Charge Transfer Excitations, Pair Density Waves, and Superconductivity in Moiré Materials

Kevin Slagle[1,2] and Liang Fu[3]

[1] *Department of Physics and Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*

[2] *Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, California 91125, USA*

[3] *Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA*

(Dated: April 1, 2020)

Transition metal dichalcogenide (TMD) heterobilayers are a new class of tunable moiré systems attracting interest as quantum simulators of strongly-interacting electrons in two dimensions. In particular, recent theory predicts that the correlated insulator observed in WSe_2/WS_2 at half filling is a charge-transfer insulator similar to cuprates, and upon further hole doping, exhibits a transfer of charge from anion-like to cation-like orbitals at different locations in the moiré unit cell. In this work, we demonstrate that in this doped charge-transfer insulator, tightly-bound charge-$2e$ excitations can form to lower the total electrostatic repulsion. This composite excitation, which we dub a *trimer*, consists of a pair of holes bound to a charge-transfer exciton. When the bandwidth of doped holes is small, trimers crystallize into insulating pair density waves at a sequence of commensurate doping levels. When the bandwidth becomes comparable to the pair binding energy, itinerant holes and charge-$2e$ trimers interact resonantly, leading to unconventional superconductivity similar to superfluidity in an ultracold Fermi gas near Feshbach resonance.

Moiré superlattices [1–3] can be viewed as magnified crystals whose unit cell is of nanometer instead of angstrom size. Correspondingly, the relevant electronic phenomena in moiré superlattices is governed by the coarse-grained moiré potential and the extended Coulomb repulsion, with characteristic energy scale on the order of meV instead of eV. Thanks to the increase of length scale and the reduction of energy scale, moiré systems feature remarkable tunability through the control of twist angle and displacement field. A variety of moiré materials have emerged as exciting venues for studying and designing correlated electron phenomena with an unprecedented level of control. [4–15]

Recently, moiré superlattices of transition metal dichalcogenides (TMD) [16] have attracted great interest as quantum simulators of strongly-correlated electron systems in two dimensions. [17–22] By varying the twist angle, the relative strength of the bandwidth and electron interaction can be tuned, and a rich quantum phase diagram can potentially be realized. [23–25] Encouragingly, transport and optical experiments are starting to observe correlated insulating states in the TMD heterobilayer WSe_2/WS_2 with $n \leq 1$ holes per moiré unit cell. [26, 27] In particular, the insulating state at $n = 1$ is theoretically identified as a charge-transfer insulator with a cation and an anion at different locations in the moiré unit cell, corresponding to localized Wannier orbitals at the primary and secondary energy minimum of the moiré potential, respectively [28]. While a charge-transfer insulator is similar to a Mott insulator in terms of ground state properties, the key difference is that upon doping a charge-transfer insulator to $n > 1$, the additional charges fill a higher-energy orbital in order to avoid double occupancy [29]. A famous example of charge-transfer insulators is undoped cuprates [30, 31], for which a link between charge-transfer physics and high-temperature superconductivity upon doping has long been proposed and studied [32, 33].

In this work, we present a microscopic theory of charge pairing by Coulomb repulsion in TMD heterobilayers under a range of fillings $n > 1$. This counter-intuitive phenomenon occurs when the charge-transfer gap at $n = 1$ is small, so that two doped charges can lower their energy by polarizing their surroundings to form a tightly-bound charge-$2e$ "trimer" that consists of three holes on adjacent cations surrounding an electron on an anion. We show that the trimer costs less energy than two spatially separated holes for realistic forms of electron-electron interaction. When the single-particle bandwidth is small, we predict the formation of periodic density waves of trimers at certain doping levels $n = 1 + p/q > 1$ (p, q are integers), whose periodicity is commensurate with the moiré lattice. As the bandwidth of holes increases and becomes comparable with the binding energy of trimers, holes and trimers coexist and interact resonantly to form a strong-coupling superconductor, similar to a strongly-interacting superfluid in a Fermi gas near Feshbach resonance [34, 35]. Our theory of pair density waves and superconductivity in TMD heterobilayers is asymptotically exact in a certain regime of strong interaction and small doping.

We start by describing the single-particle electronic

structure of TMD heterobilayers (e.g., WSe_2/WS_2). Here, the topmost valence band of WSe_2 is reconstructed into a set of moiré bands by the periodic moiré potential resulting from the 4% lattice mismatch with WS_2. Importantly, the moiré potential has two inequivalent local minima located at the AA and AB stacking regions, giving rise to two sets of localized Wannier orbitals. The AA (AB) orbital has a lower (higher) on-site energy for holes and can be regarded as anion (cation) like.

Thus, the low-energy physics of a TMD heterobilayer can be faithfully mapped onto a two-dimensional diatomic crystal with one cation and one anion per unit cell. The effective Hamiltonian takes the form of an extended Hubbard model on the honeycomb lattice [28]:

$$H = H_K + H_0 \qquad H_K = \sum_{ij,s=\uparrow\downarrow} t_{ij}\, c_{is}^\dagger c_{js}$$
$$H_0 = \sum_{j\in B} \Delta\, n_j + \sum_i U\, n_{i\uparrow} n_{i\downarrow} + \frac{1}{2}\sum_{i\neq j} V_{ij}\, n_i n_j \tag{1}$$

where the A and B sublattices (colored black and red in the figures) correspond to the anion and cation, respectively. $c_i^\dagger$ creates a hole in the moiré valence band with charge $e>0$, and $n_i = n_{i\uparrow}+n_{i\downarrow}$. $\Delta>0$ is the energy difference of cation and anion orbitals. For WSe_2/WS_2, $\Delta = 14.9$meV is extracted from first-principles band structure calculations [28, 36].

Trimers—The hopping integrals t_{ij} decrease exponentially as the moiré period $L_{\rm M}$ increases. The on-site and extended two-body interactions $U, V_{ij} > 0$ are given by Coulomb integrals in Wannier basis [28, 36]. Since U and V_{ij} decrease as power law functions of $L_{\rm M}$, electron-electron interactions dominate over single-particle hopping at large $L_{\rm M}$. In this strong-coupling regime, the on-site repulsion U is the largest relevant energy scale. At the filling $n=1$, the system is in an insulating state where all anions are singly occupied and cations unoccupied; i.e., $n_i^0 = 1$ for $i\in A$ and $n_i^0 = 0$ for $i\in B$.

Upon doping to $n>1$, the additional $n-1$ charges have to occupy the cations in order to avoid the large energy cost of double occupancy. To study the many-body physics at finite doping, we shall first identify the relevant charged excitations at $n=1$. For this purpose, it is useful to rewrite H_0 in terms of $\delta n_i \equiv n_i - n_i^0$, the change of occupation relative to the ground state:

$$H_0 = \sum_{i\in A} V_a \delta n_i + \sum_{j\in B}(\Delta+V_c)\delta n_j + \frac{1}{2}\sum_{i\neq j} V_{ij}\,\delta n_i \delta n_j. \tag{2}$$

We have taken the $U=\infty$ limit so that double occupancy is forbidden. Then, there exist two types of elementary excitations: (1) electrons on the A sublattice of anions with $\delta n = -1$ (charge $-e$), and (2) holes on the B sublattice of cations with $\delta n = 1$ (charge $+e$). V_c (V_a) is the self-energy of a hole (electron) due to its electrostatic interaction with all other charges in the ground state, defined by V_c (V_a) $=\sum_{j\in A} V_{ij}$ with $i\in B$ (A) [37]. The Coulomb energies V_{ij} decrease rapidly due to screening effects when the distance between sites i and j exceeds the distance to nearby metallic gates.

