I. INTRODUCTION

Spin-triplet superconductors display a plethora of unconventional phenomena, including multi-component order parameter, fractional vortices [11, 12], Majorana fermions [3] and topological boundary modes [4, 5]. Further interest in spin-triplet superconductors is fueled by their prospect as a material platform for topological quantum computing, such as SrTiO$_3$ and UTe$_2$ [13, 14]. In recent years, evidence of anisotropic spin-triplet pairing has also been discovered in superconducting doped topological insulators such as Sr$_2$RuO$_4$ [11], UPt$_3$ [12] and UTe$_2$ [13, 14]. In particular, we highlight the observed increase of $T_c$ in WTe$_2$ under a small in-plane magnetic field as a strong evidence for spin-triplet superconductivity.

Spin-triplet superconductivity from interband effect in doped insulators

Valentin Crépel, Liang Fu

Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, USA

Despite being of fundamental importance and potential interest for topological quantum computing, spin-triplet superconductors remain rare in solid state materials after decades of research. In this work, we present a general mechanism for spin-triplet superconductivity in multi-band systems, where a non-retarded pairing interaction between conduction electrons is produced by their electronic repulsion to a third-electron undergoing a virtual interband transition. Our theory is analytically controlled by an interband hybridization parameter, and explicitly demonstrated in doped band insulators with the example of an extended Hubbard model. In light of this theory, we propose that recently discovered dilute superconductors such as ZrNCl and WTe$_2$ are spin-triplet, and compare the expected consequences of our theory with experimental data.
II. MODEL

To illustrate the interband electronic mechanism for superconductivity, we consider a two-band Hubbard model on the honeycomb lattice with staggered potential on A/B sites and extended interactions. The Hamiltonian takes the form

$$\mathcal{H} = -t_0 \sum_{\langle r, r' \rangle, \sigma} (c_{r, \sigma}^\dagger c_{r', \sigma} + \text{h.c}) + \frac{\Delta_0}{2} \left[ \sum_{r \in B} n_r - \sum_{r \in A} n_r \right] + U_A \sum_{r \in A} n_{r, \uparrow} n_{r, \downarrow} + U_B \sum_{r \in B} n_{r, \uparrow} n_{r, \downarrow} + V_0 \sum_{\langle r, r' \rangle} n_{r, \uparrow} n_{r', \downarrow}, \quad (1)$$

where $\Delta_0$ is the staggered sublattice potential, $U_A$ and $U_B$ are on-site interactions, and $V_0$ the nearest-neighbor repulsion. We consider this two band model at or slightly above the filling of $n = 2$ electrons per unit cell.

To controllably describe the effects of interband processes on doped electrons, we will consider situations where the two bands are weakly coupled. This is realized in the two following regimes, that we respectively develop in Secs. III and IV:

- When $t_0 = 0$, the lower and upper bands of our model are formed by A and B sublattice states respectively. A small tunneling amplitude $t_0 \ll \Delta_0$ induces weak hybridization between these sublattices/bands, which we treat as a perturbation to derive an effective model for doped electrons.

- When the interactions are small compared to the single particle band gap $\Delta_0$, the two bands are weakly coupled by the interaction terms. This regime, often studied with standard many-body perturbation theory, offers a stringent test of our kinetic expansion method since their domains of validity overlap when $(t_0, V_0) \ll \Delta_0$.

In both limits, the same essential physics – attractive interaction induced by virtual interband processes, sketched in Fig. 1 – is at play, as highlighted by the consistency between kinetic and interaction based perturbative methods.

III. KINETIC ENERGY EXPANSION

A. Effective dynamics of charge carriers

When $\Delta \equiv \Delta_0 - U_A$ is much greater than $t_0$, the ground state of Eq. (1) at $n = 2$ is an insulator, which is adiabatically connected to the noninteracting band insulator obtained when $U_A = U_B = V_0 = 0$. This state is also adiabatically connected to the $t_0 = 0$ limit where A sites of the honeycomb lattice are doubly occupied and B sites are empty. Pauli exclusion principle requires doped electrons above $n = 2$ to live on B sites, which form a triangular lattice. Our following analysis is based on a perturbative expansion in the tunneling term around the atomic limit with $t_0 = 0$ [23]. Since direct tunneling to doubly occupied A sites is prohibited, the effective Hamiltonian governing the dynamics of the $x \equiv n - 2$ doped electrons, denoted as $\mathcal{H}_f$, arises from second-order tunneling processes where the first step necessarily creates a hole on an A site, resulting in intermediate states involving excitons (see Fig. 1).

To account for all such processes, we perform a Schrieffer-Wolff transformation that integrates out high-energy degrees of freedom [31], and obtain the effective Hamiltonian

$$\mathcal{H}_f = t \sum_{(i,j), \sigma} (f_{i, \sigma}^\dagger f_{i, \sigma} + \text{h.c}) + U \sum_i n_{i, \uparrow} n_{i, \downarrow} + V \sum_{(i,j)} n_i n_j + \frac{\lambda}{2} \sum_{(i,j), \sigma} (f_{i, \sigma}^\dagger f_{i, \sigma} + \text{h.c})(n_i + n_j) + \lambda \sum_{ijk \in \Delta, \sigma} \left[ f_{j, \sigma}^\dagger n_k f_{i, \sigma} + P_{ijk} \right] \quad (2)$$

where the fermion operators $f_i$ describe electrons on the triangular lattice formed by B sites. Besides single-particle hopping $(t)$, on-site and nearest-neighbor repulsion $(U, V)$, this effective Hamiltonian contains two types of correlated hopping terms: an electron hops between $i$ and $j$ only when either site is occupied by a second electron of opposite spin $(i \neq j)$, or when their common neighbor $k$ is occupied $(\lambda)$. The second type – hereafter referred to as $\lambda$-hopping – acts on upper triangles $(i j k \in \Delta)$. $P_{ijk}$ denotes all possible permutations of the vertices $i j k$. $\mathcal{H}_f$ also contains three-body interactions (see App. [31]), but their effect is negligible in the low density limit.

$\mathcal{H}_f$ is exact up to second order in the hybridization parameter $t_0/\Delta$ that governs the perturbative Schrieffer-Wolff transformation, holds at all dopings, and features...
interactions between doped electrons that are instantaneous on the time scale set by their kinetic energy $t$. The density-density interaction and correlated hopping terms between doped electrons arise because the bare interactions $(U_A, U_B, V_0)$ affect the energy of intermediate exciton states. For example, the process leading to $\lambda$-hopping is shown in Fig. 1b. The presence of a charge at site $k$ in the upper triangle $ijk$ decreases the energy of the intermediate state in the electron tunneling event $f_{j,\alpha}^\dagger f_{i,\sigma}$ from $\Delta + 4V_0$ to $\Delta + 3V_0$ [32]. The hopping amplitude accordingly increases from $t$ to $t + \lambda$, with

$$t = \frac{t_0^2}{\Delta + 4V_0},$$

$$\lambda = \frac{t_0^2}{\Delta + 3V_0} - \frac{t_0^2}{\Delta + 4V_0}. \tag{3b}$$

Similar considerations yield the other coefficients of $\mathcal{H}_f$ (see App. S1), and give $(t, V) \propto t_0^2$ and $U \simeq U_B$ when $U_B \gg t_0^2/\Delta$. These parameters are plotted in Fig. 1c. The bandwidth $W = 9t$ largely dominates $t, \lambda, V$, even for rather large bare repulsion $V_0 = 10\Delta$.

Our derivation can be straightforwardly extended to account for longer-range hopping between $B$ orbitals, allowing us to describe materials with a larger bandwidth $W$. The effects of longer range interactions can also be included in a similar manner [23].

B. Two-particle bound state

Remarkably, $\mathcal{H}_f$ features an attractive interaction between conduction electrons forming a spin-triplet state. This effective attraction surprisingly emerges in the purely repulsive Hubbard model Eq. 1 due to $\lambda$-hopping, as it lowers the energy of two electrons on adjacent sites compared to the case when they are far apart. In our model, this attraction competes with and wins over the repulsion $V > 0$, which disfavors nearest-neighbor pairing.

To demonstrate pairing, we consider the analog of Cooper’s problem in a doped band insulator and solve Eq. 2 for two doped electrons (see App. S2). Bound states, signaled by a positive energy $\varepsilon_b$, are found in the spin-triplet channel in a very large parameter range. In Fig. 2, we represent such triplet bound state as a function of the relative position between the two particles. This wavefunction exhibits f-wave symmetry, i.e. it is symmetric under three-fold rotation and changes sign under any reflection flipping one of the primitive vectors $\mathbf{a}_i$. In Fig. 2a, we show the binding energy $\varepsilon_b = 6t - E_2$, with $E_2$ the two-electron ground state energy, as a function of the original model parameters $U_B, \Delta$ and $V_0$. The binding energy reaches the maximum value $\varepsilon_b = 0.89t$ when the condition $V_0 \gg \Delta, U_B$ is satisfied. Increasing $U_B$ suppresses the contribution of intermediate states involving doubly occupied $B$ sites, thus leading to a reduced binding energy. In the case $U_B = 1.5V_0 \gg \Delta$, we have $\varepsilon_b = 0.088t$. In the limit $U_B \gg V_0$ and $\Delta$, bound state formation requires $V_0 > \Delta$. In contrast to triplet pairs, no bound state is found for spin-singlet pairs.

It is important to note that in our model, triplet pairing already occurs at infinitesimal carrier doping in a band insulator with purely repulsive interaction. It is remarkable that electron pairing by repulsion occurs in such a simple system, without any connection to resonating valence bond or quantum spin liquid.

We further show in App. S2 that the formation of two-particle bound state is perturbatively stable against the inclusion of longer range part of the Coulomb repulsion. Moreover, at finite doping concentration, Coulomb interaction is dynamically screened by free carriers and effectively becomes short-ranged. Our pairing mechanism relies on short-range repulsion (such as $V_0$ between $B$ sites) that couples itinerant electrons in the conduction band with core electrons in the filled band. Such interband effects produces attractive pairing interaction, as we have demonstrated rigorously above. Thanks to the beneficial effect of dynamical screening, our interband electronic mechanism for superconductivity can work at finite doping, even when two-particle bound state does not exist [24].

C. Superconductivity in the dilute limit

The existence of spin-triplet bound states is suggestive of superconductivity at finite doping. We now study the Hamiltonian $\mathcal{H}_f$ for doped electrons at finite density on the triangular lattice. The single-particle dispersion, given by $\varepsilon_k = 2t \sum_{j=1}^3 \cos(k \cdot \mathbf{a}_j)$, gives rise to two degenerate minima at $K$ and $K'$ points of the Brillouin zone. At sufficiently small density $x \ll 1$, electrons primarily occupy states near these minima. Thus, our system at small doping is an interacting two-valley electron liquid.

