

# CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

## Reentrant superconductivity through a quantum Lifshitz transition in twisted trilayer graphene

Ethan Lake and T. Senthil

Phys. Rev. B **104**, 174505 — Published 8 November 2021

DOI: [10.1103/PhysRevB.104.174505](https://doi.org/10.1103/PhysRevB.104.174505)

# Re-entrant superconductivity through a quantum Lifshitz transition in twisted trilayer graphene

Ethan Lake and T. Senthil

*Department of Physics, Massachusetts Institute of Technology, Cambridge, MA, 02139*

A series of recent experiments has demonstrated robust superconductivity in magic-angle twisted trilayer graphene (TTG). In particular, a recent work by Cao et al. (arxiv:2103.12083) studies the behavior of the superconductor in an in-plane magnetic field and out-of-plane displacement field, finding that the superconductor is unlikely to be spin-singlet. This work also finds that at high magnetic fields and a smaller range of dopings and displacement fields, it undergoes a transition to a distinct field-induced superconducting state. Inspired by these results, we develop an understanding of the superconductivity in TTG using a combination of phenomenological reasoning and microscopic theory. We describe role that an in-plane field plays in TTG, and use this understanding to argue that the re-entrant transition may be associated with a quantum Lifshitz phase transition, with the high-field phase possessing finite-momentum pairing. We argue that the superconductor is likely to involve a superposition of singlet singlet and triplet pairing, and describe the structure of the normal state. We also draw lessons for twisted bilayer graphene (TBG), and explain the differences in the phenomenology with TTG despite their close microscopic relationship. We propose that a singlet-triplet superposition is realized in the TBG superconductor as well, and that the  $\nu = -2$  correlated insulator may be a time reversal protected  $\mathbb{Z}_2$  topological insulator obtained through spontaneous spin symmetry breaking.

## I. INTRODUCTION

A remarkable series of developments have brought two-dimensional graphene-based Moire materials to the spotlight of condensed matter physics. Many of these systems host a fascinating array of strongly correlated insulating and magnetic states [1–15], and twisted bilayer graphene (TBG) in particular displays robust unconventional superconductivity [16–20]. Recently, another Moire system, twisted trilayer graphene (TTG), was also predicted and subsequently found to exhibit robust superconductivity [21–24]. A very recent work [25] studies superconductivity in near-magic angle TTG in the presence of an in-plane magnetic field  $\mathbf{B}$  and an out-of-plane electric displacement field  $D$ , and finds some striking phenomena as we review below. In the present paper we will synthesize the existing experimental results into a coherent theoretical picture of the superconductivity in TTG. We will explain the differences with the closely related twisted bilayer graphene (TBG), and will show how comparing phenomena in TTG and TBG can give us important clues to understand both systems.

We begin with a discussion of some of the salient experimental results on superconductivity in near magic-angle TTG [23–25]. Our focus will be on the robust superconductivity found for moire lattice fillings in the range  $-3 < \nu < -2$ .

1. The ‘normal’ state out of which the superconductivity emerges has a carrier density that is set by the deviation from  $\nu = -2$ . This normal state has a Landau fan degeneracy of 2, rather than the 4-fold degeneracy naively expected on the basis of the available flavor degrees of freedom (spin and valley). Furthermore the strength of the superconductivity does not seem to correlate with the sin-

gle particle density of states inferred from indirect measurements.

2. The superconductivity occurs even without the external perturbations of the perpendicular displacement field  $D$  or the in-plane magnetic field  $B_{\parallel}$ . Weak  $D$  and  $B_{\parallel}$  have contrasting effects on the superconductivity - the former enhances it while the latter suppresses it. Strikingly Ref. [25] finds that the superconducting state in a range of fillings survives in  $B_{\parallel}$  fields far in excess of the Pauli limit, suggesting that the pairing is not spin-singlet in nature. This behavior is also notably different from that seen in TBG, where the superconductivity is suppressed by a  $B_{\parallel}$  roughly at the Pauli limit.
3. Near optimal doping and displacement field strength, the superconductivity is seen to exhibit a ‘re-entrant’ phenomenon at large magnetic fields, with a phase transition that separates two distinct superconductors occurring at a field  $B_{\text{re-entrant}} \sim 8$  T. Interestingly, while the temperature  $T_{BKT}$  at which superconductivity occurs goes to zero at  $B_{\text{re-entrant}}$ , the temperature  $T_{\text{onset}}$  at which the resistance reaches a fixed fraction of its normal-state value is smooth across the transition, and decreases monotonically with  $B$ .

In what follows, by combining the input from experiments with a Ginzburg-Landau analysis, together with the insights from simple microscopic theory, we will provide an understanding of these phenomena, and make predictions for future experiments. We will be naturally led to the conclusion that the pairing in the superconductor is neither spin singlet *nor* spin triplet, but rather a linear combination of the two. We examine the role that the  $\mathbf{B}$  and  $D$  fields play in the physics of TTG,

and highlight differences with TBG. This will enable an explanation of the striking differences in the stability of the superconductor to in-plane fields in the two systems. We propose that interpreting the re-entrant transition as arising from a Lifshitz critical point, where one or both components of the superfluid stiffness vanish and where the pairing in the high-field phase occurs at finite momentum, is particularly natural in light of the existing experimental data.

Along the way we also discuss the nature of the normal state near filling  $\nu \approx -2$ . We propose that at  $\nu = -2$  the ground state is a time reversal protected topological insulator (obtained through spontaneous spin symmetry breaking) coexisting with the massless Dirac fermions corresponding to a decoupled graphene sector.

At  $D = 0, B = 0$ , the physics of TTG is very closely related to that of TBG [21], as will be reviewed shortly. This leads us to propose that our conclusions about the nature of the pairing in the superconductor and the nature of the normal state at  $\nu = -2$  hold in TBG as well, with minor but important modifications. This allows for experiments in TTG, which in some respects are easier to interpret than those in TBG, to improve our understanding about both systems.

What we do not attempt to do in this paper is a full-blown microscopic theory of TTG superconductivity. The questions we are interested in involve very complicated strongly interacting systems where the understanding from theory is only partially complete. Thus we view our semi-phenomenological approach as the safest way of proceeding, since it allows us to avoid committing to any particular calculational schemes of specific microscopic models, both of which come with uncertainties at present.

## II. BASIC MODEL

Let us first review the basic theoretical description of TTG [21]. We will take the graphene layers to lie normal to the  $\hat{\mathbf{z}}$  axis, and will gauge-fix the in-plane vector potential so that  $\mathbf{A}_l = (2-l)\mathbf{A}$  on layer  $l = 1, 2, 3$ , where  $\mathbf{A} = \delta\mathbf{B} \times \hat{\mathbf{z}}$  with  $\delta$  the interlayer separation. The displacement field  $D$  enters the Hamiltonian through an onsite potential of  $+U/2, 0, -U/2$  on layers 1, 2, and 3 respectively, where  $U \propto 2\delta D/\epsilon_0$ . The exact proportionality constant depends on details of how the screening works, and will not be important for the present discussion.

To proceed it is most convenient to work in a basis of fields that have definite eigenvalues under the action of mirror  $M$  [21, 26], which exchanges layers 1 and 3 and which is a symmetry when  $U = \mathbf{B} = 0$ . In the basis  $(\psi_+, \psi_2, \psi_-)^T$  with  $\psi_{\pm} = \frac{1}{\sqrt{2}}(\psi_1 \pm \psi_3)$ , where  $\psi_l$  annihilate electrons on layer  $l$ , the single-particle Hamiltonian

is

$$H_0 = \begin{pmatrix} -iv_D\boldsymbol{\sigma}_v \cdot \nabla & \sqrt{2}T & V \\ \sqrt{2}T^\dagger & -iv_D\boldsymbol{\sigma}_v \cdot \nabla & 0 \\ V^\dagger & 0 & -iv_D\boldsymbol{\sigma}_v \cdot \nabla \end{pmatrix} + H_Z, \quad (1)$$

where the Zeeman energy  $H_Z = -\frac{\mu_B g}{2}\mathbf{B} \cdot \mathbf{s}$ . Here  $v_D$  is the monolayer Dirac velocity,  $T$  is the hopping matrix familiar from the bilayer problem (see appendix E),  $V \equiv U/2 - v_D\boldsymbol{\sigma}_v \cdot \mathbf{A}$ , and we have used the notation  $\boldsymbol{\sigma}_v \equiv (\tau^z\sigma^x, \sigma^y)$ , with  $\tau^a, \sigma^a, s^a$  Pauli matrices for valley, sublattice, and physical spin, respectively. For simplicity we have also ignored a rotation of the  $\boldsymbol{\sigma}_v$  on layers 1 and 3 by the twist angle, since its effects are small [27, 28]. Typical orbital magnetic energies are on the order of a few meV, with Zeeman energies smaller by a factor of  $\sim 3$ .

We note in passing that the Hamiltonian (1) is the correct starting point only when the top and bottom graphene layers are not displaced relative to one another in the xy plane. A nonzero relative displacement  $\mathbf{r}$  breaks mirror symmetry, and can have strong effects on the flat band physics [29]. However, a vanishing relative displacement is likely to be favored energetically [30], and at  $\mathbf{B} = 0$  the data of Ref. [25] appear to be mirror-symmetric to a good approximation. We will therefore set  $\mathbf{r} = 0$  in the following.

From (1), we see that in the absence of  $U$  or  $\mathbf{B}$ ,  $H_0$  decouples into a TBG Hamiltonian for the mirror-even fields  $\psi_+, \psi_2$ , which form a set of flat bands near the magic angle, together with a decoupled Dirac cone formed by the mirror-odd field  $\psi_-$  [21]. Given that  $U$  plays a crucial role in the observed superconductivity, the physics of the superconductivity is likely to be closely related to effects caused by the mirror-odd  $\psi_-$  Dirac cone.

The mirror symmetry of the  $U = \mathbf{B} = 0$  Hamiltonian has other important consequences for how  $\mathbf{B}$  affects the single-particle physics, as both  $U$  and  $\mathbf{B}_{\text{orb}}$  are odd under mirror (here  $\mathbf{B}_{\text{orb}}$  denotes the orbital field). In particular, at  $U = 0$  the spectrum at each momentum must be an even function of  $B_{\text{orb}}$ . This means that unlike in TBG, at  $U = 0$  there will be no field-induced de-pairing of intervalley Cooper pairs due to misalignment of the Fermi surfaces in the two valleys [11], as here the single-particle energy satisfies  $\varepsilon_K(\mathbf{k}) = \varepsilon_{K'}(-\mathbf{k})$  even when  $B_{\text{orb}} \neq 0$ . The fact that the leading effects are at order  $B_{\text{orb}}^2$  also means that at small  $U, B$ , the orbital effects of the magnetic field are rather weak. Diagonalizing  $H_0$  at realistic values of  $U, \mathbf{B}$ , one finds two Dirac cones at the moire  $K$  point at energies separated by an amount which increases with  $U, B$ , as well as a Dirac cone located very near the  $K'$  point. Details on the single-particle physics can be found in appendix E.

One observation from Ref. [25] is that despite the sensitivity of the observed superconductivity to  $U$  and  $\mathbf{B}$ , superconductivity near optimal doping is seen even when  $U = \mathbf{B} = 0$ . Therefore the mirror-even fermions  $\psi_+, \psi_2$  must be superconducting on their own, independent of

their coupling to the mirror-odd  $\psi_-$ .<sup>1</sup> We are therefore prompted to analyze the system by considering the Lagrangian (suppressing all indices)

$$\begin{aligned}\mathcal{L} &= \mathcal{L}_{SC}[\psi_+, \psi_2, \Delta] + \mathcal{L}_-[\psi_-, \Delta] + \mathcal{L}_\pm[\psi_+, \psi_-], \\ \mathcal{L}_- &= \psi_-^\dagger (\partial_\tau + v_D \mathbf{k} \cdot \boldsymbol{\sigma}_v) \psi_- + \lambda (\psi_- \Delta^\dagger \psi_- + \psi_-^\dagger \Delta \psi_-^\dagger) \\ &\quad - \frac{\mu B g}{2} \psi_-^\dagger \mathbf{B} \cdot \mathbf{s} \psi_- \\ \mathcal{L}_\pm &= \psi_\pm^\dagger V \psi_- + \psi_\pm^\dagger V^\dagger \psi_+, \end{aligned} \quad (2)$$

where  $\Delta$  is the superconducting order parameter and  $\mathcal{L}_{SC}[\psi_+, \psi_2, \Delta]$  is some unknown Lagrangian capturing the physics of the  $U = \mathbf{B} = 0$  superconductor, together with the Zeeman energy for the mirror-even fermions. The term proportional to  $\lambda$  arises from a mean-field decoupling of an interaction of the form  $\psi_{+/2}^\dagger \psi_{+/2}^\dagger \psi_- \psi_- + h.c.$  Such a term does not arise from a standard density-density interaction, but as it is allowed by symmetries we will keep it.

Importantly, the modifications to the superconductivity by  $U, B_{\text{orb}}$  can be reliably addressed by integrating out the  $\psi_-$  fermions perturbatively, as it is only through the mixing with these fermions that these perturbations have any effect.

### III. THE NORMAL STATE NEAR $\nu = -2$

Before attempting to obtain an effective action for  $\Delta$  by integrating out the fermions in (2), it is helpful to think about constraints on the matrix structure of  $\Delta$  by considering the likely nature of the  $D = 0, B = 0$  normal state near  $\nu = -2$ . As we will see, known properties of the superconductor will help us place nontrivial constraints on the nature of the normal state. Because this discussion takes place at  $D = 0, B = 0$ , we may temporarily ignore the mirror-odd  $\psi_-$  fermions: this is because in this limit the leading coupling of the  $\psi_-$  fermions to the mirror-even  $\psi_+, \psi_2$  fields vanishes,<sup>2</sup> rendering the problem equivalent to that of TBG [21]. This simplification is quite useful, as it allows us to draw on insights gained in the study of TBG. Aspects of the physics particular to TTG will arise when we turn on nonzero  $D, B$ , which is discussed in the following sections.

Existing experiments on TTG [23, 24] indicate that the normal state at  $\nu = -2$  is likely highly resistive (and

presumably would be insulating in the absence of the  $\psi_-$  Dirac cone). Additionally, the doped state at  $\nu = -2 - \delta$  is seen to possess a Landau level degeneracy reduced from that of the charge neutrality point by a factor of two. The simplest way to produce a ground state in accordance with these observations is to polarize a flavor, and then to open an interaction-induced gap in the remaining polarized subspace. We will be making the assumption that the superconductor pairs fields related by time reversal, implying that both  $K, K'$  valleys participate in the superconducting state. We will also be assuming (at least at small  $B, T$ ) that the superconductor inherits the flavor polarization of the  $\nu = -2 - \delta$  normal state. This implies that the normal state is not valley polarized, and in what follows we will assume that the polarized flavor is spin. Spin polarization here means that the ground state is spin-valley locked (SVL), with the spins in a given valley fixed to lie along some (valley-dependent) direction.

The assumption of spin-valley locking is reasonable from a theoretical point of view, as polarizing spin (as opposed to valley) allows the system to take advantage of the intervalley Hund's coupling. This assumption is also motivated by experiment, as SVL allows the superconductor to achieve the strong Pauli limit violation observed in Ref. [25].

The simplest way to open up a gap is to give some fermion bilinear  $c_\alpha^\dagger Q_{\alpha\beta} c_\beta$  a nonzero expectation value, where the indices  $\alpha, \beta$  run over the states in the low-energy SVL subspace. To determine what we should take for the form of  $Q$ , we will borrow results from the strong-coupling picture proposed for twisted bilayer graphene in Refs. [32, 33]. In this picture, the four flat bands in the SVL subspace can be labeled as  $|K, A, \xi\rangle, |K', B, \eta\rangle$  and  $|K', A, \eta\rangle, |K, B, \xi\rangle$ , where  $|\xi\rangle, |\eta\rangle$  denote normalized spinors determining the directions along which the spins are locked in the  $K, K'$  valleys respectively, and where  $A/B$  is the sublattice index (which is a good index for the flat bands in the strong-coupling limit [32]). The first set of two bands has Chern number  $C = \sigma^z \tau^z = +1$ , and the second set of two have  $C = -1$ . The strong-coupling analysis indicates that  $Q$  must mix the two valleys, and must be proportional to  $\sigma^y$  in sublattice space, this being the only matrix structure which allows the ordered state at  $\nu = -2$  to completely fill two of the Chern bands and minimize the kinetic and potential energy in the strong-coupling limit (see [32] for the full logic). Any  $Q$  which fulfills these properties and survives projection onto the occupied SVL subspace may be written as

$$Q = \sigma^y \begin{pmatrix} & |\eta\rangle\langle\xi| e^{i\varphi} \\ |\xi\rangle\langle\eta| e^{-i\varphi} & \end{pmatrix}, \quad (3)$$

where the matrix written out explicitly is in valley space and where the phase  $\varphi$  is arbitrary. A nonzero expectation value for  $c_\alpha^\dagger Q_{\alpha\beta} c_\beta$  breaks time reversal symmetry  $\mathcal{T} = s^y \tau^x \mathcal{K}$  unless the spins in each valley are antiferromagnetically aligned.

At zero magnetic field, the relative orientation of  $|\xi\rangle, |\eta\rangle$  is dictated by the sign of the intervalley Hund's

<sup>1</sup> One possibility is that dissipation from the  $\psi_-$  fermions promotes superconductivity in the mirror-even sector, along the lines of the mechanism in Ref. [31]. In what follows we will however assume that the dominant effects of the  $\psi_-$  fermions arise from the terms written in (2).

<sup>2</sup> In principle there are also higher-body interactions between the  $\psi_-$  fermions and the mirror even sector — in what follows we will assume that these interactions do not qualitatively change the nature of the normal state, which seems reasonable on account of the low density of the  $\psi_-$  fermions.

coupling  $J_H$ . The sign of  $J_H$  is not known a priori, and is determined by a delicate competition between various effects, e.g. Coulomb interactions and coupling to phonons [34]. However, as we will see momentarily, the fact that the superconductor seen in experiment has a nonzero BKT transition temperature at zero field means that  $J_H$  must favor antiferromagnetic alignment of the spins in the two valleys (for the case without SVL, see appendix B). For definiteness we will take the zero-field normal state to have  $|\xi\rangle = |\uparrow\rangle, |\eta\rangle = |\downarrow\rangle$ , so that

$$Q = \sigma^y (\tau^+ s^+ e^{i\varphi} + \tau^- s^- e^{-i\varphi}). \quad (4)$$

There are two important things to note about  $Q$ . The first is that it is invariant under time reversal symmetry, as the spins are aligned antiferromagnetically. Given that the zero-field state at  $\nu = -2$  contains two filled bands with Chern numbers  $\pm 1$ , this antiferromagnetic SVL state is, remarkably, therefore a topological insulator protected by the unbroken  $U(1)$  and  $\mathcal{T}$  symmetries.<sup>3</sup> The second is that  $SO(2)_s$  rotations of the spin about  $\hat{\mathbf{z}}$  act on  $Q$  in the same way as the group  $U(1)_o$  of opposite phase rotations in the two valleys. Therefore the order parameter manifold of the insulator is  $(U(1)_o \times SO(3)_s)/U(1)_{o+s} = SO(3)$ . As  $\pi_2(SO(3)) = 0$  there are no skyrmions in this state (only  $\mathbb{Z}_2$  vortices), complicating a skyrmion-driven mechanism for superconductivity along the lines of that proposed in Refs [33, 34] (although depending on the strength of  $J_H$ , a description in terms of skyrmions that exist at intermediate scales may still be useful).

