PNAS
www.pnas.org

Supplementary Information for

Spin-triplet superconductivity from excitonic effect in doped insulators

Valentin Crépel and Liang Fu

E-mail: vcrepel@mit.edu, liangfu@mit.edu

This PDF file includes:

Supplementary text
Figs. S1 to S2 (not allowed for Brief Reports)
SI References

Supporting Information Text

1. Effective model in the atomic limit

A. General framework. In this appendix, we derive the effective Hamiltonian used in the main text by treating the tunneling terms

$$\mathcal{H}_t = t_0 \sum_{\langle r,r'\rangle,\sigma} \left(c^\dagger_{r,\sigma} c_{r',\sigma} + hc \right), \quad [1]$$

as a perturbation to the 'classical' part $\mathcal{H}_0 = \mathcal{H} - \mathcal{H}_t$ (1), which is justified by the large gap $\Delta \gg t_0$. To that purpose, we apply the unitary transformation $\mathcal{H}' = e^{iS}\mathcal{H}e^{-iS}$, with S Hermitian and satisfying (2)

$$[\mathcal{H}_0, iS] = \mathcal{H}_t. \quad [2]$$

It leads to the following approximation of the Hamiltonian

$$\mathcal{H}' = \mathcal{H}_0 + \frac{1}{2}[iS, \mathcal{H}_t] + \mathcal{O}(\mathcal{H}_t S^2), \quad [3]$$

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

$$\mathcal{H}_t = \sum_{d=\pm 1} \sum_{v=-5}^{5} \sum_{u=-1}^{1} T_{d,v,u}, \quad [4]$$

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

$$S = -i \sum_{d,v,u} \frac{T_{d,v,u}}{d(\Delta_0 - U_A) + vV_0 + uU_B}. \quad [5]$$

Plugging this expression in Eq. 3, we obtain

$$\mathcal{H}' = \mathcal{H}_0 + \frac{1}{2} \sum_{\substack{d,v,u \\ d',v',u'}} \frac{[T_{d',v',u'}, T_{d,v,u}]}{d'(\Delta_0 - U_A) + v'V_0 + u'U_B}, \quad [6]$$

which is valid up to $\mathcal{O}\left(t_0^3/\Delta^2\right)$ 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 $E_f = \Delta + 3V_0$. 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 Δ. 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. 6.

Projecting $\mathcal{H}'$ onto the f-band requires to have $d' = -d$ and $v' = -v$ in Eq. 6. 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

$$\mathcal{H}' \simeq \mathcal{H}_0 - \frac{1}{2} \sum_{v,u,u' \geq 0} T_{-1,-u',-v} T_{1,u,v} \left[\frac{1}{\Delta_0 - U_A + vV_0 + uU_B} + \frac{1}{\Delta_0 - U_A + vV_0 + u'U_B} \right]. \quad [7]$$

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:

$$\mathcal{H}' = U_B \sum_{i \in B} \frac{n_i(n_i - 1)}{2} + \sum_{ijk \in \triangle} \left[\tilde{V}_{ij,k} + \circlearrowleft_{ijk} \right] + \sum_{ijk \in \triangle, \sigma} \left[f^\dagger_{j,\sigma} \tilde{T}_{ij,k} f_{i,\sigma} + P_{ijk} \right], \quad [8]$$

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

$$\tilde{V}_{ij,k} = -\frac{t_0^2(2 - n_k)}{\Delta + (2 - n_\triangle)V_0 + n_k U_B}, \quad \tilde{T}_{ij,k} = \frac{t_{\Delta,i} + t_{\Delta,j}}{2}, \quad t_{\Delta,\ell} = \frac{t_0^2}{\Delta + (1 - n_\triangle)V_0 + n_\ell U_B}, \quad [9]$$

with $n_\triangle = n_i + n_j + n_k$, 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 $\tilde{T}_{ij,k}$ 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

$$\tilde{T}_{ij,k} = T_{0,0} + n_k[T_{0,1}-T_{0,0}] + (n_i+n_j)[T_{1,0}-T_{0,0}]. \quad [10]$$

Summing over all possible triangle, we can rewrite

$$\sum_{ijk\in\triangle,\sigma} \left[f^\dagger_{j,\sigma}\tilde{T}_{ij,k}f_{i,\sigma} + P_{ijk}\right] = \sum_{\langle i,j\rangle,\sigma}\left[f^\dagger_{j,\sigma}\left[t+\tilde{t}\frac{n_i+n_j}{2}\right]f_{i,\sigma}+hc\right] + \lambda\sum_{ijk\in\triangle,\sigma}\left[f^\dagger_{j,\sigma}n_k f_{i,\sigma}+P_{ijk}\right], \quad [11]$$

where we have introduced the coefficients

$$t=\frac{t_0^2}{\Delta+V_0},\quad \tilde{t}=\frac{t_0^2}{\Delta}+\frac{t_0^2}{\Delta+U_B}-\frac{2t_0^2}{\Delta+V_0},\quad \lambda=\frac{t_0^2}{\Delta}-\frac{t_0^2}{\Delta+V_0}. \quad [12]$$

