Editors' Suggestion   Featured in Physics

Unconventional superconductivity due to interband polarization

Valentin Crépel [1], Tommaso Cea,[2] Liang Fu,[1] and Francisco Guinea[2,3,4]

[1]*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[2]*Imdea Nanoscience, Faraday 9, 28015 Madrid, Spain*
[3]*Donostia International Physics Center, Paseo Manuel de Lardizábal 4, 20018 San Sebastián, Spain*
[4]*Ikerbasque, Basque Foundation for Science, 48009 Bilbao, Spain*

We analyze in detail the superconductivity that arises in an extended Hubbard model describing a multiband system with repulsive interactions. We show that virtual interband processes induce an effective attractive interaction for small momentum transfers, a situation not found in most models of superconductivity from repulsion. This attraction can be traced back, in real space, to the presence of correlated hopping terms induced by interband polarization. We reveal this physics with both strong-coupling expansion and many-body perturbation theory, supplemented by numerical calculations. Finally, we point out interesting similarities with the problem of interacting electrons in twisted bilayer graphene, suggesting the importance of the interband contribution to superconductivity.

I. INTRODUCTION

Superconductivity in low carrier density materials, such as WTe_2 [1–3], $SrTiO_3$ [4,5], ZrNCl [6,7], and twisted bilayer [8–10] and trilayer graphene [11–13], is a subject of intense interest. In conventional electron-phonon superconductors, the vast difference between Fermi and Debye energies is responsible for the retarded nature of the phonon-mediated attraction. This allows mobile electrons to interact without being close to each other at the same time, effectively reducing their mutual Coulomb repulsion [14]. In contrast, low-density systems lack this large separation of electron and phonon energies, which motivates the study of superconductivity due to the electron-electron repulsion itself, instead of the electron-phonon interaction.

A well-known electronic mechanism for superconductivity has been introduced by Kohn and Luttinger (KL) [15], in which an effective attraction arises from dynamical screening of the bare repulsion by the highly polarizable electronic Fermi sea [16,17]. The KL theory is analytically controlled in the weak-interaction regime, and yields a superconducting transition temperature T_c that is orders of magnitude smaller than the Fermi temperature T_F. In contrast, the experimentally determined T_c/T_F in WTe_2, ZrNCl, and magic-angle graphene is remarkably high and reaches up to ∼0.1, calling for an electronic mechanism of *strong-coupling* superconductivity.

Strong-coupling superconductivity can arise from repulsive interactions, as a result of enhanced fluctuations near insulating ordered phases driven by these interactions. This picture has been extensively discussed in connection with the superconducting cuprates [18], pnictides [19–21], and also twisted bilayer graphene and other graphitic systems [22–32]. Superconductivity can also arise from repulsive interactions through virtual transition of electron pairs from the Fermi surface to distant or incipient bands [33,34].

Recently, a new mechanism for unconventional superconductivity from repulsive interactions in doped insulators has been proposed, where the attraction between two conduction electrons arises from virtual interband transition of a third electron in the filled band [35–37]. The pairing interaction obtained by integrating out high-energy interband excitations is nonretarded and acts on all conduction electrons rather than a small fraction near the Fermi level. As a result, the gap-T_c ratio $\Delta/(k_B T_c)$ is significantly higher than the standard value for electron-phonon superconductors, which clearly indicates unconventional strong-coupling superconductivity. This "three-particle mechanism" for superconductivity was rigorously demonstrated with a hybridization expansion method [35], which is *nonperturbative* in interaction strength and provides an analytically controlled theory of strong-coupling superconductivity at low carrier density.

In this work, we present a comprehensive study of the emergence of superconductivity from repulsive interactions in a two-band Hubbard model of spinless fermions introduced by two of us in Ref. [35]. By combining a variety of analytical and numerical methods, our study covers a wide range of parameters and electron fillings. We show that, unlike the Kohn-Luttinger mechanism or fluctuation-induced superconductivity discussed above, the screened interaction, when projected onto the states near the Fermi surface, becomes *attractive* for small momentum transfers. As a result, a nodeless order parameter, similar to the the one induced by phonons, is possible. The existence of this attractive interaction comes from interband screening effects of the bare repulsion, and its effect can be understood in terms of a new interaction, correlated hopping [35] (also referred to as electron-assisted hopping), between conduction electrons. Focus is not only given to the physical origin for electron pairing, but also to the methodology allowing us to controllably capture this pairing in various parameter regimes. We finally discuss the

connections of this type of superconductivity with the phase diagram of twisted bilayer graphene [38–40].

II. ILLUSTRATIVE MODEL AND OUTLINE

We consider spinless fermions on the honeycomb lattice with nearest-neighbor (NN) tunneling t, sublattice potential difference Δ, and NN repulsion V,

$$\mathcal{H} = \Delta \sum_{r \in B} n_r - t \sum_{\langle r, r' \rangle} (c_r^\dagger c_{r'} + c_{r'}^\dagger c_r) + V \sum_{\langle r, r' \rangle} n_r n_{r'}, \quad (1)$$

above unit filling, i.e., for $n = 1 + x$ electrons per unit cell on average, with $0 < x \leqslant 1$. Reference [35] rigorously proved that this model hosts a superconducting phase at low doping above unit filling, assuming $t \ll \Delta$. One goal of this work is to demonstrate the emergence of attractive interaction from bare repulsion through interband polarization in a wide range of dopings x and model parameters t/Δ, V/Δ. Our work not only extends the results of Ref. [35] (obtained from kinetic energy expansion) beyond the regime $t \ll \Delta$, but also elucidates the origin of electron pairing from a band picture viewpoint.

To achieve this goal, we first provide, for any parameter set (t, V, Δ), at least one analytically exact result establishing the presence of attractive interactions between conduction electrons at small doping concentrations. When $t \ll \Delta$ or $V \ll \Delta$, we respectively rely on the previously mentioned hybridization expansion or on standard many-body perturbation theory. Away from these perturbative regimes, we can still obtain analytical results if we extend Eq. (1) to the case where fermions carry an additional flavor index $\sigma = 1, \dots, N_f$:

$$\mathcal{H} = \Delta \sum_{r \in B} n_r - t \sum_{\langle r, r' \rangle, \sigma} (c_{r,\sigma}^\dagger c_{r',\sigma} + hc) + \frac{V}{N_f} \sum_{\langle r, r' \rangle} n_r n_{r'}, \quad (2)$$

where $n_r = \sum_\sigma n_{r,\sigma}$ now denotes the total density at site r. Note that we have scaled the NN interaction strength to keep the same balance between kinetic and interaction energies at unit filling $n = 1$, which now corresponds to N_f fermions per unit cells. When the number of fermionic flavors goes to infinity ($N_f \gg 1$), the random phase approximation (RPA) provides exact results on the pairing of electrons in the conduction band, for any values of the ratios t/Δ or V/Δ.

This enterprise is the focus of Secs. III, IV, and V, where we derive effective models for doped charges and prove the existence of attractive interaction between them due to the interband screening of the bare repulsion. The results obtained in the three regimes of Fig. 1 agree in their overlapping region of validity, which provides a stringent test of our calculations and confirms the emergence of superconductivity in our repulsive model.

We then numerically study the superconducting properties of Eq. (1) as a function of doping in Sec. VI. Our simulations include both Hartree-Fock (HF) corrections and dynamical screening effects within the RPA. They show a robust superconducting phase with dome-shaped T_c as a function of doping and a nodeless order parameter for $x < 25\%$. Finally, in Sec. VII, we highlight features that transcend our specific model, and discuss their relevance for the superconducting phase of twisted bilayer graphene. Section IX summarizes and closes the discussion.
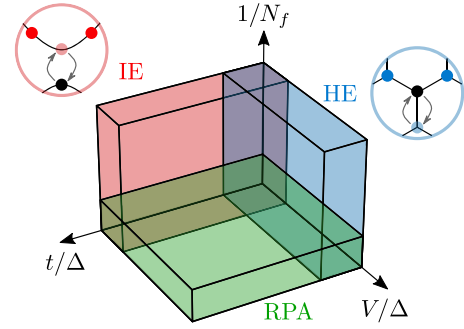

FIG. 1. Domains of validity of the different analytical methods. The hybridization expansion (HE), interaction expansion (IE), and random phase approximation (RPA) have overlapping regions of validity, allowing for stringent consistency checks. The insets depict the virtual particle-hole processes responsible for pairing, in real space for the HE ($t \ll \Delta$) and momentum space for the IE ($V \ll \Delta$).

III. HYBRIDIZATION EXPANSION

We start with the weak-tunneling limit $t \ll \Delta$, and briefly summarize the hybridization expansion introduced in Refs. [35,36], to which we refer for more details.

A. Effective model for doped charges

At unit filling and for $t = 0$, the system forms a charge density wave in which A sites are occupied with one fermion of each flavor, while the B sites remain empty. States with holes on A sites lie higher in energy due to the large single-particle gap Δ. Due to the Pauli principle, the x charges added above $n = 1$ must live on B sites. For the same reason, direct tunneling of these doped charges to filled adjacent sites is forbidden, and the doped charges' dynamics involves multiple tunneling processes with a high-energy intermediate state having holes on A sites.

Using a unitary transformation, we can trace over these high-energy excitons to obtain an effective Hamiltonian for the doped charges on the B triangular lattice, thereby including A-B hybridization effects perturbatively in t/Δ. To second order, this hybridization expansion (HE) yields

$$\mathcal{H}^{HE} = \sum_{ijk \in \triangle, \sigma} [c_{i,\sigma}^\dagger T(n_\triangle) c_{j,\sigma} + P_{ijk}] + E(n_\triangle), \quad (3)$$

where the sum runs over upper triangles $\triangle$ of the B lattice (see Appendix A). In Eq. (3), $n_\triangle = n_i + n_j + n_k$ is the number of fermions in the triangle $\triangle$, and P_{ijk} describes summation over all possible permutations of the indices ijk. We have also defined the functions

$$T(n) = \frac{t^2}{\Delta + [3 - (2 + n)/N_f]V}, \quad (4a)$$

$$E(n) = \frac{-t^2(3N_f - n)}{\Delta + [3 - (1 + n)/N_f]V}, \quad (4b)$$

which contain the typical energy difference denominators of second-order perturbation theory.

At low density, we can simplify these coefficients assuming that no more than two doped charges simultaneously occupy the same triangle:

$$T(n) \simeq T(0) + n[T(1) - T(0)] + \cdots, \tag{5a}$$

$$E(n) \simeq E(0) + n[E(1) - E(0)]$$
$$+ \frac{n(n-1)}{2}[E(2) - 2E(1) + E(0)] + \cdots. \tag{5b}$$

Defining $t_f = T(0)$, $\lambda = T(1) - T(0)$, and $V_f = E(2) - 2E(1) + E(0)$, we can rewrite the effective Hamiltonian for moderate doping as

$$\mathcal{H}^{\mathrm{HE}} \simeq \sum_{ijk \in \triangle, \sigma} [c_{i,\sigma}^{\dagger}(t_f + \lambda n_{\triangle})c_{j,\sigma} + P_{ijk}]$$
$$+ 3V_f \sum_i \frac{n_i(n_i - 1)}{2} + V_f \sum_{\langle i,j \rangle_B} n_i n_j. \tag{6}$$

Explicit evaluation of the parameters in this Hamiltonian gives

$$t_f = \frac{t^2}{\Delta + [3 - 2/N_f]V}, \tag{7a}$$

$$\frac{\lambda}{t_f} = \frac{V/N_f}{\Delta + 3(1 - 1/N_f)V}, \tag{7b}$$

$$\frac{V_f}{t_f} = \frac{2(V/N_f)(\Delta - V/N_f)}{[\Delta + (3 - 1/N_f)V][\Delta + 3(1 - 1/N_f)V]}. \tag{7c}$$

They agree with the expression given in Refs. [35] and [36] for $N_f = 1$ and $N_f = 2$, respectively.

B. Dilute limit

In the dilute limit, doped electrons mostly occupy states at the band bottom and their physics is governed by long-wavelength properties that transcend details on the lattice scale. To derive this continuum description, we project out the fermionic modes with large kinetic energy and retain only those near the band minima.