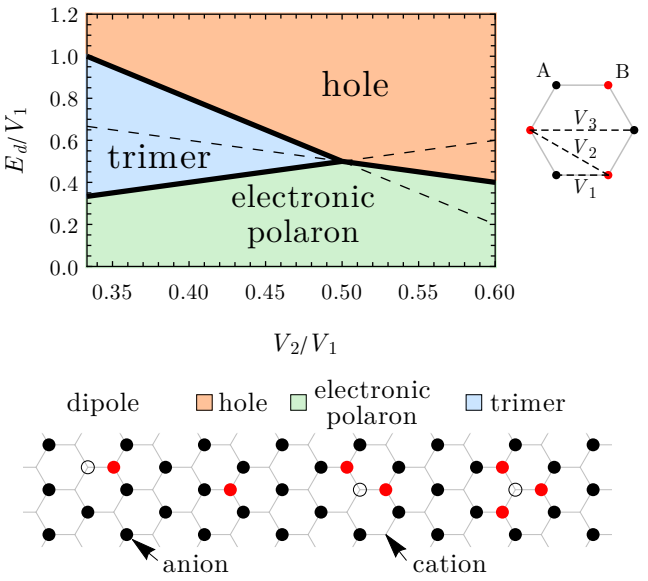


FIG. 1. Phase diagram showing which of the three charged excitations (hole, electronic polaron, or trimer) in the charge-transfer insulator at $n=1$ has the lowest energy per unit charge as a function of the fundamental gap E_d [defined in Eq. (3)] and the ratio of next-nearest-neighbor to nearest-neighbor repulsion: V_2/V_1. The dashed lines separate regions where the second lowest energy per charge excitation changes. See Fig. 4 in the appendix for more information.

It follows from Eq. (2) that adding a hole to the charge-transfer insulator costs energy $E_e-\mu$, where $E_e = \Delta+V_c$ and μ is chemical potential for holes. Likewise, adding an electron costs energy $E_{-e}+\mu$ with $E_{-e} = -V_a$. Transferring a charge from an anion to its adjacent cation creates a *charge-transfer exciton*, which carries an electric dipole. Its energy cost E_d is less than that the sum of the electron and the hole energies:

$$E_d = E_e + E_{-e} - V_1 = \Delta + V_c - V_a - V_1, \tag{3}$$

where V_1 denotes the nearest-neighbor repulsion (see Fig. 1). $E_d > 0$ defines the fundamental gap of the charge-transfer insulator to local neutral excitations.

Two holes cost energy $2E_e$ when they are spatially separated. Alternatively, consider binding two adjacent holes with a neutral charge-transfer dipole. The result is a charge-$2e$ composite excitation consisting of three adjacent holes on cations surrounding an electron at the center anion, a "trimer" (see Fig. 1). Its energy cost, written as $E_t = 2E_e - \epsilon_b$, differs from the two separate

holes by a pair binding energy ϵ_b:

$$\begin{aligned}\epsilon_b &= -E_d + 2V_1 - 3V_2 \\ &= -\Delta + 6V_2' - 3V_2 - 3V_3 + \cdots \\ &\approx -\Delta + 3V_2 - 3V_3 + \cdots , \end{aligned} \tag{4}$$

where the second-nearest neighbor interactions V_2' and V_2 are within the A and B sublattice respectively, and $\cdots$ denote interactions at larger distance. In the second and third equalities, we have used Eq. (3) and $V_2' \approx V_2$.

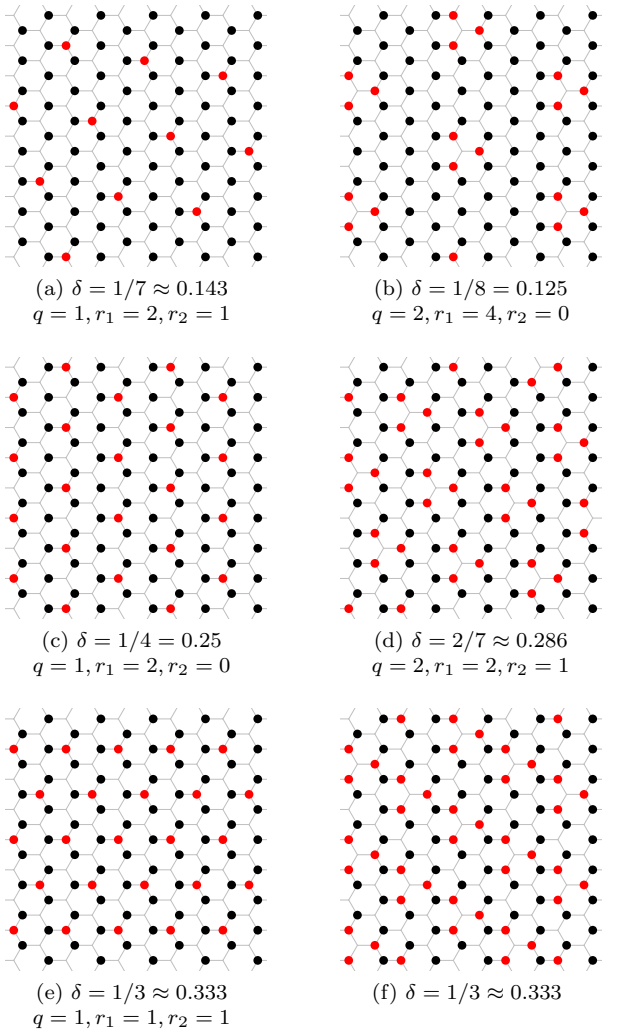
Importantly, in a range of realistic material parameters Δ and V_n (see Fig. 1), ϵ_b is positive so that a charge-$2e$ trimer costs less energy than two individual holes. The energy gain here comes mostly from the simple fact that pairing two holes into a trimer frees up three second-neighbor pairs in the system, which results in the $-3V_2$ energy reduction in Eq. (4). It is remarkable that despite the direct mutual repulsion, two doped holes can, at the expense of energy Δ, tightly bind together with a charge-transfer exciton to lower the total electrostatic repulsion energy.

As an example, for slightly twisted WSe_2/WS_2 with a moiré period $L_M = 7$nm and a distance to top and bottom gates equal to L_M, a calculation using Wannier functions finds $V_1, V_2', V_2, V_3 = 1.2998, 0.4599, 0.4780, 0.3239$ in units of $e^2/(\epsilon L_M)$, where ϵ is the permittivity of the dielectric environment [36]. The Coulomb energies at larger separation are much smaller. The trimer binding energy is then found to be $\epsilon_b = -14.9 + \frac{72.8}{\epsilon}$ meV.

The finding of charge pairing from Coulomb repulsion in a moiré superlattice is our first main result, which forms the basis for our theory at finite doping. Notably, previous works found an effective attraction between two added charges in small Hubbard-model clusters at intermediate U/t [38–40] or with extended interactions [41]. In this study of extended moiré systems, the pair binding energy is already manifest in the strong-coupling limit $t_{ij} = 0$ without explicitly invoking any charge fluctuation or weakly coupled clusters. Note however, the fact crucial to our analysis that doped charges occupy quantized orbitals localized around discrete lattice points, rather than taking arbitrary positions in the continuum. It is this quantum-mechanical effect that leads to the quantized energy of trimers.

We also mention in passing that besides holes and trimers, other composite excitations can be energetically favorable in certain parameter ranges. These include (1) the electronic polaron ($q = e$), which is a bound state of a hole and a dipole, and (2) higher-charge excitations with $q \geq 3e$. By comparing the energy cost of different types of charged excitations at filling $n = 1$, we identify the excitation with the least energy per unit charge—the cheapest charge excitation; see Fig. 1.

Pair Density Waves—Building on these results on few-body excitations, we now study the many-body physics of TMD heterobilayers at fillings $n = 1 + \delta$ with small $\delta > 0$, which are doped charge-transfer insulators. In particular, we shall develop an analytically-controlled theory to predict pair density waves and superconductivity under appropriate conditions.