We now derive an effective continuum theory that captures the long-wavelength behavior of this liquid. Re-
taining only fermionic modes $f_x(\tau K + k) \equiv \psi_{\sigma \tau}(k)$, \(\sigma \in \{\uparrow, \downarrow\}\), close to either of the two valleys \(\tau \in \{K, K'\}\), we derive the effective continuum description of the lattice model \(\mathcal{H}_f\)

\[
\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_i, \quad \mathcal{H}_i = \int dx \sum_{\sigma \tau} \psi_{\sigma \tau}^\dagger \left[ -\frac{\nabla^2}{2m} \right] \psi_{\sigma \tau}, \quad (4)
\]

with \(m = 2/(3a^2)\) the effective mass at \(K\) and \(K'\). \(\mathcal{H}_i\) consists of three symmetry-allowed contact interactions:

\[
\mathcal{H}_i = \int dx \left( \rho_0 (\rho_{K1} \rho_{K1} + \rho_{K'1} \rho_{K'1}) + g_1 \rho_{K1} \rho_{K'1} + g_2 s_{K} \cdot s_{K'} \right), \quad (5)
\]

where \(\rho_{\sigma \tau} = \psi_{\sigma \tau}^\dagger \psi_{\alpha \tau}\) and \(s_{\sigma \tau} = \psi_{\sigma \tau}^\dagger \sigma_{\alpha \beta} \psi_{\beta \tau}\) respectively denote the density and spin at valley \(\tau\). Here, \(g_0\) and \(g_1\) respectively correspond to intra- and inter-valley repulsion, while \(g_2\) describes inter-valley exchange interactions.

By expressing the microscopic lattice model Eq. 2 in terms of low-energy modes, we derive the three coupling constants (see App. [33]):

\[
\begin{align}
 g_0 &= (U + 6V - 6\lambda - 6\delta)/\mathcal{A} > 0, \quad (6a) \\
 g_1 &= (U + 15V - 24\lambda - 6\delta)/(2\mathcal{A}) > 0, \quad (6b) \\
 g_2 &= -2(U - 3V + 12\lambda - 6\delta)/\mathcal{A} < 0, \quad (6c)
\end{align}
\]

with \(A = 2/\sqrt{3a^2}\) the Brillouin zone area. The emergence of ferromagnetic inter-valley exchange interaction \(g_2 < 0\) can be intuitively understood as Hund’s rules applied to the valley degree of freedom.

The continuum theory Eq. 3 predicts the existence of spin-triplet valley-singlet bound states in the \(s\)-wave channel when

\[ g = g_1 + g_2/4 = 9(V - 2\lambda)/\mathcal{A} < 0. \quad (7) \]

Importantly, these valley-singlet bound states correspond to f-wave electron pairs on the lattice, as they are invariant under threefold rotation \(\psi_{\sigma \tau} \rightarrow e^{i2\pi/3} \psi_{\sigma K}, \psi_{\sigma K'} \rightarrow e^{-i2\pi/3} \psi_{\sigma K'}\), but change sign under reflections exchanging the two valleys \(\psi_{K} \leftrightarrow \psi_{K'}\). Their binding energy is given by [33]

\[ \varepsilon_b = \Lambda \left[ \exp \left( \frac{4\pi}{m|g|} \right) - 1 \right], \quad (8) \]

with \(\Lambda\) a UV energy cutoff. As shown in Fig. 2b, this formula almost perfectly reproduces our solution of the two-body problem on the lattice with \(\Lambda \simeq 1.75t\) as the only fitting parameter. This proves the validity of our derivation of the continuum theory Eq. 3 from the microscopic lattice model.

The condition \(g < 0\) provides an analytical criterion for the appearance of pairing in terms of the original lattice parameters \(\Delta_0, U, U_B\) and \(V_0\). Using Eqs. 4 we observe that this criterion is satisfied in a wide parameter window, which is shown in Fig. 3a. Note that in the presence of on-site repulsion \(U_B \neq 0\), pairing occurs only when the nearest-neighbor repulsion \(V_0\) exceeds a critical value that depends on \(U_B\) and \(\Delta\). To capture such effect, it is thus essential that our theory of exciton-mediated pairing is non-perturbative in interaction strength.

## D. \(T_c\) versus doping

At finite density, the spin-triplet valley-singlet pairing leads to a superconducting state with a vector order parameter

\[ \mathbf{d}_c = \langle \psi_{K,\alpha} [\sigma(\sigma_g)]_{\alpha,\beta} \psi_{K',\beta} \rangle. \quad (9) \]

This state is fully gapped and isotropic, i.e. the order parameter \(\mathbf{d}\) is constant (see Fig. 3b). This property is unusual for triplet superconductors, where the antisymmetry of the Cooper pair wavefunction implies \(\mathbf{d}(\mathbf{k}) = -\mathbf{d}(-\mathbf{k})\). For a singly connected Fermi surface centered at \(\mathbf{k} = 0\), the condition necessarily requires strong variations of the \(\mathbf{d}\)-vector over the Fermi surface, as exemplified in the A and B phases of \(^3\text{He}\). In stark contrast, in our case the presence of two disconnected Fermi surfaces centered around \(K\) and \(K'\) enables inter-valley spin-triplet pairing with a constant and opposite \(\mathbf{d}\)-vector over the Fermi surface of each valley.

Another important feature of our theory is the non-retarded pairing interaction, which spreads over the entire bandwidth of doped charges. This sharply contrasts with electron-phonon superconductors, where the attractive interaction is cut off by the Debye energy, and thus limited to the vicinity of the Fermi surface. The absence of energy cutoff changes the expression of the superconducting critical temperature \(T_c\). We find that the meanfield \(T_c\) depends strongly on the carrier density \(x\), despite the constant density of states [34]

\[ k_B T_c^{\text{MF}} = \frac{\gamma}{\pi} \sqrt{\varepsilon F \varepsilon_b} \propto \sqrt{x W} \exp \left( -\frac{4\pi}{m|g|} \right). \quad (10) \]
where $\varepsilon_F$ is the noninteracting Fermi energy. As a consequence, the gap to $T_c$ ratio exceeds the BCS value of 1.764 by almost a factor of 3 [23]. At very low density where $\varepsilon_F < \varepsilon_h$, the BCS mean-field approximation is inadequate. The superconducting state is instead a Bose-Einstein condensate of spin-triplet pairs.

At high density, lattice effects become important. We have performed mean-field calculations for the lattice model Eq. 2 (see App. S1). In agreement with the previous discussion, we find that the spin-triplet f-wave pairing channel is the leading instability in our model. The mean-field order parameter takes the form

$$d_k = d \sum_j \sin(k \cdot a_j) \equiv s_k d.$$  \hspace{1cm} (11)

Its dependence on the crystal momentum, shown in Fig. 3a, is determined by the form of interactions. The overall amplitude of $d$ is obtained from the spin-triplet f-wave gap equation

$$\frac{3}{2|\lambda - V|} = \int \frac{d^2 k}{4\pi^2} s_k^2 \tanh \left( \frac{E_k}{2k_B T} \right),$$  \hspace{1cm} (12)

which always has a solution in the pairing region identified by Eq. 7 (see Fig. 3a). Here, the quasi-particle energy spectrum

$$E_k = \sqrt{(\varepsilon_k - \mu)^2 + |d_k|^2},$$  \hspace{1cm} (13)

with $\mu$ the chemical potential, remains gapped at the Fermi level throughout the transition from the band insulator to the spin-triplet superconductor (recall that $|d_k| = |d_k'| \neq 0$).

In Fig. 2, we show the critical temperature $T_c$ extracted from Eq. 12 as a function of $x$ and $V_0/\Delta$. At small doping, the f-wave pairing vector $|d_k|$ is nearly constant around $K$ and $K'$. $T_c$ sharply rises with $\sqrt{T}$ dependence, in agreement with the continuum prediction Eq. 8. As doping increases, the Fermi surface approaches the $\Gamma-M$ lines, where the order parameter $d_k \propto s_k$ vanishes. Its small amplitude close to these lines leads to a reduction of $T_c$ for $x > 0.1$, as can be seen in Fig. 2.

Our results for $x > 0.5$ shows that $T_c$ almost vanishes in the strongly doped regime where the f-wave superconducting order parameter has nodes, see Fig. 2. In this limit, other studies based on weak-interaction expansions suggest that f-wave and d-wave superconductivity compete [23 56]. Our strong coupling analysis also finds an attractive d-wave pairing channel in a small region of the parameter space (see App. S3). While, for the parameters considered, this d-wave amplitude is always subleading, it can win over the f-wave superconductivity when longer range repulsion are included. For instance, a next-nearest neighbor $V_2 > 0$ automatically penalizes the f-wave $B - B$ next-nearest neighbor pairing, but can be accommodated by a d-wave superconducting state which features $A - B$ nearest neighbor pairing. Similarly, when the sublattice potential $\Delta_0$ decreases, doped charges start to populate $A$ sites, which may again favor a nearest neighbor d-wave order. Both effects are further enhanced near Van Hove doping, where nodeless d-wave states are provably more stable than nodal f-wave order in the limit $\Delta_0 = 0$ [27].

IV. INTERACTION EXPANSION

In this section, we show that the three-particle mechanism for superconductivity described in Sec. III is not restricted to the atomic limit $t_0 \ll \Delta$, but extends at any value of the tunneling parameter. To preserve analytical control of interband effects, we work in the small interaction limit, such that interband hybridization remains small. In this regime, we find evidence of attractive interband interaction between conduction electrons. Using a similar method than in Sec. III. We follow the methodology of Sec. III. i.e. we first derive an effective model for doped electrons with an unitary transformation, and then study the obtained model for small doping concentrations.

A. Unitary transformation

We start from the Hamiltonian Eq. 1 written in momentum space as $\mathcal{H} = \mathcal{H}_0 + \mathcal{V}$ with

$$\mathcal{H}_0 = \sum_i \varepsilon_i c_i^\dagger c_i, \quad \mathcal{V} = \frac{1}{N_s} \sum_{1234} V_{i2}^{12} c_{i1}^\dagger c_{i3} c_{i4} c_{i2},$$  \hspace{1cm} (14)

where we have used a generalized index $i = (k_i, b_i, \sigma_i)$ gathering the momentum $k_i$, band $b_i = \pm$, and spin $\sigma_i$ labels. Our goal is to derive an effective Hamiltonian for the upper band assuming the lower one fully filled. We do so by eliminating the direct interband mixing interaction terms in the Hamiltonian with the help of a unitary transformation.

This transformation is carried out explicitly in App. S5, where we find the leading corrections to the dispersion relation $\varepsilon$ and the scattering vertex $V$. The
former is akin to the Hartree-Fock correction of standard many-body perturbation theory:

\[
\delta\varepsilon_{k,+} = \frac{1}{N_s} \sum_{q,\sigma} \left[ V^{(k,+\uparrow)(q,-\sigma)} + V^{(q,-\sigma)(k,+\uparrow)} \right] - \frac{1}{N_s} \sum_q \left[ V^{(k,+\uparrow)(q,-\uparrow)} + V^{(q,-\uparrow)(k,+\uparrow)} \right],
\]

where we have used the spin-independence of \( V \), and choose \( \uparrow \) as a preferred spin index. This small correction does not change the position of the band minima, which remain degenerate at the \( K \) and \( K' \) points. Expanding around these minima in the limit \( t_0 \ll \Delta_0 \) gives \((\varepsilon + \delta\varepsilon)_{K+k,+} \simeq |k|^2/(2m^*)\) with

\[
\frac{1}{m^*} = \frac{3t_0^2a^2}{2\Delta_0^2} \left[ \Delta_0 + U_A - 4V_0 \right],
\]

which agrees with the result of Eq. (1) obtained with the kinetic expansion provided interactions are small compared to \( \Delta_0 \). This offers an important consistency test between the two methods in their overlapping regimes of validity.

