

Practical consequences of the Luttinger-Ward functional multivaluedness for cluster DMFT methods

J. Vučićević,^{1,2} N. Wentzell,¹ M. Ferrero,^{3,4} and O. Parcollet¹

¹*Institut de Physique Théorique (IPhT), CEA, CNRS, UMR 3681, 91191 Gif-sur-Yvette, France*

²*Scientific Computing Laboratory, Center for the Study of Complex Systems, Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia*

³*Centre de Physique Théorique, Ecole Polytechnique, CNRS, Université Paris-Saclay, 91128 Palaiseau, France*

⁴*Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France*



(Received 3 October 2017; revised manuscript received 7 February 2018; published 23 March 2018)

The Luttinger-Ward functional (LWF) has been a starting point for conserving approximations in many-body physics for 50 years. The recent discoveries of its multivaluedness and the associated divergence of the two-particle irreducible vertex function Γ have revealed an inherent limitation of this approach. Here we demonstrate how these undesirable properties of the LWF can lead to a failure of computational methods based on an approximation of the LWF. We apply the nested cluster scheme (NCS) to the Hubbard model and observe the existence of an additional stationary point of the self-consistent equations, associated with an unphysical branch of the LWF. In the strongly correlated regime, starting with the first divergence of Γ , this unphysical stationary point becomes attractive in the standard iterative technique used to solve DMFT. This leads to an incorrect solution, even in the large cluster size limit, for which we discuss diagnostics.

DOI: [10.1103/PhysRevB.97.125141](https://doi.org/10.1103/PhysRevB.97.125141)

The Luttinger-Ward functional (LWF) Φ is a central object in the quantum many-body theory of strongly correlated fermionic systems. Φ [1] is defined as the interacting part of the Legendre transform of the free energy with respect to the bare propagator G_0 [2]. It is a functional of the full propagator G , formally equal to the sum of all vacuum skeleton diagrams [3,4]. Φ has been the basis of many approximations in the field over the last decades.

Dynamical mean-field theory (DMFT) [5,6] and its cluster extensions [7–10] are a class of Φ -derivable approximations with a systematic control parameter: the size N_c of the cluster. They interpolate between DMFT ($N_c = 1$) and the exact solution of the lattice model for $N_c = \infty$. Cluster methods allow to treat the Mott physics à la DMFT and to include short-range spatial correlations. They have led to significant progress in recent years, in particular on the Hubbard model [11–58]. Cluster DMFT methods are formulated in terms of one (or a few) auxiliary quantum impurity models in a noninteracting bath encoded in the bare propagator \mathcal{G} . The bath is determined self-consistently in such a way that the impurity Green function G^{imp} coincides with some (local) components of the Green function of the lattice model G^{latt} . This *representability* property, i.e., the possibility to find \mathcal{G} for a given G^{imp} in a quantum impurity model lies at the very heart of DMFT methods [6,59].

Surprisingly, it was recently discovered [60–62] in simple strongly correlated models that the functional $\Phi[G]$ is in fact multivalued, i.e., has multiple branches. As a consequence, the relation $G[\mathcal{G}]$ cannot always be inverted in quantum impurity models as several \mathcal{G} yield the same Green function G . This has deep consequences for numerical methods in some parameter regimes. The crossing of two branches of Φ leads to divergence of the two-particle irreducible vertex Γ [61,63–66] and

therefore the breakdown of the parquet decomposition [65,67,68]. Moreover, at strong coupling, the bold diagrammatic series can converge to an incorrect result, as was checked explicitly using a bold quantum Monte Carlo algorithm [60]. Similar pathological behavior was observed in the context of GW -like approximations of Φ [69].

In this paper, we show that the multivaluedness of Φ has unexpected and severe consequences in certain cluster DMFT methods, and can potentially lead to incorrect results. Concretely, we study the *nested cluster* DMFT scheme (NCS) [5,52,70], which is an early example of the recently introduced self-energy embedding theory (SEET) [71–75]. NCS is a particularly interesting scheme since it addresses the main drawbacks of the most widely used cluster methods: cellular DMFT (CDMFT) [9] and the dynamical cluster approximation (DCA) [7]. It is a real-space cluster method which is translationally invariant (unlike CDMFT) and yields a continuous self-energy in reciprocal space (unlike DCA). In the classical limit, it reduces to the well-known Bethe-Kikuchi method of classical statistical physics [5,52].

We solve the NCS for the Hubbard model and compare it to benchmarks established with converged large DCA clusters. At weak to moderate couplings the scheme is stable and performs very well. Even at strong coupling, there is a physical solution, which is very close to the benchmarks already at moderate cluster size. However, (i) in the standard iterative method used to solve the DMFT equations, this solution is unstable towards an unphysical solution characterized by a noncausal Weiss field; (ii) as the cluster size increases, this stable unphysical solution converges to an incorrect result; and (iii) this occurs in the strong-coupling regime as delimited by the generalization of the divergences of the irreducible vertex observed in Refs. [61–65,76,77].

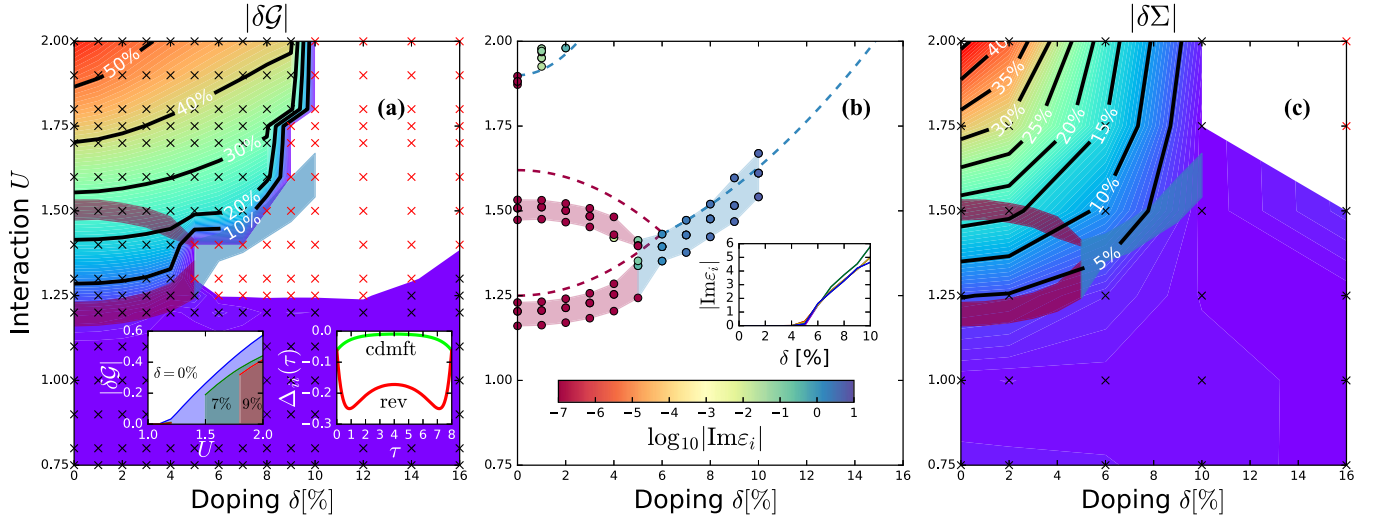


FIG. 1. For all panels, the temperature is $T/D = 0.125$. (a) Color plot of $|\delta\mathcal{G}| \equiv |\mathcal{G}_{00}^{\text{cdmft}}(i\omega_0) - \mathcal{G}_{00}^{\text{rev}}(i\omega_0)|/|\mathcal{G}_{00}^{\text{cdmft}}(i\omega_0)|$; black crosses are data points, red crosses are points where the reverse impurity solver does not converge. (Left inset) Slices for fixed doping, showing discontinuity vs U for $\delta = 7\%$ and 9% . (Right inset) Hybridization (local) of $\mathcal{G}^{\text{cdmft}}$ (top, green) and \mathcal{G}^{rev} (bottom, red), for $U = 2$ and $\delta = 0$ showing its causality violation. (b) Vertex divergences, where the real part of an eigenvalue ε_i of $\tilde{\chi}_c^{\Omega=0}$ crosses zero for single-site DMFT (dashed line) and 2×2 CDMFT (colored circles). Color encodes $\text{Im} \varepsilon_i$ at the given point; colored stripes are guides for the eyes. (Inset) $\text{Im} \varepsilon_i$ vs doping for the bottom two groups of circles for 2×2 CDMFT. (c) Color plot of $|\delta\Sigma| \equiv |\text{Im}\Sigma_{00}^{\text{cdmft}}(i\omega_0) - \text{Im}\Sigma_{00}^{\text{nested}}(i\omega_0)|/|\text{Im}\Sigma_{00}^{\text{cdmft}}(i\omega_0)|$, i.e., the difference between the imaginary part of the local self-energy for 2×2 NCS and the 2×2 CDMFT (the latter is close to the exact solution, see Appendix A1 and Fig. 3).

We consider the Hubbard model on a square lattice:

$$H = -t \sum_{\langle ij \rangle \sigma} c_{i\sigma}^\dagger c_{j\sigma} - \mu \sum_{i\sigma} n_{i\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}, \quad (1)$$

where $c_{i\sigma}^\dagger$ creates a fermion with spin σ at site i . The density operator is $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$. The nearest-neighbor hopping amplitude is t , the on-site interaction U and the chemical potential μ . $D = 4t$ is the unit of energy. We use the CT-INT algorithm to solve the quantum impurity model [78,79].

Let us first address the representability issue of the Green function G by a Weiss field \mathcal{G} in a cluster impurity model. We consider a 2×2 CDMFT calculation for $T/D = 0.125$ and various U and dopings δ , where it yields a quantitatively good solution as compared to converged large cluster DCA benchmarks (see Fig. 3). The CDMFT self-consistency equation reads [9] $G^{\text{imp}}[\mathcal{G}] = G^{\text{loc}}[\mathcal{G}]$ with

$$G^{\text{loc}}[\mathcal{G}](i\omega_n) \equiv \sum_{k \in \text{RBZ}} (i\omega_n + \mu - \hat{\varepsilon}_k - \Sigma^{\text{imp}}[\mathcal{G}](i\omega_n))^{-1},$$

where $\hat{\varepsilon}_k$ is the dispersion over the superlattice of clusters, RBZ is the reduced Brillouin zone and Σ^{imp} (resp. G^{imp}) is the impurity cluster self-energy (respectively, Green function). The CDMFT equations are solved with the usual iterative technique for DMFT; given $\mathcal{G}^{(i)}$ at iteration i , the impurity model yields $\Sigma^{\text{imp}}[\mathcal{G}^{(i)}]$ and the next iteration $\mathcal{G}^{(i+1)}$ is given by

$$\mathcal{G}^{(i+1)} = (G^{\text{loc}}[\mathcal{G}^{(i)}]^{-1} + \Sigma^{\text{imp}}[\mathcal{G}^{(i)}])^{-1}. \quad (2)$$

Starting from the converged CDMFT solution G^{cdmft} we then implement a *reverse quantum impurity solver* [60]: we seek a bare propagator \mathcal{G}^{rev} of the cluster model such that $G^{\text{imp}}[\mathcal{G}^{\text{rev}}] = G^{\text{cdmft}}$, with a similar iterative method as in

Eq. (2) but with $G^{\text{loc}}[\mathcal{G}^{(i)}]$ replaced by G^{cdmft} , which remains fixed in the calculation.

In Fig. 1(a), we present the relative difference between the local component of the converged CDMFT Weiss field $\mathcal{G}^{\text{cdmft}}$ and the result of the reverse impurity solver \mathcal{G}^{rev} . We observe three regions. At weak coupling, the reverse impurity solver yields $\mathcal{G}^{\text{cdmft}}$ as naively expected. At strong coupling and high doping, the reverse solver does not converge. At strong coupling and low doping, \mathcal{G}^{rev} progressively deviates from $\mathcal{G}^{\text{cdmft}}$, even though they both yield the exact same Green function G^{cdmft} . As soon as \mathcal{G}^{rev} is different from $\mathcal{G}^{\text{cdmft}}$ it acquires a noncausal hybridization function Δ [80] as shown in the inset of Fig. 1(a). Indeed, $\Delta(\tau)$ is not concave over the full $[0, \beta]$ interval and therefore has a corresponding spectral function with negative parts. This calculation demonstrates the existence of multiple branches of Φ for the 2×2 impurity problem by exhibiting explicitly two \mathcal{G} (and hence Σ) giving the same G , see also Refs. [60,61,63–65,77]. We will see below that a similar phenomenon occurs in NCS.

It is interesting to note that in the reverse impurity calculation at low doping $\delta < 5\%$, one first finds $\mathcal{G}^{\text{rev}} = \mathcal{G}^{\text{cdmft}}$ for small interactions $U < 1.25$ and then *continuously* switches to an unphysical solution for \mathcal{G} as U is increased. This means that the physical branch of Φ crosses the unphysical branch. As has been discussed in the particle-hole symmetric case [61], this crossing has to be accompanied by a divergence of the corresponding two-particle irreducible vertex function Γ , since it is the second derivative of Φ with respect to G . We generalize the results of Refs. [63–65] to the doped case and map these divergences of Γ in the 2×2 CDMFT case, to obtain a characterization of the strong-coupling region, which is not linked to the details of an iterative algorithm. Given the two-particle propagator

$G_{2,\sigma\sigma',ijkl}^{\omega\omega'\Omega} = \frac{1}{\beta} \langle c_{i,\sigma}^\dagger(\omega) c_{j,\sigma}(\omega + \Omega) c_{k,\sigma'}^\dagger(\omega' + \Omega) c_{l,\sigma'}(\omega') \rangle$ and the single-particle Green function G , Γ can be calculated with the inverse Bethe-Salpeter equation

$$\Gamma_{c,ijkl}^{\omega,\omega',\Omega} = \beta^2 \left[[\chi_0^\Omega]^{-1} - [\tilde{\chi}_c^\Omega]^{-1} \right]_{ij\omega,kl\omega'}, \quad (3)$$

where $\chi_{0,ijkl}^{\omega\omega'\Omega} = -G_{li}(i\omega)G_{jk}(i\omega + i\Omega)\beta\delta_{\omega,\omega'}$, and $\tilde{\chi}_{c,ijkl}^{\omega\omega'\Omega} = G_{2,\uparrow\uparrow,ijkl}^{\omega\omega'\Omega} + G_{2,\uparrow\downarrow,ijkl}^{\omega\omega'\Omega} - 2G_{ji}(\omega)G_{lk}(\omega')\beta\delta_{\Omega,0}$. The inverse is assumed to be in combined indices $(ij\omega)$ and $(lk\omega')$, where ω, ω' denote fermionic and Ω bosonic Matsubara frequencies. If $\tilde{\chi}_c(i\Omega)$ as a matrix has an eigenvalue $\varepsilon_i = 0$, it is singular and Γ diverges at the given $i\Omega$. While in single-site DMFT at particle-hole symmetry, the eigenvalues of $\tilde{\chi}_c(i\Omega = 0)$ are purely real by symmetry, it is no longer necessarily true here [65].

Figure 1(b) shows trajectories in the (δ, U) plane where the real part of an eigenvalue of $\tilde{\chi}_c$ crosses zero for single-site DMFT and 2×2 CDMFT. In single-site DMFT, at half-filling, there are three Γ divergences in the examined range of interaction, in agreement with Ref. [64]. As we go to finite doping, the divergence close to $U = 1.8$ disappears immediately as the corresponding eigenvalue acquires an imaginary part. However, the divergences close to $U = 1.2$ and 1.5 extend up to $\delta \simeq 5\%$ where they merge. For higher doping, the divergences disappear because the corresponding eigenvalues acquire an imaginary part. In CDMFT, the behavior is very similar except that each divergence is split into four, the two middle ones occurring simultaneously. Hence we conjecture (see also Refs. [61,63–65]) that the divergences in Γ are not an artifact of the single-site model but rather survive and multiply in the cluster impurity model. Finally, in the left inset of Fig. 1(a), we see that for $\delta \geq 6\%$, the unphysical solution appears discontinuously when U is increased, in agreement with the absence of a divergence in Γ .

Let us now turn to the NCS. It approximates Φ by $\Phi^{(L)}$, defined as its restriction to the set of real-space two particle irreducible (2PI) diagrams that involve lattice points lying within a box of shape $L \times L$. $\Phi^{(L)}$ can be expressed as a linear combination of the LWFs $\Phi_{L \times L}$ of a $L \times L$ cluster and the LWF of its subclusters, with appropriate weights that eliminate the double counting of diagrams. Each cluster LWF is associated to an impurity model, via the representability property. The lattice self-energy Σ^{latt} is therefore a linear combination of the self-energies of the impurities. This couples the impurity models together and the baths adjust so that, e.g., the impurity Green function is the same for every site of every cluster. This method was introduced for a two site cluster (a dimer) in Ref. [70], see also Refs. [5,52,81].

A priori, solving large nested clusters seems like a daunting task, requiring to solve a large number of coupled impurity problems, one for every subcluster of the $L \times L$ cluster. However, as shown in Appendix C2, it is sufficient to solve only *four* coupled clusters since

$$\begin{aligned} \Phi^{(L)}[G] &= \sum_i \Phi_{L \times L}[G|_{C_i^{L \times L}}] - \Phi_{L-1 \times L}[G|_{C_i^{L-1 \times L}}] \\ &\quad - \Phi_{L \times L-1}[G|_{C_i^{L \times L-1}}] + \Phi_{L-1 \times L-1}[G|_{C_i^{L-1 \times L-1}}], \end{aligned} \quad (4)$$

where $C_i^{n \times p}$ is the cluster of shape $n \times p$ whose bottom-left lattice point is i , and $G|_{C_i^{n \times p}}$ the restriction of the Green function to this cluster (i.e., the set $\{G_{lm}\}_{l,m \in C_i^{n \times p}}$). If we assume rotational invariance, the last two terms give the same contribution and the method can be solved using three coupled cluster impurity models. We present the full formalism for the NCS with several examples in Appendix C.

We solve the NCS using the standard iterative method of solution for DMFT equations as in Eq. (2). At weak coupling, the NCS yields a solution in excellent agreement with large DCA cluster benchmarks (see Fig. 3). However, at strong coupling the situation is more complex. First, in Fig. 1(c), we observe that the 2×2 NCS gives a poor result compared to CDMFT in the strong-coupling region delimited by the divergences of Γ discussed above. We then solve larger clusters $L = 2, 3, 4$, and 6 to examine the convergence of the method with the cluster size. We observe an unexpected and severe problem: *the nested cluster scheme converges for $L \rightarrow \infty$ but to an incorrect solution* even though formally $\Phi^{(L)} \xrightarrow{L \rightarrow \infty} \Phi$. In Fig. 2(a) and Appendix A4 we show the momentum dependent self-energies obtained for $L = 4$ and 6 ; they are very close to each other, indicating convergence, but quite far from the benchmarks.

To gain further insight, we study the convergence of the $L = 4$ case at strong coupling iteration by iteration. In Fig. 2(b), we plot the difference between successive G_{loc} for $U/D = 1$ and $U/D = 2$, $\delta = 0$. At $U/D = 1$ convergence is roughly exponential until the level of Monte-Carlo noise is reached. However, for $U/D = 2$, we observe that the self-consistency is almost converged after three iterations [the green point on Fig. 2(b)] to an unstable solution before finally converging to another solution of the equation (red point). Remarkably, in Fig. 2(a), we see that this unstable (green) solution is almost perfectly on top of the benchmark, contrary to the stable (red) one.

Furthermore, we observe two pathologies of the stable (red) solution, which can be used as diagnostics in the absence of benchmarks. First, the inset of Fig. 2(b) shows the local hybridization function of both solutions (at the corner of the $L \times L$ cluster), Δ_{stable} and Δ_{unstable} . Δ_{stable} clearly violates causality at $U/D = 2$, similarly to the reverse impurity solver studied above, while Δ_{unstable} is fine. Moreover, we see in Fig. 2(c) that this effect appears as a function of U for $U > 1.2$, i.e., exactly when the solution deviates from the benchmark (or CDMFT in this case). Second, the Δ_{stable} bath does not decay in the large L limit at strong coupling ($U/D = 2$), contrary to $U/D = 1$, as illustrated in the inset of Fig. 2(a). Contrary to CDMFT or DCA, the NCS does not impose $\Delta = 0$ for every converged solution for $L \rightarrow \infty$, but only the weaker condition $\Sigma^{\text{latt}} = \Sigma^{\text{imp}} + \Delta$ (see Appendix A4 and Fig. 8 for further discussion). For the physical solution, we conjecture that $\Delta \rightarrow 0$ for $L \rightarrow \infty$: the large cluster will be a Hubbard model with no bath. The unphysical solution converges on the other hand to a certain resummation of the bold diagrams series.