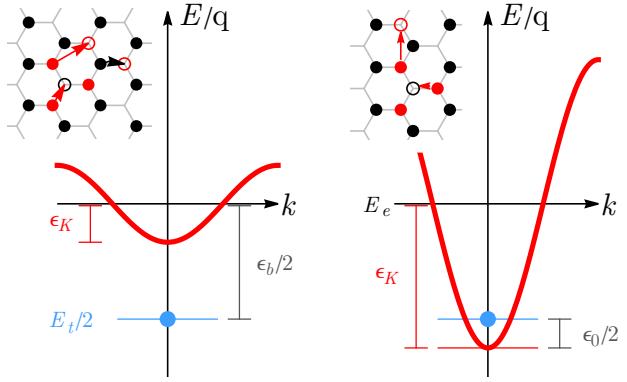


(a) $\delta = 1/7 \approx 0.143$, $q = 1, r_1 = 2, r_2 = 1$
(b) $\delta = 1/8 = 0.125$, $q = 2, r_1 = 4, r_2 = 0$
(c) $\delta = 1/4 = 0.25$, $q = 1, r_1 = 2, r_2 = 0$
(d) $\delta = 2/7 \approx 0.286$, $q = 2, r_1 = 2, r_2 = 1$
(e) $\delta = 1/3 \approx 0.333$, $q = 1, r_1 = 1, r_2 = 1$
(f) $\delta = 1/3 \approx 0.333$

FIG. 2. Examples of commensurate charge (a,c,e) and pair (b,d,f) density waves at various fractional fillings $n = 1 + \delta$. The charge q and integers r_1 and r_2 from Eq. (5) are listed for the triangular Wigner crystals of holes ($q = e$) and of trimers ($q = 2e$).

For large moiré periods, to minimize the dominant Coulomb interaction, a periodic array of charge excitations of the least cost is favored. If the cheapest excitation is a charge-e hole, then the ground state of H_0 will be a commensurate charge density wave at e.g. the dopings $\delta = 1/7$, $1/4$, or $1/3$, shown in Fig. 2(a,c,e). Similar charge density waves have been discussed in the context of twisted bilayer graphene [42].

On the other hand, if the cheapest charge excitation is a charge-$2e$ trimer, then the ground state of H_0 is a pair density wave with a commensurate periodicity at dopings such as $\delta = 1/8, 2/7$, or $1/3$, shown in Fig. 2(b,d,f). In



(a) pair density wave or BEC (b) resonantly-paired superconductivity

FIG. 3. Energy spectrum of the dispersive hole band and the trimer. (a) When the trimer energy per charge, $E_t/2$, lies below the hole band bottom, a pair density wave (PDW) or Bose-Einstein condensate of trimers is expected. (inset) A third-order trimer hopping process. (b) When $E_t/2$ lies close to the hole band bottom, low-energy holes interact resonantly with trimers, leading to superconductivity. (inset) A second-order process in which a trimer decays into two holes separated by a distance $2L_{\rm M}$.

particular, the pair density wave at $\delta = 2/7$ (shown in Fig. 2d) can be viewed as the closet packing of trimers with negligible inter-trimer interaction (involving only $V_{n\geq 5}$), in contrast to Fig. 2f (which involves $V_{n\geq 2}$).

More generally, we predict that at low temperature, clean TMD heterobilayers with a large moiré period should exhibit a sequence of insulating density wave states at the following fillings when a commensurate triangular lattice of charge excitations is formed:

$$n = 1 + \frac{q}{r_1^2 + r_1 r_2 + r_2^2} \qquad \text{(5)}$$

The integer $q > 0$ is the charge per excitation, and the integers $r_i \geq 0$ specify the Bravais lattice vector $r_1\mathbf{b}_1 + r_2\mathbf{b}_2$ of the commensurate density wave, with $\mathbf{b}_1, \mathbf{b}_2$ denoting the two lattice vectors of the moiré honeycomb lattice shown in Eq. (5).

Resonantly-Paired Superconductivity—While Coulomb repulsion favors density waves, the single-particle hopping term H_K favors charge delocalization. In the following, we study the competition between Coulomb energy and kinetic energy in the interesting and experimentally relevant parameter regime $\epsilon_b > 0$, where the trimer has a lower energy than two separated holes in the limit $t_{ij} = 0$ (the trimer region of Fig. 1). For simplicity, we will consider the scenario that the system is fully spin polarized, which is experimentally realized in WSe_2/WS_2 under a small magnetic field (less than 1T at 1.6K) [26].

Single-particle hopping t_{ij} between cation sites on the honeycomb lattice leads to a dispersive band of doped holes. Then, the lowest-energy hole excitation is in the delocalized state at the bottom of this band. Its energy is $E_e - \epsilon_K$, where $\epsilon_K > 0$ is proportional to the hopping integral t. In contrast, as a composite excitation, the trimer can only hop via a high-order process involving high-energy intermediate states (see Fig. 3a). In the strong-coupling regime, the trimer hopping integral is on the order of t^3/V^2 with $V \sim e^2/\epsilon L_{\rm M}$ and thus likely negligible.

With $t_{ij} \neq 0$, it is important to consider the hybridization between trimers and holes. This occurs when a constituent e charge in a trimer hops back to the center anion, leaving behind two adjacent holes in a high-energy state due to their strong mutual repulsion V_2. To lower their energy, these two remaining holes tend to hop away from each other. Thus, converting a trimer into two weakly-interacting holes that are sufficiently apart requires at least a second-order process, shown in Fig. 3b. Due to the large energy barrier in the pathway between the trimer and two free holes, the trimer remains a long-lived quasi-bound state.

Therefore, at finite doping, we are faced with a mixture of charge-e holes and charge-$2e$ trimers that are hybridized and at a total charge density δ. For small δ, the typical distance between doped charges is much larger than the moir'e period $L_{\rm M}$; hence the underlying moiré lattice only plays a minor role. Thus, the essential low-energy physics is captured by a boson-fermion model in the continuum, which we introduce for doped charge-transfer insulators:

$$H_\delta = \int d\mathbf{r} \sum_{\sigma=\pm} \psi_\sigma^\dagger \left(-\frac{\nabla^2}{2m}\right)\psi_\sigma + \epsilon_0\, \phi^\dagger\phi + g\,(\phi\psi_+^\dagger\psi_-^\dagger + \phi^\dagger\psi_-\psi_+) - \mu\, n(\mathbf{r}) + \frac{1}{2}\int d\mathbf{r}'\, V(|\mathbf{r}-\mathbf{r}'|)\, n(\mathbf{r})n(\mathbf{r}')$$
$$\text{with } n(\mathbf{r}) = \sum_{\sigma=\pm} \psi_\sigma^\dagger\psi_\sigma + 2\phi^\dagger\phi \qquad \text{(6)}$$

where ψ and ϕ denote the itinerant hole and immobile trimer, respectively.

According to band structure calculations for WSe_2/WS_2, $t_{ij} > 0$ between nearest-neighbor cations. Hence the band of doped holes has two degenerate minimum at corners of the moiré Brillouin zone $\pm\mathbf{K} = (0, \pm\frac{4\pi}{3L_M})$ [28], which are denoted by the valley index $\sigma = \pm$ in Eq. (6). $m \propto 1/t$ is the effective mass at the band bottom. $\epsilon_0 = -\epsilon_b + 2\epsilon_K$ with $\epsilon_K \propto t$ is the energy difference between a trimer and two delocalized holes at the band bottom. Since the hopping integral t changes significantly with the moiré period, the detuning parameter ϵ_0 is tunable by varying the twist angle.

Two types of interactions are included in our model Hamiltonian H_δ and play dominant roles in the *low density* regime $\delta \ll 1$: (1) the extended Coulomb interaction $V(r)$, whose range is determined by the distance to metallic gates, and (2) the local hybridization g between holes and trimers. The form of the hybridization is dictated by symmetry. The trimer state with maximal spin is invariant under three-fold rotation around the center anion and odd under reflection which exchanges a pair of holes (fermions). Therefore, the trimer hybridizes with a valley-singlet pair of holes $\psi_+^\dagger\psi_-^\dagger$, which transform in the same way (note that reflection interchanges $\pm\mathbf{K}$). Despite being a weak local interaction, the hole-trimer hybridization can have dramatic consequences for our system at low density.