#### IV. GINZBURG-LANDAU THEORY

We now turn to discussing the general features that we expect any Ginzburg-Landau (GL) free energy for the observed superconductor should satisfy.

As mentioned above, we are operating under the assumption that  $\Delta$  is off-diagonal in the valley index. Interaction effects similar to those mentioned above then mandate [33, 35, 36] that  $\Delta$  be proportional to the identity in sublattice space (so that pairing takes place between opposite Chern sectors). For the present discussion, we will further assume that over the full temperature range of interest, the spin-valley locking which likely occurs in the normal state is inherited by the superconductor (the case where this assumption is relaxed is treated in appendix B).

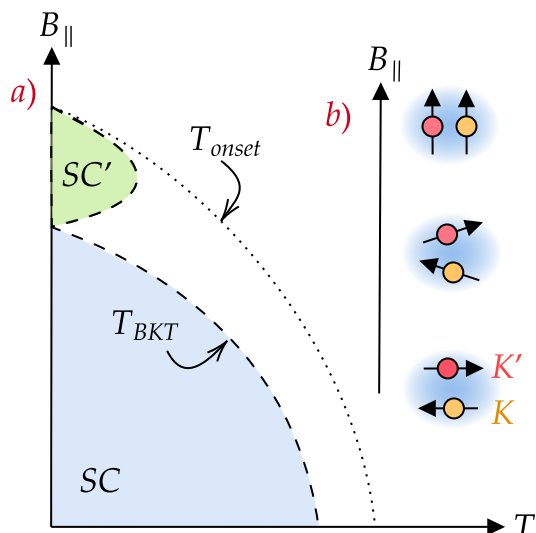


FIG. 1: a) A schematic of the phase diagram in the  $B_{\parallel}$ - $T$  plane. SC denotes the low-field superconductor, and SC' denotes the re-entrant finite-momentum pairing phase. The inset b) shows a schematic of how the pairing evolves in the field: the electrons in the Cooper pair have their spins oppositely aligned in the  $K, K'$  valleys in zero field, with the spins smoothly rotating to point along the field at large  $B_{\parallel}$ .

With these assumptions, we may write the order parameter as

$$\begin{aligned} \Delta &= \frac{|\Delta|}{2} \begin{pmatrix} -|\eta\rangle\langle\xi^*| \\ |\xi\rangle\langle\eta^*| \end{pmatrix} \\ &= \frac{|\Delta|}{2^{3/2}} (i\tau_y \mathbf{d} \cdot \mathbf{s} + \tau^x d_0) i s^y, \end{aligned} \quad (5)$$

where the matrix on the first line is in valley space, and the  $d$  four-vector has components

$$d^\mu = \frac{1}{\sqrt{2}} \langle \xi^* | i s^y s^\mu | \eta \rangle. \quad (6)$$

Note that in writing down this form for the order parameter, we have assumed that the pairing occurs in an even angular momentum channel. None of the analysis to follow is affected by this choice, and odd angular momentum pairing can be treated simply by interchanging  $\tau^x$  and  $\tau^y$  in the above expression for  $\Delta$ .

A general expression for the GL free energy in the SVL subspace can be written in terms of  $\Delta$  and the vectors  $\xi = \langle \xi | \mathbf{s} | \xi \rangle, \eta = \langle \eta | \mathbf{s} | \eta \rangle$  as

$$\begin{aligned} f &= \frac{c}{2} |\partial\Delta|^2 + r|\Delta|^2 + \frac{u}{2} |\Delta|^4 + w|\Delta|^2 \boldsymbol{\eta} \cdot \boldsymbol{\xi} \\ &\quad + \kappa |\Delta|^2 (\mathbf{B} \cdot (\boldsymbol{\xi} - \boldsymbol{\eta}))^2 - \zeta |\Delta|^2 \mathbf{B} \cdot (\boldsymbol{\xi} + \boldsymbol{\eta}), \end{aligned} \quad (7)$$

where  $-r, u, \kappa > 0$ , the coefficient  $w$  determines whether the spins in each valley prefer to align ferromagnetically or antiferromagnetically, and where all of the coefficients in the above expression contain implicit dependence on

<sup>3</sup> This should be contrasted with the analysis of the spinless model of Ref. 32 where the proposed  $K-IVC$  state is also topological, protected by  $U(1)$  and a  $\mathcal{T}$  symmetry which combines time reversal with a valley rotation. Such a symmetry is not sufficient to protect gapless edge states as the valley symmetry is broken at the edge. In contrast the spinful state discussed in this paper will have protected edge states.

$U$  and  $B$ , with this dependence determined from the Lagrangian (2) (to be discussed shortly).

Suppose first that  $w < 0$ , so that the spins in each valley prefer to align ferromagnetically along some direction  $\mathbf{n}$ , which is chosen spontaneously when  $\mathbf{B} = 0$ . The order parameter is then invariant under the combined action of  $U(1)_g$  gauge rotations and  $SO(2)_s$  spin rotations about  $\mathbf{n}$ . The order parameter manifold for the zero-field superconductor is therefore  $(U(1)_g \times SO(3)_s)/U(1)_{o+s} = SO(3)$ . Consequently, this phase does not have a finite- $T$  BKT transition at  $\mathbf{B} = 0$ . Furthermore, since this phase only has  $\mathbb{Z}_2$  vortices, the critical current in zero field vanishes [37], and when  $\zeta \neq 0$  the free energy of this state decreases linearly in the presence of a small in-plane field. All of these features are in disagreement with experiment.

When  $w > 0$  on the other hand, the spins align antiferromagnetically in the plane normal to the field direction at small  $B$ , and then gradually cant to point along the field at larger  $B$  (see Fig. 1). This canting process happens without breaking apart the Cooper pair, allowing the superconductor to survive at fields above the Pauli limit. Unlike in the  $w < 0$  case, the free energy and condensate fraction do not evolve linearly with  $B$  near  $B = 0$  (appendix B contains the details). Furthermore, since  $\Delta$  is invariant under  $SO(2)_s$  spin rotations about the ordering direction when  $\mathbf{B} = 0$ , the order parameter manifold in the zero field limit is  $(U(1)_g \times SO(3)_s)/(\mathbb{Z}_2 \times SO(2)_s) = U(1) \times S^2/\mathbb{Z}_2$ . Then there are vortices with integer-valued charge that cost logarithmic energy, leading to a nonzero  $T_{BKT}$  at zero field. All of these features are in agreement with experiment.

We are therefore forced to conclude that the zero-field superconductor favors antiferromagnetic alignment. This provides the justification for our assumption in the previous section that the normal state has antiferromagnetic spin-valley locking (see Ref. [35] for a similar argument).

It now remains to understand the mechanism producing the re-entrant phenomenon seen in experiment, as well as to understand the dependence of  $T_{\text{onset}}$  and  $T_{BKT}$  on the applied fields. This can be done by understanding the  $U, B$  dependence of the coefficients appearing in the free energy, which we now turn to discussing.

## V. QUANTUM LIFSHITZ CRITICALITY AND FINITE-MOMENTUM PAIRING

As mentioned in the introduction, a particularly noteworthy aspect of the experiment is that  $T_{\text{onset}}$  decreases more or less smoothly all the way until superconductivity is completely killed, while  $T_{BKT}$  shows non-monotonic behavior [25]. The temperature  $T_{\text{onset}}$  is set by the  $T = 0$  condensate fraction  $\langle |\Delta|^2 \rangle$ . By contrast,  $T_{BKT}$  is set by the  $T = 0$  superfluid stiffness  $K$ , viz. the coefficient of the  $(\partial\phi)^2$  term in the free energy, where  $\phi$  is the phase mode of  $\Delta$ .

A natural explanation for the observed re-entrant phenomenon is therefore that  $\rho$  decreases monotonically with  $B$  by a factor  $\propto \sqrt{1 - B^2}$ , while  $K$  decreases *faster* with  $B$ , and goes to zero when  $B = B_{\text{re-entrant}}$ , where the phase transition occurs (see the schematic phase diagram in Fig. 1.)

The IR physics of the superconductor is dominated by the phase mode  $\phi$ ,<sup>4</sup> with a Lagrangian capturing the essential physics being

$$\mathcal{L} = \frac{1}{2} \left( (\partial_t \phi)^2 + K(\partial\phi)^2 + \varsigma(\nabla^2 \phi)^2 \right), \quad (8)$$

where  $\varsigma > 0$  is included for stability. The proposed re-entrant transition is a Lifshitz point, where  $K$  vanishes.

Whether or not a Lifshitz transition can occur is determined by the field dependence of the coefficients in the GL free energy, which can be calculated by integrating out the  $\psi_-$  fermions in the Lagrangian (2). The important point here is that  $U, B$  only enter (2) in the coupling of  $\psi_+$  to  $\psi_-$ , and since  $\psi_-$  is simply a free Dirac cone, the field dependence of the terms in the free energy can be calculated in a perturbatively controlled way. We find that to quadratic order in  $U$  and  $B$ , the coupling between the  $\psi_-, \psi_+$  fermions result in the parameter  $r$  and the stiffness  $K$  appearing in the free energy (7) being renormalized by amounts

$$\begin{aligned} \delta r &= -aU^2 + bB^2 \\ \delta K &= cU^2 - dB^2, \end{aligned} \quad (9)$$

where  $a, b, c, d$  are all positive coefficients, whose detailed form is not important for the present discussion. The upshot is that (at least at small fields), the displacement field tends to promote superconductivity, while the orbital field leads to a suppression in both the condensate fraction and spin stiffness. The fact that the orbital and displacement fields have opposite effects on the superconductor (at least at small  $U, B$ ) originates from the different ways in which they couple the  $\psi_+, \psi_-$  fermions together. Importantly, this is exactly the behavior seen in experiment, with the superconductivity strengthened (suppressed) with increasing  $U$  (increasing  $B$ ), at least in the range of displacement fields for which the re-entrant transition occurs. The details of the calculation leading to (9) are standard, and are relegated to appendix A. A rough estimate of the critical in-plane field for the superconductor at  $U = 0$  in terms of the microscopic parameters which enter the coefficients appearing in (9) is given in appendix D. Our estimate is of the same order of magnitude as the critical field observed in experiment

<sup>4</sup> For nonzero field strengths which are not strong enough to fully lock the spins along  $\mathbf{B}$ , there is an additional neutral spin mode coming from  $U(1)$  spin rotations about the field direction. It however is not important for the following discussion, and will be ignored.

and we will be satisfied with this consistency check, leaving a more detailed calculation to future work.

This dependence of  $\delta r, \delta K$  on  $B^2$  is precisely of the form required for the proposed Lifshitz transition to work, with the Lifshitz transition occurring if  $d$  is large enough to drive  $K$  to zero before  $r$ . Furthermore, the fact that the condensate fraction increases with  $U$  to quadratic order provides an explanation for why, at moderate values of  $U$ , the superconductivity is seen to be stronger with increasing  $D$ . That the re-entrant transition occurs only at moderate values of  $D$  can be understood in the same way, as the Lifshitz transition can only occur if  $\langle |\Delta|^2 \rangle$  remains nonzero at the field strength where  $K$  is driven to zero. In this scenario, at small values of  $D$  it is  $\langle |\Delta|^2 \rangle$  which goes to zero first with increasing  $B$ , while at larger values of  $D$  it is  $K$  that goes to zero first.

Note that due to the rotational symmetry breaking by the in-plane field, it is possible that the stiffnesses along and normal to the field direction differ, so that

$$\mathcal{L} = \frac{1}{2}(K_{\parallel}(\partial_{\parallel}\phi)^2 + K_{\perp}(\partial_{\perp}\phi)^2) + \dots, \quad (10)$$

where  $\partial_{\parallel}$  ( $\partial_{\perp}$ ) denote derivatives along (normal to) the field direction. One could therefore consider a situation in which only one of  $K_{\parallel}, K_{\perp}$  is driven to zero at the Lifshitz transition. However, the isotropic case where  $K_{\parallel}, K_{\perp}$  go to zero simultaneously seems more likely to be realized. Indeed, consider the case where only one component vanishes at  $B_{\text{re-entrant}}$ . The critical point is an anisotropic Lifshitz model. In appendix C we show that this model orders at  $T = 0$ , and displays QLRO up to some non-universal nonzero temperature. Thus in this scenario we would *not* expect the observed  $T_{BKT}$  to go all the way to zero at  $B_{\text{re-entrant}}$  (although the temperature at which the anisotropic model loses QLRO is likely very small, and may well be below the current experimental resolution).

On the other hand consider the isotropic case, where both  $K_{\parallel}$  and  $K_{\perp}$  vanish at  $B_{\text{re-entrant}}$ . In this case the critical point has no QLRO at any temperature [38], so that  $T_{BKT} = 0$  at the critical point, as seems to be the case experimentally. Another reason for favoring the isotropic case is that to leading order in  $B$ ,  $K_{\parallel}$  and  $K_{\perp}$  are renormalized by the same amount. Our prior is therefore that near the critical point,  $K_{\perp} = K_{\parallel}$  to a good approximation. However, it may well be the case that in narrow field regimes only one of  $K_{\perp}, K_{\parallel}$  is negative, possibly explaining the rather broad nature of the transition and the observed small pockets of potentially distinct superconducting phases [25].

On the other side of the Lifshitz phase transition, the pairing occurs at finite momentum. Note that this scenario is very different in spirit to the usual FFLO mechanism [39, 40], since the latter relies on Pauli pair breaking being the dominant mechanism to suppress superconductivity. This is opposite to case of the superconductor currently under consideration, which strongly violates the

Pauli limit and in which the main effects of the field are orbital in nature. Because the present mechanism differs from that of FFLO, in order to minimize the condensation energy the order parameter will presumably live at a single nonzero wavevector  $\mathbf{q} \neq 0$ , so that that  $\Delta(\mathbf{r}) \sim e^{i\mathbf{q}\cdot\mathbf{r}}$  has a spatially uniform magnitude.

The first indications from Ref. [25] are that the re-entrant transition is first order. As written down in (8) the transition is second order, but in general it may be driven first order by other higher-order terms not included in (8) [41, 42]. One example is the term  $\lambda(\nabla\phi \cdot \nabla\phi)^2$ , which is marginally relevant if  $\lambda < 0$  [42].<sup>5</sup>

## VI. LESSONS FOR TBG

As was emphasized above, in the absence of external fields the  $\psi_{\pm}$  fermions decouple form the mirror-even sector, rendering the system nearly identical to TBG together with a decoupled Dirac fermion. In particular, this means that the physics of the superconductor and the normal state near  $\nu = -2$  are likely to be the same in both TBG and TTG. As we have argued that at  $D = 0, B = 0$  the superconductor near  $\nu = -2$  in TTG has antiferromagnetic spin-valley locking and pairing which involves a superposition of spin singlet and spin triplet, the same should be true in TBG (see also Ref. 35).

How should we reconcile this conclusion with the striking differences between TBG and TTG in the presence of a nonzero in-plane magnetic field? In TBG, the critical in-plane field in the superconducting state at optimal filling is  $B_{\parallel,c} \approx 1T$  [19], which is much smaller than in TTG, and is much closer to the Pauli limit. Since we claim that the pairing in TBG also involves a combination of spin singlet and spin triplet, the comparatively small value of  $B_{\parallel,c}$  in TBG cannot be due to Zeeman effects. Instead, the critical field is set by orbital effects. Indeed, unlike in TTG the single-particle energies in TBG do *not* satisfy  $\varepsilon_K(\mathbf{k}) = \varepsilon_{K'}(-\mathbf{k})$  in a nonzero orbital field, allowing orbital pair-breaking effects to produce a relatively lower value for  $B_{\parallel,c}$  [11]. The suppression of superconductivity at roughly the Pauli limit in an in-plane field was interpreted early on as evidence for a spin singlet superconductor. This interpretation was first called into question by the observation that the in-plane critical field is strongly angle dependent [11]. Such angle

<sup>5</sup> The cubic term  $\partial_x\phi(\partial_x\phi - \sqrt{3}\partial_y\phi)(\partial_x\phi + \sqrt{3}\partial_y\phi)$  also renders the critical point first order, but is not invariant under the  $C_6$  symmetry of the theory, which sends  $\psi(\mathbf{x}) \mapsto \tau^x \sigma^x e^{2\pi i \tau^z \sigma^z} \psi(R_{\pi/3}\mathbf{x})$  and  $\mathbf{B} \mapsto R_{\pi/3}\mathbf{B}$  [43]. Also note that nonzero  $\mathbf{B}$  permits us to write down other relevant cubic terms like  $UB^i\partial_i\phi(\partial\phi)^2$  (with the  $U$  included to preserve mirror symmetry). These terms however are not generated in the context of our model, as the effects of the orbital coupling only enter at order  $B^2$  (see appendix A).

dependence cannot arise from Zeeman effects. Within the present picture the superconductor is a superposition of singlet and triplet, its suppression by an in-plane field is entirely orbital, and the rough agreement with the Pauli limit is a coincidence coming from the rough equality of the orbital Zeeman  $g$ -factor and the spin  $g$ -factor.

Turning to the normal state of TBG, a striking consequence of the antiferromagnetic spin-valley locking is that, within the strong coupling theory, the correlated insulating state at  $\nu = -2$  is an interaction-driven  $\mathbb{Z}_2$  TI. This state will have time reversal protected edge states which will be interesting to look for in future experiments.

Finally, a puzzle which we do not resolve is that the *insulating* state at  $\nu = -2$  in TBG is also suppressed by an in-plane magnetic field [1, 18]. There are conflicting reports [9, 18, 44] on the details of how the insulating gap evolves with a field (linear in some experiments, quadratic in others), and we leave the understanding of this phenomenon to the future.

## VII. DISCUSSION AND SUMMARY

In this paper we have used a combination of phenomenological and microscopic theoretical reasoning to build a coherent picture of the superconductivity and related phenomena in TTG. We argued that the superconducting state is a superposition of spin singlet and spin triplet pairings. We showed how this explains the contrasting effects of small displacement and  $B_{\parallel}$  fields, even though both break mirror symmetry. We also showed that the mirror symmetry explains the striking difference between the stability of TTG and TBG superconductors to in-plane fields despite their close microscopic relationship. We proposed that the re-entrant superconductivity recently observed in twisted trilayer graphene [25] can be explained by a spin-valley locked superconductor which undergoes a quantum Lifshitz transition into a finite-momentum pairing state in the presence of an in-plane magnetic field. The close relationship between TBG and TTG when the displacement field and magnetic field are turned off enables insights from the bilayer system to contribute to the study of the trilayer case, and vice versa. One consequence of this is our conclusion that the normal state near  $\nu = -2$  in twisted bilayer graphene may be a  $\mathbb{Z}_2$  TI.