For CDMFT and DCA, the standard iterative method of solution is *iteratively causal* [7,9], i.e., one can prove that the bath stays causal at each iteration (and therefore at convergence). Hence the causality violation of the bath cannot occur and the solution stays on the physical branch. The NCS does

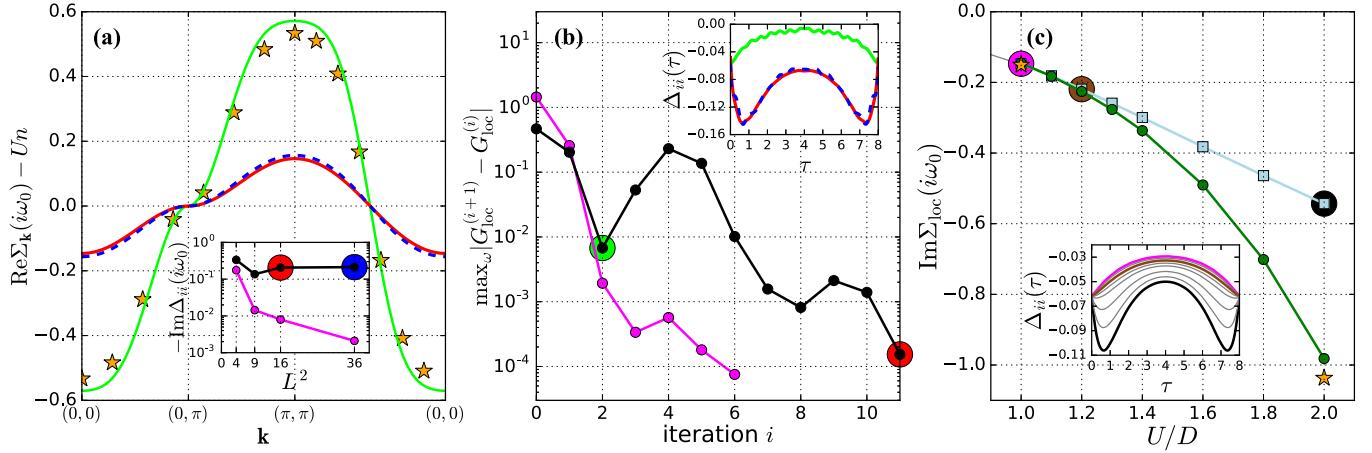


FIG. 2. (a) Real part of $\Sigma_{\mathbf{k}}(i\omega_0) - Un$ vs momentum \mathbf{k} , in NCS for $L = 4$ (solid line, red), $L = 6$ (dashed line, blue); unstable solution at the third iteration for $L = 4$ (solid line, green); the stars denote 98 sites DCA results. (Inset) Hybridization at the center of the $L \times L$ cluster at the first Matsubara frequency vs L for $U/D = 2$ (circles, black), and $U/D = 1$ (circles, magenta). (b) Norm of the difference of G_{loc} between iterations vs the iteration number i , for $L = 4$, $U/D = 2$ (circles, black), and $U/D = 1$ (circles, magenta). Green (resp. red) dot corresponds to the unstable (respectively, stable) solution, cf. text. (Inset) Hybridization at the corner of the $L \times L$ cluster vs τ for the unstable and the stable noncausal solution, with same convention as in (a). (c) $\text{Im}\Sigma_{\text{loc}}(i\omega_0)$ vs U/D for NCS $L = 2$ (square, blue online) and 2×2 CDMFT, (circle, green online). The star is 98 sites DCA. (Inset) Local hybridization for $L = 2$ vs τ .

not have this property, which, as we have seen, has drastic consequences on the stability of the physical solution in the iterative procedure. In the dimer case, NCS was already known to yield noncausal self-energies at low temperatures and strong coupling [5,70]. However, in previous works [52,82], this was simply interpreted as the signature of an insufficiently large cluster, i.e., a defect that the large L should cure.

To summarize, the nested cluster is a translationally invariant, real-space cluster method with a physical solution very close to numerically exact benchmarks already at moderate cluster sizes, both at weak and strong coupling. However, the multivaluedness of the LWF leads to an instability of the standard iterative procedure of solution in the strong-coupling region (as delimited by the divergence of the irreducible vertex Γ) towards an unphysical solution, even in the infinite cluster limit. This failure is signaled by causality violations of the hybridization function. All this points to the importance of distinguishing between a cluster method and the iterative procedure used to solve its equations. An important challenge is therefore to design new ways of solving the cluster DMFT equations that are guaranteed to stay on the physical branch of the LWF and stabilize the “hidden” physical solution, e.g., by implementing the “shifted-action” [83] proposal in this context. Alternatively, one can use cluster methods based on higher-order functionals (TRILEX [54,56,57], QUADRILEX [55]). We believe these are less likely to be multivalued, as it would require the existence of two systems with identical single-particle but also higher-order correlation functions, which is *a priori* harder to achieve. Moreover, going to higher-order functionals would correspond to adding more degrees of freedom to the solution, which in itself could remove the multivaluedness.

We are grateful to A. Georges and P. Thunström for useful insights and discussion. We further thank T. Schäfer and A. Toschi for critical reading of the manuscript. The DCA 98A and 50A data was provided by J. Leblanc. We thank Z. Mitrović

Vučičević for figure editing. This work is supported by the FP7/ERC, under Grant Agreement No. 278472-MottMetals (J.V., N.W.) and by the Simons Foundation within the Many Electron Collaboration (N.W.). Part of this work was performed using HPC resources from GENCI-TGCC (Grant No. 2016-t2016056112). The CT-INT algorithm has been implemented using the TRIQS toolbox [84].

APPENDIX A: BENCHMARKS

In this section, we present results for various cluster DMFT methods applied to the two-dimensional square-lattice Hubbard model as introduced in the main text. We pay special attention to the nested cluster scheme (NCS), which is discussed in detail in Appendix C. Detailed summary of other cluster DMFT methods is provided in Appendix D.

We first present an extensive benchmark against exact results (Appendix A1), which we use in the main text to determine the quality of solutions and to identify problematic regimes. We then address in particular the causality violations in the problematic region (subsection A2). In Appendix A3, we provide a comparison between two variants of the nested cluster scheme, differing in the nested quantity (self-energy versus cumulant). In Appendix A4, we discuss the stable and unstable solution of the nested equations.

1. Comparison against exact results

In Fig. 3, we show the results of cluster DMFT methods for the Hubbard model, at various cluster sizes, in the four corners and the center of the phase diagram examined in the main text. The temperature is $T/D = 0.125$. At half-filling, $\text{Re}\Sigma_{\text{loc}}(i\omega_n) = U/2$ by symmetry, so we omit this data. The nonlocal part $\tilde{\Sigma}_{\mathbf{k}}(i\omega_n) = \Sigma_{\mathbf{k}}(i\omega_n) - \Sigma_{\text{loc}}(i\omega_n)$ we present at the lowest Matsubara frequency, along a triangular path enclosing the irreducible Brillouin zone. With stars we denote the best available result: at half-filling, we have DCA

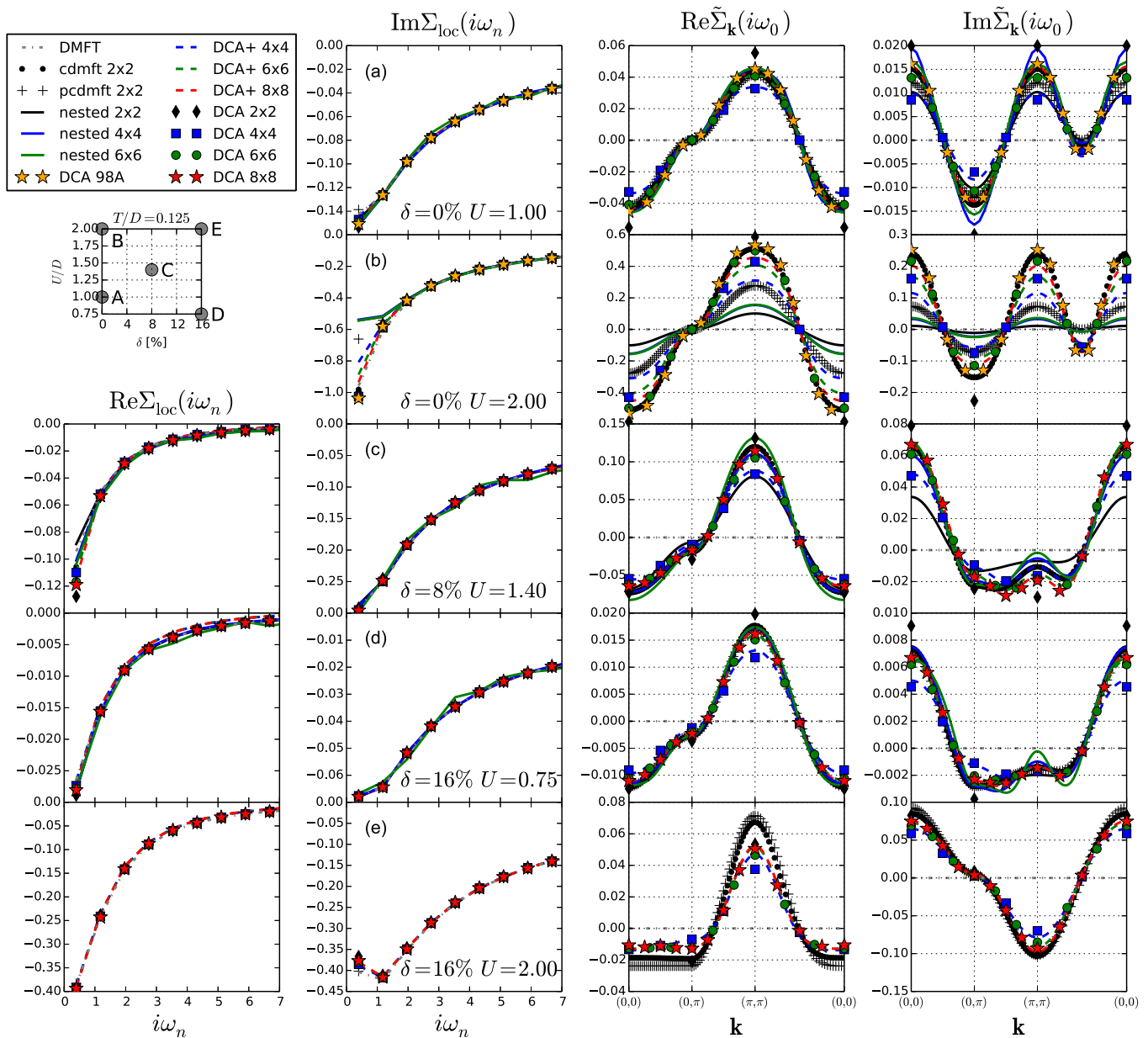


FIG. 3. Benchmark of DMFT and various cluster DMFT methods. Hubbard model square lattice, temperature $T/D = 0.125$. We present separately the local and the nonlocal parts of self-energy, $\tilde{\Sigma}_{\mathbf{k}} = \Sigma_{\mathbf{k}} - \Sigma_{\text{loc}}$. Color denotes cluster size, symbols/line styles different methods. Stars denote the best available result. Agreement is excellent in all points except point B: NCS converges to a wrong solution, PCDMFT 2×2 is considerably worse than in other points, and DCA⁺ converges very slowly with cluster size, while being a poor approximation at small cluster size. In other points, NCS 4×4 performs well, but at 6×6 amplification of statistical noise becomes a problem (see text for details).

$N_c = 98$, and away from half-filling, the biggest cluster is 8×8 ($N_c = 64$). These results are converged with respect to cluster size, and can be considered exact solutions of the Hubbard model.

The presented CDMFT result is the self-energy periodized by Eq. (D17) (in Appendix D2 below). In DCA, we are showing only the values at coarse-grained wave vectors \mathbf{K} (see Appendix D4).

We first concentrate on the points other than point B. We see excellent agreement of all methods. The local part is captured correctly already at 2×2 cluster size. DCA typically overestimates the amount of \mathbf{k} dependence at 2×2 , then underestimates it at 4×4 , and is mostly converged at 6×6 .

DCA⁺ has a similar behavior (2×2 not shown for the sake of clarity). CDMFT and PCDMFT give almost the same result, and are on top of the benchmark except for the real nonlocal part in point E, where the overall shape is correct, but the amplitude is overestimated slightly; PCDMFT also noticeably misses the local imaginary part in point A. Nested cluster performs well, and at 4×4 cluster size is even more accurate than DCA around $\mathbf{k} = (0,0)$. In point E, it does not converge at any cluster size. Away from half-filling and at cluster size 6×6 , statistical noise amplification in nested cluster becomes significant (see Appendix C1b for details). It is particularly noticeable in the local part of self-energy at high Matsubara frequencies, in points C and D. Also in these points, there is a

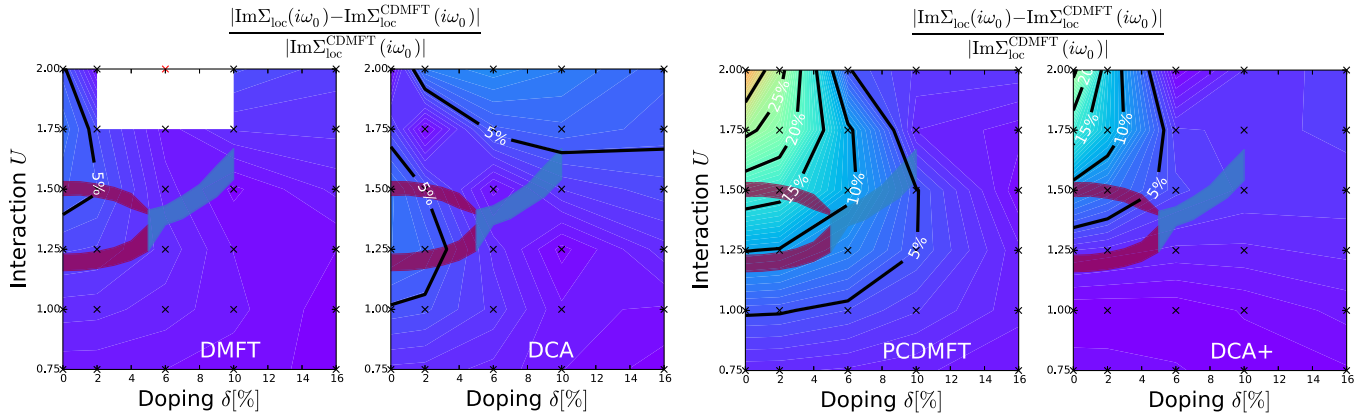


FIG. 4. Hubbard model, square lattice, temperature $T/D = 0.125$. Relative difference in $\text{Im}\Sigma_{\text{loc}}$ from the reference method (CDMFT 2×2). DCA, PCDMFT, and DCA^+ clusters are 2×2 . The red and blue stripes are the Γ divergence trajectories from CDMFT 2×2 calculation [Fig. 1(b) in main text]. In DCA and DMFT, performance is uniform across the phase diagram and appears unrelated to the divergence trajectories. In PCDMFT and DCA^+ , the result is considerably poorer in the region roughly bounded by the divergence trajectories.

peaklike feature around $\mathbf{k} = (\pi, \pi)$ in the nonlocal imaginary part. It comes from the numerous long distance self-energy components, which are small and comparable to the statistical error bar. These fine details of the solution can not be perfectly converged due to the statistical noise.

Now we turn to point B. CDMFT 2×2 is, again, on top of the benchmark. DCA behaves no differently than in other points, and is almost converged at 6×6 ; the local part is correct already at 2×2 . On the other hand, DCA^+ is not on top of the benchmark even at 8×8 , and especially the local part is strongly underestimated: at 8×8 , it is still worse than single-site DMFT. The nonlocal part is underestimated as well: the 6×6 calculation is comparable to DCA 4×4 . PCDMFT, similarly, underestimates both the local and nonlocal part. Nested cluster converges to a wrong solution with respect to N_c : the local part is indistinguishable already between 2×2 and 4×4 , and the nonlocal part between 4×4 and 6×6 . The local part is underestimated by about 50%, and imaginary nonlocal part by almost an order of magnitude. The failure of PCDMFT, NCS, and DCA^+ in this particular point is strongly reminiscent of the failure of bold-diagrammatic QMC presented in Ref. [60], for the same model parameters: the self-energy obtained in these methods is more metallic and much more local than the exact solution. We note that the similar phenomenon can also be observed in the original DCA^+ paper [48]—in the strongly coupled regime, the $N_c = 16$ DCA^+ self-energy result is much more local and metallic than that of the DCA at the same cluster size.

In conclusion, in this phase diagram, the best performing 2×2 method is CDMFT. We take it as a reference method for benchmarking on a denser (δ, U) grid [Fig. 1(c) in the main text, and Fig. 4 below]. At 4×4 cluster size, in the points where it works, NCS does have an advantage over CDMFT and DCA. DCA 4×4 coarse-graining is still quite crude—due to symmetries of the lattice, it yields only six independent self-energy components; NCS at the same size yields ten independent self-energy components, and captures longer distance processes. In DCA, interpretation of the results in real space is problematic; NCS results can be looked at in both \mathbf{r} and \mathbf{k} space. CDMFT is also problematic at 4×4 . At

this size, both the translational symmetry and the homogeneity within a supercell are broken, and the periodization becomes even less straight-forward. Finally, we note that in point B, even though NCS fails with forward substitution algorithm, there still appears to be a stationary point of the NCS equations (Fig. 2 in main text), which is in better agreement with the exact result than DCA at the same cluster size.

We finalize our analysis by a high-resolution benchmark of DMFT and 2×2 cluster methods (DCA, DCA^+ , and PCDMFT), analogous to Fig. 1(c) in the main text. In Fig. 4, we present the deviation from the exact result of these methods. DMFT and DCA perform uniformly well across the phase diagram, and are at most $\approx 5\%$ away from the correct result. No features can be associated with the Γ divergence trajectories. On the other hand, NCS, DCA^+ and PCDMFT all fail in similarly shaped regions around point B, but give good results in other regimes. In DCA^+ and PCDMFT the coincidence of the problematic region with the Γ -divergence trajectories is less conclusive, but we can similarly connect the failure with the noncausality of the hybridization function. It is, however, unclear whether a correct stationary point is present in these methods at all.

2. Causality properties

In this section, we analyze the causality properties of various quantities in the cluster methods presented above. Quantities like Green's functions, self-energies, and hybridization baths should have Lehmann spectral representation. The diagonal components of these quantities, should satisfy in real frequency

$$\text{Im}X_{ii}(\omega) < 0, \quad (\text{A1})$$

where X stands for G , Σ , \mathcal{G} , or Δ . This has implications for the shape of these objects in imaginary time:

$$X_{ii}(\tau) = \frac{1}{\pi} \int d\omega \frac{e^{-\tau\omega}}{1 + e^{-\beta\omega}} \text{Im}X_{ii}(\omega), \quad (\text{A2a})$$

$$\begin{aligned} \partial_\tau^{2n} X_{ii}(\tau) &= \frac{1}{\pi} \int d\omega \frac{\omega^{2n} e^{-\tau\omega}}{1 + e^{-\beta\omega}} \text{Im}X_{ii}(\omega) \\ &< 0, \quad n \in \mathbb{N} \end{aligned} \quad (\text{A2b})$$

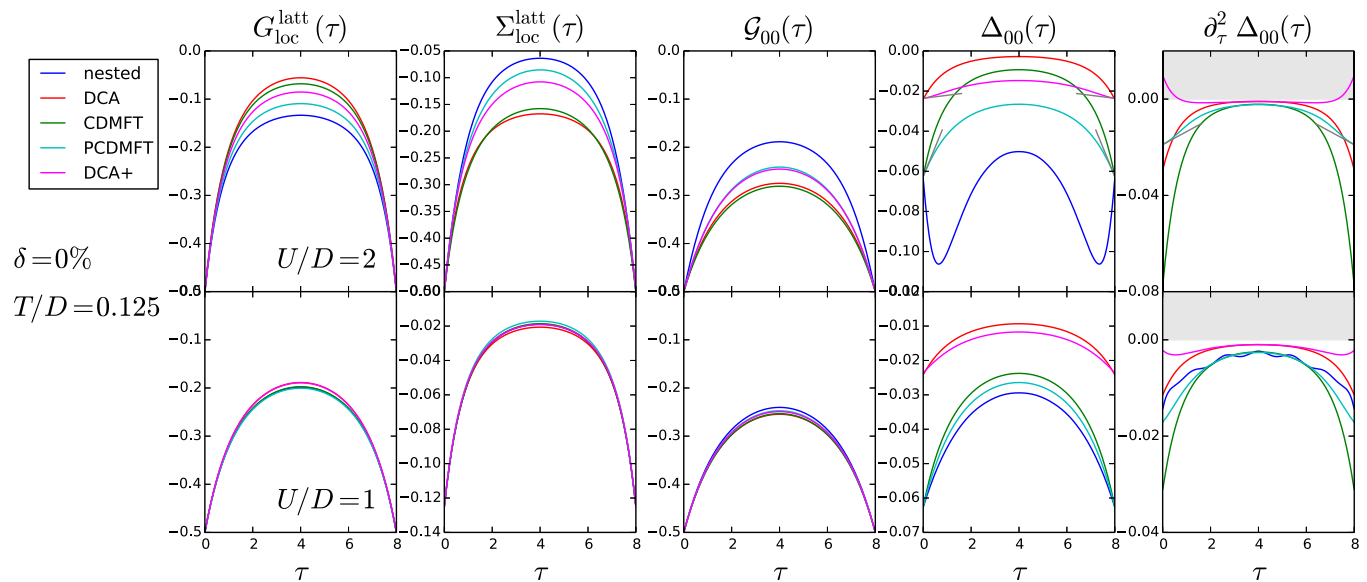


FIG. 5. Causality analysis of relevant quantities, in various cluster methods at 2×2 cluster size. Temperature $T/D = 0.125$, square lattice. At strong coupling, DCA^+ and NCS have pronounced inflection points in $\Delta(\tau)$, PCDMFT in the second derivative of $\Delta(\tau)$. Gray lines extrapolate the linear component close to $\tau = 0$ and β . Other quantities are all causal, including the bare propagator on the impurity. Increasing the cluster size in DCA^+ improves the causality in Δ , but not in NCS. In the upper right panel, NCS result is omitted for the sake of clarity [the noncausality is already obvious in $\Delta_{00}(\tau)$].