Our model exhibits an enormously rich phase diagram resulting from the interplay of (1) the kinetic energy of holes, (2) the binding energy of trimers, (3) Coulomb repulsion between charges, and (4) the hybridization between holes and trimers. In particular, we shall show that pair density wave and superconducting ground states emerge in certain parameter regimes.

If the hopping integral t in the microscopic Hamiltonian (1) is small compared to the binding energy of trimers $\epsilon_b > 0$ (see Fig. 3a), then the kinetic energy of holes is of minor importance. Therefore $\epsilon_0 \approx -\epsilon_b < 0$, and doped charges go into trimers. At finite charge density up to $\delta = 1/3$, the (screened) Coulomb repulsion $V(r)$ between charges leads to a Wigner crystal of charge-$2e$ trimers, which takes a triangular lattice structure. At the dopings specified by Eq. (5), this trimer Wigner crystal is commensurate with and pinned by the moiré potential. The resulting state is a gapped and insulating pair density wave. At sufficiently low doping where the average distance between trimers exceeds the range of $V(r)$, the density wave state becomes fragile and potentially unstable to Bose condensation of trimers when their small hopping amplitudes are taken into account.

As the hopping integral t increases, the bottom of the hole band is lowered and eventually falls below $E_t/2$, as shown in Fig. 3b. Correspondingly, the bare detuning parameter ϵ_0 changes from negative to positive. The true detuning parameter ϵ is renormalized by the hole-trimer hybridization: $\epsilon = \epsilon_0 - o(g^2)$ [43]. At negative detuning $\epsilon < 0$, there exists a true bound state of two e charges, which is a superposition of a trimer and a cloud of two holes. At positive detuning $\epsilon > 0$, no such bound state exists. However, when the detuning is small, low-energy holes and trimers are strongly hybridized. Two e charges at low energy or large de Broglie wavelength scatter *resonantly* in the s-wave valley-singlet channel. Such resonant interaction is universally characterized by a large scattering length, which can exceed the range of screened Coulomb repulsion $V(r)$. The scattering length is positive (negative) for $\epsilon > 0$ ($\epsilon < 0$) and diverges at $\epsilon = 0$.

The physics of the resonant interaction via the trimer channel is reminiscent of a Feshbach resonance in an ultracold Fermi gas [34]. Similar to the latter, under the resonant condition, the ground state of our system at low doping is a superconductor with same-spin pairing in the valley-singlet channel (corresponding to f-wave pairing symmetry on the honeycomb lattice). A crossover between the Bose-Einstein condensate (BEC) and Bardeen-Cooper-Schrieffer (BCS) regimes [35] can be achieved as a function of doping density δ and the detuning ϵ. See Appendix A for a mean-field theory analysis.

It should be noted that the origin of the resonant interaction in our system is completely different from the case of cold atoms. Unlike the molecule formed by two atoms in empty space, the trimer is a *charge-transfer excitation* in a *many-body* "vacuum" at the filling $n = 1$. It is remarkable that spin-polarized superconductivity can be realized in a solid-state system with purely Coulomb repulsion.

Discussion—To summarize, we developed a strong-coupling theory to predict electron pairing from repulsion via charge-transfer excitations in TMD heterobilayers. We further predict insulting pair density wave states at a sequence of doping levels. Finally, we show that with the increase of electron itinerancy, the resonant interaction between itinerant holes and local charge-$2e$ pairs leads to unconventional superconductivity. Our pairing mechanism may shed new insight on other moiré materials, such as twisted graphene multilayers, where charge redistribution under doping may be important [44, 45] and spin-polarized superconductivity may have been observed [46, 47]. We hope our prediction of new fascinating correlated states in moiré materials can stimulate further activity and find experimental realization soon.

We are grateful to Steve Kivelson for a stimulating and insightful discussion. We thank Jason Alicea, Chenhao Jin, Patrick Lee, Stevan Nadj-Perge, Gil Refael, Alex Thomson, and Sagar Vijay for helpful conversations. We thank Yang Zhang, Noah Yuan, and especially Tongtong Liu for collaboration on related works, and Kin Fai Mak, Jie Shan, and Feng Wang for valuable discussions on experiments. It is our pleasure to thank the organizers of the KITP conference "Topological Quantum Matter:

From Fantasy to Reality," where this work was initiated. K.S. is supported by the Walter Burke Institute for Theoretical Physics at Caltech. L.F. is supported by the DOE Office of Basic Energy Sciences under Award desc0018945, and in part by a Simons Investigator award from the Simons Foundation.