B. Attraction in dilute limit

The expressions for the corrections to the scattering vertex \( 4V \) being quite involved (see App. S5), we simply support here the emergence of attractive interactions in the \( f \)-wave scattering channel by computing the effective interaction strength in the spin-triplet valley-singlet channel, i.e., between equal-spin electrons living in opposite valleys. Writing \( V = V + \delta V \), this coefficient can be expressed as

\[
U_0 = V^{(K',+)(K'+)} + V^{(K,+)(K')},
\]

where all spin component are equal. This coefficient greatly simplifies as the Bloch states at the \( K \) and \( K' \) point in the upper band are localized on \( B \) sites, \( \Psi_{K/k',+} = 0 \). Using this simplification, we end up with

\[
U_0 = \frac{12t_0^2V_0(U_B - 3V_0)}{N_s} \sum_q \frac{|f(q)|^2}{(2\varepsilon_{q,+})^3},
\]

which also agrees with the our results of Sec. 111 in the limit where both methods are analytically controlled (see App. S5). The presence of an effective attraction \( U_0 < 0 \) in the \( f \)-wave channel when \( U_B < 3V_0 \) proves that the results of Fig. 3 hold for non-perturbative tunneling amplitudes \( t_0 \).

Although the three-particle mechanism for superconductivity introduced in this article is more easily pictured near the atomic limit (Sec. 111), it can also apply to band insulators where the conduction band’s width is substantial compared to the single particle gap. Strong of this generalization, we now compare the predictions of our theory to experimental data on two dilute superconductors.

V. APPLICATION TO DILUTE SUPERCONDUCTORS

Let us recapitulate our results so far. Using the specific model Eq. (1) as an example, we have demonstrated a general mechanism of spin-triplet superconductivity in doped band insulators with strong repulsive interaction. Focusing on the low density regime, we have derived a universal continuum model for the two-valley Fermi liquid. When either tunneling or interactions are small compared to the \( n = 2 \) insulator gap, we have shown that virtual interband transitions can produce a sufficiently strong intervalley ferromagnetic interaction that leads to spin-triplet pairing. A remarkable consequence of our theory is that spin-triplet superconductivity occurs at infinitesimal doping, with a sharp increase \( T_c \sim \sqrt{\epsilon} \).

Based on this theory, we hereby propose that the recently discovered dilute superconductors ZrNCl WTe\(_2\) are spin-triplet.

A. ZrNCl

ZrNCl shares many common features with our toy model Eq. (1). At the single particle level, it is a band insulator with a large gap \( \Delta \sim 2.5eV \) 38 39, where a single band with quadratic minima at the \( K \) and \( K' \) points is relevant to describe the physics at low carrier density 40 41. Upon electron doping, two-dimensional superconductivity arises in the ZrN planes 42 43, even for extremely dilute samples \( \epsilon = 0.0035 \) 25. The critical temperatures ranging from 11.5 K to 19 K is hard to explain by electron-phonon mechanism, which makes ZrNCl an unconventional superconductor. A compelling argument against the phonon-mechanism of pairing is the simultaneous enhancement of \( T_c \) 44 and reduction of electron-phonon interactions, probed by Raman scattering 45, when the doping decreases from 0.2 to 0.05.

We now highlight that our theory consistently captures the available experimental results on the superconducting state in ZrNCl 10, and propose that this material be regarded as a legitimate candidate for spin-triplet pairing.

First, we find pairing at infinitesimal doping above the band insulator, and predict a smooth crossover from a BEC of pairs to a BCS regime as the Fermi energy increases 24, which has recently been observed in extremely high quality ZrNCl samples 25. On the BCS side of this crossover, we expect a gap to critical temperature ratio \( 2\Delta/k_BT_c \sim 3.5 \) 17 19. This is because the induced pairing interaction in Eq. 2 is instantaneous on
the time scale of the inverse bandwidth, so that all carriers in the narrow band contribute to the superconducting gap. This contrasts with phonon-mediated retarded attractions, which only spread over a Debye width near the Fermi level and lead the universal ratio of 3.5. The gap to tractions, which only spread over a Debye width near the gap. This contrasts with phonon-mediated retarded at-

Second, our theory also allows a fine understanding of the critical temperature observed in ZrNCl. Using the DFT results for the lattice constant $a = 3.663\,\text{Å}$ and the effective mass at the $K$ point $m = 0.580m_e$ \cite{51,52}, we estimate $t \approx 0.65\,\text{eV}$. With the parameters $U_B = 4V_0 = 4\,\text{Å}$ of Fig. 4 we obtain a critical temperature $T_c \approx 0.002t \approx 15\,\text{K}$ for 9% doping, which lies very close to the experimentally measured value.

Our theory also successfully captures the non-monotonic dependence of $T_c$ on doping observed in Refs. \cite{28}. The original increase of $T_c$ following from the BEC/BCS physics, while the subsequent decay of $T_c$ when $x = 5\% \rightarrow 20\%$ \cite{44} is interpreted by the larger modulation of the form factor $s_q$ near $\Gamma - M$ lines. For the parameters of Fig. 4, $T_c$ at $x = 0.2$ equals 2/3 of its maximum value, which almost quantitatively agrees with the measured value of 40% \cite{28}.

Finally, specific heat measurements point towards a change from an almost isotropic to a highly anisotropic order parameter upon increasing the doping level \cite{53,54}, which is consistent with an $f$-wave superconducting state. The unconventional pairing symmetry is further substantiated by the lack of coherence peak in the NMR signal of Ref. \cite{55}.

B. WTe$_2$

Monolayer WTe$_2$ is a topological insulator with spin-helical edge states \cite{56,60}. Recently, two independent groups discovered a transition from insulating to superconducting state in monolayer WTe$_2$ under electron doping via electrostatic gating \cite{29,30}. Another separate work observed superconductivity in epitaxial thin film of WTe$_2$ \cite{61}. The origin of superconductivity and the nature of the superconducting state are unknown and have attracted considerable interest \cite{62,63}.

Our picture is that superconductivity in electron doped WTe$_2$ is driven by excitonic effects and exhibits spin-triplet pairing. Indeed, a recent experiment on insulating WTe$_2$ reported evidence of strong excitonic effects which significantly enhance the single-particle gap \cite{64}. While the electronic structure of WTe$_2$ is far more complicated than our model used to illustrate the exciton pairing mechanism, they both feature two valleys in the conduction band. Therefore, electron doped WTe$_2$ at low density is a two valley Fermi liquid described by our continuum theory Eq. 5. If excitons in WTe$_2$ mediate strong enough inter-valley ferromagnetic exchange, our theory should capture the essential physics of its superconducting state.

This motivates a thorough comparison between the expected consequences of spin-triplet superconductivity to experimental findings on WTe$_2$. First, differential resistance measurements \cite{29} shows that upon doping, an insulating resistance peak transforms directly into a superconducting resistance dip, consistent with our picture of an insulator-superconductor transition. The experimentally observed sharp increase of $T_c$ with doping \cite{29} agrees remarkably well with the prediction $T_c \sim \sqrt{\chi}$ in the low density regime, as shown in Fig. 5. Another prediction of our theory is that single-particle gap does not close across the doping induced insulator-superconductor transition, which can be tested in future tunneling measurements.

The scenario of spin-triplet superconductivity in WTe$_2$ is supported by the observation of an in-plane critical field much larger than the Pauli limit in both monolayers and thin films \cite{29,30,61}. The most significant experimental evidence of spin-triplet superconductivity is the initial increase of $T_c$ upon application of an in-plane magnetic field \cite{29,61}. This behavior is incompatible with $s$-wave superconductivity (even after considering the effect of spin-orbit interaction \cite{63}). On the contrary, the enhancement of $T_c$ by magnetic field follows naturally from our equal-spin triplet superconducting state in two-valley systems. This can be seen with the Landau free energy:

$$F = \alpha(d \cdot d^\ast) + \mu B \cdot (i(d \times d^\ast) + \eta) B \cdot d^2 + \chi B^2(d \cdot d^\ast),$$  \hspace{1cm} (19)

with $\alpha = \kappa(T - T_c(B = 0))$ and $\kappa, \mu, \eta, \chi > 0$ near $T_c$. Importantly, the magnetic field $\textbf{B}$ results in a linear Zeeman shift for pairs of total spin $\textbf{S} = i(d \times d^\ast)$. The $\eta$ term describes the preferred equal-spin pairing with $d \perp \textbf{B}$, i.e. the spins are aligned or anti-aligned with the field. The last term accounts for orbital effect of the in-plane $B$ field, which causes pair breaking.

From the Landau free energy, we obtain

$$\Delta T_c^B = \mu B - \chi B^2,$$  \hspace{1cm} (20)
with $\Delta T^c_B = T_c(B) - T_c(0)$. Due to the Zeeman effect on triplet pairs, $T_c$ increases linearly with $B$ at small field. This non-analytic dependence on $B$ is a consequence of the degeneracy of triplet superconducting states associated with spin degrees of freedom at zero field. At such large $B$, orbital effect dominates and reduces $T_c$. As shown in Fig. 5, the above theoretical curve $T_c(B)$ fits excellently with experimental data.

The presence of spin-orbit coupling in WTe$_2$ is expected to lift the degeneracy between different triplet states. Due to a plane mirror symmetry perpendicular to the crystallographic axis, an additional term $F_{\text{SOC}} = 2\gamma |d_x|^2$ is allowed after including the spin-orbit effect. If $\mathbf{B} \parallel b$, the linear increase of $T_c$ is rounded off by $\gamma$ according to

$$\Delta T^c_B = \sqrt{(\mu B)^2 + \gamma^2} - |\gamma| - \chi B^2. \quad (21)$$

For small spin-orbit coupling $\gamma < \mu^2/2\chi$, the Zeeman term dominates, which still enables the enhancement of $T_c$. If $\mathbf{B} \parallel a$, we should distinguish two cases. First if $\gamma > 0$, both spin-orbit and magnetic field favor the vector $\mathbf{d} \propto (0, 1, i)^T$, and we recover Eq. (20). Then for $\gamma < 0$, the original spin degeneracy is fully lifted and we observe a competition between two different triplet states with $\mathbf{d} \parallel a$ and $\mathbf{d} \perp a$, respectively favored at low and high field. This leads to

$$\Delta T^c_B = \begin{cases} - (\eta + \chi) B^2 & \text{if } B < B^* \\ \mu B - |\gamma| - \chi B^2 & \text{if } B > B^* \end{cases}, \quad (22)$$

which exhibits a kink at $B^* \simeq |\gamma|/\mu$ where the system undergo a first-order phase transition. We hope more measurements of $T_c$ as a function of both in-plane magnetic field strength and direction can be performed to establish the highly unusual behaviour predicted by Eqs. (20) and (22) which we regard as an unambiguous evidence for triplet superconductivity.

VI. ACKNOWLEDGEMENT

We thank Kin Fai Mak, Lu Li, Valla Fatemi, Pablo Jarillo-Herrero, Sanfeng Wu, Ali Yazdani, Atac Imomgul and Yoshi Iwasa for valuable discussions. We thank Tomoya Asaba and Valla Fatemi for providing raw experimental data for use in Fig. [5].

This work was supported by DOE Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under Award DE-SC0018945. VC gratefully acknowledges support from the MathWorks fellowship. LF was supported in part by a Simons Investigator Award from the Simons Foundation.