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

$$\tilde{V}_{ij,k} = \tilde{V}_{ij,0} + n_k[\tilde{V}_{ij,1}-\tilde{V}_{ij,0}] + \frac{n_k(n_k-1)}{2}[\tilde{V}_{0,0}-2\tilde{V}_{0,1}] \quad [13]$$

$$= \tilde{V}_{0,0} + (n_i+n_j)[\tilde{V}_{1,0}-\tilde{V}_{0,0}] + \frac{n_i(n_i-1)+n_j(n_j-1)+2n_in_j}{2}[\tilde{V}_{2,0}-2\tilde{V}_{1,0}+\tilde{V}_{0,0}] \quad [14]$$

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

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

$$U_B\sum_{i\in B}\frac{n_i(n_i-1)}{2} + \sum_{ijk\in\triangle}\left[\tilde{V}_{ij,k}+\circlearrowleft_{ijk}\right] = U\sum_i\frac{n_i(n_i-1)}{2} + V\sum_{\langle i,j\rangle}n_in_j, \quad [15]$$

where the coefficients read

$$U = U_B + 3[2\tilde{V}_{2,0}-4\tilde{V}_{1,0}-2\tilde{V}_{0,1}+3\tilde{V}_{0,0}],\quad V_f = \tilde{V}_{2,0}+2\tilde{V}_{1,1}-4\tilde{V}_{1,0}-2\tilde{V}_{0,1}+3\tilde{V}_{0,0}. \quad [16]$$

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

$$V = \frac{4t_0^2V_0(\Delta-V_0)}{\Delta(\Delta+V_0)(\Delta+2V_0)} - \frac{2t_0^2V_0}{(\Delta+U_B)(\Delta+V_0+U_B)},\quad U = U_B - \frac{6t_0^2(4V_0^2+V_0\Delta+\Delta^2)}{\Delta(\Delta+V_0)(\Delta+2V_0)} + \frac{6t_0^2}{\Delta+V_0+U_B}. \quad [17]$$

Gathering the various terms, we obtain the expression given in 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

$$\mathcal{H}' = \sum_{\langle i,j\rangle,\sigma} t\left(f^\dagger_{i,\sigma}f_{j,\sigma}+hc\right) + Vn_in_j + \sum_{(ijk)\in\triangle}\lambda\left(\sum_\sigma f^\dagger_{i,\sigma}n_kf_{j,\sigma}+P_{ijk}\right) + U_3n_in_jn_k, \quad [18]$$

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

$$U_3 = \frac{12t_0^2V_0^2}{\Delta(\Delta+V_0)(\Delta+2V_0)} - \frac{6t_0^2V_0^2}{(\Delta+U_B-V_0)(\Delta+U_B)(\Delta+V_0+U_B)}. \quad [19]$$

2. Two-particle lattice calculation

In this appendix, we solve the effective model obtained above for two particles – the analog of Cooper's problem on the lattice. To do so, we separate the center of mass momentum K from the relative motion with the introduction of the states

$$|\varphi_0(K,r)\rangle = \frac{1}{\sqrt{2(1+\delta_{r,0})N_s}}\sum_R e^{i(K\cdot R)}(f^\dagger_{R,\uparrow}f^\dagger_{R+r,\downarrow}+f^\dagger_{R+r,\uparrow}f^\dagger_{R,\downarrow}),\quad |\varphi_1(K,r)\rangle = \frac{1}{\sqrt{N_s}}\sum_R e^{i(K\cdot R)}f^\dagger_{R,\uparrow}f^\dagger_{R+r,\uparrow}, \quad [20]$$

where the subscripts denotes the total spin S of the state (singlet $S=0$ or triplet $S=1$). The only difference between these spin configurations is their statistic under the exchange of the two particles, which translates into the sign difference $|\varphi_S(K,-r)\rangle = (-1)^S e^{i(K\cdot r)}|\varphi_S(K,r)\rangle$. The action of the Hamiltonian Eq. 18 on this basis is

$$\mathcal{H}'|\varphi_S(K,r)\rangle = \delta_{S,0}\delta_{r,0}U|\varphi_S(K,0)\rangle + t\sum_{\substack{j=1,2,3\\ \epsilon=\pm}}\delta_{r\neq 0}\,\delta_{r+\epsilon a_j\neq 0}\left[1+e^{i\epsilon(K\cdot a_j)}\right]|\varphi_S(K,r+\epsilon a_j)\rangle \quad [21]$$

$$+\frac{\tilde{t}}{2}\sum_{\substack{j=1,2,3\\ \epsilon=\pm}}\delta_{r,\epsilon a_j}\left[1+e^{-i\epsilon(K\cdot a_j)}\right]\sqrt{2}|\varphi_S(K,0)\rangle + \delta_{r,0}\left[1+e^{i\epsilon(K\cdot a_j)}\right]\frac{1}{\sqrt{2}}|\varphi_S(K,\epsilon a_j)\rangle$$

$$+\sum_{\substack{j=1,2,3\\ \epsilon=\pm}}\delta_{r,\epsilon a_j}\left[V|\varphi_S(K,\epsilon a_j)\rangle + \lambda|\varphi_S(K,-\epsilon a_{j-\epsilon})\rangle\right] + \lambda\sum_{\substack{j=1,2,3\\ \epsilon=\pm}}\delta_{r,\epsilon a_j}e^{i\epsilon(K\cdot a_{j-\epsilon})}|\varphi_S(K,-\epsilon a_{j+\epsilon})\rangle.$$

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 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 the non-local repulsion becomes comparable to the exciton-induced short-range pairing interaction (which is much larger than the binding energy ε_b). If we take into account the direct Coulomb repulsion between nearest-neighbor B sites V', bound sate 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 $\epsilon a > 86.4$Å (a is the lattice constant), which corresponds to $e^2/\epsilon a = 16.5$meV. Thus, in order for electron pairing to occur in the limit of vanishing doping, it is helpful to have a large ϵ which can result from dielectric screening by a different band.