A possible test of our proposal for the re-entrant superconductor in TTG is to measure the Josephson current between two neighboring TTG regions maintained at different displacement fields. Due to the dependence of  $B_{\text{re-entrant}}$  on the displacement field  $D$ , it is possible to tune between the low-field and high-field phases in constant magnetic field, by varying  $D$  alone. Therefore one could imagine engineering a Josephson tunneling experiment between the low- and high-field superconducting states simply by varying the applied electric field in space. If the high-field state has finite-momentum pairing, the Josephson current should vanish, due to the

spatially oscillating phase of the high-field phase's order parameter [45]. Of course, a vanishing current could also occur by virtue of the high- and low-field phases having different pairing symmetries. However, if the pair momentum in the high-field state is a continuously varying function of  $U$  (as in our analysis), the Josephson current between two high-field superconducting regions maintained at different displacement fields should also vanish, providing a stronger test of the proposed Lifshitz scenario.

In the case which the high-field state is obtained through an (anisotropic) Lifshitz transition, one could also simply measure whether or not the critical current is direction-dependent, as near the phase transition the critical current should be highly direction-dependent (see appendix C).

Finally, the conclusion that the normal state near  $\nu = -2$  is a  $\mathbb{Z}_2$  TI is also interesting from an experimental perspective, as it will have symmetry protected gapless edge states in a single domain sample. In TTG, this TI coexists with the massless Dirac fermions of the mirror-odd sector. Complications caused by the mirror-odd Dirac cone could potentially be avoided in a setup where the sample edge is mirror symmetric and electrical contact is made only with the middle layer, enabling the transport properties of the mirror-even sector to be isolated. It may also be interesting to study the effects of induced spin orbit coupling in this system, along the lines of Ref. [46], to explicitly lock in the symmetry breaking leading to the TI sector.

The proposed  $\mathbb{Z}_2$  TI at  $\nu = -2$  may be more readily accessible to experiments in TBG, where there are no extra gapless mirror-odd fermions.

## Acknowledgements

We thank Yuan Cao, Jane Park, and Pablo Jarillo-Herrero for extensive conversations about their experiments. EL is supported by the Fannie and John Hertz Foundation and the NDSEG fellowship. TS was supported by US Department of Energy grant DE-SC0008739, and partially through a Simons Investigator Award from the Simons Foundation. This work was also partly supported by the Simons Collaboration on Ultra-Quantum Matter, which is a grant from the Simons Foundation (651440, TS).

## Appendix A: Microscopic derivation of the terms in the Ginzburg-Landau free energy

In this appendix we describe how the  $\mathbf{B}$  and  $U$  dependence of the coefficients appearing in GL free energy (B19) can be obtained microscopically, starting from the Lagrangian

$$\begin{aligned}\mathcal{L} &= \mathcal{L}_{SC}[\psi_+, \psi_2, \Delta] + \mathcal{L}_-[\psi_-, \Delta] + \mathcal{L}_\pm[\psi_+, \psi_-], \\ \mathcal{L}_-[\psi_-, \Delta] &= \psi_-^\dagger (\partial_\tau + v_D \mathbf{k} \cdot \boldsymbol{\sigma}_v) \psi_- + \lambda (\psi_- \Delta^\dagger \psi_- + \psi_-^\dagger \Delta \psi_-^\dagger) - \frac{\mu_B g}{2} \psi_-^\dagger \mathbf{B} \cdot \mathbf{s} \psi_- \\ \mathcal{L}_\pm[\psi_+, \psi_-] &= \psi_+^\dagger V \psi_- + \psi_-^\dagger V^\dagger \psi_+, \end{aligned}\tag{A1}$$

where as in the main text  $V = U/2 - v_D \mathbf{A} \cdot \boldsymbol{\sigma}_v$  (recall that  $\boldsymbol{\sigma}_v = (\tau^z \sigma^x, \sigma^y)$ ), and where  $\mathcal{L}_{SC}[\psi_+, \psi_2, \Delta]$  contains the physics of the mirror-even flat band fermions and the  $U = B = 0$  superconductor. As in the main text, we will assume even angular momentum pairing and write the order parameter as

$$\Delta = i\tau^y \mathbf{d} \cdot \mathbf{s} i s^y + \tau^x d_0 i s^y.\tag{A2}$$

The fact that  $\Delta$  is proportional to the identity in the sublattice index and off diagonal in the valley index is important, with this matrix structure being responsible for several important minus signs in what follows. However, the assumption of even angular momentum pairing is not important, and odd angular momentum pairing can be treated simply by exchanging  $\tau^x$  and  $\tau^y$  in the above expression.

In what follows we will treat the mirror-even fermions  $\psi_+, \psi_2$  by assuming that they form a flavor unpolarized Fermi liquid. Situations in which flavor polarization occurs can be easily dealt with by inserting appropriate projectors into traces. In particular, if the polarized flavor is spin, with the spins being locked to valley (which as we have argued in the main text is likely the case relevant for experiment), the effects of the polarization only enter when dealing with the Zeeman couplings, and can be dealt with in the manner described near (B8).

### 1. Orbital effects

We begin by examining the orbital effects of the magnetic field. The only place where the orbital coupling enters into the above Lagrangian is in the  $\mathcal{L}_\pm$  term, which couples the symmetric and antisymmetric sectors. This allows the effects of  $U, B$  to be treated within a controlled perturbative expansion.

Integrating out  $\psi_+$  produces the term

$$\delta\mathcal{L} = -\frac{1}{2} \begin{pmatrix} \psi_-^\dagger & \psi_- \end{pmatrix} \begin{pmatrix} V^\dagger & \\ & -V^T \end{pmatrix} \mathcal{G}_+ \begin{pmatrix} V & \\ & -V^* \end{pmatrix} \begin{pmatrix} \psi_- \\ \psi_-^\dagger \end{pmatrix},\tag{A3}$$

where  $\mathcal{G}_+$  is the  $\psi_+$  propagator. Since we are remaining largely agnostic about the physics of the symmetric flat bands, we do not know what  $\mathcal{G}_+$  looks like, other than that it involves  $\Delta$ s on the off-diagonals. In what follows we will simply assume that the most important effects of  $\delta\mathcal{L}$  can be captured by the term

$$\delta\mathcal{L} \supset h \left( \psi_-^\dagger V^\dagger \Delta V^* \psi_-^\dagger + \psi_- V^T \Delta^\dagger V \psi_- \right)\tag{A4}$$

where  $h$  is a constant. Here and below, the notation  $A \supset B$  is to be read as “ $B$  is a term appearing in  $A$ ”.

We now integrate out  $\psi_-$  and examine the structure of the resulting terms generated in the effective action  $S[\Delta]$  for  $\Delta$ . Thus the remaining task is to compute the functional determinant

$$\begin{aligned} S[\Delta] &\supset -\text{Tr} \ln \left[ \mathbf{1} + \begin{pmatrix} \frac{1}{-i\omega + \varepsilon} & \\ & \frac{1}{-i\omega - \varepsilon^T} \end{pmatrix} \begin{pmatrix} \mathcal{H} & \\ & \mathcal{H}^\dagger \end{pmatrix} \right] \\ &= \sum_{n>0} \frac{1}{2n} \text{Tr} \left[ \left( \begin{pmatrix} G_p & \\ & G_h \end{pmatrix} \begin{pmatrix} \mathcal{H} & \mathcal{H}^\dagger \end{pmatrix} \right)^{2n} \right] \\ &= \sum_{n>0} \frac{1}{n} \text{Tr} [(G_p \mathcal{H} G_h \mathcal{H}^\dagger)^{2n}], \end{aligned}\tag{A5}$$

where

$$\mathcal{H} = hV^\dagger \Delta V^* + \lambda \Delta\tag{A6}$$

and where the particle / hole Greens functions are

$$G_p = \frac{1}{-i\omega + \varepsilon}, \quad G_h = \frac{1}{-i\omega - \varepsilon^T}. \quad (\text{A7})$$

We will assume that integrating out  $\psi_+$  leaves the  $\psi_-$  fermions as continuing to be well-described by a Dirac cone, albeit one with a renormalized (possibly anisotropic) Dirac velocity and a nonzero chemical potential. Therefore we take  $\varepsilon$  to be (ignoring for now the Zeeman coupling)

$$\varepsilon = v_x k_x \tau^z \sigma^x + v_y k_y \sigma^y - \mu, \quad (\text{A8})$$

so that

$$G_p = \frac{i\omega + \mu + v_x k_x \tau^z \sigma^x + v_y k_y \sigma^y}{-(-i\omega - \mu)^2 + (v_x k_x)^2 + (v_y k_y)^2}, \quad G_h = \frac{i\omega - \mu + v_x k_x \tau^z \sigma^x - v_y k_y \sigma^y}{-(-i\omega + \mu)^2 + (v_x k_x)^2 + (v_y k_y)^2}. \quad (\text{A9})$$

In the following we will use the notation

$$\mathcal{L}[\Delta] = \sum_{n>0} \mathcal{L}_{2n}[\Delta], \quad (\text{A10})$$

with  $\mathcal{L}_{2n}$  containing  $2n$  powers of  $\Delta$ . We will first turn to examining the quadratic piece  $\mathcal{L}_2[\Delta]$ .

#### a. Quadratic terms

We will work at  $\mu = 0$  for now, as the effects of a small nonzero  $\mu$  do not qualitatively modify the structure or field dependence of the terms that are generated. We will also only focus on the behavior of  $\mathcal{L}_2[\Delta]$  in the zero-frequency limit, by taking  $\Delta = \Delta(0, \mathbf{q})$ . After re-scaling momenta by  $k_x \mapsto k_x/v_x$ ,  $k_y \mapsto k_y/v_y$ , we have

$$\mathcal{L}_2[\Delta(0, q_x/v_x, q_y/v_y)] \supset + \frac{1}{v_x v_y} T \sum_n \int_{\mathbf{k}} \text{Tr} \left[ \frac{i\omega_n + \mathbf{k} \cdot \boldsymbol{\sigma}_v}{\omega_n^2 + k^2} (hV^\dagger \Delta V^* + \lambda \Delta) \frac{i\omega_n + (\mathbf{k} + \mathbf{q}) \cdot \boldsymbol{\sigma}_v^*}{\omega_n^2 + (\mathbf{k} + \mathbf{q})^2} (hV^T \Delta^\dagger V + \lambda \Delta^\dagger) \right] \quad (\text{A11})$$

To perform the trace, we can use the explicit expression for  $V$  to write

$$\begin{aligned} \text{Tr}[\mathbf{k} \cdot \boldsymbol{\sigma}_v \Delta(\mathbf{k} + \mathbf{q}) \cdot \boldsymbol{\sigma}_v^* \Delta^\dagger] &= -\mathbf{k} \cdot (\mathbf{k} + \mathbf{q}) |\Delta|^2 \\ \text{Tr}[V^\dagger \Delta V^* \Delta^\dagger] &= \Gamma |\Delta|^2 \\ \text{Tr}[\mathbf{k} \cdot \boldsymbol{\sigma}_v V^\dagger \Delta V^* (\mathbf{k} + \mathbf{q}) \cdot \boldsymbol{\sigma}_v^* \Delta^\dagger] &= -\Gamma (\mathbf{k} + \mathbf{q}) \cdot \mathbf{k} |\Delta|^2 \\ \text{Tr}[V^\dagger \Delta V^* V^T \Delta^\dagger V] &= \Gamma^2 |\Delta|^2 \\ \text{Tr}[\mathbf{k} \cdot \boldsymbol{\sigma}_v V^\dagger \Delta V^* (\mathbf{k} + \mathbf{q}) \cdot \boldsymbol{\sigma}_v^* V^T \Delta V] &= -\Gamma^2 \mathbf{k} \cdot (\mathbf{k} + \mathbf{q}) |\Delta|^2, \end{aligned} \quad (\text{A12})$$

where in this appendix  $|\Delta|^2 = \text{Tr}[\Delta^\dagger \Delta]$ , we have defined

$$\Gamma \equiv \frac{U^2}{4} - v_D^2 A^2, \quad (\text{A13})$$

and have used  $\boldsymbol{\sigma}_v \Delta = -\Delta \boldsymbol{\sigma}_v^*$  as well as

$$\begin{aligned} \text{Tr}[\mathbf{a} \cdot \boldsymbol{\sigma} \mathbf{b} \cdot \boldsymbol{\sigma} \mathbf{c} \cdot \boldsymbol{\sigma} \mathbf{d} \cdot \boldsymbol{\sigma}] &= ((\mathbf{a} \cdot \mathbf{b})(\mathbf{c} \cdot \mathbf{d}) - (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d})) \text{Tr}[\mathbf{1}] \\ &= ((\mathbf{a} \cdot \mathbf{b})(\mathbf{c} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) + (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c})) \text{Tr}[\mathbf{1}]. \end{aligned} \quad (\text{A14})$$

The fact that  $A$  and  $U$  enter in  $\Gamma$  with opposite signs is important for the physics of the Lifshitz transition, and is a consequence of the fact that  $\Delta$  pairs opposite valleys, and hence anticommutes with the  $\tau^z$  contained in  $\boldsymbol{\sigma}_v$ . Note that this physics is independent of our assumption of even angular momentum pairing, since it only relies on the fact that  $\Delta$  is off-diagonal in the valley index.

Performing the trace, we then find

$$\mathcal{L}_2[\Delta(0, q_x/v_x, q_y/v_y)] \supset - \frac{|\Delta|^2}{v_x v_y} T \sum_n \int_{\mathbf{k}} \frac{1}{(\omega_n^2 + k^2)(\omega_n^2 + (\mathbf{k} + \mathbf{q})^2)} (\omega_n^2 (\lambda + h\Gamma)^2 + (\mathbf{k} + \mathbf{q}) \cdot \mathbf{k} (\lambda^2 + h\Gamma)^2) \quad (\text{A15})$$

Introducing a Feynman parameter  $x$  (integrated from 0 to 1),

$$\mathcal{L}_2[\Delta(0, q_x/v_x, q_y/v_y)] \supset -\frac{|\Delta|^2(\lambda + h\Gamma)^2}{v_x v_y} T \sum_n \int_{\mathbf{k}, x} \frac{\omega_n^2 + k^2 - q^2 y}{(\omega_n^2 + k^2 + q^2 y)^2} \quad (\text{A16})$$

where  $y = x(1 - x)$ .

Doing the sum over Matsubara frequencies,

$$\mathcal{L}_2[\Delta(0, q_x/v_x, q_y/v_y)] \supset -\frac{|\Delta|^2(\lambda + h\Gamma)^2}{2T v_x v_y} \int_{\mathbf{k}, x} \left( \frac{q^2 y \operatorname{sech}^2(\sqrt{k^2 + q^2 y}/2T)}{k^2 + q^2 y} + \frac{2k^2 T}{(k^2 + q^2 y)^{3/2}} \tanh(\sqrt{k^2 + q^2 y}/2T) \right). \quad (\text{A17})$$

The exact expression for this integral is not so important for the present purposes, as we are mostly just interested in the signs of the order  $q^0, q^2$  terms. We find

$$\mathcal{L}_2[\Delta(0, q_x/v_x, q_y/v_y)] \supset \frac{|\Delta|^2(\lambda + h\Gamma)^2}{2v_x v_y} (-C_1 + C_2 q^2), \quad (\text{A18})$$

where  $C_1, C_2$  are positive (temperature-dependent) coefficients. Selecting out the contribution which depends on the displacement and magnetic fields  $U$  and  $A$  by using the definition of  $\Gamma$ , and dropping terms of order  $A^4$ , we then have

$$\mathcal{L}_2[\Delta(0, \mathbf{q})] \supset \frac{2\lambda h(U^2/4 - v_D^2 A^2) + h^2(U^4/16 - U^2 v_D^2 A^2/2)}{2v_x v_y} (-C_1 |\Delta|^2 + C_2 (v_x^2 |\partial_x \Delta|^2 + v_y^2 |\partial_y \Delta|^2)). \quad (\text{A19})$$

#### b. Quartic terms

We now examine how the orbital field affects the the quartic terms  $u, v$  in the GL free energy. Using  $V^\dagger \Delta V^* = \Gamma \Delta$ , we have (now restricting to zero momentum transfer)

$$\begin{aligned} \mathcal{L}_4[\Delta] &\supset +T \sum_{\omega_n} \int_{\mathbf{k}} \frac{\operatorname{Tr}[(hV^\dagger \Delta V^* + \lambda \Delta)(i\omega_n + \mathbf{k} \cdot \boldsymbol{\sigma}_v)(hV^T \Delta^\dagger V + \lambda \Delta^\dagger)(i\omega_n + \mathbf{k} \cdot \boldsymbol{\sigma}_v^*)^2]}{(\omega_n^2 + k^2)^4} \\ &= T(h\Gamma + \lambda)^4 \sum_{\omega_n} \int_{\mathbf{k}} \frac{\operatorname{Tr}[(\Delta \Delta^\dagger)^2]}{(\omega_n^2 + k^2)^2} \\ &= (h\Gamma + \lambda)^4 \frac{\operatorname{Tr}[(\Delta \Delta^\dagger)^2]}{16\pi T} \\ &= (h\Gamma + \lambda)^4 \frac{2|\Delta|^4 - \Delta^2(\Delta^*)^2}{16\pi T}. \end{aligned} \quad (\text{A20})$$

#### c. Summary

To understand the consequences of the terms derived so far, consider the simplified free energy  $f_s = \frac{c}{2} |\partial \Delta|^2 + r |\Delta|^2 + \frac{u}{2} |\Delta|^4$ , where the phase stiffness  $K$  is related to  $c$  as  $K = |\Delta|^2 c$ . The effects derived in (A19), (A20) lead to renormalizations of  $K, r, u$  by amounts

$$\begin{aligned} \delta r &= -C_1 \frac{2\lambda h(U^2/4 - v_D^2 A^2) + h^2(U^4/16 - U^2 v_D^2 A^2/2)}{2v^2} \\ \delta K &= 2|\Delta|^2 C_2 \frac{2\lambda h(U^2/4 - v_D^2 A^2) + h^2(U^4/16 - U^2 v_D^2 A^2/2)}{2v^2} \\ \delta u &= \frac{(h(U^2/4 - v_D^2 A^2) + \lambda)^4}{4\pi T} \end{aligned} \quad (\text{A21})$$

where we have assumed  $v_x = v_y \equiv v$  for simplicity.

The condensate fraction  $\langle |\Delta|^2 \rangle = -r/u$  is renormalized by

$$\delta \langle |\Delta|^2 \rangle = -\delta r \frac{1}{u_0} + \delta u \frac{r_0}{u_0^2}, \quad (\text{A22})$$

where  $r_0, u_0$  are the values of  $r, u$  in the absence of any coupling between the mirror-even and mirror-odd fermions.