The effective band dispersion of Eq. (6) reads $\varepsilon_k = 2t_f \sum_i \cos(k \cdot a_j)$, with a_j the three primitive lattice vectors, and displays two degenerate minima at the K and $K' = -K$ points of the Brillouin zone. At low energy, doped electrons therefore acquire an additional isospin degree of freedom τ, which distinguishes electrons near K from those close to K'; i.e., the resulting continuum theory involves $2N_f$ fermionic fields $\psi_{q,\sigma,\tau} = c_{\tau K + q, B, \sigma}$, with $|qa| \ll 1$ and a the lattice constant. Near its minima, ε_k is quadratic and isotropic, giving the kinetic part of the continuum model

$$\tilde{\mathcal{H}}_{\mathrm{kin}} = \sum_{k,\sigma,\tau} \frac{|k|^2}{2m} \psi_{k,\sigma,\tau}^{\dagger} \psi_{k,\sigma,\tau}, \quad m = \frac{\Delta + (3 - 2/N_f)V}{3(at)^2}. \tag{8}$$

The interaction part of the continuum model is made of the only three symmetry-allowed contact terms—intravalley and intervalley density interactions together with valley-isospin

exchange—which can be written as

$$\tilde{\mathcal{H}}_{\mathrm{int}} = \frac{1}{N} \sum_{q,\sigma,\sigma',\tau=\pm K} g_0 \rho_{q,\sigma,\tau} \rho_{-q,\sigma',\tau}$$
$$+ \frac{1}{N} \sum_{q,\sigma,\sigma'} g_1 \rho_{q,\sigma,K} \rho_{-q,\sigma',K'} + g_2 \boldsymbol{\tau}_{q,\sigma} \cdot \boldsymbol{\tau}_{-q,\sigma'}, \tag{9}$$

where we have defined the density and valley-isospin operators

$$\rho_{q,\sigma,\tau}(q) = \sum_k \psi_{k+q,\sigma,\tau}^{\dagger} \psi_{k,\sigma,\tau}(q), \tag{10a}$$

$$\boldsymbol{\tau}_{q,\sigma}(q) = \sum_{k,\alpha,\beta} \psi_{k+q,\sigma,\alpha}^{\dagger} [\boldsymbol{\tau}]_{\alpha,\beta} \psi_{k,\sigma,\beta}, \tag{10b}$$

with $\boldsymbol{\tau}$ the set of Pauli matrices. Using the interaction derived in Eq. (6), we obtain the following g coefficients (see Appendix B):

$$g_0 = 9(V_f - 2\lambda)/2, \quad g_1 = 2g_0, \quad g_2 = 0. \tag{11}$$

This drastically simplifies the interacting part of the effective continuum theory, which now simply reads

$$\tilde{\mathcal{H}} = \sum_{k,\sigma,\tau} \frac{|k|^2}{2m} \psi_{k,\sigma,\tau}^{\dagger} \psi_{k,\sigma,\tau} + \frac{g_0}{N} \sum_q \rho_q \rho_{-q}, \tag{12}$$

with $\rho_q = \sum_{\tau,\sigma} \rho_{q,\sigma,\tau}$ the total density operator. We remark that the original SU(N_f) flavor symmetry of the model has been promoted to an enlarged SU($2N_f$) symmetry in Eq. (12), covering both flavor and the emergent valley degrees of freedom. This is due to the special form of interactions in our model, which does not include any on-site interactions. As shown in Ref. [36], such on-site repulsion would break the valley isospin rotation symmetry and produce an additional valley-ferromagnetic coupling.

C. Attraction

Remarkably, $\tilde{\mathcal{H}}$ describes a dilute Fermi gas with *attractive* interactions, as can be seen from the sign of the only relevant interaction coefficient,

$$g_0 = -\frac{27(tV)^2}{N_f P} < 0, \quad P = \prod_{k=1}^{3}[\Delta + (3 - k/N_f)V]. \tag{13}$$

We stress that this result holds for any values of N_f and is nonperturbative in the interaction parameter V/Δ. Even more importantly, this effective attractive interaction appears for two doped electrons in a band insulator. This strikingly contrasts our approach to other mechanisms for superconductivity (such as the KL mechanism) that rely on the presence of a Fermi surface. It also lead to distinctive predictions, such as a BEC-BCS crossover upon increasing the doping concentration, which has recently been observed in ZrNCl [7,36].

For $N_f = 1$, the continuum Hamiltonian for a low density of doped electrons takes the form of a pseudospin-$\frac{1}{2}$ Fermi gas with local *attractive* interaction, which is known to exhibit BEC-BCS crossover. On the BCS side and for low densities, the critical temperature satisfies

$$k_B T_c \propto \sqrt{W E_F} e^{-1/(2g)}, \tag{14}$$

where E_F denotes the Fermi energy, W the conduction bandwidth, and $g = 6V^2/[\pi \Delta(\Delta + 2V)]$ a dimensionless coupling constant [35]. Note that g is nonperturbative in the interaction V, and it does not have to be small. This allows $k_B T_c/E_F$ ratios as high as 10%. Even for small g, since the nonretarded attraction spreads over the entire conduction band, the gap-T_c ratio $\Delta/(k_B T_c)$ is 4.8, which far exceeds the standard value 1.57 for electron-phonon superconductors. Both behaviors, obtained by exact solution in the small-t/Δ and low-doping regime, clearly indicate strong-coupling superconductivity in our model [35].

It is also worth comparing Eq. (14) with the KL result for the two-dimensional repulsive Fermi gas [41,42]: $T_c \propto \exp(-1/U_{\rm eff})$, where $U_{\rm eff} \propto U^3$ and U is the strength of bare repulsion. This formula only holds at small U; hence T_c is bound to take values orders of magnitude below the ones of Eq. (14).

The square root dependence of T_c with respect to Fermi energy or doping concentration, shown in Eq. (14), may seem unusual especially given that the density of states near the band edge is constant. It originates from the nonretarded attraction between conduction electrons in two dimensions. In 2D, two-particle bound states exist for arbitrarily weak attraction, and the pair binding energy ε_b is proportional to the bandwidth, the natural cutoff of the system. At finite density, provided that the Fermi energy E_F is large compared to ε_b (the BCS regime), both T_c and the zero-temperature superconducting gap are proportional to $\sqrt{E_F \varepsilon_b}$, as shown by mean-field treatment detailed in Appendix C.

IV. INTERACTION EXPANSION

We now turn to the limit of weak interactions $V \ll \Delta$, for which electrons in the conduction band and holes in the valence band are only mixed slightly. Their small admixture, produced by off-diagonal interactions in the band basis (as opposed to sublattice basis), can be accounted for with standard many-body perturbation theory, as detailed in this section.

A. Notations

We rewrite the model Eq. (2) in momentum space,

$$\mathcal{H} = \sum_{\substack{12 \\ \sigma}} c_{2,\sigma}^\dagger h_{2,1} c_{1,\sigma} + \frac{1}{2N} \sum_{\substack{1234 \\ \sigma,\sigma'}} c_{4,\sigma}^\dagger c_{3,\sigma'}^\dagger \Gamma_{43,21} c_{2,\sigma'} c_{1,\sigma}, \quad (15)$$

with N the number of unit cells, and $i = (k_i, s_i)$ a generalized index gathering the single-particle momentum k_i, belonging to the Brillouin zone (BZ), and the sublattice index $s_i \in \{A, B\}$. Due to momentum conservation, and because our model only features density-density interactions, the parameters of the Hamiltonian can be simplified as

$$h_{2,1} = \delta_{k_1,k_2} h_{s_2,s_1}(k_1),$$
$$\Gamma_{43,21} = \delta_{s_4,s_1} \delta_{s_3,s_2} \delta_{k_1+k_2,k_3+k_4} \Gamma_{s_2,s_1}(q), \quad (16)$$

with $q = k_4 - k_1 = k_2 - k_3$ the momentum exchanged, and where we have defined the 2×2 matrices

$$h(q) = \begin{bmatrix} -\Delta/2 & -tf(q) \\ -tf^*(q) & \Delta/2 \end{bmatrix}, \quad \Gamma(q) = \frac{V}{N_f}\begin{bmatrix} 0 & f(q) \\ f^*(q) & 0 \end{bmatrix}. \quad (17)$$

We have used $f(q) = \sum_{j=1}^3 e^{i(q \cdot u_j)}$, with $u_{j=1,2,3}$ the vectors connecting B sites to their three nearest neighbors (see Appendix A). The one-body part h gives two bands with dispersion

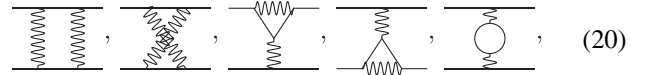$$\varepsilon_{q,n} = n\sqrt{(\Delta/2)^2 + |tf(q)|^2}, \quad (18)$$

where $n = \pm$ for the valence and conduction band, respectively. As in Sec. III, the conduction band displays degenerate minima at the K and K' points. The corresponding Bloch wave functions are

$$\Psi_{q,n} = \frac{1}{\sqrt{2\varepsilon_{q,+}(\varepsilon_{q,+} + n\Delta/2)}}\begin{bmatrix} -ntf(q) \\ \varepsilon_{q,+} + n\Delta/2 \end{bmatrix}, \quad (19)$$

whose upper/lower component corresponds to the wave function's amplitude on each sublattice $\Psi_{q,n}^A/\Psi_{q,n}^B$.

B. Kohn-Luttinger diagrams

Renormalization of the two-body scattering vertex is described, to second order in many-body perturbation theory, by the five diagrams [15]

$$(20)$$

where curvy and straight lines respectively denote interaction events and single-particle propagators. The evaluation of these diagrams is detailed in Appendix D, and produces a renormalized interaction vertex $\Gamma^{\rm IE} = \Gamma + \delta\Gamma$, with

$$\delta\Gamma_{43,21} = -\frac{1}{N}\sum_{abcd} \chi_{dc,ba}^- \Gamma_{43,ba} \Gamma_{dc,21}$$
$$+ \chi_{dc,ba}^+ [\Gamma_{4c,2a}\Gamma_{d3,b1} + \Gamma_{c3,2a}\Gamma_{d4,b1}$$
$$+ \Gamma_{4c,a1}\Gamma_{d3,b2} - N_f \Gamma_{4c,a1}\Gamma_{d3,2b}], \quad (21)$$

where the terms are ordered as in Eq. (20). We recall that the sum over the generalized indices (a, b, c, d) runs over all single-particle momenta $(k_a, k_b, k_c, k_d) \in$ BZ and orbital indices $(s_a, s_b, s_c, s_d) \in \{A, B\}$ that characterize the single-particle propagators through the Bloch eigenvectors $\Psi_{k_i,n}^{s_i}$. Finally, we have denoted as χ^- and χ^+ the particle-particle and particle-hole susceptibilities, respectively. They assume the explicit form

$$\chi_{dc,ba}^{\epsilon=\pm} = \delta_{k_a}^{k_d}\delta_{k_b}^{k_c} \sum_{n,n'=\pm} \Psi_{k_d,n}^{s_d *}\Psi_{k_c,n'}^{s_c *}\Psi_{k_b,n'}^{s_b}\Psi_{k_a,n}^{s_a}$$
$$\times \frac{f_\beta(\epsilon\xi_{k_a,n}) - f_\beta(\xi_{k_b,n'})}{\xi_{k_a,n} - \epsilon\xi_{k_b,n'}}, \quad (22)$$

where $f_\beta(x) = 1/[1 + e^{\beta x}]$ is the Fermi-Dirac distribution, and $\xi_i = \varepsilon_i - \mu$ measures energies with respect to the chemical potential μ.

The diagrams of Eq. (20) were first used by Kohn and Luttinger to show how attractive interactions could be produced in a fully repulsive Fermi sea, as a consequence of the dynamical screening of the bare repulsion by the Fermi sea [15]. This picture relies on a large particle-hole susceptibility near $2k_F$, which requires a fully formed Fermi sea. At infinitesimal doping however, this condition is not satisfied and

the susceptibility almost vanishes, leading to a breakdown of the KL mechanism.

Here, in contrast, the particle-hole susceptibility remains finite at low densities, and is dominated by interband contributions. To see this, consider a situation where the chemical potential is positioned right at the conduction band bottom, such that $n = 1 + x$ with $x \ll 1$. At low temperature, the valence band is completely filled $f_\beta(\xi_{k,-}) = 1$ and the conduction band barely populated $f_\beta(\xi_{k,+}) \simeq 0$. This forces the band indices in the particle-hole susceptibility to be opposite, thereby highlighting the crucial role of virtual interband transitions.

C. Pairing in the dilute limit

As in Sec. III, we unveil the low-density properties of the model with an effective continuum field theory that only retains the modes near the conduction band minima. Following lines similar to those above (see Appendix B), we end up with the continuum model of Eqs. (8) and (9), albeit with a new mass $m = \Delta/[3(at)^2]$, which agrees with Eq. (8) when $V \ll \Delta$, and new g coefficients

$$g_0 = \Lambda_{K,K}(0) - \Lambda_{K,K}(K)/2, \tag{23a}$$

$$g_1 = 2\Lambda_{K,K'}(0) + \Lambda_{K,K}(K), \tag{23b}$$

$$g_2 = 2\Lambda_{K,K}(K), \tag{23c}$$

where Λ denotes the effective interaction in the band basis

$$\Lambda_{k,k'}(q) = \frac{1}{2} \sum_{a,b=A/B} \Psi_{k+q,+}^{b*} \Psi_{k',+}^{a*} \Gamma_{a,b}^{\mathrm{IE}}(q) \Psi_{k'+q,+}^{a} \Psi_{k,+}^{b}. \tag{24}$$

A few simple observations allow us to greatly simplify the expressions of the g coefficients. First, electrons near the conduction band bottom (at K and K') have a strong B character since $\Psi_{\pm K,+}^A = 0$. Because all the diagrams in the g's have incoming/outgoing electrons with momentum $\pm K$, this forces all external legs in these diagrams to hold a B sublattice index. Furthermore, the bare propagator does not have any direct B-B interaction [$\Gamma_{B,B}(k) = 0$]. Together, these observations already rule out the presence of first-order interactions at the conduction band bottom, requiring us to go to second order and evaluate the diagrams of Eq. (20).