All even-order derivatives with respect to τ must be negative. This rules out the appearance of inflection points in $X_{ii}(\tau)$ and any of its even-order derivatives.

In Fig. 5, we present the results for the local G , Σ on the lattice, as well as the diagonal components of the bare propagator \mathcal{G} and the hybridization function Δ on the impurity, all in imaginary time. All methods used are at 2×2 cluster size. In NCS we present the impurity quantities only for the biggest cluster. In all methods at 2×2 , all the diagonal components of \mathcal{G} and Δ are the same by symmetry (in DCA/DCA^+ this holds at any cluster size).

We see that all the quantities except the hybridization bath are causal. At $U/D = 1$, there is a slight violation of (A2b) in the second derivative of Δ in NCS, PCDMFT and DCA^+ , but it is a tiny effect. In this regime, small fluctuations in the noncausal direction do not cause problems for these methods and the result is correct. However, it is clear that these methods do not impose causality on the hybridization function strictly, which then leads to problems at strong coupling. At $U/D = 2$, we see a strong violation of (A2b) in NCS, a clear inflection point in $\Delta(\tau)$ in DCA^+ , and in PCDMFT there is an inflection point in $\partial_\tau^2 \Delta(\tau)$. Here we observe a similar trend in DCA^+ , NCS and PCDMFT: $\Delta_{00}(\tau \sim \beta/2)$ is generically overestimated (by absolute value) with respect to DCA and CDMFT, respectively (note that the difference in the bath between DCA/DCA^+ on one side and CDMFT/PCDMFT/NCS on the other is due to a different way of closing self-consistency in these two groups of methods: \mathbf{k} -space versus \mathbf{r} -space clusters; see Appendix D). The bigger $\Delta_{00}(\tau \sim \beta/2)$ translates to having a bigger bath at the low frequency—the observed noncausal bath is also bigger, and as we see in Fig. 2 in the main text, in NCS it does not even decay with increasing cluster size.

3. Cumulant versus self-energy nesting

In Fig. 6, we compare the two variants of the NCS: one embeds either the cumulant g , or the self-energy Σ (for details see Appendix C4). The results are compared to a 50-site DCA calculation. The temperature is $T/D = 0.0625$ and the (hole) doping is 20%.

We present results for the simplest dimer calculation (2×1 , see Appendix C5a), the double dimer $2 + 2$ [see Eq. (C42) and the corresponding section], and the 2×2 calculation (Appendix C5c). We see that the result is solid already at $2 + 2$, and is overall improved at 2×2 . However, it is clearly not yet converged, and looking at the nonlocal part, the convergence is not monotonic. This is clearly expected at such small cluster size.

We observe that the cumulant variant performs slightly better, but the difference is almost negligible. We have checked that none of the features of the failure of NCS depend on the choice of the nested quantity (g or Σ). In the problematic region, the cumulant variant converges to almost exactly the same wrong solution as the self-energy variant.

4. Unstable and unphysical solutions

In Fig. 7, we present the self-energy for the apparently unstable (green line) and the stable solution (red line) in NCS 4×4 , compared to the exact benchmark (from Fig. 3). We observe that the unstable solution is in excellent agreement with the exact benchmark, even better than DCA of the same size cluster. The stable solution on the other hand, is much more metallic and much more local. However, it does have the correct asymptotics and is apparently causal (see Appendix A2).

Even in the large cluster limit, NCS does not guarantee $\Sigma^{\text{imp}} \rightarrow \Sigma^{\text{latt}}$, and therefore at large cluster size, a principal

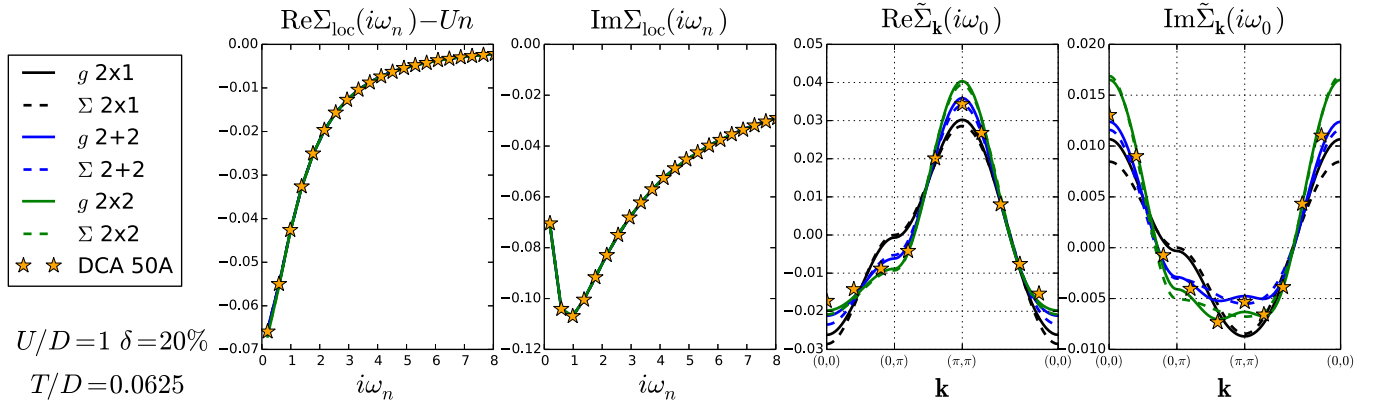


FIG. 6. Benchmark of the two variants of NCS (cumulant vs self-energy embedding) at small-cluster sizes. Stars denote the best available result (50-site DCA). The cumulant variant performs only slightly better. 2×2 calculation is already very close to the benchmark data.

solution is possible,

$$\Sigma^{\text{imp},C}[\Delta^C] + \Delta^C \approx \mathcal{F}^C \Sigma^{\text{lat}}[\Sigma^{\text{imp}}], \quad (\text{A3})$$

where \mathcal{F}^C projects a lattice quantity onto impurity degrees of freedom of the cluster C . We check this explicitly in our unphysical solution in Fig. 8 and find excellent agreement.

APPENDIX B: VERTEX DIVERGENCES

The irreducible vertex function Γ_r contains all possible two-particle scattering processes that are two-particle irreducible [2,3,85] in the given channel r (see Fig. 9 for an illustration of the two-particle reducibility concept). The reducibility channels are particle-hole (ph), transverse particle-hole (ph), and particle-particle (pp), depending on which of the external indices remain connected after cutting two propagator lines [85]. Γ_{ph} in particular corresponds to the second-order functional derivative of the LWF:

$$\begin{aligned} \Gamma_{\text{ph},\alpha\beta\gamma\delta} &= \left. \frac{\delta \Sigma_{\delta\gamma}[G]}{\delta G_{\alpha\beta}} \right|_{G=G[G_0,\Sigma]} \\ &= \left. \frac{\delta^2 \Phi[G]}{\delta G_{\alpha\beta} \delta G_{\gamma\delta}} \right|_{G=G[G_0,\Sigma]}. \end{aligned} \quad (\text{B1})$$

The greek indices combine the orbital and spin index and the imaginary time, e.g., $\alpha = (i_\alpha, \sigma_\alpha, \tau_\alpha)$. This relation is illustrated in Fig. 10 for diagrams of the second order.

The connection between Γ and Φ is the reason why Γ is sensitive to the multivaluedness of the LWF: it diverges along the lines in the phase diagram where two branches of the LWF cross [61] [see Fig. 1(a) in the main part]. However, note that also Γ_{pp} can diverge in some cases [63,64]. One can define the irreducible vertex function in the “charge” channel as $\Gamma_c = \Gamma_{\text{ph},\uparrow\uparrow\uparrow} + \Gamma_{\text{ph},\uparrow\uparrow\downarrow}$ where we have omitted the time/frequency and orbital indices for clarity.

In this paper, we are interested in identifying divergences of Γ_c . It does not appear explicitly in the cluster DMFT equations, so we only need to calculate it at the end of the self-consistency loop. Note that due to the LWF construction of the methods, we calculate it only from the correlation functions on the impurity.

1. The Bethe-Salpeter equation

In general, Γ_{ph} can be calculated from the Green’s function G and the four-point correlation function

$$\begin{aligned} \chi_{4,\sigma\sigma',ijkl}^{\omega\omega'\Omega} &= \frac{1}{\beta} \langle c_{i,\sigma}^\dagger(\omega) c_{j,\sigma}(\omega + \Omega) c_{k,\sigma'}^\dagger(\omega' + \Omega) c_{l,\sigma'}(\omega') \rangle \\ &\quad - G_{ji}(\omega) G_{lk}(\omega') \beta \delta_{\Omega,0} \\ &\quad + \delta_{\sigma,\sigma'} G_{li}(\omega) G_{jk}(\omega + \Omega) \beta \delta_{\omega,\omega'}, \end{aligned} \quad (\text{B2})$$

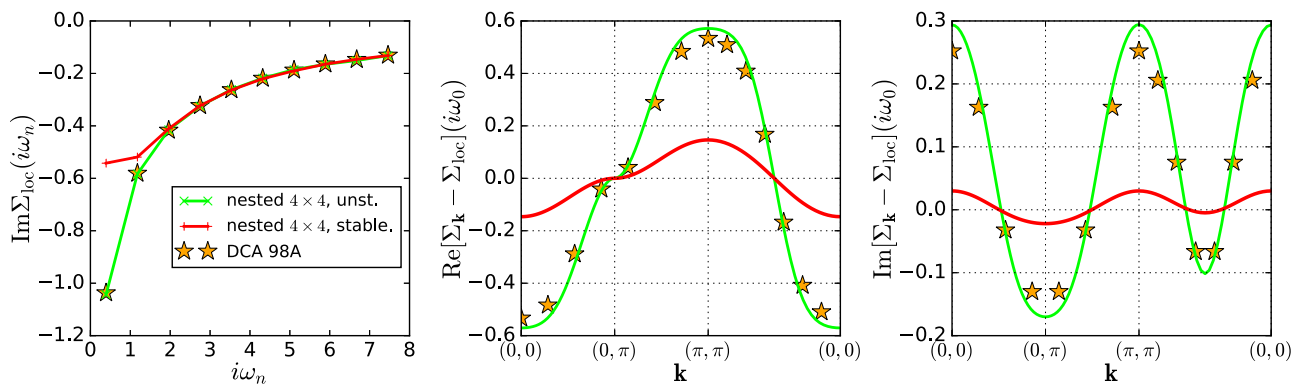


FIG. 7. Temperature $T/D = 0.125$, $U/D = 2$, $\delta = 0$, NCS 4×4 . Red line: converged solution (stable, unphysical); green line: solution after 3 iterations (almost converged, physical, unstable); stars: DCA 98A (exact benchmark).

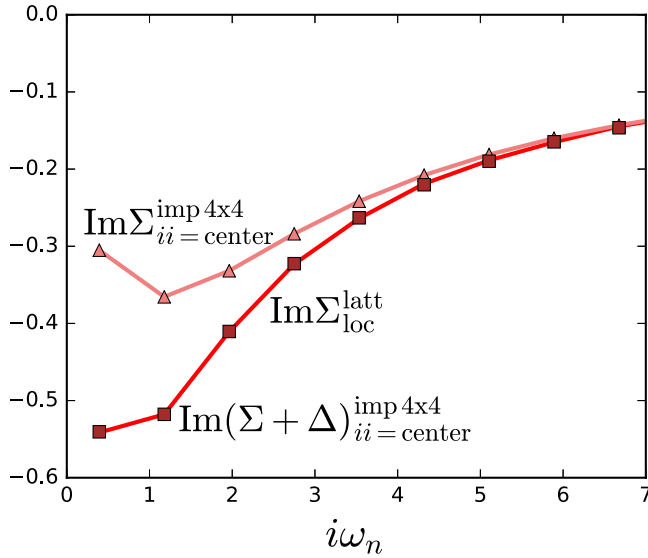


FIG. 8. Temperature $T/D = 0.125$, $U/D = 2$, $\delta = 0$, NCS 4×4 . The unphysical solution retains a finite hybridization function even in the large cluster limit, such that $\Sigma^{\text{latt}} = \Sigma^{\text{imp}} + \Delta$. Quantities are presented at the center of the biggest cluster.

where we have assumed $SU(2)$ symmetry and absence of spin-orbit interactions. First, we calculate the general χ_4 , and then calculate the charge channel simply via

$$\chi_{4,c} = \chi_{4,\uparrow\uparrow} + \chi_{4,\uparrow\downarrow}. \quad (\text{B3})$$

From this object, one can obtain the full vertex function F_c , which contains *all* the possible two-particle scattering processes (including the reducible ones). It is identical to the four-point correlation function with amputated incoming/outgoing two-point propagators

$$F_{c,ijkl}^{\omega\omega'\Omega} = \sum_{mnop} G_{mi}^{-1}(\omega) G_{ok}^{-1}(\omega' + \Omega) \chi_{4,c,mnop}^{\omega\omega'\Omega} \times G_{jn}^{-1}(\omega + \Omega) G_{lp}^{-1}(\omega'). \quad (\text{B4})$$

Γ_c is linked to F by the corresponding Bethe-Salpeter equation (BSE). The BSE can be understood as a Dyson Equation at the two-particle level [86], and it reads

$$F_{c,ijkl}^{\omega\omega'\Omega} = \Gamma_{c,ijkl}^{\omega\omega'\Omega} - \frac{1}{\beta} \sum_{\omega''} \sum_{mnop} \Gamma_{c,ijmn}^{\omega\omega''\Omega} G_{on}(\omega'') \times G_{mp}(\Omega + \omega'') F_{c,opkl}^{\omega''\omega'\Omega}. \quad (\text{B5})$$

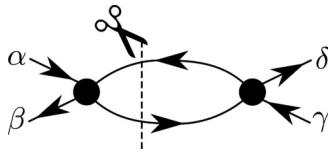


FIG. 9. Scattering diagrams can be classified according to their two-particle reducibility. If after cutting two Green's function lines, the diagram is separated into two vertex diagrams, with the external indices α and β in one and γ and δ in the other, the diagram is reducible in the ph channel.

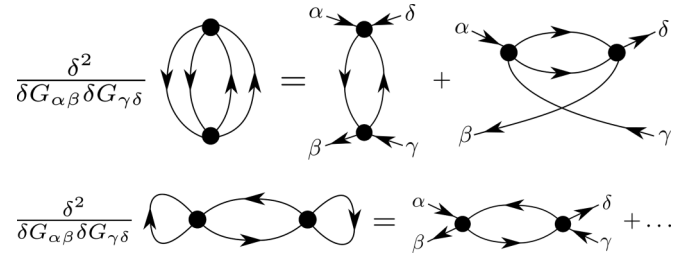


FIG. 10. (Top) Second-order functional derivative of the second-order contribution to Φ generates diagrams reducible in the pp and $\bar{p}\bar{h}$ channel. (Bottom) The ph-reducible diagrams can only be obtained by functional derivatives of nonskeleton vacuum diagrams which are not found in Φ .

The diagrammatic representation of BSE is presented in Fig. 11.

One can invert the BSE to obtain a closed expression for Γ . After rewriting the vertex quantities as matrices with respect to the properly grouped indices for a given transfer frequency Ω , $\hat{V}_{(i,j,\omega),(l,k,\omega')}^{\omega\omega'\Omega} = V_{ijkl}^{\omega\omega'\Omega}$, the BSE becomes a matrix equation

$$\hat{F}_c^\Omega = \hat{\Gamma}_c^\Omega + \frac{1}{\beta^2} \hat{\Gamma}_c^\Omega \hat{\chi}_0^\Omega \hat{F}_c^\Omega, \quad (\text{B6})$$

where

$$\chi_{0,ijkl}^{\omega\omega'\Omega} = -G_{li}(i\omega) G_{jk}(i\omega + i\Omega) \beta \delta_{\omega,\omega'}. \quad (\text{B7})$$

A few algebraic steps then yield

$$\hat{\Gamma}_c^\Omega = \beta^2 [[\hat{\chi}_0^\Omega]^{-1} - [\tilde{\chi}_c^\Omega]^{-1}], \quad (\text{B8})$$

where we have defined the so-called generalized susceptibility [85]

$$\tilde{\chi}_c = \hat{\chi}_0 + \hat{\chi}_{4,c}. \quad (\text{B9})$$

The matrix $\hat{\chi}_0^\Omega$ is always invertible. This does not necessarily hold for the generalized susceptibility $\tilde{\chi}_c$. As it approaches a singular matrix, Γ_c diverges.

While the analysis of Γ divergences can be performed for an arbitrary transfer frequency Ω , we here focus only on the $\Omega = 0$ case. $\tilde{\chi}_c^{\Omega=0}$ is a symmetric matrix. In a single-site model at particle-hole symmetry, it is also purely real, which makes it Hermitian, and its eigenvalues purely real. In cluster-impurity models, and/or away from ph symmetry, it can have complex elements, and its eigenvalues are no longer necessarily real [65].

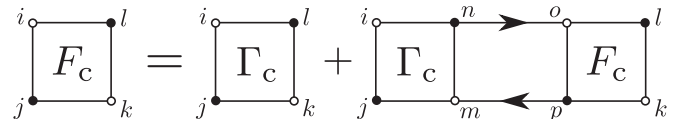


FIG. 11. Diagrammatic representation of the Bethe-Salpeter equation in the charge channel. Empty circles denote an ingoing connector of a vertex function, while black dots correspond to outgoing ones.

2. Eigenvalues and divergences

In this part, we present the procedure for determining the divergence trajectories of the irreducible vertex function, Γ_c . At a fixed temperature $T/D = 0.125$, in the (δ, U) phase diagram discussed in the main text, we determine trajectories $U^{d=1,2,3,\dots}(\delta)$, where d indexes different divergences, counting from the low U ($U^d < U^{d'}$, $d < d'$).

From Eq. (B8), it is clear that $\Gamma_c^{\Omega=0}$ diverges when an eigenvalue of $\tilde{\chi}_c^{\Omega=0}$ goes through zero. The dimension of the matrix is $M = N_\omega \times N_c^2$, where N_ω is the number of fermionic frequencies stored, and N_c is the number of sites in the cluster. We start by solving the eigenproblem for $\tilde{\chi}_c^{\Omega=0}$. We fully diagonalize this matrix at each discrete value of U (U_l) at a fixed doping δ , and obtain a set of M eigenvectors and eigenvalues $\{(\mathbf{v}_i^l, \varepsilon_i^l)\}_{i \in [0, M]}$. In single-site DMFT, at the lowest U , the real part of all eigenvalues is positive ($\text{Re}\varepsilon_i > 0$). Therefore, as we iterate over the interaction values U_l , it is straightforward to detect when the real part of an eigenvalue crosses zero—it is whenever a new eigenvalue with the negative real part appears. However, with this simple method, the error bar for $U^d(\delta)$ is given by the interaction step $U_{l+1} - U_l$. Furthermore, this method could potentially miss an event where between two U_l 's two eigenvalues cross zero, one becoming negative, the other one positive. This is particularly important in CDMFT 2×2 where there are many negative eigenvalues present already at the lowest U . Furthermore, we would like to know the exact value of the imaginary part of the eigenvalue ($\text{Im}\varepsilon_i$) when its real part is crossing zero—if it is nonzero ($\text{Im}\varepsilon_i \neq 0$), $U^d(\delta)$ at that point does not correspond to an actual divergence of Γ .

One can do better by connecting the eigenvalues ε_i^l according to matching eigenvectors and then interpolating $\varepsilon_i^l \rightarrow \varepsilon_i(U)$. $U^d(\delta)$ is then defined by $\text{Re}\varepsilon_i(U^d) = 0$. We start from the lowest U ($l = 0$), and for each eigenvector \mathbf{v}_i^l we search for an eigenvector \mathbf{v}_j^{l+1} , such that $|\mathbf{v}_i^l \cdot \mathbf{v}_j^{l+1}|$ is maximal. After this is done for all eigenvectors \mathbf{v}_i^l , one proceeds with the next l until all the eigenvector/eigenvalue pairs are connected across the entire range of U . This procedure is, however, not entirely straightforward, especially when the step in U is big. The eigenbasis rotates with changing U , and in a given U step, different eigenvectors may “exchange.” In the single-site DMFT calculation, we had to additionally require that ε_i^l is smooth to avoid getting eigenvalues mixed up. In CDMFT 2×2 , the vector space is much bigger and we encountered no such problems. Note also that, as doping is changed, the eigenvectors change considerably, and we were unable to reliably connect the eigenvalues at the same U , but different values of doping.

