

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

# Ground-state phase diagram of the square lattice Hubbard model from density matrix embedding theory

Bo-Xiao Zheng and Garnet Kin-Lic Chan

Phys. Rev. B **93**, 035126 — Published 21 January 2016

DOI: [10.1103/PhysRevB.93.035126](https://doi.org/10.1103/PhysRevB.93.035126)

# Ground-state phase diagram of the square lattice Hubbard model from density matrix embedding theory

Bo-Xiao Zheng and Garnet Kin-Lic Chan

Department of Chemistry, Princeton University, New Jersey 08544, United States

We compute the ground-state phase diagram of the Hubbard and frustrated Hubbard models on the square lattice with density matrix embedding theory using clusters of up to 16 sites. We provide an error model to estimate the reliability of the computations and complexity of the physics at different points in the diagram. We find superconductivity in the ground-state as well as competition between inhomogeneous charge, spin, and pairing states at low doping. The estimated errors in the study are below  $T_c$  in the cuprates and on the scale of contributions in real materials that are neglected in the Hubbard model.

The Hubbard model<sup>1–3</sup> is one of the simplest quantum lattice models of correlated electron materials. Its one-band realization on the square lattice plays a central role in understanding the essential physics of high temperature superconductivity<sup>4,5</sup>. Rigorous, near exact results are available in certain limits<sup>6</sup>: at high temperatures from series expansions<sup>7–10</sup>, in infinite dimensions from converged dynamical mean-field theory<sup>11–14</sup>, and at weak coupling from perturbation theory<sup>15</sup> and renormalization group analysis<sup>16,17</sup>. Further, at half-filling, the model has no fermion sign problem, and unbiased determinantal quantum Monte Carlo simulations can be converged<sup>18</sup>. Away from these limits, however, approximations are necessary. Many numerical methods have been applied to the model at both finite and zero temperatures, including fixed-node, constrained path, determinantal, and variational quantum Monte Carlo (QMC)<sup>19–29</sup>, density matrix renormalization group (DMRG)<sup>30–32</sup>, dynamical cluster (DCA)<sup>33,34</sup>, (cluster) dynamical mean-field theories (CDMFT)<sup>35,36</sup>, and variational cluster approximations (VCA)<sup>37,38</sup>. (We refer to DCA/CDMFT/VCA collectively as Green function cluster theories). These pioneering works have suggested rich phenomenology in the phase diagram including metallic, antiferromagnetic,  $d$ -wave (and other kinds of) superconducting phases, a pseudogap regime, and inhomogeneous orders such as stripes, and charge, spin, and pair-density waves, as well as phase separation<sup>6,19,20,24,25,27–29,32,35,39–58</sup>. However, as different numerical methods have yielded different pictures of the ground-state phase diagram, a precise quantitative picture of the ground-state phase diagram has yet to emerge.

It is the goal of this Letter to produce such a quantitative picture as best as possible across the *full* Hubbard model phase diagram below  $U=8$ . Our method of choice is density matrix embedding theory (DMET), which is very accurate in this regime<sup>59–66</sup>, employed together with clusters of up to 16 sites and thermodynamic extrapolation. We carefully calibrate errors in our calculations, giving error bars to quantify the remaining uncertainty in our phase diagram. These error bars also serve, by proxy, to illustrate the relative complexity of the underlying physics for different Hubbard parameters. The accuracy we achieve is significantly higher than attained by earlier comparable Green function cluster calculations for the ground-state. We also carefully estimate the finite size effects, which we find to have a crucial impact on the location of the phase boundaries of the antiferromagnetic and  $d$ -wave SC orders, in contrast to some early ground-state studies<sup>41</sup>.

The one-band (frustrated) Hubbard model on the  $L \times L$  square lattice is

$$H = -t \sum_{\langle ij \rangle \sigma} a_{i\sigma}^\dagger a_{j\sigma} - t' \sum_{\langle\langle ij \rangle\rangle \sigma} a_{i\sigma}^\dagger a_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} \quad (1)$$

where  $\langle \dots \rangle$  and  $\langle\langle \dots \rangle\rangle$  denote nearest and next-nearest neighbors, respectively,  $a_{i\sigma}^\dagger$  destroys (creates) a particle on site  $i$  with spin  $\sigma$ , and  $n_{i\sigma} = a_{i\sigma}^\dagger a_{i\sigma}$  is the number operator. We study the standard Hubbard model with  $t' = 0$  (with  $t = 1$ ), and frustrated models with  $t' = \pm 0.2$ .

DMET is a cluster impurity method which is exact for weak coupling ( $U = 0$ ) and weak hybridization ( $t = 0$ ) and becomes exact for arbitrary  $U$  as the cluster size  $N_c$  increases. It differs from Green function impurity methods such as the DCA or (C)DMFT, as it is a wavefunction method, with a *finite* bath constructed to reproduce the entanglement of the cluster with the remaining lattice sites. DMET has recently been applied and benchmarked in a variety of settings from lattice models<sup>59,61–63</sup> to *ab-initio* quantum chemistry calculations<sup>64,65</sup>, and for ground-state and spectral quantities<sup>66</sup>. In its ground-state formulation, the use of wavefunctions substantially lowers the cost relative to Green function impurity methods, allowing larger clusters to become computationally affordable.

We briefly summarize the method here, with details in SI and Ref.<sup>59,60</sup>. DMET maps the problem of solving for the bulk ground-state  $|\Psi\rangle$  to solving for the ground-state of an impurity model with  $2N_c$  sites consisting of impurity (physical) and bath (auxiliary) sites. The mapping is defined via the Schmidt decomposition<sup>67</sup> of bulk ground-state  $|\Psi\rangle$ , and is exact as long as  $|\Psi\rangle$  is exact. In practice, however, the exact  $|\Psi\rangle$  is, of course, unknown! DMET therefore solves an approximate impurity problem defined from a non-interacting *model* bulk wavefunction  $|\Phi\rangle$ , the ground-state of a quadratic Hamiltonian  $h = h_0 + u$ , where  $h_0$  is hopping part of the Hubbard Hamiltonian, and  $u$  is a one-body operator to be determined. Via  $|\Phi\rangle$  we define an impurity model Hamiltonian and ground-state  $|\Psi'\rangle$  (now an approximation to the exact  $|\Psi\rangle$ ) and from which energies and local observables are measured. The one-body operator  $u$  (analogous to the impurity self-energy in DMFT) is determined self-consistently by matching the one-body density matrix of the impurity wavefunction  $|\Psi'(u)\rangle$ , and the lattice wavefunction  $|\Phi(u)\rangle$ , projected to the impurity model subspace. In this work, we modified the original

TABLE I. Energy of the Hubbard model from various methods. All numbers are extrapolated to the thermodynamic limit. (CP-)AFQMC results are from Zhang<sup>71</sup>. Note that the half-filling results are numerically exact<sup>72</sup>. DMRG results are from White<sup>73</sup>.

U/t	Filling	DMET	AFQMC	CP-AFQMC	DMRG
2	1.0	-1.1764(3)	-1.1763(2)	-	-1.176(2)
4	1.0	-0.8604(3)	-0.8603(2)	-	-0.862(2)
6	1.0	-0.6561(5)	-0.6568(3)	-	-0.658(1)
8	1.0	-0.5234(10)	-0.5247(2)	-	-0.5248(2)
12	1.0	-0.3686(10)	-0.3693(2)	-	-0.3696(3)
4	0.8	-1.108(2)	-	-1.110(3)	-1.1040(14)
4	0.6	-1.1846(5)	-	-1.185(1)	-
4	0.3	-0.8800(3)	-	-0.879(1)	-

DMET procedure slightly. First, we allowed  $u$  to vary over pairing terms, thus allowing  $|\Psi(u)\rangle$  to be a BCS state. Second, we introduced a self-consistent chemical potential to ensure that the filling factor for  $|\Phi\rangle$  and  $|\Psi'\rangle$  exactly match.

To obtain the ground-state phase diagram, we carried out DMET calculations using  $2\times 2$ ,  $4\times 2$ ,  $8\times 2$ , and  $4\times 4$  impurity clusters, cut from a bulk square lattice with  $L = 72$ . We considered  $t' = 0, \pm 0.2$ , and  $U = 2, 4, 6, 8$ , and various densities between  $n = 0.6 - 1$ . The impurity model ground-state  $|\Psi'\rangle$  was determined using a DMRG solver<sup>68</sup> with a maximum number of renormalized states  $M = 2000$ , allowing for  $U(1)$  and  $SU(2)$  spin symmetry breaking. The energy, local moment  $m = \frac{1}{2}(n_{i\uparrow} - n_{i\downarrow})$ , double occupancy  $D = \langle n_{i\uparrow}n_{i\downarrow} \rangle$ , and local  $d$ -wave pairing  $d_{sc} = \frac{1}{\sqrt{2}}(\langle a_{i\uparrow}a_{j\downarrow} \rangle + \langle a_{j\uparrow}a_{i\downarrow} \rangle)$  were measured from  $|\Psi'\rangle$ .

The finite cluster DMET energies and measurements contain 3 sources of error relative to the exact thermodynamic limit. These are from (i) DMET self-consistency, (ii) finite  $M$  in the DMRG solver (only significant for the 16-site impurity clusters), which also induces error in the correlation potential  $u$ , (iii) finite *impurity cluster* size. (The error from the use of a finite  $L = 72$  bulk lattice, is so small as to not affect any of the significant digits presented here). To estimate the thermodynamic result, we (i) estimated DMET self-consistency error by the convergence of expectation values in the last iterations, (ii) extrapolated DMRG energies and observables with the DMRG density matrix truncation error<sup>69</sup>, (iii) estimated the error in  $u$  due to finite  $M$ , by analyzing expectation values from self-consistent  $u(M)$  obtained with different  $M$ , (iv) extrapolated cluster size to infinite size, with the  $N_c^{-1/2}$  scaling<sup>70</sup> appropriate to a non-translationally-invariant impurity. Each of (i) to (iv) gives an estimate of an uncertainty component ( $1\sigma$  for linear extrapolations), which we combined to obtain a single error bar on the DMET thermodynamic estimates. Details of the error estimation and a discussion of the complete data (of which only a fraction is presented here) are given in SI.

We first verify the accuracy of our thermodynamic estimates and error bars by comparing to benchmark data available at half-filling. The data in Table I and Fig. 1(a) shows the high accuracy of the DMET energies at half-filling. The error bars from DMET, AFQMC, and DMRG are all consistent

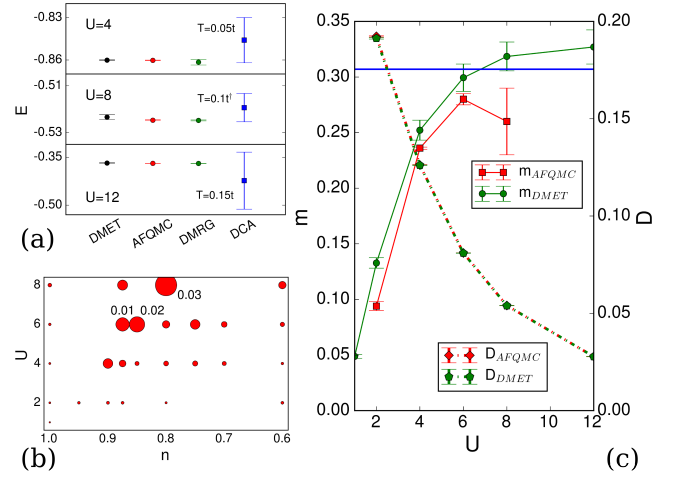


FIG. 1. (color online) Benchmark for  $t' = 0$  Hubbard model. (a) Ground state DMET, AFQMC<sup>71</sup>, DMRG<sup>73</sup> and low temperature DCA<sup>74</sup> energies at half-filling. <sup>†</sup> DCA data at  $U=8$  is from a 50-site finite cluster calculation. (b) DMET Energy uncertainties. The areas of the circles are proportional to the estimated uncertainties. (c) Staggered magnetization ( $m$ ) and double occupancy ( $D$ ) at half-filling.

with an accuracy better than  $0.001t$ . Indeed, the DMET error bars are competitive with the exact "statistical" error bars of AFQMC up to  $U=6$ . As a point of reference, the DMET uncertainty is 1-2 orders of magnitude smaller than finite temperature contributions to recent low-temperature benchmark DCA calculations (Fig. 1(a)), and is similarly 2-3 orders of magnitude smaller than energy errors in earlier zero-temperature Green function cluster calculations<sup>75</sup>.

Figure 1(c) further gives the half-filling staggered magnetization and double occupancies computed with DMET, as compared with AFQMC. The DMET double occupancies are obtained with similar error bars to the AFQMC estimates. The staggered magnetization exhibits larger errors at the smallest  $U = 2$  (a cluster size effect) but for  $U > 4$  appears similarly, or in fact more accurate than the AFQMC result. At the largest value  $U = 12$ , we find  $m = 0.327(15)$ , slightly above the exact Heisenberg value  $m = 0.3070(3)$ <sup>76</sup>.

The half-filling benchmarks lend confidence to the DMET thermodynamic estimates of the energy and observables, and their associated error bars. We therefore use the same error model away from half-filling, in the absence of benchmark data. We can verify our error model by comparing to constrained path (CP) AFQMC, a sign-free QMC with a bias that disappears at low density and moderate  $U$ <sup>23,24</sup>. For  $U = 4$  and  $n \leq 0.6$ , a parameter regime where CP-AFQMC is very accurate, the DMET and CP-AFQMC energies agree to  $0.001t$  (Table I). Fig. 1(b) shows the energy uncertainties across the phase diagram for  $t' = 0$  (The same figure for  $t' = \pm 0.2$  is given in SI). As expected, the accuracy away from half-filling is significantly lower than at half-filling, with the largest errors found in the underdoped region of  $n=0.8-0.9$ . The main source of error is from cluster size extrapolation, especially in the underdoped region, where the solution is sensitive to cluster shapes because of phase boundaries and/or the onset

of competing inhomogeneous orders.