[1] R. Bistritzer and A. H. MacDonald, Proceedings of the National Academy of Science **108**, 12233 (2011), arXiv:1009.4203.
[2] L. Balents, SciPost Physics **7**, 048 (2019), arXiv:1909.01545.
[3] J. M. B. Lopes Dos Santos, N. M. R. Peres, and A. H. Castro Neto, Phys. Rev. Lett. **99**, 256802 (2007), arXiv:0704.2128.
[4] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Nature (London) **556**, 43 (2018a), arXiv:1803.02342.
[5] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, R. C. Ashoori, and P. Jarillo-Herrero, Nature (London) **556**, 80 (2018b), arXiv:1802.00553.
[6] X. Lu, P. Stepanov, W. Yang, M. Xie, M. A. Aamir, I. Das, C. Urgell, K. Watanabe, T. Taniguchi, G. Zhang, A. Bachtold, A. H. MacDonald, and D. K. Efetov, Nature (London) **574**, 653 (2019), arXiv:1903.06513.
[7] G. Chen, L. Jiang, S. Wu, B. Lyu, H. Li, B. L. Chittari, K. Watanabe, T. Taniguchi, Z. Shi, J. Jung, Y. Zhang, and F. Wang, Nature Physics **15**, 237 (2019a), arXiv:1803.01985.
[8] Y. Choi, J. Kemmer, Y. Peng, A. Thomson, H. Arora, R. Polski, Y. Zhang, H. Ren, J. Alicea, G. Refael, F. von Oppen, K. Watanabe, T. Taniguchi, and S. Nadj-Perge, Nature Physics **15**, 1174 (2019), arXiv:1901.02997.
[9] C. Shen, N. Li, S. Wang, Y. Zhao, J. Tang, J. Liu, J. Tian, Y. Chu, K. Watanabe, T. Taniguchi, R. Yang, Z. Y. Meng, D. Shi, and G. Zhang, (2019), arXiv:1903.06952.
[10] S. Carr, S. Fang, H. C. Po, A. Vishwanath, and E. Kaxiras, Physical Review Research **1**, 033072 (2019), arXiv:1907.06282.
[11] H. C. Po, L. Zou, T. Senthil, and A. Vishwanath, Phys. Rev. B **99**, 195455 (2019), arXiv:1808.02482.
[12] M. Koshino, N. F. Q. Yuan, T. Koretsune, M. Ochi, K. Kuroki, and L. Fu, Physical Review X **8**, 031087 (2018), arXiv:1805.06819.
[13] J. Kang and O. Vafek, Physical Review X **8**, 031088 (2018), arXiv:1805.04918.
[14] G. Tarnopolsky, A. J. Kruchkov, and A. Vishwanath, Phys. Rev. Lett. **122**, 106405 (2019), arXiv:1808.05250.
[15] X.-C. Wu, C.-M. Jian, and C. Xu, Phys. Rev. B **99**, 161405 (2019a), arXiv:1811.08442.
[16] S. Manzeli, D. Ovchinnikov, D. Pasquier, O. V. Yazyev, and A. Kis, Nature Reviews Materials **2**, 1–15 (2017).
[17] L. Wang, E.-M. Shih, A. Ghiotto, L. Xian, D. A. Rhodes, C. Tan, M. Claassen, D. M. Kennes, Y. Bai, B. Kim, K. Watanabe, T. Taniguchi, X. Zhu, J. Hone, A. Rubio, A. Pasupathy, and C. R. Dean, (2019), arXiv:1910.12147.
[18] Q. Shi, E.-M. Shih, M. V. Gustafsson, D. A. Rhodes, B. Kim, K. Watanabe, T. Taniguchi, Z. Papić, J. Hone, and C. R. Dean, (2019), arXiv:1911.04428.
[19] J. Wang, Q. Shi, E.-M. Shih, L. Zhou, W. Wu, Y. Bai, D. A. Rhodes, K. Barmak, J. Hone, C. R. Dean, and X. Y. Zhu, (2020), arXiv:2001.03812.
[20] C. Jin, E. C. Regan, A. Yan, M. Iqbal Bakti Utama, D. Wang, S. Zhao, Y. Qin, S. Yang, Z. Zheng, S. Shi, K. Watanabe, T. Taniguchi, S. Tongay, A. Zettl, and F. Wang, Nature (London) **567**, 76 (2019), arXiv:1812.09815.
[21] Y. Shimazaki, I. Schwartz, K. Watanabe, T. Taniguchi, M. Kroner, and A. Imamoğlu, (2019), arXiv:1910.13322.
[22] J. Sung, Y. Zhou, G. Scuri, V. Zólyomi, T. I. Andersen, H. Yoo, D. S. Wild, A. Y. Joe, R. J. Gelly, H. Heo, D. Bérubé, A. M. Mier Valdivia, T. Taniguchi, K. Watanabe, M. D. Lukin, P. Kim, V. I. Fal'ko, and H. Park, (2020), arXiv:2001.01157.
[23] Z. Bi and L. Fu, (2019), arXiv:1911.04493.
[24] F. Wu, T. Lovorn, E. Tutuc, and A. H. MacDonald, Phys. Rev. Lett. **121**, 026402 (2018), arXiv:1804.03151.
[25] F. Wu, T. Lovorn, E. Tutuc, I. Martin, and A. H. MacDonald, Phys. Rev. Lett. **122**, 086402 (2019b), arXiv:1807.03311.
[26] Y. Tang, L. Li, T. Li, Y. Xu, S. Liu, K. Barmak, K. Watanabe, T. Taniguchi, A. H. MacDonald, J. Shan, and K. F. Mak, Nature **579**, 353–358 (2020), arXiv:1910.08673.
[27] E. C. Regan, D. Wang, C. Jin, M. I. Bakti Utama, B. Gao, X. Wei, S. Zhao, W. Zhao, K. Yumigeta, M. Blei, J. Carlstroem, K. Watanabe, T. Taniguchi, S. Tongay, M. Crommie, A. Zettl, and F. Wang, Nature **579**, 359–363 (2020), arXiv:1910.09047.
[28] Y. Zhang, N. F. Q. Yuan, and L. Fu, (2019), arXiv:1910.14061.
[29] J. Zaanen, G. A. Sawatzky, and J. W. Allen, Phys. Rev. Lett. **55**, 418 (1985).
[30] V. J. Emery, Phys. Rev. Lett. **58**, 2794 (1987).
[31] F. C. Zhang and T. M. Rice, Phys. Rev. B **37**, 3759 (1988).
[32] E. Abrahams, S. Schmitt-Rink, and C. Varma, Physica B+C **148**, 257 (1987).
[33] M. Grilli, R. Raimondi, C. Castellani, C. Di Castro, and G. Kotliar, Phys. Rev. Lett. **67**, 259 (1991).
[34] V. Gurarie and L. Radzihovsky, Annals of Physics **322**, 2 (2007), arXiv:cond-mat/0611022.
[35] Q. Chen, J. Stajic, S. Tan, and K. Levin, Phys. Rep. **412**, 1 (2005), arXiv:cond-mat/0404274.
[36] T. Liu and Y. Zhang, private communication.
[37] $V_c = 3V_1 + 3V_3 + 6V_4 + \cdots$ and $V_a = 6V_2 + 6V_5 + \cdots$ where V_n is the repulsion between n^{th} nearest-neighbors, shown in Fig. 5 of the Appendix.
[38] S. R. White, S. Chakravarty, M. P. Gelfand, and S. A. Kivelson, Phys. Rev. B **45**, 5062 (1992).
[39] H. Yao, W.-F. Tsai, and S. A. Kivelson, Phys. Rev. B **76**, 161104 (2007), arXiv:0706.0761.
[40] L. Isaev, G. Ortiz, and C. D. Batista, Phys. Rev. Lett. **105**, 187002 (2010), arXiv:1003.5207.
[41] K. Slagle and Y. B. Kim, SciPost Phys. **6**, 16 (2019), arXiv:1805.05331.
[42] B. Padhi, C. Setty, and P. W. Phillips, Nano Letters **18**, 6175 (2018), arXiv:1804.01101.
[43] See e.g. Appendix B.1 of Ref. 34.
[44] L. Rademaker and P. Mellado, Phys. Rev. B **98**, 235158

(2018), arXiv:1805.05294.
[45] F. Guinea and N. R. Walet, Proceedings of the National Academy of Science **115**, 13174 (2018), arXiv:1806.05990.
[46] X. Liu, Z. Hao, E. Khalaf, J. Y. Lee, K. Watanabe, T. Taniguchi, A. Vishwanath, and P. Kim, (2019), arXiv:1903.08130.
[47] G. Chen, A. L. Sharpe, P. Gallagher, I. T. Rosen, E. Fox, L. Jiang, B. Lyu, H. Li, K. Watanabe, T. Taniguchi, J. Jung, Z. Shi, D. Goldhaber-Gordon, Y. Zhang, and F. Wang, Nature **572**, 215–219 (2019b), arXiv:1901.04621.
[48] N. Luo, (2000), arXiv:cond-mat/0101004.
[49] C. M. Varma, Phys. Rev. Lett. **61**, 2713 (1988).

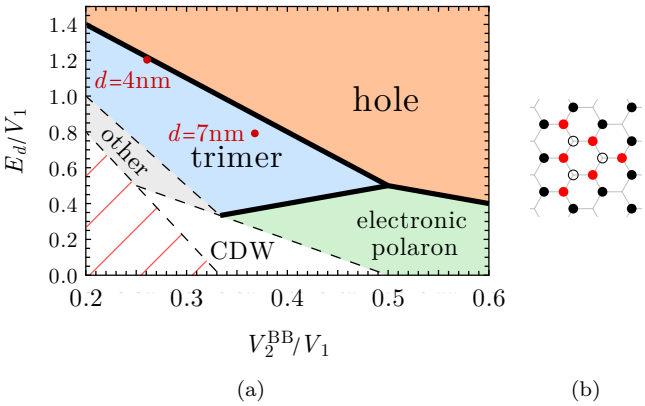


FIG. 4. (a) A detailed excitation phase diagram of the lowest energy per charge (E/q) excitation. The red dots mark estimates for a slightly twisted WSe_2/WS_2 inside a dielectric environment with permittivity $\epsilon = 3$ and for different distances d to metallic gates. In the "other" region, other excitations have the least E/q, such as the charge-3e excitation shown in (b). The bottom-left region (white with red stripes) is inaccessible (when $V_{n\geq 3} = 0$) as it would require $\Delta < 0$. The dashed lines are drawn for $V_{n\geq 3} = 0$. See Appendix A 1 for more details.

Appendix A: Additional Details

In this appendix, we will provide extra details on the low-energy excitations, followed by a mean-field theory analysis of the superconducting states starting from the lattice model in Eq. (1). We will consider the low density and strongly interacting limit:

$$n - 1 = \delta \ll 1 \quad \text{and} \quad t_{ij} \ll V_2 \tag{A1}$$

where V_n is the Coulomb repulsion between n-th nearest-neighbors (see Fig. 5). We also assume that the spins are fully polarized, e.g. by an external magnetic field.

