

Dynamical cluster approximation: Nonlocal dynamics of correlated electron systems

M. H. Hettler*

*Department of Physics, University of Cincinnati, Cincinnati, Ohio 45221
and Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439*

M. Mukherjee[†] and M. Jarrell

Department of Physics, University of Cincinnati, Cincinnati, Ohio 45221

H. R. Krishnamurthy

*Department of Physics, University of Cincinnati, Cincinnati, Ohio 45221
and Department of Physics, Indian Institute of Science, Bangalore 560012, India*

(Received 18 March 1999; revised manuscript received 10 December 1999)

We recently introduced the dynamical cluster approximation (DCA), a technique that includes short-ranged dynamical correlations in addition to the local dynamics of the dynamical mean-field approximation while preserving causality. The technique is based on an iterative self-consistency scheme on a finite-size periodic cluster. The dynamical mean-field approximation (exact result) is obtained by taking the cluster to a single site (the thermodynamic limit). Here, we provide details of our method, explicitly show that it is causal, systematic, Φ derivable, and that it becomes conserving as the cluster size increases. We demonstrate the DCA by applying it to a quantum Monte Carlo and exact enumeration study of the two-dimensional Falicov-Kimball model. The resulting spectral functions preserve causality, and the spectra and the charge-density-wave transition temperature converge quickly and systematically to the thermodynamic limit as the cluster size increases.

I. INTRODUCTION

Strongly correlated-electron systems have been at the center of theoretical and experimental research interest for several decades. This interest was greatly intensified by the discovery of heavy fermion metals and superconductors, and recently of the high- T_c superconductors. The observation of non-Fermi-liquid behavior first in the Cuprates and later even in some heavy fermion systems has given further impetus. Away from a transition, these materials are characterized by short-ranged dynamical correlations such as the local correlations responsible for the Kondo effect. In addition, the doped cuprates display short-ranged antiferromagnetic dynamical correlations thought to be responsible for pair formation. Some of this physics is captured by the simplest models of strongly correlated electrons, such as the Hubbard model (HM) and the periodic Anderson model (PAM). Despite the short range of the dynamical correlations and numerous sophisticated techniques introduced since the inception of the models, they remain unsolved.

However, recently Metzner and Vollhardt showed¹ that these models undergo significant simplification in the limit of infinite dimensions, $D = \infty$. In this limit, provided the kinetic energy is scaled as $1/\sqrt{D}$, the self-energy and vertex functions may be taken to be purely local in space although they retain a nontrivial frequency dependence. Consequently, the HM and PAM can be mapped onto a self-consistently embedded Anderson impurity problem; i.e., a single correlated site subject to a self-consistently determined energy-dependent hybridization with a conduction electron “bath” or “host” representing the remaining sites of the lattice, or equivalently (on eliminating this bath), to a dynamical mean

field.^{2,3} The resulting dynamical mean-field theory (DMFT) is exact in infinite dimensions and has been used to establish the thermodynamic properties and phase diagrams of these models using quantum Monte Carlo (QMC) and other methods.^{2,4,5}

A similar self-consistent single site theory can be obtained by *assuming* a purely local self-energy (and vertex functions) even in finite dimensions. This yields the natural mean-field theory for correlated lattice systems and is called the dynamical mean-field approximation (DMFA). While it has been shown that this approximation captures many key features of strongly correlated systems even in a finite-dimensional context, the DMFA has some obvious and significant limitations. For example, the only dynamical correlations present are those that may be properly treated on a single site. Therefore, there are no nonlocal dynamical correlations. These are necessary, for example, to describe phases with explicitly nonlocal order parameters or those with lower symmetry than the lattice, of which d -wave superconductivity is perhaps the most prominent example. But even phases with local order parameters (e.g., commensurate magnetism) will certainly be affected by the nonlocal dynamical correlations (spin waves) neglected by the DMFA. In addition, as we show in this paper, the DMFA is not a conserving approximation, with violations of the Ward identity associated with current conservation (the equation of continuity) for any D , including the limit $D \rightarrow \infty$.