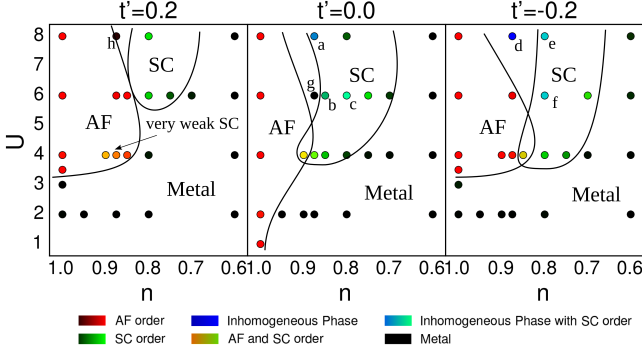


FIG. 2. (color online) Phase diagrams of the Hubbard model. Orders are represented with three primary colors: red (antiferromagnetism), green (d-wave superconductivity) and blue (inhomogeneity), with the brightness proportional to the robustness of the order. The points highlighted with letters: (a) local phase separation; (b) d-wave SC with a slight modulation in the  $(\pi, \pi)$  direction; (c) SC with a weak spin density wave (SDW); (d) a “classic” stripe phase; (e) stripe with pair-density wave (PDW) coexisting with SC; (f) CDW and spin  $\pi$ -phase shift; (g-h) intermediate points between AF and SC where both order parameters extrapolate to zero. Phase boundaries are guides only.

We present the DMET phase diagrams in Fig. 2. Interestingly, they feature many behaviours previously proposed in different studies, although to the best of our knowledge, previous work has generally only seen a subset of the orders that we observe. In particular, we observe (i) an AF phase at half-filling, (ii) a metallic phase at large dopings and at small  $U$ , enhanced by frustration, (iii) a region of d-wave SC order at intermediate dopings and sufficiently large  $U$ , (iv) a region of coexisting AF and SC order, (v) a region rich with inhomogeneous charge, spin, and superconducting orders that are very sensitive to the Hubbard parameters, (vi) points in between the AF and SC phase where the AF and SC orders extrapolate to zero. (The metallic phase is predicted to be unstable at weak coupling and large dopings from weak coupling expansions<sup>17,77</sup>, but the relevant parameter region is outside the scope of this Letter). At  $t' = 0$ , for  $U = 8$ ,  $n = 0.875$ , a SC state with strong inhomogeneity appears which creates large uncertainties in the extrapolated order parameters, thus the precise location of the SC phase boundary at  $U = 8$  is uncertain.

Fig. 3 shows the average AF and d-wave SC order parameters as a function of filling for  $U=4$ . We find that for  $t' = 0$ , the peak in SC order is around  $\langle n \rangle = 0.9$  and SC extends to  $\langle n \rangle \sim 0.8$ . The figures also show that next-nearest-neighbor hopping  $t' = 0.2$  stabilizes AF versus SC, and the reverse is true for  $t' = -0.2$ . The suppression (enhancement) of SC order with positive (negative)  $t'$  is consistent with the stronger superconductivity found in hole-doped materials<sup>78–80</sup>.

The presence of SC in Hubbard model ground-state has previously been much discussed. The strongest SC order found in DMET roughly occurs in the same region as seen in earlier Green function cluster calculations<sup>41,47</sup>. However, this re-

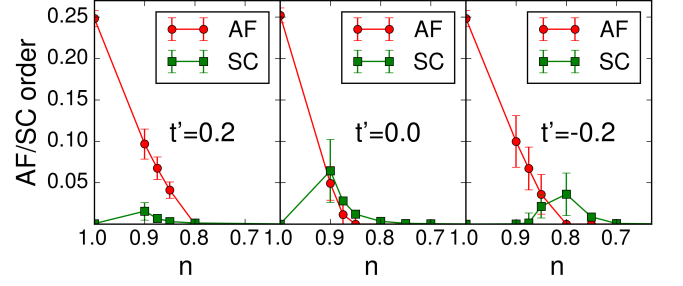


FIG. 3. (color online) Antiferromagnetic (red circle) and (d-wave) superconducting (green square) order parameters at  $U=4$ .

gion is not typically found to be superconducting in ground-state wavefunction calculations using DMRG and AFQMC on finite lattices, even though such calculations achieve significantly higher energy accuracies than the Green function cluster studies<sup>25,32,81,82</sup>. The significance of the DMET result is that the energy error bar in this region (e.g.  $0.001t$  for  $U=4$ ,  $n=0.8$ ,  $t'=-0.2$ ) is comparable to, or better than the accurate ground-state wavefunction calculations, yet SC order is still seen. This strongly suggests that SC is in fact the ground-state order.

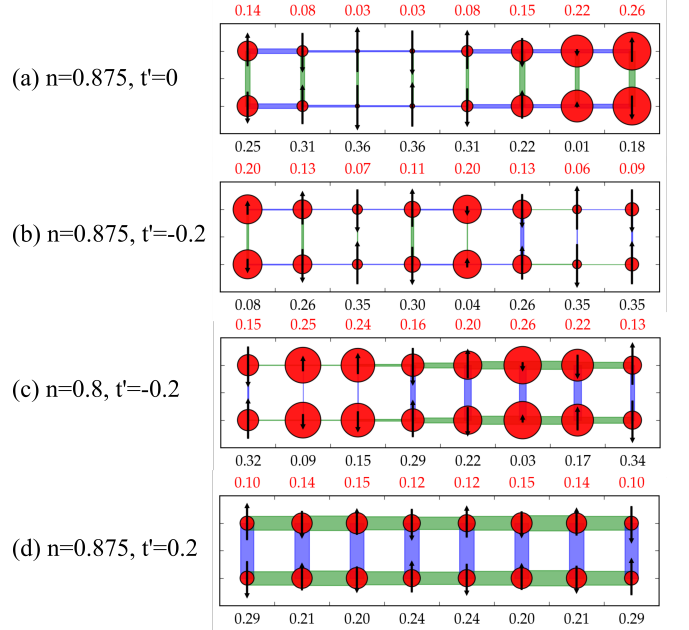


FIG. 4. (color online) Local order parameters in the (frustrated) Hubbard model at selected points at  $U=8$ .

We now further discuss the intermediate region between the AF and SC phases. In this region, a variety of spin-density<sup>25,43,45,46,49,83–85</sup> charge-density<sup>25,86–88</sup>, pair-density wave<sup>88–91</sup>, and stripe orders<sup>30,32,51,52,85,92–95</sup>, have been posited in both the Hubbard model and the simpler  $t$ - $J$  model, with different types of orders seen in different simulation methods. These inhomogeneous phases are proposed to be relevant in the pseudogap physics<sup>89,90,96–100</sup>. Recent projected

entangled pair state (PEPS) studies of the  $t$ - $J$  model and Hubbard model at large  $U \geq 8$  suggest that inhomogeneous and homogeneous states are near degenerate at low doping and can be stabilized with small changes in the model parameters<sup>95,101</sup>. Our work indicates that the Hubbard model behaves similarly. Although we show here in Figs. 4, 5 only the  $8 \times 2$  clusters, it is important to note that the  $8 \times 2$  cluster geometry does not always lead to inhomogeneity, nor are all the  $4 \times 4$  clusters homogeneous. Instead, at points where the tendency towards inhomogeneity is strong, we find a significant lowering of the energy associated with the inhomogeneous order, reflected either in a much lower energy of an inhomogeneous  $8 \times 2$  cluster relative to the  $4 \times 4$  cluster (Fig. 5), or inhomogeneity in *both*  $8 \times 2$  and  $4 \times 4$  clusters. Thus while it is not possible with our cluster sizes to extrapolate details of the inhomogeneities in the thermodynamic limit (for example, the particular wavelengths of the spin, charge, and pairing instabilities, or diagonal versus vertical stripe patterns), the evidence points strongly to some forms of inhomogeneity surviving in the thermodynamic limit at the indicated parts of the phase diagram. Interestingly, the kinds of inhomogeneity we observe are extremely rich, and some representative examples are shown in Fig. 4. These correspond to (i) a local phase separation between a half-filled, antiferromagnetic phase and a superconducting ribbon (Fig. 4(a)), (ii) a classic stripe phase order (Fig. 4(b)) very similar to as seen in earlier DMRG ladder studies<sup>32</sup>. There is also a coexisting weak PDW (exhibiting a sign change across the cell), consistent with earlier stripe proposals<sup>91</sup>. (iii) Inhomogeneities in the pairing order coexisting with the charge and spin orders in, eg., Fig. 4(c), similar to a recent theoretical proposal (see e.g. Ref.<sup>90</sup>). The inhomogeneity is mainly observed with  $t' \leq 0$ , corresponding to the hole-doped cuprates. Fig. 4(d) shows an example at  $1/8$  doping with positive  $t'$ , where the inhomogeneity is much weaker.

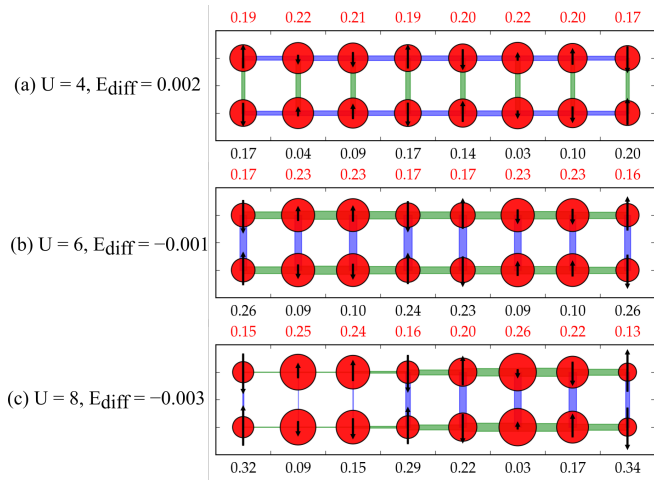


FIG. 5. (color online) Evolution of inhomogeneous patterns and stabilities for  $n = 0.8$ ,  $t' = -0.2$  at various coupling strength.  $E_{\text{diff}}$  is the energy difference between  $8 \times 2$  and  $4 \times 4$  impurity cluster calculations. At  $U = 8$ , both 16-site cluster solutions are inhomogeneous.

To summarize, we have computed a ground-state phase diagram for the Hubbard model on the square lattice using cluster

DMET. The accuracy achieved by DMET appears competitive with the exact ground-state benchmarks available at half-filling, while away from half-filling our error model suggests that the calculations remain very accurate. We observe AF and metallic phases and robust d-wave pairing. Further in parts of the phase space, our calculations strongly suggest that inhomogeneous phases are a feature of the thermodynamic limit, although the precise inhomogeneous patterns require larger clusters to resolve and reflect competition between different orders at very low energy scales. However, for real materials such as the cuprates ( $t \approx 3000\text{K}$ ), the energy resolution achieved here for most of the phase diagram is already below the superconducting  $T_c$ , suggesting that the near degeneracy of competing orders will be lifted by terms beyond those in the Hubbard model, such as long-range charge and hopping terms, multi-orbital effects, and interlayer coupling. Moving beyond the Hubbard model to more realistic material models thus now appears of principal relevance.

## ACKNOWLEDGMENTS

We acknowledge funding from the US Department of Energy, Office of Science, through DE-SC0008624 and DE-SC0010530. This work was also performed as part of the Simons Collaboration on the Many Electron Problem, sponsored by the Simons Foundation. We thank Steven White and Shiwei Zhang for providing unpublished data, and Emanuel Gull for helpful comments. We also thank Sandeep Sharma for discussion on implementing DMRG with broken particle number symmetry. Further discussion of the methodology and results can be found in Appendix A.

## Appendix A: Summary of DMET

Fig. 6 illustrates the computational flow of a DMET calculation. A DMET self-consistency cycle consists of (i) solving for the ground-state of the DMET lattice Hamiltonian, (ii) building the impurity Hamiltonian, (iii) solving for the impurity Hamiltonian ground-state and observables, and (iv) fitting the DMET correlation potential. As discussed in the main text, in this work we allow the DMET solutions to spontaneously break particle number and spin symmetry. We also include a chemical potential in the self-consistency. Here, we explain some general aspects of practical DMET calculations which have not been discussed in detail in the existing literature, as well as describe the technical extensions to broken particle number symmetry, and the self-consistency procedure for the additional chemical potential.

### 1. DMET correlation potential

A general DMET correlation potential is a quadratic operator. It is local in the sense that it does not have cross terms between different images of the impurity cluster on the lattice.

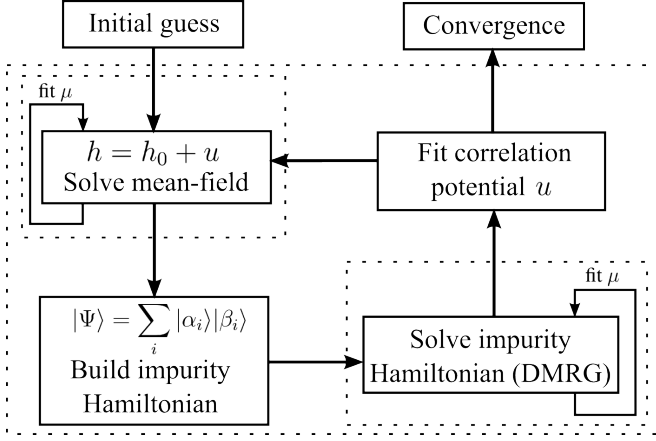


FIG. 6. Graphical representation of the DMET procedure.