1. Electrostatics

In Fig. 1 of the main text, we showed which of the following three excitations has the smallest energy per charge, E/q:

$$\begin{aligned} \text{hole } (q=1): &\quad E_{-e} \\ \text{trimer } (q=2): &\quad E_t = 2E_e + E_d - 2V_1 + 3V_2^{BB} \\ \text{electronic polaron } (q=1): &\quad E_{ep} = E_e + E_d - V_1 + V_2^{BB} \end{aligned} \tag{A2}$$

Remarkably, for the above three excitations, this only requires knowing E_d/V_1 and V_2^{BB}/V_1; Δ and all V_n are effectively absorbed into these two ratios.

It is remarkable that charge-2e pairing can occur from just repulsive interactions. The nontrivial charge-transfer insulating background is essential for the trimer

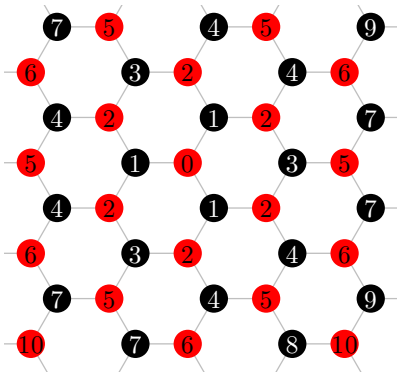
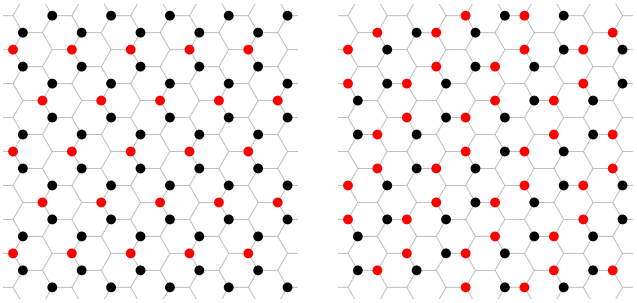


FIG. 5. n^{th} nearest-neighbors on a honeycomb lattice. Throughout this work, V_n denotes the Coulomb repulsion between n^{th} nearest-neighbors. We sometimes differentiate between V_n^{AA} and V_n^{BB} for repulsions between a pair of sites on the A or B sublattice. n^{th} nearest-neighbors are separated by a_n with $a_{n=1,2,3,4,5}/a_1 = 1, \sqrt{3}, 2, \sqrt{7}, 3$. $L_{\text{M}} = a_2$ is the moiré period.



(a) (b)

FIG. 6. Possible instabilities of the charge transfer insulator. (a) A Wigner crystal of dipoles, which is a possible state when the dipole energy is negative, $E_d < 0$. (b) The charge density wave that occurs in the CDW region (white) of Fig. 4 when $V_{n\geq 4} = 0$ and $V_3 \ll V_2$.

stability; a trimer is not stable in the vacuum [48]. As an aid to intuition, we give an even simpler example of how this can occur in a finite system in Appendix B.

In Fig. 4, we show a more detailed diagram of the smallest E/q excitation when arbitrary excitations are considered. We also check for instabilities (shown in Fig. 6) of the charge transfer insulator, which occur in the "CDW" region of Fig. 4 and when $E_d < 0$. The energy of other excitations and these instabilities depend on more than just E_d/V_1 and V_2^{BB}/V_1. Therefore, we show the locations of these other excitations and instabilities for the simple case when $V_{n\geq 3} = 0$. In Fig. 4, dashed lines are used to depict boundaries that assume $V_{n\geq 3} = 0$.

We also estimate where in the phase diagram a slightly twisted WSe_2/WS_2 with a moiré period $L_{\text{M}} = 7$nm could be, which we show in Fig. 4 using red dots. The locations are calculated using $\Delta = 14.9$meV and the values of V_n shown in Tab. I. Δ and V_n were calculated using Wannier

n	0	1	2	3
V_n^{AA}	3.7769		0.2292	
V_n^{AB}		0.9479		0.1340
V_n^{BB}	2.9828		0.2472	

(a) $d = 4$nm

n	0	1	2	3
V_n^{AA}	4.2407		0.4599	
V_n^{AB}		1.2998		0.3239
V_n^{BB}	3.4132		0.4780	

(b) $d = 7$nm

TABLE I. The values of V_n (in units of $\frac{e^2}{\epsilon L_{\text{M}}} = \frac{205.7}{\epsilon}$meV with $L_{\text{M}} = 7$nm) used to estimate the location (denoted by red dots in Fig. 4) of slightly twisted WSe_2/WS_2 in the excitation phase diagram.

orbitals and a Coulomb interaction $V(r)$ that is screened by a pair of metallic gates, each a distance $\pm d$ from the TMD hererobilayer. [36] By modeling the gates as perfect conductors, the screened Coulomb interaction can be calculated using the method of image charges, yielding[1]

$$V(r) = \frac{e^2}{\epsilon} \sum_{z\in\mathbb{Z}} \frac{(-1)^z}{\sqrt{r^2 + (2dz)^2}} \tag{A3}$$

When $r \gg d$, $V(r)$ decays exponentially.

In Fig. 7, we point out possible experimental evidence of insulating pair density waves of trimers from recently observed resistivity peaks [27].

2. Mean Field Theory of Superconductivity

Here, we study the mean field theory of trimer superconductivity. Suppose that we are near the edge of the trimer region of the phase diagram, so that the trimer binding energy, ϵ_b [Eq. (4)], is small:

$$0 < \epsilon_b \ll V_2 \tag{A4}$$

We will also assume that excitations other than the charge-e hole and charge-$2e$ trimers (such as the dipole and electronic polaron) have a large energy cost $\sim V_2$ so that they can be ignored.

The low-energy Hamiltonian thus consists of only the mobile holes $c_{\mathbf{k}}$ on the cations, and bosonic trimers b_a centered on the anions:

$$H_{\text{eff}} = \sum_{\mathbf{k}} \epsilon(\mathbf{k}) c_{\mathbf{k}}^\dagger c_{\mathbf{k}} + (\epsilon_0 - 2\mu) \sum_a b_a^\dagger b_a - \tilde{g} \sum_{\mathbf{k},\mathbf{k}'} (b_{\mathbf{k}+\mathbf{k}'}^\dagger c_{\mathbf{k}} c_{\mathbf{k}'} + h.c.) + \cdots \tag{A5}$$

[1] A single parallel conductor results in $V(r) \propto \frac{1}{r} - 1/\sqrt{r^2 + (2d)^2}$.