Due to the special form of interactions [Eq. (16)], the orbital index can only change along the single-particle propagators (straight lines) of these diagrams. This readily shows that the first four KL diagrams vanish in the dilute limit. To see this, we can draw all diagrams compatible with the above orbital rules, using red/green for A/B indices,
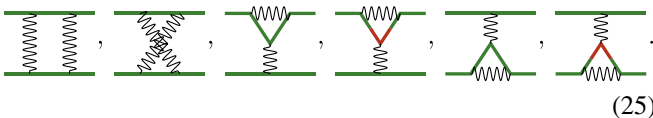
$$\tag{25}$$

They all feature at least one B-B bare interaction, reducing them to zero. As a result, only the bubble polarization diagram,
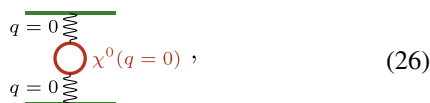
$$\tag{26}$$

remains in the weakly interacting and low-density limit. Its interaction vertices can, in principle, involve the momentum exchange $q = 0$ or $q = \pm K$. However, because the bare interaction vanishes for $\Gamma_{A,B}(\pm K) = V f(\pm K)/N_f = 0$, only the diagram with $q = 0$, shown in Eq. (26), is nonvanishing. We extend the discussion on the contributions of the five KL diagrams in the low-density limit in Appendix E.

This shows that $\Lambda_{K,K}(K) = 0$, implying $g_2 = 0$. Because the bubble contribution only depends on the momentum exchanged q, we find $\Lambda_{K,K}(0) = \Lambda_{K,K'}(0)$ and therefore $g_1 = 2g_0$. As a consequence, the effective continuum theory is the same as Eq. (12). The sign of $g_0 = \Lambda_{K,K}(0) = \Gamma_{B,B}^{\mathrm{IE}}(0)/2 = \delta\Gamma_{B,B}^{\mathrm{bbl}}(0)/2$ determines the nature of interactions, attractive or repulsive.

To distinguish between these two possibilities, we evaluate the contribution of the bubble diagram in Eq. (26). We recall that the interactions of our original model can be written as $\mathcal{H}_{\mathrm{int}} = (2N)^{-1} \sum_{s_1,s_2 \in \{A,B\}} \Gamma_{s_2,s_1}(q) n_{s_1}(q) n_{s_2}(-q)$ with Γ_{s_1,s_2} a 2×2 matrix. The bubble contribution preserves this matrix structure, yielding

$$\delta\Gamma_{s_2,s_1}^{\mathrm{bbl}}(q) = N_f \sum_{s_a,s_b} \Gamma_{s_2,s_b}(q) \chi_{s_b,s_a}^0(q) \Gamma_{s_a,s_1}(q), \tag{27}$$

with

$$\chi_{s',s}^0(q) = \frac{1}{N} \sum_k \chi_{(k,s')(k+q,s),(k+q,s')(k,s)}^+. \tag{28}$$

More succinctly, it can be written as a matrix product $\delta\Gamma^{\mathrm{bbl}}(q) = N_f \Gamma(q) \chi^0(q) \Gamma(q)$. As described above, in the dilute limit, the bubble polarization χ^0 is dominated by interband contributions, which give

$$\chi \equiv \chi_{A,A}^0(0) = -\chi_{A,B}^0(0) = -\frac{1}{N} \sum_k \frac{|tf(k)|^2}{4\varepsilon_{k,+}^3}, \tag{29}$$

and $\chi_{B,B}^0(0) = \chi_{A,A}^0(0)$, $\chi_{A,B}^0(0) = \chi_{A,B}^0(0)$.

Putting all pieces together, we find—as originally announced—an effective *attractive* interaction in the dilute limit

$$g_0 = -\frac{9V^2|\chi|}{2N_f}. \tag{30}$$

As a consistency check of this result, we can apply it to the case $t \ll \Delta$ where our hybridization expansion result also holds. In that limit, we use the identity $N^{-1} \sum_k |f(k)|^2 = 3$ to get $g_0 \simeq -27(tV)^2/[N_f \Delta^3]$, which perfectly matches Eq. (13) when $V \ll \Delta$ (for which $P \simeq \Delta^3$).

V. RANDOM PHASE APPROXIMATION

The hybridization and interaction expansions, which are real-space and momentum-space based methods, respectively, allow us to solve our extended Hubbard model in different parts of the parameter space. We can further bridge between the two approaches and obtain analytical results for any values of t, Δ, and V in the limit where $N_f \to \infty$ using the random phase approximation (RPA).
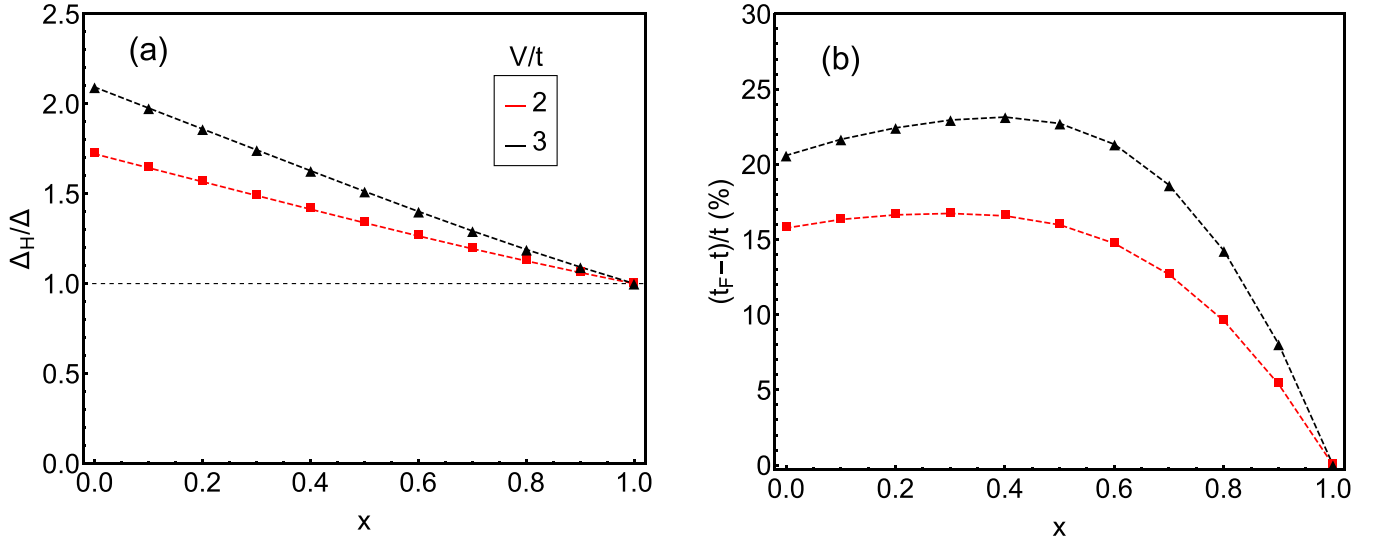
FIG. 2. (a) $\Delta_{\mathrm{H}}/\Delta$ and (b) the relative deviation $(t_{\mathrm{F}} - t)/t$ obtained for $\Delta = 8t$ and $V/t = 2, 3$, as a function of doping.

A. Hartree-Fock and RPA diagrams

The essence of RPA lies in the possibility to exactly sum the entire series of diagrams that dominate in the $N_f \gg 1$ limit. Such dominating diagrams maximize the number of bubbles for a given number of interaction vertices because, for flavor-conserving interactions, each bubble multiplies its diagram's contribution by N_f, while other configurations yield smaller combinatorial factors. We have already encountered this feature in Eq. (21), where only the bubble diagram is multiplied by a factor N_f. Generalizing to all orders, the only relevant diagrams in the $N_f \to \infty$ limit are

$$\overline{\underset{\underline{\hspace{0.5cm}}}{\gtrless}} \, , \, \overline{\underset{\underline{\hspace{0.5cm}}}{\bigcirc}} \, , \, \overline{\underset{\underline{\hspace{0.5cm}}}{\bigcirc\!\!\sim\!\!\bigcirc}} \, , \, \overline{\underset{\underline{\hspace{0.5cm}}}{\bigcirc\!\!\sim\!\!\bigcirc\!\!\sim\!\!\bigcirc}} \, , \, \cdots . \tag{31}$$

This diagrammatic series can be recast in a geometric form, enabling its exact summation. To see this, recall that the one-bubble correction takes the simple matrix product form $\delta\Gamma^{\mathrm{bbl}}(q) = N_f \Gamma(q)\chi^0(q)\Gamma(q)$. Its obtaining can be generalized to the m-bubble diagram, giving $(N_f \Gamma \chi^0)^m \Gamma$. Summing all these corrections yields the renormalized RPA scattering vertex

$$\Gamma^{\mathrm{RPA}}(q) = [1 - N_f \Gamma(q)\chi^0(q)]^{-1}\Gamma(q), \tag{32}$$

which is valid when $N_f \gg 1$. Performing the matrix inversion explicitly, we get

$$\Gamma^{\mathrm{RPA}} = \frac{1}{D}\left(\Gamma + \frac{|Vf|^2}{N_f}\begin{bmatrix} \chi^0_{BB} & -\chi^0_{AB} \\ -\chi^0_{BA} & \chi^0_{AA} \end{bmatrix}\right), \tag{33}$$

with $D = (1 - Vf\chi^0_{BA})(1 - Vf^*\chi^0_{AB}) - |Vf|^2\chi^0_{AA}\chi^0_{BB}$. For convenience of notation, the q dependence is not explicitly shown.

Having described, within the RPA, the renormalization of the interaction kernel, we now turn to the renormalization of the single-particle Green's function by interactions. These corrections come from the Hartree and Fock diagrams, and can be efficiently accounted for by a redefinition of the sublattice potential difference $\Delta \to \Delta_{\mathrm{H}}$ and of the NN tunneling amplitude $t \to t_{\mathrm{F}}$, with the value of Δ_{H} and t_{F} determined

self-consistently (see Appendix F). In the large-N_f limit, the Hartree correction dominates and we find $\Delta_{\mathrm{H}} \simeq \Delta + 3V$ and $t_{\mathrm{F}} \simeq t$. Away from this analytical limit, we obtain Δ_{H} and t_{F} numerically. As an illustration, Fig. 2 shows these two quantities obtained for $\Delta = 8t$ and $V/t = 2, 3$ as a function of doping. We observe that Δ_{H} approaches linearly its bare value Δ as $x \to 1$, where the Hartree-Fock corrections vanish since both the A and B are completely filled. In addition, we find a nonzero doping-dependent correction to the hopping amplitude, t_{F}, up to 20%–25% of its bare value, t, for $V = 3t$.

B. Pairing in dilute limit

The argument given in Sec. IV can be repeated to find that the dilute effective model in the RPA regime also has $g_2 = 0$ and $g_1 = 2g_0$, such that it is described by Eq. (12). Thus, the sign of g_0 determines the nature of interaction. This coefficient takes the form

$$g_0 = \Gamma^{\mathrm{RPA}}_{B,B}(0) = \frac{9V^2}{N_f}\frac{\chi}{1 + 6V\chi}, \tag{34}$$

and indeed describe an attractive interaction since $\chi < 0$.

Our different analytical methods prove the presence of an effective attraction and the formation of bound electronic pairs in all the regimes highlighted in Fig. 1 for low enough densities. The reliability of our approach can be demonstrated by comparing the result of all three approaches in the joint region of applicability. It is clear that Eq. (34) reduces to the results of the interaction expansion when $V \ll \Delta$, since they include the exact same diagrams. We now compare the RPA and HE results assuming both $N_f \gg 1$ and $t \ll \Delta$. In that case, the Hartree-Fock coefficients read $\Delta_{\mathrm{H}} \simeq \Delta + 3V$ and $t_{\mathrm{F}} \simeq t$ (Appendix F) and the conduction band is almost flat $\varepsilon_{k,+} \simeq \Delta_{\mathrm{H}}/2$. Finally, using $\sum_k |f(k)|^2 = 3N$, the RPA result can be written as

$$g_0 \simeq -\frac{27t^2V^2}{N_f\Delta_{\mathrm{H}}^3} = -\frac{27t^2V^2}{N_f(\Delta + 3V)^3}, \tag{35}$$

which perfectly agrees with the HE result of Eq. (13), because $P \simeq (\Delta + 3V)^3$ when $N_f \gg 1$. This provides a stringent test of our methods, and rigorously demonstrates the presence of pairing between doped charges in our model.