The exact field dependence of  $\delta\langle|\Delta|^2\rangle$  depends on the magnitude of  $\lambda$  and the sign of  $h\lambda$ , which are not things that can be derived microscopically within the current approach. First consider the case where  $\lambda$  is small enough so that terms of order  $\lambda h$  may be dropped compared to those of order  $h^2$ , and the contribution from the  $\delta u$  term can be ignored. This very well may be the case microscopically, as the interaction which produces the  $\lambda$  term does not appear in the (presumably dominant) density-density interaction (see the discussion following (2)). In this case we have

$$\delta\langle|\Delta|^2\rangle \approx \frac{C_1 h^2}{2v^2 u_0} (U^4/16 - U^2 v_D^2 A^2/2). \quad (\text{A23})$$

Therefore the displacement field increases the condensate fraction, while the magnetic field decreases it. The same dependence also holds for  $K$ . Note in particular that in this limit, the condensate fraction is unchanged by the magnetic field when  $U = 0$  (at least to order  $B^2$ ), implying that to this order the critical in-plane magnetic field  $B_c$  is infinite.

When  $\lambda h$  is not small compared to  $h^2$ , it is possible to engineer a situation where the behavior of the condensate fraction is reversed from the discussion above (provided that  $\lambda h < 0$ ). This however can be ruled out from experiment, given that the condensate fraction is seen to increase with  $U$  at  $B = 0$  (at least for the range of displacement fields we are interested in). Therefore in this case the leading shifts of the condensate fraction and of  $K$  are both proportional to  $U^2/4 - v_D^2 A^2$ , with positive constants of proportionality. Note that this gives a finite value of  $B_c$  even when  $U = 0$ , as in experiment. This could either be because the experimental system has a nonzero (though possibly small) value of  $\lambda$ , or because the mirror symmetry is not perfect.

Summarizing, we conclude that the displacement field and (orbital) magnetic field have *opposite* effects: the displacement field  $U$  *increases* both the condensate fraction and the superfluid stiffness, while the magnetic field *decreases* both the condensate fraction and the superfluid stiffness. Because  $K$  and  $r$  are renormalized by different amounts, it is reasonable to consider a scenario where the superfluid stiffness goes to zero with increasing  $B$  before the condensate fraction does.

## 2. Zeeman effects

We now turn to examining contributions to the effective action for  $\Delta$  arising from the Zeeman coupling of the magnetic field to the microscopic fermions. These contributions come from the term

$$\mathcal{L}_Z = -\mathbf{g} \left( \psi_+^\dagger \mathbf{B} \cdot \mathbf{s} \psi_+ + \psi_2^\dagger \mathbf{B} \cdot \mathbf{s} \psi_2 + \psi_-^\dagger \mathbf{B} \cdot \mathbf{s} \psi_- \right), \quad (\text{A24})$$

where for brevity we have defined  $\mathbf{g} \equiv \mu_B \mathbf{g}/2$ .

### a. Response from a Fermi liquid

Modeling the mirror-even fermions as a Fermi liquid as described at the beginning of this appendix, we can incorporate the Zeeman effects they induce by using the propagators

$$\begin{aligned} G_p &= \frac{i\omega - \xi_{\mathbf{k}} - \mathbf{g}\mathbf{B} \cdot \mathbf{s}}{-(-i\omega + \xi_{\mathbf{k}})^2 + \mathbf{g}^2 B^2} \\ G_h &= \frac{i\omega + \xi_{\mathbf{k}} + \mathbf{g}\mathbf{B} \cdot \mathbf{s}^*}{-(-i\omega - \xi_{\mathbf{k}})^2 + \mathbf{g}^2 B^2}, \end{aligned} \quad (\text{A25})$$

where  $\xi_{\mathbf{k}}$  is the energy measured relative to the Fermi energy. This then provides the term

$$\begin{aligned} \mathcal{L}_2[\Delta(0, \mathbf{q})] &\supset \text{Tr}[G_p \Delta G_h \Delta^\dagger] \\ &= T \sum_{\omega_n} \int_{\mathbf{k}} \frac{\text{Tr}[(i\omega_n - \xi_{\mathbf{k}} - \mathbf{g}\mathbf{B} \cdot \mathbf{s}) \Delta (i\omega_n + \xi_{\mathbf{k}+\mathbf{q}} + \mathbf{g}\mathbf{B} \cdot \mathbf{s}^*) \Delta^\dagger]}{(-(-i\omega_n + \xi_{\mathbf{k}})^2 + \mathbf{g}^2 B^2) (-(-i\omega_n - \xi_{\mathbf{k}+\mathbf{q}})^2 + \mathbf{g}^2 B^2)} \end{aligned} \quad (\text{A26})$$

At linear order in  $\mathbf{B}$ , this gives the contribution

$$\begin{aligned} \mathcal{L}_2[\Delta(0, \mathbf{q})] &\supset gT \sum_{\omega_n} \int_{\mathbf{k}} \text{Tr} \left[ \frac{\mathbf{B} \cdot \mathbf{s} \Delta \Delta^\dagger}{(-i\omega_n + \xi_{\mathbf{k}})^2 (-i\omega_n - \xi_{\mathbf{k}+\mathbf{q}})} - \frac{\Delta \mathbf{B} \cdot \mathbf{s}^* \Delta^\dagger}{(-i\omega_n + \xi_{\mathbf{k}}) (-i\omega_n - \xi_{\mathbf{k}+\mathbf{q}})^2} \right] \\ &= gT \sum_{\omega_n} \int_{\mathbf{k}} \frac{\text{Tr}[\mathbf{B} \cdot \mathbf{s} \Delta \Delta^\dagger + \Delta \mathbf{B} \cdot \mathbf{s}^* \Delta^\dagger]}{(-i\omega_n + \xi_{\mathbf{k}})^2 (-i\omega_n - \xi_{\mathbf{k}+\mathbf{q}})} \end{aligned} \quad (\text{A27})$$

where we shifted the momentum integration and used that the response will only be a function of  $q^2$ , so that  $\xi_{\mathbf{k}-\mathbf{q}}$  can be replaced with  $\xi_{\mathbf{k}+\mathbf{q}}$ . Evaluating the trace, we find

$$\mathcal{L}_1[\Delta(0, \mathbf{q})] = -8g\mathbf{B} \cdot \mathbf{M} I(q), \quad (\text{A28})$$

where as before  $\mathbf{M} = \mathbf{d}_1 \times \mathbf{d}_2$ , and where

$$I(q) = T \sum_{\omega_n} \int_{\mathbf{k}} \frac{1}{(-i\omega_n + \xi_{\mathbf{k}})^2 (-i\omega_n - \xi_{\mathbf{k}+\mathbf{q}})} \quad (\text{A29})$$

For simplicity we will only explicitly evaluate the  $O(q^0)$  term. In the limit where we can replace  $\int_{\mathbf{k}}$  with  $N(0) \int_{-\Lambda}^{\Lambda} d\xi$ , this term vanishes due to particle-hole symmetry about the Fermi level. The leading contribution is thus proportional to  $N'(0)$  [47], giving

$$I(0) \approx -\frac{N'(0)}{4} \int \frac{d\xi}{\xi} \tanh(\xi/2T) \approx \frac{N'(0)}{4} \ln\left(\frac{\Lambda}{T}\right). \quad (\text{A30})$$

Due to the smallness of  $N'(0)$  in a generic Fermi liquid scenario, this term would normally be neglected. However, it is possible that the displacement field modifies the density of states significantly enough to make this term important at moderate  $U$ .

Before moving on, we note that the above  $\mathbf{B} \cdot \mathbf{M}$  term was derived by performing an expansion in powers of  $\Delta$ , using the normal state form of the fermion propagators. In the limit where  $T \ll \Delta$ , we should instead perform an expansion about the  $B = 0$  superconductor. In this case the  $\mathbf{B} \cdot \mathbf{M}$  term simply vanishes, as there is no magnetization for the field to couple to.

Now we look at effects from (A26) which enter at order  $B^2$ . Using

$$\text{Tr}[\mathbf{B} \cdot \mathbf{s} \Delta \mathbf{B} \cdot \mathbf{s}^* \Delta^\dagger] = -2|\mathbf{B} \cdot \mathbf{d}|^2 + B^2(|\mathbf{d}|^2 - |d_0|^2), \quad (\text{A31})$$

the contribution at order  $B^2$  is

$$\begin{aligned} \mathcal{L}_2[\Delta(0, \mathbf{q})] &\supset g^2 T \sum_{\omega_n} \int_{\mathbf{k}} \left[ \frac{2|\mathbf{B} \cdot \mathbf{d}|^2 - B^2(|\mathbf{d}|^2 - |d_0|^2)}{(-i\omega_n + \xi_{\mathbf{k}})^2 (-i\omega_n - \xi_{\mathbf{k}+\mathbf{q}})^2} \right. \\ &\quad \left. + \frac{B^2|\Delta|^2}{(-i\omega_n + \xi_{\mathbf{k}}) (-i\omega_n - \xi_{\mathbf{k}+\mathbf{q}})} \left( \frac{1}{(-i\omega_n + \xi_{\mathbf{k}})^2} + \frac{1}{(-i\omega_n - \xi_{\mathbf{k}+\mathbf{q}})^2} \right) \right] \\ &= g^2 T \sum_{\omega_n} \int_{\mathbf{k}} \left[ \frac{2|\mathbf{B} \cdot \mathbf{d}|^2 - B^2(|\mathbf{d}|^2 - |d_0|^2)}{(-i\omega_n + \xi_{\mathbf{k}})^2 (-i\omega_n - \xi_{\mathbf{k}+\mathbf{q}})^2} + \frac{2B^2|\Delta|^2}{(-i\omega_n + \xi_{\mathbf{k}})^3 (-i\omega_n - \xi_{\mathbf{k}+\mathbf{q}})} \right] \end{aligned} \quad (\text{A32})$$

where the second term in the first line comes from the expansion of the  $B^2$ s in the denominators of the propagators, and we have again used that the response is only a function of  $q^2$  when going to the second line.

The general expressions that appear after doing the frequency sum are rather complicated. For our purposes it suffices to expand in  $q^2$ , and write

$$\mathcal{L}_2[\Delta(0, \mathbf{q})] \supset g^2 (|\mathbf{B} \cdot \mathbf{d}|^2 (F_1 - F_2 q^2) + B^2 |\mathbf{d}|^2 (G_1 - G_2 q^2) + B^2 |d_0|^2 (H_1 - H_2 q^2)), \quad (\text{A33})$$

where the  $F_i, G_i, H_i$  are all positive functions of  $T$ , with all but  $F_1, H_1$  vanishing in the limit where we may perform the  $\mathbf{k}$  integrals as  $N(0) \int_{-\infty}^{\infty} d\xi$ .

*b. Response from a Dirac cone*

We now compute the Zeeman effects induced by the mirror-odd Dirac fermion  $\psi_-$  (these are expected to be smaller, but we include the discussion for completeness). To do this, we need to compute the propagator for  $\psi_-$  in the absence of hybridization with  $\Delta$ . To this end, consider the matrix

$$M = a\mathbf{1} + \mathbf{b} \cdot \mathbf{s} + \mathbf{c} \cdot \boldsymbol{\sigma}_v. \quad (\text{A34})$$

One can show that the inverse is

$$M^{-1} = \frac{a(a^2 - b^2 - c^2) + (b^2 - a^2 - c^2)\mathbf{b} \cdot \mathbf{s} + (c^2 - a^2 - b^2)\mathbf{c} \cdot \boldsymbol{\sigma}_v + 2a\mathbf{b} \cdot \mathbf{s} \otimes \mathbf{c} \cdot \boldsymbol{\sigma}_v}{a^2(a^2 - b^2 - c^2) + b^2(b^2 - a^2 - c^2) + c^2(c^2 - a^2 - b^2)}. \quad (\text{A35})$$

If we apply this to the case of  $a = -i\omega - \mu$ ,  $\mathbf{b} = -g\mathbf{B}$ ,  $\mathbf{c} = \mathbf{k}$ , we expand to order  $B^2$  and get

$$\begin{aligned} G_p &= \frac{i\omega_- + \mathbf{k} \cdot \boldsymbol{\sigma}_v}{\omega_-^2 + k^2} + g\mathbf{B} \cdot \mathbf{s} \frac{-\omega_-^2 + k^2 + 2i\omega_- \mathbf{k} \cdot \boldsymbol{\sigma}_v}{(\omega_-^2 + k^2)^2} - g^2 B^2 \frac{i\omega_-^3 - 3i\omega_- k^2 + (3\omega_-^2 - k^2)\mathbf{k} \cdot \boldsymbol{\sigma}_v}{(\omega_-^2 + k^2)^3} \\ G_h &= \frac{i\omega_+ + \mathbf{k} \cdot \boldsymbol{\sigma}_v^*}{\omega_+^2 + k^2} - g\mathbf{B} \cdot \mathbf{s}^* \frac{-\omega_+^2 + k^2 + 2i\omega_+ \mathbf{k} \cdot \boldsymbol{\sigma}_v^*}{(\omega_+^2 + k^2)^2} - g^2 B^2 \frac{i\omega_+^3 - 3i\omega_+ k^2 + (3\omega_+^2 - k^2)\mathbf{k} \cdot \boldsymbol{\sigma}_v^*}{(\omega_+^2 + k^2)^3}, \end{aligned} \quad (\text{A36})$$

where  $\omega_{\pm} = \omega \pm i\mu$ . We then use these propagators to integrate out  $\psi_-$  and compute the resulting terms in the effective action for  $\Delta$ , following the same approach as when considering orbital effects.

The contribution linear in  $g\mathbf{B}$  vanishes if  $\mu = 0$  due to particle-hole symmetry, but is nonzero otherwise. To leading (linear) order in  $\mu$ ,

$$\begin{aligned} \mathcal{L}_2[\Delta(0, \mathbf{q})] &= g(\lambda + h\Gamma)^2 T \sum_{\omega_n} \int_{\mathbf{k}} \left[ \frac{\text{Tr}[\mathbf{B} \cdot \mathbf{s}(-\omega_-^2 + k^2 + 2i\omega_- \mathbf{k} \cdot \boldsymbol{\sigma}_v)\Delta(i\omega_+ + \mathbf{k} \cdot \boldsymbol{\sigma}_v^*)\Delta^\dagger]}{(\omega_-^2 + k^2)^2(\omega_+^2 + k^2)} \right. \\ &\quad \left. - \frac{\text{Tr}[(i\omega_- + \mathbf{k} \cdot \boldsymbol{\sigma}_v)\Delta\mathbf{B} \cdot \mathbf{s}^*(-\omega_+^2 + k^2 + 2i\omega_+ \mathbf{k} \cdot \boldsymbol{\sigma}_v^*)\Delta^\dagger]}{(\omega_-^2 + k^2)(\omega_+^2 + k^2)^2} \right] \\ &= -L\mu(\lambda + h\Gamma)^2 g \text{Tr}[\mathbf{B} \cdot \mathbf{s}\Delta\Delta^\dagger + \Delta\mathbf{B} \cdot \mathbf{s}^*\Delta^\dagger] \\ &= 4L\mu(\lambda + h\Gamma)^2 g\mathbf{B} \cdot \mathbf{M}, \end{aligned} \quad (\text{A37})$$

where  $L$  is a positive constant.

Contributions at order  $(gB)^2$  give the term (now dropping the chemical potential for simplicity)

$$\begin{aligned} \mathcal{L}_2[\Delta(0, \mathbf{q})] &\supset -g^2(\lambda + h\Gamma)^2 T \sum_{\omega_n} \int_{\mathbf{k}} \left[ \frac{\text{Tr}[\mathbf{B} \cdot \mathbf{s}(-\omega_n^2 + (\mathbf{k} + \mathbf{q})^2 + 2i\omega_n(\mathbf{k} + \mathbf{q}) \cdot \boldsymbol{\sigma}_v)\Delta\mathbf{B} \cdot \mathbf{s}^*(-\omega_n^2 + k^2 + 2i\omega_n \mathbf{k} \cdot \boldsymbol{\sigma}_v^*)\Delta^\dagger]}{(\omega_n^2 + (\mathbf{k} + \mathbf{q})^2)^2(\omega_n^2 + k^2)^2} \right. \\ &\quad \left. + 2B^2 \frac{\text{Tr}[(i\omega_n + (\mathbf{k} + \mathbf{q}) \cdot \boldsymbol{\sigma}_v)\Delta(i\omega_n^3 - 3i\omega_n k^2 + (3\omega_n^2 - k^2)\mathbf{k} \cdot \boldsymbol{\sigma}_v^*)\Delta^\dagger]}{(\omega_n^2 + (\mathbf{k} + \mathbf{q})^2)(\omega_n^2 + k^2)^3} \right] \\ &= -g^2(\lambda + h\Gamma)^2 T \sum_{\omega_n} \int_{\mathbf{k}} \left[ \frac{-2|\mathbf{B} \cdot \mathbf{d}|^2 + B^2(|\mathbf{d}|^2 - |d_0|^2)}{(\omega_n^2 + (\mathbf{k} + \mathbf{q})^2)^2(\omega_n^2 + k^2)^2} ((-\omega_n^2 - (\mathbf{k} + \mathbf{q})^2)(-\omega_n^2 + k^2) + 4\omega_n^2(\mathbf{k} + \mathbf{q}) \cdot \mathbf{k}) \right. \\ &\quad \left. - 2B^2|\Delta|^2 \left( \frac{\omega_n^4 - 3k^2\omega_n^2 + (\mathbf{k} + \mathbf{q}) \cdot \mathbf{k}(3\omega_n^2 - k^2)}{(\omega_n^2 + (\mathbf{k} + \mathbf{q})^2)(\omega_n^2 + k^2)^3} \right) \right] \end{aligned} \quad (\text{A38})$$

Evaluating the integrals gives

$$\mathcal{L}_2[\Delta(0, \mathbf{q})] \supset (\lambda + h\Gamma)^2 g^2 (|\mathbf{B} \cdot \mathbf{d}|^2(P_1 - P_2q^2) - B^2|\mathbf{d}|^2(Q_1 - Q_2q^2) + B^2|d_0|^2(R_1 - R_2q^2)), \quad (\text{A39})$$

where as usual  $P_i, Q_i, R_i$  are positive (temperature dependent) constants.

### Appendix B: Details on the Ginzburg-Landau theory

In this section we detail the phases described by the Ginzburg-Landau theory discussed in the main text. See [48] for a similar analysis in a more general context.

### 1. Fully spin-valley locked limit

We will first consider the case where the spin is fully locked to valley, so that the order parameter is allowed to be nonzero only in a subset of the Hilbert space.

As was discussed above near (B22), in the case of full SVL, we have  $\det(\Delta_{KK'}) = 0$  and may write the order parameter as

$$\Delta = \frac{|\Delta|}{2} \begin{pmatrix} & -|\eta\rangle\langle\xi^*| \\ |\xi\rangle\langle\eta^*| & \end{pmatrix}, \quad (\text{B1})$$

where as before  $|\Delta| = \sqrt{\text{Tr}[\Delta^\dagger\Delta]}$ , the matrix written out explicitly is in valley space, and where  $\xi, \eta$  are normalized spinors which determine the directions along which the spins in the  $K, K'$  valleys are polarized.