Consequently, there have been efforts to extend the DMFA by inclusion of nonlocal correlations, which would correspond to $1/D$ corrections to the self-energy of the $D = \infty$ models.^{6,7} These efforts have failed to construct a causal theory, one that preserves spectral weight and which retains

positive semidefinite spectral functions, out of nonlocal Green functions. Such violations of positivity have been seen explicitly and discussed in the work by van Dongen.⁶ Even in the sophisticated Φ derivable technique developed by Schiller and Ingersent,⁷ violations of the sum rules occurred for moderately large values of the interaction strength in the Falicov-Kimball model (FKM).

A different approach by Smith and Si⁸ allows for the incorporation of nonlocal interactions in the original Hamiltonian (beyond the Hartree level) by rescaling them with the same $1/\sqrt{D}$ factor in the limit $D=\infty$ as the kinetic energy. The resulting self-energy remains local, and the system maps to an impurity model coupled to both a Fermionic bath (the electrons on the host) as well as a bosonic bath (the two-particle interactions). While this approach is attractive we believe that this scaling is difficult to justify formally. In addition, since the resulting effective theory is still a single site theory, it does not allow one to address some of the problems discussed above.

In a recent paper⁹ we introduced the dynamical cluster approximation (DCA), an iterative self-consistency scheme on a finite-size periodic cluster of size N_c . It extends the DMFA through the inclusion of short-ranged dynamical correlations, remains fully causal, and restores the conservation laws of Ward¹⁰ and Baym¹¹ when the cluster becomes large. The essential approximation of the DCA is to take the irreducible self-energy $\Sigma^c(\mathbf{K}, \omega)$ of the cluster as a good approximation to the self-energy of the real system at the cluster momenta \mathbf{K} . When N_c , the number of cluster momenta in the first Brillouin zone, is relatively small, this approximation can only be justified if the self-energy of the real system is weakly momentum dependent. Such a weak momentum dependence is realized in high dimensions (there is *no* momentum dependence in $D=\infty$). Then, a coarse grid of \mathbf{K} points is sufficient to capture all the short-ranged (but non-local) dynamics. In low dimensions, the validity of the approximation is less clear. However, in many correlated systems the momentum dependence of the self-energy is less important than its frequency dependence, for example in the effective-mass corrections arising from interactions in Fermi liquids and the marginal Fermi liquid¹² (MFL) or nearly antiferromagnetic Fermi-liquid¹³ (NAFL) phenomenology of high- T_c superconductors. In addition, because of the coupling of the cluster to a much larger host, the method allows for a systematic finite size study that is likely to converge faster than standard methods such as exact diagonalization, lattice QMC, and the fluctuation exchange approximation¹⁴ (FLEX).

In this work we present the first detailed discussion of the DCA. The paper is organized as follows: First, we review the DMFA and discuss its limitations. Then, we review the steps of the DCA and discuss the details of the formalism. We then apply the DCA to the half-filled FKM using quantum Monte Carlo and exact enumeration for the cluster problem to obtain self-energies and Green functions. For simplicity, we consider only the single-band model with nearest-neighbor hopping on a periodic square lattice with N sites. We demonstrate that the DCA algorithm converges systematically with increasing cluster size and remains fully causal. We then discuss the results and their implications. In the Appendixes, we provide the formalism needed to calculate

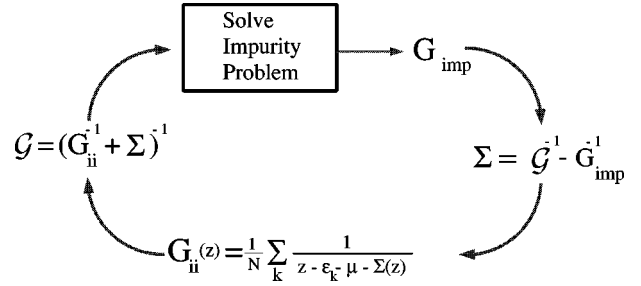


FIG. 1. Sketch of the DMFA algorithm.

the two-particle properties, generalize our formalism to models with extended range interactions, prove that it is causal, and discuss its conserving properties.