[1] V. Vakaryuk and A. J. Leggett, Spin polarization of half-quantum vortex in systems with equal spin pairing, Physical review letters 103, 057003 (2009).
[2] M. Salomaa and G. Volovik, Quantized vortices in superfluid he 3, Reviews of modern physics 59, 533 (1987).
[3] N. Read and D. Green, Paired states of fermions in two dimensions with breaking of parity and time-reversal symmetries and the fractional quantum hall effect, Physical review B 61, 10267 (2000).
[4] A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. Ludwig, Classification of topological insulators and superconductors in three spatial dimensions, Physical Review B 78, 195125 (2008).
[5] T. H. Hsieh and L. Fu, Majorana fermions and exotic surface andreev bound states in topological superconductors: application to cu x bi 2 se 3, Physical review letters 108, 107005 (2012).
[6] J. Alicea, New directions in the pursuit of majorana fermions in solid state systems, Reports on progress in physics 75, 076501 (2012).
[7] C. Beenakker, Search for majorana fermions in superconductors, Annu. Rev. Condens. Matter Phys. 4, 113 (2013).
[8] S. D. Sarma, M. Freedman, and C. Nayak, Majorana zero modes and topological quantum computation, npj Quantum Information 1, 1 (2015).
[9] V. Crepel, B. Estienne, and N. Regnault, Variational ansatz for an abelian to non-abelian topological phase transition in $\nu = 1/2 \pm 1/2$ bilayers, Physical review letters 123, 126804 (2019).
[10] A. J. Leggett, A theoretical description of the new phases of liquid he 3, Reviews of Modern Physics 47, 331 (1975).
[11] A. P. Mackenzie and Y. Maeno, The superconductivity of sr 2 ruo 4 and the physics of spin-triplet pairing, Reviews of Modern Physics 75, 657 (2003).
[12] J. Sauls, The order parameter for the superconducting phases of upt3, Advances in Physics 43, 113 (1994).
[13] S. Ran, C. Eckberg, Q.-P. Ding, Y. Furukawa, T. Metz, S. R. Saha, I.-L. Liu, M. Zic, H. Kim, J. Paglione, et al., Nearly ferromagnetic spin-triplet superconductivity, Science 365, 684 (2019).
[14] L. Jiao, S. Howard, S. Ran, Z. Wang, J. O. Rodriguez, M. Sigrist, Z. Wang, N. P. Butch, and V. Madhavan, Chiral superconductivity in heavy-fermion metal ute 2, Nature 579, 523 (2020).
[15] K. Matano, M. Kriener, K. Segawa, Y. Ando, and G.-q. Zheng, Spin-rotation symmetry breaking in the superconducting state of cu x bi 2 se 3, Nature physics 12, 852 (2016).
[16] S. Yonezawa, K. Tajiri, S. Nakata, Y. Nagai, Z. Wang, K. Segawa, Y. Ando, and Y. Maeno, Thermodynamic evidence for nematic superconductivity in cu x bi 2 se 3, Nature Physics 13, 123 (2017).
[17] K. Willa, R. Willa, K. W. Song, G. Gu, J. A. Schneeloch, R. Zhong, A. E. Koshelev, W.-K. Kwok, and U. Welp, Nanocalorimetric evidence for nematic superconductivity in the doped topological insulator sr 0.1 bi 2 se 3, Physical Review B 98, 184509 (2018).
[18] C.-w. Cho, J. Shen, J. Lyu, O. Atanov, Q. Chen, S. H. Lee, Y. San Hor, D. J. Gawryluk, E. Pomjakushina, M. Bartkowiak, et al., Z 3-vestigial nematic order due to superconducting fluctuations in the doped topological insulators nb x b i 2 se 3 and cu x b i 2 se 3, Nature communications 11, 1 (2020).

[19] L. Fu and E. Berg, Odd-parity topological superconductors: theory and application to cu x b i 2 se 3, Physical review letters 105, 097001 (2010).

[20] T. Hashimoto, K. Yada, A. Yamakage, M. Sato, and Y. Tanaka, Bulk electronic state of superconducting topological insulator, Journal of the Physical Society of Japan 82, 044704 (2013).

[21] L. Fu, Odd-parity topological superconductor with nematic order: Application to cu b i 2 se 3, Phys. Rev. B 90, 100509 (2014).

[22] X. Wan and S. Y. Savrasov, Turning a band insulator into an exotic superconductor, Nature communications 5, 1 (2014).

[23] V. Crèpel and L. Fu, New mechanism and exact theory of superconductivity from strong repulsive interaction, Science Advances 7, eabg2233 (2021).

[24] W. Little, Possibility of synthesizing an organic superconductor, Physical Review 134, A1416 (1964).

[25] V. L. Ginzburg and D. Kirzhnits, On the problem of high temperature superconductivity, Physics Reports 4, 343 (1972).

[26] D. Allender, J. Bray, and J. Bardeen, Model for an excitation mechanism of superconductivity, Physical Review B 7, 1020 (1973).

[27] J. Hirsch and D. Scalapino, Enhanced superconductivity in quasi two-dimensional systems, Physical review letters 56, 2732 (1986).

[28] Y. Nakagawa, Y. Kasahara, T. Nomoto, R. Arita, T. Nonjima, and Y. Iwasa, Gate-controlled bcs-bec crossover in a two-dimensional superconductor, Science 372, 190 (2021).

[29] 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, 922 (2018).

[30] 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).

[31] J. R. Schrieffer and P. A. Wolff, Relation between the anderson and kondo hamiltonians, Physical Review 149, 491 (1966).

[32] K. Slagle and L. Fu, Charge transfer excitations, pair density waves, and superconductivity in moiré materials, Physical Review B 102, 235423 (2020).

[33] J. Levinse and M. M. Parish, Strongly interacting two-dimensional fermi gases, in Annual review of cold atoms and molecules (World Scientific, 2015) pp. 1–75.

[34] M. Randeria, J.-M. Duan, and L.-Y. Shieh, Superconductivity in a two-dimensional fermi gas: Evolution from cooper pairing to bose condensation, Physical Review B 41, 327 (1990).

[35] L.-Y. Xiao, S.-L. Yu, W. Wang, Z.-J. Yao, and J.-X. Li, Possible singlet and triplet superconductivity on honeycomb lattice, EPL (Europhysics Letters) 115, 27008 (2016).

[36] K. Kuroki, Spin-fluctuation-mediated d+ i d’ pairing mechanism in doped beta - m ncl (m= hf, zn) superconductors, Physical Review B 81, 104502 (2010).

[37] R. Nandkishore, R. Thomale, and A. V. Chubukov, Superconductivity from weak repulsion in hexagonal lattice systems, Physical Review B 89, 144501 (2014).

[38] R. Weht, A. Filippetti, and W. Pickett, Electron doping in the honeycomb bilayer superconductors (zr, hf) ncl, EPL (Europhysics Letters) 48, 320 (1999).

[39] I. Hase and Y. Nishihara, Electronic structure of superconducting layered zirconium and hafnium nitride, Physical Review B 60, 1573 (1999).

[40] R. Heid and K.-P. Bohnen, Ab initio lattice dynamics and electron-phonon coupling in li x zncl, Physical Review B 72, 134527 (2005).

[41] Z. Yin, A. Kutepov, and G. Kotliar, Correlation-enhanced electron-phonon coupling: Applications of g w and screened hybrid functional to bismuthates, chloronitriles, and other high-t c superconductors, Physical Review X 3, 021011 (2013).

[42] H. Tou, Y. Maniwa, K. Kohwasaki, and S. Yamanaka, Evidence for quasi-two-dimensional superconductivity in electron-doped li 0.48 (thf) y hfncl, Physical Review B 63, 020508 (2000).

[43] H. Tou, Y. Tanaka, M. Sera, Y. Taguchi, T. Sasaki, Y. Iwasa, L. Zhu, and S. Yamanaka, Upper critical field in the electron-doped layered superconductor zn cl 0.7: Magnetoresistance studies, Physical Review B 72, 020501 (2005).

[44] T. Takano, T. Kishiume, Y. Taguchi, and Y. Iwasa, Interlayer-spacing dependence of t c in li x m ncl (m= zn, br, i), Physica C: Superconductivity and its Applications 514, 354 (2015).

[45] M. M. Parish, The bcs-bec crossover, in Quantum Gas Experiments: Exploring Many-Body States (World Scientific, 2015) pp. 179–197.

[46] L. Gor’kov and T. Melik-Barkhudarov, Contribution to the theory of superfluidity in an imperfect fermi gas, Sov. Phys. JETP 13, 1018 (1961).

[47] A. V. Chubukov, I. Eremin, and D. V. Efremov, Superconductivity versus bound-state formation in a two-band superconductor with small fermi energy: Applications to fe pnictides/chalcogenides and doped srtio 3, Physical Review B 93, 174516 (2016).

[48] T. Ekino, A. Sugimoto, A. M. Gabovich, Z. Zheng, and S. Yamanaka, Superconducting β-zrncx probed by scanning-tunnelling and break-junction spectroscopy, Physica C: Superconductivity 494, 89 (2013).

[49] H. Tanaka, K. Suzuki, H. Usui, and K. Kuroki, Minimal electronic model for a layered nitride halide superconductor β-zrnc, Journal of the Physical Society of Japan 84, 124706 (2015).

[50] W. S. Yun and J. D. Lee, Two-dimensional semiconductors zncl and hfncl: Stability, electric transport, and thermoelectric properties, Scientific reports 7, 1 (2017).

[51] Y. Kasahara, T. Kishiume, T. Takano, K. Kobayashi,
E. Matsuoka, H. Onodera, K. Kuroki, Y. Taguchi, and Y. Iwasa, Enhancement of pairing interaction and magnetic fluctuations toward a band insulator in an electron-doped Li$_x$ZrNiCl$_2$ superconductor, Physical review letters 103, 077004 (2009).

[54] Y. Taguchi, M. Hisakabe, and Y. Iwasa, Specific heat measurement of the layered nitride superconductor Li$_x$ZrNiCl$_2$, Physical review letters 94, 217002 (2005).

[55] H. Kotegawa, S. Oshiro, Y. Shimizu, H. Tou, Y. Kasahara, T. Kishitome, Y. Taguchi, and Y. Iwasa, Strong suppression of coherence effect and appearance of pseudogap in the layered nitride superconductor Li$_x$ZrNiCl$_2$, Physical Review B 90, 020503 (2014).

[56] X. Qian, J. Liu, L. Fu, and J. Li, Quantum spin hall effect in two-dimensional transition metal dichalcogenides, Science 346, 1344 (2014).

[57] Z. Fei, T. Palomaki, S. Wu, W. Zhao, X. Cai, B. Sun, P. Nguyen, J. Finney, X. Xu, and D. H. Cobden, Edge conduction in monolayer WTe$_2$, Nature Physics 13, 677 (2017).

[58] S. Tang, C. Zhang, D. Wong, Z. Pedramrazi, H.-Z. Tsai, C. Jia, B. Moritz, M. Claassen, H. Ryu, S. Kahn, et al., Quantum spin hall state in monolayer 1t'-WTe$_2$, Nature Physics 13, 683 (2017).

[59] S. Wu, V. Fatemi, Q. D. Gibson, K. Watanabe, T. Taniguchi, R. J. Cava, and P. Jarillo-Herrero, Observation of the quantum spin hall effect up to 100 kelvin in a monolayer crystal, Science 359, 76 (2018).

[60] A. Lau, R. Ray, D. Varjas, and A. R. Akhmerov, Influence of lattice termination on the edge states of the quantum spin hall insulator monolayer 1 t'-wte 2, Physical Review Materials 3, 054206 (2019).

[61] 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, Scientific reports 8, 1 (2018).

[62] Y.-T. Hsu, W. S. Cole, R.-X. Zhang, and J. D. Sau, Inversion-protected higher-order topological superconductivity in monolayer wte 2, Physical Review Letters 125, 097001 (2020).