In Fig. 12, we present the results from the single-site DMFT calculation. Here we have data at $\delta = 0, 2\%, 6\%, 10\%$, and 16% . On the top left panel, results for $U^d(\delta)$ are presented with colored circles; the color represents the imaginary part of the eigenvalue crossing zero (color code is in the inset). The dashed lines are guides for the eye, and are also presented on Fig. 1(a) in the main text. The total count of negative eigenvalues as a function of U_l is given on the top right panel. We see that at $\delta = 0, 2\%$, eigenvalues cross zero one by one. Then at $\delta = 6\%$ and 10% , we see that two eigenvalues cross zero in the same U step. In the bottom panels, we plot the interpolation $\varepsilon_i(U)$ obtained after connecting the

eigenvalues at different values of U . We present only the first two eigenvalues to cross zero in the examined range of U . We note that these eigenvalues are the highest valued ones at the lowest U . At $\delta = 16\%$, no eigenvalues have the real part cross zero, and instead we present the two mutually complex conjugate eigenvalues, which are the biggest ones at the lowest U , and thus apparently correspond to the two eigenvalues crossing zero at the lower δ 's. We see that at low δ we have two separate eigenvalues which are purely real and cross zero at different values of U . Then at 6% doping, the two eigenvalues crossing zero are mutually complex conjugate, and cross zero at the same time, but with finite imaginary parts of opposite signs. As doping is further increased, the two eigenvalues remain mutually complex conjugate and have the real part grow towards positive values such that at $\delta = 16\%$ they no longer cross zero. The imaginary part grows with both doping and interaction.

In Fig. 13, we present the result from CDMFT 2×2 . We show the result for the first eight eigenvalues to cross zero. These are separated in two groups of four (yielding $U^d(\delta)$ with $d = 1-4$ and $d = 5-8$), and each group apparently corresponds to one of the two eigenvalues crossing zero in single-site DMFT. At higher U , there is another group of four eigenvalues crossing zero ($d = 8-11$, not shown), apparently corresponding to the 3rd divergence in single-site DMFT. The two middle eigenvalues in all groups are mutually complex conjugate [the ones yielding $U^d(\delta)$ with $d = 2, 3$, $d = 6, 7$, and $d = 10, 11$]. At $\delta = 3\%$, we see that the first two groups merge at around $U = 1.8$ ($d = 1$ with $d = 5$, 2 with $d = 6$, and so on). This point is denoted with the vertical gray dashed line. The merging of eigenvalues occurs at different U for various dopings, along the gray dashed line on the phase diagram in the inset. At $\delta = 8\%$, there are still eight eigenvalues crossing zero, but they have only three distinct real parts: first and last doubly degenerate and the middle one is four times degenerate. After merging, the imaginary part of the eigenvalues grows from zero, both with U and δ , similarly to the single-site DMFT case.

Note that we have performed the analogous analysis also in DCA, DCA^+ , and PCDMFT. The overall picture is very similar. The only qualitative difference is the presence of additional crossings of zero at low U in DCA/ DCA^+ . These, however, occur with a very big imaginary part and do not correspond to singular behavior of Γ .

APPENDIX C: NESTED CLUSTER SCHEME

In this section, we present the fully general formalism of the self-energy embedding theory (SEET) and then focus on its application to infinite lattice systems (NCS). The main idea is to approximate the Luttinger-Ward functional Φ (LWF) by a sum of functionals, including counter terms to cancel double counting of diagrams. By now it is clear that combining different LW functionals is a very general approach, and can lead to a great variety of approximations. For example, one can rederive within the SEET framework also the GW+EDMFT method [75,87]. Moreover, CDMFT can be viewed as a special case of NCS, where no counter terms are needed in the construction of the LWF.

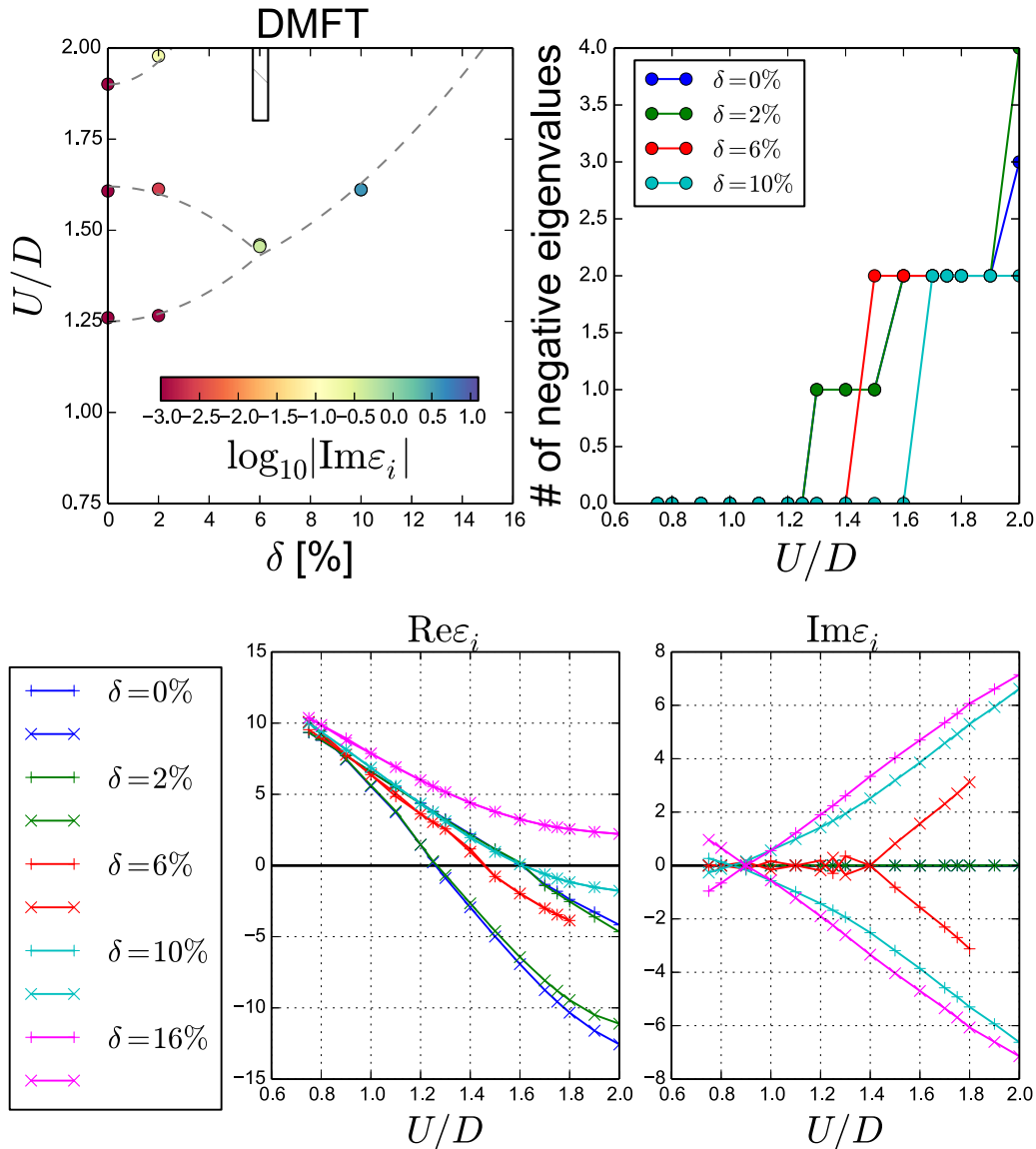


FIG. 12. Single-site DMFT calculation. Temperature $T/D = 0.125$, square lattice. Upper left: the colored points denote where the real part of an eigenvalue of $\tilde{\chi}_c(i\Omega = 0)$ crosses zero; the color denotes the imaginary part at that point, with respect to the color bar in the inset; gray dashed lines are eye-guides, used also in the Fig. 1 in the main text. Upper right: the number of negative eigenvalues; at $\delta > 5\%$ two eigenvalues cross zero at the same time. Bottom panels: evolution of the first two eigenvalues crossing zero, with U and δ . They remain purely real before becoming mutually complex conjugate. No eigenvalues cross zero at $\delta = 16\%$.

We develop a general algorithm to obtain NCS based LWF approximations and the corresponding self-energy expressions with no doubly counted diagrams, given a set of independent clusters one wishes to solve. Also, it was not clear previously whether pushing the cluster size will also increase the number of impurity problems one needs to solve. Here we prove that in the simplest scheme (square clusters), one needs to solve only three impurity problems, regardless of the cluster size.

1. General formulation

Consider a system with single-particle degrees of freedom $i \in \mathcal{L}$. At this point these may be lattice sites, or more general orbitals, and the system may or may not be infinite. The exact Luttinger-Ward functional depends on all the components of

the Green's function

$$\Phi[G] \equiv \Phi[\{G_{ij}\}_{i,j \in \mathcal{L}}]. \quad (\text{C1})$$

Consider now an approximation of the Luttinger-Ward functional, such that it is a sum of functionals, each depending on components of G that connect only a certain subset C of orbitals i , i.e., components of G within a ‘‘cluster’’ $C \subset \mathcal{L}$:

$$\Phi \approx \sum_{C \in \mathcal{C}} \Phi_C[G|_C] \quad (\text{C2})$$

with

$$G|_C \equiv \{G_{ij}\}_{ij \in C}, \quad (\text{C3})$$

where $|_C$ denotes the restriction of the orbital-space domain of the Green's function to the cluster C . It is assumed that the

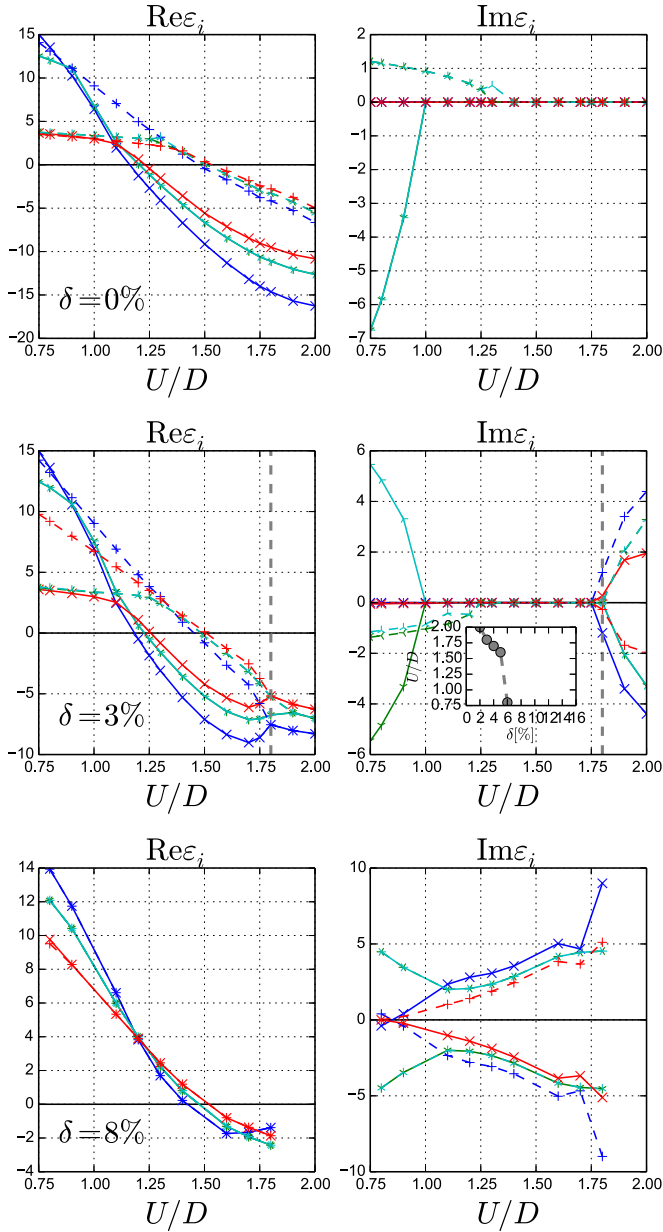


FIG. 13. 2x2 CDMFT calculation. Temperature $T/D = 0.125$, square lattice. Evolution of first eight eigenvalues to cross zero, with U and δ .

clusters are mutually independent,

$$C \not\subseteq C', \forall C, C' \in \mathcal{C},$$

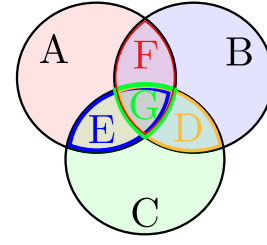
and cover the entire system

$$\bigcup_{C \in \mathcal{C}} C = \mathcal{L}.$$

However, if any of the clusters are overlapping,

$$\exists C, C' \in \mathcal{C} : C \cap C' \neq \{\}, \quad (\text{C4})$$

then we are double-counting diagrams constructed entirely from G components connecting the orbitals present in both C and C' . To avoid this, we need to add functional counter terms, each dependent only on G_{ij} within an overlap of clusters in \mathcal{C} .



$$\begin{aligned} \mathcal{C} &= \{A, B, C\} \\ \mathcal{O} &= \{D, E, F, G\} \\ \mathcal{U} &= \{A, D, G\} \end{aligned}$$

FIG. 14. Illustration of reasoning in SEET. \mathcal{C} sets of orbitals are chosen. \mathcal{O} are all possible overlaps. \mathcal{U} are sets independent by symmetry that one needs to solve in practice.

In general,

$$\Phi[G] \approx \sum_{C \in \mathcal{C}} \Phi_C[G|_C] + \sum_{C \in \mathcal{O}} p_C \Phi_C[G|_C] \quad (\text{C5})$$

where \mathcal{O} is the set of all possible overlaps between any number of nonidentical clusters in set \mathcal{C} , i.e.,

$$\mathcal{O} = \bigcup_{n \in [2, N_C]} \left\{ \bigcap_{a=1}^n C_a \right\}_{C_a \in \mathcal{C}} \setminus \mathcal{C},$$

which is illustrated on Fig. 14. N_C is the size of the set \mathcal{C} . p_C are appropriately chosen integer prefactors, possibly negative or even zero.

We emphasize that NCS is defined only by the choice of \mathcal{C} ; the set \mathcal{O} and prefactors p_C are then determined uniquely by the requirement that no diagrams are counted more than once. We can rewrite more simply

$$\Phi[G] \approx \sum_{C \in \mathcal{C} \cup \mathcal{O}} p_C \Phi_C[G|_C], \quad (\text{C6})$$

where $p_{C \in \mathcal{C}} = 1$. Hereinafter, summation \sum_C is assumed to go over $\mathcal{C} \cup \mathcal{O}$ unless stated differently. Finally, the prefactors p_C must satisfy for each C

$$\sum_{\substack{C' \in \mathcal{C} \cup \mathcal{O} \\ C' \supseteq C}} p_{C'} = 1, \quad (\text{C7})$$

which means that the contribution of diagrams that involve orbitals from a given cluster C are taken into account exactly once. In Appendix C3, we present an algorithm to find p_C which satisfy this requirement.

a. Obtaining self-energy from the functional

Anticipating that the formalism will be used for the lattice systems below, and to avoid introducing new notation, here we assume that the system is a lattice, with lattice sites $\{i\}$. Note, however, that the considerations presented here are still fully general.

The self-energy is given by the functional derivative with respect to the Green's function

$$\Sigma_{\alpha\beta} = \sum_C p_C \left. \frac{\partial \Phi[\{G_{\gamma\delta}\}_{i_\gamma, i_\delta \in C}]}{\partial G_{\beta\alpha}} \right|_{G=G[G_0, \Sigma]}, \quad (\text{C8})$$

where α, β, \dots are combined indices defined in Appendix B. We can obtain it by solving a set of impurity problems

$$S^C = \sum_{IJ \in C, \sigma} \iint d\tau d\tau' c_{\sigma, I}^+(\tau) [-(\hat{G}^C)^{-1}]_{IJ}(\tau - \tau') c_{\sigma, J}(\tau') + U \sum_{I \in C} \int d\tau c_{\uparrow, I}^+(\tau) c_{\downarrow, I}^+(\tau) c_{\downarrow, I}(\tau) c_{\uparrow, I}(\tau) \quad (C9)$$

corresponding to each cluster C , under the condition that

$$G_{IJ}^{\text{imp}C}(i\omega_n) = G_{i_C, I, j_C, J}^{\text{latt}}(i\omega_n), \quad (C10)$$

where i_C, I is the mapping between the index I of a site within a cluster C and its index i within the lattice. The Green's function on the lattice is a matrix in site indices

$$\hat{G}^{\text{latt}}(i\omega_n) = [(i\omega_n + \mu)\hat{I} - \hat{\Sigma}^{\text{latt}}]^{-1} \quad (C11)$$

and the self-energy approximation on the lattice is given by

$$\Sigma_{ij}^{\text{latt}} = \sum_{C \supseteq \{i, j\}} p_C \Sigma_{I_C, I, j_C, J}^{\text{imp}C}, \quad (C12)$$

where I_C, I is the mapping between the site index on the lattice and in the cluster C , inverse of the previously defined i_C, I . Note that up to now we have not used any lattice symmetries. Therefore, this prescription can be used for solving (finite-size) disordered and inhomogenous lattice models (e.g., one could write down a cluster extension of real-space DMFT).

b. Application to lattice models

When there are symmetries in the system Hamiltonian, one should choose \mathcal{C} in a way that does not artificially break those symmetries. For example, if there is translational symmetry on the lattice, clusters must be arranged uniformly across the entire lattice; if there is rotational symmetry, the arrangement must be the same along equivalent directions. A simple realization of a translationally and rotationally invariant set \mathcal{C} for a square lattice would include 2×2 plaquettes on all possible positions on the lattice. On the contrary, if the plaquettes are only tiled over the system, with no overlaps (as is the case in CDMFT), the translational symmetry is artificially broken.

If translational, rotational, and mirror symmetry are present, the number of clusters one actually needs to solve is reduced—one solves only one cluster of each different shape and/or size. Due to translational invariance, the position of the cluster on the lattice does not make a difference, just its shape/size. Due to rotational symmetry, quantities on clusters of the same (nonsquare) shape, but different orientation, can be inferred one from another.

Translational symmetry also allows to rewrite the lattice quantities as functions of the real-space vector rather than matrices in site index. The self-consistency condition can be rewritten as

$$G_{IJ}^{\text{imp}C}(i\omega_n) = G_{\mathbf{r}=\mathbf{r}_{i_C, I} - \mathbf{r}_{j_C, J}}^{\text{latt}}(i\omega_n). \quad (C13)$$

The Green's function on the lattice, again, is calculated from the approximated self-energy

$$G_{\mathbf{r}}^{\text{latt}} = \sum_{\mathbf{k} \in BZ} e^{-i\mathbf{k} \cdot \mathbf{r}} G_{\mathbf{k}}^{\text{latt}} = \sum_{\mathbf{k} \in BZ} \frac{e^{-i\mathbf{k} \cdot \mathbf{r}}}{G_{0, \mathbf{k}}^{-1}(i\omega_n) - \Sigma_{\mathbf{k}}^{\text{latt}}(i\omega_n)}, \quad (C14)$$

$$\Sigma_{\mathbf{k}}^{\text{latt}} = \sum_{\mathbf{r} \in BL} e^{i\mathbf{k} \cdot \mathbf{r}} \Sigma_{\mathbf{r}}^{\text{latt}}, \quad (C15)$$

which is given by a general expression

$$\Sigma_{\mathbf{r}}^{\text{latt}} = \sum_{C \in \mathcal{U}} \sum_{IJ} a_{\mathbf{r}, C, I, J} \Sigma_{IJ}^{\text{imp}C}. \quad (C16)$$

The sum runs only over a set of clusters $\mathcal{U} \subset \mathcal{C}$ independent by lattice symmetry, as illustrated on Fig. 14. If both translational and rotational symmetries are present, \mathcal{U} contains a single choice of a cluster, of each size and shape, and the sum over IJ accounts for all the shifts and rotations of the same cluster on the lattice. Note that $\sum_{C, I, J} a_{\mathbf{r}, C, I, J} = 1$ and $a_{\mathbf{r}, C, I, J} \sim \delta_{\mathbf{r}, \mathbf{r}_{i_C, I} - \mathbf{r}_{j_C, J}}$. Because some bonds on the cluster correspond to the same real-space vector and can have the same self-energy due to the symmetries of the cluster, one is free to choose which one to use, so $a_{\mathbf{r}, C, I, J}$ is not uniquely defined. More importantly, $\sum_{C, I, J} |a_{\mathbf{r}, C, I, J}| \sim N_c$. This is a problematic property of the method and is the reason why the limit $N_c \rightarrow \infty$ does not guarantee the exact solution, and is the reason for an undesirable amplification of statistical noise when clusters are big.