3. Continuum Limit

As shown in the main text, the kinetic part of the effective Hamiltonian 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 the effective Hamiltonian $\mathcal{H}'$

$$\mathcal{H}_f = \sum_{k,\sigma} \varepsilon_k f^\dagger_{k,\sigma} f_{k,\sigma} + \frac{1}{2N_s} \sum_{\substack{k,q,p \\ \sigma,\sigma'}} V_{k,q} f^\dagger_{k,\sigma} f^\dagger_{q+p,\sigma'} f_{k+p,\sigma'} f_{q,\sigma} \tag{22}$$

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

$$V_{k,q} = U + 2V \sum_j \cos[(k-q)\cdot a_j] + 2\lambda \sum_j (e^{ika_j + iqa_{j-1}} + e^{-ika_{j-1} - iqa_j}) + 2\tilde{t} \sum_j [\cos(k\cdot a_j) + \cos(q \cdot a_j)]. \tag{23}$$

Due to the quadratic band dispersion near the K and K' points, low energy fermions acquire an effective mass $m = 2/(3ta^2)$. They also carry an additional index $\{\uparrow K, \downarrow K, \uparrow K', \downarrow K'\}$ 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

$$V_c = V_{K,K} = V_{K',K'} = 6(V - \lambda - \tilde{t}) + U. \tag{24}$$

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

$$V_x = V_{K,K'} = V_{K',K} = 3(4\lambda - 2\tilde{t} - V) + U. \tag{25}$$

Introducing different fields for the two valleys

$$f_{k,\sigma} = \begin{cases} \psi_{k,\sigma,K} & \text{if } k \text{ near } K \\ \psi_{k,\sigma,K'} & \text{if } k \text{ near } K' \end{cases}, \tag{26}$$

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

$$\begin{aligned} \mathcal{H}_{\text{int}} = &\frac{V_c}{N_s} \sum_{\substack{k,q,p \\ V=K,K'}} \psi^\dagger_{k,\uparrow,V} \psi^\dagger_{p-k,\downarrow,V} \psi_{p-q,\downarrow,V} \psi_{q,\uparrow,V} + \frac{V_c - V_x}{N_s} \sum_{\substack{k,q,p \\ \sigma=\uparrow,\downarrow}} \psi^\dagger_{k,\sigma,K} \psi^\dagger_{p-k,\sigma,K'} \psi_{p-q,\sigma,K'} \psi_{q,\sigma,K} \\ &+ \frac{V_c}{N_s} \sum_{\substack{k,q,p \\ V=K,K'}} \psi^\dagger_{k,\uparrow,V} \psi^\dagger_{p-k,\downarrow,\bar{V}} \psi_{p-q,\downarrow,\bar{V}} \psi_{q,\uparrow,V} + \frac{V_x}{N_s} \sum_{\substack{k,q,p \\ V=K,K'}} \psi^\dagger_{k,\uparrow,V} \psi^\dagger_{p-k,\downarrow,\bar{V}} \psi_{p-q,\downarrow,V} \psi_{q,\uparrow,\bar{V}}. \end{aligned} \tag{27}$$

Let us rearrange these terms in terms of pair operators to make their physical meaning clearer. Valley-polarized spin-singlet electron pairs $S_V = f_{V,\downarrow} f_{V,\uparrow}$ 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, $T_\sigma = f_{K',\sigma} f_{K,\sigma}$ 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 (see main text). To summarize, we can rewrite the different contact interaction terms as

$$\tilde{H} = \int \mathrm{d}x \sum_{\sigma,V} \psi^\dagger_{\sigma,V} \left[\frac{-\nabla^2}{2m}\right] \psi_{\sigma,V} + \int \frac{\mathrm{d}x}{\mathcal{A}} [(V_c - V_x)(T^\dagger_\downarrow T_\downarrow + T^\dagger_0 T_0 + T^\dagger_\uparrow T_\uparrow) + V_c(S^\dagger_{K'} S_{K'} + S^\dagger_K S_K) + (V_c + V_x) S^\dagger_0 S_0], \quad [28]$$

with $\mathcal{A} = \sqrt{3}/a^2$ 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 $V_c - V_x < 0$.