[63] Y.-M. Xie, B. T. Zhou, and K. T. Law, Spin-orbit-parity-coupled superconductivity in topological monolayer wte 2, Physical Review Letters 125, 107001 (2020).

[64] Y. Jia, P. Wang, C.-L. Chiu, Z. Song, G. Yu, B. Jäck, S. Lei, S. Klemenz, F. A. Cevallos, M. Onyszczak, et al., Evidence for a monolayer excitonic insulator, arXiv preprint arXiv:2010.05390 (2020).
S1. EFFECTIVE MODEL IN THE ATOMIC LIMIT

A. General framework

In this appendix, we derive the effective Hamiltonian Eq. 2 from the original model Eq. 1 by treating the tunneling terms

$$\mathcal{H}_t = t_0 \sum_{\langle r,r' \rangle, \sigma} \left( \frac{1}{2} c_{r,\sigma}^\dagger c_{r',\sigma} + \frac{1}{2} \right),$$

(E1)

as a perturbation of the ‘classical’ part 32

$$\mathcal{H}_0 = \sum_{r \in B} \frac{U_B}{2} n_r (n_r - 1) + \Delta n_r + \sum_{\langle r,r' \rangle} V_0 n_r n_r',$$

(E2)

thanks to the large staggered potential $\Delta \gg t_0$. To that purpose, we apply the unitary transformation $\mathcal{H}' = e^{iS}He^{-iS}$, with $S$ Hermitian and satisfying 31

$$[\mathcal{H}_0, iS] = \mathcal{H}_t.$$

(E3)

It leads to the following approximation of the Hamiltonian

$$\mathcal{H}' = \mathcal{H}_0 + \frac{1}{2} [iS, \mathcal{H}_t] + \mathcal{O}(\mathcal{H}_t S^2),$$

(E4)

obtained with the Baker-Campbell-Haussdorf formula. To find an explicit representation of $S$, we decompose the tunneling Hamiltonian as

$$\mathcal{H}_t = \sum_{d = \pm 1} \sum_{v = -5}^5 \sum_{u = -1}^1 T_{d,v,u},$$

(E5)

where $T_{d,v,u}$ gathers all tunneling events that change the number of occupied $A$ sites by $d$, the number of nearest neighbor pairs by $v$ and the number of doubly occupied $B$ sites by $u$. In terms of these operators, we find

$$S = -i \sum_{d,v,u} \frac{T_{d,v,u}}{d\Delta + vV_0 + uU_B}. $$

(E6)

Plugging this expression in Eq. (E4) we obtain

$$\mathcal{H}' = \mathcal{H}_0 + \frac{1}{2} \sum_{d,v,u} \sum_{d',v',u'} \frac{T_{d',v',u'} T_{d,v,u}}{d'\Delta + v'V_0 + u'U_B},$$

(E7)

which is valid up to $O(t_0^3/\Delta^2)$ corrections.

B. Projection

The ground state of $\mathcal{H}_0$ for two electrons per unit cell has a singlet on all $A$-sites and $B$-sites completely empty. Due to Pauli exclusion principle, the $x = n - 2$ doped electrons above this insulating state are placed at the $B$-sites. They have an energy per particle $E_f = \Delta + 6V_0$. This low energy manifold, named $f$-band, hybridizes with local excitation having a hole on a $A$-site due to the tunneling part $\mathcal{H}_t$. Such local excitation are separated from the low-energy band by an energy of at least $\Delta$. They are only virtually occupied due to the small ratio $t_0 \ll \Delta$, and their effects on the $f$-electrons’ dynamics can be obtained with Eq. (E7)

Projecting $\mathcal{H}'$ onto the $f$-band requires to have $d' = -d$ and $v' = -v$ in Eq. (E7). Furthermore, the first operator acting on the $f$-band should move an electron from an $A$ to a $B$ site, i.e. the rightmost $T_{d,v,u}$ must have $d = 1$ and $v, u \geq 0$. This gives

$$\mathcal{H}' \simeq \mathcal{H}_0 - \frac{1}{2} \sum_{v,u,u' \geq 0} \sum_{d = 1} \sum_{v' = -v} T_{-1,-u',-v} T_{1,u,v}$$

$$\times \left[ \frac{1}{\Delta + vV_0 + uU_B} + \frac{1}{\Delta + vV_0 + u'U_B} \right].$$

(E8)

This Hamiltonian can be recast as a tight binding Hamiltonian for the $f$-electrons on the triangular lattice, with density-assisted hopping and local interactions:

$$\mathcal{H}' = U_B \sum_{i \in B} \frac{n_i (n_i - 1)}{2} + \sum_{ijk \in \triangle} \left[ \tilde{V}_{ijk} + \mathcal{O}_{ijk} \right]$$

$$+ \sum_{ijk \in \triangle, \sigma} \left[ f_{j,\sigma} \tilde{T}_{ijk} f_{i,\sigma} + P_{ijk} \right],$$

(E9)

where sums run over upper triangles with vertices $ijk$, while $\mathcal{O}_{ijk}$ and $P_{ijk}$ respectively denote cyclic and all permutations of $ijk$. The density dependent interaction and tunneling operators read

$$\tilde{V}_{ijk} = -t_0^2 \frac{(2 - n_k)}{\Delta + (5 - n_{\Delta}V_0 + n_kU_B)},$$

(E10)

$$\tilde{T}_{ijk} = \frac{t_{\Delta,i} + t_{\Delta,j}}{2}, \quad t_{\Delta,\ell} = \frac{t_0^2}{\Delta + (4 - n_{\Delta})V_0 + n_{\Delta}U_B},$$

with $n_{\Delta} = n_i + n_j + n_k$, which are symmetric under the exchange $i \leftrightarrow j$. These interaction coefficients and density-dependent tunneling amplitudes can be expressed in terms of sum and product of density operators. We now simplify their expression in two particular cases.
C. Dilute limit

At small doping concentration, we can discard states with more than two fermions on the same triangle, which only appear with negligible probability. The tunneling coefficient $\hat{T}_{ij,k}$ is thus restricted to cases where $(n_i + n_j = 0, n_k = 0)$, $(n_i + n_j = 1, n_k = 0)$ or $(n_i + n_j = 0, n_k = 1)$. Projecting on these configurations, we find the equivalent representation

$$\hat{T}_{ij,k} = T_{0,0} + n_k[T_{0,1} - T_{0,0}] + (n_i + n_j)[T_{1,0} - T_{0,0}].$$

Summing over all possible triangle, we can rewrite

$$\sum_{ijk \in \Delta} \left[ f_{j,\sigma}^\dagger \hat{T}_{ij,k} f_{i,\sigma} + P_{ijk} \right] = \sum_{(i,j),\sigma} \left[ f_{j,\sigma}^\dagger \left[ t + \frac{n_i + n_j}{2} \right] f_{i,\sigma} + hc \right] + \lambda \sum_{ijk \in \Delta,\sigma} \left[ f_{j,\sigma}^\dagger n_k f_{i,\sigma} + P_{ijk} \right],$$

where we have introduced the coefficients

$$t = \frac{t_0^2}{\Delta + 4V}, \quad \tilde{t} = \frac{t_0^2}{\Delta + 3V_0 + U_B}, \quad \frac{t_0^2}{\Delta + 3V_0}, \quad \frac{2t_0^2}{\Delta + 4V_0}, \quad \lambda = \frac{t_0^2}{\Delta + 3V_0} - \frac{t_0^2}{\Delta + 4V_0}.$$  \hfill (E13)

Similarly, we can project the interaction terms on configurations having $n_k, (n_i + n_j), (n_i + n_j + n_k) \leq 2$:

$$\hat{V}_{ij,k} = \hat{V}_{i,j,0} + n_k[\hat{V}_{i,j,1} - \hat{V}_{i,j,0}] + \frac{n_k(n_k - 1)}{2}[\hat{V}_{0,0} - 2\hat{V}_{0,1}]$$

$$= \hat{V}_{0,0} + (n_i + n_j)[\hat{V}_{1,0} - \hat{V}_{0,0}] + \frac{n_i(n_i - 1) + n_j(n_j - 1) + 2n_i n_j}{2}[\hat{V}_{2,0} - 2\hat{V}_{1,0} + \hat{V}_{0,0}]$$

$$+ n_k[\hat{V}_{0,1} - \hat{V}_{0,0} + (n_i + n_j)(\hat{V}_{1,1} - \hat{V}_{1,0} - \hat{V}_{0,1} + \hat{V}_{0,0})] + \frac{n_k(n_k - 1)}{2}[\hat{V}_{0,0} - 2\hat{V}_{0,1}].$$

Up to a global constant and a shift of chemical potential, this expansion leads to

$$U_B \sum_{i \in \mathbb{B}} \frac{n_i(n_i - 1)}{2} + \sum_{ijk \in \Delta} \left[ \hat{V}_{ij,k} + \psi_{ijk} \right] = U \sum_i \frac{n_i(n_i - 1)}{2} + V \sum_{(i,j)} n_i n_j,$$

where the coefficients read

$$U_f = U + 2\hat{V}_{2,0} - 4\hat{V}_{1,0} - 2\hat{V}_{0,0} + 3\hat{V}_{0,1}, \quad V_f = \hat{V}_{2,0} + 2\hat{V}_{1,1} - 4\hat{V}_{1,0} - 2\hat{V}_{0,1} + 3\hat{V}_{0,0}.$$  \hfill (E18)

In terms of the lattice parameters $t_0, \Delta, V_0$ and $U_B$, we find them equal to

$$V = \frac{4t_0^2 V_0(2V_0 + \Delta)}{(\Delta + 3V_0)(\Delta + 4V_0)}, \quad \lambda = \frac{2t_0^2 V_0}{(\Delta + 3V_0)(\Delta + 4V_0)},$$

$$U = U_B - \frac{2t_0^2 (16V_0^2 + 7V_0 \Delta + \Delta^2)}{(\Delta + 3V_0)(\Delta + 4V_0)(\Delta + 5V_0)} + \frac{2t_0^2}{\Delta + 4V_0 + U_B}. $$

Gathering the various terms, we obtain Eq. 2 of the main text.

D. Large $U$ limit

Assuming $U_B \gg t_0$, we can project the effective Hamiltonian to the $f$-band with no double occupancy. Restricting to $n_i \leq 1$ in the above equations yields the effective Hamiltonian

$$\mathcal{H}' = \sum_{(i,j),\sigma} t \left( f_{j,\sigma}^\dagger f_{j,\sigma} + hc \right) + V n_i n_j + \lambda \sum_{(ijk) \in \Delta} \left( \sum_{\sigma} f_{j,\sigma}^\dagger n_k f_{i,\sigma} + P_{ijk} \right) + U_3 n_i n_j n_k,$$

where the coefficients $t$, $\lambda$ and $V$ have the same form as above. The three-body interaction terms read

$$U_3 = \frac{12t_0^2 V_0^2}{(\Delta + 3V_0)(\Delta + 4V_0)(\Delta + 5V_0)} - \frac{6t_0^2 V_0^2}{(\Delta + 2V_0 + U_B)(\Delta + 3V_0 + U_B)(\Delta + 4V_0 + U_B)}.$$  \hfill (E22)

S2. TWO-PARTICLE LATTICE CALCULATION

In this appendix, we solve the effective model Eq. 2 for two particles – the analog of Cooper’s problem on the lattice. To do so, we separate the center of mass moment-
Due to the quadratic band dispersion near the $K$ point with $N$ the sign difference: the exchange of the two particles, which translates into between these spin configurations is their statistic under the states $S$ of the system when fermions remain close to these two valleys. 