VI. NUMERICAL RESULTS

To go beyond infinitesimal doping and reach regimes where none of the perturbative parameters t/Δ, V/Δ, or $1/N_f$ are small, we now perform extensive numerical simulations of the original model Eq. (1). We include Hartree-Fock corrections and the repulsive interaction screened by particle-hole excitations described by $\Gamma^{\mathrm{RPA}}(q)$ for $N_f = 1$.

The method then assumes the superconductivity to be driven by the charge fluctuations and is analogous to the one that two of us used recently for studying the origin of superconductivity in the twisted bi- and trilayer graphene [39,40], as well in the rhombohedral trilayer graphene [43]. Within this framework, the linearized gap equation is

$$\Delta_{n,k} = -\frac{1}{N} \sum_{k',n'} V_{n',k'}^{n,k} \frac{\tanh(\beta\xi_{n',k'}/2)}{2\xi_{n',k'}} \Delta_{n',k'},$$

$$V_{n',k'}^{n,k} = \sum_{a,b=A,B} \Gamma_{a,b}^{\mathrm{RPA}}(k-k') \Psi_{k,n}^{a,*} \Psi_{k',n'}^{a} \Psi_{k,n}^{b} \Psi_{k',n'}^{b,*}. \quad (36)$$

For convenience, we rewrite Eq. (36) in terms of a Hermitian operator, as

$$\tilde{\Delta}_{n,k} = -\sum_{k',n'} \tilde{\Gamma}_{(n,k),(n',k')} \tilde{\Delta}_{n',k'}, \quad (37)$$

where

$$\tilde{\Delta}_{n,k} \equiv \Delta_{n,k} \times \sqrt{\frac{\tanh(\beta\xi_{n,k}/2)}{2\xi_{n,k}}},$$

$$\tilde{\Gamma}_{(n,k),(n',k')} = \frac{1}{N}\sqrt{\frac{\tanh(\beta\xi_{n,k}/2)\tanh(\beta\xi_{n',k'}/2)}{4\xi_{n,k}\xi_{n',k'}}} V_{n',k'}^{n,k}. \quad (38)$$

The critical temperature, T_c, is defined as the largest value of T such that the kernel $\tilde{\Gamma}$ has the eigenvalue 1, in order for Eq. (37) to have solutions. The corresponding eigenvector provides the symmetry of the superconducting order parameter. In general, Eq. (37) admits both symmetric (A_1 irrep of the lattice point group symmetry) and antisymmetric (A_2 irrep) solutions upon exchanging $k \to -k$. However, only the latter satisfies the Pauli principle imposed by the statistics of spinless fermions. The A_1 solutions being unphysical, we discard them by projecting Eq. (37) on the A_2 subspace:

$$\tilde{\Delta}_{n,k} = -\sum_{k',n'} \tilde{\Gamma}_{(n,k),(n',k')}^{A_2} \tilde{\Delta}_{n',k'}, \quad (39)$$

with

$$\tilde{\Gamma}_{(n,k),(n',k')}^{A_2} \equiv \frac{\tilde{\Gamma}_{(n,k),(n',k')} - \tilde{\Gamma}_{(n,k),(n',-k')}}{2}. \quad (40)$$

We now present our numerical solutions of Eq. (39), with the band structure and the eigenfunctions accounting for the HF corrections (see Fig. 2). We sample the BZ with $N = 2700$ points, set $\Delta = 8t$, and consider different values of V upon varying the doping x. Figure 3 shows T_c as function of
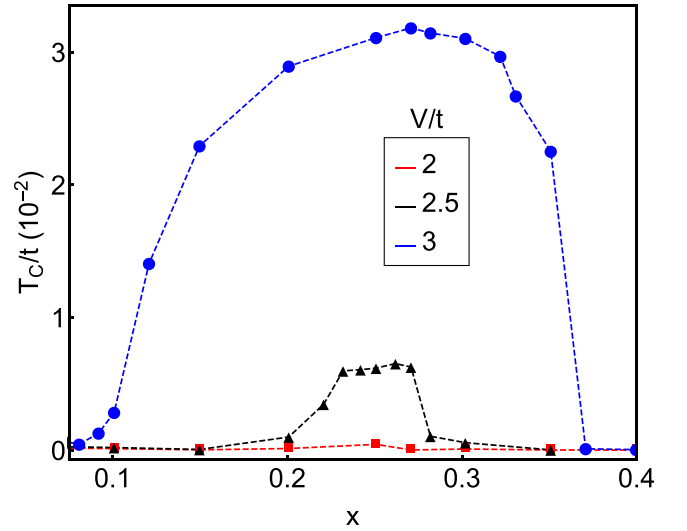

FIG. 3. Values of the critical temperature as a function of doping, obtained by solving numerically Eq. (39), for $\Delta = 8t$ and $V/t = 2, 2.5, 3$. We use $N = 2700$ unit cells and include the HF corrections to the band structure and the eigenfunctions.

x for $V/t = 2, 2.5, 3$. The critical temperature displays a non-monotonic behavior, and reaches a maximum at an optimal doping $x \simeq 0.25$. As expected when the superconductivity is driven by electronic interactions, T_c increases with the interaction's strength V, reaching values up to $\sim(3\text{–}3.5) \times 10^{-2}t$ for $V/t = 3$.

Figure 4 shows the symmetry of the superconducting order parameter in the BZ, obtained for $\Delta = 8t$, $V = 3t$, and four different values of the doping. The continuum black lines represent the Fermi surface. Comparing Figs. 4 and 3, we note
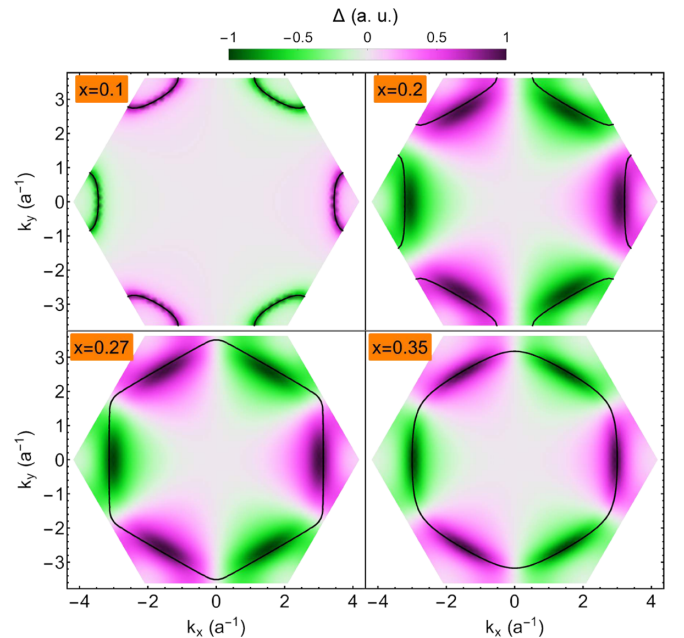

FIG. 4. Symmetry of the superconducting order parameter in the BZ, obtained for $\Delta = 8t$, $V = 3t$, and four different values of doping. The continuum black lines represent the Fermi surface.

that T_c increases with the size of the Fermi surface. The latter is a nonmonotonic function of the doping, increasing with x up to $x \sim 0.26$ and then shrinking around the center of the BZ. We also remark that T_c falls down abruptly when $x \gtrsim 0.26$, which can be understood from the emergence of nodes in the superconducting order parameter, as can be seen in Fig. 4.

VII. DISCUSSION

A. Interband polarization

Our analytic arguments establishing the presence of attraction and the formation of bound pairs in the low carrier density limit (see Sec. III) offer a new perspective from many-body perturbation theory on the fully electronic mechanism for superconductivity introduced in Ref. [35]. More precisely, the pairing interaction emerges from the polarization bubble Eq. (26), which is dominated by interband contributions (see Sec. IV). Therefore the emergence of superconductivity in our model is due to interband polarization. We further substantiate this statement in Appendix E, where we analyze the diagrams leading to attractive interaction at small momenta in more detail.

The interband process leading to pairing interaction corresponds to a virtual state having one hole in the valence band and one additional conduction electron. This perfectly connects to the "three-particle mechanism" introduced in Refs. [35,36], where the pairing between two conduction electrons is mediated by a third electron in the valence band undergoing a virtual interband transition. Importantly, this mechanism differs from other interband mechanisms of superconductivity relying on virtual *pair* scattering of *two* electrons from the Fermi surface to distant or incipient bands [33,34], as described by the ladder diagram of Eq. (20) rather than the bubble diagram.

To get more insights into the effective attraction between electrons leading to the formation of the Cooper pairs, we study in more the details the effective Cooper interaction $V_{n',k'}^{n,k}$ in the upper band $n = n' = +$ obtained within RPA; see Eq. (36). For illustration purposes, we fix the incoming vector $k' = K$ at the center of the Fermi sea, and plot the interaction strength as a function of k in Fig. 5 in the dilute limit $x \simeq 0\%$, obtained by fixing μ at the valence band bottom, and for a substantial doping concentration $x = 10\%$. First and foremost, we observe that the effective interactions in the upper band are attractive, as expected from our analytical and numerical results of Sec. III–VI. At infinitesimal doping $x \simeq 0\%$, obtained by fixing μ at the valence band bottom, we observe that this attraction is entirely due to interband polarization effects. Indeed, it almost perfectly agrees with our analytical prediction for g_0 in Eq. (35) when the momentum transfer $|k' - k| \to 0$ (squares in Fig. 5). On the other hand, for larger doping concentrations, exemplified here with $x = 10\%$, the intraband susceptibilities become non-negligible and start to strongly renormalize the effective RPA interactions. These intraband contributions are similar to those appearing in the standard Kohn-Luttinger mechanism. For the values chosen here, they increase the strength of the attractive potential by up to 30% compared to the pure interband polarization effects. The increase of T_c for $x = 0.1$–0.25 observed in Fig. 3 is
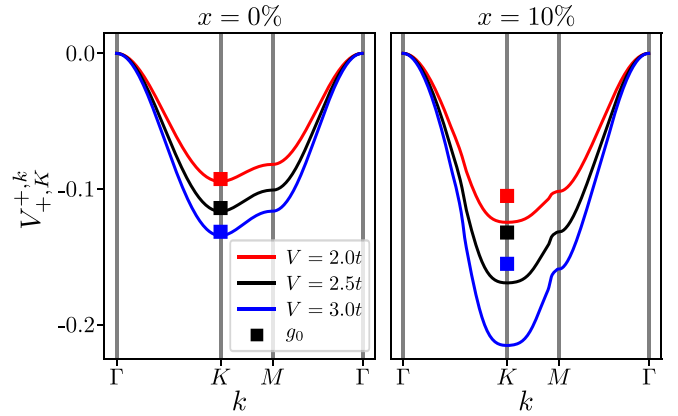

FIG. 5. The effective Cooper interaction $V_{n',k'}^{n,k}$ [see Eq. (36)] in the upper band $n = n' = +$ is attractive, as shown here using the incoming momentum $k' = K$ as reference, with $\Delta = 8t$ and two different doping concentrations. Squares show the analytical prediction g_0 [Eq. (35)] obtained by neglecting intraband contributions.

thus due to a combined effects of the interband and intraband polarization.

B. Contrast with Kohn-Luttinger mechanism

It is worthwhile to compare the mechanism for superconductivity studied here with the traditional Kohn-Luttinger mechanism. From a technical viewpoint, we highlighted in Sec. IV that the susceptibilities consist of two different contributions: intraband, covered by the standard Kohn-Luttinger theory, and interband, which greatly dominates at low doping (see Fig. 5). The nonzero interband polarization has important physical consequences.

First, our screened interaction, once projected near the Fermi energy, is attractive for small momentum transfers $q \ll k_F$. This is not the case in the KL mechanism, which relies on screening processes near $q = 2k_F$. These high-momentum oscillations of the screened Coulomb potential can, at sufficiently large distances, reach negative values. This produces an effective attraction at momentum transfer close to $2k_F$ and pairing in high orbital angular channels, which often leads to nodal superconducting order parameters. Here, in contrast, the interband polarization bubble is already attractive at small q (see also Appendix E), which enables a nodeless order parameter at small densities.

Second, the KL mechanism requires a finite carrier density leading to a Fermi surface [16,17,44–48] to produce screened interaction. To put it another way, the Kohn-Luttinger mechanism does not produce any bound state when only two electrons are present and subject to mutual repulsion. In contrast, our system consists of a completely filled band of "core" electrons, which induces the pairing interaction between doped electrons in the conduction band. Remarkably, we find that two doped electrons already attract and form a bound state, and this attractive interaction is the result of the interband electronic effect in a band insulator with repulsive bare interaction. This novel interband mechanism distinguishes our work from the Kohn-Luttinger mechanism, and produces robust pairing at *infinitesimal* doping.