Large-cluster limit. As cluster size increases, the difference in self-energy between different clusters becomes smaller, and the self-energy on the clusters becomes more uniform. On the other hand, the coefficients $a_{\mathbf{r}, C, I, J}$ grow by absolute value roughly proportionally to N_c , while their total sum remains 1. This means that in the limit $N_c \rightarrow \infty$, an infinitesimal difference between the self-energies in different clusters and at different positions in the same cluster, all corresponding to the same real-space vector, can in principle be amplified such that

$$\Sigma_{\mathbf{r}=\mathbf{r}_{i_C, I} - \mathbf{r}_{j_C, J}}^{\text{latt}} - \Sigma_{IJ}^{\text{imp}C} \sim 1. \quad (C17)$$

Whether this happens or not depends on whether the coefficients $a_{\mathbf{r}, C, I, J}$ grow more quickly than do decay the difference between clusters and the inhomogeneity within them. On the other hand, in the $N_c \rightarrow \infty$ limit, we have $G_0^{\text{imp}} \rightarrow G_0$, where G_0^{imp} denotes the static part of the bare propagator on the impurity ($\mathcal{G} = [[G_0^{\text{imp}}]^{-1} - \Delta]^{-1}$). So, if $\Sigma^{\text{latt}} \neq \Sigma^{\text{imp}}$, we must have a nonzero Δ to satisfy the self-consistency condition [recall Eq. (C10)]

$$[[G_0^{\text{imp}}]^{-1} - \Delta - \Sigma^{\text{imp}}]^{-1} = [[G_0]^{-1} - \Sigma^{\text{latt}}]^{-1}. \quad (C18)$$

Because of this, NCS does not guarantee that in the $N_c \rightarrow \infty$ limit we arrive at the exact solution. A way of checking is to see whether the hybridization function falls off with increasing cluster size.

Amplification of noise. Having the property $\sum_{C, I, J} a_{\mathbf{r}, C, I, J} = 1$, when coefficients are large by absolute value, leads to amplification of QMC statistical error. The problem can be reduced by using symmetries of the clusters,

but may prove prohibitive at very large cluster sizes. On the other hand, an approximate solution not involving a stochastic impurity solver, can be safely pushed to bigger cluster sizes.

2. Square cluster case

For the special case that \mathcal{C} contains all $L \times L$ square clusters of the lattice, the nested cluster approximation for Φ can be written down explicitly for arbitrary size L . It turns out that the only overlaps $C \in \mathcal{O}$ with $p_C \neq 0$ are the clusters of shape $L - 1 \times L$, $L \times L - 1$ and $L - 1 \times L - 1$, i.e., Φ is approximated by [Eq. (4) from the main text]

$$\begin{aligned} \Phi^{(L)} = & \sum_i \Phi_{L \times L} [G|_{C_i^{L \times L}}] + \Phi_{L-1 \times L-1} [G|_{C_i^{L-1 \times L-1}}] \\ & - \Phi_{L-1 \times L} [G|_{C_i^{L-1 \times L}}] - \Phi_{L \times L-1} [G|_{C_i^{L \times L-1}}]. \end{aligned} \quad (\text{C19})$$

Here, $G|_C$ denotes the Green's function with the orbital-domain restricted to the sites within cluster C [recall Eq. (C3)]. The notation $C_i^{L_x \times L_y}$ denotes a rectangular cluster with width L_x and height L_y with its bottom left site sitting at lattice site i . In the following, we prove that $\Phi^{(L)}$ contains only the diagrams which can fit in a cluster $L \times L$, and counts each exactly once.

a. Proof of Eq. (4)

Let us consider any one diagram of Φ^{latt} in real space. This defines the (finite) set of lattice sites $D = \{\mathbf{i}\}$ contained in it. Denoting the coordinate as $\mathbf{i} = (\mathbf{i}_x, \mathbf{i}_y)$, we define

$$n = \max_{\mathbf{i} \in D} (\mathbf{i}_x) - \min_{\mathbf{i} \in D} (\mathbf{i}_x) + 1, \quad (\text{C20})$$

$$p = \max_{\mathbf{i} \in D} (\mathbf{i}_y) - \min_{\mathbf{i} \in D} (\mathbf{i}_y) + 1. \quad (\text{C21})$$

Then (n, p) is the shape of the smallest rectangular cluster containing the diagram (with $n = p = 1$ in the local case).

Let us first count the number of times the diagram appears in $\sum_i \Phi_{L_x \times L_y} [G|_{C_i^{L_x \times L_y}}]$. This count is identical to the number of ways to place a cluster of shape (n, p) into one of shape (L_x, L_y) , i.e., $f(L_x + 1 - n)f(L_y + 1 - p)$, where $f(x) = x\theta(x)$ and θ is the Heaviside function. Therefore the number of times the diagram appears in $\Phi^{(L)}$, with proper weights, is given by

$$\begin{aligned} R = & f(L + 1 - n)f(L + 1 - p) + f(L - n)f(L - p) \\ & - f(L - n)f(L + 1 - p) - f(L + 1 - n)f(L - p). \end{aligned}$$

Whenever $n \leq L$ and $p \leq L$ we have (denoting $a \equiv L + 1 - n, b \equiv L + 1 - p$)

$$\begin{aligned} R = & (L + 1 - n)(L + 1 - p) + (L - n)(L - p) \\ & - (L - n)(L + 1 - p) - (L + 1 - n)(L - p) \\ = & ab + (a - 1)(b - 1) - (a - 1)b - a(b - 1) \\ = & 1 \end{aligned}$$

while otherwise $R = 0$ by the definition of f . *QED.*

Note that even with the knowledge of p_C for all subclusters, one still needs to write down the expression for $\Sigma^{\text{latt}}[\Sigma^{\text{imp}}]$. We discuss the way this is done in the following sections, including nested schemes more general than the square cluster case discussed here.

3. Algorithm for self-energy coefficients

Deriving expressions for the self-energy when \mathcal{C} clusters are taken to be bigger than 2×2 becomes very cumbersome, and should not be done by hand. Here we present a general algorithm for a uniform and rotationally invariant arrangement of solid rectangular clusters (solid meaning there are no sites missing in the rectangle; a more general algorithm can be devised, but we don't present it here). No symmetries are assumed in the beginning, and the first part of the algorithm gives the fully general expression for $\Sigma_{ij}^{\text{latt}}$ at a given choice of ij . In the second part, the symmetries of the lattice and the clusters are used to fully simplify the expressions.

The algorithm finds the subset of clusters and the corresponding coefficients p_C that appear in the expression Eq. (C12), for a given ij on the lattice. The prefactors p_C are determined so as to satisfy Eq. (C7). The algorithm finds all the clusters in \mathcal{C} and their overlaps \mathcal{O} containing the given two sites i and j ($i = j$ allowed), orders them by size, and then assigns the prefactors starting from biggest clusters, i.e., the ones in \mathcal{C} for which we know $p_{C \in \mathcal{C}} = 1$. For the rest of the clusters C , the prefactors $p_{C'}$ of their superclusters $C' \supset C$ are taken into account to ensure that the contribution of C is taken exactly once. The procedure is “one pass” because the coefficients of smaller clusters cannot affect the coefficients for the bigger ones.

(1) Define the nested scheme by picking a set of independent rectangular clusters, defined by the size in each direction (L_x, L_y) (independent meaning no cluster can be fit into another). Note that the placement of these clusters on all possible positions on the lattice, with all possible orientations, constructs the set of clusters \mathcal{C} .

(2) For each pair of the lattice indices ij (“bond” if $i \neq j$ or “site” if $i = j$), perform the following (if you know there are symmetries, this part can be performed for only the independent bonds/sites).

(i) Determine all possible positions of all the clusters such that they contain the bond/site in question ij . These form a set of clusters defined by size and position (x, y, L_x, L_y) , and the position is assumed to correspond to the left-bottom site of the cluster.

(ii) Determine all the overlaps between the clusters obtained in the previous step. Overlaps themselves form a set of clusters defined by size and position (x, y, L_x, L_y) . Note that under present assumptions, any overlap of clusters from \mathcal{O} is also an overlap of clusters in \mathcal{C} ($C \cap C' \in \mathcal{O}, \forall C, C' \in \mathcal{O}$).

(iii) Group by shape all the clusters obtained in the previous two steps, independently of position and rotation, i.e., (x_1, y_1, L_x, L_y) goes together with (x_2, y_2, L_y, L_x) .

(iv) Order the groups according to $N_c = L_x L_y$ [or $\max(L_x, L_y)$], from biggest to smallest clusters, and place them “left to right,” so that no cluster contains a cluster to the left of it, but may or may not contain clusters to the right of it. Because clusters in the same group are of the same size, but different position and/or orientation, no cluster can contain a different cluster in the same group.

(v) Assign a prefactor $p(c)$ to each cluster c in each group.

(vi) For each group g , starting from biggest clusters (left-most). (a) For each cluster c in the group g , do a weighted count of how many times it is contained in the clusters in the groups

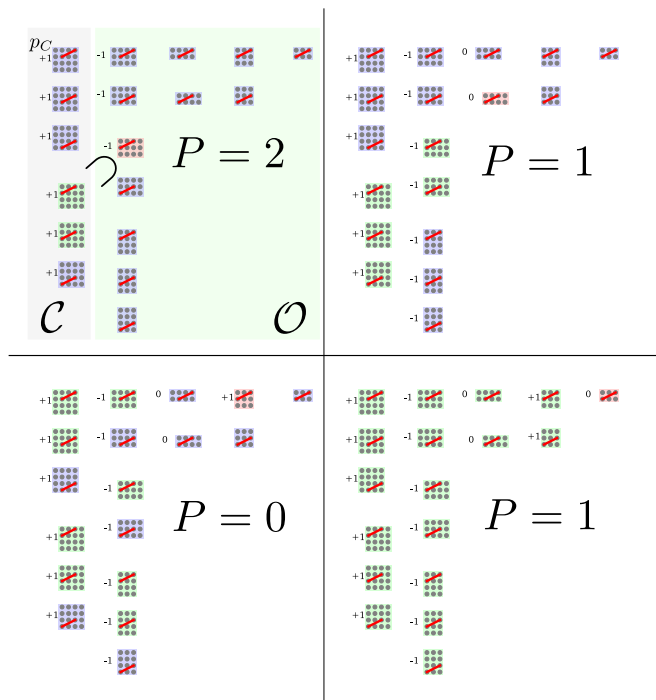


FIG. 15. Four snapshots of the algorithm for getting the self-energy coefficients; 4×4 nested scheme; coefficients are calculated for an ij bond (red line) corresponding to the real-space vector $\mathbf{r} = (2, 1)$; clusters are sorted by size; the prefactors p_C are determined starting from the biggest clusters for which we know $p_C = 1$; the red cluster is the one for which p_C is determined at the given step; green clusters are the ones that contain the red cluster. See text for details.

left to it. Weighted means to take into account the prefactor of the cluster in which c is found to be contained. In other words, obtain the number $P = \sum_{c'} p(c')$, where the sum goes over all clusters c' , which contain c . Then set the prefactor of c to be $p(c) = 1 - P$. This assures that the total contribution of the cluster c is 1. The coefficients of the smaller clusters, which are yet to change cannot affect this value. By construction, the clusters in \mathcal{C} have $p(c) = 1$; they are not contained in any other clusters, so $P = 0$.

(vii) For the bond ij , the expression for self-energy is now

$$\Sigma_{ij} = \sum_c p(c) \Sigma_{I_{c,i}, J_{c,j}}^c,$$

where c runs over all clusters in all groups, and $I_{c,i}$ and $J_{c,j}$ are determined trivially for each cluster c . The algorithm is visualized in Fig. 15.

a. Use of symmetries

When there are symmetries, we want to simplify the expression for self-energy by identifying identical contributions in the sum over clusters. First, if there is translational symmetry, clusters of the same shape but different position will have the same self-energy. If there is rotational symmetry, again, clusters of the same shape but different orientation must have the same $\hat{\Sigma}^C$. We therefore only solve clusters of different shape/size, and the sum over all C is replaced by the sum over independent clusters and a sum over the bonds IJ [recall

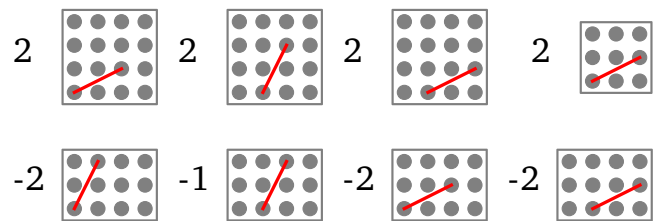


FIG. 16. After the use of symmetries, the result from the final step in Fig. 15 simplifies to what is shown in this figure. One needs to solve only three different clusters, and take into account a smaller number of bonds on each cluster, for a given real-space vector.

Eq. (C16)]. Second, $\Sigma_{IJ}^{\text{imp}C}$ may not be the same for every IJ corresponding to the same real-space vector, but clusters will in general have some symmetries, and one is able to use them to simplify the expressions further. It is straightforward to identify groups of identical bonds/sites. Then, the sum over all IJ is replaced by a sum over only the independent bonds/sites IJ on a given cluster, and the prefactors are adjusted accordingly.

Recall now the self-consistency condition in the nested cluster scheme Eq. (C13). Unlike \mathcal{G}^C and $\Sigma^{\text{imp}C}$, when convergence is reached, $G_{IJ}^{\text{imp}C}$ will be the same for any choice of IJ corresponding to the same real-space vector. We find it beneficial for the stability of the loop and the maximal level of convergence reached if this symmetry of $G^{\text{imp}C}$ is imposed in each iteration, and if the cluster symmetries are imposed on \mathcal{G}^C and $\Sigma^{\text{imp}C}$. The simplification of the self-energy expression one obtains after using cluster symmetries is visualized in Fig. 16 (see caption).

4. Nesting the cumulant

Here we discuss a different variant of the nested cluster approximation, corresponding to cumulant embedding rather than self-energy embedding theory. The benchmark of this method in comparison with the self-energy nesting variant is presented in Fig. 6.

One can in principle define a functional Γ of the Green's function such that its derivative yields the cumulant, instead of self-energy

$$\hat{g} = \frac{\partial \Gamma[\hat{G}]}{\partial \hat{G}^\top}. \quad (\text{C22})$$

The cumulant is the full-Green's function stripped of the bare hopping processes, so that

$$\hat{G}(i\omega_n) = \hat{g}(i\omega_n) + \hat{g}(i\omega_n) \hat{t} \hat{G}(i\omega_n), \quad (\text{C23})$$

i.e.,

$$\hat{G}(i\omega_n) = [\hat{g}^{-1}(i\omega_n) - \hat{t}]^{-1}, \quad (\text{C24})$$

where \hat{t} is the hopping matrix. In \mathbf{k} space,

$$G_{\mathbf{k}}(i\omega_n) = g_{\mathbf{k}}(i\omega_n) + g_{\mathbf{k}}(i\omega_n) \varepsilon_{\mathbf{k}} G_{\mathbf{k}}(i\omega_n) \quad (\text{C25})$$

$$G_{\mathbf{k}}(i\omega_n) = \frac{1}{g_{\mathbf{k}}^{-1}(i\omega_n) - \varepsilon_{\mathbf{k}}}, \quad (\text{C26})$$

which leads to the identity

$$\hat{g}(i\omega_n) = [(i\omega_n + \mu)\hat{I} - \hat{\Sigma}(i\omega_n)]^{-1} \quad (\text{C27})$$

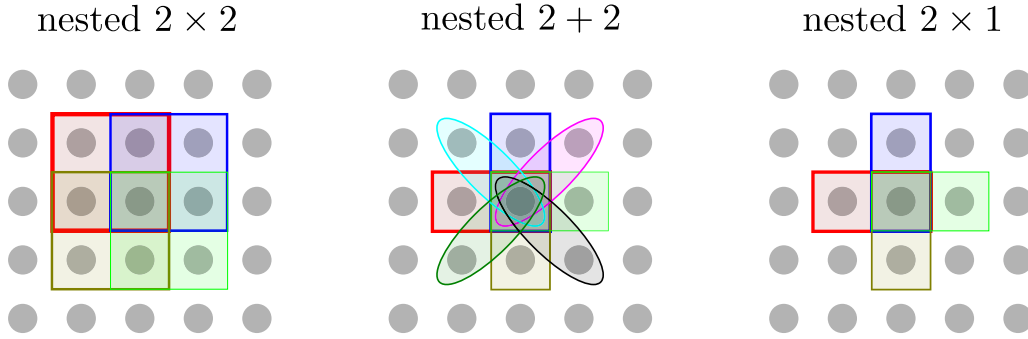


FIG. 17. Clusters contributing to the local self-energy in different kinds of simple nested schemes.

and the inverse is

$$\hat{\Sigma}(i\omega_n) = (i\omega_n + \mu)\hat{I} - \hat{g}^{-1}(i\omega_n). \quad (\text{C28})$$

So, we can construct the cumulant on the lattice $g_{\mathbf{k}}$ from the cumulants on the impurities, the same way we did for the self-energy. The self-energy on the lattice can be obtained as

$$\Sigma_{\mathbf{k}}^{\text{latt}}(i\omega_n) = i\omega_n + \mu - (g_{\mathbf{k}}^{\text{latt}}(i\omega_n))^{-1}, \quad (\text{C29})$$

but this expression is ill-defined at high frequency, so it is important to avoid using it in the DMFT loop. Therefore, in each iteration, we construct G^{latt} directly from the cumulant using Eq. (C26), and calculate the self-energy only in the post-processing of the results.

We expect that the cumulant variant works better whenever the cumulant is shorter ranged than the self-energy. In practice, we find that the cumulant version does a slightly better job, but the difference is not big (see Appendix A 3).

5. Simple examples and summary of equations

a. 2×1

This example was originally presented in Ref. [70]. We rederive it only for pedagogical purposes.

For simplicity, we introduce a shorthand notation $i+x$ to denote the index of the nearest neighbor of the site i in the $+x$ direction, and similarly $i-x, i+y, \dots$. Recall also the cluster notation $C_i^{2 \times 1} \equiv \{i, i+x\}$, $C_i^{1 \times 2} \equiv \{i, i+y\}$ and $C_i^{1 \times 1} \equiv \{i\}$.

Let us approximate LW functional such that it contains diagrams that involve at most two nearest-neighboring lattice sites

$$\mathcal{C} = \{C_i^{2 \times 1}\}_{\forall i} \cup \{C_i^{1 \times 2}\}_{\forall i}. \quad (\text{C30})$$

This means we want to solve at most a two-site impurity problem. As for all possible overlaps of the clusters in \mathcal{C} , one can easily verify

$$\mathcal{O} = \{C \cap C'\}_{\forall C, C' \in \mathcal{C}} = \{C_i^{1 \times 1}\}_{\forall i}, \quad (\text{C31})$$

which means we will need to take care of double counting. Each overlap cluster is contained in four clusters in \mathcal{C} ,

$$\begin{aligned} C_i^{1 \times 1} &\subset C_i^{2 \times 1}, & C_i^{1 \times 1} &\subset C_{i-x}^{2 \times 1}, \\ C_i^{1 \times 1} &\subset C_i^{1 \times 2}, & C_i^{1 \times 1} &\subset C_{i-y}^{1 \times 2}, \end{aligned} \quad (\text{C32})$$

which means that we are counting diagrams which involve only the local Green's function four times at each site. To have them

taken into account only once, we need to subtract the DMFT functional (D1) three times at each site, i.e., $p_{C \in \mathcal{O}} = -3$,

$$\Phi \approx \sum_i (\Phi_2[G|_{C_i^{2 \times 1}}] + \Phi_2[G|_{C_i^{1 \times 2}}] - 3\Phi_1[G|_{C_i^{1 \times 1}}]). \quad (\text{C33})$$

Now we write the clusters explicitly to perform the derivatives that yield the self-energy. The local component is given by

$$\begin{aligned} \Sigma_{ii} &= \frac{\partial}{\partial G_{ii}} \sum_l \left(\sum_{\delta \in \{x, y\}} \Phi_2[\{G_{l'm'}\}_{l'm' \in \{l, l+\delta\}}] - 3\Phi_1[G_{ll}] \right) \\ &= \sum_{\delta \in \{x, y, -x, -y\}} \frac{\partial \Phi_2[\{G_{lm}\}_{lm \in \{i, i+\delta\}}]}{\partial G_{ii}} - 3 \frac{\partial \Phi_1[G_{ii}]}{\partial G_{ii}} \end{aligned} \quad (\text{C34})$$

and the nearest-neighbor components (with $\delta = x, y$)

$$\begin{aligned} \Sigma_{i, i+\delta} &= \frac{\partial}{\partial G_{i+\delta, i}} \left(\sum_{l'} \sum_{\delta' \in \{x, y\}} \Phi_2[\{G_{lm}\}_{lm \in \{l', l'+\delta'\}}] \right. \\ &\quad \left. - 3 \sum_l \Phi_1[G_{ll}] \right) \\ &= \frac{\partial \Phi_2[\{G_{lm}\}_{lm \in \{i, i+\delta\}}]}{\partial G_{i+\delta, i}}. \end{aligned} \quad (\text{C35})$$

When there is translational, mirror, and rotational symmetry, the contribution to the local part coming from four different nearest-neighbor pairs will be the same, and the self-energy on all nearest-neighbor bonds will be the same (see Fig. 17),

$$\Sigma_{\mathbf{r}=(0,0)}^{\text{latt}} = 4\Sigma_{00}^{\text{imp } 2 \times 1} - 3\Sigma_{00}^{\text{imp } 1 \times 1}, \quad \Sigma_{\mathbf{r}=(0,1)}^{\text{latt}} = \Sigma_{01}^{\text{imp } 2 \times 1}, \quad (\text{C36})$$

and the self-consistency is

$$G_{00/11}^{\text{imp } 2 \times 1} = G_{00}^{\text{imp } 1 \times 1} = G_{\mathbf{r}=(0,0)}^{\text{latt}}, \quad G_{01/10}^{\text{imp } 2 \times 1} = G_{\mathbf{r}=(1,0)}^{\text{latt}} \quad (\text{C37})$$

b. Long-distance dimers

In this section, we present a nested cluster scheme where self-energy at an arbitrary real-space vector is approximated by the self-energy of a corresponding two-site impurity problem. The expression for Φ and Σ^{latt} can be worked out analytically.