As shown in Fig. 1c, the kinetic part of the effective Hamiltonian Eq. 2 dominates over interactions. Thus, low-energy fermions live near the two degenerate minima of the single-particle dispersion relation located at the $K$ points in the Brillouin Zone. Our goal here is to derive an effective continuum field theory capturing the physics of the system when fermions remain close to these two valleys.

We then solve this equation numerically for large enough system sizes to extract the ground state energy in each spin sector. Our solution are shown in Fig. 2 of the main text.

While our original model does not include any direct repulsion between $B$ sites on the honeycomb lattice, the two-particle bound state we have established is robust against longer range interactions. To study their effect, we further add non-local interaction between conduction electrons to the effective Hamiltonian $\mathcal{H}'$ and re-solve the two-particle problem. We find that bound state is destroyed only when the non-local repulsion becomes comparable to the exciton-induced short-range pairing interaction (which is much larger than the binding energy $\epsilon_a$). If we take into account the direct Coulomb repulsion between nearest-neighbor $B$ sites $V'$, bound state persists for $V' < 2\lambda - V$, or $0.25eV$ when the parameters mentioned in the main text are used. If we include the long-range Coulomb interaction $\frac{e^2}{\epsilon r}$ fully, bound state exists for $ca > 86.4\text{Å} (a$ is the lattice constant), which corresponds to $e^2/\epsilon a = 16.5mEV$. Thus, in order for electron pairing to occur in the limit of vanishing doping, it is helpful to have a large $\epsilon$ which can result from dielectric screening by a different band.

**S3. CONTINUUM LIMIT**

As shown in Fig. 1, the kinetic part of the effective Hamiltonian Eq. 2 dominates over interactions. Thus, low-energy fermions live near the two degenerate minima of the single-particle dispersion relation located at the $K$ and $K'$ points in the Brillouin Zone. Our goal here is to derive an effective continuum field theory capturing the physics of the system when fermions remain close to these two valleys.

We start with the momentum representation of Eq. 2

$$\mathcal{H}_f = \sum_{k,\sigma} \varepsilon_k f^\dagger_{k,\sigma} f_{k,\sigma} + \frac{1}{2N_s} \sum_{k,q,p,\sigma,\sigma'} V_{k,q} f^\dagger_{k,\sigma} f^\dagger_{q+p,\sigma'} f_{q+p,\sigma'} f_{q,\sigma}$$

(E26)

with $N_s$ the number of unit cells in the lattice, $\varepsilon_k = 2t \sum_{j=1}^3 \cos(k \cdot a_j)$ and

$$V_{k,q} = U + 2V \sum_j \cos[(k - q) \cdot a_j] + 2\lambda \sum_j (e^{ika_j + iqa_{j-1}} + e^{-ika_{j-1} - iqa_j}) + 2t \sum_j [\cos(k \cdot a_j) + \cos(q \cdot a_j)].$$

(E27)

Due to the quadratic band dispersion near the $K$ and $K'$ points, low energy fermions acquire an effective mass $m = 2/(3t_f a^2)$. They also carry an effective $SU(4)$ $\{\uparrow K, \downarrow K, \uparrow K', \downarrow K'\}$ that distinguishes both their spin and their
valley degeneracy and enable contact interactions between fermions with the same spin, provided they have opposite valley index.

Let us now focus on the scattering properties of these low energy fermions. Due to momentum conservation, two incoming low-energy fermions from the same valley can only scatter into a pair of fermions living in the same valley. The corresponding vertex interaction reads

\[ V_c = V_{K,K} = V_{K',K'} = 6(V - \lambda - \tilde{t}) + U. \]  

(E28)

When the electrons are in opposite valley \( K \) and \( K' \), they can scatter to a pair in \( K \) and \( K' \) with the same valley preserving interaction strength \( V_c \), or exchange valley to end up in \( K' \) and \( K \) through the vertex

\[ V_x = V_{K,K'} = V_{K',K} = 3(4\lambda - 2\tilde{t} - V) + U. \]  

(E29)

Introducing different fields for the two valleys

\[ f_{k,\sigma} = \begin{cases} \psi_{k,\sigma,K} & \text{if } k \text{ near } K, \\ \psi_{k,\sigma,K'} & \text{if } k \text{ near } K', \end{cases} \]  

(E30)

and accounting for the \( V_c \) and \( V_x \) terms, we find that the following effective interacting Hamiltonian

\[ \mathcal{H}_{\text{int}} = \frac{V_c}{N_s} \sum_{k,q,p} \psi_{k,\uparrow,V}^\dagger \psi_{p-k,\downarrow,V} \psi_{p-q,\downarrow,V} \psi_{q,\uparrow,V} + \frac{V_c - V_x}{N_s} \sum_{k,q,p} \psi_{k,\sigma,K}^\dagger \psi_{p-k,\sigma,K'}^\dagger \psi_{p-q,\sigma,K'} \psi_{q,\sigma,K} + \frac{V_c}{N_s} \sum_{k,q,p} \psi_{k,\uparrow,V}^\dagger \psi_{p-k,\downarrow,V} \psi_{p-q,\downarrow,V} \psi_{q,\uparrow,V}. \]  

(E31)

Let us rearrange these terms in terms of pair operators to make their physical meaning clearer. Valley-polarized spin-singlet electron pairs \( S_V = f_{\uparrow,\downarrow}^\dagger f_{\downarrow,\uparrow}^\dagger \) with \( V = K, K' \) only feel the valley conserving term and therefore exhibit repulsive interaction (\( V_c > 0 \)). When incoming electrons occupy opposite valleys, the ferromagnetic exchange leads to a total interaction strength \( V_c - (-1)^S V_x \) depending on the total spin \( S \) of the pair. As a consequence, the last spin-singlet valley-triplet channel \( S_0 = (f_{K',\downarrow} f_{K,\uparrow} - f_{K,\uparrow} f_{K',\downarrow})/\sqrt{2} \) is also repulsive (\( V_c + V_x > 0 \)). On the contrary, the three valley-singlet spin-triplet pair states, \( T_\sigma = f_{K,\sigma}^\dagger f_{K,\sigma} \) with \( \sigma = \uparrow, \downarrow \) and \( T_0 = (f_{K',\downarrow} f_{K,\uparrow} + f_{K',\uparrow} f_{K,\downarrow})/\sqrt{2} \), all display a low-energy interaction strength \( V_c - V_x = 9(V - 2\lambda) \), which is negative for a wide range of parameter, as shown in Fig. 3b. To summarize, we can rewrite the different contact interaction terms as

\[ \tilde{H} = \int dx \sum_{\sigma,V} \psi_{\sigma,V}^\dagger \left[ -\frac{\nabla^2}{2m} + \frac{1}{2} \rho_{\sigma,K} \rho_{K'} \right] \psi_{\sigma,V} + \int dx \frac{A}{\mathcal{A}} [(V_c - V_x) (T_\uparrow^\dagger T_\uparrow + T_\downarrow^\dagger T_\downarrow + T_0^\dagger T_0) + V_c (S_{K'}^x S_{K'} + S_{K}^x S_{K}) + (V_c + V_x) S_{K'}^x S_{K}] \]  

(E32)

with \( \mathcal{A} = 2/\sqrt{3}a^2 \) the Brillouin zone area. This effective field theory describes a four-component Fermi liquid with repulsive interactions in the spin-singlet channel, owing to the large on-site interaction \( U \) which appears in both \( V_c \) and in \( V_c + V_x \), and attractive interaction between fermions with total spin one when \( V_c - V_x < 0 \) (see Fig. 3b). This is the conclusion drawn in Eq. 4 of the main text.

Alternatively, we can replace pair operators by more physical quantities, such as the total density on each valley \( \rho_V = \psi_{\uparrow,V}^\dagger \psi_{\uparrow,V} + \psi_{\downarrow,V}^\dagger \psi_{\downarrow,V} \) and the total spin on each valley \( s_V = \psi_{\alpha,V}^\dagger \sigma_{\alpha,\beta} \psi_{\beta,V} \). Together, they allow to represent the exchange term as

\[ T_\uparrow^\dagger T_\uparrow + T_\downarrow^\dagger T_\downarrow + T_0^\dagger T_0 - S_{0}^x S_0 = 2s_K \cdot s_{K'} + \frac{1}{2} \rho_K \rho_{K'}. \]  

(E33)

The valley conserving terms present in all interaction channels can be simply with the total density \( \rho_\text{tot} = \rho_K + \rho_{K'} \)

\[ T_\uparrow^\dagger T_\uparrow + T_\downarrow^\dagger T_\downarrow + T_0^\dagger T_0 + S_0^x S_0 + S_K^x S_K + S_{K'}^x S_{K'} = \frac{1}{2} \rho_\text{tot} (\rho_\text{tot} - 1). \]  

(E34)

Together, they allow to rewrite the interaction part of the continuum Hamiltonian as

\[ \tilde{H}_i = \frac{1}{2A} \int dx [V_c n_{\text{tot}} (n_{\text{tot}} - 1) - V_x (4s_K \cdot s_{K'} + n_K n_{K'})]. \]  

(E35)
This forms makes clear the ferromagnetic interactions between opposite valleys, which are responsible for the formation of triplet pairs. Expanding the total density as a function of $\rho_K$ and $\rho_{K'}$, we find the three coupling constant given in Eq. [37] of the main text

$$g_0 = V_e/(2A), \quad g_1 = (2V_e - V_s)/(2A), \quad g_2 = -2V_s/A.$$  \hfill (E36)

### S4. MEAN-FIELD THEORY OF SUPERCONDUCTIVITY

In this appendix, we carry out a mean-field treatment of the effective Hamiltonian Eq. [226] to investigate its superconducting behavior. With the mean-field substitution $f_{q',\sigma'} f_{q,\sigma} \simeq \delta_{q+q'} \langle f_{q',\sigma'} f_{q,\sigma} \rangle$, we get the following quadratic mean-field approximation:

$$\mathcal{H}_{mf} = \sum_{k,\sigma} \xi_k f_{q,\sigma} f_{q,\sigma} + \frac{1}{2} \sum_{k,\sigma,\sigma'} \left[ \tilde{\Delta}_{k,\sigma\sigma'} f_{q,\sigma}^\dagger f_{q,\sigma'}^\dagger + h.c. \right],$$

$$\tilde{\Delta}_{k,\sigma\sigma'} = -\frac{1}{N_s} \sum_q V_{k,q} \langle f_{q,\sigma} f_{-q,\sigma'} \rangle,$$  \hfill (E37)

with $\xi_k = \xi_{-k} = \varepsilon_k - \mu$ and $\mu$ the chemical potential. It can be rewritten as a sum over a halved Brillouin Zone (denoted with primed sums and products below):

$$\mathcal{H}_{mf} = \sum_k \left[ f_{k}^\dagger f_{-k}^\dagger \right] \left[ \xi_k - \tilde{\Delta}_{k} \right] \left[ f_{k} f_{-k} \right].$$  \hfill (E38)

The order parameters have been gathered in a $2 \times 2$ matrix

$$\Delta_k = \frac{\tilde{\Delta}_k - \tilde{\Delta}_k^T}{2} = -\frac{1}{N_s} \sum_q \text{Re} (V_{k,q}) \langle f_{q,\sigma} f_{-q,\sigma'} \rangle,$$  \hfill (E39)

and should be computed self-consistently.