C. Role of correlated hopping

In the real-space picture offered by the hybridization expansion (Sec. III), the emergence of superconductivity crucially relies on correlated hopping processes of doped electrons—the λ term Eq. (6). Historically, correlated hopping was discussed in connection with superconductivity in the cuprates [49,50] (see also [51,52]). However, the correlated hopping term discussed for cuprates differs from the λ term that emerges in the low-energy limit of our model. The former describes electrons hopping between two sites, one of which being already occupied by another electron of opposite spin. This process necessarily involves double occupation and therefore is strongly suppressed by large on-site repulsion U. In contrast, our λ term describes electron hopping between two sites in the presence of another electron occupying a third site nearby, a process that remains active and leads to attractive interactions even at large U [36].

These density-dependent hopping terms manifest themselves as an increase of the effective tunneling strength for doped charges as a function of their density. This behavior is also captured by the sublattice potential in the Hartree approximation $\Delta_H \approx \Delta + 3V(1 - x)$ (see Fig. 2). Notably, the interaction-induced renormalization of Δ_H makes the effective hopping amplitude doping dependent:

$$t_{\text{eff}} = \frac{t^2}{\Delta_H} \approx \frac{t^2}{\Delta + 3V} + \frac{3t^2 V x}{(\Delta + 3V)^2}. \quad (41)$$

This effect is captured by the correlated hopping term in the effective Hamiltonian after downfolding, and it is essential for the emergence of superconductivity in our model.

VIII. RELEVANCE FOR TWISTED BILAYER GRAPHENE

Throughout this paper, we have highlighted that superconductivity at low doping above integer filling can arise from repulsive interactions in multiband systems due to a nonzero virtual interband polarization. The cause of the attractive interactions is the presence of correlated hopping terms for doped electrons, which can be identified by a filling-dependent Hartree potential Δ_H.

Similar correlated hopping terms have also been observed in twisted bilayer graphene (TBG), where the Hartree correction experiences significant changes as a function of band filling [38]. Drawing an analogy between the two systems, we expect the correlated hopping terms to emerge in the low-energy theory of twisted bilayer graphene and to be responsible for superconductivity.

To place this analogy on firmer grounds, we now examine and compare effective interactions in TBG and in our model using the RPA of Sec. V, and show the evident similarity between the two.

Let us first briefly summarize our results for the spinless model Eq. (1). The renormalization of the sublattice potential difference is the only inhomogeneous on-site potential that respects the translational symmetry of the system. It takes the form $V_H(r) = \Delta_H[\rho_A(r) - \rho_B(r)]/2$, with $\rho_{A/B}(r)$ the density on each orbital in the unit cell located at r. Because of this orbital structure, all quantities involved in the renormalization
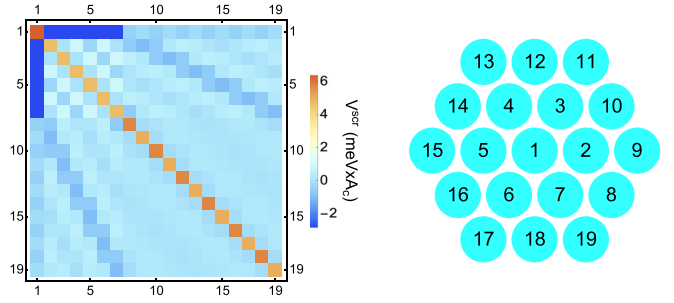

FIG. 6. Some matrix elements $\Gamma_{G,G'}^{\text{RPA}}(q \to 0)$ of the screened potential for twisted bilayer graphene, obtained for the twist angle $\theta = 1.05°$ and for $x = -1$ electrons per moiré unit cell. The index G denotes moiré reciprocal lattice vectors, as defined on the right.

of the scattering vertex, such as the polarizability $\chi_{a,b}$, are promoted to matrix form with orbital indices [see Eq. (28)].

In TBG, the Hartree term contains all possible inhomogeneous on-site potentials with the periodicity of the moiré unit cell. Its Fourier components are

$$V_H(r) = \sum_G \frac{2\pi e^2}{|G|} \rho_G(r), \quad (42)$$

where $G \neq 0$ is a moiré reciprocal lattice vector, and ρ_G is a Fourier component of the total charge density. The latter can be written as

$$\rho_G(r) = \sum_{k,n \in \text{occ}} \langle k, n | e^{iGr} | k, n \rangle, \quad (43)$$

where contraction over the layer and sublattice indices is implied in the expectation value of the single-particle wave functions $|k, n\rangle$, which are now four component spinors.

Hence, the Hartree potential can be written in terms of the ρ_G, in the same manner in which the Hartree potential of the model discussed in Sec. II can be written in terms of $\rho_{A/B}$, the role of the sublattice index being replaced by the reciprocal lattice vectors. The analog goes deeper. In the honeycomb model the two sublattices are connected by electron tunneling t, and similarly, electron states of graphene at wave vectors k and $k + G$ are connected by the interlayer tunneling. In both cases, single-particle bands are formed by the hybridization between multiple components within a unit cell.

Furthermore, as the number of flavors is larger in TBG (spin, layer, sublattice), the RPA treatment becomes more and more justified. Following the reasoning of Sec. V, we find that the susceptibility and RPA vertex are now promoted to matrices, $\chi_{G,G'}(q)$ and $\Gamma_{G,G'}^{\text{RPA}}(q)$, indexed by the moiré reciprocal lattice vectors, in the same way that the repulsive potential and the susceptibility of the model in Sec. II can be written as a matrix using sublattice indices.

To compare the physics of the two models, we plot in Fig. 6 the RPA vertex $\Gamma_{G,G'}^{\text{RPA}}(q \to 0)$ obtained for TBG near the magic angle, at filling $x = -1$ (measured with respect to charge neutrality). We observe that the *off-diagonal* elements of this kernel, not initially present in the interaction, are now negative, signaling attractive interactions.

The analogy between the spinless model Eq. (1) and the interacting model of TBG consolidates. In both cases, the matrix which defines the screened RPA vertex acquires new
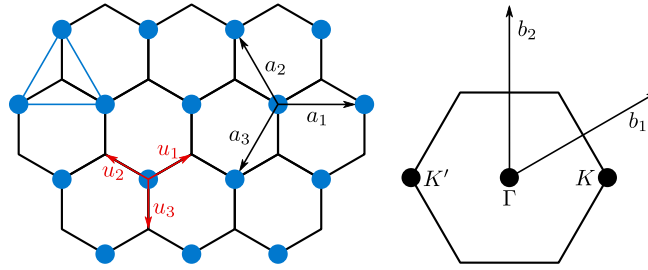
FIG. 7. On the left, primitive lattice vectors $a_{j=1,2,3}$ and nearest-neighbor vectors $u_{j=1,2,3}$ on the honeycomb lattice. Highlighted in blue is the triangular B lattice and the upper triangle centered on an A site summed over in the effective HE model [see Eq. (6)]. On the right, first Brillouin zone and high-symmetry points Γ, K, and K'.

elements, which are not present in the bare interaction, but emerge from *interband* polarization. Furthermore, the order parameter for TBG induced by these new attractive terms shows few or no sign changes throughout the miniature Brillouin zone of the moiré structure [39], as in the spinless model described here [35,36].

We emphasize that the attractive part of the interaction in TBG, shown in Fig. 6, is largely due to interband processes. The essential role of the interband contribution to superconductivity is further testified by considering an approximation method that discards it. This approximation, known as the "flat-metric condition," is convenient for the analytical study of the effect of interactions in TBG, and has been extensively applied in Refs. [53–58]. As described in Appendix G, we show that this approximation however leads to a purely repulsive interaction $V_{c,k'}^{c,k}$ in both TBG and our honeycomb model, thus failing to capture the emergence of superconductivity.

IX. CONCLUSION

In this article, we comprehensively studied one of the simplest electronic mechanisms for superconductivity in doped band insulators. Using a minimal two-band model for illustration, we firmly established the emergence of attractive interactions between doped charges with the help of three analytically controlled methods, whose results perfectly agree in their overlapping domains of validity. All three methods convey a common understanding for the origin of the effective attraction: it arises from the coupling of conduction electrons pairs to virtual excitons. In hybridization expansion in the real-space picture, this effective attraction is described by a correlated hopping term in the effective Hamiltonian for doped electrons, while in the RPA framework based on the k-space band picture, it arises from interband contribution to the particle-hole polarization bubble.

The pairing interaction described here shows a feature not commonly found in models of superconductivity from bare repulsive interactions: the screened interaction, projected onto the states near the Fermi energy, is attractive for small momentum transfer [59], $|q| \ll k_F$.

We have characterized the superconducting order parameter using full-fledged numerical calculations. The small momentum attraction mentioned above translates into a nodeless order parameter at the Fermi surface pockets at the corners of the Brillouin zone, the points K and K'. We found that the critical temperature T_c exhibits a dome as a function of doping, whose maximum reaches up to 3% of the nearest-neighbor hopping amplitude. For a realistic value $t = 0.1$ eV, this already leads to a substantial $T_c \simeq 35$ K. This maximum appears near $x = 1/4$, above which the superconducting gap acquires nodes, leading to a reduction of T_c.

Finally, we highlighted some consequences of our work regarding the theoretical modeling of interactions in twisted bilayer graphene. In particular, we showed the crucial role of correlated hopping terms arising from interband polarization in the understanding of superconductivity. This hints that approximation schemes discarding their effects will probably also underestimate the strength of attractive interaction—if they find any. Remarkably, the structures of the interaction matrix in our simple model and in twisted bilayer graphene are analogous, with the attractive off-diagonal driving the transition to a superconducting state. This shows the importance of the electronic mechanism of superconductivity studied in this article for the understanding of more complex systems.

APPENDIX A: CONVENTIONS FOR THE LATTICE

For completeness, we provide in this Appendix one possible parametrization of the honeycomb lattice and of its first Brillouin zone. The explicit expressions given below ease the analytical evaluation of the g coefficients appearing in the continuum model of Sec. III, for instance.

We define, as shown in Fig. 7, the following: (a_1, a_2) the two primary lattice basis vectors; $a_3 = -a_2 - a_1$ the third primitive lattice vector; $a = |a_1| = |a_2| = |a_3|$ the lattice constant; (u_1, u_2, u_3) the three nearest-neighbor vectors pointing toward A sites; (b_1, b_2) the reciprocal lattice basis vectors; Γ the center of the Brillouin zone; and (K, K') its two nonequivalent corners. Their expression and properties can be straightforwardly obtained from Fig. 7. We also gather for convenience some useful relations,

v	a_1	a_2	a_2	u_1	u_2	u_3
$K \cdot v$	$-2\pi/3$	$4\pi/3$	$-2\pi/3$	0	$2\pi/3$	$-2\pi/3$
$K' \cdot v$	$2\pi/3$	$-4\pi/3$	$2\pi/3$	0	$-2\pi/3$	$2\pi/3$

$$(A1)$$

which show—for instance—that $f(K) = f(K') = 0$, or $\Lambda_{K,K}^{\mathrm{HE}}(0) = \Lambda_{K,K'}^{\mathrm{HE}}(0) = 7V_f/2 - 9\lambda$ and $\Lambda_{K,K}^{\mathrm{HE}}(K) = -V_f$ using Eq. (6).