Alternatively, we can replace pair operators by more physical quantities, such as the total density on each valley $\rho_V = \psi^\dagger_{\uparrow,V}\psi_{\uparrow,V} + \psi^\dagger_{\downarrow,V}\psi_{\downarrow,V}$ and the total spin on each valley $\boldsymbol{s}_V = \psi^\dagger_{\alpha,V}\boldsymbol{\sigma}_{\alpha,\beta}\psi_{\beta,V}$. Together, they allow to represent the exchange term as

$$T^\dagger_\downarrow T_\downarrow + T^\dagger_\uparrow T_\uparrow + T^\dagger_0 T_0 - S^\dagger_0 S_0 = 2\boldsymbol{s}_K \cdot \boldsymbol{s}_{K'} + \frac{1}{2}\rho_K \rho_{K'}. \quad [29]$$

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

$$T^\dagger_\downarrow T_\downarrow + T^\dagger_\uparrow T_\uparrow + T^\dagger_0 T_0 + S^\dagger_0 S_0 + S^\dagger_K S_K + S^\dagger_{K'} S_{K'} = \frac{1}{2}\rho_{\mathrm{tot}}(\rho_{\mathrm{tot}} - 1). \quad [30]$$

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

$$\tilde{H}_i = \frac{1}{2\mathcal{A}} \int \mathrm{d}x \left[V_c\, n_{\mathrm{tot}}(n_{\mathrm{tot}} - 1) - V_x \left(4\boldsymbol{s}_K \cdot \boldsymbol{s}_{K'} + n_K n_{K'}\right)\right]. \quad [31]$$

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 ρ_K and $\rho_{K'}$, we find the three coupling constant given in the main text

$$g_0 = V_c/(2\mathcal{A}), \quad g_1 = (2V_c - V_x)/(2\mathcal{A}), \quad g_2 = -2V_x/\mathcal{A}. \quad [32]$$

4. Mean-field theory of superconductivity

In this appendix, we carry out a mean-field treatment of the effective Hamiltonian Eq. 22 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:

$$\mathcal{H}_{\mathrm{mf}} = \sum_{k,\sigma} \xi_q f^\dagger_{q,\sigma} f_{q,\sigma} + \frac{1}{2}\sum_{k,\sigma,\sigma'} \left[\tilde{\Delta}_{k,\sigma\sigma'} f^\dagger_{k,\sigma} f^\dagger_{-k,\sigma'} + hc\right], \quad \tilde{\Delta}_{k,\sigma\sigma'} = -\frac{1}{N_s}\sum_q V_{k,q} \langle f_{q,\sigma} f_{-q,\sigma'}\rangle, \quad [33]$$

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

$$\mathcal{H}_{\mathrm{mf}} = \sum_k{}' \begin{bmatrix} f^\dagger_k & f_{-k} \end{bmatrix} \begin{bmatrix} \xi_k & \Delta_k \\ \Delta^\dagger_k & -\xi_k \end{bmatrix} \begin{bmatrix} f_k \\ f^\dagger_{-k} \end{bmatrix}. \quad [34]$$

The order parameters have been gathered in a 2×2 matrix

$$\Delta_k = \frac{\tilde{\Delta}_k - \tilde{\Delta}^T_{-k}}{2} = \frac{-1}{N_s}\sum_q \mathrm{Re}\,(V_{k,q}) \langle f_{q,\sigma} f_{-q,\sigma'}\rangle, \quad [35]$$

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

$$\Delta_k = \Delta'_s + \sum_{j=1}^{3} \Delta^s_j \cos(k \cdot a_j) + \Delta^t_j \sin(k \cdot a_j), \quad [36]$$

which respectively read:

$$\begin{aligned}
\Delta'_s &= \frac{-1}{N_s}\sum_q \left[U + 2\tilde{t}(c_1 + c_2 + c_3)\right] \langle f_{q,\sigma} f_{-q,\sigma'}\rangle, \\
\Delta^s_j &= \frac{-2}{N_s}\sum_q \left[\lambda(c_{j-1} + c_{j+1}) + V c_j + \tilde{t}\right] \langle f_{q,\sigma} f_{-q,\sigma'}\rangle, \\
\Delta^t_j &= \frac{2}{N_s}\sum_q \left[\lambda(s_{j-1} + s_{j+1}) - V s_j\right] \langle f_{q,\sigma} f_{-q,\sigma'}\rangle,
\end{aligned} \quad [37]$$

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

$$\Delta'_s, \quad \Delta_{s/f} = \frac{1}{3}\left[\Delta_1^{s/t} + \Delta_2^{s/t} + \Delta_3^{s/t}\right], \tag{38}$$

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

$$\Delta_{d_{x^2-y^2}/p_x} = \frac{1}{6}\left[\Delta_1^{s/t} + \Delta_2^{s/t} - 2\Delta_3^{s/t}\right], \quad \Delta_{d_{xy}/p_y} = \frac{1}{2}\left[\Delta_1^{s/t} - \Delta_2^{s/t}\right]. \tag{39}$$

The inverse transformations are

$$\Delta_1 = \Delta_{s/f} + \Delta_{d_{x^2-y^2}/p_x} + \Delta_{d_{xy}/p_y}, \quad \Delta_2 = \Delta_{s/f} + \Delta_{d_{x^2-y^2}/p_x} - \Delta_{d_{xy}/p_y}, \quad \Delta_3 = \Delta_{s/f} - 2\Delta_{d_{x^2-y^2}/p_x}. \tag{40}$$

Finally, singlet pairs cannot be of equal spin, and we can therefore express them as a scalar times the 2×2 matrix $i\sigma_y$, *e.g.* $\Delta_s = d_0(i\sigma_y)$ with δ_s a complex number. Triplet on the other hand, take the form of a Pauli vector multiplied by $i\sigma_y$, *e.g.* $\Delta_f = (\boldsymbol{d}_f \cdot \boldsymbol{\sigma})(i\sigma_y)$.