II. DYNAMICAL MEAN-FIELD APPROXIMATION

The DMFA (Ref. 1) may be derived in any dimension by disregarding momentum conservation at the internal vertices of the self-energy.¹⁵ This approximation becomes exact in the limit of infinite dimensions $D \rightarrow \infty$, provided that the near-neighbor electronic hopping integral is rescaled so that $t \sim D^{-1/2}$. Then, the single-particle Green function $G(r) \sim t^r \sim D^{-r/2}$ and the self-energy becomes a purely local functional of the local Green function only, $\Sigma_{i,j} = \Sigma_{i,i}(G_{i,i}) \delta_{i,j}$, which is momentum independent $\Sigma(\mathbf{k}, \omega) = \Sigma_{i,i}(\omega) + \mathcal{O}(1/\sqrt{D})$. The lattice problem may then be mapped onto a self-consistently embedded impurity problem. The resulting DMFA algorithm, illustrated in Fig. 1, has the following steps: (1) The procedure starts with a guess for $\Sigma_{ii}(\omega)$, usually zero. (2) Then, we calculate the local lattice Green function $G_{i,i}(\omega) = 1/N \sum_{\mathbf{k}} [G_o^{-1}(\mathbf{k}, \omega) - \Sigma_{i,i}(\omega)]^{-1}$, where $G_o(\mathbf{k}, \omega)$ is the bare lattice Green function and N is the (infinite) number of points of the lattice. (3) Next, we compute $\mathcal{G}(\omega)$ which includes self-energy processes at all lattice sites except at the ‘‘impurity’’ site i under consideration, $\mathcal{G}^{-1}(\omega) = G_{i,i}^{-1}(\omega) + \Sigma_{i,i}(\omega)$. This step corresponds to a site exclusion to prevent the overcounting of self-energy diagrams on site i . $\mathcal{G}(\omega)$ defines the undressed Green function of a generalized Anderson impurity model. (4) We solve the associated impurity problem with some technique, e.g., the QMC method, which produces $G_{imp}(\omega)$, the Green function of the generalized Anderson impurity model. (5) Then $\Sigma_{i,i}(\omega) = \mathcal{G}^{-1}(\omega) - G_{imp}^{-1}(\omega)$. $\Sigma_{i,i}(\omega)$ may be used in (2) to continue the procedure. The iteration typically continues until $G_{i,i}(\omega) = G_{imp}(\omega)$ to within the desired accuracy, and the procedure may be shown to be completely causal.

This DMFA algorithm may be applied in any dimension, but it is only exact for $D=\infty$. In finite dimensions, it is very difficult to formulate $1/D$ corrections to the DMFA which are both causal and systematic. For example, consider the first nontrivial correction to the self-energy of a Hubbard model on a hypercubic lattice given by the self-energy diagrams evaluated between nearest-neighbor sites i and j . This contributes a term of order $\mathcal{O}(1/\sqrt{D})$ to the self-energy which then assumes the form $\Sigma(\mathbf{k}, \omega) = \Sigma_{ii}(\omega) + \epsilon_{\mathbf{k}} \Sigma_{ij}(\omega)/t$, where t is the hopping matrix element and $\epsilon_{\mathbf{k}}$ the bare electronic dispersion. Note that when $\Sigma_{ij}(\omega)$ and/or $\epsilon_{\mathbf{k}}$ is large, it is possible for the imaginary part of the self-energy $\text{Im} \Sigma(\mathbf{k}, \omega) > 0$, for some (ω, \mathbf{k}) . The corresponding

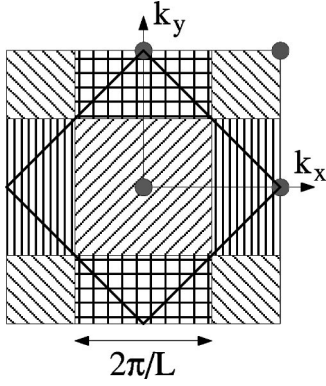


FIG. 2. The cluster momenta and coarse-graining cells for a $N_c=2 \times 2$ cluster covering the Brillouin zone of a real two-dimensional square lattice. The cluster momenta are indicated by filled circles, and the cells by different fill patterns. The solid line in the shape of a diamond is the Fermi surface of the noninteracting system at half filling. The cells adjacent to the BZ boundary extend periodically to the opposite side.

quasiparticle excitations grow exponentially in time; a clear violation of causality.

III. DYNAMICAL CLUSTER APPROXIMATION

For this reason we formulated the DCA approach which includes systematic nonlocal corrections to the DMFA but is not systematic in $1/D$. Like the DMFA, the DCA is a self-consistency scheme, although in the DCA the ‘‘impurity’’ is replaced by a finite-sized cluster. In Appendix C we prove that the DCA is causal, a feature of fundamental importance.¹⁶ The DCA also restores momentum conservation as well as the Ward identities systematically as the cluster size becomes large.