APPENDIX B: PROJECTION OF INTERACTION

We now detail the projection of the interaction terms of the HE effective model near the band bottom, used to derive the continuum model Eq. (9) of the main text. We express the interacting part of our effective Hamiltonian as

$$H_{\mathrm{int}} = \frac{1}{N} \sum_{\substack{k,k',q \\ \sigma,\sigma'}} \Lambda_{k,k'}^{\mathrm{HE}}(q) c_{k+q,B,\sigma}^{\dagger} c_{k',B,\sigma'}^{\dagger} c_{k'+q,B,\sigma'} c_{k,B,\sigma}, \quad (B1)$$

and Fourier-transform Eq. (6) to get the interaction coefficients in the HE regime ($t \ll \Delta$),

$$\Lambda_{k,k'}^{\mathrm{HE}}(q) = V_f \left[\frac{3}{2} + \sum_{j=1}^{3} \cos(q \cdot a_j) \right] + \lambda \sum_{j=1}^{3} 2\cos(k \cdot a_j) + 2\cos[(q+k) \cdot a_j] + [e^{i(q \cdot a_j - k \cdot a_{j+1})} + e^{-i(q \cdot a_j - k \cdot a_{j-1})}]. \quad (B2)$$

The projection restricts the sum to contributions where the four fermionic operators lie close to the K and K' valleys. Among these contributions, we can also discard the ones that cannot satisfy momentum conservation, which either involve three electrons in the same valley (e.g., $\psi_K^{\dagger} \psi_K^{\dagger} \psi_K \psi_{K'}$) or scatter pairs between the two valleys (e.g., $\psi_K^{\dagger} \psi_K^{\dagger} \psi_{K'} \psi_{K'}$). We get

$$
\begin{aligned}
H_{\mathrm{int}} \simeq \frac{1}{N} \sum_{\substack{k,k',q \\ \sigma,\sigma'}} \big[& \Lambda_{K,K}^{\mathrm{HE}}(0) \psi_{k+q,\sigma,K}^{\dagger} \psi_{k',\sigma',K}^{\dagger} \psi_{k'+q,\sigma',K} \psi_{k,\sigma,K} \\
& + \Lambda_{K',K'}^{\mathrm{HE}}(0) \psi_{k+q,\sigma,K'}^{\dagger} \psi_{k',\sigma',K'}^{\dagger} \psi_{k'+q,\sigma',K'} \psi_{k,\sigma,K'} \\
& + \Lambda_{K,K'}^{\mathrm{HE}}(0) \psi_{k+q,\sigma,K}^{\dagger} \psi_{k',\sigma',K'}^{\dagger} \psi_{k'+q,\sigma',K'} \psi_{k,\sigma,K} \\
& + \Lambda_{K',K}^{\mathrm{HE}}(0) \psi_{k+q,\sigma,K'}^{\dagger} \psi_{k',\sigma',K}^{\dagger} \psi_{k'+q,\sigma',K} \psi_{k,\sigma,K'} \\
& + \Lambda_{K,K}^{\mathrm{HE}}(K) \psi_{k+q,\sigma,K'}^{\dagger} \psi_{k',\sigma',K}^{\dagger} \psi_{k'+q,\sigma',K'} \psi_{k,\sigma,K} \\
& + \Lambda_{K',K'}^{\mathrm{HE}}(K') \psi_{k+q,\sigma,K}^{\dagger} \psi_{k',\sigma',K'}^{\dagger} \psi_{k'+q,\sigma',K} \psi_{k,\sigma,K'} \big]. \quad (B3)
\end{aligned}
$$

Using the Hamiltonian's Hermiticity and time-reversal invariance (or the explicit expression above), we get the relations $[\Lambda_{k,k'}(q)]^* = \Lambda_{k+q,k'+q}(-q) = \Lambda_{-k,-k'}(-q)$ allowing us to equate some of the coefficients above, and to find

$$
\begin{aligned}
H_{\mathrm{int}} \simeq & \frac{1}{N} \sum_{\substack{q,\tau \\ \sigma,\sigma'}} \Lambda_{K,K}^{\mathrm{HE}}(0) \rho_{q,\sigma,\tau} \rho_{-q,\sigma',\tau} \\
& + \frac{1}{N} \sum_{q,\sigma,\sigma'} 2\Lambda_{K,K'}^{\mathrm{HE}}(0) \rho_{q,\sigma,K} \rho_{-q,\sigma',K'} \\
& + \frac{1}{N} \sum_{q,\sigma,\sigma'} \Lambda_{K,K}^{\mathrm{HE}}(K) [\tau_{q,\sigma}^{+} \tau_{-q,\sigma'}^{-} + \tau_{q,\sigma}^{-} \tau_{-q,\sigma}^{-}]. \quad (B4)
\end{aligned}
$$

Using $\tau_{\sigma}^{+} \tau_{\sigma'}^{-} + \tau_{\sigma}^{-} \tau_{\sigma'}^{+} = [2\boldsymbol{\tau}_{\sigma} \cdot \boldsymbol{\tau}_{\sigma'} - (\rho_{K,\sigma} - \rho_{K',\sigma})(\rho_{K,\sigma'} - \rho_{K',\sigma'})/2]$, we end up with the continuum model of Eq. (9) where the g coefficients take the explicit form

$$g_0 = \Lambda_{K,K}^{\mathrm{HE}}(0) - \Lambda_{K,K}^{\mathrm{HE}}(K)/2, \quad (B5a)$$

$$g_1 = 2\Lambda_{K,K'}^{\mathrm{HE}}(0) + \Lambda_{K,K}^{\mathrm{HE}}(K), \quad (B5b)$$

$$g_2 = 2\Lambda_{K,K}^{\mathrm{HE}}(K), \quad (B5c)$$

which has been reproduced in Eq. (23) concerning the interaction expansion—the derivation being identical in that case. Using the explicit form of Λ^{HE}, we find

$$g_2 = 0, \quad g_1 = 2g_0 = 9(V_f - 2\lambda), \quad (B6)$$

as announced in the main text.

APPENDIX C: FORMULA FOR T_c

We briefly reproduce the calculation of Refs. [60,61] leading to the explicit expression of T_c in Eq. (14). When $N_f = 1$, the continuum model of Eq. (12) is a two-component Fermi gas in two dimensions with attractive interactions $g_0 < 0$, and a quadratic band dispersion with effective mass m. This theory hosts a bound state of energy $-\varepsilon_b$ determined as a pole of the two-body scattering T matrix

$$\frac{1}{|g_0|} = \mathcal{D}_0 \int d\varepsilon \, \frac{1}{\varepsilon_b + 2\varepsilon}, \quad (C1)$$

with $\mathcal{D}_0 = Am/(\pi\hbar^2)$ the constant density of state in two dimensions and A the area of the unit cell in real space. Its solution gives $\varepsilon_b \propto We^{-1/g}$ when g, given below Eq. (14), is large [35].

We then solve the gap and number equations

$$\frac{1}{|g_0|} = \mathcal{D}_0 \int_0^\infty d\varepsilon \, \frac{\tanh(\beta E/2)}{2E}, \tag{C2a}$$

$$2E_F = \int_0^\infty d\varepsilon \left[1 - \frac{\varepsilon - \mu}{E} \tanh\left(\frac{\beta E}{2}\right) \right], \tag{C2b}$$

where $E = \sqrt{(\varepsilon - \mu)^2 + \Delta^2}$ for $T = 0$ and $T = T_c$ by direct evaluation of the integrals above. At $T = 0$, the number equation gives

$$\sqrt{\mu^2 + \Delta^2} + \mu = 2E_F, \tag{C3}$$

where μ is the chemical potential, while the gap equation can be recast as

$$\int_0^\infty d\varepsilon \left[\frac{1}{2E_k} - \frac{1}{2\varepsilon + \varepsilon_b} \right] = 0, \tag{C4}$$

which is convergent and yields a second relation between μ and Δ:

$$\sqrt{\mu^2 + \Delta^2} - \mu = \varepsilon_b. \tag{C5}$$

Together, they give

$$\mu = E_F - \varepsilon/2, \quad \Delta = \sqrt{2E_F \varepsilon_b}. \tag{C6}$$

Because the two-dimensional density of states is constant, we can reproduce the standard BCS integrals without assuming any separation of scale between the UV cutoff ε_Λ and the Fermi energy. This yields

$$k_B T_c = K\Delta \propto \sqrt{E_F \varepsilon_b} \propto \sqrt{E_F W} e^{-1/(2g)}, \tag{C7}$$

as announced in the main text. Note that this equation applies to the dilute limit of all regimes considered (HE, KL, and RPA), with the g_0 respectively given by Eqs. (13), (30), and (34).

For instance, let us highlight how the RPA linearized gap equation Eq. (36) gets mapped onto Eq. (C2) in the dilute limit. In this regime electrons mostly live near $\pm K$, and interact through small momentum transfers $q \simeq 0$ since $\Gamma^{\text{RPA}}(\pm K) = 0$. The kernel in Eq. (36) then becomes

$$\tilde{\Gamma}_{k,k'} \simeq \Gamma_{BB}^{\text{RPA}}(0) \frac{\varepsilon_{k,+} \varepsilon_{k',+} + [\Delta/2]^2}{2\varepsilon_{k,+} \varepsilon_{k',+}}$$
$$+ \Gamma_{AB}^{\text{RPA}}(0) \frac{\text{Re}\,[t_F^2 f^*(k') f(k)]}{2\varepsilon_{k,+} \varepsilon_{k',+}}, \tag{C8}$$

which holds up to q^2 corrections. Then, expanding the functions of k and k' near the $\pm K$ points, we find

$$\tilde{\Gamma}_{k,k'} \simeq \Gamma_{BB}^{\text{RPA}}(0) + 0, \tag{C9}$$

which also holds up to quadratic corrections of the form $|k \pm K|^2$ or $|k' \pm K|^2$. This gives the effective attractive interaction $\tilde{\Gamma}_{k,k'} \simeq -g_0$ between the two valleys. Finally, we need to expand the dispersion relation near the $\pm K$ points to find the effective mass

$$\varepsilon_{K+k,+} \simeq \frac{\Delta}{2} + \frac{|k|^2}{2m}, \quad m = \Delta_{\text{H}}/[3(at_F)^2], \tag{C10}$$

which makes the connection with Eq. (C2) complete. Note that we focus on intervalley pairing as intravalley in s-wave is forbidden for our spinless model ($N_f = 1$).

APPENDIX D: EVALUATION OF KL DIAGRAMS

The two-body vertex corrections $\delta\Gamma$ obtained from many-body perturbation theory are described by the diagrams in Eq. (20), where we have used curvy lines for interaction events

$$(k_a, s_a) = a \quad\quad d = (k_d, s_d) \\ \qquad\qquad q = k_d - k_a \quad \equiv \Gamma_{dc,ba}, \\ (k_b, s_b) = b \quad\quad c = (k_c, s_c) \tag{D1}$$

and straight lines for the bare HF propagators

$$\underline{\quad a, b \quad} \equiv G_{b,a}^{(0)}(i\omega_m) = \delta_{k_a, k_b} \sum_{n=\pm} \frac{\Psi_{k_b, n}^{s_b\,*} \Psi_{k_a, n}^{s_a}}{\xi_{k_a, n} - i\omega_m}, \tag{D2}$$

with ω_m a Matsubara frequency, and where $\xi_{k_a, n} = \varepsilon_{k_a, n} - \mu$ measures energies with respect to the chemical potential μ. They are both expressed in the orbitals basis, leading to the important conservation $s_a = s_d$, $s_b = s_c$ in Γ, and to the single-particle dressing $\Psi_{k_b, n}^{s_b\,*} \Psi_{k_a, n}^{s_a}$ in $G^{(0)}$.

The five diagrams of Eq. (20)—referred to as ladder (ldr), cross (crs), up and down wine glasses (uwg/dwg), and bubble (bbl) diagrams due to their form—contribute to the renormalization of the two-body scattering vertex [15–17]:

$$\delta\Gamma_{43,21}^{\text{ldr}} = \frac{-1}{N} \sum_{abcd} \chi_{dc,ba}^- \Gamma_{43,ba} \Gamma_{dc,21}, \tag{D3a}$$

$$\delta\Gamma_{43,21}^{\text{crs}} = \frac{-1}{N} \sum_{abcd} \chi_{dc,ba}^+ \Gamma_{4c,2a} \Gamma_{d3,b1}, \tag{D3b}$$

$$\delta\Gamma_{43,21}^{\text{uwg}} = \frac{-1}{N} \sum_{abcd} \chi_{dc,ba}^+ \Gamma_{c3,2a} \Gamma_{d4,b1}, \tag{D3c}$$

$$\delta\Gamma_{43,21}^{\text{dwg}} = \frac{-1}{N} \sum_{abcd} \chi_{dc,ba}^+ \Gamma_{4c,a1} \Gamma_{d3,b2}, \tag{D3d}$$

$$\delta\Gamma_{43,21}^{\text{bbl}} = \frac{N_f}{N} \sum_{abcd} \chi_{dc,ba}^+ \Gamma_{4c,a1} \Gamma_{d3,2b}, \tag{D3e}$$

where we recall that the sums labeled by the generalized indices $abcd$ run over all sublattice indices $s_a, s_b, s_c, s_d \in \{A, B\}$ and all momenta $k_a, k_b, k_c, k_d \in$ BZ. The particle-particle and particle-hole susceptibilities are computed by the Matsubara sums

$$\chi_{dc,ba}^{\epsilon=\pm} = \frac{1}{\beta} \sum_{\omega_m} G_{d,a}^{(0)}(i\epsilon\omega_m) G_{c,b}^{(0)}(i\omega_m), \tag{D4}$$
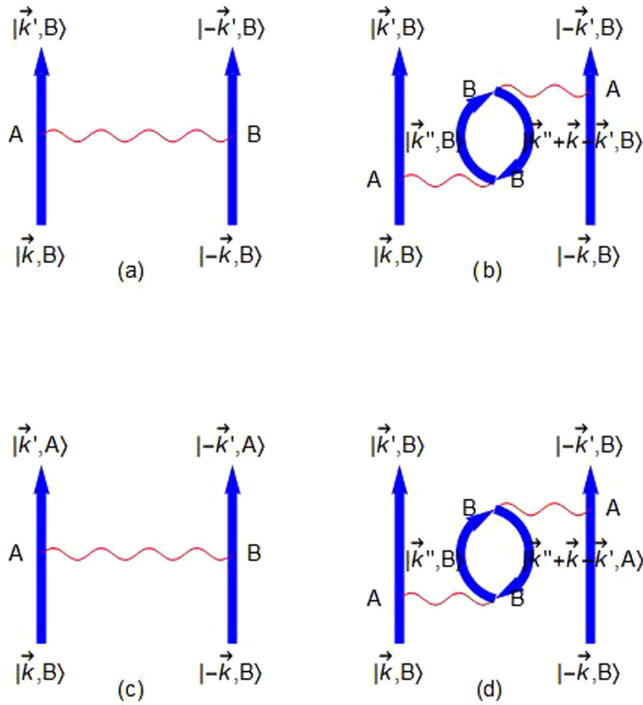
FIG. 8. Some diagrams which describe the scattering of Cooper pairs, and appear in the superconducting kernel in Eq. (38). The band labels A and B stand for the sublattice where the band is mostly localized. States $|k, A\rangle$, $|k', A\rangle$ lie near the Fermi surface of one valley, and states $|-k, A\rangle$, $-|k', A\rangle$ lie near the Fermi surface of the other valley. The states within the bubbles in diagrams (b) and (d) belong to the same valley, either K or K'. The overall momentum transfer is such that $|q| = |k - k'| \ll k_F$.

which gives the result of Eq. (22). The resulting vertex corrections

$$\delta\Gamma = \delta\Gamma^{\text{ldr}} + \delta\Gamma^{\text{crs}} + \delta\Gamma^{\text{uwg}} + \delta\Gamma^{\text{dwg}} + \delta\Gamma^{\text{bbl}} \qquad \text{(D5)}$$

appear in Eq. (21) of the main text.