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

$$\begin{aligned}
\Delta'_s &= \frac{-1}{N_s}\sum_q \left[U + 2\tilde{t}(c_1 + c_2 + c_3)\right]\langle f_{q,\sigma} f_{-q,\sigma'}\rangle, \\
\Delta_s &= \frac{-2}{N_s}\sum_q \left[\tilde{t} + \frac{2\lambda + V}{3}(c_1 + c_2 + c_3)\right]\langle f_{q,\sigma} f_{-q,\sigma'}\rangle, \\
\Delta_{p_x} &= \frac{-2(\lambda + V)}{N_s}\sum_q \frac{s_1 + s_2 - 2s_3}{6}\langle f_{q,\sigma} f_{-q,\sigma'}\rangle, \\
\Delta_{p_y} &= \frac{-2(\lambda + V)}{N_s}\sum_q \frac{s_1 - s_2}{2}\langle f_{q,\sigma} f_{-q,\sigma'}\rangle, \\
\Delta_{d_{x^2-y^2}} &= \frac{2(\lambda - V)}{N_s}\sum_q \frac{c_1 + c_2 - 2c_3}{6}\langle f_{q,\sigma} f_{-q,\sigma'}\rangle, \\
\Delta_{d_{xy}} &= \frac{2(\lambda - V)}{N_s}\sum_q \frac{c_1 - c_2}{2}\langle f_{q,\sigma} f_{-q,\sigma'}\rangle, \\
\Delta_f &= \frac{2(2\lambda - V)}{N_s}\sum_q \frac{s_1 + s_2 + s_3}{3}\langle f_{q,\sigma} f_{-q,\sigma'}\rangle.
\end{aligned} \tag{41}$$

B. Self-consistent conditions. The mean field quadratic Hamiltonian can be diagonalized by a Bogoliubov transformation. Writing the hermitian matrix

$$\Delta_q \Delta_q^\dagger = a_0 + \boldsymbol{a} \cdot \boldsymbol{\sigma} \tag{42}$$

as a Pauli vector, the eigen-energies read

$$E_{q,\pm} = \sqrt{\xi_q^2 + a_0 \pm |\boldsymbol{a}|}. \tag{43}$$

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

$$\langle f_q f_{-q}^T\rangle = \left[g_q^+ + g_q^- \frac{\boldsymbol{a} \cdot \boldsymbol{\sigma}}{|a|}\right]\frac{\Delta_q}{2}, \tag{44}$$

with

$$g_q^\pm = \frac{g(E_{q,+}) \pm g(E_{q,-})}{2}, \; g(E) = \frac{\tanh(\beta E/2)}{E}. \tag{45}$$

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. 41 with the help of Eq. 44 allows to determine the nature of the superconducting state. We now consider each pairing channel separately to check whether a superconducting phase can fully form.

C.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/Δ. The coupled gap equations for Δ_s and Δ_s' linearized near T_c read

$$\begin{bmatrix}\Delta_s'\\ \Delta_s\end{bmatrix} = \begin{bmatrix} UI_0 + 2\tilde{t}I_1 & UI_1 + 2\tilde{t}I_2 \\ 2\tilde{t}I_0 + \frac{2(2\lambda+V)}{3}I_1 & 2\tilde{t}I_1 + \frac{2(2\lambda+V)}{3}I_2\end{bmatrix}\begin{bmatrix}\Delta_s'\\ \Delta_s\end{bmatrix}, \quad [46]$$

where $I_k = -\sum_q \tanh(\beta E_q/2)(c_1+c_2+c_3)^k/(2N_s E_q)$ and β 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

$$U\left[\frac{2(2\lambda+V)}{3}(I_0I_2 - I_1^2) - I_0\right] = 0. \quad [47]$$

This equation has no solution because $(-I_0) > 0$ has the same sign as $I_0I_2 - I_1^2 > 0$ (we used Cauchy-Schwarz inequality for the last inequality). Thus, s-wave singlet pairing does not happen in our model. It could nevertheless appear for smaller ratios Δ/t_0 where our perturbation theory breaks down.

C.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 $\boldsymbol{a} \simeq \mathbf{0}$ and $E_{q,\pm} \simeq |\xi_q|$, we find the coupled equations

$$\begin{bmatrix}\Delta_{p_x}\\ \Delta_{p_y}\end{bmatrix} = -\frac{V+\lambda}{N_s}\sum_q \frac{\tanh(\beta E_q/2)}{E_q}\begin{bmatrix}(s_1+s_2-2s_3)^2/6 & (s_1+s_2-2s_3)(s_1-s_2)/6\\ (s_1-s_2)(s_1+s_2-2s_3)/2 & (s_1-s_2)^2/2\end{bmatrix}\begin{bmatrix}\Delta_{p_x}\\ \Delta_{p_y}\end{bmatrix}. \quad [48]$$

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

$$\frac{-6}{V+\lambda} = \frac{1}{N_s}\sum_q \frac{\tanh(\beta E_q/2)}{E_q}\sum_j (s_j - s_{j+1})^2, \quad [49]$$

which does not have any solution since the left and right hand sides have opposite signs since $V+\lambda > 0$.