The general form of the DCA was given in Ref. 9. Here, we briefly review the formalism, and then give a more detailed description of the method and its approximations. For simplicity, we consider a single-band model with a local Hubbard-like interaction on a periodic hypercubic lattice with N sites. This is mapped onto a self-consistently embedded periodic cluster of size $N_c=L^D$. As illustrated in Fig. 2, the corresponding crystal momenta \mathbf{K} of the cluster are at the centers of a set of N_c cells of size $(2\pi/L)^D$ inside the first Brillouin zone (BZ) for the lattice. Although there is considerable latitude in the choice of \mathbf{K} , we typically choose $K_{\alpha l} = \pi(2l/L - 1)$ (where l is an integer $1 \leq l \leq L$, and α indicates spatial direction).¹⁷

The crucial assumption of the DCA is that the irreducible self-energy of the cluster $\Sigma^c(\mathbf{K}, \omega)$ and the two-particle irreducible vertex functions of the cluster are good approximations to the irreducible self-energy and vertex functions of the real lattice for values of the lattice momenta inside the cells around the cluster momenta. This assumption is justified if the momentum dependence of the irreducible self-energy and vertex functions of the real system is sufficiently weak; or equivalently, if the dynamical nonlocal correlations have a short range $b \leq L/2$. If this is the case, then, according to Nyquist’s sampling theorem,¹⁸ to reproduce these correlations in the self-energy and vertex functions, we need only sample the reciprocal space at an interval of $\Delta k \approx 2\pi/L$; i.e.,

on a set of $N_c=L^D$ points within the first Brillouin zone. Therefore, $\Sigma(\mathbf{K} + \tilde{\mathbf{k}}, \omega) \approx \Sigma(\mathbf{K}, \omega)$ for each $\tilde{\mathbf{k}}$ within a cell of size $(\pi/b)^D$ about \mathbf{K} , so the lattice self-energy is well approximated by the self-energy $\Sigma^c(\mathbf{K})$ obtained from the cluster. Similar arguments can be made for the vertex functions as well.

Next, within the spirit of the same approximation, the cluster self-energies and vertex functions can be equated with the *coarse-grained averages* of the lattice self-energies and vertex functions over these momentum cells around the cluster momenta. For example, for the self-energy,

$$\Sigma^c(\mathbf{K}, \omega) = \bar{\Sigma}(\mathbf{K}, \omega) = \frac{N_c}{N} \sum_{\tilde{\mathbf{k}}} \Sigma(\mathbf{K} + \tilde{\mathbf{k}}, \omega), \quad (1)$$

where the $\tilde{\mathbf{k}}$ summation runs over the N/N_c momenta of the cell about the cluster momentum \mathbf{K} . This assumption is consistent with that made in the previous paragraph, and ensures that all the states of the full system are represented once the problem is reduced to the cluster. Similar equations can be written down for the vertex functions.

The above two (related) sets of assumptions completely prescribe the DCA and ensure that it reduces to an effective, self-consistently embedded cluster problem for any lattice problem with *local* interactions. For Hubbard-like models such as the HM, PAM, and FKM, within a diagrammatic framework it is not hard to see that the skeleton graph expansions for the coarse-grained self-energies and vertex functions defined above are then the same as the skeleton graph expansions on a finite periodic cluster of size N_c . The cluster Green function $G^c(\mathbf{K}, \omega)$ is given by the *coarse-grained average* of the Green function of the real lattice,

$$G^c(\mathbf{K}, \omega) = \bar{G}(\mathbf{K}, \omega) = \frac{N_c}{N} \sum_{\tilde{\mathbf{k}}} \frac{1}{\omega - \epsilon_{\mathbf{K} + \tilde{\mathbf{k}}} + \mu - \Sigma^c(\mathbf{K}, \omega)}. \quad (2)$$

Here, $\epsilon_{\mathbf{k}}$ is the dispersion for the noninteracting lattice problem and μ is the chemical potential. The DCA assumption that $\Sigma(\mathbf{K} + \tilde{\mathbf{k}}, \omega) \approx \Sigma^c(\mathbf{K}, \omega)$ has been explicitly put in for the lattice Green function.