#### A. Pairing symmetries

The explicit expression of $V_{k,q}$ allows to decompose this order parameter into spin-singlet and spin-triplet components

$$\Delta_k = \Delta'_s + \sum_{j=1}^3 \Delta'_j \cos(k \cdot a_j) + \Delta'_j \sin(k \cdot a_j),$$  \hfill (E40)

which respectively read:

$$\Delta'_s = -\frac{1}{N_s} \sum_q \left[ U + 2\tilde{\tau} (c_{1} + c_{2} + c_{3}) \right] \langle f_{q,\sigma} f_{-q,\sigma'} \rangle,$$

$$\Delta'_j = -\frac{2}{N_s} \sum_q \left[ \lambda (c_{j-1} + c_{j+1}) + V c_{j} + \tilde{\tau} \right] \langle f_{q,\sigma} f_{-q,\sigma'} \rangle,$$

$$\Delta'_j = \frac{2}{N_s} \sum_q \left[ \lambda (s_{j-1} + s_{j+1}) - V s_{j} \right] \langle f_{q,\sigma} f_{-q,\sigma'} \rangle,$$  \hfill (E41)

with $c_j = \cos(q \cdot a_j)$ and $s_j = \sin(q \cdot a_j)$. We can further split these order parameters in terms of irreducible representation of $C_3v$ that they represent on the triangular lattice. For the singlet and triplet case, there are two one-dimensional irrep, only one of which can be obtained because of the particular form of $V_{k,q}$ and one two dimensional irrep. The two former measure the strength of s-wave and f-wave pairing

$$\Delta_{s} = \frac{1}{3} \left[ \Delta_{1}^{s/f} + \Delta_{2}^{s/f} + \Delta_{3}^{s/f} \right],$$  \hfill (E42)

while the two dimensional irrep are related to d-wave and p-wave pairing

$$\Delta_{d,x} = \frac{1}{3} \left[ \Delta_{1}^{d/x} + \Delta_{2}^{d/x} - 2\Delta_{3}^{d/x} \right],$$  \hfill (E43)

$$\Delta_{d,y} = \frac{1}{3} \left[ \Delta_{1}^{d/y} - \Delta_{2}^{d/y} \right].$$  \hfill (E44)

The inverse transformations are

$$\Delta_{1} = \Delta_{s} + \Delta_{d,2x-y^2/p_x} + \Delta_{d,y/p_y},$$

$$\Delta_{2} = \Delta_{s} - \Delta_{d,2x-y^2/p_x} - \Delta_{d,y/p_y},$$

$$\Delta_{3} = -2\Delta_{d,2x-y^2/p_x}.$$  \hfill (E45)

Finally, singlet pairs cannot be of equal spin, and we can therefore express them as a scalar times the $2 \times 2$ matrix $i\sigma_y$, e.g. $\Delta_s = d_{0}(i\sigma_y)$ with $d_{0}$ a complex number. Triplet on the other hand, take the form of a Pauli vector multiplied by $i\sigma_y$, e.g. $\Delta_f = (d_f \cdot \sigma)(i\sigma_y)$.

Our mean-field treatment relies on the self-consistent computation of four scalars related to singlet pairing in s-wave ($\Delta_s, \Delta'_s$), d-wave ($\Delta_{d,2x-y^2}, \Delta_{d,y}$), and three vectors describing f-wave ($\Delta_f$) or p-wave ($\Delta_{p_x}, \Delta_{p_y}$) pairs.
The corresponding self-consistent equations become

\[ \Delta_s' = \frac{-1}{N_s} \sum_q \left[ U + 2\tilde{t}(c_1 + c_2 + c_3) \right] \langle f_q, \sigma f_{-q}, \sigma' \rangle, \]

\[ \Delta_s = \frac{-2}{N_s} \sum_q \left[ \tilde{t} + \frac{2\lambda + V}{3} (c_1 + c_2 + c_3) \right] \langle f_q, \sigma f_{-q}, \sigma' \rangle, \]

\[ \Delta_{p_x} = \frac{-2(\lambda + V)}{N_s} \sum_q \frac{s_1 + s_2 - 2s_3}{6} \langle f_q, \sigma f_{-q}, \sigma' \rangle, \]

\[ \Delta_{p_y} = \frac{-2(\lambda + V)}{N_s} \sum_q \frac{s_1 - s_2}{2} \langle f_q, \sigma f_{-q}, \sigma' \rangle, \]

\[ \Delta_{d_{12}} = \frac{2(\lambda - V)}{N_s} \sum_q \frac{c_1 - c_2}{6} \langle f_q, \sigma f_{-q}, \sigma' \rangle, \]

\[ \Delta_{d_{xy}} = \frac{2(\lambda - V)}{N_s} \sum_q \frac{s_1 + s_2 + s_3}{3} \langle f_q, \sigma f_{-q}, \sigma' \rangle. \] (E46)

### B. Self-consistent conditions

The mean field quadratic Hamiltonian can be diagonalized by a Bogoliubov transformation. Writing the hermitian matrix

\[ \Delta_q \Delta_q^\dagger = a_0 + \mathbf{a} \cdot \mathbf{\sigma} \] (E47)

as a Pauli vector, the eigen-energies read

\[ E_{q,\pm} = \sqrt{\xi_q^2 + a_0 \pm |\mathbf{a}|}. \] (E48)

The corresponding eigenvectors lead to the following expression for the anomalous correlator

\[ \langle f_q f_{-q}^T \rangle = \left[ g_q^+ + g_q^a \cdot \mathbf{\sigma} \right] \Delta_q \frac{1}{2}, \] (E49)

with

\[ g_q^\pm = \frac{g(E_{q,+}) \pm g(E_{q,-})}{2}, \quad g(E) = \frac{\tanh(\beta E/2)}{E}. \] (E50)

Noting that \( E_q = |\xi_q| \) is \( C_3 \) invariant, while the off-diagonal terms of the equation are not, we can rewrite

When \( \Delta_q \Delta_q^\dagger \) is simply proportional to the identity, for instance when pairing occurs for spin-singlet, \( E_{q,+} = E_{q,-} = E_q \) and the previous expression simply becomes \( \langle f_q f_{-q}^T \rangle = g(E_q) \Delta_q / 2 \).

### C. Critical temperature

Solving the self-consistent relations of Eq. E46 with the help of Eq. E49 allows to determine the nature of the superconducting state. We now consider each pairing channel separately to check a superconducting phase can fully form.

1. **s-wave**

The possibility of an s-wave SC order can be ruled out because \( U \) is much larger than all the other terms scaling as \( t_0^2 / \Delta \). The coupled gap equations for \( \Delta_s \) and \( \Delta_s' \) linearized near \( T_c \) read

\[ \begin{bmatrix} \Delta_s' \\ \Delta_s \end{bmatrix} = \begin{bmatrix} \frac{2\lambda + V}{3} & \alpha I_0 + 2\tilde{t}I_1 \\ \alpha I_1 + 2\tilde{t}I_0 + \frac{2(\lambda + V)}{3} I_2 \end{bmatrix} \begin{bmatrix} \Delta_s' \\ \Delta_s \end{bmatrix}, \] (E51)

where \( I_k = -\sum_j \tanh(\beta E_j/2)(c_1 + c_2 + c_3)^k/(2N_s E_q) \) and \( \beta \) the inverse temperature. The matrix in the previous equation must have at least one eigenvalue equal to one for the system to exhibit s-wave symmetry. However, to leading order in \( U \), this requires to have

\[ U \left[ \frac{2(\lambda + V)}{3} (I_0 I_2 - I_1^2) - I_0 \right] = 0. \] (E52)

This equation has no solution because \( -I_0 > 0 \) has the same sign as \( I_0 I_2 - I_1^2 > 0 \) (due to Cauchy-Schwarz inequality). Thus, singlet pairing does not occur happen. It could nevertheless appear for smaller ratios \( \Delta / t_0 \) where our perturbation theory breaks down.

2. **p-wave**

The possibility of p-wave pairing can be ruled out as well. Indeed, let’s assume a p-wave SC order and compute the critical temperature \( T_c \) of that state. Linearizing the gap equation, such that \( a \simeq 0 \) and \( E_{q,\pm} \simeq |\xi_q| \), we find the coupled equations

\[ \begin{bmatrix} \Delta_{p_x} \\ \Delta_{p_y} \end{bmatrix} = \frac{V + \lambda}{N_s} \sum_q \frac{\tanh(\beta E_q/2)}{E_q} \left[ \frac{(s_1 + s_2 - 2s_3)^2}{6} \right] \begin{bmatrix} \Delta_{p_x} \\ \Delta_{p_y} \end{bmatrix}. \] (E53)

the diagonal terms as:

\[ -6 \frac{V + \lambda}{N_s} \sum_q \frac{\tanh(\beta E_q/2)}{E_q} \sum_j (s_j - s_{j+1})^2, \] (E54)
which does not have any solution since the left and right hand sides have opposite signs.

3. \textit{d-} and \textit{f-wave}

We can similarly derive an implicit equation for the critical temperature in the d-wave channel

\[
\frac{6}{\lambda - V} = \frac{1}{N_s} \sum_q \frac{\tanh(\beta E_q/2)}{E_q} \sum_j (c_j - c_{j+1})^2. \tag{E55}
\]

Contrary to p-wave pairing, \(\lambda - V\) can be positive for \(U_B\) not too large and \(V_0/\Delta\) large enough. We highlight the region where \(\lambda > V\) in Fig. S1.

The same situation arise for f-wave pairing, whose critical temperature follows from

\[
\frac{3}{2\lambda - V} = \frac{1}{N_s} \sum_q (s_1 + s_2 + s_3)^2 \frac{\tanh(\beta E_q/2)}{E_q}. \tag{E56}
\]

As explained in the main text, \(2\lambda > V\) for most choice of parameter, leading to a superconducting order with f-wave symmetry in most cases.

When f- and d- wave superconductivity may coexist, we generically expect the f-wave critical temperature to dominate due to a larger interaction strength \(2\lambda - V > \lambda - V\). This is confirmed by our numerical solution of Eqs. \(\text{E56}\) and \(\text{E55}\) where we observe that the f-wave critical temperature dominates. At large doping concentration \(x > 0.5\), the f-wave superconducting order parameter exhibits nodes along the \(\Gamma - M\) lines, while d-wave paired state can avoid the presence of nodes. The latter may therefore be energetically favored. To investigate this competition, we solve the zero temperature mean-field equations assuming \(\Delta_f, \Delta_d \neq 0\) for \(\Delta_{d, \sigma} \neq 0\). Our results presented in Fig. S1 show that the f-wave superconducting state still wins over d-wave order, despite its nodes, even when \(x > 0.5\). Numerical data right at the van Hove singularity \(x = 0.5\) show a more subtle competition between the two orders. At that point, other instabilities, e.g. charge density wave, may arise and our mean-field treatment should be supplemented with more precise analytical tools. We leave these details for a future work, simply noting for now that f-wave superconductivity largely dominate in the entire phase diagram of our model, except maybe at the van Hove singularity.