C.3. d- and f-wave. We can similarly derive an implicit equation for the critical temperature in the d-wave channel

$$\frac{6}{\lambda - V} = \frac{1}{N_s}\sum_q \frac{\tanh(\beta E_q/2)}{E_q}\sum_j (c_j - c_{j+1})^2. \quad [50]$$

Contrary to p-wave pairing, $\lambda - V$ can be positive for U_B not too large and V_0/Δ large enough. The region where $\lambda > V$ is explicitly shown in the main text. The two independent order parameters $\Delta_{d_{xy}}$ and $\Delta_{d_{x^2-y^2}}$ are shown in Fig. S1. They exhibits nodal lines crossing the corners of the Brillouin zone, leading to a nodal superconducting order parameter at low doping when time reversal symmetry is not explicitly broken.

Turning to f-wave pairing, we find the effective gap equation

$$\frac{3}{2\lambda - V} = \frac{1}{N_s}\sum_q (s_1+s_2+s_3)^2 \frac{\tanh(\beta E_q/2)}{E_q}. \quad [51]$$

As explained in the main text, $2\lambda > V$ for most choice of parameter, leading to a superconducting order with f-wave symmetry that we extensively study in the main text. The competition between f- and d-wave superconducting state in regions where both of them are allowed is also discussed in the main text. In our model, f-wave pairing strongly dominate, but additional terms in the Hamiltonian may work in favor of the d-wave paired state.

5. Weakly interacting regime

A. Explicit unitary transformation. We start by isolating the band mixing (or off-diagonal) interaction elements from the others $\mathcal{V} = \mathcal{V}_{\rm od} + \mathcal{V}_{\rm d}$. More explicitly, we write $\mathcal{V}_{\rm od} = \frac{1}{N_s}\sum_{C(1234)} V_{43}^{21}\delta_{43}^{21} c_4^\dagger c_3^\dagger c_2 c_1$, with $C(1234)$ restricting the sum to terms satisfying either $b_1b_2b_3b_4 = -1$, 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 $\mathcal{H}' = e^S \mathcal{H} e^{-S}$, with S anti-hermitian. This unitary transformation can be carried order by order in the small parameter $|\mathcal{V}|/\Delta$, and we write $S = S_1 + S_2 + \cdots$ with $S_n = \mathcal{O}(|\mathcal{V}|^n/\Delta^n)$. Requiring

$$[H_0, S_1] = \mathcal{V}_{\rm od}, \quad [52]$$

gets rid of the direct band-mixing terms in $\mathcal{H}'$:

$$\mathcal{H}' = \mathcal{H} + [S,\mathcal{H}] + \frac{1}{2}[S,[S,\mathcal{H}]] + \mathcal{O}(S^2\mathcal{H}) = \mathcal{H}_0 + \mathcal{V}_{\rm d} + \left[S_1, \mathcal{V}_{\rm d} + \frac{\mathcal{V}_{\rm od}}{2}\right] + [S_2, H_0] + \mathcal{O}\left(\frac{|\mathcal{V}|^3}{\Delta^2}\right). \quad [53]$$

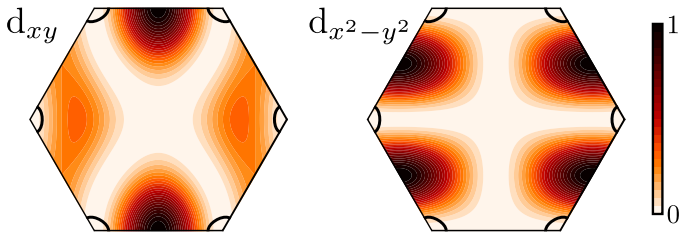

Fig. S1. Normalized d-wave superconducting order parameter amplitudes. The Fermi surface for $x = 0.1$ is indicated with a solid black line.

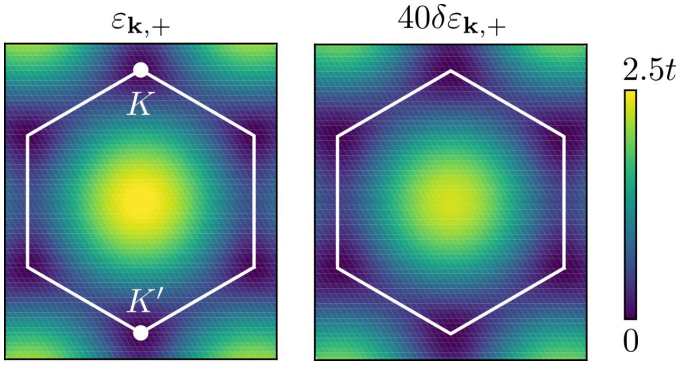

Fig. S2. Bare dispersion $\varepsilon_{\mathbf{k},+}$ and its leading correction $\delta\varepsilon_{\mathbf{k},+}$ (multiplied 40 times for visibility) in the weakly interacting limit, shown for $\Delta_0 = t = 2U_A = 2U_B = 10V_0$.