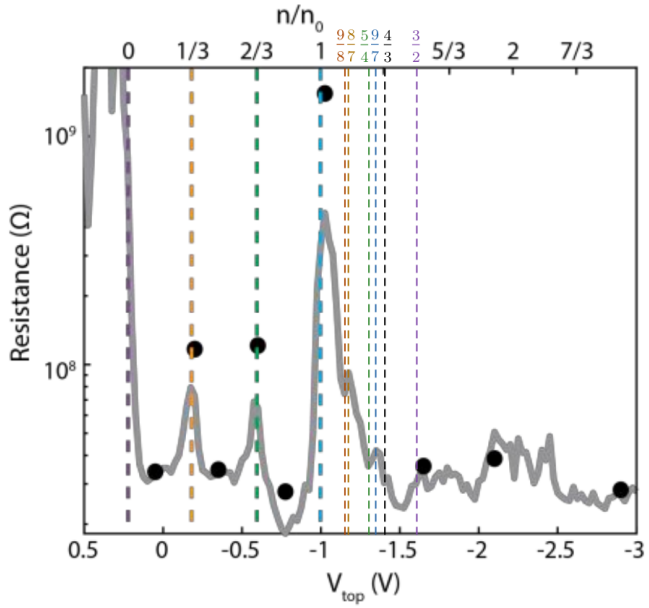
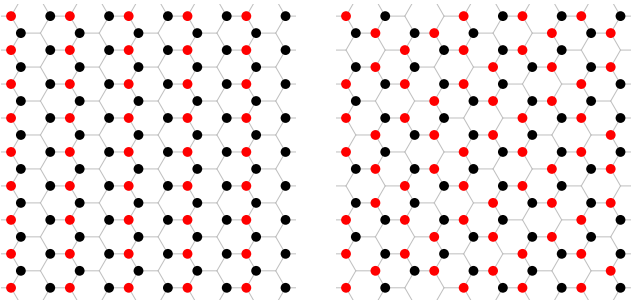


FIG. 7. Resistance of the Berkeley group's WSe_2/WS_2 moiré device as a function of gate voltage, which determines the electron filling, n. The figure is copied from Ref. 27. We add vertical lines to indicate various Wigner crystal fillings. The blue line at filling $n = 9/7$ is particularly interesting because it shows a large resistance peak at the same filling as the pair density wave in Fig. 2d. However, these peaks were not observed by the Cornell group [26]. Energetically favorable charge density waves at filling $n = 3/2$ are shown in Fig. 8.



(a) $\delta = 1/2$ (b) $\delta = 1/2$

FIG. 8. A charge density wave state at doping $\delta = 1/2$ that is typical of the (a) hole and (b) trimer regimes of the excitation phase diagram in Fig. 4.

The first term gives the dispersion of the holes, which hop on the triangular lattice of cations. At low density, $\delta \ll 1$, $\epsilon(\mathbf{k})$ can be expanded about its band minima $\pm\mathbf{K} = (0, \pm\frac{4\pi}{3L_M})$:

$$\epsilon(\pm\mathbf{K} + \mathbf{k}) = \frac{1}{2m}k^2 - \mu + O(k^3) \tag{A6}$$

$m \approx 2/9t_2$ where t_n is the hopping energy between n-th nearest-neighbors. We have shifted $\epsilon(\mathbf{k})$ so that $\epsilon(\pm\mathbf{K}) = 0$.

The second term sets the energy cost of trimers. $b_a^\dagger$ creates a trimer centered at the anion a, and $b_{\mathbf{k}}^\dagger$ creates a trimer with momentum $\mathbf{k}$. $\epsilon_0 \approx 6t_2 - \epsilon_b$ is the energy difference between a trimer and two holes at the band minima $\pm\mathbf{K}$.

The last term accounts for the conversion between a trimer and a pair of holes in the s-wave channel. The conversion amplitude $\tilde{g}$ can be calculated perturbatively; the leading contribution comes from the process shown in Fig. 3b and has amplitude $\tilde{g} \sim t_1 t_2/V_2$.[2]

The "$\cdots$" in H_{eff} denotes other terms that could be included in H_{eff}. We will ignore these terms in the following mean-field analysis because we do not expect these terms to be relevant in the resonantly-paired superconductivity regime of interest. To justify this, consider two potentially important kinds terms that we are omitting. The first is a trimer kinetic energy term, $-t_t \sum_{a'a} b_{a'}^\dagger b_a$, where $t_t \sim t^3/V^2$ is the trimer hopping energy, resulting from the perturbative process shown in Fig. 3a. However, we expect that near resonance, this term is negligible compared to the effective boson mass resulting from the coupling $\tilde{g}$ to the fermions. The second potentially important terms are 4-fermion interactions, such as $V_{ij} n_i n_j$. However, we will soon see [from Eq. (A16)] that near resonance, the fermion and boson operators scale as $c \sim b \sim O(\sqrt{\delta})$ at low density δ. Therefore the first term in H_{eff} is $O(t\delta)$; the third term contributes $O(\delta^{3/2}\tilde{g}) \sim O(\delta^{3/2}t^2/V_2)$; and a 4-fermion interaction would contribute $O(\delta^2 V)$. Thus, we expect that the 4-fermion interaction is negligible when $\delta^2 V \ll \delta^{3/2}t^2/V$, i.e. at the sufficiently-low doping $\delta \ll (t/V)^4$.

To make a connection to H_δ in Eq. (6) of the main text, note that $\psi^\dagger \sim L_M^{-1} c^\dagger$ and $\phi^\dagger \sim L_M^{-1} b^\dagger$. Then g in H_δ and $\tilde{g}$ in H_{eff} [Eq. (A5)] are related by $g \sim L_M \tilde{g}$.

To make analytical progress, we consider the following mean-field approximation:

$$\begin{aligned} b_a^\dagger b_a &= b_a^\dagger \langle b_a\rangle + \langle b_a^\dagger\rangle b_a - \langle b_a^\dagger\rangle\langle b_a\rangle + (b_a^\dagger - \langle b_a^\dagger\rangle)(b_a - \langle b_a\rangle) \\ &\approx b_a^\dagger \langle b_a\rangle + \langle b_a^\dagger\rangle b_a - \langle b_a^\dagger\rangle\langle b_a\rangle \end{aligned} \tag{A7}$$

and $b_{\mathbf{k}+\mathbf{k}'}^\dagger c_{\mathbf{k}} c_{\mathbf{k}'} \approx \langle b_{\mathbf{k}+\mathbf{k}'}^\dagger\rangle c_{\mathbf{k}} c_{\mathbf{k}'}$. With this approximation, the low-energy Hamiltonian becomes quadratic:

$$H_{\text{MF}} = \sum_{\mathbf{k}} \begin{pmatrix} c_{\mathbf{k},+} \\ c_{\mathbf{k},-}^\dagger \end{pmatrix}^\dagger \begin{pmatrix} +\epsilon_k & -\Delta_b \\ -\Delta_b & -\epsilon_k \end{pmatrix} \begin{pmatrix} c_{\mathbf{k},+} \\ c_{\mathbf{k},-}^\dagger \end{pmatrix} + \frac{\epsilon_0 - 2\mu}{\tilde{g}^2}\Delta_b^2$$
$$\Delta_b = \tilde{g}\langle b_a\rangle \tag{A8}$$

where we take the dispersion to be $\epsilon_k = \frac{1}{2m}k^2 - \mu$. Δ_b is the superconducting order parameter. $\Delta_b > 0$ will be assumed to be positive, without loss of generality.

[2] The perturbative approximation $\tilde{g} \sim t_1 t_2/V_2$ formally also requires assuming $V_4 \ll V_2$ so that V_4 only results in a perturbative correction to the final state energy in Fig. 3b.

The ground state energy can be written as

$$E = -\int_E D(E)\sqrt{E^2+\Delta_b^2} + \frac{\epsilon_0 - 2\mu}{\tilde{g}^2}\Delta_b^2 \tag{A9}$$

$$D(E) = \int_{\mathbf{k}} \delta(E-\epsilon_k) = \begin{cases} 2\pi m & -\mu < E < W \\ 0 & \text{otherwise} \end{cases} \tag{A10}$$

where $D(E)$ is the density of single-particle states and $\int_{\mathbf{k}} = \int \frac{d^3\mathbf{k}}{(2\pi)^2}\Theta(W-\epsilon_k)$ integrates over momentum states with energy $\epsilon_k < W$. W is a UV cutoff which should be taken to be roughly equal to the bandwidth: $W \sim 9t_2 \approx 2m^{-1}$. Evaluating the integral yields

$$\begin{aligned} E = -\pi m \Bigg[& W\sqrt{W^2+\Delta_b^2} + \Delta_b^2 \log\left(W + \sqrt{W^2+\Delta_b^2}\right) \\ & \mu\sqrt{\mu^2+\Delta_b^2} + \Delta_b^2 \log\left(\mu + \sqrt{\mu^2+\Delta_b^2}\right) \\ & - 2\Delta_b^2 \log \Delta_b \Bigg] + \frac{\epsilon_0 - 2\mu}{\tilde{g}^2}\Delta_b^2 \end{aligned} \tag{A11}$$