By decomposing the outer products appearing in  $\Delta$  in terms of Pauli matrices, we may write (assuming even angular momentum pairing for concreteness, as in the main text)

$$\begin{aligned} \Delta &= \frac{|\Delta|}{4} (i\tau^y(\mathbf{s} \cdot i\mathbf{s}^y) \cdot \langle\bar{\xi}|\mathbf{s}|\eta\rangle + \tau^x i\mathbf{s}^y \langle\bar{\xi}|\eta\rangle) \\ &= \frac{|\Delta|}{2^{3/2}} (i\tau_y \mathbf{d} \cdot \mathbf{s} + \tau^x d_0) i\mathbf{s}^y, \end{aligned} \quad (\text{B2})$$

where the  $d$  four-vector is *defined* in terms of  $|\xi\rangle, |\eta\rangle$ , and has components

$$d_\mu = \frac{1}{\sqrt{2}} \langle\bar{\xi}|s^\mu|\eta\rangle, \quad (\text{B3})$$

where we have defined the adjoint spinor<sup>6</sup>

$$|\bar{\xi}\rangle = -i\mathbf{s}^y |\xi\rangle^*. \quad (\text{B5})$$

In what follows it will be helpful to define, for any spinor  $|\varsigma\rangle$ , the three-vector

$$\boldsymbol{\varsigma} = \langle\varsigma|\mathbf{s}|\varsigma\rangle \implies |\varsigma\rangle\langle\varsigma| = \frac{1}{2}(1 + \boldsymbol{\varsigma} \cdot \mathbf{s}). \quad (\text{B6})$$

Taking the adjoint in the manner above simply reverses the direction of the 3-vector, so that  $\bar{\boldsymbol{\varsigma}} = -\boldsymbol{\varsigma}$ . This notation lets us write the magnitudes of  $\mathbf{d}$  and  $d_0$  as

$$\begin{aligned} |\mathbf{d}|^2 &= \frac{1}{2} \langle\bar{\xi}|s^j|\eta\rangle\langle\eta|s^j|\bar{\xi}\rangle \\ &= \frac{1}{4}(3 + \boldsymbol{\eta} \cdot \boldsymbol{\xi}), \\ |d_0|^2 &= \frac{1}{4} |\langle\bar{\xi}|\eta\rangle|^2 \\ &= \frac{1}{4}(1 - \boldsymbol{\xi} \cdot \boldsymbol{\eta}). \end{aligned} \quad (\text{B7})$$

The norm of the full  $d$  vector is thus  $|d_\mu|^2 = 1$ , giving  $\text{Tr}[\Delta^\dagger\Delta] = |\Delta|^2$  as required. Also as expected for a spin-polarized state, we see that the weight of the triplet part  $|\mathbf{d}|^2$  is always nonzero, being minimal for antiferromagnetic alignment between the valleys and maximal for ferromagnetic alignment, with the opposite being true for the weight of the singlet part  $|d_0|^2$ .

---

<sup>6</sup> The adjoint of a given spinor is obtained by inverting the spinor through the Bloch sphere and changing the handedness used to define its overall phase. In terms of the Euler angle parametrization

$$|\xi\rangle = e^{i\gamma/2} \begin{pmatrix} e^{i\phi/2} \cos(\theta/2) \\ e^{-i\phi/2} \sin(\theta/2) \end{pmatrix}, \quad (\text{B4})$$

the adjoint  $|\bar{\xi}\rangle$  can be obtained by sending  $\gamma \mapsto -\gamma + \pi$ ,  $\phi \mapsto \phi + \pi$ ,  $\theta \mapsto -\theta + \pi$ .

When deriving the coefficients in the GL free energy microscopically, SVL can be taken into account simply by replacing the particle and hole propagators by

$$G_p \mapsto \mathcal{P}G_p\mathcal{P}, \quad G_h \mapsto \mathcal{P}^T G_h \mathcal{P}^T, \quad (\text{B8})$$

where we have defined the projector

$$\mathcal{P} = |\eta\rangle \langle \eta| \oplus |\xi\rangle \langle \xi|, \quad (\text{B9})$$

with the  $\oplus$  in valley space. Note that  $\mathcal{P}$  satisfies

$$\mathcal{P}\Delta = \Delta, \quad \Delta\mathcal{P}^T = \Delta. \quad (\text{B10})$$

The form of terms not involving the Zeeman coupling of the magnetic field are not modified by the projection. For terms involving the Zeeman coupling, we simply make the replacement

$$\mathbf{B} \cdot \mathbf{s} \mapsto \mathcal{P}\mathbf{B} \cdot \mathbf{s}\mathcal{P} = \mathbf{B} \cdot \begin{pmatrix} \boldsymbol{\eta} |\eta\rangle \langle \eta| \\ \boldsymbol{\xi} |\xi\rangle \langle \xi| \end{pmatrix} \equiv \mathcal{P}_B. \quad (\text{B11})$$

The structure of the Zeeman terms quadratic in the magnetic field are determined by following the analysis of appendix A, and using the traces

$$\begin{aligned} \text{Tr}[\mathcal{P}_B \Delta \Delta^\dagger + \Delta \mathcal{P}_B^* \Delta^\dagger] &= 2|\Delta|^2 \mathbf{B} \cdot (\boldsymbol{\eta} + \boldsymbol{\xi}), \\ \text{Tr}[\Delta^\dagger \mathcal{P}_B^2 \Delta + \Delta^\dagger \Delta (\mathcal{P}_B^*)^2] &= 2|\Delta|^2 ((\mathbf{B} \cdot \boldsymbol{\eta})^2 + (\mathbf{B} \cdot \boldsymbol{\xi})^2) \\ \text{Tr}[\Delta^\dagger \mathcal{P}_B \Delta \mathcal{P}_B^*] &= 2|\Delta|^2 (\mathbf{B} \cdot \boldsymbol{\eta})(\mathbf{B} \cdot \boldsymbol{\xi}). \end{aligned} \quad (\text{B12})$$

For example, for a spin-valley locked Fermi liquid, the contributions to the free energy from the Zeeman couplings are, to quadratic order in  $\Delta$  and  $B$ ,

$$\zeta |\Delta|^2 \mathbf{B} \cdot (\boldsymbol{\xi} + \boldsymbol{\eta}) + \kappa |\Delta|^2 (\mathbf{B} \cdot (\boldsymbol{\xi} - \boldsymbol{\eta}))^2, \quad (\text{B13})$$

where  $\zeta$  is small in  $N'(0)/N(0)$  and  $\kappa > 0$  (see appendix A). As expected, we may avoid Pauli suppression of the pairing either by having the spins in each valley align in the plane normal to  $\mathbf{B}$ , or by having the two spins ferromagnetically align along  $\mathbf{B}$ , both of which prevent pair-breaking by the field from occurring.

The free energy in general may be written as a function of  $|\Delta|, \boldsymbol{\xi}, \boldsymbol{\eta}$  as (ignoring derivatives of  $\boldsymbol{\xi}, \boldsymbol{\eta}$  for simplicity)

$$f = K|\partial\Delta|^2 + r|\Delta|^2 + \frac{u}{2}|\Delta|^4 + w|\Delta|^2 \boldsymbol{\eta} \cdot \boldsymbol{\xi} + \kappa |\Delta|^2 (\mathbf{B} \cdot (\boldsymbol{\xi} - \boldsymbol{\eta}))^2 - \zeta |\Delta|^2 \mathbf{B} \cdot (\boldsymbol{\xi} + \boldsymbol{\eta}) \quad (\text{B14})$$

where the coefficient  $w$  determines whether the spins in each valley prefer to align ferromagnetically or antiferromagnetically. The explicit  $\mathbf{B}$  dependence arises from the Zeeman coupling terms, while the orbital effects of the field provide implicit dependence of the various coefficients on  $B^2$  in the manner described in appendix A.

As discussed in the main text, the system observed in experiments likely has  $w > 0$ . When  $B = 0$ , the minimum of  $f$  occurs for  $\boldsymbol{\xi} = -\boldsymbol{\eta}$ .  $\Delta$  is now invariant under  $SO(2)$  rotations about  $\boldsymbol{\xi}$ , and hence the order parameter manifold is  $(U(1) \times S^2)/\mathbb{Z}_2$ .

Now we will examine what happens to this state at nonzero  $\mathbf{B}$ . Taking  $\mathbf{B}$  to lie along  $\hat{\mathbf{z}}$  in spin space, upon minimizing  $f$  we find

$$\boldsymbol{\xi} = R_z(\varphi) \begin{pmatrix} 0 \\ \sin \theta \\ \cos \theta \end{pmatrix}, \quad \boldsymbol{\eta} = R_z(\varphi) \begin{pmatrix} 0 \\ -\sin \theta \\ \cos \theta \end{pmatrix}, \quad (\text{B15})$$

where  $R_z(\varphi)$  represents a rotation by an arbitrary angle  $\varphi$  about  $\hat{\mathbf{z}}$ , and where

$$\theta = \begin{cases} \arccos(|\zeta B|/2w), & |\zeta B|/2w \leq 1 \\ 0, & |\zeta B|/2w > 1 \end{cases}. \quad (\text{B16})$$

Thus for  $0 < |\zeta B|/2w < 1$  the spins are partially magnetized and there is an  $U(1)$ s worth of degenerate minima, and as in the partially polarized phase of the following subsection the order parameter manifold is  $(U(1) \times U(1))/\mathbb{Z}_2$ . For

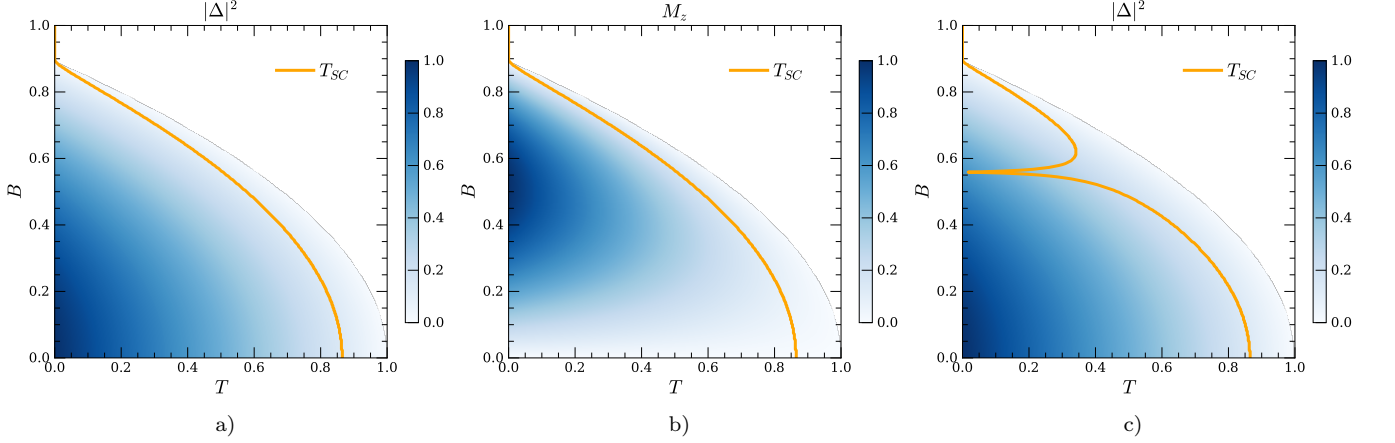


FIG. 2: a) condensate fraction in the spin-valley locked case (normalized by  $|\Delta(B=0, T=0)|^2$ ) for a typical choice of parameters, and ignoring dependence of  $K$  on  $B$ .  $T_{SC}$  marks the BKT transition temperature for the superconductor. b) expected value of the component of the magnetization along the field direction, normalized to  $M_z(B=0, T=0)$ . c) the same as a), but taking into account a downward renormalization of  $K$  by a term proportional to  $B^2$ , leading to a field-induced Lifshitz transition. Here we have taken  $\zeta < 2w/B_c$ , where  $B_c$  is the critical field for the superconductor; as such the system never reaches the fully magnetized regime.

$|\zeta B|/2w > 1$  the spins are ferromagnetically aligned and fully magnetized, and the order parameter manifold is  $U(1)$ . Given this solution for  $\boldsymbol{\xi}, \boldsymbol{\eta}$ , the magnitude of the order parameter is consequently determined to be

$$|\Delta| = \sqrt{-\tilde{r}/u}, \quad \tilde{r} = r + w(2 \cos^2(\theta) - 1) - 2|\zeta B| \cos(\theta). \quad (\text{B17})$$

The condensate fraction and its first derivative with respect to  $B$  are continuous across the transition into the high-field ferromagnetic state, and  $(\partial(|\Delta|)/\partial B)|_{B=0} = 0$ , implying that unlike in the non-locked case to be discussed in the next subsection,  $T_{SC}$  will not increase linearly with  $B$  at small  $B$ .

Figure 2 shows the condensate fraction (left panel) and magnetization (center panel) for a typical choice of parameters (here  $\zeta$  is small enough so that the system never reaches the fully magnetized regime). The right panel is the same as the left, but shows the BKT temperature after taking into account orbital effects which suppress  $K$ .

## 2. No spin-valley locking

Now we consider the slightly simpler case in which there is no spin-valley locking (SVL), with the order parameter allowed to have components in the entire fermionic Hilbert space. This discussion will therefore only be germane for the problem considered in this paper at values of  $T$  and / or  $B$  large enough to eliminate any SVL which occurs in the  $T=0$  normal state near  $\nu = -2$ .

We start by breaking the order parameter up into triplet and singlet parts as

$$\Delta = i\tau^y \mathbf{d} \cdot \mathbf{s} i s^y + \tau^x d_0 i s^y, \quad (\text{B18})$$

where  $d_\mu = (d_0, \mathbf{d})$  is a complex four-vector. Up to quartic order in the order parameter, the free energy can be written as

$$\begin{aligned} f &= f_t + f_s + f_{ts} \\ f_t &= \sum_{i=1,2} \left( K(\partial \mathbf{d}_i)^2 + r d_i^2 + \frac{u}{2} d_i^4 + \kappa(\mathbf{B} \cdot \mathbf{d}_i)^2 \right) + (u+v) d_1^2 d_2^2 - v(\mathbf{d}_1 \cdot \mathbf{d}_2)^2 + \zeta \mathbf{B} \cdot (\mathbf{d}_1 \times \mathbf{d}_2) \\ f_s &= K'(\partial d_0)^2 + r'|d_0|^2 + \frac{u'}{2}|d_0|^2 + \kappa' B^2 |d_0|^2 \\ f_{ts} &= q|d_0|^2(d_1^2 + d_2^2) + w|d_0^2 - \mathbf{d}^2|^2 \end{aligned} \quad (\text{B19})$$

where we have used the standard notation  $\mathbf{d} = \mathbf{d}_1 + i\mathbf{d}_2$ . As usual, the terms which depend explicitly on the magnetic field arise from Zeeman effects, while orbital effects are responsible for giving the remaining coefficients implicit dependence on  $B^2$ .

When minimizing  $f$  in the presence of the couplings to the magnetic field, it is often helpful to write the  $\mathbf{d}$  vector in terms of components which diagonalize the  $\mathbf{B} \cdot (\mathbf{d}_1 \times \mathbf{d}_2)$  coupling [49]. Taking the magnetic field to point in the  $\hat{\mathbf{z}}$  direction in spin space, this is done through the unitary transformation

$$\begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \\ \sqrt{2} & \end{pmatrix} \begin{pmatrix} d_+ \\ d_z \\ d_- \end{pmatrix}, \quad (\text{B20})$$

In this representation, the  $f_t$  and  $f_{ts}$  components of the free energy are now

$$\begin{aligned} f_t &= K \sum_{\sigma=+,-,z} (|\nabla d_\sigma|^2 + r|d_\sigma|^2) + \frac{u+v/2}{2} (|d_+|^2 + |d_-|^2 + |d_z|^2)^2 - v|d_+d_- + d_z^2/2|^2 \\ &\quad + \kappa B^2 |d_z|^2 - \frac{\zeta B}{2} (|d_+|^2 - |d_-|^2) \\ f_{ts} &= q|d_0|^2 (|d_+|^2 + |d_-|^2 + |d_z|^2) + p|d_0^2 - (2d_+d_- + d_z^2)|^2. \end{aligned} \quad (\text{B21})$$

The effects of the  $p$  term can be understood as a way of softly implementing SVL. Indeed, if we write the order parameter as

$$\Delta = \begin{pmatrix} & -\Delta_{KK'}^T \\ \Delta_{KK'} & \end{pmatrix}, \quad (\text{B22})$$

(where the  $2 \times 2$  matrix is in valley space) we see that we may write

$$|d_0^2 - \mathbf{d}^2|^2 = |d_0^2 - (d_z^2 + 2d_+d_-)|^2 = |\det(\Delta_{KK'})|^2. \quad (\text{B23})$$

Therefore large positive  $p$  enforces that the determinant of  $\Delta_{KK'}$  vanish. In this case we may write

$$\Delta_{KK'} \propto |\xi\rangle\langle\eta^*| \quad (\text{B24})$$

for two spinors  $|\xi\rangle, |\eta\rangle$ , which have the interpretation as the directions along which the spins in the  $K, K'$  valleys are polarized. For simplicity of the analytical results, and because the fully spin-valley locked case was discussed in the last subsection, in what follows we will simply set  $p = 0$ .

Extremizing the free energy for uniform solutions then gives

$$\begin{aligned} rd_+ + (u+v/2)d_+d^2 - vd_-^*(d_+d_- + d_z^2/2) - \zeta Bd_+/2 + q|d_0|^2d_+ &= 0 \\ rd_- + (u+v/2)d_-d^2 - vd_+^*(d_+d_- + d_z^2/2) + \zeta Bd_-/2 + q|d_0|^2d_- &= 0 \\ rd_z + (u+v/2)d_zd^2 + vd_z^*(d_+d_- + d_z^2/2) + \kappa B^2d_z + q|d_0|^2d_z &= 0 \\ (r' + \kappa' B^2 + u'|d_0|^2 + q(|d_+|^2 + |d_-|^2 + |d_z|^2))d_0 &= 0 \end{aligned} \quad (\text{B25})$$

One can show that when  $\kappa > 0$ , which is the case we are interested in, when  $B \neq 0$  the global minima of the free energy always have  $d_z = 0$ , while when  $B = 0$  we have  $SO(3)$  spin symmetry and may fix  $d_z = 0$  without loss of generality. Setting  $d_z = 0$  then, when  $d_0 \neq 0$  we just have to solve

$$\begin{aligned} \tilde{r}d_+ + (\tilde{u} + v/2)d_+d^2 - vd_+|d_-|^2 - \zeta Bd_+/2 &= 0 \\ \tilde{r}d_- + (\tilde{u} + v/2)d_-d^2 - vd_-|d_+|^2 + \zeta Bd_-/2 &= 0 \end{aligned} \quad (\text{B26})$$

where

$$\tilde{r} = r - \frac{q}{u'}(r' + \kappa' B^2), \quad \tilde{u} = u - q^2/u' \quad (\text{B27})$$

and<sup>7</sup>

$$|d_0|^2 = \frac{-r' - \kappa' B^2 - q(|d_+|^2 + |d_-|^2)}{u'}. \quad (\text{B28})$$

<sup>7</sup> When the RHS of (B28) becomes negative, the minimum of  $f$  has  $d_0 = 0$ . The solutions in this case are obtained simply by replacing  $\tilde{r}, \tilde{u}$  with  $r, u$ .