In the original DMET paper<sup>59</sup>, it took the form

$$u = \sum_C u_C = \sum_C \sum_{i,j \in C, \sigma} v_{ij\sigma} a_{i\sigma}^\dagger a_{j\sigma} \quad (\text{A1})$$

In Eq. (A1),  $C$  ranges over all impurity cluster supercells within the (large) lattice,  $i, j$  range over sites in the same cluster  $C$ , and  $\sigma \in \{\alpha, \beta\}$  denotes the two flavors of spin. In this form, the correlation potential has  $N_c(N_c + 1)/2$  free parameters (here and later on, we assume real potentials) where  $N_c$  is the number of sites in the impurity cluster. For spontaneously broken particle number and spin symmetry, the correlation potential acquires additional terms,

$$u = \sum_C \sum_{i,j \in C, \sigma} v_{ij,\sigma} a_{i\sigma}^\dagger a_{j\sigma} + \Delta_{ij} a_{i\alpha}^\dagger a_{j\beta}^\dagger + h.c. \quad (\text{A2})$$

In this work, we only allow singlet pairing (strictly speaking,  $S_z = 0$  pairing) but it is straightforward to extend the above to triplet pairing. The normal part  $v$  has two spin components. The pairing term  $\Delta$  has  $N_c^2$  free parameters (it is symmetric when spin symmetry is preserved, but we allow for spin symmetry breaking). In total, the correlation potential  $u$  has  $N_c(2N_c + 1)$  degrees of freedom.

## 2. DMET lattice Hamiltonian

The DMET lattice Hamiltonian (including a chemical potential  $\mu n$ ) is

$$h' = h + u - \mu n = \sum_{ij\sigma} h_{ij\sigma} a_{i\sigma}^\dagger a_{j\sigma} + \Delta_{ij} a_{i\alpha}^\dagger a_{j\beta}^\dagger + c.c. \quad (\text{A3})$$

where  $h = t + v - \mu$  is the normal one-body term from the hopping, correlation potential and chemical potential.  $h'$  can be rewritten in the form of a spin-unrestricted Bogoliubov-de Gennes (BdG)<sup>102,103</sup> equation,

$$\begin{pmatrix} h'_\alpha & \Delta \\ \Delta^\dagger & -h'_\beta \end{pmatrix} \begin{pmatrix} U_\alpha \\ V_\beta \end{pmatrix} = \begin{pmatrix} U_\alpha \\ V_\beta \end{pmatrix} \varepsilon_\alpha \quad (\text{A4})$$

$$\begin{pmatrix} h'_\beta & -\Delta^\dagger \\ -\Delta & -h'_\alpha \end{pmatrix} \begin{pmatrix} U_\beta \\ V_\alpha \end{pmatrix} = \begin{pmatrix} U_\beta \\ V_\alpha \end{pmatrix} \varepsilon_\beta$$

These coupled equations are expressed concisely as

$$\begin{pmatrix} h'_\alpha & \Delta \\ \Delta^\dagger & -h'_\beta \end{pmatrix} \begin{pmatrix} U_\alpha & V_\alpha \\ V_\beta & U_\beta \end{pmatrix} = \begin{pmatrix} U_\alpha & V_\alpha \\ V_\beta & U_\beta \end{pmatrix} \begin{pmatrix} \varepsilon_\alpha & \\ & -\varepsilon_\beta \end{pmatrix} \quad (\text{A5})$$

where  $\varepsilon_\alpha$  and  $\varepsilon_\beta$  are both positive.  $h'$  is diagonalized by transforming to the Bogoliubov quasiparticles,

$$\begin{aligned} c_{i\alpha}^\dagger &= u_{ji}^\alpha a_{j\alpha}^\dagger + v_{ji}^\beta a_{j\beta} \\ c_{i\beta}^\dagger &= u_{ji}^\beta a_{j\beta}^\dagger + v_{ji}^\alpha a_{j\alpha} \end{aligned} \quad (\text{A6})$$

Note that the number of  $\{c_\alpha^\dagger\}$  and  $\{c_\beta^\dagger\}$  quasiparticles will differ if  $S_z \neq 0$  in the physical ground-state.

In terms of the quasiparticles, the lattice Hamiltonian in Eq. (A3) is diagonalized as

$$h' = E_0 + \sum_{i\sigma} \varepsilon_{i\sigma} c_{i\sigma}^\dagger c_{i\sigma} \quad (\text{A7})$$

and the (ground state) quasiparticle vacuum  $|- \rangle$ , defined by  $c_{i\sigma} |- \rangle = 0$ , has energy  $E_0$ . The quasiparticle vacuum is also known as the Bardeen-Cooper-Schrieffer (BCS) ground-state<sup>104</sup>.

## 3. DMET impurity model Schmidt subspace

We now discuss how to define the impurity model Schmidt subspace corresponding to a BCS ground-state of the lattice Hamiltonian in Eq. (A3). To start, we review the “product space” construction of the impurity model Schmidt subspace, starting from the lattice Hamiltonian Slater determinant ground-state, as used in the original DMET<sup>59,60</sup>.

The original DMET impurity model consists of a set of impurity sites augmented by a set of bath modes. In Ref.<sup>60</sup>, the bath modes are defined through the *projected overlap matrix* of the Slater determinant. We compute the projected overlap matrix from the Slater determinant coefficient matrix,

$$C_0 = \begin{pmatrix} M \\ N \end{pmatrix}_{N \times n} \quad (\text{A8})$$

where the rows denote physical sites ( $N$  sites in total), and columns are occupied modes (orbitals). The upper part  $M$  has  $N_c$  rows, which correspond to the  $N_c$  impurity sites. The projected overlap matrix is

$$S = M^T M \quad (\text{A9})$$

From the singular value decomposition (SVD) of  $M$  as  $M = L \Sigma R^T$  (where we use the “full” form of the SVD,  $L$  is  $N_c \times N_c$ ,  $\Sigma$  is  $N_c \times n$ , and  $R$  is  $n \times n$ ) then  $S = R(\Sigma^T \Sigma)R^T$ , i.e.  $R$  is the eigenvector matrix of the projected overlap matrix.  $R$  defines a unitary transformation of the occupied modes in  $C_0$ , giving a new coefficient matrix  $C = C_0 R$ , where

$$C = \begin{pmatrix} L \Sigma \\ N R \end{pmatrix} = \begin{pmatrix} A & 0 \\ B & D \end{pmatrix} \quad (\text{A10})$$



and the second equality follows because  $\Sigma$  is a rectangular matrix of the form  $(\text{diag}(\sigma), \mathbf{0}, \mathbf{0}, \dots, \mathbf{0})$ , where the first  $N_c$  columns constitute a diagonal matrix, and the remaining  $n - N_c$  columns are zero columns. The first  $N_c$  columns of  $C$ ,  $\begin{pmatrix} A \\ B \end{pmatrix}$  define the *embedding modes*, which have non-zero weight on the impurity sites. The matrix  $B$  defines the *bath modes*, which may be orthonormalized using the  $QR$  decomposition,  $B = QR$ . The remaining columns in  $C$  define the *core modes*, which have no weight on the impurity. The Schmidt subspace is then  $\mathcal{F}(a_i^\dagger) \otimes \mathcal{F}(b_i^\dagger) \otimes |e_1 \dots e_{n-N_c}\rangle$ , where  $\{a_i^\dagger\}$  create electrons in the impurity modes,  $\{b_i^\dagger\}$  create electrons in the bath modes (from the columns of  $Q$ ), and  $|e_1 \dots e_{n-N_c}\rangle$  is the core state, defined by the columns of  $D$ . The coefficients defining  $\{a_i^\dagger\}$ ,  $\{b_i^\dagger\}$  can be gathered in the columns of a matrix  $C_1$ ,

$$C_1 = \begin{pmatrix} I_{N_c} & \\ & Q \end{pmatrix} \quad (\text{A11})$$

where  $I_{N_c}$  is an  $N_c \times N_c$  identity matrix.

$A$  and  $B$  can also be obtained directly from the one-particle density matrix. The rotation between  $C$  and  $C_0$  leaves the one-body density matrix invariant, thus

$$\begin{aligned} \rho &= \langle a_i^\dagger a_j \rangle = C_0 C_0^T = C C^T \\ &= \begin{pmatrix} A A^T & A B^T \\ B A^T & B B^T + D D^T \end{pmatrix} \equiv \begin{pmatrix} \rho_{\text{imp}} & \rho_c^T \\ \rho_c & \rho_{\text{env}} \end{pmatrix} \end{aligned} \quad (\text{A12})$$

Defining the eigendecomposition  $\rho_{\text{imp}} = U \Lambda U^T$ , we find

$$A = U \Lambda^{\frac{1}{2}} \text{ and } B = \rho_c (A^T)^{-1} \quad (\text{A13})$$

The above defines the impurity model Schmidt subspace as a tensor product of the impurity site space, and a bath space, thus we refer to it as a “product-space” embedding construction. However, for the BCS state, it is easier to use a slightly different, but equivalent construction. We explain this first for the Slater determinant. Here we build an  $L \times 2N_c$  matrix  $C_2$ , whose columns span the same vector space as  $C_1$  in Eq. (A11), but which does not have the block structure. We start with the “hole” one-particle density matrix

$$\rho_h = \langle a_i a_j^\dagger \rangle = I - \rho^T = I - \rho \quad (\text{A14})$$

We can replace  $\rho$  with  $\rho_h$  in Eqs. (A12) and (A13) and compute an analogous set of coefficients  $A'$  and  $B'$ . Taking  $A$ ,  $B$ , and  $A'$ ,  $B'$  gives  $C_2$ ,

$$C_2 = \begin{pmatrix} A & A' \\ B & B' \end{pmatrix}_{N \times 2N_c} \quad (\text{A15})$$

The  $2N_c$  columns of  $C_2$  span *exactly the same space* as  $C_1$  (proved in the Sec. A 7). Thus, we can equivalently define the Schmidt subspace from the columns of  $C_2$  as we can from  $C_1$ . Transforming to the quasiparticle vacuum of the Slater determinant,  $|- \rangle$ , the columns of  $C_2$  define a set of  $2N_c$  quasipar-

ticle creation operators

$$c_{i\sigma}^\dagger = \sum_{j \in \text{imp}} A_{ji} a_{j\sigma'} + \sum_{j \in \text{env}} B_{ji} a_{j\sigma'} \quad (\text{A16})$$

$$c_{i\bar{\sigma}}^\dagger = \sum_{j \in \text{imp}} A'_{ji} a_{j\sigma}^\dagger + \sum_{j \in \text{env}} B'_{ji} a_{j\sigma}^\dagger \quad (\text{A17})$$

that yields the Schmidt subspace as  $\mathcal{F}(\{\bar{c}_{i(i)}^\dagger\}) \otimes |- \rangle$ . As the impurity model Schmidt subspace here does not (transparently) separate between the impurity sites and environment sites, but rather involves a set of modes which are a linear transformation of both the occupied and virtual modes in the Slater determinant, we refer to this as a “quasiparticle embedding” construction. This provides an alternative view of the DMET embedding as an active space method that uses the embedding quasiparticles defined from  $C_2$  as the active space, while freezing other excitations that involve only the environment.

Extending the quasiparticle embedding construction to BCS states is straightforward. By analogy with the one-particle density matrix of a Slater determinant, we define the generalized one-body density matrix for BCS states,

$$G_\sigma = \begin{pmatrix} U_\sigma \\ V_\sigma \end{pmatrix} \begin{pmatrix} U_\sigma^T & V_\sigma^T \end{pmatrix} = \begin{pmatrix} 1 - \rho_\sigma & \kappa_\sigma \\ \kappa_\sigma^T & \rho_\sigma \end{pmatrix} \quad (\text{A18})$$

where the normal one-particle density matrices  $\rho_\sigma = \langle a_{i\sigma}^\dagger a_{j\sigma} \rangle = V_\sigma V_\sigma^T = 1 - U_\sigma U_\sigma^T$ , and the pairing density matrix  $\kappa = \langle a_{i\alpha} a_{j\beta} \rangle = \kappa_\alpha = -\kappa_\beta^T = U_\alpha V_\beta^T$ . The diagonal of  $G$  is formed by the hole and particle density matrices, and the off-diagonals by the pairing matrix. When the BCS state degenerates to a Slater determinant,  $\kappa = 0$ .

We reorganize the generalized density matrix  $G$  into impurity and environment blocks, placing the impurity (environment) submatrices of  $\rho$  and  $\kappa$  together.

$$G = \begin{pmatrix} G_{\text{imp}} & G_c^T \\ G_c & G_{\text{env}} \end{pmatrix} \quad (\text{A19})$$

For instance, the impurity block is

$$G_{\text{imp}} = \begin{pmatrix} 1 - \rho_\sigma^{\text{imp}} & \kappa_\sigma^{\text{imp}} \\ (\kappa_\sigma^{\text{imp}})^T & \rho_\sigma^{\text{imp}} \end{pmatrix}_{2N_c \times 2N_c} \quad (\text{A20})$$

Then, similar to the treatment in Eq. (A13), we rewrite the impurity part of the generalized density matrix  $G_{\text{imp}} = \bar{A} \bar{A}^T$ , and define a rotated quasiparticle coefficient matrix

$$C_2 = \begin{pmatrix} \bar{U}_{\sigma, \text{imp}} \\ \bar{V}_{\sigma, \text{imp}} \\ \bar{U}_{\sigma, \text{env}} \\ \bar{V}_{\sigma, \text{env}} \end{pmatrix} = \begin{pmatrix} G_{\text{imp}} \\ G_c \end{pmatrix} (\bar{A}^T)^{-1} = \begin{pmatrix} \bar{A} \\ \bar{B} \end{pmatrix}_{2N \times 2N_c} \quad (\text{A21})$$

where  $\bar{B} = G_c (\bar{A}^T)^{-1}$ . Eq. (A21) defines a new set of quasiparticles (with associated quasiparticle creation operators  $\{\bar{c}_{i\sigma}^\dagger\}$ ) in Eq. (A6) through the coefficients  $\bar{U}_\sigma, \bar{V}_\sigma$ . These are a unitary rotation of the original  $2L$  quasiparticles such that only  $2N_c$  of them have non-zero overlap with the impurity. As the rotation does not mix the quasiparticle creation

and annihilation operators, the vacuum of  $\bar{c}_{i\sigma}$  is still the BCS ground state  $|- \rangle$ . In analogy to the embedding for Slater determinants, the Schmidt subspace is now spanned by the embedding quasiparticles,  $\mathcal{F}(\{\bar{c}_{i\sigma}^\dagger\}) \otimes |- \rangle$ .