A possible S_1 satisfying this condition is

$$S_1 = \sum_{C(1234)} \frac{V_{43}^{21}\delta_{43}^{21}}{\varepsilon_4+\varepsilon_3-\varepsilon_2-\varepsilon_1} c_4^\dagger c_3^\dagger c_2 c_1. \quad [54]$$

We can then choose $[S_2, H_0]$ to remove the band mixing terms of $[S_1, \mathcal{V}_\mathrm{d} + \mathcal{V}_\mathrm{od}/2]$, and so on. Note that, because $\mathcal{V}_\mathrm{d}$ contains no band mixing terms, such that $[S_1, \mathcal{V}_\mathrm{d}]$ is purely off diagonal and therefore completely eliminated by $[S_2, H_0]$. The second order corrections read $[S_1, \mathcal{V}_\mathrm{od}]/2$, which as promised are of order $|\mathcal{V}|^2/\Delta$.

Finally, the effective Hamiltonian for doped charge is obtained by projecting $\mathcal{H}'$ to the subspace where the lower band is fully filled, which amounts to pairing up lower band indices as $c_a^\dagger c_b \to \tilde{\delta}_{a,b}$.

B. Leading corrections. To make analytical progress, we need the explicit form of the dispersion and scattering vertex. The single particle energy dispersion reads

$$\varepsilon_{\mathbf{k},\pm,\sigma} = \varepsilon_{\mathbf{k},\pm} = \pm\sqrt{(\Delta_0/2)^2 + |t_0 f(\mathbf{k})|^2}, \quad [55]$$

with $f(\mathbf{k}) = \sum_{j=1}^3 e^{i(\mathbf{k}\cdot\mathbf{u}_j)}$, and $\mathbf{u}_{j=1,2,3}$ the vectors connecting B sites to their three nearest neighbors. The corresponding Bloch eigenvectors are

$$\Psi_{\mathbf{k},\pm} = \frac{1}{\sqrt{2\varepsilon_{\mathbf{k},+}(\varepsilon_{\mathbf{k},+} \pm \Delta_0/2)}} \begin{bmatrix} \mp t_0 f(\mathbf{k}) \\ \varepsilon_{\mathbf{k},+} \pm \Delta_0/2 \end{bmatrix}, \quad [56]$$

They allow us to write the interaction vertex in the band basis as

$$V_{43}^{21} = V_0 f(\mathbf{k}_4 - \mathbf{k}_1) \left[\Psi^A_{\mathbf{k}_4,b_4}\Psi^B_{\mathbf{k}_3,b_3}\right]^* \Psi^B_{\mathbf{k}_2,b_2}\Psi^A_{\mathbf{k}_1,b_1} + \sum_{\tau=A/B} \frac{U_\tau}{2}\delta_{(\sigma_4=\sigma_1)\neq(\sigma_3=\sigma_2)} \left[\Psi^\tau_{\mathbf{k}_4,b_4}\Psi^\tau_{\mathbf{k}_3,b_3}\right]^* \Psi^\tau_{\mathbf{k}_2,b_2}\Psi^\tau_{\mathbf{k}_1,b_1}. \quad [57]$$

The leading corrections to the band dispersion come from $\mathcal{V}_\mathrm{d}$ and are in the main text. The Hartree-like part takes the explicit form

$$\delta\varepsilon^\mathrm{H}_{\mathbf{k},+} = (6V_0C_- + U_A C_+)\frac{\varepsilon_{\mathbf{k},+} - \Delta_0/2}{2\varepsilon_{\mathbf{k},+}} + (6V_0C_+ + U_B C_-)\frac{\varepsilon_{\mathbf{k},+} + \Delta_0/2}{2\varepsilon_{\mathbf{k},+}}, \quad [58]$$

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

$$\delta\varepsilon^\mathrm{F}_{\mathbf{k},+} = \frac{t_0^2 V_0}{2N_s}\mathrm{Re}\left[\frac{f^*(\mathbf{k})}{\varepsilon_{\mathbf{k},+}}\sum_\mathbf{q} f(\mathbf{k}-\mathbf{q})\frac{f(\mathbf{q})}{\varepsilon_{\mathbf{q},+}}\right]. \quad [59]$$

These corrections are plotted together with the bare band dispersion for $\Delta_0 = t = 2U_A = 2U_B = 10V_0$ in Fig. S2, where we observe that they both admit degenerate minima at the K and K' points.

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

$$\delta\varepsilon^\mathrm{H}_{\mathbf{k},+} = \frac{|t_0 f(\mathbf{k})|^2}{\Delta_0^2}(U_A - 6V_0), \quad \delta\varepsilon^\mathrm{F}_{\mathbf{k},+} = \frac{2t_0^2 V_0}{\Delta_0^2 N_s}\mathrm{Re}\left[f^*(\mathbf{k})\sum_\mathbf{q} f(\mathbf{k}-\mathbf{q})f(\mathbf{q})\right], \quad [60]$$