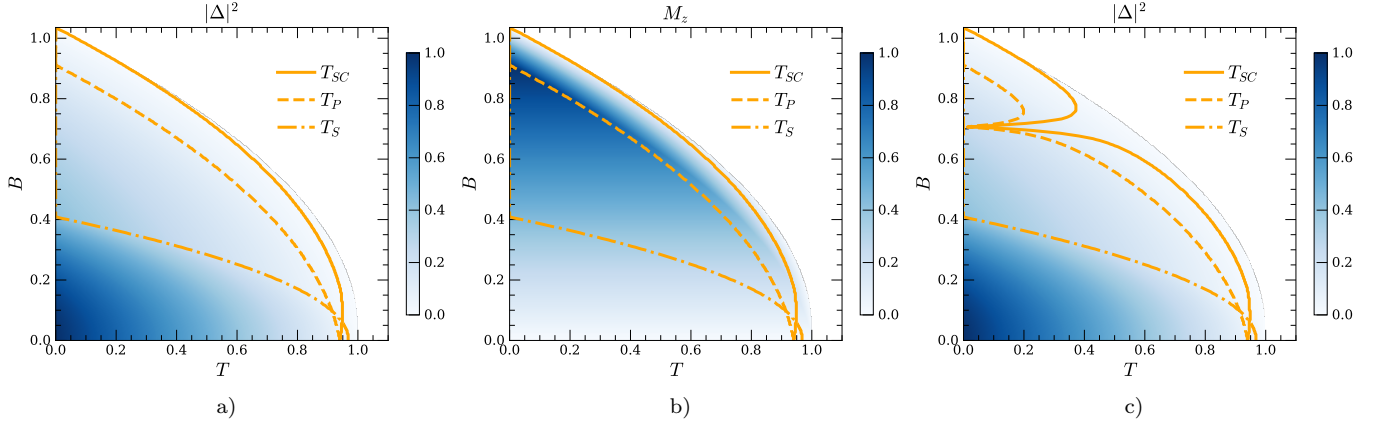


FIG. 3: a) condensate fraction with no spin-valley locking (normalized by  $|\Delta(B=0, T=0)|^2$ ) for a typical choice of parameters, and ignoring dependence of  $K$  on  $B$ .  $T_S$  denotes the transition temperature when the singlet component of the order parameter becomes zero, the line  $T_P$  denotes the transition from the partially to fully polarized state, and  $T_{SC}$  marks the superconducting transition temperature. b) expected value of the component of the magnetization  $\mathbf{M} = \mathbf{d}_1 \times \mathbf{d}_2$  along the field direction, normalized to  $M_z(B=0, T=0)$ . c) the same as a), but taking into account a downward renormalization of  $K$  by a term proportional to  $B^2$ , leading to a field-induced Lifshitz transition. These types of phase diagrams are obtained by taking  $-r, -r', u, u', \kappa, \kappa' > 0$ ,  $v, q \geq 0$ , and  $\zeta$  small and nonzero.

There are now two types of solutions. The first is with  $d_+ = 0$  or  $d_- = 0$ , corresponding to fully spin-polarized triplets. Choosing  $\zeta B > 0$ , the favored solution will be the one with  $d_- = 0$ . This gives the fully polarized solution

$$\mathbf{d}_{fp} = \left( e^{i(\phi+\gamma)} \sqrt{\frac{\zeta B - 2\tilde{r}}{2\tilde{u} + v}}, 0, 0 \right), \quad (\text{B29})$$

where  $\phi$  ( $\gamma$ ) shifts under  $U(1)_g$  gauge transformations ( $SO(2)_s$  spin rotations about  $\hat{\mathbf{z}}$ ). This order parameter is therefore invariant under combined  $U(1)_g$  and  $SO(2)_s$  rotations. As a result, when  $B = 0$  the order parameter manifold is  $\mathcal{M} = (U(1)_g \times SO(3)_s) / U(1)_{g+s} = SO(3)$  [37], since the  $U(1)_g$  subgroup can be absorbed into the  $SO(3)$  part. As discussed in the main text, this means that at  $B = 0$  there are only  $\mathbb{Z}_2$  vortices and that  $T_{BKT} = 0$ , which is incompatible with experiment. When  $B \neq 0$  the order parameter manifold is instead  $\mathcal{M} = (U(1)_g \times U(1)_s) / U(1)_{g+s} = U(1)$ , so that  $\mathbb{Z}$  vortices (and a nonzero  $T_{BKT}$ ) occur at finite fields.

When both  $d_+$  and  $d_-$  are nonzero, we find the partially polarized solution

$$\mathbf{d}_{pp} = \frac{1}{\sqrt{2}} \left( e^{i(\phi+\gamma)} \sqrt{\zeta B/v - \tilde{r}/\tilde{u}}, 0, e^{i(\phi-\gamma)} \sqrt{-\zeta B/v - \tilde{r}/\tilde{u}} \right), \quad (\text{B30})$$

which is only a solution provided that the field is less than the critical value  $B_c$ , where

$$B_c = \frac{v|\tilde{r}|}{\zeta\tilde{u}}, \quad (\text{B31})$$

which only implicitly determines  $B_c$  due to the implicit  $B^2$  dependence of the coefficients on the RHS. When  $B = 0$ ,  $\mathbf{d}_{pp}$  is invariant under  $SO(2)_d$  spin rotations about  $\mathbf{d}$ , but unlike the fully polarized case the only  $SO(3)$  rotation which acts in the same way as a gauge transformation is a spin flip, which acts in the same way as a gauge rotation by  $\pi$ . The order parameter manifold when  $B = 0$  is therefore  $\mathcal{M} = (U(1)_g \times SO(3)_s) / (SO(2)_d \times \mathbb{Z}_2) = (U(1) \times S^2) / \mathbb{Z}_2$ . Since  $\pi_1([U(1) \times S^2] / \mathbb{Z}_2) = \mathbb{Z}$ , the  $B = 0$  phase has vortices and a nonzero  $T_{BKT}$  [50]. When  $B \neq 0$ , the order parameter manifold is instead  $(U(1)_g \times U(1)_s) / \mathbb{Z}_2$ .

When  $v < 0$  the fully polarized solution is favored regardless of  $B$ , while when  $v > 0$  the partially polarized solution is favored as long as  $B < B_c$ . The transition at  $B_c$  occurs when vortices in the spin mode  $\gamma$  condense, and is of BKT type within the context of the present discussion. Since we know from experiments that the low-field phase is not fully polarized, we may therefore assume  $v > 0$  in what follows.

Figure 3 shows the condensate fraction  $\langle |\Delta|^2 \rangle$  and magnetization  $\langle M_z \rangle$  for a typical choice of parameters. Panels a) and b) show the condensate fraction and magnetization, with the lines drawn at the locations of the BKT transitions separating the various phases. Panel c) shows where the transitions occur upon accounting for a further suppression

of the stiffness  $K$  by orbital effects (see appendix A), with the Lifshitz transition occurring at the field where  $T_{SC}$  first goes to zero.

Note that for any nonzero magnetic field, the transition into the fully polarized state always occurs before superconductivity is killed altogether ( $T_P < T_{SC}$  in figure 3). This is because the fully polarized state is favored by a term which is linear in  $|\Delta|^2$  (the  $\mathbf{B} \cdot (\mathbf{d}_1 \times \mathbf{d}_2)$  term), while the partially polarized state is favored by a term quadratic in  $|\Delta|^2$  (the  $v$  term); hence when  $|\Delta|^2$  is small the fully polarized solution always wins. This means that in the  $w = 0$  limit, near the boundary of the superconducting dome the superconductor will always be in the fully polarized phase, regardless of the strength of the magnetic field. As a result,  $T_{SC}$  *increases* linearly with  $B$  at small  $B$ , by an amount which is small in  $\zeta$ . Since in the context of magic-angle TTG we expect  $\zeta$  to be small, especially so at small  $U$  (again, see appendix A), this may or may not be incompatible with experiment.

### Appendix C: Nature of the Lifshitz critical points

In this appendix we elaborate on the nature of the Lifshitz critical points discussed in section V. Throughout we will write down free energies which involve only the phase mode  $\phi$ , with the fluctuations in the spin part of the order parameter ignored on grounds that they flow to zero coupling in the IR.

#### a. Isotropic case

We first review the isotropic case, where the Lagrangian at the  $T = 0$  critical point contains no quadratic derivative terms:

$$\mathcal{L} = \frac{1}{2} ((\partial_\tau \phi)^2 + \zeta (\nabla^2 \phi)^2). \quad (\text{C1})$$

A complete discussion of the finite-temperature physics of this model can be found in [38]; here we will review the results of [38] which are relevant for the present discussion.

The spatial correlator  $\langle \phi(x)\phi(0) \rangle$  calculated from (C1) is logarithmically divergent in the IR at  $T = 0$ , and the model has QLRO. At nonzero  $T$  the divergence is power law, so that the correlation functions of  $e^{i\phi}$  are ultra short-ranged. This latter fact does not mean that the model is necessarily disordered at all  $T > 0$  however, since as stressed in [38] irrelevant terms will generate a nonzero stiffness (i.e, a nonzero  $(\nabla\phi)^2$  term), allowing the  $e^{i\phi}$  correlators to decay as power laws at  $T > 0$ .

Accounting for these effects at  $T > 0$  is done by starting instead with the Lagrangian

$$\mathcal{L} = \frac{1}{2} ((\partial_\tau \phi)^2 + K^0 (\nabla \phi)^2 + \zeta (\nabla^2 \phi)^2 + \eta (\nabla \phi)^4), \quad (\text{C2})$$

where the marginally irrelevant  $\eta$  term is responsible for generating a nonzero stiffness at  $T > 0$ , and  $K^0$  is a counterterm ensuring that the renormalized stiffness  $K$  vanish at  $T = 0$ .

The renormalized stiffness  $K$  can be computed within the framework of an RPA (large- $N$ ) approximation. In this approximation,  $K$  is given by the bare value  $K^0$  in addition to a contribution coming from a geometric sum of 1-loop diagrams involving the  $\eta$  interaction. Writing the renormalized  $\phi$  propagator as  $(\omega^2 + Kk^2 + \zeta k^4)^{-1}$ , we have

$$\begin{aligned} K &= K^0 + 4\eta T \sum_{\omega_n} \int_{\mathbf{k}} \frac{k^2}{\omega_n^2 + Kk^2 + \zeta k^4} \\ &= 2\eta \int_{\mathbf{k}} k^2 \left( \frac{1}{\sqrt{Kk^2 + \zeta k^4}} - \frac{1}{\sqrt{\zeta} k^2} + 2 \frac{1}{\sqrt{Kk^2 + \zeta k^4}} \frac{1}{e^{\sqrt{Kk^2 + \zeta k^4}/T} - 1} \right) \end{aligned} \quad (\text{C3})$$

where we have used the requirement that  $K_y = 0$  at  $T = 0$  to solve for  $K_y^0$ .

Performing this integral and solving for  $K$  at low  $T$  gives [38]

$$K \approx 4T\sqrt{\zeta} \frac{\ln^2(\Lambda^2/T)}{\ln(\Lambda^2/T)}, \quad (\text{C4})$$

so that  $K/T\sqrt{\zeta} \ll 1$  at low  $T$ .

The value of  $K$  determines  $T_{BKT}$  by way of the Kosterlitz-Thouless criterion  $K\pi/2 = T$ . Since in the present case  $K/T \rightarrow 0$  at low  $T$ , there will *not* be a BKT transition at finite  $T$ , and the theory is disordered for all non-zero temperatures [38].

b. *Anisotropic case*

Let us now apply a similar analysis to the anisotropic case, where only one component of the stiffness vanishes at the  $T = 0$  critical point. At  $T = 0$ , we may work with the Lagrangian

$$\mathcal{L} = \frac{\lambda}{2} \left( (\partial_\tau \phi)^2 + K_x (\partial_x \phi)^2 + \varsigma (\partial_y^2 \phi)^2 \right), \quad (\text{C5})$$

where  $\varsigma$  is needed for stability. We take  $\phi$  to be dimensionless, with the dimensionful factor of  $\lambda$  out front having dimensions of  $\partial_y$ . At zero temperature the  $\phi$  correlator along e.g. the  $\hat{x}$  direction goes as

$$\langle \phi(x) \phi(0) \rangle = \frac{1}{\lambda} \int_{\omega, \mathbf{k}} \frac{e^{ik_x x}}{\omega^2 + K_x k_x^2 + \varsigma k_y^4} = \frac{1}{\lambda} \frac{2\Gamma(5/4)^2}{\pi^{3/2} (\varsigma K_x)^{1/4}} \int \frac{dk_x}{2\pi} \frac{e^{ik_x x}}{\sqrt{k_x}}, \quad (\text{C6})$$

which is IR finite. The theory therefore orders at  $T = 0$ , unlike in the isotropic case [51]. However, like the isotropic case, the finite- $T$  correlator calculated using the above Lagrangian (C5) has a power-law divergence in the system size: for example, at large  $T$ ,

$$\langle \phi(x) \phi(0) \rangle = \frac{T}{2^{5/2} K_x^{3/4} \lambda} \int \frac{dk_x}{2\pi} \frac{e^{ik_x x}}{k_x^{3/2}}, \quad (\text{C7})$$

which diverges as  $\sqrt{L}$ , with  $L$  the system size.

As in the isotropic case, this divergence does not rule out the existence of a finite- $T$  phase with QLRO, as we must account for the presence of other terms which generate a nonzero  $y$  component of the stiffness at  $T > 0$ .

At  $T > 0$ , the correct starting point is therefore instead, analogously to (C2),

$$\mathcal{L} = \frac{\lambda}{2} \left( (\partial_\tau \phi)^2 + K_x (\partial_x \phi)^2 + K_y^0 (\partial_y \phi)^2 + \varsigma (\partial_y^2 \phi)^2 + \eta (\partial_y \phi)^4 \right). \quad (\text{C8})$$

The renormalized  $y$  stiffness  $K_y$  can be computed using the same 1-loop diagram as before. We have

$$\begin{aligned} K_y &= K_y^0 + \frac{4\eta T}{\lambda} \sum_{\omega_n} \int_{\mathbf{k}} \frac{k_y^2}{\omega_n^2 + K_x k_x^2 + K_y k_y^2 + \varsigma k_y^4} \\ &= \frac{2\eta}{\lambda} \int_{\mathbf{k}} k_y^2 \left( \frac{1}{\sqrt{K_x k_x^2 + K_y k_y^2 + \varsigma k_y^4}} - \frac{1}{\sqrt{K_x k_x^2 + \varsigma k_y^4}} + 2 \frac{1}{\sqrt{K_x k_x^2 + K_y k_y^2 + \varsigma k_y^4}} \frac{1}{e^{\sqrt{K_x k_x^2 + K_y k_y^2 + \varsigma k_y^4}/T} - 1} \right). \end{aligned} \quad (\text{C9})$$

Assuming  $K_y \ll \sqrt{\varsigma T}$ , we can ignore the  $K_y$  dependence of the third term in the integral. We then solve for  $K_y$ , finding

$$K_y \approx \Upsilon T^{3/2}, \quad \Upsilon = \frac{C}{\varsigma^{3/4} \sqrt{K_x} (2\pi^2 \lambda / \eta + 2\Lambda / \sqrt{K_x \varsigma})} \quad (\text{C10})$$

where  $C \approx 11$ . At small  $T$  we indeed have  $K_y \ll \sqrt{\varsigma T}$ , justifying our earlier assumption. Note that unlike in the isotropic case, here we are not in the upper critical dimension of the  $T = 0$  theory, and the term  $\eta$  which generates  $K_y$  is not marginal. Therefore there is no logarithmic dependence of  $K_y$  on  $T$ , and the dependence of  $K_y$  on the value of  $\eta$  can not a priori be neglected at low  $T$ .

In the anisotropic case, the stiffness which determines the location of a possible BKT transition is set by  $K_{\text{eff}} = \lambda \sqrt{K_x K_y}$ . Since  $K_{\text{eff}} \propto T^{3/4}$ ,  $K_{\text{eff}}$  will always be larger than  $2T/\pi$  at low enough  $T$ , meaning that there will always be a finite-temperature BKT transition. Explicitly, in the present approximation the transition occurs at

$$T_{BKT} \approx \left( \frac{\lambda \pi}{2} \sqrt{K_x \Upsilon} \right)^4. \quad (\text{C11})$$

The value of  $T_{BKT}$  is of course non-universal, and depending on the microscopics may very well be small enough to be essentially indistinguishable in experiments from zero.

c. *Critical currents*

In this section we compute the critical current in the region of parameter space near the Lifshitz transition.

