Wang, K. Watanabe, T. Taniguchi, O. Vafek, and J. I. A. Li, Science **371,** 1261 (2021).

[12] G. Chen, A. L. Sharpe, E. J. Fox, Y.-H. Zhang, S. Wang, L. Jiang, B. Lyu, H. Li, K. Watanabe, T. Taniguchi, Z. Shi, T. Senthil, D. Goldhaber-Gordon, Y. Zhang, and F. Wang, Nature (London) **579,** 56 (2020).

[13] M. Serlin, C. L. Tschirhart, H. Polshyn, Y. Zhang, J. Zhu, K. Watanabe, T. Taniguchi, L. Balents, and A. F. Young, Science **367,** 900 (2020).

[14] P. Stepanov, M. Xie, T. Taniguchi, K. Watanabe, X. Lu, A. H. MacDonald, B. A. Bernevig, and D. K. Efetov, Phys. Rev. Lett. **127,** 197701 (2021).

[15] A. Uri, S. Grover, Y. Cao, J. A. Crosse, K. Bagani, D. Rodan-Legrain, Y. Myasoedov, K. Watanabe, T. Taniguchi, P. Moon, M. Koshino, P. Jarillo-Herrero, and E. Zeldov, Nature (London) **581,** 47 (2020).

[16] Y. Xie, A. T. Pierce, J. M. Park, D. E. Parker, E. Khalaf, P. Ledwith, Y. Cao, S. H. Lee, S. Chen, P. R. Forrester, K. Watanabe, T. Taniguchi, A. Vishwanath, P. Jarillo-Herrero, and A. Yacoby, Nature (London) **600,** 439 (2021).

[17] J. M. Park, Y. Cao, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, Nature (London) **592,** 43 (2021).

[18] N. Bultinck, E. Khalaf, S. Liu, S. Chatterjee, A. Vishwanath, and M. P. Zaletel, Phys. Rev. X **10,** 031034 (2020).

[19] H. Ochoa, Phys. Rev. B **102,** 201107(R) (2020).

[20] J. S. Hofmann, E. Khalaf, A. Vishwanath, E. Berg, and J. Y. Lee, Phys. Rev. X **12,** 011061 (2022).

[21] G. Shavit, E. Berg, A. Stern, and Y. Oreg, Phys. Rev. Lett. **127,** 247703 (2021).

[22] Y. H. Kwan, G. Wagner, T. Soejima, M. P. Zaletel, S. H. Simon, S. A. Parameswaran, and N. Bultinck, Phys. Rev. X **11,** 041063 (2021).

[23] D. Călugăru, N. Regnault, M. Oh, K. P. Nuckolls, D. Wong, R. L. Lee, A. Yazdani, O. Vafek, and B. A. Bernevig, Phys. Rev. Lett. **129,** 117602 (2022).

[24] J. P. Hong, T. Soejima, and M. P. Zaletel, Phys. Rev. Lett. **129,** 147001 (2022).

[25] J. Yu, B. A. Foutty, Y. H. Kwan, M. E. Barber, K. Watanabe, T. Taniguchi, Z.-X. Shen, S. A. Parameswaran, and B. E. Feldman, arXiv:2206.11304.

[26] P. Stepanov, I. Das, X. Lu, A. Fahimniya, K. Watanabe, T. Taniguchi, F. H. L. Koppens, J. Lischner, L. Levitov, and D. K. Efetov, Nature (London) **583,** 375 (2020).

[27] J. H. Wilson, Y. Fu, S. Das Sarma, and J. H. Pixley, Phys. Rev. Res. **2,** 023325 (2020).

[28] A. Ramires and J. L. Lado, Phys. Rev. B **99,** 245118 (2019).

[29] A. Thomson and J. Alicea, Phys. Rev. B **103,** 125138 (2021).

[30] G. Shavit and Y. Oreg, Phys. Rev. Lett. **128,** 156801 (2022).

[31] J. Shi, J. Zhu, and A. H. MacDonald, Phys. Rev. B **103,** 075122 (2021).

[32] S. Grover, M. Bocarsly, A. Uri, P. Stepanov, G. Di Battista, I. Roy, J. Xiao, A. Y. Meltzer, Y. Myasoedov, K. Pareek, K. Watanabe, T. Taniguchi, B. Yan, A. Stern, E. Berg, D. K. Efetov, and E. Zeldov, Nat. Phys. **18,** 885 (2022).

[33] P. W. Anderson, J. Phys. Chem. Solids **11,** 26 (1959).

[34] K. Maki, in *Superconductivity*, edited by R. D. Parks (Marcel Dekker, New York, 1969), Vol. 2, p. 1035.

[35] E. I. Timmons, S. Teknowijoyo, M. Kończykowski, O. Cavani, M. A. Tanatar, S. Ghimire, K. Cho, Y. Lee, L. Ke, N. H. Jo, S. L. Bud'ko, P. C. Canfield, P. P. Orth, M. S. Scheurer, and R. Prozorov, Phys. Rev. Res. **2,** 023140 (2020).

[36] L. Andersen, A. Ramires, Z. Wang, T. Lorenz, and Y. Ando, Sci. Adv. **6,** eaay6502 (2020).

[37] L. Yu, Acta Phys. Sin. **21,** 75 (1965).

[38] H. Shiba, Prog. Theor. Phys. **40,** 435 (1968).

[39] A. I. Rusinov, Zh. Eksp. Teor. Fiz. **9,** 146 (1969) [JETP Lett. **9,** 85 (1969)], http://jetpletters.ru/ps/1658/article_25295.shtml.

[40] A. V. Balatsky, I. Vekhter, and J.-X. Zhu, Rev. Mod. Phys. **78,** 373 (2006).

[41] J. Kang and O. Vafek, Phys. Rev. X **8,** 031088 (2018).

[42] Z. Song, Z. Wang, W. Shi, G. Li, C. Fang, and B. A. Bernevig, Phys. Rev. Lett. **123,** 036401 (2019).

[43] B. A. Bernevig, Z.-D. Song, N. Regnault, and B. Lian, Phys. Rev. B **103,** 205413 (2021).

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

[45] D. E. Parker, T. Soejima, J. Hauschild, M. P. Zaletel, and N. Bultinck, Phys. Rev. Lett. **127,** 027601 (2021).

[46] L. Balents, SciPost Phys. **7,** 048 (2019).

[47] G. Wagner, Y. H. Kwan, N. Bultinck, S. H. Simon, and S. A. Parameswaran, Phys. Rev. Lett. **128,** 156401 (2022).

[48] M. Xie and A. H. MacDonald, Phys. Rev. Lett. **124,** 097601 (2020).

[49] S. Liu, E. Khalaf, J. Y. Lee, and A. Vishwanath, Phys. Rev. Res. **3,** 013033 (2021).

[50] B. Lian, Z.-D. Song, N. Regnault, D. K. Efetov, A. Yazdani, and B. A. Bernevig, Phys. Rev. B **103,** 205414 (2021).

[51] H. C. Po, L. Zou, A. Vishwanath, and T. Senthil, Phys. Rev. X **8,** 031089 (2018).

[52] G. Tarnopolsky, A. J. Kruchkov, and A. Vishwanath, Phys. Rev. Lett. **122,** 106405 (2019).

[53] G. Shavit, K. Kolář, C. Mora, F. von Oppen, and Y. Oreg, arXiv:2208.03655.