APPENDIX E: DIAGRAMMATIC DESCRIPTION OF THE ORIGIN OF AN ATTRACTIVE INTERACTION

In the following, we give a simplified description of the origin of attractive terms at small momentum transfer, $|q| \ll k_F$, in the superconducting kernel of the model Eq. (38). We assume, for simplicity, a small hole doping, which leads to two isotropic bands, one per valley, with the filling determined by the Fermi wave vector, k_F. The partially occupied bands are mostly localized in sublattice B, and there are two empty bands at energies of order Δ.

Some low-order ladder and bubble diagrams which describe interaction-induced scattering of Cooper pairs and which contribute to the superconducting kernel in Eq. (38) are shown in Fig. 8. The bare interaction at small momenta is $\lim_{|q| \to 0} V(q) = 3V$, where the factor 3 arises from the number of nearest neighbors. The interaction is between electrons residing at A and B sites, so that the vertices involve form factors describing the amplitude of the wave functions of the A and B bands at the A and B sites. The amplitude of a wave function of the A band at the A sublattice is ≈ 1, and the

amplitude on the B sublattice is, on average, of order t/Δ, and of order $(t/\Delta) \times (k_F a) \sim (t/\Delta) \times \sqrt{n}$ for states near the points K and K' [k_F is the Fermi wavelength, a is the lattice constant, and $n \sim (k_F a)^2 \ll 1$ is the number of electrons per unit cell in the partially occupied band]. The reverse holds for wave functions in the B band.

Diagram (a) in Fig. 8 is due to the bare interaction. It is repulsive, of order $\sim 3V \times (t/\Delta)^2 \times n$. It has a multiplicity of 2, as the two interaction vortices can be exchanged. Diagram (c) describes interband transitions; see [34]. It gives a contribution also of order $\sim 3V \times (t/\Delta)^2 \times n$. It has a multiplicity of 2.

The bubble diagrams, (b) and (d), describe the screening of the bare interactions. Both are second-order processes in perturbation theory, and lead to an attractive interaction. Diagram (b) describes the contribution of the polarizability of the partially occupied valence band, modulated by the weight of the the wave functions of this band on the B sublattice. Its contribution is $\sim -(3V)^2 \times (t/\Delta)^4 \times (\Delta/t^2)$, where the factor Δ/t^2 describes approximately the polarization of the valence band for $|q| \to 0$. The contribution of diagram (d) is $\sim -(3V)^2 \times (t/\Delta)^2 \times (1/\Delta)$, where the factor $1/\Delta$ describes the energy cost of making a transition to the conduction band. The multiplicity of these diagrams is the number of electron flavors times 4, from the all possible exchanges of vertices.

The analysis of the diagrams in Fig. 8 suggests that the superconducting kernel, for $|q| \ll k_F$, becomes attractive for $3V \sim \Delta$, in agreement with the results in the main text. For $n \ll 1$, the leading diagram is (b), leading to a pairing interaction of order $(V^2 t^2)/\Delta^3$, as in Eq. (35).

APPENDIX F: HARTREE AND FOCK CORRECTIONS

The Hartree and Fock diagrams are graphically represented as

$$\text{(F1)}$$

Up to a global shift of chemical potential, the Hartree contribution reads

$$h^{\text{H}}(q) = \frac{3V}{2} \begin{bmatrix} -\delta & 0 \\ 0 & \delta \end{bmatrix}, \qquad \text{(F2)}$$

where $\delta = (NN_f)^{-1} \langle \sum_{r \in A} n_r - \sum_{r \in B} n_r \rangle$ is the sublattice polarization. The Hartree contribution can thus be simply taken into account through a redefinition of the sublattice potential difference $\Delta_{\text{H}} = \Delta + 3V\delta$. The Fock term is purely off-diagonal and assumes the form

$$h^{\text{F}}(q) = -V \sum_{j=1}^{3} \begin{bmatrix} 0 & e^{i(q \cdot u_j)} t_j \\ e^{-i(q \cdot u_j)} t_j^* & 0 \end{bmatrix}, \qquad \text{(F3)}$$

with $t_j = (NN_f^2)^{-1} \sum_{r \in B} \langle c_r^\dagger c_{r+u_j} \rangle$. As long as the system does not spontaneously break the C_3 symmetry of the original model, the Fock correction can be accounted for by a redefinition of the NN tunneling amplitude $t_{\text{F}} = t + Vt_0$, with $t_0 = (t_1 + t_2 + t_3)/3$.

Note that these contributions could have equivalently been derived by performing the standard Hartree-Fock substitution,

$$n_{r,\sigma} n_{r',\sigma'} = \langle n_{r',\sigma'} \rangle n_{i,\sigma} + \langle n_{r,\sigma} \rangle n_{j,\sigma'}$$
$$- \langle c_{r,\sigma}^\dagger c_{r',\sigma'} \rangle c_{r',\sigma'}^\dagger c_{r,\sigma} - \langle c_{r',\sigma'}^\dagger c_{r,\sigma} \rangle c_{r,\sigma}^\dagger c_{r',\sigma'}$$
$$- \langle n_{r,\sigma} \rangle \langle n_{r',\sigma'} \rangle + |\langle c_{r,\sigma}^\dagger c_{r',\sigma'} \rangle|^2, \tag{F4}$$

in the interaction term of Eq. (2), and then discarding constant terms.

In conclusion, the Hartree-Fock corrections to the single-particle Hamiltonian can be simply accounted for by a redefinition of the sublattice potential difference $\Delta \to \Delta_H$ and of the NN tunneling amplitude $t \to t_F$. These new parameters should be self-consistently computed from

$$\Delta_H = \Delta + \frac{3V \langle \sum_{r \in A} n_r - \sum_{r \in B} n_r \rangle}{N N_f}, \tag{F5a}$$

$$t_F = t + \frac{V \sum_{j=1}^3 \langle \sum_{r \in B} c_r^\dagger c_{r+u_j} \rangle}{3 N N_f^2}. \tag{F5b}$$

The renormalized band dispersion and Bloch vectors follow from performing this substitution in Eqs. (18) and (19).

In the weakly doped regime $x \ll 1$, the conduction band is almost empty, such that the expectation values in Eq. (F5) are dominated by valence band contributions. This simplifies the self-consistent equations as

$$\frac{\Delta_H - \Delta}{\Delta_H} = \frac{3V}{2N} \sum_q \frac{1}{\varepsilon_{q,+}}, \quad \frac{t_F - t}{t_F} = \frac{V}{6N N_f} \sum_q \frac{|f(q)|^2}{\varepsilon_{q,+}}. \tag{F6}$$

They allow us to obtain approximate solutions for the Hartree and Fock corrections in the perturbative regimes studied in the main text. For instance, we observe that $|t_F - t| \to 0$ when $N_f \to \infty$, as expected by inspection of the diagrams of Eq. (F1). Indeed, the Hartree bubble can hold any of the N_f fermionic flavors, while the flavor of the intermediate line in the Fock diagram is fixed. The other interesting limit is $t \ll \Delta$ used for the HE. In that regime, we find that electrons of the $n = 1$ insulating state mostly live on A sites, as expected from Sec. III and described by $\Delta_H = \Delta + 3V$ and $t_F = t[1 + V/(N_f \Delta_H)]$.

APPENDIX G: OVERSIMPLIFYING APPROXIMATION

We now detail the "flat-metric condition" introduced in the main text, and show that it leads to purely repulsive interaction, both in the spinless model Eq. (1) and in the interaction model for TBG.

In more detail, we assume that the form factors appearing in the definition of the susceptibility are diagonal in the band basis and do not vary over the BZ, as they would if correlated hopping terms were present. For the spinless model Eq. (1) and twisted bilayer graphene, this respectively corresponds to

$$\Psi_{k,m}^{i*} \Psi_{k,n}^i \approx \delta_{n,m} \overline{\Psi_{k,n}^{i*} \Psi_{k,n}}\big|_{k \in BZ} = \delta_{n,m} f_n^i,$$
$$\langle k, m | e^{iGr} | k, n \rangle \approx \delta_{n,m} \overline{\langle k, n | e^{iGr} | k, n \rangle}\big|_{k \in BZ} = \delta_{n,m} f_n^G. \tag{G1}$$

Focusing on the conduction band, the static susceptibility coefficients can be written as

$$\chi_{\alpha,\beta}^c(q) = \bar\chi(q) f_c^\alpha f_c^\beta, \quad \bar\chi(q) = \sum_{k \in BZ} \frac{f_\beta(\xi_{k+q,n}) - f_\beta(\xi_{k,n})}{\xi_{k+q,n} - \xi_{k,n}}, \tag{G2}$$

with c referring to the conduction band, and where, depending on the model, α and β either denote orbital indices (A/B) or moiré reciprocal lattice vectors (G).

In the spinless model of Eq. (1), this approximation leads to a screened RPA vertex

$$\Gamma^{RPA}(q) \approx \frac{1}{1 - 2V \text{Re}\,[f(q)] f_c^A f_c^B \bar\chi(q)} \begin{bmatrix} |Vf(q) f_c^B|^2 \bar\chi(q) & Vf(q)[1 - Vf^*(q) f_c^A f_c^B \bar\chi(q)] \\ Vf^*(q)[1 - Vf(q) f_c^A f_c^B \bar\chi(q)] & |Vf(q) f_c^A|^2 \bar\chi \end{bmatrix}. \tag{G3}$$

In turn, this vertex produces an effective Cooper interaction [defined in Eq. (36)]

$$V_{c,k'}^{c,k} \approx \frac{2V \text{Re}\,[f(k'-k)] f_c^A f_c^B}{1 - 2\bar\chi(k'-k) V \text{Re}\,[f(k'-k)] f_c^A f_c^B}, \tag{G4}$$

which is positive, i.e., repulsive, for small momentum difference since $\bar\chi(q \to 0) < 0$, $f(q = 0) = 3$ and $f_c^A, f_c^B > 0$. As a result, washing out the interband contributions and the correlated hopping terms is too restrictive to explain the results obtained throughout this paper.

Similarly, for twisted bilayer graphene, following similar steps, we find that the approximation of Eq. (G1) leads to

$$\Gamma_{G,G'}^{RPA}(q) \approx V_G^0(q) \delta_{G,G'} + \frac{V_G^0(q) V_{G'}^0(q) \bar\chi(q) f_G f_{G'}}{1 - \bar\chi(q) \sum_{G''} V_{G''}^0(q) f_{G''}^2}, \tag{G5}$$

where V^0 denotes the bare Coulomb repulsion, and $\bar{\chi}(q)$ is given by an approximation similar to that in Eq. (G2). Projecting the latter in the conduction band, we find an effective Cooper potential

$$V_{c,k'}^{c,k} \approx g^2(k,k') \frac{\sum_G V_G^0(k'-k)f_G^2}{1 - \bar{\chi}(k'-k)\sum_G V_G^0(k'-k)f_G^2}, \tag{G6}$$

where $g(k,k')$ is a band structure dependent coefficient (see Ref. [39] for more details). Because $\bar{\chi}(q) < 0$, all the matrix elements of this potential are positive, describing a net repulsion.