One can now ask what bare Green function $\mathcal{G}(\mathbf{K})$ on the cluster this skeleton graph expansion corresponds to. The answer is determined by the Dyson equation on the cluster used in reverse,

$$\mathcal{G}^{-1}(\mathbf{K}, \omega) = \bar{G}^{-1}(\mathbf{K}, \omega) + \Sigma^c(\mathbf{K}, \omega). \quad (3)$$

This step corresponds to a ‘‘cluster exclusion’’ to prevent overcounting of self-energy contributions from the interactions on the sites belonging to the cluster, analogous to the ‘‘site exclusion’’ of the DMFA (which is the DCA if the cluster consists of a single site only). It is this step that determines the self-consistent embedding of the cluster, since \mathcal{G} includes the effects of self-energy processes at sites of the lattice other than the cluster sites, and thus has strong retardation effects. The retardation effects can be interpreted in terms of hybridization of the cluster (cells) to ‘‘conduction electron baths’’ (one for each \mathbf{K}) analogously to the interpretation of the single site in DMFA in terms of an Anderson impurity problem.

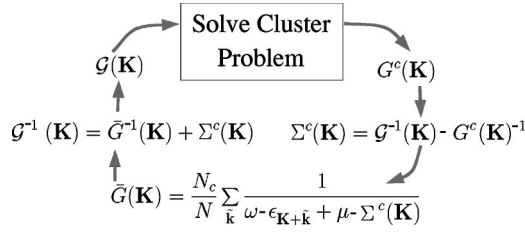


FIG. 3. Sketch of the DCA algorithm.

The DCA iteration procedure is now easily prescribed. It is started by guessing an initial $\Sigma^c(\mathbf{K}, \omega)$, usually zero, which is used to calculate the coarse-grained Green function $\bar{G}(\mathbf{K}, \omega)$ using Eq. (2). The cluster problem is then set up with the bare Green function $\mathcal{G}(\mathbf{K}, \omega)$ given by Eq. (3) and interactions on the cluster sites. $\Sigma^c(\mathbf{K}, \omega)$ may then be calculated using any of a variety of methods, including perturbation theory, QMC, the noncrossing approximation, etc., as appropriate. (If a skeletal graph perturbation expansion is used for the calculation, then the cluster exclusion step may be skipped.) For Green-function techniques, such as QMC, which produce the fully dressed cluster Green function $G^c(\mathbf{K}, \omega)$ rather than the self-energy, the cluster self-energy is calculated as

$$\Sigma^c(\mathbf{K}, \omega) = \mathcal{G}^{-1}(\mathbf{K}, \omega) - G^c(\mathbf{K}, \omega)^{-1}. \quad (4)$$

The iteration closes by calculating a new $\bar{G}(\mathbf{K}, \omega)$ with Eq. (2), and the iteration is continued until $\bar{G}(\mathbf{K}, \omega) = G^c(\mathbf{K}, \omega)$ to within the desired accuracy. The self-consistency loop for the DCA is illustrated in Fig. 3.

In analogous fashion we can also provide prescriptions for calculating two-particle properties of the lattice from the irreducible cluster two-particle self-energies (or vertex functions). Again, the basic assumption is that the momentum dependence of the irreducible vertex function of the real lattice is weak. This is elaborated on in more detail in Appendix A.

For lattice problems with nonlocal interactions such as the extended Hubbard model, the problem is first converted into one that has only local interactions by introducing auxiliary Hubbard-Stratonovich bosonic fields. The DCA can then be prescribed in a straightforward way for this interacting Fermionic-bosonic problem with local interactions. The effective cluster problem will necessarily involve coarse-grained bosonic Green functions as well. The details are given in Appendix B.

IV. DISCUSSION OF THE DCA

In this section we provide a detailed discussion of some of the features of the DCA. We discuss the coarse-graining procedure and offer a simple diagrammatic interpretation. For large but finite D , we show that the DCA includes short-ranged dynamical correlations without resorting to Nyquist's theorem, and we give a simple argument showing its causality.