First consider the normal phase where the pairing occurs at zero momentum. In this case we may examine the free energy density

$$f[\nabla\phi] = \frac{\rho}{2} (\gamma_x(\partial_x\phi)^2 + \gamma_y(\partial_y\phi)^2) + r\rho + \frac{u}{2}\rho^2, \quad (\text{C12})$$

where  $\gamma_x, \gamma_y, u, -r > 0$ . To determine the critical current, we follow the standard procedure [52] of Legendre transforming to a thermodynamic potential  $g$  which is a function of the current  $\mathbf{I}$ , satisfying (working in units where  $2e = 1$ )

$$g[\mathbf{I}] = -\nabla\phi \cdot \mathbf{I} + f[\nabla\phi], \quad \frac{\delta g}{\delta \mathbf{I}} = -\nabla\phi, \quad \frac{\delta f}{\delta \nabla\phi} = \mathbf{I}. \quad (\text{C13})$$

Therefore

$$g[\mathbf{I}] = -\frac{I_x^2}{2\rho\gamma_x} - \frac{I_y^2}{2\rho\gamma_y} + r\rho + \frac{u}{2}\rho^2. \quad (\text{C14})$$

Minimizing over  $\rho$  gives

$$\frac{I_x^2}{2\gamma_x} + \frac{I_y^2}{2\gamma_y} = |r|\rho^2 - u\rho^3. \quad (\text{C15})$$

Maximizing the RHS, we then see that the critical currents along the  $x$  and  $y$  directions are given by

$$I_{c,x} = \sqrt{\frac{8\gamma_x}{27} \frac{|r|^3}{u}}, \quad I_{c,y} = \sqrt{\frac{\gamma_y}{\gamma_x}} I_{c,x}. \quad (\text{C16})$$

Now we consider a situation which favors finite-momentum pairing along the  $\hat{\mathbf{y}}$  direction. We consider the free energy density

$$f[\Psi] = \frac{1}{2} (\gamma_x |\partial_x \Psi|^2 - \gamma_y |\partial_y \Psi|^2 + \frac{\varkappa}{2} |\partial_y^2 \Psi|^2) + r |\Psi|^2 + \frac{u}{2} |\Psi|^4, \quad (\text{C17})$$

with  $\gamma_y, \varkappa > 0$ . We then consider an FF state, with

$$\Psi = \sqrt{\rho} e^{iq_0 y + i\phi}, \quad q_0 = \sqrt{\gamma_y / \varkappa}, \quad (\text{C18})$$

and assume that  $\phi$  varies on scales much longer than  $1/q_0$ . We can then consider the approximate free energy

$$f[\nabla\phi] \approx \frac{\rho}{2} (\gamma_x (\partial_x\phi)^2 + \frac{\gamma_y}{2} (\partial_y\phi)^2) + (r - \gamma_y^2/4\varkappa)\rho + \frac{u}{2}\rho^2. \quad (\text{C19})$$

Repeating the same steps as above leads to critical currents of

$$I_{c,x} = \sqrt{\frac{8\gamma_x}{27u^2} |r - \gamma_y^2/4\varkappa|^3}, \quad I_{c,y} = \sqrt{\frac{\gamma_y}{2\gamma_x}} I_{c,x}. \quad (\text{C20})$$

Consider finally the scenario where both quadratic derivatives in the free energy density have negative coefficients [53, 54], with

$$f[\Psi] = \frac{1}{2} (-\gamma_x |\partial_x \Psi|^2 - \gamma_y |\partial_y \Psi|^2 + \frac{\varkappa_x}{2} |\partial_x^2 \Psi|^2 + \frac{\varkappa_y}{2} |\partial_y^2 \Psi|^2 + \varkappa_{xy} |\partial_x \partial_y \Psi|^2) + r |\Psi|^2 + \frac{u}{2} |\Psi|^4, \quad (\text{C21})$$

with all coefficients except  $r$  positive. Consider again an FF state  $\Psi = \sqrt{\rho} e^{iq_0 y + i\phi}$ , with finite momentum in the  $\hat{\mathbf{y}}$  direction. Assuming  $\gamma_y \varkappa_{xy} / \varkappa_y - \gamma_x \geq 0$  and making the same approximations as above, this leads to the critical currents

$$I_{c,x} = \sqrt{\frac{8(\gamma_y \varkappa_{xy} / \varkappa_y - \gamma_x)}{27u^2} |r - \gamma_y^2/4\varkappa_y|^3}, \quad I_{c,y} = \sqrt{\frac{\gamma_y}{2(\gamma_y \varkappa_{xy} / \varkappa_y - \gamma_x)}} I_{c,x}. \quad (\text{C22})$$

### Appendix D: Estimate of the in-plane critical field $B_c$

In this appendix we provide a crude estimate of the critical in-plane field  $B_c$  of the superconductor at zero displacement field and optimal doping. Given the phenomenological nature of our analysis and the fact that  $B_c$  may very well not be small enough for leading-order perturbation theory in  $B$  to be reliable, this calculation is not meant to be taken as a sharp prediction — rather it is meant simply as a consistency check to make sure that our proposal is not grossly incompatible with experiment.

From equation (A21) of appendix A, the stiffness at zero displacement field and small  $B$  is given by<sup>8</sup> (assuming for simplicity that the velocity of the mirror-odd Dirac cone after integrating out the mirror-even fermions is isotropic and equal to the monolayer value of  $v_D$ )

$$K_{eff} = K_0 - 2C_2|\Delta|^2\lambda hv_D^2\delta^2 B^2 + \dots, \quad (D1)$$

where  $\dots$  are terms higher order in  $B$ ,  $\delta \approx 0.3\text{nm}$  is the interlayer separation,  $K_0$  is the  $B = 0$  phase stiffness, and where the definitions of  $C_2, \lambda, h$  will be recapitulated shortly. Estimating the zero-field stiffness as  $K_0 \approx 2T_{BKT,0}/\pi$  where  $T_{BKT,0} \approx 1.5\text{K}$  [25] is the zero-field transition temperature, we have

$$B_c \approx \frac{1}{v_D\delta} \sqrt{\frac{T_{BKT,0}}{\pi C_2|\Delta|^2\lambda h}} \approx 1.2 \sqrt{\frac{\mathcal{E}}{1\text{meV}}} \text{T}, \quad (D2)$$

where the energy scale  $\mathcal{E}$  is defined as

$$\mathcal{E} \equiv \frac{1}{\pi C_2|\Delta|^2\lambda h}. \quad (D3)$$

We now need to estimate the quantities appearing in  $\mathcal{E}$ .

The coupling  $h$  is associated with the pairing terms for the mirror-odd fermions that arise upon integrating out the mirror-even sector (see (A4)), and has dimension of inverse energy squared. While we are not in a position to calculate  $h$  exactly, on general grounds we expect  $h \sim 1/\varepsilon_F^2$ , where  $\varepsilon_F$  is the Fermi energy of the mirror-even bands which become superconducting.  $\varepsilon_F$  is not known exactly and its determination depends on knowledge of what happens to the correlated insulator at  $\nu = -2$  upon hole doping, but a value in the ballpark of a few tens of meV seems appropriate.

$C_2$  can be read off from an expansion of (A17), giving

$$C_2 = -\frac{3}{T} \int_{\mathbf{k},x} y \text{sech}^2(k/2T) \frac{k - T \sinh(k/T)}{2k^3} = \frac{1}{8\pi T}, \quad (D4)$$

where  $x$  is integrated from 0 to 1 and  $y = x(1-x)$ . For the present purposes of estimating  $B_c$ , we will take  $T \rightarrow T_{BKT,0} \sim T_{onset}$  in the above.

The appropriate value to take for  $\Delta$  (equal to the superconducting gap in our mean-field treatment) is not currently known experimentally. We will leave  $\Delta$  as a free parameter for now, and note only that due to the strong-coupling nature of the problem and the energy scales involved, a value of  $\Delta \sim 1\text{meV}$  does not seem unreasonable.

Finally, we need to estimate the dimensionless parameter  $\lambda$ , which appears in our Lagrangian as  $\lambda\psi_- \Delta^\dagger \psi_- + h.c.$  and originates from a mean-field decoupling of interactions of the form  $\psi_-^\dagger \psi_-^\dagger \psi_+ \psi_+ + h.c.$  (all indices suppressed). These interaction only conserve the number of mirror-even and mirror-odd fermions modulo 2, and they do not appear directly in the regular density-density interaction, as the total fermion density is  $\delta\rho = \delta\rho_{even} + \delta\rho_{odd}$  which separately conserves the number of mirror-even and mirror-odd fermions. These interactions furthermore do not arise at first order upon projecting the density-density interaction to the active bands, since at  $D = 0, B = 0$  the form factors of the active bands do not mix the mirror-even and mirror-odd sectors (by symmetry). Therefore after projecting the interaction to the active bands,  $\lambda$  can only arise from an interaction corresponding to a second-order process involving

---

<sup>8</sup> Note that we are ignoring contributions to the stiffness which arise from Zeeman effects. Within the framework of appendix A, Zeeman effects in the mirror-even sector do not contribute to the stiffness at leading order, while contributions from the mirror-odd sector are small (being suppressed by a factor of  $\lambda(\mu_B/v_D\delta)^2 \ll 1$ ), and will be ignored.

tunneling to remote bands. The strength of these higher-order interactions is roughly  $U^2/\Delta E$ , where  $\Delta E$  is the gap to the remote bands and  $U$  is the Coulomb interaction strength;  $\lambda$  is accordingly  $\lambda \sim (U^2/\Delta E)/U = U/\Delta E$ . Estimating  $U \sim e^2/(4\pi\epsilon_0\epsilon_r\sqrt{A_{\text{moire}}}) \sim 30$  meV (with the relative dielectric constant  $\epsilon_r \sim 5$ ) and  $\Delta E \sim 50$  meV (based on a naive reading of the band structure [21, 55]) gives  $\lambda \sim 3/5$ .

With the above estimates in hand, we may plug in numbers and write

$$B_c \approx 28 \frac{1 \text{ meV}}{|\Delta|} \frac{\epsilon_F}{20 \text{ meV}} \sqrt{\frac{3/5}{\lambda}} \sqrt{\frac{C_2^{-1}}{8\pi T_{BKT,0}}} \text{ T}. \quad (\text{D5})$$

The value of  $B_c$  at zero displacement field and optimal doping is experimentally determined to be around 12 T [25]. This is rather close to the value we would read off from (D5) given the estimates for the parameters on the RHS as discussed above, and given the crudeness of the present estimate (especially in the determination of e.g.  $h$  and  $C_2$ ) this is really the best we may hope to do. A more careful analysis is left to future work, but we find it encouraging that the rough estimate here at least seems to be in the right ballpark.

## Appendix E: Single-particle physics

In this appendix we provide a few details on the single-particle properties of TTG in the presence of an in-plane field.

### 1. Hamiltonian and symmetries

In the basis  $(\psi_1, \psi_2, \psi_3)$ , with  $\psi_l$  the annihilation operator for fermions on layer  $l$ , the single-particle Hamiltonian is

$$H_0 = \begin{pmatrix} -iv_D \boldsymbol{\sigma}_{v,\theta/2} \cdot (\nabla - i\mathbf{A}_1) + U/2 & T & & \\ T^\dagger & -iv_D \boldsymbol{\sigma}_{v,-\theta/2} \cdot (\nabla - i\mathbf{A}_2) & & \\ & T & T^\dagger & \\ & & -iv_D \boldsymbol{\sigma}_{v,\theta/2} \cdot (\nabla - i\mathbf{A}_3) - U/2 & \end{pmatrix} + H_Z, \quad (\text{E1})$$

where  $\boldsymbol{\sigma}_{v,\theta/2}$  denotes a rotation of  $\boldsymbol{\sigma}_v = (\tau^z \sigma^x, \sigma^y)$  by  $\theta/2$ , where  $\theta$  is the twist angle.  $\mathbf{A}_l$  and  $H_Z$  are as in the main text.

We will use the form of the  $T$  matrix given in e.g. [28], which we take to be

$$T(\mathbf{x}) = \sum_{i=1}^3 e^{-i\tau^z \mathbf{q}_i \cdot \mathbf{x}} T_i, \quad (\text{E2})$$

with the tunneling matrices

$$T_i = w_{AA} \mathbf{1} + w_{AB} (\sigma^x \cos[2\pi(i-1)/3] + \tau^z \sigma^y \sin[2\pi(i-1)/3]) \quad (\text{E3})$$

and where

$$\mathbf{q}_1 = k_\theta(0, -1), \quad \mathbf{q}_2 = k_\theta(\sqrt{3}/2, 1/2), \quad \mathbf{q}_3 = k_\theta(-\sqrt{3}/2, 1/2), \quad (\text{E4})$$

with

$$k_\theta \equiv 2k_D \sin(\theta/2), \quad k_D \equiv \frac{4\pi}{3a_0}, \quad (\text{E5})$$

and with  $a_0 \approx 0.246$  nm the graphene lattice constant. We will work in conventions where the Moire  $K$  and  $K'$  points are given as  $K = -\mathbf{q}_2$  and  $K' = \mathbf{q}_3$ . In the limit where the graphene layers are decoupled, the band structure then consists of two Dirac cones at  $K$ , and one at  $K'$ .

In the limit  $w_{AA} = 0$  the rotations of the  $\boldsymbol{\sigma}_v$  matrices in (E1) can be removed by a unitary transformation [28], but even away from this limit their effects can usually be ignored on the grounds that  $\theta$  is small [27, 28]. We will therefore simply ignore the rotations in what follows.

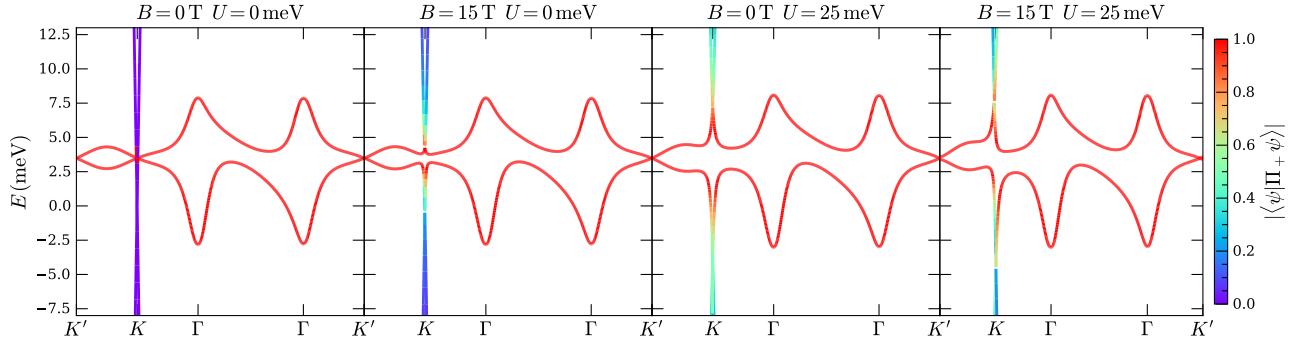


FIG. 4: The structure of the central bands for several values of  $U$  and  $B$  (here  $\mathbf{B} \parallel \hat{\mathbf{x}}$ ). The color scale denotes the extent to which the wavefunctions are mirror-even, with  $\Pi_+ = (1 + M)/2$ .

The symmetries of the  $U = \mathbf{B} = 0$  problem include [21, 26]<sup>9</sup>

$$\begin{aligned}
 C_{2z} &= \sigma^x \tau^x R_\pi \\
 C_3 &= e^{2\pi i \sigma^z \tau^z / 3} R_{2\pi/3} \\
 C_{2x} &= M \sigma^x R_y \\
 \mathcal{T} &= \tau^x \mathcal{K},
 \end{aligned} \tag{E6}$$

where the RHSs are the matrices which appear in the adjoint action on the  $\psi_l$  fields and  $\mathcal{K}$  is complex conjugation. Here  $R_\phi$  is a rotation through  $\phi$  acting in the  $xy$  plane,  $R_y$  implements a reflection of the  $y$  coordinate, and the mirror symmetry  $M$  acts in layer space as

$$M = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}. \tag{E7}$$

When written in terms of fields which diagonalize  $M$ , the above expression for  $H_0$  in (E1) reduces to the form given in the main text (1).

As in the bilayer case, in the  $w_{AA} = U = 0$  limit we have  $\{\sigma^z, H\} = 0$ . Unlike in TBG, away from this limit there is no unitary chiral symmetry which arises if we ignore the  $\pm\theta/2$  rotations of  $\sigma_v$  (which in TBG acts as  $i\sigma^x \mu^y \mathcal{K}$ , with  $\mu^y$  in layer space).

Turning on a nonzero  $U$  breaks  $M$  and  $C_{2x}$ , and turning on a nonzero  $\mathbf{B}$  breaks  $C_3, \mathcal{T}, M$ , and  $C_{2z}$ , and breaks  $C_{2x}$  unless  $\mathbf{B} \parallel \hat{\mathbf{x}}$ . Any direction of field preserves  $C_{2z}M, C_{2z}\mathcal{T}$ , and  $M\mathcal{T}$ . The combination of  $C_{2x}$  and one of  $\mathcal{T}, M, C_{2z}$  is also preserved if  $\mathbf{B} \parallel \hat{\mathbf{y}}$ . Like in the bilayer case,  $C_{2z}\mathcal{T}$  is unbroken for any choices of  $U, \mathbf{B}$ , and prevents the innermost bands from separating from one another [43].

## 2. Band structure

We now briefly discuss the band structure obtained from (E1). For concreteness we will fix  $\theta = 1.57^\circ, w_{AB} = 110\text{meV}$ , and  $w_{AA} = 0.8w_{AB}$ . Since we are mainly interested in the orbital effects of the field, we will ignore the Zeeman energy.

Similarly to  $U$ , the main orbital effects of  $\mathbf{B}$  on the band structure are to hybridize the mirror-odd Dirac fermion  $\psi_-$  with the mirror-even flat bands. Examples of the band structure in the graphene  $K$  valley for various values of  $U, \mathbf{B} \parallel \hat{\mathbf{x}}$  are shown in figure 4, where the color denotes the magnitude of the mirror-even component of the band wavefunctions.

The orbital effects of the magnetic field are rather different than in the case of TBG [56]. First, when  $U = 0$ , mirror symmetry implies that the spectrum at each  $\mathbf{k}$  point is an even function of  $B$ . This means that at small  $U$

<sup>9</sup>  $C_{2x}$  is technically only a symmetry in the limit where the  $\pm\theta/2$  rotations of  $\sigma_v$  can be ignored.

and small  $\mathbf{B}$ , the orbital effects on the band structure are rather weak. Secondly, despite the explicit breaking of  $C_3$  by the magnetic field, Dirac points remain at the Moire  $K, K'$  points to an extremely good approximation for all experimentally relevant field strengths.

### 3. Effects of the field near the $K$ and $K'$ points

In this section we adopt the approach used in [27] to make some simple statements about the effects of the field on the band structure near the Moire  $K$  and  $K'$  points, at small values of  $U, B$ . Throughout we will focus on the graphene  $K$  valley, and will neglect the physical spin degree of freedom. We will also find it convenient to follow [28] and use the notation  $\alpha \equiv w_{AB}/v_D k_\theta$  and  $\kappa \equiv w_{AA}/w_{AB}$ .

#### a. $K'$ point

We can understand the (lack of) effects of the field at the  $K'$  point in a rather pedestrian manner as follows. We follow [27] and introduce a minimal model for the modes right at  $K'$  by writing down the Hamiltonian

$$H_{K'} = \begin{pmatrix} 0 & T_1 & T_2 & T_3 & T_1^\dagger & T_2^\dagger & T_3^\dagger \\ T_1^\dagger & h_1 & & & & & \\ T_2^\dagger & & h_2 & & & & \\ T_3^\dagger & & & h_3 & & & \\ T_1 & & & & -h_1 & & \\ T_2 & & & & & -h_2 & \\ T_3 & & & & & & -h_3 \end{pmatrix}, \quad (\text{E8})$$

where we are in the basis  $(\psi_2, \psi_1^1, \psi_1^2, \psi_1^3, \psi_2^1, \psi_2^2, \psi_2^3)^T$ , with the lower index denoting the layer index, the upper index denoting a reciprocal space lattice index, and where (working in units where  $v_D = 1$ )

$$h_i = (\mathbf{q}_i - \mathbf{A}) \cdot \boldsymbol{\sigma} + U/2. \quad (\text{E9})$$

As in the analysis of the analogous bilayer problem, there are two zero modes in this simplified 14-band problem. Let  $\psi_2^\uparrow = (1, 0)^T$  and  $\psi_2^\downarrow = (0, 1)^T$ . We will then have a zero mode provided we take

$$\psi_1^j = -h_j^{-1} T_j^\dagger \psi_\sigma^2, \quad \psi_3^j = +h_j^{-1} T_j \psi_\sigma^2, \quad (\text{E10})$$

and that

$$\sum_j (T_j h_j^{-1} T_j^\dagger - T_j^\dagger h_j^{-1} T_j) \psi_2^\sigma = 0, \quad (\text{E11})$$

with this last equation trivially satisfied as our  $T_j$  matrices are Hermitian.