Let us define $i+n_x+n_y$ to be the index of the lattice site at the real-space vector $\mathbf{r} = \mathbf{r}_i + n_x \mathbf{e}_x + n_y \mathbf{e}_y$. We can

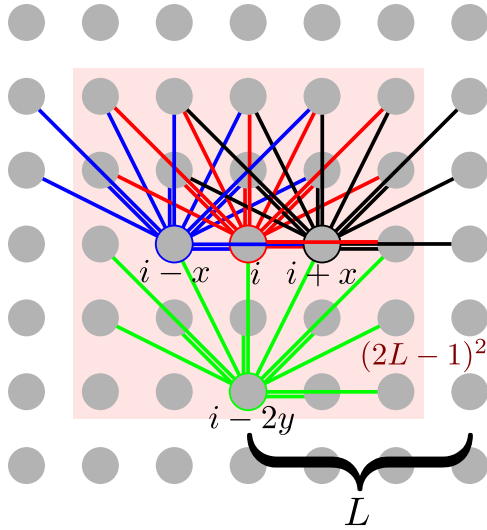


FIG. 18. Illustration of Eq. (C38). By placing the green set of dimers starting from each site, one covers all the dimers on the lattice, up to Manhattan distance $2(L - 1)$. Each site is involved in $(2L - 1)^2 - 1$ different dimers (red site goes with all the sites in the red square).

approximate the LW functional in such a way that it contains diagrams which involve at most two sites, but at a distance not bigger than $L - 1$ in both directions (maximum Manhattan distance $2L - 2$). We define the cluster notation $C_i^{(n_x, n_y)} \equiv \{i, i + n_x + n_y\}$. Analogously to Eq. (C33), one can prove the following approximation has no double counting:

$$\Phi \approx \sum_i \left(\sum_{\substack{n_x \in (-L, L) \\ n_y \in (0, L)}} + \sum_{\substack{n_x \in (0, L) \\ n_y = 0}} \right) \Phi_2[\{G|_{C_i^{(n_x, n_y)}}\}] + (2 - (2L - 1)^2) \sum_i \Phi_1[G|_{C_i^{1 \times 1}}]. \quad (\text{C38})$$

This is illustrated in Fig. 18. For example, the site with red outline is involved in dimers with all the sites within the red square, of which there are $(2L - 1)^2 - 1$.

With translational/rotational/mirror symmetry, we get for the self-energy

$$\begin{aligned} \Sigma_{\mathbf{r}=(0,0)}^{\text{latt}} &= \sum_{\substack{n_x \in (0, L) \\ n_y \in [0, n_x]}} m_{n_x, n_y} \Sigma_{00}^{\text{imp}(n_x, n_y)} \\ &+ (2 - (2L - 1)^2) \Sigma_{00}^{\text{imp} 1 \times 1}, \\ \Sigma_{\mathbf{r}=(n_x, n_y)}^{\text{latt}} &= \Sigma_{01}^{\text{imp}(n_x, n_y)}, \end{aligned} \quad (\text{C39})$$

where in the bottom row, $0 < n_x < L$, $0 \leq n_y \leq n_x$. m_{n_x, n_y} is the multiplicity of the (nonzero) real-space vector $\mathbf{r} = n_x \mathbf{e}_x + n_y \mathbf{e}_y$

$$m_{n_x, n_y} = \begin{cases} 4, & n_x = 0 \vee n_y = 0 \vee n_x = n_y \\ 8, & \text{otherwise} \end{cases} \quad (\text{C40})$$

and the self-consistency reads

$$G_{00/11}^{\text{imp}(n_x, n_y)} = G_{00}^{\text{imp} 1 \times 1} = G_{\mathbf{r}=(0,0)}^{\text{latt}}, \quad G_{01/10}^{\text{imp}(n_x, n_y)} = G_{\mathbf{r}=(n_x, n_y)}^{\text{latt}}. \quad (\text{C41})$$

The simplest example is the $2 + 2$ scheme corresponding to $L = 2$, where we just take the dimer as in the previous example, and add the diagonal one $\{i, i + x, i + y\}$ (see Fig. 17):

$$\begin{aligned} \Sigma_{\mathbf{r}=(0,0)}^{\text{latt}} &= 4 \Sigma_{00}^{\text{imp}(0,1)} + 4 \Sigma_{00}^{\text{imp}(1,1)} - 7 \Sigma_{00}^{\text{imp} 1 \times 1}, \\ \Sigma_{\mathbf{r}=(0,1)}^{\text{latt}} &= \Sigma_{01}^{\text{imp}(0,1)}, \\ \Sigma_{\mathbf{r}=(1,1)}^{\text{latt}} &= \Sigma_{01}^{\text{imp}(1,1)}. \end{aligned} \quad (\text{C42})$$

Here we are solving three impurity problems: two of them two-site and one single-site.

c. 2×2

Here we discuss the special case of the square clusters scheme presented in Appendix C2, with $L = 2$. It corresponds to placing a square 2×2 cluster on all possible positions on the lattice. The Φ approximation is given by Eq. (4) with $L = 2$.

We can write it more explicitly

$$\begin{aligned} \Phi^{(L=2)} &= \sum_i \Phi_4[\{G|_{lm \in \{i, i+x, i+y, i+x+y\}}\}] \\ &- \sum_i \sum_{\delta \in \{x, y\}} \Phi_2[\{G|_{lm \in \{i, i+\delta\}}\}] + \sum_i \Phi_1[G|_{ii}]. \end{aligned} \quad (\text{C43})$$

Now let us apply the derivative with respect to different components of the Green's function to get the expressions for self-energy:

$$\begin{aligned} \Sigma_{ii} &= \frac{\partial}{\partial G_{ii}} \left[\sum_{l \in \{i-x, i-y, i-x-y\}} \Phi_4[\{G|_{l'm' \in \{l, l+x, l+y, l+x+y\}}\}] \right. \\ &- \left. \sum_{\delta \in \{x, y\}} \sum_{l \in \{i-\delta\}} \Phi_2[\{G|_{l'm' \in \{l, l+\delta\}}\}] + \Phi_1[\{G|_{ii}\}] \right], \end{aligned} \quad (\text{C44})$$

$$\begin{aligned} \Sigma_{i, i+x} &= \frac{\partial}{\partial G_{i+x, i}} \left[\sum_{l \in \{i-y\}} \Phi_4[\{G|_{l'm' \in \{l, l+x, l+y, l+x+y\}}\}] \right. \\ &- \left. \Phi_2[\{G|_{lm \in \{i, i+x\}}\}] \right]. \end{aligned}$$

$$\Sigma_{i, i+x+y} = \frac{\partial}{\partial G_{i+x+y, i}} \Phi_4[\{G|_{lm \in \{i, i+x, i+y, i+x+y\}}\}]. \quad (\text{C45})$$

With full translational/rotational/mirror symmetry, clusters with same size and shape must give identical contributions to the self-energy. Following considerations analogous to Eqs. (C34) and (C35), we arrive at the final expression which connects the self energy on the lattice with the one in three different impurity problems (2×2 , 2×1 , and 1×1 , see Fig. 17):

$$\begin{aligned} \Sigma_{\mathbf{r}=(0,0)}^{\text{latt}} &= 4 \Sigma_{00}^{\text{imp} 2 \times 2} - 4 \Sigma_{00}^{\text{imp} 2 \times 1} + \Sigma_{00}^{\text{imp} 1 \times 1}, \\ \Sigma_{\mathbf{r}=(1,0)}^{\text{latt}} &= 2 \Sigma_{01}^{\text{imp} 2 \times 2} - \Sigma_{01}^{\text{imp} 2 \times 1}, \\ \Sigma_{\mathbf{r}=(1,1)}^{\text{latt}} &= \Sigma_{03}^{\text{imp} 2 \times 2}, \end{aligned} \quad (\text{C46})$$

and the self-consistency condition is given by

$$G_{00/11/22/33}^{\text{imp } 2 \times 2} = G_{00/11}^{\text{imp } 2 \times 1} = G_{00}^{\text{imp } 1 \times 1} = G_{\mathbf{r}=(0,0)}^{\text{latt}}, \quad G_{01/13/32/20/10/\dots}^{\text{imp } 2 \times 2} = G_{01/10}^{\text{imp } 2 \times 1} = G_{\mathbf{r}=(0,1)}^{\text{latt}}, \quad G_{03/30/12/21}^{\text{imp } 2 \times 2} = G_{\mathbf{r}=(1,1)}^{\text{latt}}. \quad (\text{C47})$$

d. 4×4

Using the algorithm (C3) and lattice symmetries, we can now automatize the derivation of expressions for the self energy. Here we present as an example the expressions for the 4×4 nested scheme, where \mathcal{C} contains all possible positions of a 4×4 cluster, and Φ approximation is given by Eq. (4) with $L = 4$.

$$\begin{array}{|c|c|c|} \hline 6 & 7 & 8 \\ \hline 3 & 4 & 5 \\ \hline 0 & 1 & 2 \\ \hline \end{array} \quad \begin{array}{|c|c|c|c|} \hline 12 & 13 & 14 & 15 \\ \hline 8 & 9 & 10 & 11 \\ \hline 4 & 5 & 6 & 7 \\ \hline 0 & 1 & 2 & 3 \\ \hline \end{array} \quad \begin{array}{|c|c|c|c|} \hline 8 & 9 & 10 & 11 \\ \hline 4 & 5 & 6 & 7 \\ \hline 0 & 1 & 2 & 3 \\ \hline \end{array}$$

$$\begin{aligned} \Sigma_{\mathbf{r}=(0,0)}^{\text{latt}} &= +4\Sigma_{00}^{\text{imp } 4 \times 4} + 8\Sigma_{11}^{\text{imp } 4 \times 4} + 4\Sigma_{55}^{\text{imp } 4 \times 4} - 8\Sigma_{00}^{\text{imp } 4 \times 3} - 4\Sigma_{44}^{\text{imp } 4 \times 3} + 4\Sigma_{00}^{\text{imp } 3 \times 3} - 8\Sigma_{11}^{\text{imp } 4 \times 3} \\ &\quad + 4\Sigma_{11}^{\text{imp } 3 \times 3} - 4\Sigma_{55}^{\text{imp } 4 \times 3} + \Sigma_{44}^{\text{imp } 3 \times 3}, \\ \Sigma_{\mathbf{r}=(1,0)}^{\text{latt}} &= +4\Sigma_{01}^{\text{imp } 4 \times 4} + 4\Sigma_{15}^{\text{imp } 4 \times 4} + 2\Sigma_{12}^{\text{imp } 4 \times 4} + 2\Sigma_{56}^{\text{imp } 4 \times 4} - 4\Sigma_{01}^{\text{imp } 4 \times 3} - 2\Sigma_{45}^{\text{imp } 4 \times 3} \\ &\quad - 4\Sigma_{04}^{\text{imp } 4 \times 3} + 4\Sigma_{01}^{\text{imp } 3 \times 3} - 4\Sigma_{15}^{\text{imp } 4 \times 3} + 2\Sigma_{14}^{\text{imp } 3 \times 3} - 2\Sigma_{12}^{\text{imp } 4 \times 3} - \Sigma_{56}^{\text{imp } 4 \times 3}, \\ \Sigma_{\mathbf{r}=(1,1)}^{\text{latt}} &= +2\Sigma_{05}^{\text{imp } 4 \times 4} + 4\Sigma_{16}^{\text{imp } 4 \times 4} + 2\Sigma_{14}^{\text{imp } 4 \times 4} + \Sigma_{510}^{\text{imp } 4 \times 4} - 4\Sigma_{05}^{\text{imp } 4 \times 3} - 4\Sigma_{14}^{\text{imp } 4 \times 3} + 2\Sigma_{04}^{\text{imp } 3 \times 3} - 4\Sigma_{16}^{\text{imp } 4 \times 3} + 2\Sigma_{13}^{\text{imp } 3 \times 3}, \\ \Sigma_{\mathbf{r}=(2,0)}^{\text{latt}} &= +4\Sigma_{02}^{\text{imp } 4 \times 4} + 4\Sigma_{19}^{\text{imp } 4 \times 4} - 4\Sigma_{02}^{\text{imp } 4 \times 3} - 2\Sigma_{46}^{\text{imp } 4 \times 3} - 2\Sigma_{08}^{\text{imp } 4 \times 3} + 2\Sigma_{02}^{\text{imp } 3 \times 3} - 2\Sigma_{19}^{\text{imp } 4 \times 3} + \Sigma_{17}^{\text{imp } 3 \times 3}, \\ \Sigma_{\mathbf{r}=(2,1)}^{\text{latt}} &= +2\Sigma_{06}^{\text{imp } 4 \times 4} + 2\Sigma_{110}^{\text{imp } 4 \times 4} + 2\Sigma_{17}^{\text{imp } 4 \times 4} - 2\Sigma_{06}^{\text{imp } 4 \times 3} - 2\Sigma_{17}^{\text{imp } 4 \times 3} - 2\Sigma_{09}^{\text{imp } 4 \times 3} + 2\Sigma_{05}^{\text{imp } 3 \times 3} - \Sigma_{110}^{\text{imp } 4 \times 3}, \\ \Sigma_{\mathbf{r}=(2,2)}^{\text{latt}} &= +2\Sigma_{010}^{\text{imp } 4 \times 4} + 2\Sigma_{111}^{\text{imp } 4 \times 4} - 4\Sigma_{010}^{\text{imp } 4 \times 3} + \Sigma_{08}^{\text{imp } 3 \times 3}, \\ \Sigma_{\mathbf{r}=(3,0)}^{\text{latt}} &= +2\Sigma_{03}^{\text{imp } 4 \times 4} + 2\Sigma_{113}^{\text{imp } 4 \times 4} - 2\Sigma_{03}^{\text{imp } 4 \times 3} - \Sigma_{47}^{\text{imp } 4 \times 3}, \\ \Sigma_{\mathbf{r}=(3,1)}^{\text{latt}} &= +2\Sigma_{07}^{\text{imp } 4 \times 4} + \Sigma_{114}^{\text{imp } 4 \times 4} - 2\Sigma_{07}^{\text{imp } 4 \times 3}, \\ \Sigma_{\mathbf{r}=(3,2)}^{\text{latt}} &= +2\Sigma_{011}^{\text{imp } 4 \times 4} - \Sigma_{011}^{\text{imp } 4 \times 3}, \\ \Sigma_{\mathbf{r}=(3,3)}^{\text{latt}} &= +\Sigma_{015}^{\text{imp } 4 \times 4}. \end{aligned} \quad (\text{C48})$$

In practice, we calculate self-energy for all vectors $\mathbf{r} = (x, y)$ such that $y \leq x$, up to $x = \max_C L_x(C) - 1$, and the rest is filled by lattice symmetry (analogously to Eq. (D17d) further below). In Eq. (C48), we have also used the symmetries of the clusters. The groups of equivalent bonds on all three clusters are given below in curly brackets (inversion symmetry $ij = ji$ is implicit):

$$3 \times 3 : \begin{array}{|c|c|c|} \hline 6 & 7 & 8 \\ \hline 3 & 4 & 5 \\ \hline 0 & 1 & 2 \\ \hline \end{array}$$

$$\begin{aligned} &\{(0,0), (2,2), (6,6), (8,8)\}; \quad \{(0,1), (0,3), (2,1), (2,5), (6,3), (6,7), (8,5), (8,7)\}; \\ &\{(0,2), (0,6), (2,8), (6,8)\}; \quad \{(0,4), (2,4), (6,4), (8,4)\}; \quad \{(0,5), (0,7), (2,3), (2,7), (6,1), (6,5), (8,1), (8,3)\}; \\ &\{(0,8), (2,6)\}; \quad \{(1,1), (3,3), (5,5), (7,7)\}; \quad \{(1,3), (1,5), (3,7), (5,7)\}; \\ &\{(1,4), (3,4), (5,4), (7,4)\}; \quad \{(1,7), (3,5)\}; \quad \{(4,4)\}. \end{aligned} \quad (\text{C49})$$

$$4 \times 3 : \begin{array}{|c|c|c|c|} \hline 8 & 9 & 10 & 11 \\ \hline 4 & 5 & 6 & 7 \\ \hline 0 & 1 & 2 & 3 \\ \hline \end{array}$$

$$\begin{aligned} &\{(0,0), (3,3), (8,8), (11,11)\}; \quad \{(0,1), (3,2), (8,9), (11,10)\}; \quad \{(0,2), (3,1), (8,10), (11,9)\}; \\ &\{(0,4), (3,7), (8,4), (11,7)\}; \quad \{(0,5), (3,6), (8,5), (11,6)\}; \quad \{(0,6), (3,5), (8,6), (11,5)\}; \\ &\{(0,7), (3,4), (8,7), (11,4)\}; \quad \{(0,8), (3,11)\}; \quad \{(0,3), (8,11)\}; \quad \{(0,9), (3,10), (8,1), (11,2)\}; \\ &\{(0,10), (3,9), (8,2), (11,1)\}; \quad \{(0,11), (3,8)\}; \quad \{(1,2), (9,10)\}; \quad \{(1,1), (2,2), (9,9), (10,10)\}; \\ &\{(1,4), (2,7), (9,4), (10,7)\}; \quad \{(1,5), (2,6), (9,5), (10,6)\}; \quad \{(1,6), (2,5), (9,6), (10,5)\}; \end{aligned}$$

$$\begin{aligned} & \{(1,7), (2,4), (9,7), (10,4)\}; \{(1,9), (2,10)\}; \{(1,10), (2,9)\}; \{(4,4), (7,7)\}; \{(4,5), (7,6)\}; \\ & \{(4,6), (7,5)\}; \{(4,7)\}; \{(5,5), (6,6)\}; \{(5,6)\}. \end{aligned} \quad (\text{C50})$$

$$4 \times 4: \begin{array}{|c|c|c|c|} \hline 12 & 13 & 14 & 15 \\ \hline 8 & 9 & 10 & 11 \\ \hline 4 & 5 & 6 & 7 \\ \hline 0 & 1 & 2 & 3 \\ \hline \end{array}$$

$$\begin{aligned} & \{(0,0), (3,3), (12,12), (15,15)\}; \{(0,1), (0,4), (3,2), (3,7), (12,8), (12,13), (15,11), (15,14)\}; \\ & \{(0,2), (0,8), (3,1), (3,11), (12,4), (12,14), (15,7), (15,13)\}; \{(0,3), (0,12), (3,15), (12,15)\}; \\ & \{(0,5), (3,6), (12,9), (15,10)\}; \{(0,6), (0,9), (3,5), (3,10), (12,5), (12,10), (15,6), (15,9)\}; \\ & \{(0,7), (0,13), (3,4), (3,14), (12,1), (12,11), (15,2), (15,8)\}; \{(0,10), (3,9), (12,6), (15,5)\}; \\ & \{(0,11), (0,14), (3,8), (3,13), (12,2), (12,7), (15,1), (15,4)\}; \{(0,15), (3,12)\}; \\ & \{(1,1), (2,2), (4,4), (7,7), (8,8), (11,11), (13,13), (14,14)\}; \{(1,2), (4,8), (7,11), (13,14)\}; \\ & \{(1,4), (2,7), (8,13), (11,14)\}; \{(1,5), (2,6), (4,5), (7,6), (8,9), (11,10), (13,9), (14,10)\}; \\ & \{(1,6), (2,5), (4,9), (7,10), (8,5), (11,6), (13,10), (14,9)\}; \{(1,7), (2,4), (4,13), (7,14), (8,1), (11,2), \\ & (13,11), (14,8)\}; \{(1,9), (2,10), (4,6), (7,5), (8,10), (11,9), (13,5), (14,6)\}; \\ & \{(1,10), (2,9), (4,10), (7,9), (8,6), (11,5), (13,6), (14,5)\}; \\ & \{(1,11), (2,8), (4,14), (7,13)\}; \{(1,13), (2,14), (4,7), (8,11)\}; \{(1,14), (2,13), (4,11), (7,8)\}; \\ & \{(5,5), (6,6), (9,9), (10,10)\}; \{(5,6), (5,9), (6,10), (9,10)\}; \{(5,10), (6,9)\}. \end{aligned} \quad (\text{C51})$$

APPENDIX D: CLUSTER DMFT METHODS

Here we summarize the (cluster) DMFT methods used in this paper.