To connect with Eq. (A15), note that when the BCS state is a Slater determinant,  $G_{\text{imp}}$  and  $G_c$  are both block diagonal, and thus,  $\bar{A} = \text{diag}(A'^\sigma, A^{\bar{\sigma}})$ ,  $\bar{B} = \text{diag}(B'^\sigma, B^{\bar{\sigma}})$ , and Eq. (A21) becomes

$$\bar{U}_\sigma = \begin{pmatrix} A'^\sigma & 0 \\ B'^\sigma & 0 \end{pmatrix} \quad (\text{A22})$$

$$\bar{V}_{\bar{\sigma}} = \begin{pmatrix} 0 & A^{\bar{\sigma}} \\ 0 & B^{\bar{\sigma}} \end{pmatrix} \quad (\text{A23})$$

Combining both sets of spins, the quasiparticles in Eq. (A23) then span exactly the same Hilbert space as the basis defined in Eq. (A15). For general BCS ground states, however,  $\bar{A}$  and  $\bar{B}$  are not block diagonal, and the embedding quasiparticles are mixtures of particles and holes.

The above gives the quasiparticle embedding construction for the BCS state. In the case of the Slater determinant, we started with the equivalent product space embedding, the relation between the two being given by the unitary transformation between the matrix  $C_2$  and  $C_1$ , which has block diagonal form. In the case of the  $C_2$  matrix in Eq. (A21), there does not exist a *unitary* transformation that separates the quasiparticles into  $N_c$  impurity modes and  $N_c$  bath modes. One can, however, use a *Bogoliubov* transformation to separate the impurity and bath degrees of freedom, accompanied by a renormalization of the vacuum. This then gives rise to a “product space” embedding for the BCS state.

To show this, we only need to write the BCS state as a product state, as described in Ref.<sup>105</sup>. We define a simple vacuum  $|\text{vac}\rangle$  as a ferromagnetic state where all  $L$  physical sites are occupied with a spin-down ( $\beta$ ) electron, and let

$$d_{i\alpha}^\dagger = c_{i\beta} \quad (\text{A24})$$

where the  $\{c_{i\sigma}\}$  are the quasiparticle creation operators defined in Eq. (A6).  $\{d_{i\alpha}^\dagger\}$  are correspondingly quasiparticles for the ferromagnetic vacuum  $|\text{vac}\rangle$ , since  $d_{i\alpha}|\text{vac}\rangle = 0$  since  $|\text{vac}\rangle$  is already the lowest eigenstate of  $S_z$ .

The BCS ground state  $|- \rangle$  is the quasiparticle vacuum of  $\{c_{i\sigma}\}$ . This means that it can be written as (up to a phase)

$$|- \rangle = \prod_i d_{i\alpha}^\dagger |\text{vac}\rangle \quad (\text{A25})$$

since  $c_{i\sigma}d_{i\alpha}^\dagger = 0$ . Eq. (A25) rewrites the BCS state into a product state representation (the vacuum  $|\text{vac}\rangle$  is also a product state). We can then take the “occupied” modes ( $\{d_{i\alpha}\}$ ), and use the standard product space construction for Slater determinants to write down the Schmidt decomposition,

$$|- \rangle = \prod_{i \in N_c} (p_i a_{i\alpha}^\dagger + q_i b_{i\alpha}^\dagger) |\text{vac}\rangle_{\text{imp}} \otimes \prod_{j \in L - N_c} \bar{d}_{j\alpha}^\dagger |\text{vac}\rangle_{\text{env}} \quad (\text{A26})$$

where  $\{a_{i\alpha}^\dagger\}$  and  $\{b_{i\alpha}^\dagger\}$  are impurity and bath modes, and  $\{\bar{d}_{j\alpha}^\dagger\}$  are the “core” environment modes. In fact,  $b_{i\alpha}^\dagger$  is simply the (normalized) environment part of  $\bar{c}_{i\beta}$ , and the Schmidt

space is  $\mathcal{F}(\{a_{i\alpha}^\dagger\}) \otimes \mathcal{F}(\{b_{i\alpha}^\dagger\}) \otimes \prod_{j \in L - N_c} \bar{d}_{j\alpha}^\dagger |\text{vac}\rangle_{\text{env}}$ . Since the core wavefunction now becomes  $\prod_{j \in L - N_c} \bar{d}_{j\alpha}^\dagger |\text{vac}\rangle_{\text{env}}$  instead of the BCS ground state  $|- \rangle$ , one has to explicitly include the contributions of the core state in observables.

The quasiparticle embedding and the product state embedding are equivalent theoretically, and in this work we use the quasiparticle approach. However, the resulting embedding modes are delocalized, which increases the entanglement of the impurity model which needs to be captured in the DMRG solver. Thus, for larger clusters, we believe the product space approach may prove favorable from a computational point of view.

#### 4. DMET impurity Hamiltonian and DMRG solver

Once the Schmidt subspace has been defined, the DMET Hamiltonian is formally obtained by projecting an interacting lattice Hamiltonian into the subspace as  $H_{\text{imp}} = PH'P$ , with the many-particle projector defined as

$$P = \sum_{\vec{n}_{i\sigma}} |\Psi_{\vec{n}_{i\sigma}}\rangle \langle \Psi_{\vec{n}_{i\sigma}}|, \quad (\text{A27})$$

where  $\vec{n}_{i\sigma}$  is a vector of occupation numbers of the embedding quasiparticles, and  $|\Psi_{\vec{n}_{i\sigma}}\rangle = \prod_{n_{i\sigma}} (c_{i\sigma}^\dagger)^{n_{i\sigma}} |- \rangle$ . In earlier DMET work, two choices of lattice Hamiltonian were used in the projection: the original interacting lattice Hamiltonian  $H$  (in this case the original Hubbard Hamiltonian) and a modified interacting lattice Hamiltonian  $H'$ , where the interaction term  $U$  is only used in the impurity sites. As in earlier DMET work on lattice models, here we use the latter simpler Anderson-like lattice Hamiltonian  $H'$ . In  $H'$ , on the environment sites (outside of the impurity cluster) the Coulomb interaction  $U$  is replaced by the correlation potential  $u$ , giving

$$H' = h + \sum_{C \neq \text{imp}} u_C + \sum_{i \in \text{imp}} U n_{i\alpha} n_{i\beta} - \mu n \quad (\text{A28})$$

The projection defined in Eq. (A27) reduces to transforming  $\{a_{i\sigma}^\dagger\}$  to the embedding quasiparticle basis using the inverse Bogoliubov transformation,

$$a_{i\sigma}^\dagger = u_{ij}^\sigma c_{j\sigma}^\dagger + v_{ij}^\sigma c_{j\bar{\sigma}} \quad (\text{A29})$$

and replacing the pure environment quasiparticle operators with their expectation values with the BCS ground state  $|- \rangle$ .

After projection, we can write  $H_{\text{imp}}$  as a sum of one- and two-particle parts,  $H_{\text{imp}} = h_{\text{imp}} + V_{\text{imp}}$ , where  $h_{\text{imp}}$  is

$$h_{\text{imp}} = \bar{h}_{ij}^\sigma c_{i\sigma}^\dagger c_{j\sigma} + \bar{\Delta}_{ij} c_{i\alpha}^\dagger c_{j\beta}^\dagger + c.c. + E_0 \quad (\text{A30})$$

and  $c_{i\sigma}^{(\dagger)}$  here denote the embedding quasiparticles. In terms of the Bogoliubov coefficients of the embedding quasiparticles



$U_\sigma, V_\sigma$ , the components of  $h_{\text{imp}}$  are defined as

$$\begin{aligned}\bar{h}^\alpha &= U_\alpha^T h_\alpha U_\alpha - V_\beta^T h_\beta V_\beta + U_\alpha^T \Delta V_\beta + V_\beta^T \Delta^T U_\alpha \\ \bar{h}^\beta &= U_\beta^T h_\beta U_\beta - V_\alpha^T h_\alpha V_\alpha - U_\beta^T \Delta^T V_\alpha - V_\alpha^T \Delta U_\beta \\ \bar{\Delta} &= U_\alpha^T \Delta U_\beta + V_\beta^T \Delta^T V_\alpha - V_\beta^T h_\beta U_\beta + U_\alpha^T h_\alpha V_\alpha \\ E_0 &= \text{Tr}(V_\beta^T h_\beta V_\beta + V_\alpha^T h_\alpha V_\alpha + V_\alpha^T \Delta U_\beta + U_\beta^T \Delta^T V_\alpha)\end{aligned}\quad (\text{A31})$$

where  $h_\sigma$  and  $\Delta$  are the one-particle and pairing terms in the lattice Hamiltonian respectively. The two-particle part  $V$  contains many contributions due to the breaking of particle number symmetry in the quasiparticle formulation. These have the form

$$\begin{aligned}V_{\text{imp}} &= \frac{1}{2} \sum_{pqsr, \sigma\mu} w_{pqsr, \sigma\mu} c_{p\sigma}^\dagger c_{q\mu}^\dagger c_{s\mu} c_{r\sigma} \\ &+ \sum_{pq, \sigma} h_{pq, \sigma} c_{p\sigma}^\dagger c_{q\sigma} + E_1 \\ &+ \frac{1}{4} \sum_{pqsr} x_{pqsr} c_{p\alpha}^\dagger c_{q\alpha}^\dagger c_{s\beta}^\dagger c_{r\beta} \\ &+ \frac{1}{2} \sum_{pqsr, \sigma} \tilde{v}_{pqsr, \sigma} c_{p\sigma}^\dagger c_{q\sigma}^\dagger c_{s\bar{\sigma}}^\dagger c_{r\sigma} \\ &+ \sum_{pq} \Delta_{pq} c_{p\alpha}^\dagger c_{q\beta}^\dagger + \text{c.c.}\end{aligned}\quad (\text{A32})$$

$V_{\text{imp}}$  connects  $N$  particle states with  $N, N \pm 2, N \pm 4$  states. For brevity, we do not give the formulae for the coefficients explicitly (which are obtained by simple algebra from Eq. (A29)). The scalar and one-particle terms in  $V_{\text{imp}}$  contain contributions from pure environment quasiparticles, and can be absorbed into  $h_{\text{imp}}$ .

We have adapted our quantum chemistry density matrix renormalization group (DMRG) code BLOCK<sup>106–108</sup> to break  $U(1)$  particle number symmetry and to incorporate the Hamiltonian terms in Eq. (A30) and (A32). While the full wavefunction is not restricted to  $U(1)$  symmetry, the particle quantum number is still used in the calculations in the sense that the renormalized states are required to carry a definite particle number. This allows us to use the block-sparsity of the Hamiltonian to tackle larger numbers of renormalized states.

The quasiparticle basis associated with  $c_{i\sigma}^{(\dagger)}$  is not localized to a site, thus we use a localization and ordering procedure as used in quantum chemistry DMRG calculations to reduce long-range entanglement between the embedding quasiparticles. We define a center for each quasiparticle, and minimize the metric

$$d = \sum_i \langle [p(r_i) - p(r_{i0})]^2 \rangle \quad (\text{A33})$$

where  $r_{i0}$  is the center for  $c_{i\sigma}^{(\dagger)}$ . The centers are assigned to the impurity sites, with two centers per site. A simple mapping  $p(r)$  deforms the lattice (Fig. 7) to prioritize the localization of the sites within and nearby the impurity cluster, which are more entangled because of the interaction. The localized

quasiparticles are then reordered according to the position of their centers  $r_{i0}$ . We find that the localization and reordering significantly reduce the DMRG truncation error, by up to a factor of 10.

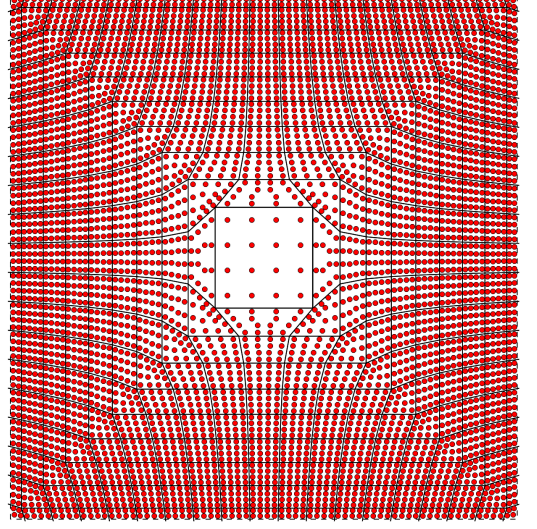


FIG. 7. The effect of the deform function  $p(r)$  used in the quasiparticle localization procedure. The figure shows a deformed  $72 \times 72$  (periodic) lattice with a  $4 \times 4$  impurity cluster.

## 5. Expectation values

As discussed in the original papers on DMET<sup>60</sup>, the DMET energy of  $H_{\text{imp}}$  defined in Eq. (A30) does not correspond to the ground-state energy of the impurity cluster. This is because the impurity Hamiltonian contains three types of energy contributions: pure impurity, impurity-bath interactions, and pure bath (environment) parts. The proper DMET energy should exclude the pure environment contributions and include only part of the impurity-bath interaction energy. Therefore, the DMET energy is evaluated as a *partial* trace of the one- and two-particle reduced density matrices of the impurity wavefunction. This partial trace can be equivalently implemented as a full trace, with appropriate scaling factors for terms in the Hamiltonian which couple the impurity and environment. For each class of term in the Hamiltonian, this scaling factor is given by the number of indices in the impurity, divided by the total number of indices. (For example, for the one-particle terms in the Hamiltonian, the contribution of the impurity-bath block to the total trace is scaled by a factor of  $\frac{1}{2}$ ).

An equivalent formulation for the Hubbard Hamiltonian (which contains no long-range Coulomb terms) is to evaluate the two-particle part of DMET energy as

$$E_2 = \langle \Psi | V_{\text{imp}} | \Psi \rangle = E_{\text{DMRG}} - \langle \Psi | h_{\text{imp}} | \Psi \rangle, \quad (\text{A34})$$

where  $|\Psi\rangle$  is the DMRG ground state. Since  $h_{\text{imp}}$  is a quadratic operator,  $E_2$  can be computed only with knowledge

of the DMRG energy and the one-particle (and pairing) density matrix, avoiding explicitly evaluating  $\langle \Psi | V_{\text{imp}} | \Psi \rangle$  through the two-particle density matrix.