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 $f(K+\mathbf{k}) \simeq \sqrt{3}(k_x + ik_y)a/2$ and furthermore find that

$$\sum_\mathbf{q} f(K+\mathbf{k}-\mathbf{q})f(\mathbf{q}) = \sum_\mathbf{q} f(\mathbf{k}-\mathbf{q})f(\mathbf{q}+K) \simeq \sum_\mathbf{q} f(-\mathbf{q})f(\mathbf{q}+K) + \mathbf{k}\cdot\sum_\mathbf{q}(\nabla f)(-\mathbf{q})f(\mathbf{q}+K) = 0 + \sqrt{3}(k_x+ik_y)a/2, \quad [61]$$

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

$$(\varepsilon + \delta\varepsilon^\mathrm{H} + \delta\varepsilon^\mathrm{F})_{K+\mathbf{k},+} \simeq \frac{3t_0^2 a^2|\mathbf{k}|^2}{4\Delta_0^2}(\Delta_0 + U_A - 4V_0), \quad [62]$$

yielding the effective mass given in the main text.

We now turn to the corrections to the two-body scattering vertex, which are contained in the second order term $X = [S_1, \mathcal{V}_\mathrm{od}]/2$. Direct evaluation of the commutator using the explicit expression of S_1 (Eq. 54) gives

$$X = \frac{1}{N_s}\sum_{123456} \Gamma^{321}_{654}\delta^{321}_{654} c_6^\dagger c_5^\dagger (2c_4^\dagger c_3 - \delta_{3,4})c_2 c_1, \quad [63]$$

where the indices 653 and 421 originates from the same interaction elements and thus satisfy the conditions given above for the elements of $\mathcal{V}_{\rm od}$ and S_1. The three-body tensor Γ takes the form

$$\Gamma^{321}_{654} = \frac{1}{N_s}\sum_{0} \delta^{30}_{65}\delta^{21}_{40}(V^{03}_{65} - V^{30}_{65})(V^{21}_{04} - V^{21}_{40}) \left[\frac{1}{\varepsilon_6+\varepsilon_5-\varepsilon_3-\varepsilon_0} + \frac{1}{\varepsilon_2+\varepsilon_1-\varepsilon_4-\varepsilon_0}\right]. \quad [64]$$

For later use, we also define $\tilde{\Gamma}$ and $\hat{\Gamma}$, 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

$$\delta V^{21}_{43} = 2\sum_{i,b_i=-} \Gamma^{i21}_{43i} + \hat{\Gamma}^{2i1}_{i43} + \hat{\Gamma}^{21i}_{i43} + \hat{\Gamma}^{2i1}_{4i3} + \hat{\Gamma}^{21i}_{4i3} + 2\sum_{i,b_i=-} \tilde{\Gamma}^{i21}_{i43} + \tilde{\Gamma}^{i21}_{4i3} + \tilde{\Gamma}^{2i1}_{43i} + \tilde{\Gamma}^{21i}_{43i} - \sum_i \tilde{\Gamma}^{i21}_{43i}. \quad [65]$$

This expression greatly simplifies when we assume that the four momenta $\mathbf{k}_{1,2,3,4}$ 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 $\Psi^A_{K/K',+} = 0$, any term of the form Γ^{321}_{654} with $6 = (K/K',+)$ or $1 = (K/K',+)$ appearing in U_0 vanishes (see Eq. 64). 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

$$U_0^{(1)} = 2\sum_q \hat{\Gamma}^{(K'+)(K+)(q-)}_{(q-)(K+)(K'+)} + \hat{\Gamma}^{(K+)(K'+)(q-)}_{(q-)(K'+)(K+)} - \hat{\Gamma}^{(K+)(K'+)(q-)}_{(q-)(K+)(K'+)} - \hat{\Gamma}^{(K'+)(K+)(q-)}_{(q-)(K'+)(K+)}. \quad [66]$$

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. 57). The other terms contribute equally, and we finally obtain

$$U_0^{(1)} = -\frac{36t_0^2V_0^2}{N_s}\sum_{\mathbf{q}} \frac{|f(\mathbf{q})|^2}{(2\varepsilon_{\mathbf{q},+})^3}. \quad [67]$$

It is not difficult to check that the contribution proportional to U_B^2 vanishes, and we now turn to the crossed V_0U_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

$$U_0^{(2)} = \frac{12t_0^2V_0U_B}{N_s}\sum_{\mathbf{q}} \frac{|f(\mathbf{q})|^2}{(2\varepsilon_{\mathbf{q},+})^3}. \quad [68]$$

For $t_0 \ll \Delta_0$, we use $\sum_{\mathbf{q}} |f(\mathbf{q})|^2 = 3N_s$ to find the simpler form

$$U_0 = \frac{36t_0^2V_0(U_B - 3V_0)}{\Delta_0^3}, \quad [69]$$

which – remarkably – exactly match the result of the kinetic expansion in the weakly interacting regime

$$U_0^{\rm KE} = V_c - V_x \overset{(U_A,U_B,V_0\ll\Delta_0)}{\simeq} \frac{36t_0^2V_0(U_B - 3V_0)}{\Delta_0^3}. \quad [70]$$

References

1. K Slagle, L Fu, Charge transfer excitations, pair density waves, and superconductivity in moiré materials. *Phys. Rev. B* **102**, 235423 (2020).
2. JR Schrieffer, PA Wolff, Relation between the anderson and kondo hamiltonians. *Phys. Rev.* **149**, 491 (1966).