A. Coarse graining

One can think of other (perhaps more *ad hoc*) prescriptions for the calculation of the cluster self-energies and ver-

tex functions, e.g., using a modified \bar{G} where the coarse graining over \mathbf{k} involves a positive semidefinite weight function $f_w(\mathbf{k}, \mathbf{K})$ which we can choose,

$$\bar{G}(\mathbf{K}, \omega) = \frac{1}{N} \sum_{\mathbf{k}} \frac{f_w(\mathbf{k}, \mathbf{K})}{\omega - \epsilon_{\mathbf{k}} + \mu - \Sigma^c(\mathbf{K}, \omega)}, \quad (5)$$

where the sum on \mathbf{k} is now over the whole Brillouin zone. Our choice of

$$f_w(\mathbf{k}, \mathbf{K}) = N_c \prod_l \Theta\left(\frac{\Delta k}{2} - |k_l - K_l|\right), \quad (6)$$

where $\Delta k = 2\pi/L$ will reproduce the DMFA if the cluster is a single site. In addition, even for larger clusters, the local lattice Green function and the local cluster Green function will be identical given our choice. We note that the choice $f_w(\mathbf{k}, \mathbf{K}) = N\delta(\mathbf{k} - \mathbf{K})$ corresponds to evaluating the system on the finite size cluster without any feedback of the host. For a cluster of one site this is identical to the atomic limit. One could also imagine forms of f_w that allow for overlap of the cells in the Brillouin zone, such as products of Gaussians. However, most $f_w(\mathbf{k}, \mathbf{K})$ different from the two specified above will lead to a calculation which does not have an obvious physical limit for the case of a single site "cluster."

The DCA also has a simple diagrammatic interpretation. For Hubbard-like models, the local Hubbard U is unchanged by the coarse graining, and thus the momentum dependence of each vertex is completely characterized¹⁵ by the Laue function,

$$\Delta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) = \sum_{\mathbf{r}} e^{i(\mathbf{k}_1 - \mathbf{k}_2 + \mathbf{k}_3 - \mathbf{k}_4) \cdot \mathbf{r}}, \quad (7)$$

which expresses the conservation of momenta \mathbf{k}_1 and \mathbf{k}_3 (\mathbf{k}_2 and \mathbf{k}_4) entering (leaving) each vertex. For example, in the conventional diagrammatic approach $\Delta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) = N\delta_{\mathbf{k}_1 + \mathbf{k}_3, \mathbf{k}_2 + \mathbf{k}_4}$. If we reintroduce the cluster and cell momenta, such that $\mathbf{k}_i = \mathbf{K}_i + \tilde{\mathbf{k}}_i$, $i = 1, 4$, then

$$\begin{aligned} \Delta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) &= \sum_{\mathbf{r}} e^{i(\tilde{\mathbf{k}}_1 - \tilde{\mathbf{k}}_2 + \tilde{\mathbf{k}}_3 - \tilde{\mathbf{k}}_4 + \mathbf{K}_1 - \mathbf{K}_2 + \mathbf{K}_3 - \mathbf{K}_4) \cdot \mathbf{r}} \\ &= N_c \sum_n \frac{1}{n!} [(\tilde{\mathbf{k}}_1 - \tilde{\mathbf{k}}_2 + \tilde{\mathbf{k}}_3 - \tilde{\mathbf{k}}_4) \cdot \nabla_{\mathbf{K}_1}]^n \\ &\quad \times \delta_{\mathbf{K}_1 + \mathbf{K}_3, \mathbf{K}_2 + \mathbf{K}_4}. \end{aligned} \quad (8)$$

Within the DCA, only the first term in the sum ($n=0$) is kept so

$$\begin{aligned} \Delta_{DCA}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) &= N_c \delta_{\mathbf{M}(\mathbf{k}_1) + \mathbf{M}(\mathbf{k}_3), \mathbf{M}(\mathbf{k}_2) + \mathbf{M}(\mathbf{k}_4)} \\ &= \Delta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) + \mathcal{O}(\Delta k), \end{aligned} \quad (9)$$

where $\mathbf{M}(\mathbf{k})$ is a function which maps \mathbf{k} onto the momenta label \mathbf{K} of the cell containing \mathbf{k} . Note that with this choice of Laue function the momenta of each internal leg may be freely summed over the cell. Thus, each internal leg $G(\mathbf{k}_1, \omega)$ in the diagram is replaced by $\bar{G}(\mathbf{M}(\mathbf{k}_1), \omega)$ defined by Eq. (2). Furthermore, since each external momenta \mathbf{k} also enters the diagram only through $\mathbf{M}(\mathbf{k})$, the self-energy be-