The forward-substitution algorithm for the generic cluster DMFT scheme is presented in Fig. 19. Cluster DMFT methods differ in the cluster-impurity action, self-consistency condition, and the self-energy mapping $\Sigma^{\text{latt}}[\Sigma^{\text{imp}}]$ —these properties we state for each method in the following sections. Where possible, we also state the LW functional approximation which leads to the given method.

1. Single-site DMFT

Single-site DMFT [5] is the limiting case of all cluster DMFT methods, corresponding to cluster size $N_c = 1$. It can be derived as the local approximation of the LW functional. While the exact LW functional depends on all components of the Green's function, in DMFT it depends only on the local components G_{ii} :

$$\Phi[\{G_{ij}\}_{\forall i,j}] \approx \Phi[\{G_{ii}\}_{\forall i}] = \sum_i \Phi[G_{ii}]. \quad (\text{D1})$$

The second step is specific to local interactions, and is crucial to obtain a self-consistent scheme involving a single-site impurity problem.

The impurity action involves degrees of freedom of a single lattice site

$$\begin{aligned} S = & \sum_{\sigma} \iint d\tau d\tau' c_{\sigma}^{\dagger}(\tau) [-\mathcal{G}_{\sigma}^{-1}](\tau - \tau') c_{\sigma}(\tau') \\ & + U \int d\tau c_{\uparrow}^{\dagger}(\tau) c_{\downarrow}^{\dagger}(\tau) c_{\downarrow}(\tau) c_{\uparrow}(\tau). \end{aligned} \quad (\text{D2})$$

The self-consistency condition requires that the local Green's function on the lattice is the same as the one on the impurity,

$$G^{\text{imp}}(i\omega_n) = G_{ii}^{\text{latt}}(i\omega_n) \equiv \sum_{\mathbf{k} \in \text{BZ}} G_{\mathbf{k}}^{\text{latt}}(i\omega_n), \quad (\text{D3})$$

where

$$G_{\mathbf{k}}^{\text{latt}}(i\omega_n) = (G_{0,\mathbf{k}}^{-1}(i\omega_n) - \Sigma_{\mathbf{k}}^{\text{latt}}(i\omega_n))^{-1}. \quad (\text{D4})$$

The self-energy approximation reads

$$\Sigma_{\mathbf{k}}^{\text{latt}}(i\omega_n) \approx \Sigma^{\text{imp}}(i\omega_n) \quad (\text{D5})$$

as $\partial \sum_l \Phi[G_{ll}]/\partial G_{ij} \sim \delta_{ij} \partial \Phi[G_{ii}]/\partial G_{ii}$. The bare-propagator on the lattice

$$G_{0,\mathbf{k}}(i\omega_n) = \frac{1}{i\omega_n + \mu - \varepsilon_{\mathbf{k}}} \quad (\text{D6})$$

is determined by the chemical potential μ and the bare dispersion $\varepsilon_{\mathbf{k}}$. On the square-lattice with only nearest-neighbor hopping, it is given by

$$\varepsilon_{\mathbf{k}} = -2t(\cos k_x + \cos k_y). \quad (\text{D7})$$

2. Cellular DMFT (CDMFT)

Cellular DMFT rewrites the lattice problem in terms of supercells, as illustrated on Fig. 20 [9,11–26,88]. The lattice-site index is replaced by a double index: the index of the supercell and the index of the site within the supercell:

$$i \rightarrow (i, I), \quad G_{ij} \rightarrow G_{iI, jJ}.$$

We denote with i, j, \dots the index of the supercell, and with I, J, \dots the index of the site within the supercell. From here, the derivation proceeds just as in single-site DMFT—one may

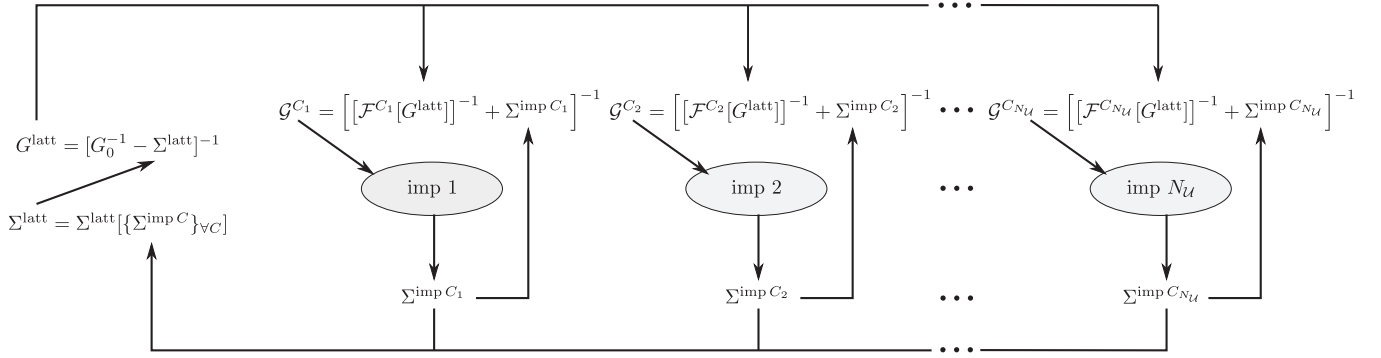


FIG. 19. Forward substitution solution of a generic cluster DMFT method. N_U is the number of independent impurity problems one needs to solve. In all methods except nested cluster, $N_U = 1$. In nested cluster, it can be any number, but in the simplest scheme $N_U = 3$ independently from the cluster size. The loop starts with an initial guess for Σ^{imp} . $\mathcal{F}^C[G^{latt}]$ project G^{latt} onto degrees of freedom of impurity C . Convergence is reached when $G^{imp C} = \mathcal{F}^C[G^{latt}]$ for each C .

view CDMFT as “single-supercell DMFT.” The approximated LW functional then depends only on Green’s function components within a single supercell

$$\Phi[\{\hat{G}_{ij}\}_{\forall i,j}] \approx \Phi[\{\hat{G}_{ii}\}_{\forall i}] = \sum_i \Phi[\hat{G}_{ii}], \quad (D8)$$

where with “hat” we denote matrix objects: \hat{G}_{ij} is a matrix in the space of the I, J indices. Impurity action is given by

$$S = \sum_{I,J,\sigma} \iint d\tau d\tau' c_{\sigma,I}^+(\tau) [-\hat{G}_{\sigma}^{-1}]_{IJ}(\tau - \tau') c_{\sigma,J}(\tau') + U \sum_I \int d\tau c_{\uparrow,I}^+(\tau) c_{\downarrow,I}^+(\tau) c_{\downarrow,I}(\tau) c_{\uparrow,I}(\tau) \quad (D9)$$

and the self-consistency condition reads

$$\hat{G}^{imp}(i\omega_n) = \hat{G}_{ii}^{latt}(i\omega_n) \equiv \sum_{\mathbf{k} \in \text{RBZ}} \hat{G}_{\mathbf{k}}^{latt}(i\omega_n), \quad (D10)$$

where RBZ stands for “reduced Brillouin zone.” Note however, that in the derivation below we also rescale the lattice constant $a \rightarrow a/2$ so that no extra prefactors appear in the expressions,

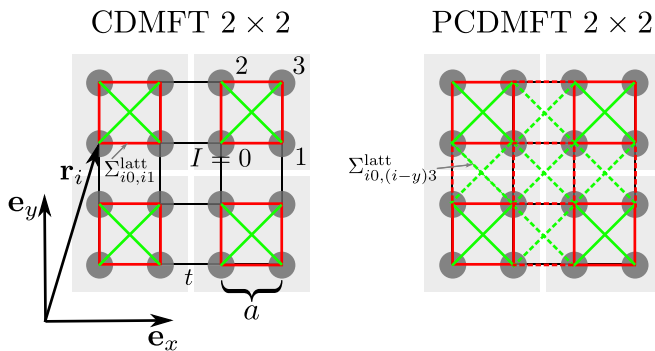


FIG. 20. Gray circles are lattice sites. Light gray squares are supercells. A site is denoted by the supercell index i and its index within the supercell I . In CDMFT, the self-energy is nonzero only between sites within a single supercell (green and red lines). In PCDMFT, it is copied by hand onto intercell bonds (dashed green and red lines). t is the hopping amplitude, a the lattice spacing, and $\mathbf{e}_{x,y}$ are the superlattice vectors.

and the RBZ extends from 0 to 2π along both axes. The lattice Dyson equation now involves a matrix inversion

$$\hat{G}_{\mathbf{k}}^{latt}(i\omega_n) = (\hat{G}_{0,\mathbf{k}}^{-1}(i\omega_n) - \hat{\Sigma}_{\mathbf{k}}^{latt}(i\omega_n))^{-1}. \quad (D11)$$

The self-energy approximation is simply

$$\hat{\Sigma}_{\mathbf{k}}^{latt}(i\omega_n) \approx \hat{\Sigma}^{imp}(i\omega_n). \quad (D12)$$

Note that, physically, the self-energy is put exclusively on bonds *within* a supercell, and *not* on bonds between supercells. This artificially breaks the translational symmetry of the lattice.

The bare propagator and the dispersion need to be rewritten in the supercell language. Here we present the expressions in the simple 2×2 tiling ($\mathbf{e}_{x,y}/a \rightarrow 2\mathbf{e}_{x,y}/a$):

$$\hat{G}_{0,\mathbf{k}}(i\omega_n) = [(i\omega_n + \mu)\hat{I} - \hat{\epsilon}_{\mathbf{k}}]^{-1}, \quad (D13)$$

$$\hat{\epsilon}_{\mathbf{k}} = t \cdot \hat{u}_{\mathbf{k}}, \quad (D14)$$

$$\hat{u}_{\mathbf{k}} = \begin{bmatrix} u(k_x) & u(k_y) & & \\ u^*(k_x) & & u(k_y) & \\ u^*(k_y) & & & u(k_x) \\ & u^*(k_y) & u^*(k_x) & \end{bmatrix}, \quad (D15)$$

$$u(k) = 1 + e^{-ik}. \quad (D16)$$

The drawback of this approach is that no simple interpretation of the result in terms of the original, translationally invariant lattice is possible. To obtain a translationally invariant self-energy, which can be plotted in the original BZ requires a post-processing step, or “periodization.” In the present case,

$$\Sigma_{\mathbf{r}=(0,0)}^{\text{per}} = \Sigma_{00}^{\text{imp}}, \quad (D17a)$$

$$\Sigma_{\mathbf{r}=(0,1)}^{\text{per}} = \Sigma_{01}^{\text{imp}}, \quad (D17b)$$

$$\Sigma_{\mathbf{r}=(1,1)}^{\text{per}} = \Sigma_{03}^{\text{imp}}. \quad (D17c)$$

The real-space vectors \mathbf{r} are given in the basis of the original lattice vectors. The rest of the real-space vectors can be filled in by symmetry

$$\Sigma_{\mathbf{r}=(x,y)}^{\text{per}} = \Sigma_{\mathbf{r}=(\pm x, \pm y)}^{\text{per}} = \Sigma_{\mathbf{r}=(\pm y, \pm x)}^{\text{per}} \quad (D17d)$$

and then we can Fourier transform to \mathbf{k} space:

$$\Sigma_{\mathbf{k}}^{\text{per}} = \sum_{\mathbf{r}} e^{i\mathbf{k}\cdot\mathbf{r}} \Sigma_{\mathbf{r}}^{\text{per}}. \quad (\text{D17e})$$

Note that periodization is an *ad hoc* procedure that does not have a clear physical interpretation in terms of the LW approximation. Also, the physical quantity that is being periodized can be chosen arbitrarily, and different choices will in general lead to different results.

3. Periodized CDMFT (PCDMFT)

The idea of PCDMFT [6,8,49,50,52] is that the periodization should be performed in each DMFT iteration, and that the self-consistency should be closed using the translationally invariant Green's function, rather than the superlattice one. This scheme cannot be simply derived from an approximation of the LW functional. The impurity action remains the same as in CDMFT, Eq. (D9).

The idea of PCDMFT can be achieved either by placing the missing self-energies on the superlattice (see Fig. 20)

$$\begin{aligned} \hat{\Sigma}_{\mathbf{k}}^{\text{latt}}(i\omega_n) &\approx \Sigma_{00}^{\text{imp}}(i\omega_n)\hat{I} + \Sigma_{01}^{\text{imp}}(i\omega_n)\hat{u}_{\mathbf{k}} + \Sigma_{03}^{\text{imp}}(i\omega_n)\hat{w}_{\mathbf{k}} \\ &= \hat{\Sigma}^{\text{imp}}(i\omega_n) \circ (\hat{I} + \hat{u}_{\mathbf{k}} + \hat{w}_{\mathbf{k}}), \end{aligned} \quad (\text{D18})$$

where \circ denotes element-wise product, and

$$\hat{w}_{\mathbf{k}} = \begin{bmatrix} & & w_1(\mathbf{k}) \\ & w_2(\mathbf{k}) & \\ w_1^*(\mathbf{k}) & w_2^*(\mathbf{k}) & \end{bmatrix}, \quad (\text{D19})$$

$$w_1(\mathbf{k}) = 1 + e^{-ik_x} + e^{-ik_y} + e^{-i(k_x+k_y)}, \quad (\text{D20})$$

$$w_2(\mathbf{k}) = 1 + e^{-ik_x} + e^{-ik_y} + e^{-i(k_x-k_y)}, \quad (\text{D21})$$

or, equivalently, by periodizing the self-energy with Eq. (D17) and rewriting the self-consistency condition with

$$\hat{G}^{\text{imp}} = \begin{bmatrix} G_{\mathbf{r}=(0,0)}^{\text{per}} & G_{\mathbf{r}=(0,1)}^{\text{per}} & G_{\mathbf{r}=(0,1)}^{\text{per}} & G_{\mathbf{r}=(1,1)}^{\text{per}} \\ G_{\mathbf{r}=(0,1)}^{\text{per}} & G_{\mathbf{r}=(0,0)}^{\text{per}} & G_{\mathbf{r}=(1,1)}^{\text{per}} & G_{\mathbf{r}=(0,1)}^{\text{per}} \\ G_{\mathbf{r}=(0,1)}^{\text{per}} & G_{\mathbf{r}=(1,1)}^{\text{per}} & G_{\mathbf{r}=(0,0)}^{\text{per}} & G_{\mathbf{r}=(0,1)}^{\text{per}} \\ G_{\mathbf{r}=(1,1)}^{\text{per}} & G_{\mathbf{r}=(0,1)}^{\text{per}} & G_{\mathbf{r}=(0,1)}^{\text{per}} & G_{\mathbf{r}=(0,0)}^{\text{per}} \end{bmatrix}, \quad (\text{D22})$$

where

$$G_{\mathbf{r}}^{\text{per}} = \sum_{\mathbf{k} \in \text{BZ}} e^{-i\mathbf{k}\cdot\mathbf{r}} G_{\mathbf{k}}^{\text{per}} = \sum_{\mathbf{k} \in \text{BZ}} \frac{e^{-i\mathbf{k}\cdot\mathbf{r}}}{G_{0,\mathbf{k}}^{-1}(i\omega_n) - \Sigma_{\mathbf{k}}^{\text{per}}(i\omega_n)}. \quad (\text{D23})$$

The final result is the translationally invariant self-energy, which solves Eq. (D22).

Note there is another variant of PCDMFT method (proposed in Ref. [51]) where the self-energy is periodized with additional coefficients, so that it is rigorously causal. In the present case, this method would correspond to restoring translational invariance on the lattice the following way:

$$\hat{\Sigma}_{\mathbf{k}}^{\text{latt}}(i\omega_n) \approx \hat{\Sigma}^{\text{imp}}(i\omega_n) \circ (\hat{I} + \frac{1}{2}\hat{u}_{\mathbf{k}} + \frac{1}{4}\hat{w}_{\mathbf{k}}).$$

We observe that this method corrects the local part of self-energy in the difficult regime compared to regular PCDMFT,

but the nonlocal part is strongly underestimated throughout the phase diagram (results not shown).

4. Dynamical cluster approximation (DCA)

In DCA [7,27–45] method, the conservation of momentum in LW diagrams is approximated by [27,33]

$$\mathbf{k}' - \mathbf{k} = \mathbf{q} \longrightarrow \mathbf{K}(\mathbf{k}') - \mathbf{K}(\mathbf{k}) = \mathbf{K}(\mathbf{q}), \quad (\text{D24})$$

where \mathbf{K} represents the ‘‘coarse-grained’’ BZ points, which is illustrated on Fig. 21. The coarse-grained BZ contains only a certain finite and discrete subset of wave vectors. The notation $\mathbf{K}(\mathbf{k})$ means ‘‘the coarse grained wave vector closest to the wave vector \mathbf{k} .’’ Because of the relaxation of momentum conservation, the diagrams factorize: the LW functional still depends on all G components, but only through their sums:

$$\Phi[\{G_{\mathbf{k}}\}_{\forall \mathbf{k}}] \approx \Phi \left[\left\{ \sum_{\mathbf{k} \in \mathcal{P}(\mathbf{K})} G_{\mathbf{k}} \right\}_{\forall \mathbf{K}} \right] \equiv \Phi[\{G_{\mathbf{K}}\}_{\forall \mathbf{K}}]. \quad (\text{D25})$$

Here, $\mathcal{P}(\mathbf{K})$ is the set of fine-grain wave vectors \mathbf{k} that are closest to the coarse-grained wave vector \mathbf{K} (Voronoi patch [89] around \mathbf{K}). This approximation leads to a piecewise-constant self-energy in \mathbf{k} space, because of

$$\frac{\partial \Phi[G_{\mathbf{K}}]}{\partial G_{\mathbf{k}}} = \frac{\partial \Phi[G_{\mathbf{K}}]}{\partial G_{\mathbf{K}}} \frac{\partial G_{\mathbf{K}}}{\partial G_{\mathbf{k}}} = \frac{\partial \Phi[G_{\mathbf{K}}]}{\partial G_{\mathbf{K}}} \delta_{\mathbf{k} \in \mathcal{P}(\mathbf{K})}. \quad (\text{D26})$$

The impurity action is given by

$$\begin{aligned} S &= \sum_{\mathbf{K}, \sigma} \iint d\tau d\tau' c_{\sigma, \mathbf{K}}^+(\tau) [-\mathcal{G}_{\sigma, \mathbf{K}}^{-1}(\tau - \tau')] c_{\sigma, \mathbf{K}}(\tau') \\ &+ U \sum_{\mathbf{K}, \mathbf{K}', \mathbf{Q}} \int d\tau c_{\uparrow, \mathbf{K}+\mathbf{Q}}^+(\tau) c_{\downarrow, \mathbf{K}'-\mathbf{Q}}^+(\tau) c_{\downarrow, \mathbf{K}'}(\tau) c_{\uparrow, \mathbf{K}}(\tau), \end{aligned} \quad (\text{D27})$$

and it corresponds to a finite cyclic cluster in real space \mathbf{R} .

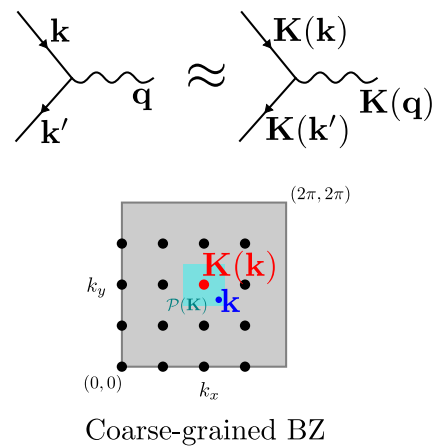


FIG. 21. DCA approximates momentum-conservation. Lattice self-energy at wave vector \mathbf{k} is obtained from the impurity self-energy at the closest coarse grained wave vector \mathbf{K} . Therefore it is constant within each Voronoi patch $\mathcal{P}(\mathbf{K})$. The example presented is the 4×4 scheme.

Self-consistency condition reads

$$G_{\mathbf{K}}^{\text{imp}}(i\omega_n) = G_{\mathbf{K}}^{\text{latt}}(i\omega_n) \equiv \sum_{\mathbf{k} \in \mathcal{P}(\mathbf{K})} G_{\mathbf{k}}^{\text{latt}}(i\omega_n). \quad (\text{D28})$$

As already mentioned, the self-energy on the lattice is simply

$$\Sigma_{\mathbf{k}}^{\text{latt}}(i\omega_n) = \Sigma_{\mathbf{K}(\mathbf{k})}^{\text{imp}}(i\omega_n). \quad (\text{D29})$$

Note that more general coarse-graining schemes exist, and that $\mathcal{P}(\mathbf{K})$ does not necessarily present a Voronoi patch around the wave vector \mathbf{K} . Patches may have different shapes [29], and may even be interlaced [46]. In the present paper, we only use the simplest scheme where patches are Voronoi patches, and all have the same shape.