D. \textit{Zero temperature superconducting gap}

As discussed in the main text, the superconducting gap for f-wave and spin-triplet symmetry may be written as \(\Delta_k = (d_k \cdot \sigma) (\sigma_y)\) with \(d_k = [\sum_j \sin(k \cdot a_j)]d_f\) and \(d_f\) a constant vector that only depends on temperature. Owing to the spin-rotation symmetry of our model, we can choose \(d_f\) to lie in the \(xy\) plane at on point of the Brillouin zone. Moreover, we can eliminate any phase difference between \(\Delta_k^f\) and \(\Delta_k^d\), which yields \(d_f = e^{i\theta} (d_x, id_y, 0)^T\) with \(d_{x,y}\) real. With this order parameter, eigen-energies of the mean-field Hamiltonian take the rather simple form

\[
E_{k, \pm} = \sqrt{\xi_k^f + s_k^2} |d_x \pm d_y|^2. \tag{E57}
\]

The self-consistent equation for such symmetry (see Eq. \(\text{E49}\)) gives

\[
\begin{bmatrix}
{d_x} \\
{d_y}
\end{bmatrix} = \frac{2\lambda - V}{3} \begin{bmatrix} a^+ & a^- \\
{a^-} & {a^+}
\end{bmatrix} \begin{bmatrix}
{d_x} \\
{d_y}
\end{bmatrix}, \tag{E58}
\]

with

\[
a^\pm = \frac{1}{N_s} \sum_q s_q^2 g_q^\pm, \quad s_q = \sum_j \sin(q \cdot a_j). \tag{E59}
\]

To have a non-zero solution, we must have

\[
|a^+ - \frac{3}{2\lambda - V}| = |a^-|. \tag{E60}
\]

The solution of this equation with the large total gap \(|d|\) necessarily has \(a^- = 0\). Because \(s_q^2 g_q^\pm\) has the same sign in the entire Brillouin zone, this implies that \(E_{q, +} = E_{q, -}\) for all momenta. Hence, \(d_x = 0\) or \(d_y = 0\) in the entire Brillouin zone. In both cases, the net spin carried by the Cooper pairs is zero \[10\] :

\[
\langle S \rangle = \langle i d \times d^\ast \rangle = 0. \tag{E61}
\]

As a consistency check, we have numerically solved the self-consistent Eq. \(\text{E58}\) and reached the same conclusion. For all parameters considered, our results show \(d_y = 0\), as exemplified in Fig. S2 for \(U_B/V_0 = 2\) and \(U_B/V_0 = 4\).

S5. \textbf{WEAKLY INTERACTING REGIME}

A. \textbf{Explicit unitary transformation}

We start by isolating the band mixing (or off-diagonal) interaction elements from the others \(V = V_{od} + V_{dd}\). More explicitly, we write \(V_{od} = \frac{1}{N}\sum_{C(1234)} V_{234} \delta_{43} \delta_{a_4}^c \delta_{a_2}^c \delta_{a_1}^c \delta_{a_1}^c\) with \(C(1234)\) restricting the sum to terms satisfying either \(b_1 b_2 b_3 b_4 = -1\), or having \(b_1 = b_2\) and \(b_3 = b_4\) with \(b_1 \neq b_2\).

To eliminate these band mixing interaction terms, we use a Schrieffer-Wolff transformation \(H' = e^S H e^{-S}\), with \(S\) anti-hermitian. This unitary transformation can be carried order by order in the small parameter \(|V|/\Delta\), and we write \(S = S_1 + S_2 + \cdots\) with \(S_n = O(|V|^n/\Delta^n)\). Requiring

\[
[H_0, S_1] = V_{od}, \tag{E62}
\]
The leading corrections to the band dispersion come from $V \delta \epsilon_c$, and are given by Eq. in the main text. The Hartree-like part, corresponding to the first line of Eq. takes the explicit form

$$
\delta \epsilon_{k,\sigma}^H = (6V_0C_- + U_A C_+) \frac{\epsilon_{k,\sigma} - \Delta_0/2}{2\epsilon_{k,\sigma}} + (6V_0C_+ + U_B C_-) \frac{\epsilon_{k,\sigma} + \Delta_0/2}{2\epsilon_{k,\sigma}}.
$$

with $C_{\pm} = (2N_s)^{-1} \sum_{\mathbf{q}} (\epsilon_{\mathbf{q},+,\pm} \Delta_0/2) / \epsilon_{\mathbf{q},\pm}$. The Fock-like term, second line of Eq. reads

$$
\delta \epsilon_{k,\sigma}^F = \frac{t_0^2 V_0}{2N_s} \text{Re} \left[ \frac{f^\ast (\mathbf{k})}{\epsilon_{k,\sigma}^F} \sum_{\mathbf{q}} f (\mathbf{k} - \mathbf{q}) \frac{f (\mathbf{q})}{\epsilon_{\mathbf{q},\sigma}} \right].
$$

These corrections are plotted together with the bare band dispersion for $\Delta_0 = t = 2U_A = 2U_B = 10V_0$ in Fig. S3.
Where we observe that they both admit degenerate minima at the K and K' points.

When \( t_0 \ll \Delta_0 \), we can approximate \( \varepsilon_{k,+} \approx \Delta_0/2 + |t_0 f(k)|^2/\Delta_0 \) and expand these corrections in powers of \( t_0/V_0 \). Without too much difficulty, we end up with

\[
\delta \varepsilon_{k,+}^H = \frac{|t_0 f(k)|^2}{\Delta_0} (U_A - 6V_0), \tag{E70}
\]

\[
\delta \varepsilon_{k,+}^F = \frac{2t_0^2 V_0}{\Delta_0^2 N_s} \text{Re} \left[ f^*(k) \sum_q f(k-q) f(q) \right], \tag{E71}
\]

up to an overall global constant. These expressions allow to obtain the effective mass by expanding around the K or K' point. For that purpose, we recall \( f(K+k) \approx \sqrt{3}(k_x + ik_y)a/2 \) and furthermore find that

\[
\sum_{\mathbf{q}} f(K+k-\mathbf{q}) f(\mathbf{q}) = \sum_{\mathbf{q}} f(\mathbf{k}-\mathbf{q}) f(\mathbf{q} + \mathbf{k})
\]

\[
\approx \sum_{\mathbf{q}} f(-\mathbf{q}) f(\mathbf{q} + \mathbf{k}) + \mathbf{k} \cdot \sum_{\mathbf{q}} (\nabla f)(-\mathbf{q}) f(\mathbf{q} + \mathbf{k})
\]

\[
= 0 + \sqrt{3}(k_x + ik_y)a/2,
\]

where the last equality is easy to check numerically. We end up with

\[
(\varepsilon + \delta \varepsilon^H + \delta \varepsilon^F)_{K,k,+} \approx \frac{3t_0^2 a^2 |\mathbf{k}|^2}{4\Delta_0^2} (\Delta_0 + U_A - 4V_0), \tag{E72}
\]

yielding the effective mass given in Eq. 16.

We now turn to the corrections to the two-body scattering vertex, which are contained in the second order term \( X = [S_1, V_{od}]/2 \). Direct evaluation of the commutator using the explicit expression of \( S_1 \) (Eq. 64) gives

\[
X = \frac{1}{N_s} \sum_{123456} \Gamma_{654}^{21} \delta \varepsilon_{k,+}^{40} (2c_1^\dagger c_3 - \delta_{3,4}) c_2 c_1, \tag{E73}
\]

where the indices 653 and 421 originates from the same interaction elements and thus satisfy the conditions given above for the elements of \( V_{od} \) and \( S_1 \). The three-body tensor \( \Gamma \) takes the form

\[
\Gamma_{654}^{21} = \frac{1}{N_s} \sum_0 \delta_{65}^{30} \delta_{46}^{31} (V_{65}^{30} - V_{65}^{31} - V_{65}^{21} + V_{65}^{21}) \tag{E74}
\]

\[
\times \left[ \frac{1}{\varepsilon_0 + \varepsilon_3 - \varepsilon_0} + \frac{1}{\varepsilon_2 + \varepsilon_1 - \varepsilon_0} \right].
\]

For later use, we also define \( \hat{\Gamma} \) and \( \tilde{\Gamma} \), which have the same explicit representation except that the sum over 0 is restricted to states in lower band \( b_0 = - \) and the upper band \( b_0 = + \), respectively. To find the two-body corrections from \( X \), we can select the terms where two of the indices belong to the lower band and contract them. Considering all possible pairs compatible with the constraints on 653 and 421, we end up with

\[
\delta V_{43}^{21} = 2 \sum_{i,b_i=-} \hat{\Gamma}_{i}^{43} + \hat{\Gamma}_{i}^{21} + \hat{\Gamma}_{i}^{43} + \hat{\Gamma}_{i}^{21} + \hat{\Gamma}_{i}^{43} \tag{E75}
\]

\[
+ 2 \sum_{i,b_i=-} \hat{\Gamma}_{i}^{21} + \hat{\Gamma}_{i}^{21} + \hat{\Gamma}_{i}^{21} + \hat{\Gamma}_{i}^{21} + \hat{\Gamma}_{i}^{21} + \hat{\Gamma}_{i}^{21}.
\]

This expression greatly simplifies when we assume that the four momenta \( k_{1,2,3,4} \) are equal to \( K \) or \( K' \), as we do to determine the effective interaction strength in the spin-triplet valley singlet channel \( U_0 \) (see main text). Because \( \Psi_{K,K',+} = 0 \), any term of the form \( \hat{\Gamma}_{i}^{43} \), with \( 6 = (K/K',+) \) or \( 1 = (K/K',+) \) appearing in \( U_0 \) vanishes (see [E73]). This completely removes any contribution from on-site interactions on \( A \) sites \( U_A \). Let us first focus on the \( V_0^2 \) contributions and set \( U_B = 0 \) for a moment. This largely simplifies the expression of \( U_0 \), which now reads

\[
U_0^{(1)} = 2 \sum_q \hat{\Gamma}_{q}^{(K')}^{(K)+}(q) \hat{\Gamma}_{q}^{(K)+}(q) + \hat{\Gamma}_{q}^{(K)+}(K')^{(K)+}(q) \tag{E76}
\]

\[
- \hat{\Gamma}_{q}^{(K)+}(K')^{(K)+}(q) - \hat{\Gamma}_{q}^{(K)+}(K')^{(K)+}(q)
\]

The two first terms of the sum involve interaction coefficients with momentum transfer \( K \) and \( K' \), respectively, which makes them vanish as \( f(K) = f(K') = 0 \) (see Eq. [E67]). The other terms contribute equally, and we finally obtain

\[
U_0^{(1)} = -\frac{36t_0^2 V_0^2}{N_s} \sum_q \frac{|f(q)|^2}{(2\varepsilon_{q,+})^3}. \tag{E77}
\]

It is not difficult to check that the contribution proportional to \( U_B^2 \) vanishes, and we now turn to the crossed \( V_0 U_B \) corrections. The calculation proceeds in a similar way, except that one of the interaction element \( 3V_0 \) is replaced by \( U_B \) and the signs needs to be flipped, such that the exchange part gives

\[
U_0^{(2)} = \frac{12t_0^2 V_0 U_B}{N_s} \sum_q \frac{|f(q)|^2}{(2\varepsilon_{q,+})^3}. \tag{E78}
\]

For \( t_0 \ll \Delta_0 \), we use \( \sum_q |f(q)|^2 = 3N_s \) to find the simpler form

\[
U_0 = \frac{36t_0^2 V_0 (U_B - 3V_0)}{\Delta_0}. \tag{E79}
\]
which – remarkably – exactly match the result of the kinetic expansion in the weakly interacting regime

\[ U_{0}^{\text{KE}} = V_c - V_x \overset{ (U_A, U_B, V_0 \ll \Delta_0) }{\approx} \frac{36t_0^2 V_0 (U_B - 3V_0)}{\Delta_0^3}. \] (E80)