[1] V. Fatemi, S. Wu, Y. Cao, L. Bretheau, Q. D. Gibson, K. Watanabe, T. Taniguchi, R. J. Cava, and P. Jarillo-Herrero, Electrically tunable low-density superconductivity in a monolayer topological insulator, Science **362**, 926 (2018).

[2] E. Sajadi, T. Palomaki, Z. Fei, W. Zhao, P. Bement, C. Olsen, S. Luescher, X. Xu, J. A. Folk, and D. H. Cobden, Gate-induced superconductivity in a monolayer topological insulator, Science **362**, 922 (2018).

[3] T. Asaba, Y. Wang, G. Li, Z. Xiang, C. Tinsman, L. Chen, S. Zhou, S. Zhao, D. Laleyan, Y. Li *et al.*, Magnetic field enhanced superconductivity in epitaxial thin film WTe$_2$, Sci. Rep. **8**, 6520 (2018).

[4] J. Schooley, W. Hosler, and M. L. Cohen, Superconductivity in Semiconducting SrTiO$_3$, Phys. Rev. Lett. **12**, 474 (1964).

[5] M. N. Gastiasoro, J. Ruhman, and R. M. Fernandes, Superconductivity in dilute SrTiO$_3$: A review, Ann. Phys. **417**, 168107 (2020).

[6] H. Kawaji, K.-i. Hotehama, and S. Yamanaka, Superconductivity of alkali metal intercalated β-zirconium nitride chloride, A$_x$ZrNCl (A = Li, Na, K), Chem. Mater. **9**, 2127 (1997).

[7] Y. Nakagawa, Y. Kasahara, T. Nomoto, R. Arita, T. Nojima, and Y. Iwasa, Gate-controlled BCS-BEC crossover in a two-dimensional superconductor, Science **372**, 190 (2021).

[8] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Unconventional superconductivity in magic-angle graphene superlattices, Nature (London) **556**, 43 (2018).

[9] X. Lu, P. Stepanov, W. Yang, M. Xie, M. A. Aamir, I. Das, C. Urgell, K. Watanabe, T. Taniguchi, G. Zhang *et al.*, Superconductors, orbital magnets and correlated states in magic-angle bilayer graphene, Nature (London) **574**, 653 (2019).

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

[11] G. Chen, A. L. Sharpe, P. Gallagher, I. T. Rosen, E. J. Fox, L. Jiang, B. Lyu, H. Li, K. Watanabe, T. Taniguchi *et al.*, Signatures of tunable superconductivity in a trilayer graphene moiré superlattice, Nature (London) **572**, 215 (2019).

[12] J. M. Park, Y. Cao, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, Tunable strongly coupled superconductivity in magic-angle twisted trilayer graphene, Nature (London) **590**, 249 (2021).

[13] H. Zhou, T. Xie, T. Taniguchi, K. Watanabe, and A. F. Young, Superconductivity in rhombohedral trilayer graphene, Nature (London) **598**, 434 (2021).

[14] N. Bogoljubov, V. V. Tolmachov, and D. Širkov, A new method in the theory of superconductivity, Fortschr. Phys. **6**, 605 (1958).

[15] W. Kohn and J. Luttinger, New Mechanism for Superconductivity, Phys. Rev. Lett. **15**, 524 (1965).

[16] S. Maiti and A. V. Chubukov, Superconductivity from repulsive interaction, in *Lectures on the Physics of Strongly Correlated Systems XVII: Seventeenth Training Course in the Physics of Strongly Correlated Systems*, edited by A. Avella and F. Mancini, AIP Conf. Proc. 1550 (AIP, Melville, NY, 2013), p. 3–73.

[17] M. Y. Kagan, V. Val'kov, V. Mitskan, and M. Korovushkin, The Kohn-Luttinger effect and anomalous pairing in new superconducting systems and graphene, J. Exp. Theor. Phys. **118**, 995 (2014).

[18] B. Keimer, S. A. Kivelson, M. R. Norman, S. Uchida, and J. Zaanen, From quantum matter to high-temperature superconductivity in copper oxides, Nature (London) **518**, 179 (2015).

[19] G. R. Stewart, Superconductivity in iron compounds, Rev. Mod. Phys. **83**, 1589 (2011).

[20] Z. P. Yin, K. Haule, and G. Kotliar, Spin dynamics and orbital-antiphase pairing symmetry in iron-based superconductors, Nat. Phys. **10**, 845 (2014).

[21] R. M. Fernandes and A. V. Chubukov, Low-energy microscopic models for iron-based superconductors: A review, Rep. Prog. Phys. **80**, 014503 (2017).

[22] V. Kozii, H. Isobe, J. W. F. Venderbos, and L. Fu, Nematic superconductivity stabilized by density wave fluctuations: Possible application to twisted bilayer graphene, Phys. Rev. B **99**, 144507 (2019).

[23] Y.-Z. You and A. Vishwanath, Superconductivity from valley fluctuations and approximate SO(4) symmetry in a weak coupling theory of twisted bilayer graphene, npj Quantum Mater. **4**, 16 (2019).

[24] J. Y. Lee, E. Khalaf, S. Liu, X. Liu, Z. Hao, P. Kim, and A. Vishwanath, Theory of correlated insulating behaviour and spin-triplet superconductivity in twisted double bilayer graphene, Nat. Commun. **10**, 5333 (2019).

[25] F. Wu and S. Das Sarma, Collective Excitations of Quantum Anomalous Hall Ferromagnets in Twisted Bilayer Graphene, Phys. Rev. Lett. **124**, 046403 (2020).

[26] E. Khalaf, N. Bultinck, A. Vishwanath, and M. P. Zaletel, Soft modes in magic angle twisted bilayer graphene, arXiv:2009.14827.

[27] G. Sharma, M. Trushin, O. P. Sushkov, G. Vignale, and S. Adam, Superconductivity from collective excitations in

magic-angle twisted bilayer graphene, Phys. Rev. Research **2**, 022040(R) (2020).

[28] A. Kumar, M. Xie, and A. H. MacDonald, Lattice collective modes from a continuum model of magic-angle twisted bilayer graphene, Phys. Rev. B **104**, 035119 (2021).

[29] M. Christos, S. Sachdev, and M. S. Scheurer, Correlated insulators, semimetals, and superconductivity in twisted trilayer graphene, arXiv:2106.02063.

[30] H. Dai, J. Hou, X. Zhang, Y. Liang, and T. Ma, Mott insulating state and $d + id$ superconductivity in an ABC graphene trilayer, Phys. Rev. B **104**, 035104 (2021).

[31] S. Chatterjee, T. Wang, E. Berg, and M. P. Zaletel, Inter-valley coherent order and isospin fluctuation mediated superconductivity in rhombohedral trilayer graphene, arXiv:2109.00002.

[32] C. Huang, N. Wei, W. Qin, and A. H. MacDonald, Pseudospin paramagnons and the superconducting dome in magic angle twisted bilayer graphene, arXiv:2110.13351.

[33] X. Chen, S. Maiti, A. Linscheid, and P. J. Hirschfeld, Electron pairing in the presence of incipient bands in iron-based superconductors, Phys. Rev. B **92**, 224514 (2015).

[34] Z. Dong and L. Levitov, Activating superconductivity in a repulsive system by high-energy degrees of freedom, arXiv:2103.08767.

[35] V. Crépel and L. Fu, New mechanism and exact theory of superconductivity from strong repulsive interaction, Sci. Adv. **7**, eabh2233 (2021).

[36] V. Crépel and L. Fu, Spin-triplet superconductivity from interband effect in doped insulators [Proc. Natl. Acad. Sci. USA (to be published)], arXiv:2103.12060.

[37] K. Slagle and L. Fu, Charge transfer excitations, pair density waves, and superconductivity in moiré materials, Phys. Rev. B **102**, 235423 (2020).

[38] F. Guinea and N. R. Walet, Electrostatic effects, band distortions, and superconductivity in twisted graphene bilayers, Proc. Natl. Acad. Sci. U.S.A. **115**, 13174 (2018).

[39] T. Cea and F. Guinea, Coulomb interaction, phonons, and superconductivity in twisted bilayer graphene, Proc. Natl. Acad. Sci. U.S.A. **118**, e2107874118 (2021).

[40] V. Tien Phong, P. A. Pantaleón, T. Cea, and F. Guinea, Band structure and superconductivity in twisted trilayer graphene, Phys. Rev. B **104**, L121116 (2021).

[41] A. V. Chubukov, Kohn-Luttinger effect and the instability of a two-dimensional repulsive fermi liquid at $t = 0$, Phys. Rev. B **48**, 1097 (1993).

[42] S. Raghu, S. A. Kivelson, and D. J. Scalapino, Superconductivity in the repulsive Hubbard model: An asymptotically exact weak-coupling solution, Phys. Rev. B **81**, 224505 (2010).

[43] T. Cea, P. A. Pantaleón, V. Tien Phong, and F. Guinea, Superconductivity from repulsive interactions in rhombohedral trilayer graphene: A Kohn-Luttinger-like mechanism, arXiv:2109.04345.

[44] R. Nandkishore, R. Thomale, and A. V. Chubukov, Superconductivity from weak repulsion in hexagonal lattice systems, Phys. Rev. B **89**, 144501 (2014).

[45] Y.-P. Lin and R. M. Nandkishore, Kohn-Luttinger superconductivity on two orbital honeycomb lattice, Phys. Rev. B **98**, 214521 (2018).

[46] R. N. Nandkishore, L. S. Levitov, and A. V. Chubukov, Chiral superconductivity from repulsive interactions in doped graphene, Nat. Phys. **8**, 158 (2012).

[47] J. González and T. Stauber, Kohn-Luttinger Superconductivity in Twisted Bilayer Graphene, Phys. Rev. Lett. **122**, 026801 (2019).

[48] N. Yuan, H. Isobe, and L. Fu, Magic of high-order Van Hove singularity, Nat. Commun. **10**, 5769 (2019).

[49] J. E. Hirsch and F. Marsiglio, Superconductivity in an oxygen hole metal, Phys. Rev. B **41**, 2049 (1990).

[50] F. Marsiglio and J. E. Hirsch, Hole superconductivity and the high-t_c oxides, Phys. Rev. B **41**, 6435 (1990).

[51] S. Schubin and S. Vonchowsky, On the electron theory of metals, Proc. R. Soc. London A **145**, 159 (1934).

[52] S. V. Vonsovsky and M. I. Katsnelson, Some types of instabilities in the electron energy spectrum of the polar model of the crystal. I. The maximum-polarity state, J. Phys. C: Solid State Phys. **12**, 2043 (1979).

[53] B. A. Bernevig, Z.-D. Song, N. Regnault, and B. Lian, Twisted bilayer graphene. I. Matrix elements, approximations, perturbation theory, and a $k \cdot p$ two-band model, Phys. Rev. B **103**, 205411 (2021).

[54] Z.-D. Song, B. Lian, N. Regnault, and B. A. Bernevig, Twisted bilayer graphene. II. Stable symmetry anomaly, Phys. Rev. B **103**, 205412 (2021).

[55] B. A. Bernevig, Z.-D. Song, N. Regnault, and B. Lian, Twisted bilayer graphene. III. Interacting Hamiltonian and exact symmetries, Phys. Rev. B **103**, 205413 (2021).

[56] B. Lian, Z.-D. Song, N. Regnault, D. K. Efetov, A. Yazdani, and B. A. Bernevig, Twisted bilayer graphene. IV. Exact insulator ground states and phase diagram, Phys. Rev. B **103**, 205414 (2021).

[57] B. A. Bernevig, B. Lian, A. Cowsik, F. Xie, N. Regnault, and Z.-D. Song, Twisted bilayer graphene. V. Exact analytic many-body excitations in Coulomb Hamiltonians: Charge gap, Goldstone modes, and absence of Cooper pairing, Phys. Rev. B **103**, 205415 (2021).

[58] F. Xie, A. Cowsik, Z.-D. Song, B. Lian, B. A. Bernevig, and N. Regnault, Twisted bilayer graphene. VI. An exact diagonalization study at nonzero integer filling, Phys. Rev. B **103**, 205416 (2021).

[59] Reference [32] also obtains, for certain parameter values, an attractive interaction for $q \to 0$. This result is achieved, as in our work, because of a large contribution from the sublattice polarizability in a model of spinless electrons. In the model studied here the sublattice degree of freedom is the only internal index available. The extension of the results to twisted bilayer graphene is based on a susceptibility where charge fluctuations within the unit cell, labeled by reciprocal lattice vectors, are crucial. See Sec. VIII.

[60] K. Miyake, Fermi liquid theory of dilute submonolayer ^{3}He on thin ^{4}He II film: Dimer bound state and Cooper pairs, Prog. Theor. Phys. **69**, 1794 (1983).

[61] M. Randeria, J.-M. Duan, and L.-Y. Shieh, Superconductivity in a two-dimensional Fermi gas: Evolution from Cooper pairing to Bose condensation, Phys. Rev. B **41**, 327 (1990).