For simplicity we will now work perturbatively in  $\mathbf{A}$ . At  $\mathbf{A} = 0$ , the zero modes  $|\psi_{zm;\sigma}\rangle$  have norm

$$\begin{aligned} \langle \psi_{zm;\sigma} | \psi_{zm;\sigma} \rangle &= 1 + 2 \sum_j T_j (h_j^{-1})^2 T_j \\ &= 1 + 6\alpha^2 \varpi (\kappa^2 + 1), \quad \varpi \equiv \frac{1 + (U/k_\theta)^2/4}{(1 - (U/k_\theta)^2/4)^2}. \end{aligned} \quad (\text{E12})$$

After correctly normalizing the wavefunctions, the effective Hamiltonian in the zero-mode subspace near the  $K'$  point then reads

$$\begin{aligned} H_{eff, K'+\mathbf{k}} &= \frac{1}{1 + 6\alpha^2 \varpi (\kappa^2 + 1)} \left( \boldsymbol{\sigma} \cdot \mathbf{k} + \sum_j (T_j h_j^{-1} \boldsymbol{\sigma} \cdot (\mathbf{k} - \mathbf{A}) h_j^{-1} T_j + T_j h_j^{-1} \boldsymbol{\sigma} \cdot (\mathbf{k} + \mathbf{A}) h_j^{-1} T_j) \right) \\ &= \frac{1 - 6\alpha^2}{1 + 6\alpha^2 \varpi (\kappa^2 + 1)} \boldsymbol{\sigma} \cdot \mathbf{k}, \end{aligned} \quad (\text{E13})$$

with the Fermi velocity out front confirming that when  $U = 0$ ,  $\theta_1^3 = \sqrt{2}\theta_1^2$ , with  $\theta_i^l$  the  $i$ th magic angle for twisted  $l$ -layer graphene [27, 28]. As expected from symmetry considerations,  $H_{eff, K'+\mathbf{k}}$  is independent of the magnetic field to this order.

b. *K* point

The effects of the magnetic field near the *K* point are slightly more interesting, due to the hybridization between the flat bands and the mirror-odd Dirac cone.

We proceed as above: near *K*, the analogous minimal model operates in the space  $(\psi_1, \psi_3, \psi_2^1, \psi_2^2, \psi_2^3)$ , where again the upper index is the layer index and the lower index is the reciprocal lattice index. Right at *K*, the Hamiltonian is

$$H_K = \begin{pmatrix} 0 & T_1^\dagger & T_2^\dagger & T_3^\dagger \\ & 0 & T_1^\dagger & T_2^\dagger & T_3^\dagger \\ T_1 & T_1 & h_1 & & \\ T_2 & T_2 & & h_2 & \\ T_3 & T_3 & & & h_3 \end{pmatrix} \quad (\text{E14})$$

There are now four zero modes. We write the first four components of each zero mode (in the above basis) as the vector  $(\psi_1^\sigma, \psi_3^\rho)$  ( $\sigma, \rho = \uparrow / \downarrow$  as before). In order to have a zero mode, the  $\psi_2^i$  components must satisfy

$$\psi_2^i = -h_i^{-1} T_i (\psi_1^\sigma + \psi_3^\rho) \quad (\text{E15})$$

and

$$\sum_i T_i^\dagger \psi_i^2 = -\sum_i T_i^\dagger h_i^{-1} T_i (\psi_1^\sigma + \psi_3^\rho) = \frac{3\alpha^2 U}{2} \frac{1 + \kappa^2}{1 - (U/k_\theta)^2/4} (\psi_1^\sigma + \psi_3^\rho) = 0. \quad (\text{E16})$$

Evidently we only have zero modes if  $U = 0$  (which we already know to be the case by looking at the band structure).

To isolate the effects of the magnetic field on the Dirac cones, we will accordingly now set  $U = 0$ . Treating  $\mathbf{k}$  and  $\mathbf{A}$  as perturbations, the effective Hamiltonian in the zero mode subspace works out to be

$$H_{eff, K+\mathbf{k}} = \frac{1}{1 + 3\alpha^2(1 + \kappa^2)} \begin{pmatrix} \boldsymbol{\sigma} \cdot (\mathbf{k}(1 - 3\alpha^2) + \mathbf{A}) & -3\alpha^2 \boldsymbol{\sigma} \cdot \mathbf{k} \\ -3\alpha^2 \boldsymbol{\sigma} \cdot \mathbf{k} & \boldsymbol{\sigma} \cdot (\mathbf{k}(1 - 3\alpha^2) - \mathbf{A}) \end{pmatrix} \quad (\text{E17})$$

When  $\mathbf{k} = 0$ , we solve for the spectrum and find two doubly degenerate states at energies (as before,  $\delta$  is the interlayer separation)

$$\varepsilon_K = \pm \delta |B|. \quad (\text{E18})$$

These doubly degenerate states form Dirac cones which live at the *K* point above / below zero energy.

By computing the band structure numerically (see figure 5), one finds that when  $\theta \gtrsim \theta_1$  (with  $\theta_1$  the first magic angle), two Dirac cones split off from the *K* point and move along the direction normal to  $\mathbf{B}$ . To demonstrate this within the context of the present ten-band model we can consider e.g.  $\mathbf{A}, \mathbf{k} \parallel \hat{\mathbf{x}}$ , for which the effective Hamiltonian is

$$H_{eff, K+\mathbf{k}} \propto \sigma^x \otimes \begin{pmatrix} k(1 - 3\alpha^2) + \delta B & -3\alpha^2 k \\ -3\alpha^2 k & k(1 - 3\alpha^2) - \delta B \end{pmatrix} \quad (\text{E19})$$

The determinant of this matrix vanishes at  $k = \pm k_*$ , where

$$k_* = \sqrt{\frac{2}{1 - 6\alpha^2}} \delta |B|, \quad (\text{E20})$$

which evidently is only a solution provided that  $1 - 6\alpha^2 > 0$ , i.e. if  $\alpha < 1/\sqrt{6}$ , which in this approximation is the same as  $\theta > \theta_1$ . Therefore when  $\theta > \theta_1$  we have two Dirac cones located at the *K* point at energies  $\pm \delta B$ , and two at zero energy at the momenta  $(k_*, 0)$ . Since the orbital magnetic energy is much smaller than  $k_\theta$ , this shift is rather small at experimentally relevant field strengths.

On the other hand, when  $\theta \lesssim \theta_1$ , one finds that the two Dirac cones which split off of the *K* point move off parallel to  $\hat{\mathbf{B}}$ . To demonstrate this we may take  $\mathbf{k} \parallel \hat{\mathbf{x}}, \mathbf{A} \parallel \hat{\mathbf{y}}$ . The Hamiltonian is then

$$H_{eff, K+\mathbf{k}} \propto \sigma^x \otimes \begin{pmatrix} k(1 - 3\alpha^2) & -3\alpha^2 k \\ -3\alpha^2 k & k(1 - 3\alpha^2) \end{pmatrix} + \delta B \sigma^y \otimes \sigma^z. \quad (\text{E21})$$

Taking the determinant, one finds two doubly degenerate zero modes when  $k = \pm k_*$ , where now

$$k_* = \frac{\delta |B|}{\sqrt{6\alpha^2 - 1}}. \quad (\text{E22})$$

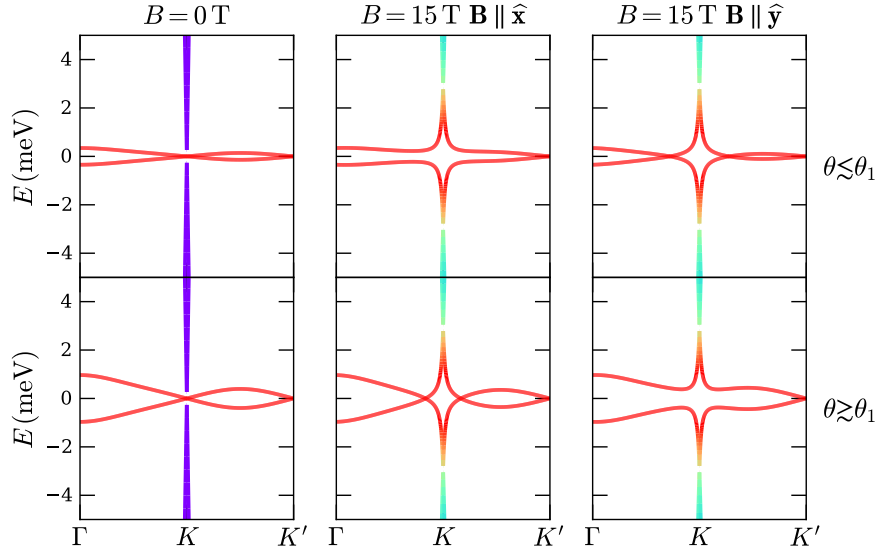


FIG. 5: Field effects of the band structure near the  $K$  point, shown in the limit  $w_{AA} = U = 0$ . The first row of plots are at angles slightly less than the magic angle  $\theta_1$ , while the plots in the second row are slightly above. For  $\theta < \theta_1$  ( $\theta > \theta_1$ ), two Dirac cones move off from  $K$  along (normal to) the field direction (note that the momentum cut is along  $\hat{y}$ ). The color coding indicates mirror projection, as in figure 4.

This is only a solution if  $\alpha > 1/\sqrt{6}$ , i.e if  $\theta < \theta_1$ . Note that the square root dependence of  $k_*$  on  $1/|1 - 6\alpha^2|$  and the linear dependence on  $B$  holds on both sides of the magic angle.

- 
- [1] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, et al., *Nature* **556**, 80 (2018).
  - [2] G. Chen, L. Jiang, S. Wu, B. Lyu, H. Li, B. L. Chittari, K. Watanabe, T. Taniguchi, Z. Shi, J. Jung, et al., *Nature Physics* **15**, 237 (2019).
  - [3] A. L. Sharpe, E. J. Fox, A. W. Barnard, J. Finney, K. Watanabe, T. Taniguchi, M. Kastner, and D. Goldhaber-Gordon, *Science* **365**, 605 (2019).
  - [4] G. Chen, A. L. Sharpe, E. J. Fox, Y.-H. Zhang, S. Wang, L. Jiang, B. Lyu, H. Li, K. Watanabe, T. Taniguchi, et al., *Nature* **579**, 56 (2020).
  - [5] M. Serlin, C. Tschirhart, H. Polshyn, Y. Zhang, J. Zhu, K. Watanabe, T. Taniguchi, L. Balents, and A. Young, *Science* **367**, 900 (2020).
  - [6] U. Zondiner, A. Rozen, D. Rodan-Legrain, Y. Cao, R. Queiroz, T. Taniguchi, K. Watanabe, Y. Oreg, F. von Oppen, A. Stern, et al., *Nature* **582**, 203 (2020).
  - [7] Y. Saito, J. Ge, K. Watanabe, T. Taniguchi, E. Berg, and A. F. Young, arXiv preprint arXiv:2008.10830 (2020).
  - [8] Y. Saito, J. Ge, L. Rademaker, K. Watanabe, T. Taniguchi, D. A. Abanin, and A. F. Young, *Nature Physics* pp. 1–4 (2021).
  - [9] S. Wu, Z. Zhang, K. Watanabe, T. Taniguchi, and E. Y. Andrei, *Nature Materials* pp. 1–7 (2021).
  - [10] S. Chen, M. He, Y.-H. Zhang, V. Hsieh, Z. Fei, K. Watanabe, T. Taniguchi, D. H. Cobden, X. Xu, C. R. Dean, et al., *Nature Physics* pp. 1–7 (2020).
  - [11] Y. Cao, D. Rodan-Legrain, J. M. Park, F. N. Yuan, K. Watanabe, T. Taniguchi, R. M. Fernandes, L. Fu, and P. Jarillo-Herrero, arXiv preprint arXiv:2004.04148 (2020).
  - [12] A. Rozen, J. M. Park, U. Zondiner, Y. Cao, D. Rodan-Legrain, T. Taniguchi, K. Watanabe, Y. Oreg, A. Stern, E. Berg, et al., arXiv preprint arXiv:2009.01836 (2020).
  - [13] H. Zhou, T. Xie, A. Ghazaryan, T. Holder, J. Ehrets, E. M. Spanton, T. Taniguchi, K. Watanabe, E. Berg, M. Serbyn, et al., arXiv preprint arXiv:2104.00653 (2021).
  - [14] L. Balents, C. R. Dean, D. K. Efetov, and A. F. Young, *Nature Physics* **16**, 725 (2020).
  - [15] E. Y. Andrei, D. K. Efetov, P. Jarillo-Herrero, A. H. MacDonald, K. F. Mak, T. Senthil, E. Tutuc, A. Yazdani, and A. F. Young, *Nature Reviews Materials* pp. 1–6 (2021).
  - [16] X. Liu, Z. Hao, E. Khalaf, J. Y. Lee, Y. Ronen, H. Yoo, D. H. Najafabadi, K. Watanabe, T. Taniguchi, A. Vishwanath, et al., *Nature* **583**, 221 (2020).
  - [17] X. Lu, P. Stepanov, W. Yang, M. Xie, M. A. Aamir, I. Das, C. Urgell, K. Watanabe, T. Taniguchi, G. Zhang, et al., *Nature*

- 574, 653 (2019).
- [18] 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).
- [19] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, *Nature* **556**, 43 (2018).
- [20] H. S. Arora, R. Polski, Y. Zhang, A. Thomson, Y. Choi, H. Kim, Z. Lin, I. Z. Wilson, X. Xu, J.-H. Chu, et al., *Nature* **583**, 379 (2020).
- [21] E. Khalaf, A. J. Kruchkov, G. Tarnopolsky, and A. Vishwanath, *Physical Review B* **100**, 085109 (2019).
- [22] X. Li, F. Wu, and A. H. MacDonald, arXiv preprint arXiv:1907.12338 (2019).
- [23] Z. Hao, A. Zimmerman, P. Ledwith, E. Khalaf, D. H. Najafabadi, K. Watanabe, T. Taniguchi, A. Vishwanath, and P. Kim, *Science* **371**, 1133 (2021).
- [24] J. M. Park, Y. Cao, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, *Nature* pp. 1–7 (2021).
- [25] Y. Cao, J. M. Park, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, *Nature* **595**, 526 (2021).
- [26] D. Calugaru, F. Xie, Z.-D. Song, B. Lian, N. Regnault, and B. A. Bernevig, arXiv preprint arXiv:2102.06201 (2021).
- [27] R. Bistritzer and A. H. MacDonald, *Proceedings of the National Academy of Sciences* **108**, 12233 (2011).
- [28] G. Tarnopolsky, A. J. Kruchkov, and A. Vishwanath, *Physical review letters* **122**, 106405 (2019).
- [29] C. Lei, L. Linhart, W. Qin, F. Libisch, and A. H. MacDonald, arXiv preprint arXiv:2010.05787 (2020).
- [30] S. Carr, C. Li, Z. Zhu, E. Kaxiras, S. Sachdev, and A. Kruchkov, *Nano letters* **20**, 3030 (2020).
- [31] A. Vishwanath, J. E. Moore, and T. Senthil, *Physical Review B* **69**, 054507 (2004).
- [32] N. Bultinck, E. Khalaf, S. Liu, S. Chatterjee, A. Vishwanath, and M. P. Zaletel, *Physical Review X* **10**, 031034 (2020).
- [33] E. Khalaf, S. Chatterjee, N. Bultinck, M. P. Zaletel, and A. Vishwanath, arXiv preprint arXiv:2004.00638 (2020).
- [34] S. Chatterjee, N. Bultinck, and M. P. Zaletel, *Physical Review B* **101**, 165141 (2020).
- [35] E. Khalaf, P. Ledwith, and A. Vishwanath, arXiv preprint arXiv:2012.05915 (2020).
- [36] J. Y. Lee, E. Khalaf, S. Liu, X. Liu, Z. Hao, P. Kim, and A. Vishwanath, *Nature communications* **10**, 1 (2019).
- [37] E. Cornfeld, M. S. Rudner, and E. Berg, arXiv preprint arXiv:2006.10073 (2020).
- [38] P. Ghaemi, A. Vishwanath, and T. Senthil, *Physical Review B* **72**, 024420 (2005).
- [39] P. Fulde and R. A. Ferrell, *Phys. Rev.* **135**, A550 (1964), URL <https://link.aps.org/doi/10.1103/PhysRev.135.A550>.
- [40] A. Larkin and I. Ovchinnikov, *Soviet Physics-JETP* **20**, 762 (1965).
- [41] E. Fradkin, D. A. Huse, R. Moessner, V. Oganesyan, and S. L. Sondhi, *Physical Review B* **69**, 224415 (2004).
- [42] A. Vishwanath, L. Balents, and T. Senthil, *Physical Review B* **69**, 224416 (2004).
- [43] H. C. Po, L. Zou, A. Vishwanath, and T. Senthil, *Physical Review X* **8**, 031089 (2018).
- [44] J. M. Park, Y. Cao, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, *Nature* **592**, 43 (2021).
- [45] K. Yang and D. Agterberg, *Physical review letters* **84**, 4970 (2000).
- [46] J.-X. Lin, Y.-H. Zhang, E. Morissette, Z. Wang, S. Liu, D. Rhodes, K. Watanabe, T. Taniguchi, J. Hone, and J. Li, arXiv preprint arXiv:2102.06566 (2021).
- [47] T. Sugiyama and T. Ohmi, *Journal of the Physical Society of Japan* **64**, 2746 (1995).
- [48] M. S. Scheurer and R. Samajdar, *Physical Review Research* **2**, 033062 (2020).
- [49] Y. Kawaguchi and M. Ueda, *Physics Reports* **520**, 253 (2012).
- [50] S. Mukerjee, C. Xu, and J. Moore, *Physical review letters* **97**, 120406 (2006).
- [51] R. D. Pisarski and A. M. Tsvelik, arXiv preprint arXiv:2103.15835 (2021).
- [52] M. Tinkham, *Introduction to superconductivity* (Courier Corporation, 2004).
- [53] A. Buzdin, Y. Matsuda, and T. Shibauchi, *EPL (Europhysics Letters)* **80**, 67004 (2007).
- [54] K. Samokhin and B. Truong, *Physical Review B* **96**, 214501 (2017).
- [55] A. Fischer, Z. A. Goodwin, A. A. Mostofi, J. Lischner, D. M. Kennes, and L. Klebl, arXiv preprint arXiv:2104.10176 (2021).
- [56] Y. H. Kwan, S. Parameswaran, and S. Sondhi, *Physical Review B* **101**, 205116 (2020).
-