The local spin moments and pairing are both one-particle quantities. We therefore obtain them from the one-particle and pairing density matrix  $\rho = \langle c_{i\sigma}^\dagger c_{j\sigma} \rangle$ ,  $\kappa = \langle c_{i\alpha} c_{j\beta} \rangle$  of the DMRG wavefunction  $|\Psi\rangle$ , transformed back to the lattice site basis  $\{a_{i\sigma}^{(\dagger)}\}$  using Eq. (A29). Note that  $\rho$  and  $\kappa$  are defined not only for quasiparticles inside the impurity Schmidt subspace, but also for core quasiparticles. (In the quasiparticle approach, although  $\rho$  and  $\kappa$  are themselves zero in the core, terms such as  $c_i c_j^\dagger$  can appear in the expansion using Eq. (A29) and result in non-zero expectation values). If one is interested only in impurity cluster expectation values, or for DMET lattice Hamiltonians without broken symmetry, the contribution of the core quasiparticles is strictly zero and may thus be omitted. However, for ordered (e.g. magnetic or superconducting) states, the core contribution does not vanish and therefore *cannot* be neglected. Doing so would produce for example, the strange result of vanishing long-range correlations even in a long-range ordered DMET state.

In this study, when a single value of the order parameter is given, it is computed using the  $2 \times 2$  plaquette at the center of the impurity cluster, to minimize the boundary effects. The antiferromagnetic order parameter is defined as

$$m = \frac{1}{4}(m_{0,0} + m_{1,1} - m_{0,1} - m_{1,0}) \quad (\text{A35})$$

and the d-wave parameter as

$$d = \frac{1}{4}[d_{(0,0),(0,1)} + d_{(1,0),(1,1)} - d_{(0,0),(1,0)} - d_{(0,1),(1,1)}] \quad (\text{A36})$$

where  $m_i = \frac{1}{2}(n_{i\alpha} - n_{i\beta})$  and  $d_{ij} = \frac{1}{\sqrt{2}}(\langle a_{i\alpha} a_{j\beta} \rangle + \langle a_{j\alpha} a_{i\beta} \rangle)$  as defined in the main text. At some points in the phase diagram there are also inhomogeneous states. When the inhomogeneity is strong, we report here the full distribution of local order parameters.

## 6. DMET self-consistency

The DMET embedding constructs the impurity model via the model ground-state of the DMET lattice Hamiltonian, however, this state (and the lattice Hamiltonian) are functions of the correlation potential  $u$ .  $u$  is determined by the self-consistency procedure, which aims to minimize the difference between the embedding wavefunction and the DMET mean-field wavefunction, as measured by their (generalized) one-particle density matrix difference. In the quasiparticle embedding space, the one-particle and pairing density matrices of the mean-field wavefunction  $|\Phi\rangle$  are simply zero. Conceptually, the simplest technique is to define  $u$  to minimize the Frobenius norm,

$$\begin{aligned} \min_u \|G_{\Psi(u)} - G_{\Phi(u)}\|_F \\ = \min_u \sum_{ij} (|\rho_{\Psi,ij}^\alpha|^2 + |\rho_{\Psi,ij}^\beta|^2 + 2|\kappa_{\Psi,ij}|^2) \end{aligned} \quad (\text{A37})$$

However, as the derivative of the correlated wavefunction  $|\Psi\rangle$  with respect to  $u$  is expensive, the above is solved in a two-step procedure consisting of an inner and outer loop. In the inner loop, we carry out  $\min_u \|G_{\Psi} - G_{\Phi(u)}\|_F$ , i.e. the correlated wavefunction is held fixed, while in the outer loop, the updated  $u$  leads to a new impurity model, and a new correlated wavefunction  $\Psi$ . Since the correlation potential is restricted to the impurity, while the generalized density matrices range over the whole embedded system (impurity+bath), the Frobenius norm does not vanish and is minimized in the least squares sense in Eq. (A37). Other choices of self-consistency condition, where the metric vanishes, can also be formulated but are not used here.

If the total particle number  $n$  is allowed to fluctuate, as in a superconducting state, then one of the conjugate pairs (chemical potential)  $\mu$  or (particle density)  $\langle n \rangle$  must be fixed. We usually want to express the observables as a function of doping, or occupation, thus we fix  $\langle n \rangle$  and determine the appropriate  $\mu$ . Since the diagonal elements of the correlation potential and chemical potential appear redundant, how can one determine the chemical potential? Formally, at the DMET mean-field level (Eq. (A3)), there is a gauge freedom between  $u$  and  $\mu$ , namely

$$\mu' = \mu + \phi, u' = u + \phi \sum_{i\sigma} a_{i\sigma}^\dagger a_{i\sigma} \quad (\text{A38})$$

however, this gauge freedom is lost at the embedding stage (Eq. (A28)), because  $u$  is only added to the environment (sites outside of the impurity) while  $\mu$  affects *every* site in the lattice, including the impurity. This difference allows us to use the two-step self-consistency scheme to determine  $\mu$ , as shown in Fig. 6. Specifically, we first fit  $\mu$  at the mean-field stage, to ensure  $\langle n \rangle$  is correct. Then at the embedding stage, we vary  $\mu$  and  $u$  simultaneously, following Eq. (A38). This means that the DMET mean-field solution (and thus definition of the impurity model) stays the same, but the relative energy levels of the impurity change as compared to the bath, which allows us to adjust the filling on the impurity.

Fitting at the embedding stage means we need to solve the correlated impurity problem more than once in a single DMET self-consistency iteration. This increases the computational cost. Our strategy is to allow only one iteration of chemical potential fitting in each DMET iteration, corresponding to at most three DMRG calculations. Because fitting  $\mu$  is a one dimensional search, even with this crude approach, we can usually control the relative deviation of  $\langle n \rangle$  to less than  $10^{-4}$ .

## 7. Proofs

Here we prove the equivalence of the Fock spaces spanned by  $C_1$  and  $C_2$  in the construction of the impurity Schmidt subspace, as defined in section (A 3). Precisely, we need to prove

1.  $C_2$  is orthonormal,  $C_2^T C_2 = I$ . (It is easy to see  $C_1$  is orthonormal, because  $Q$  is a unitary matrix from QR decomposition).

2.  $C_2 = C_1 V$ , which is equivalent to  $C_1^T C_2 = V$ , where  $V$  is unitary.

To prove (1)  $C_2^T C_2 = I$ , we need the idempotency of density matrices  $\rho^2 = \rho$ . Considering only the upper-left block of  $\rho$ , we have

$$\rho_{\text{imp}}^2 + \rho_c^T \rho_c = \rho_{\text{imp}} \quad (\text{A39})$$

$$\begin{aligned} C_2^T C_2 &= \begin{pmatrix} A^T & A^{-1} \rho_c^T \\ A'^T & -(A')^{-1} \rho_c^T \end{pmatrix} \begin{pmatrix} A & A' \\ \rho_c (A^T)^{-1} & -\rho_c (A'^T)^{-1} \end{pmatrix} = \begin{pmatrix} A^T A + A^{-1} \rho_c^T \rho_c (A^T)^{-1} & A^T A' - A^{-1} \rho_c^T \rho_c (A'^T)^{-1} \\ A'^T A - (A')^{-1} \rho_c^T \rho_c (A^T)^{-1} & A'^T A' + (A')^{-1} \rho_c^T \rho_c (A'^T)^{-1} \end{pmatrix} \\ &= \begin{pmatrix} \Lambda + \Lambda^{-\frac{1}{2}} \Lambda (I - \Lambda) \Lambda^{-\frac{1}{2}} & \Lambda^{\frac{1}{2}} (I - \Lambda)^{\frac{1}{2}} - \Lambda^{-\frac{1}{2}} \Lambda (I - \Lambda) (I - \Lambda)^{-\frac{1}{2}} \\ (I - \Lambda)^{\frac{1}{2}} \Lambda^{\frac{1}{2}} - (I - \Lambda)^{-\frac{1}{2}} \Lambda (I - \Lambda) \Lambda^{-\frac{1}{2}} & I - \Lambda + (I - \Lambda)^{-\frac{1}{2}} \Lambda (I - \Lambda) (I - \Lambda)^{-\frac{1}{2}} \end{pmatrix} = I \end{aligned} \quad (\text{A40})$$

For (2), since

$$V = C_1^T C_2 = \begin{pmatrix} A & A' \\ Q^T \rho_c (A^T)^{-1} & -Q^T \rho_c (A'^T)^{-1} \end{pmatrix} \quad (\text{A41})$$

we have

$$\begin{aligned} V V^T &= \begin{pmatrix} A & A' \\ Q^T \rho_c (A^T)^{-1} & -Q^T \rho_c (A'^T)^{-1} \end{pmatrix} \begin{pmatrix} A^T & A^{-1} \rho_c^T Q \\ A'^T & -A'^{-1} \rho_c^T Q \end{pmatrix} \\ &= \begin{pmatrix} A A^T + A' A'^T & \rho_c^T Q - \rho_c^T Q \\ Q^T \rho_c - Q^T \rho_c & Q^T \rho_c (A A^T)^{-1} \rho_c^T Q + Q^T \rho_c (A' A'^T)^{-1} \rho_c^T Q \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & R[A^{-1} + (I - A)^{-1}] R^T \end{pmatrix} \end{aligned} \quad (\text{A42})$$

In the bottom-right block

$$\begin{aligned} A^{-1} + (I - A)^{-1} &= U \Lambda^{-1} U^T + U (I - \Lambda)^{-1} U^T \\ &= U \Lambda^{-1} (I - \Lambda)^{-1} U^T \\ &= [A(I - A)]^{-1} \\ &= (B^T B)^{-1} = (R^T R)^{-1} = R^{-1} (R^T)^{-1} \end{aligned} \quad (\text{A43})$$

So  $V V^T = I$ . Here we assume  $R$  is invertible, which is true if and only if we have the full set of  $N_c$  bath orbitals coupled to the impurity. This is generally true in lattice settings where the impurity and the environment are strongly coupled. Sometimes the bath can be smaller than the impurity in molecules and when we use a large basis set, and in these cases, special treatment is needed.

## Appendix B: Error Model

As described in the main text, we consider 3 sources of error: (i) errors in DMET self-consistency, (ii) finite  $M$  in the DMRG solver, and (iii) finite impurity cluster size. The DMET self-consistency error is estimated as  $\frac{1}{2} |E^{(n-1)} - E^{(n)}|$ , where  $E^{(n)}$  and  $E^{(n-1)}$  are the energies of the last two DMET self-consistency iterations. A typical DMET calculation oscillates between two slightly different solutions with

From Eq. (A13) and (A14), we know  $A' = U(I - \Lambda)^{\frac{1}{2}}$ ,  $B' = -\rho_c (A'^T)^{-1}$ . Therefore

the magnitude of the oscillations decreasing with the number of iterations. We use the range of oscillation as a representation of the self-consistency error. The error distributions across the range of calculations in this work are shown in Fig. 8, with the average values in the inset. For most points in the phase diagram, and for all cluster sizes, the self-consistency error is less than  $0.0005t$ . For  $4 \times 4$  clusters DMET calculations are the harder to converge, due to larger error in the embedded calculations, giving a largest error of up to  $0.002t$ , and an average self-consistency error approximately twice as large as that for the other cluster shapes.

For impurity clusters larger than the  $2 \times 2$  cluster (where our DMRG solver is not exact), there is error due to using finite  $M$  in the DMRG impurity solver. The error due to finite  $M$  has two components:

1. variational error in the DMRG calculation, which is usually assumed proportional to the density matrix truncation weight  $\delta_w$ ,
2. the DMET correlation potential error  $\delta_u$ , as  $\delta_u$  is a function of the impurity density matrices, and these have an error for finite  $M$ .

For the  $4 \times 2$  and  $8 \times 2$  clusters,  $\delta_u$  appears negligible. For these clusters, we carry out the DMET self-consistency with lower  $M$  to obtain the DMET correlation potential  $u$ , then do a few final DMRG calculations at large  $M$  to extrapolate to the  $M \rightarrow \infty$  exact solver limit. For  $4 \times 4$  clusters, the  $U = 2$  data

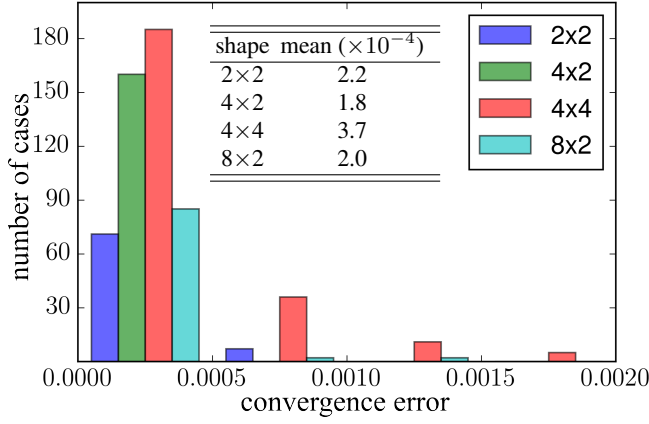


FIG. 8. Distribution and average value (inset) of the DMET self-consistency error in the energy (units of  $t$ ) for each cluster size.

is processed in this way as well. However, for other values of  $U$  using the  $4 \times 4$  clusters, the DMRG truncation weight is as large as  $10^{-3}$  for low to intermediate doping with our accessible  $M$ , thus making the contribution of  $\delta_u$  also significant. To compensate for this, we first carry out the DMET self-consistency with a series of different  $M$  up to 1200, and linearly extrapolate the energy to the  $M = \infty$  limit,  $E_1$ . This thus extrapolates errors from *both* source 1 and 2, assuming  $\delta_u \propto \delta_w$ . Another further set of DMRG calculations are then done with  $M$  up to 2000, using the converged correlation potential from the DMET self-consistency with the largest  $M$ . This second set of results are then extrapolated again against the truncated weight to obtain an energy  $E_2$ , which only accounts for the error from source 1. Although the linear relation between the source 2 error and the truncation weight need not hold in general, in practice, we find that  $\delta_u = \frac{1}{2}|E_1 - E_2|$  gives a crude estimate of  $\delta_u$ . Therefore, we report the  $4 \times 4$  cluster energy as  $E_{4 \times 4} = \frac{1}{2}(E_1 + E_2)$ , with a final uncertainty of  $\delta E_{4 \times 4}^2 = \delta_u^2 + \delta E_1^2 + \delta E_2^2$ , where  $\delta E_1$  is a combination of the linear regression uncertainty and the uncertainties of the original data points (from DMET self-consistency error), while  $E_2$  does not have any self-consistency error. Fig. 9 illustrates the set of computations and linear extrapolations performed with each  $4 \times 4$  cluster to obtain the  $4 \times 4$  cluster energy and error estimate.