5. Continuous self-energy DCA (DCA⁺)

DCA⁺ [46–48] aims at improving the interpretation of the impurity self-energy in terms of the self-energy on the lattice. A piecewise constant self-energy is strongly counter intuitive and hard to compare to other methods. It is a natural step to try and interpolate the coarse-grained self-energy to obtain a smooth self-energy on the lattice. However, this scheme does not have a clear derivation as a LW functional approximation. Furthermore, the interpolation can be done in various ways, and the method is not uniquely defined. In this paper we implement (and present here) the version of the method as proposed in the original paper, Ref. [48].

The impurity action is the same as in DCA, Eq. (D27).

The difference from DCA is the addition of a self-consistency condition that needs to be satisfied:

$$\Sigma_{\mathbf{K}}^{\text{imp}}(i\omega_n) = \Sigma_{\mathbf{K}}^{\text{latt}}(i\omega_n) \equiv \sum_{\mathbf{k} \in \mathcal{P}(\mathbf{K})} \Sigma_{\mathbf{k}}^{\text{latt}}(i\omega_n). \quad (\text{D30})$$

Here, Σ^{latt} is a smooth function of \mathbf{k} . Note that for a given $\Sigma_{\mathbf{K}}^{\text{imp}}$, Σ^{latt} is not uniquely defined. This self-consistency condition imposes

$$\Sigma_{\text{loc}}^{\text{latt}} = \Sigma_{\text{loc}}^{\text{imp}}, \quad (\text{D31})$$

but in general

$$\Sigma_{\mathbf{k}=\mathbf{K}}^{\text{latt}} \neq \Sigma_{\mathbf{K}}^{\text{imp}}. \quad (\text{D32})$$

While a general interpolation of $\Sigma_{\mathbf{K}}^{\text{imp}}$ is unlikely to satisfy the condition (D30), a Bayesian approach can be employed to find the most probable interpolation that does satisfy it. The method used is Richardson-Lucy deconvolution, and it is performed with respect to an interpolation of $\Sigma_{\mathbf{K}}^{\text{imp}} \rightarrow \bar{\Sigma}_{\mathbf{k}}^{\text{imp}}$ such that

$$\bar{\Sigma}_{\mathbf{k}=\mathbf{K}}^{\text{imp}} = \Sigma_{\mathbf{K}}^{\text{imp}}. \quad (\text{D33})$$

One starts from an initial guess for Σ^{latt} (say, $\Sigma_{\mathbf{k}}^{\text{latt}} = \bar{\Sigma}_{\mathbf{k}}^{\text{imp}}$), and iterates

$$\Sigma_{\mathbf{k}}^{\text{latt}} \leftarrow \Sigma_{\mathbf{k}}^{\text{latt}} \sum_{\mathbf{k}' \in \mathcal{P}(\mathbf{k})} \frac{\bar{\Sigma}_{\mathbf{k}'}^{\text{imp}}}{\sum_{\mathbf{k}'' \in \mathcal{P}(\mathbf{k}')} \Sigma_{\mathbf{k}''}^{\text{latt}}} \quad (\text{D34})$$

until convergence is reached. Here, $\mathcal{P}(\mathbf{k})$ denotes a patch of the same shape/size as the Voronoi patches of the coarse-grained BZ, but centered at the fine-grain wave vector \mathbf{k} . The final

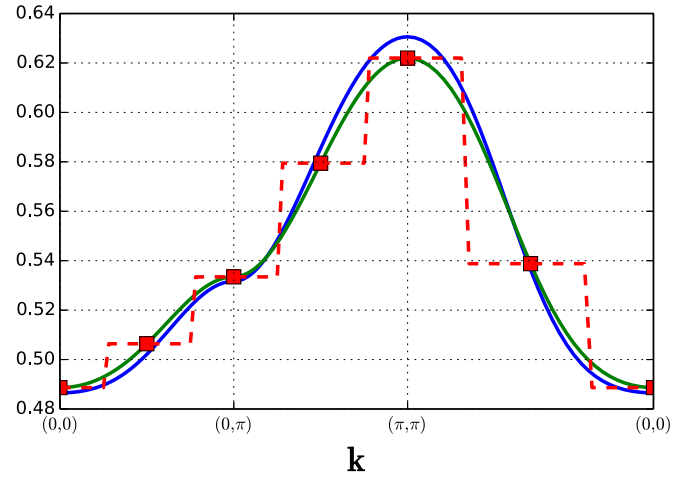


FIG. 22. Example of various quantities appearing in DCA⁺. Parameters of the calculation: $U/D = 1.4$, $\delta = 8\%$, and $T/D = 0.125$.

result has the property

$$\bar{\Sigma}_{\mathbf{k}}^{\text{imp}} = \sum_{\mathbf{k}' \in \mathcal{P}(\mathbf{k})} \Sigma_{\mathbf{k}'}^{\text{latt}}, \quad \forall \mathbf{k}, \quad (\text{D35})$$

which satisfies a stronger requirement than necessary.

Note also that the actual interpolation is performed not on Σ^{imp} , but on an auxiliary quantity Ξ which is by construction more local than the self-energy. The method of interpolation proposed is the Wannier interpolation

$$\Xi_{\mathbf{K}}(i\omega_n) = (\Sigma_{\mathbf{K}}^{\text{imp}}(i\omega_n) - \text{sgn}(\omega_n)i\alpha)^{-1}, \quad (\text{D36})$$

$$\Xi_{\mathbf{R}} = \sum_{\mathbf{K}} e^{-i\mathbf{K}\cdot\mathbf{R}} \Xi_{\mathbf{K}}, \quad (\text{D37})$$

$$\bar{\Xi}_{\mathbf{k}} = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \Xi_{\mathbf{R}}, \quad (\text{D38})$$

$$\bar{\Sigma}_{\mathbf{k}}^{\text{imp}}(i\omega_n) = \bar{\Xi}_{\mathbf{k}}^{-1}(i\omega_n) + \text{sgn}(\omega_n)i\alpha, \quad \alpha > 0. \quad (\text{D39})$$

Note that $\bar{\Xi}_{\mathbf{k}}$ does not necessarily satisfy all the lattice symmetries. One way to restore lattice symmetries is to calculate it as

$$\bar{\Xi}_{\mathbf{k}} = \frac{1}{N_M} \sum_{\hat{M}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\hat{M}\mathbf{R}} \Xi_{\mathbf{R}}, \quad (\text{D40})$$

where \hat{M} runs over all the symmetry operations on the lattice, of which there are N_M . On the square lattice there are $N_M = 8$ operations ($\mathbf{R}_{x,y} \rightarrow \pm\mathbf{R}_{(x,y),(y,x)}$), which restore the eightfold symmetry in $\bar{\Xi}_{\mathbf{k}}$. Examples of different self-energy quantities appearing in DCA⁺ are given in Fig. 22 and compared to the piecewise-constant interpolation that is used in standard DCA.

- [1] G. Baym and L. P. Kadanoff, *Phys. Rev.* **124**, 287 (1961).
- [2] C. D. Dominicis and P. C. Martin, *J. Math. Phys.* **5**, 14 (1964).
- [3] C. D. Dominicis and P. C. Martin, *J. Math. Phys.* **5**, 31 (1964).
- [4] P. Nozieres, *Theory Of Interacting Fermi Systems (Advanced Books Classics)* (Westview Press, Boulder, Colorado, USA, 1997).
- [5] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996).
- [6] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, *Rev. Mod. Phys.* **78**, 865 (2006).
- [7] M. H. Hettler, A. N. Tahvildar-Zadeh, M. Jarrell, T. Pruschke, and H. R. Krishnamurthy, *Phys. Rev. B* **58**, R7475 (1998).
- [8] A. I. Lichtenstein and M. I. Katsnelson, *Phys. Rev. B* **62**, R9283 (2000).
- [9] G. Kotliar, S. Y. Savrasov, G. Pálsson, and G. Biroli, *Phys. Rev. Lett.* **87**, 186401 (2001).
- [10] T. A. Maier, M. Jarrell, T. Pruschke, and M. H. Hettler, *Rev. Mod. Phys.* **77**, 1027 (2005).
- [11] M. Civelli, *Phys. Rev. B* **79**, 195113 (2009).
- [12] M. Civelli, M. Capone, A. Georges, K. Haule, O. Parcollet, T. D. Stanescu, and G. Kotliar, *Phys. Rev. Lett.* **100**, 046402 (2008).
- [13] M. Civelli, M. Capone, S. S. Kancharla, O. Parcollet, and G. Kotliar, *Phys. Rev. Lett.* **95**, 106402 (2005).
- [14] L. De Leo, M. Civelli, and G. Kotliar, *Phys. Rev. B* **77**, 075107 (2008).
- [15] S. S. Kancharla, B. Kyung, D. Sénéchal, M. Civelli, M. Capone, G. Kotliar, and A.-M. S. Tremblay, *Phys. Rev. B* **77**, 184516 (2008).
- [16] B. Kyung, S. S. Kancharla, D. Sénéchal, A.-M. S. Tremblay, M. Civelli, and G. Kotliar, *Phys. Rev. B* **73**, 165114 (2006).
- [17] B. Kyung, G. Kotliar, and A.-M. S. Tremblay, *Phys. Rev. B* **73**, 205106 (2006).
- [18] S. Okamoto, D. Sénéchal, M. Civelli, and A.-M. S. Tremblay, *Phys. Rev. B* **82**, 180511 (2010).
- [19] Ž. Osolin and R. Žitko, *Phys. Rev. B* **95**, 035107 (2017).
- [20] O. Parcollet, G. Biroli, and G. Kotliar, *Phys. Rev. Lett.* **92**, 226402 (2004).
- [21] H. Park, K. Haule, and G. Kotliar, *Phys. Rev. Lett.* **101**, 186403 (2008).
- [22] S. Sakai, M. Civelli, and M. Imada, *Phys. Rev. B* **94**, 115130 (2016).
- [23] G. Sordi, K. Haule, and A.-M. S. Tremblay, *Phys. Rev. B* **84**, 075161 (2011).
- [24] G. Sordi, K. Haule, and A.-M. S. Tremblay, *Phys. Rev. Lett.* **104**, 226402 (2010).
- [25] G. Sordi, P. Sémon, K. Haule, and A.-M. S. Tremblay, *Phys. Rev. B* **87**, 041101 (2013).
- [26] Y. Z. Zhang and M. Imada, *Phys. Rev. B* **76**, 045108 (2007).
- [27] K. Aryanpour, M. H. Hettler, and M. Jarrell, *Phys. Rev. B* **65**, 153102 (2002).
- [28] H. T. Dang, X. Y. Xu, K.-S. Chen, Z. Y. Meng, and S. Wessel, *Phys. Rev. B* **91**, 155101 (2015).
- [29] M. Ferrero, P. S. Cornaglia, L. De Leo, O. Parcollet, G. Kotliar, and A. Georges, *Phys. Rev. B* **80**, 064501 (2009).
- [30] M. Ferrero, P. S. Cornaglia, L. De Leo, O. Parcollet, G. Kotliar, and A. Georges, *Europhys. Lett.* **85**, 57009 (2008).
- [31] M. Ferrero, O. Parcollet, A. Georges, G. Kotliar, and D. N. Basov, *Phys. Rev. B* **82**, 054502 (2010).
- [32] E. Gull, M. Ferrero, O. Parcollet, A. Georges, and A. J. Millis, *Phys. Rev. B* **82**, 155101 (2010).
- [33] M. H. Hettler, M. Mukherjee, M. Jarrell, and H. R. Krishnamurthy, *Phys. Rev. B* **61**, 12739 (2000).
- [34] C. Huscroft, M. Jarrell, T. Maier, S. Moukouri, and A. N. Tahvildarzadeh, *Phys. Rev. Lett.* **86**, 139 (2001).
- [35] M. Jarrell, T. A. Maier, C. Huscroft, and S. Moukouri, *Phys. Rev. B* **64**, 195130 (2001).
- [36] P. R. C. Kent, M. Jarrell, T. A. Maier, and T. Pruschke, *Phys. Rev. B* **72**, 060411 (2005).
- [37] J. P. F. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, G. K.-L. Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jiménez-Hoyos, E. Kozik, X.-W. Liu, A. J. Millis, N. V. Prokof'ev, M. Qin, G. E. Scuseria, H. Shi, B. V. Svistunov, L. F. Tocchio, I. S. Tupitsyn, S. R. White, S. Zhang, B.-X. Zheng, Z. Zhu, and E. Gull, (Simons Collaboration on the Many-Electron Problem) *Phys. Rev. X* **5**, 041041 (2015).
- [38] A. Macridin, M. Jarrell, and T. A. Maier, *Phys. Rev. B* **70**, 113105 (2004).
- [39] A. Macridin, M. Jarrell, T. Maier, P. R. C. Kent, and E. D'Azvedo, *Phys. Rev. Lett.* **97**, 036401 (2006).
- [40] A. Macridin, M. Jarrell, T. Maier, and D. J. Scalapino, *Phys. Rev. Lett.* **99**, 237001 (2007).
- [41] T. A. Maier, M. Jarrell, A. Macridin, and C. Slezak, *Phys. Rev. Lett.* **92**, 027005 (2004).
- [42] T. A. Maier, M. Jarrell, T. C. Schulthess, P. R. C. Kent, and J. B. White, *Phys. Rev. Lett.* **95**, 237001 (2005).
- [43] T. A. Maier, M. S. Jarrell, and D. J. Scalapino, *Phys. Rev. Lett.* **96**, 047005 (2006).
- [44] T. A. Maier and D. J. Scalapino, *Phys. Rev. B* **84**, 180513 (2011).
- [45] T. A. Maier, P. Staar, and D. J. Scalapino, [arXiv:1507.06206](https://arxiv.org/abs/1507.06206).
- [46] P. Staar, M. Jiang, U. R. Hähner, T. C. Schulthess, and T. A. Maier, *Phys. Rev. B* **93**, 165144 (2016).
- [47] P. Staar, T. Maier, and T. C. Schulthess, *Phys. Rev. B* **89**, 195133 (2014).
- [48] P. Staar, T. Maier, and T. C. Schulthess, *Phys. Rev. B* **88**, 115101 (2013).
- [49] Q.-X. Li, R.-Q. He, and Z.-Y. Lu, *Phys. Rev. B* **92**, 155127 (2015).
- [50] T. D. Stanescu and G. Kotliar, *Phys. Rev. B* **74**, 125110 (2006).
- [51] G. Biroli and G. Kotliar, *Phys. Rev. B* **65**, 155112 (2002).
- [52] G. Biroli, O. Parcollet, and G. Kotliar, *Phys. Rev. B* **69**, 205108 (2004).
- [53] S. Sakai, G. Sangiovanni, M. Civelli, Y. Motome, K. Held, and M. Imada, *Phys. Rev. B* **85**, 035102 (2012).
- [54] T. Ayrál and O. Parcollet, *Phys. Rev. B* **93**, 235124 (2016).
- [55] T. Ayrál and O. Parcollet, *Phys. Rev. B* **94**, 075159 (2016).
- [56] T. Ayrál and O. Parcollet, *Phys. Rev. B* **92**, 115109 (2015).
- [57] T. Ayrál, J. Vucicevic, and O. Parcollet, *Phys. Rev. Lett.* **119**, 166401 (2017).
- [58] J. Vučićević, T. Ayrál, and O. Parcollet, *Phys. Rev. B* **96**, 104504 (2017).
- [59] A. Georges, in *Lectures on the Physics of Highly Correlated Electron Systems VIII: Eighth Training Course in the Physics of Correlated Electron Systems and High-Tc Superconductors*, edited by A. Avella and F. Mancini, AIP Conf. Proc. No. 715 (AIP, New York, 2004), pp. 3–74.
- [60] E. Kozik, M. Ferrero, and A. Georges, *Phys. Rev. Lett.* **114**, 156402 (2015).
- [61] O. Gunnarsson, G. Rohringer, T. Schäfer, G. Sangiovanni, and A. Toschi, *Phys. Rev. Lett.* **119**, 056402 (2017).

- [62] G. Rohringer, H. Hafermann, A. Toschi, A. A. Katanin, A. E. Antipov, M. I. Katsnelson, A. I. Lichtenstein, A. N. Rubtsov, and K. Held, [arXiv:1705.00024](https://arxiv.org/abs/1705.00024).
- [63] T. Schäfer, G. Rohringer, O. Gunnarsson, S. Ciuchi, G. Sangiovanni, and A. Toschi, *Phys. Rev. Lett.* **110**, 246405 (2013).
- [64] T. Schäfer, S. Ciuchi, M. Wallerberger, P. Thunström, O. Gunnarsson, G. Sangiovanni, G. Rohringer, and A. Toschi, *Phys. Rev. B* **94**, 235108 (2016).
- [65] O. Gunnarsson, T. Schäfer, J. P. F. LeBlanc, J. Merino, G. Sangiovanni, G. Rohringer, and A. Toschi, *Phys. Rev. B* **93**, 245102 (2016).
- [66] T. Ribic, G. Rohringer, and K. Held, *Phys. Rev. B* **93**, 195105 (2016).
- [67] O. Gunnarsson, T. Schäfer, J. P. F. LeBlanc, E. Gull, J. Merino, G. Sangiovanni, G. Rohringer, and A. Toschi, *Phys. Rev. Lett.* **114**, 236402 (2015).
- [68] W. Wu, M. Ferrero, A. Georges, and E. Kozik, *Phys. Rev. B* **96**, 041105 (2017).
- [69] F. Tandetky, J. K. Dewhurst, S. Sharma, and E. K. U. Gross, *Phys. Rev. B* **92**, 115125 (2015).
- [70] A. Schiller and K. Ingersent, *Phys. Rev. Lett.* **75**, 113 (1995).
- [71] A. A. Kananenka, E. Gull, and D. Zgid, *Phys. Rev. B* **91**, 121111 (2015).
- [72] T. N. Lan, A. Shee, J. Li, E. Gull, and D. Zgid, *Phys. Rev. B* **96**, 155106 (2017).
- [73] T. N. Lan and D. Zgid, [arXiv:1703.06981](https://arxiv.org/abs/1703.06981).
- [74] M. Motta, D. M. Ceperley, G. K.-L. Chan, J. A. Gomez, E. Gull, S. Guo, C. Jimenez-Hoyos, T. N. Lan, J. Li, F. Ma, A. J. Millis, N. V. Prokof'ev, U. Ray, G. E. Scuseria, S. Sorella, E. M. Stoudenmire, Q. Sun, I. S. Tupitsyn, S. R. White, D. Zgid, and S. Zhang, *Phys. Rev. X* **7**, 031059 (2017).
- [75] D. Zgid and E. Gull, *New J. Phys.* **19**, 023047 (2017).
- [76] V. Janiš and V. Pokorný, *Phys. Rev. B* **90**, 045143 (2014).
- [77] A. Stan, P. Romaniello, S. Rigamonti, L. Reining, and J. A. Berger, *New J. Phys.* **17**, 093045 (2015).
- [78] A. I. Rubtsov and A. N. Lichtenstein, *J. Exp. Theor. Phys. Lett.* **80**, 61 (2004).
- [79] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, *Rev. Mod. Phys.* **83**, 349 (2011).
- [80] The hybridization Δ is defined by $\Delta(i\omega_n) \equiv i\omega_n + C - \mathcal{G}^{-1}$, where C is a constant such that $\Delta(i\omega_n) \xrightarrow{n \rightarrow \infty} 0$.
- [81] T. Jabben, N. Grewe, and S. Schmitt, *Phys. Rev. B* **85**, 165122 (2012).
- [82] S. Okamoto, A. J. Millis, H. Monien, and A. Fuhrmann, *Phys. Rev. B* **68**, 195121 (2003).
- [83] R. Rossi, F. Werner, N. Prokof'ev, and B. Svistunov, *Phys. Rev. B* **93**, 161102 (2016).
- [84] O. Parcollet, M. Ferrero, T. Ayrál, H. Hafermann, P. Seth, and I. S. Krivenko, *Comput. Phys. Commun.* **196**, 398 (2015).
- [85] G. Rohringer, A. Valli, and A. Toschi, *Phys. Rev. B* **86**, 125114 (2012).
- [86] N. Bickers, in *Theoretical Methods for Strongly Correlated Electrons*, edited by D. Sènèchal, A.-M. Tremblay, and C. Bourbonnais, CRM Series in Mathematical Physics (Springer, New York, 2004), pp. 237–296.
- [87] P. Sun and G. Kotliar, *Phys. Rev. B* **66**, 085120 (2002).
- [88] D. Tanasković, K. Haule, G. Kotliar, and V. Dobrosavljević, *Phys. Rev. B* **84**, 115105 (2011).
- [89] F. Aurenhammer, *ACM Comput. Surv.* **23**, 345 (1991).