The superconducting order parameter Δ_b can be calculated by minimizing the energy as a function of Δ_b, which yields

$$\Delta_b = \frac{\sqrt{W^2+\mu^2+2W\mu\cosh\frac{\epsilon_0-2\mu}{\pi m\tilde{g}^2}}}{\sinh\frac{\epsilon_0-2\mu}{\pi m\tilde{g}^2}} \tag{A12}$$

Δ_b depends strongly on the chemical potential μ, which can be obtained from the filling constraint:

$$\begin{aligned} \delta &= \delta_c + 2\delta_b \\ \delta_c &= \langle c_i^\dagger c_i\rangle = 4\pi m\mu \\ \delta_b &= \langle b_a^\dagger b_a\rangle \approx \frac{\Delta_b^2}{\tilde{g}^2} \end{aligned} \tag{A13}$$

δ_c is the density of holes. δ_b is the density of bosonic trimers, which we approximate at the mean-field level: $\langle b_a^\dagger b_a\rangle \approx \langle b_a^\dagger\rangle\langle b_a\rangle = \Delta_b^2/\tilde{g}^2$.

There are two regimes: (1) BCS superconductivity when $\frac{\epsilon_0-2\mu}{\pi m\tilde{g}^2} \gg 1$, and (2) resonantly-paired superconductivity when $\epsilon_0 \approx 2\mu$.

BCS Superconductivity Regime

When $\frac{\epsilon_0-2\mu}{\pi m\tilde{g}^2} \gg 1$, the order parameter can be approximated as

$$\Delta_b \approx 2\sqrt{W\mu}\,\exp\left(-\frac{\epsilon_0-2\mu}{2\pi m\tilde{g}^2}\right) \tag{A14}$$

and a BCS superconductivity regime occurs where Δ_b is very small.[3] As a result, the boson density is very small, $\delta_b \ll \delta$, which allows us to approximately solve for the chemical potential from Eq. (A13):

$$\mu \approx \frac{\delta}{4\pi m} \tag{A15}$$

This regime is very similar to BCS superconductivity. This can be understood by integrating out the boson to obtain a 4-fermion interaction $\tilde{g}' c_+^\dagger c_-^\dagger c_- c_+$ with $\tilde{g}' \sim \frac{\tilde{g}^2}{\epsilon_0-2\mu}$. In terms of $\tilde{g}'$, the order parameter Δ_b scales exactly the same as the BCS order parameter (in two spatial dimensions), $\Delta_b \sim \Delta_{\text{BCS}} \sim \sqrt{W\mu}e^{-1/D\tilde{g}}$, where the density of states is $D = 2\pi m$ from Eq. (A10).

Note that in this regime, the boson density is very small, so the $\tilde{g}$ coupling term in H_{eff} [Eq. (A5)] contributes very little to the energy. Therefore in this regime, the terms in the "$\cdots$" of H_{eff} are likely to play an important role and possibly result in other kinds of symmetry breaking. So although BCS-like superconductivity results when the "$\cdots$" terms are dropped, a more detailed analysis is needed to determine the true ground state in this regime when the "$\cdots$" terms are included.

Resonantly-Paired Superconductivity Regime

At higher doping, the chemical potential approaches its maximum value, $\mu \approx \epsilon_0/2$, at which point the boson density begins to increase significantly. Setting $\mu \approx \epsilon_0/2$ allows us to solve for the boson density δ_b in Eq. (A13), which can be used to express the order parameter:[4]

$$\begin{aligned} \Delta_b &\approx \tilde{g}\sqrt{\delta_b} \\ \delta_b &\approx \frac{1}{2}\delta - \pi m\epsilon_0 \end{aligned} \tag{A16}$$

Due to the significantly larger boson density, the order parameter $\Delta_b \sim \tilde{g}$ is immensely larger in this resonantly-paired regime than in the BCS regime where $\Delta_b \sim e^{-1/\tilde{g}^2}$ [Eq. (A14)].

[3] The $\sqrt{W\mu}$ prefactor in Δ_b [in Eq. (A14)] comes from the limits of integration ($-\mu$ to W) in Eq. (A10). Eq. (A14) is only valid when $\mu > 0$. If $\mu = 0$, then note that the ground state energy in the $\mu = 0$ limit is equivalent to the energy in the $\mu = W$ limit if the mass is halved; i.e. $E_{\mu=0} = E_{\mu=W}^{m\to m/2}$ in Eq. (A11). Therefore if $\mu = 0$, then $\Delta_b \approx 2W\exp\left(-\frac{\epsilon_0-2\mu}{\pi m\tilde{g}^2}\right)$ [by replacing $\mu \to W$ and $m \to m/2$ in Eq. (A14)], which is significantly smaller than the expression in Eq. (A14) when $0 < \mu$ and $\frac{\epsilon_0-2\mu}{\pi m\tilde{g}^2} \gg 1$ due to the missing factor of $\frac{1}{2}$ in the exponent.

[4] A similar equation for the boson density in three spatial dimensions appears in Eq. (6.8) of Ref. 34.

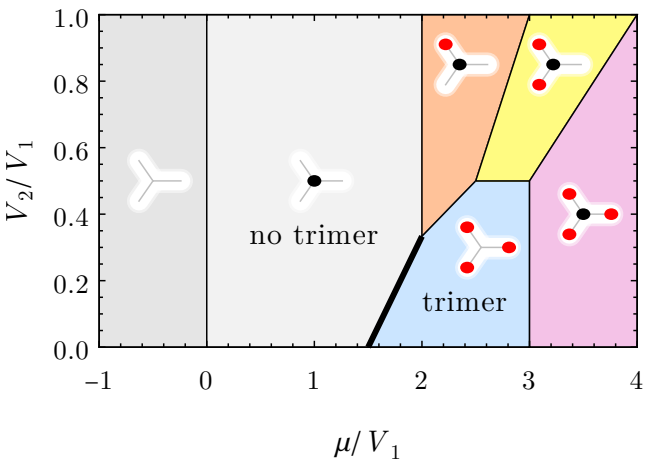
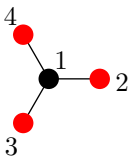


FIG. 9. Ground states of the 4-site cluster model in Eq. (B1) with $\Delta = V_1$. Valence skipping (where charge of the ground state jumps by two) [49] occurs along the thick line. Black and red dots denote filled orbitals with $n_i = 1$.

Appendix B: Valence Skipping in a 4-site Cluster

A toy model for trimer stability is obtained by considering a 4-site cluster of the Hamiltonian H_0 in Eq. (1) at a chemical potential μ:



$$\begin{aligned} H_4 &= \Delta\,(n_2+n_3+n_4) \\ &\quad + V_1\,n_1(n_2+n_3+n_4) \\ &\quad + V_2\,(n_2n_3+n_3n_4+n_4n_2) \\ &\quad - \mu\,(n_1+n_2+n_3+n_4) \end{aligned} \tag{B1}$$

Each site either has 0 or 1 fermions, $n_i = 0, 1$, which is physically relevant when a large on-site Hubbard interaction prevents double occupancy.

The ground state phase diagram of the 4-site cluster is shown in Fig. 9. The "no trimer" region is analogous to a change transfer insulator, while the "trimer" region is analogous to a trimer excitation of the change transfer insulator.

The ground states are easiest to understand in an ideal limit where $V_2 = 0$, $\Delta = V_1$, and $\mu = \frac{3}{2}V_1$. Then $H_4 = V_1\,(n_1 - \frac{1}{2})(n_2+n_3+n_4-\frac{3}{2})$, and it's simple to see that the "no trimer" and "trimer" states are degenerate ground states.