After obtaining the energy and observables for each cluster size, we extrapolate to the thermodynamic limit using the relation  $\Delta E_{N_c} \propto N_c^{-1/2}$ . Since both the  $4 \times 4$  and  $2 \times 8$  clusters are 16 site clusters, we must choose which one to use in the extrapolation to the thermodynamic limit. We believe that  $4 \times 4$  clusters have less finite size error than the  $8 \times 2$  clusters, and thus we generally use these in the extrapolation. However, at certain points in the phase diagram (e.g. at strong coupling, or negative  $t'$ ) there is a strong tendency to inhomogeneity, and the  $4 \times 4$  clusters cannot necessarily accommodate the new order parameter, resulting in a much higher energy than for the  $8 \times 2$  cluster. In such cases, namely, when (a)  $4 \times 4$  and  $8 \times 2$  clusters show different orders, and (b) the  $8 \times 2$  cluster is lower in energy, we use the  $8 \times 2$  cluster energy for

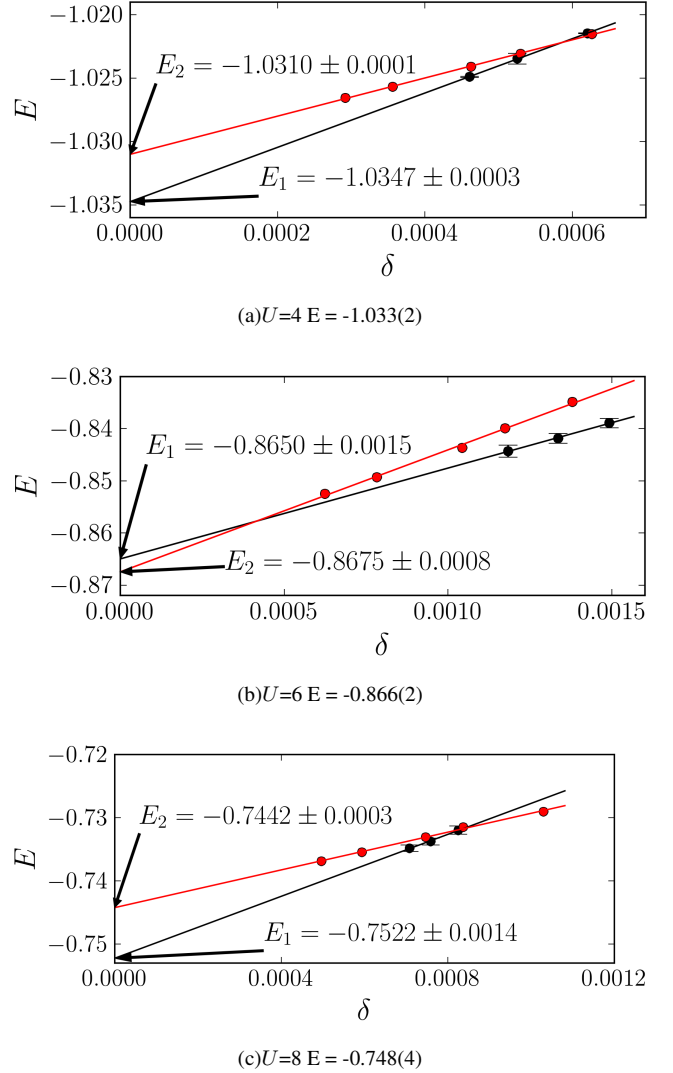


FIG. 9. Computations involved in the estimate of the  $4 \times 4$  cluster DMET energy. The black dots (with error bars for the self-consistency error) are DMET self-consistent results using different DMRG  $M$ . These points are extrapolated to obtain  $E_1$ . The red dots are DMRG results using the "best" self-consistent correlation potential, which are then extrapolated to obtain  $E_2$ . The final  $4 \times 4$  cluster DMET energies are reported as  $E = \frac{1}{2}(E_1 + E_2)$ . The plots are shown for  $t' = 0$ ,  $n = 0.875$  and (a)  $U=4$  (b)  $U=6$  (c)  $U=8$ .

the extrapolation.

The cluster size extrapolation works surprisingly well given the limited number and small sizes of the clusters, although it contributes the main source of error in the final uncertainty. In Fig. 10 we show some of the extrapolation results at  $U = 4$ . At half-filling and in the overdoped region ( $n < 0.8$ ), the linear relation used in the cluster size extrapolation appears quite good even for these small clusters. In the underdoped region, however, the energy is more strongly dependent on the cluster shape, often because the system has a strong tendency to establish an inhomogeneous phase. In Fig. 11, we plot the local order parameters at  $n = 0.875$ , where the  $8 \times 2$  clus-

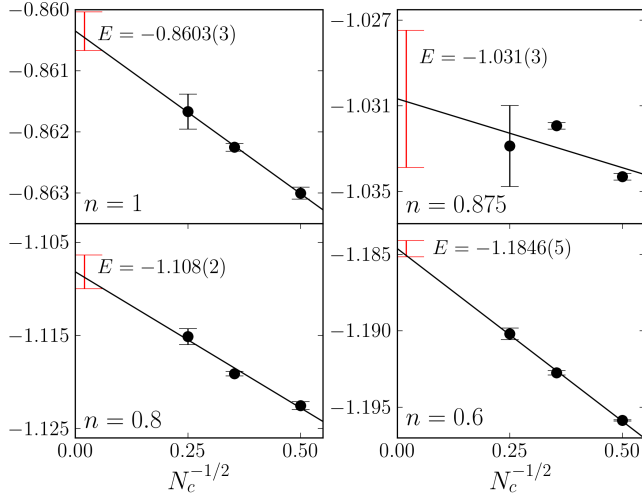


FIG. 10. Cluster size extrapolation for  $U = 4, t' = 0$  at various fillings. The black dots are finite size results. The red error bars are the confidence intervals for the thermodynamic limit.

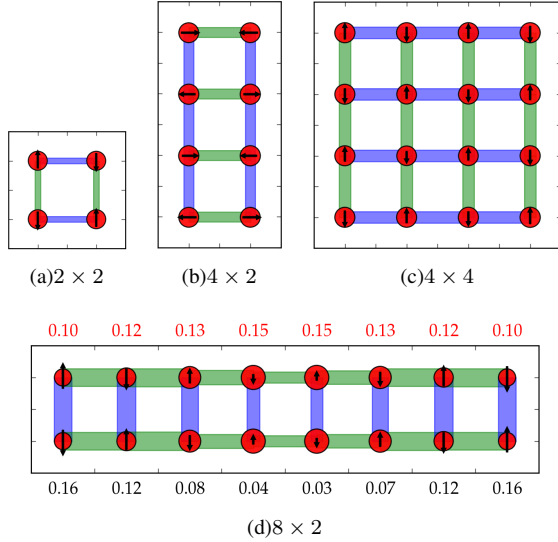


FIG. 11. Local order parameters for  $U = 4, n = 0.875, t' = 0$ . The legend is the same as for Fig. 4 in the main text.

ter calculation gives an incommensurate antiferromagnetic order. Although the  $8 \times 2$  cluster energy ( $-1.0288$ ) is slightly higher than the  $4 \times 4$  cluster result ( $-1.033$ ), its inhomogeneity suggests the existence of a low-lying inhomogeneous state that can be (relatively) stabilized by special cluster shapes. Nonetheless, even in the underdoped region, the error model appears to give a reliable estimate of the energy at the thermodynamic limit, albeit with a large uncertainty.

Fig. 12 shows the final energy errors for  $t' = \pm 0.2$  across the phase diagram. The same plot for  $t' = 0$  is shown in Fig. 1(b) in the main text. The overall uncertainty for  $t' = 0.2$  is smaller than for  $t' = 0$  (see Fig. 1(b) in the main text) and  $t' = -0.2$ , as is the maximum uncertainty ( $0.01t$  compared to

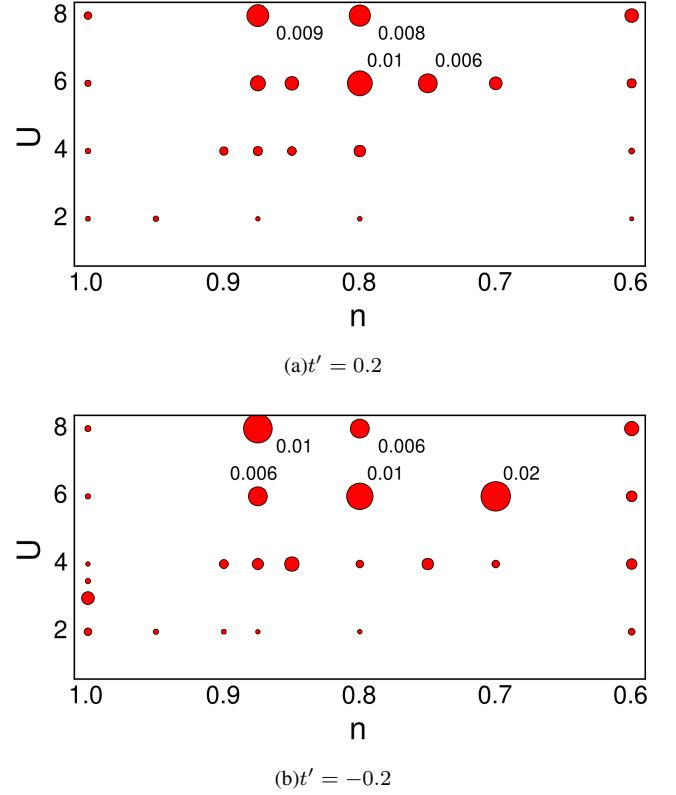


FIG. 12. DMET energy uncertainty plot for the frustrated Hubbard model with  $t' = \pm 0.2$ . Refer to Fig. 1(b) in the main text for the legend.

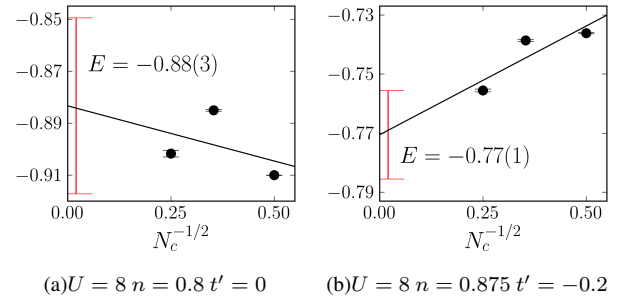


FIG. 13. Examples of thermodynamic extrapolations where the energy is sensitive to cluster shape.

$0.03t$  and  $0.02t$ , respectively). As mentioned before, the main source of error is the cluster size extrapolation. Two examples of large uncertainties due to cluster size (and shape) effect are shown in Fig. 13. The largest uncertainties are observed at  $U = 6$  and moderate doping.

### Appendix C: Further results

In this section, we will expand on the determination of the phase diagram (Fig. 2 in the main text).

The staggered magnetization for the frustrated Hubbard

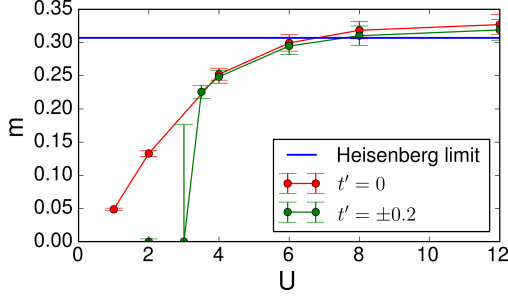


FIG. 14. Staggered magnetization ( $m$ ) of the half-filled Hubbard model for  $t' = \pm 0.2$  and  $t' = 0$ .

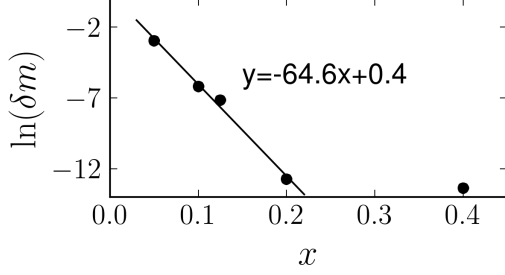


FIG. 15. At  $U = 2$ , the uncertainty of antiferromagnetic order parameter decreases exponentially with doping. The exponent is  $65 \pm 4$ .

model at half-filling (compared to the  $t' = 0$  model and the Heisenberg limit) is shown in Fig. 14. Due to particle-hole symmetry, the plot is identical for  $t' = \pm 0.2$ . The onset of antiferromagnetism is at finite  $U$  in the frustrated model, between  $U = 2$  and  $3.5$ , consistent with previous quantum Monte Carlo simulations<sup>109</sup>. The large error bar at  $U = 3$  indicates the sensitivity to impurity cluster sizes near the phase boundary, resulting in a large uncertainty in the thermodynamic extrapolation.

At weak coupling  $U = 2$ , we find that the antiferromagnetism (in the non-frustrated model) is destroyed already at small doping  $x = 0.05$ , where the staggered magnetization is  $m = 0.00 \pm 0.05$ . Although the expectation value is 0, the relatively large uncertainty  $\delta m$  reflects that short-range spin fluctuations are still significant, although long-range order does not exist. As we increase doping,  $\delta m$  decreases exponentially (Fig. 15). At  $U = 2$ , we do not find d-wave superconductivity, to within numerical precision.