comes momentum independent within each cell, i.e., it obtains the coarse-grained form defined in Eq. (1) and the approximation $\Sigma(\mathbf{k}, \omega) \approx \bar{\Sigma}(\mathbf{M}(\mathbf{k}), \omega)$ follows as a natural consequence. In the DMFA, the cell momenta extend over the entire Brillouin zone, so that $\Delta_{DMFA}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) = 1$ and momentum conservation is neglected.¹⁵ Thus, the above choices of the Laue function serve as microscopic definitions of the DCA, and of the DMFA. To interpret the choice for the DCA, note that small changes in each of the internal momentum labels will not affect Δ_{DCA} . Thus, momentum conservation for small momentum transfers less than $\Delta k = 2\pi/N_c^{1/D}$ is neglected. However, note that for momentum transfers larger than Δk momentum conservation is (partially) observed at the vertex. Thus, the DCA systematically restores the momentum conservation relinquished by the DMFA as the cluster size increases.

B. Nonlocal corrections

The range of the dynamical correlations included in the DCA is dictated by the cluster size and by the range of the Green functions used to calculate the irreducible graphs. In the DMFA, the self-energy is a functional of the local Green function, but in the DCA nonlocal Green functions also are used. Thus, the DMFA incorporates only local dynamical correlations which occur on the effective impurity, whereas the DCA incorporates nonlocal dynamical correlations which occur on the cluster.

This may be seen by exploring the coarse-graining step in detail, and in real space. For this purpose, we consider a lattice in large but finite D which we divide into L^D -sized clusters. Let \mathbf{r} denote vectors within a cluster, and \mathbf{R} the vectors between the centers of the clusters. The points of the original lattice can be represented as $\mathbf{R} + \mathbf{r}$. The relation between the real Green function $G(\mathbf{R} + \mathbf{r}, \omega)$ and the cluster Green function $\bar{G}(\mathbf{r}, \omega)$ is given by

$$\bar{G}(\mathbf{r}, \omega) = \frac{1}{N} \sum_{\mathbf{K}, \mathbf{k}} \sum_{\mathbf{R}, \mathbf{r}'} e^{i\mathbf{K} \cdot (\mathbf{r} - \mathbf{r}')} e^{-i\mathbf{k} \cdot (\mathbf{R} + \mathbf{r}')} G(\mathbf{R} + \mathbf{r}', \omega). \quad (10)$$

The sum over \mathbf{K} forces $\mathbf{r}' = \mathbf{r}$. For $\mathbf{R} = 0$ the additional phase factor $e^{-i\mathbf{k} \cdot \mathbf{r}}$ is essentially 1 over the entire range of \mathbf{k} for short distances on the cluster $r \ll 2\pi/\Delta k$, which leads to a contribution to $\bar{G}(\mathbf{r}, \omega) \approx G(\mathbf{r}, \omega)$. Contributions from larger \mathbf{R} are suppressed both by the oscillations in the phase factor which suppresses the integral and from the smallness of $G(\mathbf{R} + \mathbf{r}', \omega)$ itself. More precisely, with the choice $K_{\alpha l} = \pi(2l/L - 1)$ (where l is an integer $1 < l < L$, and α indicates spatial direction), we can complete the sums on momenta exactly to obtain

$$\bar{G}(\mathbf{r}, \omega) = \sum_{\mathbf{R}} \prod_{l=1}^D \left(\frac{\sin[\pi(x_l + X_l)/L]}{\pi(x_l + X_l)/L} \right) G(\mathbf{R} + \mathbf{r}, \omega), \quad (11)$$

where x_l (X_l) is the l th component of the vector \mathbf{r} (\mathbf{R}). Thus, $\bar{G}(\mathbf{r}, \omega)$ is composed of a sum over $G(\mathbf{r} + \mathbf{R}, \omega)$ with each term weighted by a sinusoidal prefactor that falls off like $|\mathbf{r} + \mathbf{R}|^{-D}$. For small r , the leading term in the sum comes from $\mathbf{R} = 0$. Then, by expanding the sinusoidal prefactor, we can