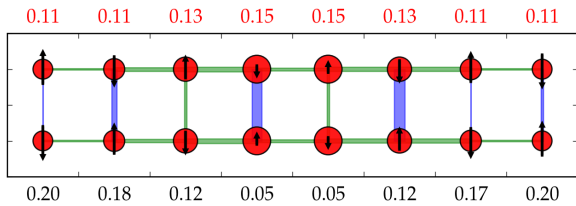


FIG. 16. Inhomogeneous order from  $8 \times 2$  cluster calculations at  $U = 4$ ,  $t' = -0.2$   $n = 0.875$ .

TABLE II. Energy comparison for different 16-site impurity clusters at  $U = 4$  and  $t' = -0.2$ .

$n$	$E_{8 \times 2}$	$E_{4 \times 4}$
0.8	-1.10483(6)	-1.0507(4)
0.85	-1.0162(1)	-1.020(2)
0.875	-0.9966(1)	-0.9989(7)

We now discuss  $U = 4$ . We have already shown the order parameters, and the observed thermodynamic extrapolated ground state orders are all homogeneous. However, for  $t' = -0.2$ , the  $8 \times 2$  cluster calculations result in an inhomogeneous state at doping  $n = 0.8 - 0.875$ , although the energy is significantly higher than obtained with the  $4 \times 4$  clusters at the same fillings. An example of inhomogeneous patterns is shown in Fig. 16, where one can see a pair density wave and incommensurate magnetic order. In Table II, we compare the energies between the  $8 \times 2$  cluster and  $4 \times 4$  cluster results at relevant points in the phase diagram for  $U = 4$ . In all these cases, the  $8 \times 2$  cluster has a higher energy, suggesting that the ground state at  $U = 4$  is homogeneous, or inhomogeneous with a very long wavelength that does not fit in our cluster shapes.

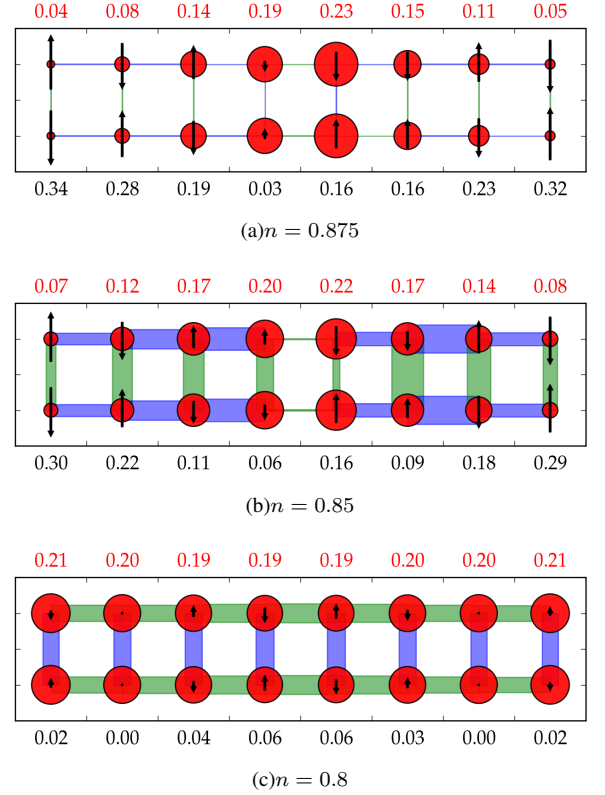


FIG. 17. Inhomogeneous order from  $8 \times 2$  cluster calculations at  $U = 6$  and  $t' = 0$  with fillings 0.875 to 0.8.

At  $U = 6$ , more interesting inhomogeneous orders start to appear. At  $t' = 0$ ,  $8 \times 2$  clusters result in various orders (Fig. 17). At both  $n = 0.875$  and  $n = 0.85$ ,  $4 \times 4$  clusters are



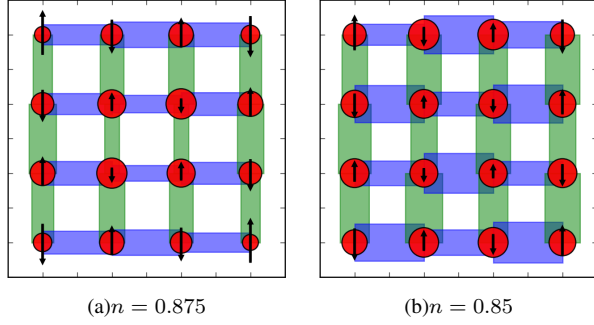


FIG. 18. Local order parameters from  $4 \times 4$  cluster calculations at  $U = 6$  and  $t' = 0$  with fillings 0.875 and 0.85.

significantly lower in energy, suggesting the charge, spin and pairing orders shown in Fig. 17(a) and 17(b) are not stable. At  $n = 0.875$ , a homogeneous solution with both superconductivity and antiferromagnetism is found (Fig. 18(a)). However, the thermodynamic extrapolation gives zero for both AF and SC order parameters. At  $n = 0.85$ , the  $4 \times 4$  cluster result also shows slight inhomogeneity, with a  $(\pi, \pi)$  modulation of the d-wave order parameter (Fig. 18(b)). At  $n = 0.8$ , where the  $8 \times 2$  impurity cluster gives a slightly lower energy ( $\Delta E = 0.003(2)$ ), DMET calculations indicate a weak spin density wave (Fig. 17(c)). This spin density wave may still exist in the thermodynamic limit because the amplitude is comparable to the staggered magnetization in smaller clusters (eg.  $m = 0.04$  for  $2 \times 2$  clusters).

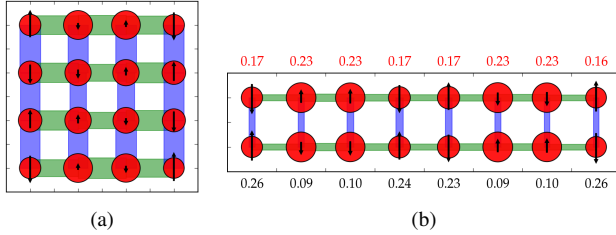


FIG. 19. Local order parameters for  $U = 6, n = 0.8, t' = -0.2$ .

We now turn to  $t' = -0.2$ . At  $n = 0.8$  and 0.875,  $8 \times 2$  cluster calculations show inhomogeneous orders. At  $n = 0.875$ , the pattern is similar to what we observed for  $t' = 0$  at the same filling, and its energy  $E_{8 \times 2} = -0.8402(4)$  is much higher than that of the  $4 \times 4$  homogeneous solution  $E_{4 \times 4} = -0.850(3)$ . At  $n = 0.8$  (Fig. 19), both  $4 \times 4$  and  $8 \times 2$  cluster calculations show  $\pi$ -phase shifts in the spin density and d-wave order, while the  $8 \times 2$  cluster has an additional charge density wave. They are very similar in energy, with  $E_{8 \times 2} = -0.9283(2)$  and  $E_{4 \times 4} = -0.927(3)$ . This suggests that the ground state here is superconducting with a superimposed spin density wave.

Most results for the underdoped region at  $U = 8$  are already shown in the main text (Fig. 4). In Table III, we compare energies for the two 16-site clusters. At all the points shown in the table, the  $8 \times 2$  cluster gives a lower energy. An unusual

result is that at  $n = 0.8, t' = 0$ , the  $8 \times 2$  cluster shows a homogeneous solution, while both the  $4 \times 4$  and  $4 \times 2$  clusters give a spin  $\pi$ -phase shift. This unusual behavior, where the  $8 \times 2$  solution favors homogeneity while the other cluster shapes do not, make thermodynamic extrapolation unreliable and thus gives a large error bar in the thermodynamic estimate of the energy ( $\delta E = 0.03$ ).

TABLE III. Energy comparison for different 16-site impurity clusters at  $U = 8$ .

$t'$	$n$	$E_{8 \times 2}$	$E_{4 \times 4}$
0	0.8	-0.9018(13)	-0.873(6)
0	0.875	-0.7548(4)	-0.748(4)
-0.2	0.8	-0.8487(4)	-0.846(10)
-0.2	0.875	-0.7556(5)	-0.737(7) <sup>a</sup>

<sup>a</sup> The error estimate may not be reliable at this point, because we have only two self-consistent DMET calculations with  $M=1000$  and 1200.

Finally, we end our discussion on the results by showing the energies across the phase space in Fig. 20. At half-filling, the energy in the frustrated model  $t' = \pm 0.2$  is slightly below  $t' = 0$ , while the difference becomes negligible at large  $U$ . At large doping, eg.  $n \leq 0.8$ , the energy order is dominated by the kinetic effects, i.e.  $E_{t'=-0.2} > E_{t'=0} > E_{t'=0.2}$ . The energy curves show more complicated behaviour in the underdoped region, especially for  $t' = 0$  and  $t' = -0.2$ .

#### Appendix D: Data set

In the attached *TDL.csv* file, we present the energy, chemical potential and (averaged) order parameters computed and their uncertainties at the thermodynamic limit. Since the averaged order parameters are meaningless when inhomogeneity dominates, we have removed these entries from the table.

In the file *clusters.csv*, we present the results for finite impurity clusters. In addition to the results available at thermodynamic limit, we also present the local order parameters. The local order parameters are encoded in an 1D array, which is explained in Fig. The errors shown only include the DMET convergence error, as the other sources of error can be deduced using the procedures described above, from the raw data. We also include the local orders (charge, spin and pairing strength) in this table as a 1D array. The order of the sites and pairs are shown in Fig. 21.



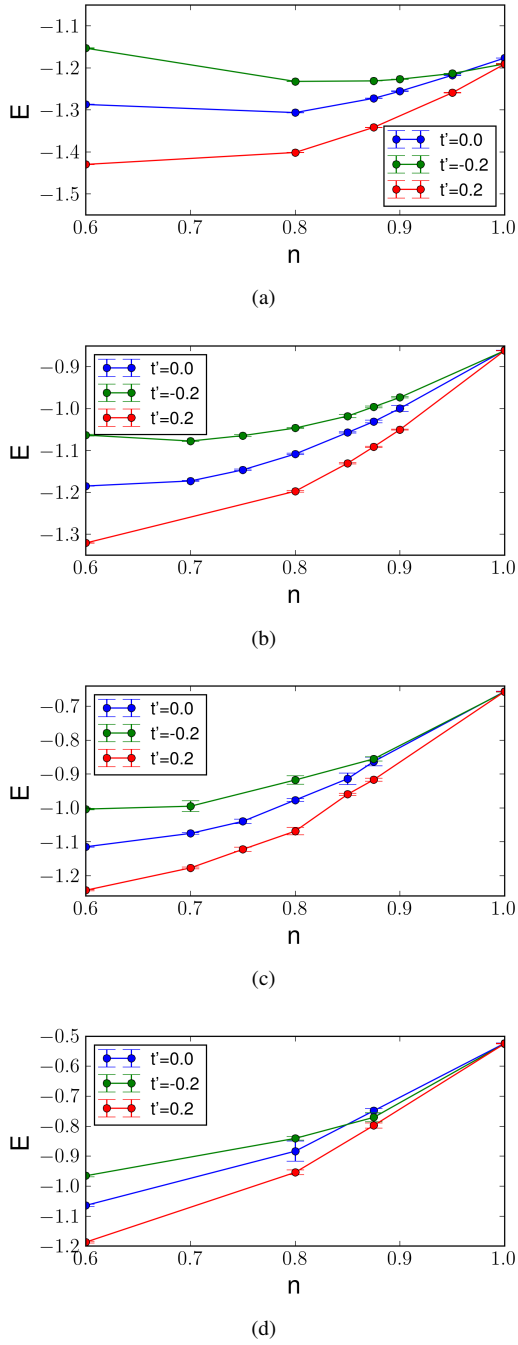


FIG. 20. DMET thermodynamic energy over the phase space. (a)  $U=2$  (b)  $U=4$  (c)  $U=6$  (d)  $U=8$

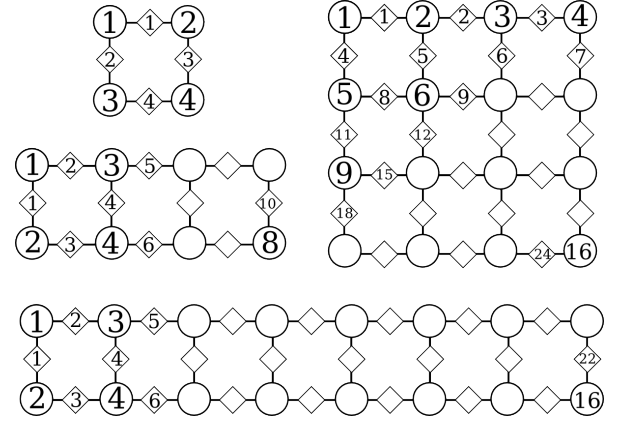


FIG. 21. The encoding of local order parameters for all impurity clusters. Numbers shown in the circles represent the order of sites, which is associated with labelling the charge density and spin density. The numbers in the rhombuses represent the order of bonds, or pairs between neighbor sites, which is associated with labelling the pairing strength. Some numbers are omitted since they are simple to deduce.

- <sup>1</sup> M. C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963).
- <sup>2</sup> J. Kanamori, Progress of Theoretical Physics **30**, 275 (1963).
- <sup>3</sup> J. Hubbard, in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, Vol. 276 (The Royal Society, 1963) pp. 238–257.
- <sup>4</sup> F. C. Zhang and T. M. Rice, Phys. Rev. B **37**, 3759 (1988).
- <sup>5</sup> E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994).
- <sup>6</sup> D. Scalapino, in *Handbook of High-Temperature Superconductivity* (Springer, 2007) pp. 495–526.
- <sup>7</sup> M. Rigol, T. Bryant, and R. R. P. Singh, Phys. Rev. Lett. **97**, 187202 (2006).
- <sup>8</sup> E. Khatami and M. Rigol, Phys. Rev. A **84**, 053611 (2011).
- <sup>9</sup> E. Khatami, R. T. Scalettar, and R. R. Singh, arXiv:1503.06213 (2015).
- <sup>10</sup> J. E. Hirsch, Phys. Rev. B **31**, 4403 (1985).
- <sup>11</sup> A. Georges and G. Kotliar, Phys. Rev. B **45**, 6479 (1992).
- <sup>12</sup> A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).
- <sup>13</sup> A. N. Rubtsov, M. I. Katsnelson, and A. I. Lichtenstein, Phys. Rev. B **77**, 033101 (2008).
- <sup>14</sup> G. Rohringer, A. Valli, and A. Toschi, Phys. Rev. B **86**, 125114 (2012).
- <sup>15</sup> H. Schweitzer and G. Czycholl, Zeitschrift für Physik B Condensed Matter **83**, 93 (1991).
- <sup>16</sup> C. J. Halboth and W. Metzner, Phys. Rev. B **61**, 7364 (2000).
- <sup>17</sup> S. Raghu, S. Kivelson, and D. Scalapino, Phys. Rev. B **81**, 224505 (2010).
- <sup>18</sup> C. N. Varney, C.-R. Lee, Z. J. Bai, S. Chiesa, M. Jarrell, and R. T. Scalettar, Phys. Rev. B **80**, 075116 (2009).
- <sup>19</sup> A. C. Cosentini, M. Capone, L. Guidoni, and G. B. Bachelet, Phys. Rev. B **58**, R14685 (1998).
- <sup>20</sup> F. Becca, M. Capone, and S. Sorella, Phys. Rev. B **62**, 12700 (2000).
- <sup>21</sup> H. J. M. van Bemmelen, D. F. B. ten Haaf, W. van Saarloos, J. M. J. van Leeuwen, and G. An, Phys. Rev. Lett. **72**, 2442 (1994).
- <sup>22</sup> L. F. Tocchio, F. Becca, A. Parola, and S. Sorella, Phys. Rev. B **78**, 041101 (2008).
- <sup>23</sup> S. Zhang, J. Carlson, and J. E. Gubernatis, Phys. Rev. B **55**, 7464 (1997).
- <sup>24</sup> C.-C. Chang and S. Zhang, Phys. Rev. B **78**, 165101 (2008).
- <sup>25</sup> C.-C. Chang and S. Zhang, Phys. Rev. Lett. **104**, 116402 (2010).
- <sup>26</sup> H. Yokoyama and H. Shiba, Journal of the Physical Society of Japan **56**, 1490 (1987).
- <sup>27</sup> D. Eichenberger and D. Baeriswyl, Phys. Rev. B **76**, 180504 (2007).
- <sup>28</sup> K. Yamaji, T. Yanagisawa, T. Nakanishi, and S. Koike, Physica C: Superconductivity **304**, 225 (1998).
- <sup>29</sup> T. Giamarchi and C. Lhuillier, Phys. Rev. B **43**, 12943 (1991).
- <sup>30</sup> S. R. White and D. J. Scalapino, Phys. Rev. B **61**, 6320 (2000).
- <sup>31</sup> D. J. Scalapino and S. R. White, Foundations of Physics **31**, 27 (2001).
- <sup>32</sup> S. R. White and D. J. Scalapino, Phys. Rev. Lett. **91**, 136403 (2003).
- <sup>33</sup> M. H. Hettler, A. N. Tahvildar-Zadeh, M. Jarrell, T. Pruschke, and H. R. Krishnamurthy, Phys. Rev. B **58**, R7475 (1998).
- <sup>34</sup> M. H. Hettler, M. Mukherjee, M. Jarrell, and H. R. Krishnamurthy, Phys. Rev. B **61**, 12739 (2000).
- <sup>35</sup> A. I. Lichtenstein and M. I. Katsnelson, Phys. Rev. B **62**, R9283 (2000).
- <sup>36</sup> G. Kotliar, S. Y. Savrasov, G. Pálsson, and G. Biroli, Phys. Rev. Lett. **87**, 186401 (2001).
- <sup>37</sup> M. Potthoff, M. Aichhorn, and C. Dahnken, Phys. Rev. Lett. **91**, 206402 (2003).
- <sup>38</sup> C. Dahnken, M. Aichhorn, W. Hanke, E. Arrigoni, and M. Potthoff, Phys. Rev. B **70**, 245110 (2004).
- <sup>39</sup> M. Aichhorn and E. Arrigoni, EPL (Europhysics Letters) **72**, 117 (2005).
- <sup>40</sup> D. Sénéchal, P.-L. Lavertu, M.-A. Marois, and A.-M. S. Tremblay, Phys. Rev. Lett. **94**, 156404 (2005).
- <sup>41</sup> M. Aichhorn, E. Arrigoni, M. Potthoff, and W. Hanke, Phys. Rev. B **74**, 024508 (2006).
- <sup>42</sup> C. J. Halboth and W. Metzner, Phys. Rev. Lett. **85**, 5162 (2000).
- <sup>43</sup> H. J. Schulz, Phys. Rev. Lett. **64**, 1445 (1990).
- <sup>44</sup> S. R. White, D. J. Scalapino, R. L. Sugar, E. Y. Loh, J. E. Gubernatis, and R. T. Scalettar, Phys. Rev. B **40**, 506 (1989).
- <sup>45</sup> A. V. Chubukov and K. A. Musaelian, Phys. Rev. B **51**, 12605 (1995).
- <sup>46</sup> P. A. Igoshev, M. A. Timirgazin, A. A. Katanin, A. K. Arzhnikov, and V. Y. Irkhin, Phys. Rev. B **81**, 094407 (2010).
- <sup>47</sup> M. Capone and G. Kotliar, Phys. Rev. B **74**, 054513 (2006).
- <sup>48</sup> E. Gull, O. Parcollet, and A. J. Millis, Phys. Rev. Lett. **110**, 216405 (2013).
- <sup>49</sup> M. Kato, K. Machida, H. Nakanishi, and M. Fujita, Journal of the Physical Society of Japan **59**, 1047 (1990).
- <sup>50</sup> A. Moreo, D. J. Scalapino, R. L. Sugar, S. R. White, and N. E. Bickers, Phys. Rev. B **41**, 2313 (1990).
- <sup>51</sup> M. Miyazaki, K. Yamaji, and T. Yanagisawa, Journal of Physics and Chemistry of Solids **63**, 1403 (2002).
- <sup>52</sup> T. Mizusaki and M. Imada, Phys. Rev. B **74**, 014421 (2006).
- <sup>53</sup> T. A. Maier, M. Jarrell, and D. J. Scalapino, Phys. Rev. B **74**, 094513 (2006).
- <sup>54</sup> T. A. Maier, M. Jarrell, T. C. Schulthess, P. R. C. Kent, and J. B. White, Phys. Rev. Lett. **95**, 237001 (2005).
- <sup>55</sup> A. Macridin, M. Jarrell, and T. Maier, Phys. Rev. B **70**, 113105 (2004).
- <sup>56</sup> M. Jarrell, T. Maier, M. H. Hettler, and A. N. Tahvildarzadeh, EPL (Europhysics Letters) **56**, 563 (2001).
- <sup>57</sup> J. Otsuki, H. Hafermann, and A. I. Lichtenstein, Phys. Rev. B **90**, 235132 (2014).
- <sup>58</sup> K.-S. Chen, Z. Y. Meng, T. Pruschke, J. Moreno, and M. Jarrell, Phys. Rev. B **86**, 165136 (2012).
- <sup>59</sup> G. Knizia and G. K.-L. Chan, Phys. Rev. Lett. **109**, 186404 (2012).
- <sup>60</sup> G. Knizia and G. K.-L. Chan, J. Chem. Theory Comput. **9**, 1428 (2013).
- <sup>61</sup> Q. Chen, G. H. Booth, S. Sharma, G. Knizia, and G. K.-L. Chan, Phys. Rev. B **89**, 165134 (2014).
- <sup>62</sup> I. W. Bulik, G. E. Scuseria, and J. Dukelsky, Phys. Rev. B **89**, 035140 (2014).
- <sup>63</sup> Z. Fan and Q.-I. Jie, Phys. Rev. B **91**, 195118 (2015).
- <sup>64</sup> Q. Sun and G. K.-L. Chan, J. Chem. Theory Comput. **10**, 3784 (2014).
- <sup>65</sup> I. W. Bulik, W. Chen, and G. E. Scuseria, J. Chem. Phys. **141**, 054113 (2014).
- <sup>66</sup> G. H. Booth and G. K.-L. Chan, Phys. Rev. B **91**, 155107 (2015).
- <sup>67</sup> I. Peschel, Brazilian Journal of Physics **42**, 267 (2012).
- <sup>68</sup> G. K.-L. Chan and S. Sharma, Annual Review of Physical Chemistry **62**, 465 (2011), pMID: 21219144, <http://www.annualreviews.org/doi/pdf/10.1146/annurev-physchem-032210-103338>.
- <sup>69</sup> S. R. White and A. L. Chernyshev, Phys. Rev. Lett. **99**, 127004 (2007).

- <sup>70</sup> T. A. Maier and M. Jarrell, Phys. Rev. B **65**, 041104 (2002).
- <sup>71</sup> M. Qin and S. Zhang, (2015), private communication.
- <sup>72</sup> S. Zhang and H. Krakauer, Phys. Rev. Lett. **90**, 136401 (2003).
- <sup>73</sup> S. White, (2014), private communication.
- <sup>74</sup> J. P. F. LeBlanc and E. Gull, Phys. Rev. B **88**, 155108 (2013).
- <sup>75</sup> M. Aichhorn, E. Arrigoni, M. Potthoff, and W. Hanke, Phys. Rev. B **76**, 224509 (2007).
- <sup>76</sup> A. W. Sandvik, Phys. Rev. B **56**, 11678 (1997).
- <sup>77</sup> W. Metzner and D. Vollhardt, Phys. Rev. B **39**, 4462 (1989).
- <sup>78</sup> E. Pavarini, I. Dasgupta, T. Saha-Dasgupta, O. Jepsen, and O. K. Andersen, Phys. Rev. Lett. **87**, 047003 (2001).
- <sup>79</sup> Z. B. Huang, H. Q. Lin, and J. E. Gubernatis, Phys. Rev. B **64**, 205101 (2001).
- <sup>80</sup> A. Eberlein and W. Metzner, Phys. Rev. B **89**, 035126 (2014).
- <sup>81</sup> S. Zhang, J. Carlson, and J. E. Gubernatis, Phys. Rev. Lett. **78**, 4486 (1997).
- <sup>82</sup> J. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, G. K. Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jiménez-Hoyos, *et al.*, arXiv:1505.02290 (2015).
- <sup>83</sup> D. J. Scalapino, E. Loh, and J. E. Hirsch, Phys. Rev. B **34**, 8190 (1986).
- <sup>84</sup> R. Peters and N. Kawakami, Phys. Rev. B **89**, 155134 (2014).
- <sup>85</sup> A. Leprévost, O. Juillet, and R. Frésard, arXiv:1503.04664 (2015).
- <sup>86</sup> D. Poilblanc and T. M. Rice, Phys. Rev. B **39**, 9749 (1989).
- <sup>87</sup> M. Vojta and S. Sachdev, Phys. Rev. Lett. **83**, 3916 (1999).
- <sup>88</sup> A. Melikyan and Z. Tešanović, Phys. Rev. B **71**, 214511 (2005).
- <sup>89</sup> H.-D. Chen, O. Vafek, A. Yazdani, and S.-C. Zhang, Phys. Rev. Lett. **93**, 187002 (2004).
- <sup>90</sup> P. A. Lee, Phys. Rev. X **4**, 031017 (2014).
- <sup>91</sup> E. Berg, E. Fradkin, and S. A. Kivelson, Nature Physics **5**, 830 (2009).
- <sup>92</sup> S. R. White and D. J. Scalapino, Phys. Rev. Lett. **80**, 1272 (1998).
- <sup>93</sup> C. S. Hellberg and E. Manousakis, Phys. Rev. Lett. **83**, 132 (1999).
- <sup>94</sup> G. Hager, G. Wellein, E. Jeckelmann, and H. Fehske, Phys. Rev. B **71**, 075108 (2005).
- <sup>95</sup> P. Corboz, S. R. White, G. Vidal, and M. Troyer, Phys. Rev. B **84**, 041108 (2011).
- <sup>96</sup> V. V. Moshchalkov, J. Vanacken, and L. Trappeniers, Phys. Rev. B **64**, 214504 (2001).
- <sup>97</sup> M. Fleck, A. I. Lichtenstein, and A. M. Oleś, Phys. Rev. B **64**, 134528 (2001).
- <sup>98</sup> T. Valla, A. Fedorov, J. Lee, J. Davis, and G. Gu, Science **314**, 1914 (2006).
- <sup>99</sup> J.-X. Li, C.-Q. Wu, and D.-H. Lee, Phys. Rev. B **74**, 184515 (2006).
- <sup>100</sup> T. A. Sedrakyan and A. V. Chubukov, Phys. Rev. B **81**, 174536 (2010).
- <sup>101</sup> P. Corboz, arXiv:1508.04003 (2015).
- <sup>102</sup> P.-G. de Gennes, *Superconductivity Of Metals And Alloys* (Benjamin, New York, 1966).
- <sup>103</sup> D. Yamaki, T. Ohsaku, H. Nagao, and K. Yamaguchi, International Journal of Quantum Chemistry **96**, 10 (2004).
- <sup>104</sup> J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).
- <sup>105</sup> S. Datta and P. F. Bagwell, Superlattices and Microstructures **25**, 1233 (1999).
- <sup>106</sup> G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **116**, 4462 (2002).
- <sup>107</sup> G. K.-L. Chan, J. Chem. Phys. **120**, 3172 (2004).
- <sup>108</sup> S. Sharma and G. K.-L. Chan, J. Chem. Phys. **136**, 124121 (2012).
- <sup>109</sup> H. Q. Lin and J. E. Hirsch, Phys. Rev. B **35**, 3359 (1987).