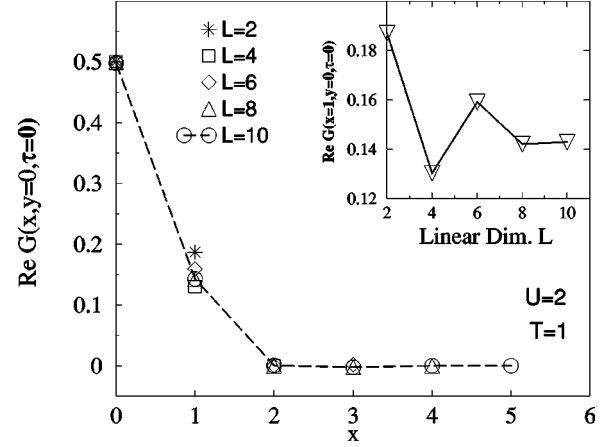


FIG. 4. $\text{Re } \bar{G}(x, y=0, \tau=0)$ versus x for various cluster sizes, obtained from QMC simulations of the FKM. In the inset, $\text{Re } \bar{G}(x=1, y=0, \tau=0)$ is plotted versus cluster size (periodic clusters of size $L \times L$).

see that for $\mathbf{r} = 0$, $\bar{G}(\mathbf{0}, \omega) = G(\mathbf{0}, \omega)$, and for $r \ll L/2$, $\bar{G}(\mathbf{r}, \omega) \approx G(\mathbf{r}, \omega) + \mathcal{O}((r\Delta k)^2)$. Contributions from $G(\mathbf{r} + \mathbf{R}, \omega)$ for finite values of \mathbf{R} are cutoff by the sinusoidal prefactor and the exponential fall-off of the Green function itself, since for large distances $G(r) \sim D^{-r/2}$. Thus, short-ranged correlations are accurately represented by $\bar{G}(\mathbf{r}, \omega)$, and longer-ranged contributions are cut off.

This behavior is seen even in two-dimensional systems, as shown in Fig. 4 where $\bar{G}(x, y=0, \tau=0)$ calculated with a QMC simulation of the two-dimensional half-filled FKM (see Sec. V) is plotted versus x for various cluster sizes. The $\mathbf{r} = 0$ result is fixed by the filling, $\bar{G}(x=0, y=0, \tau=0) = 0.5$; however, the near-neighbor result shows some significant dependence on the cluster size. $\bar{G}(x=1, y=0, \tau=0)$ is plotted versus the linear cluster size in the inset to Fig. 4. Note that it quickly converges to $\bar{G}(x=1, y=0, \tau=0) \approx 0.143$ as the cluster size increases, indicating that short-ranged correlations are correctly described by the DCA for this model. For larger x , $\bar{G}(x, y=0, \tau=0)$ falls quickly to nearly zero.

C. The role of reducible and irreducible quantities

In Appendix D we show that the DCA (and the DMFA) is not conserving, thus the calculations of different measurable quantities are not unique. For example, we approximate the lattice self-energy $\Sigma(\mathbf{k}, \omega) \approx \bar{\Sigma}(\mathbf{M}(\mathbf{k}), \omega)$, and calculate the Green function using $1/G(\mathbf{k}, \omega) = 1/G^0(\mathbf{k}, \omega) - \bar{\Sigma}(\mathbf{M}(\mathbf{k}), \omega)$; however, a different approximation, corresponding to a different implicit choice for $\Sigma(\mathbf{k}, \omega)$ would be to approximate $G(\mathbf{k}, \omega) \approx \bar{G}(\mathbf{k}, \omega)$. We show in Appendix A that the former prescription is the unique choice which minimizes the DCA free energy, and thus is the correct choice. A similar problem exists for the calculation of two-particle properties such as the magnetic susceptibility. However, as discussed in Appendix A, the approximation $\Gamma \approx \bar{\Gamma} \equiv \delta \bar{\Sigma} / \delta \mathbf{G}$ for the lattice two-particle vertex yields an estimate for the susceptibility [Eq. (A15)] equivalent to that calculated from the second derivative of the free energy with respect to the external field.

Thus, in general, *the cluster calculation should only be used to provide the irreducible quantities*. These, together with the bare real-lattice Green functions, may be used to construct the corresponding reducible quantities.

At least for the single-particle Green functions, this prescription may also be motivated physically. Short-ranged correlations are accurately represented by the cluster irreducible single-particle self-energy. Following the discussion of the preceding section, one may show that for $r \ll L/2$, $\Sigma^c(\mathbf{r}, \omega) \approx \Sigma(\mathbf{r}, \omega) + O((r\Delta k)^2)$, since it is calculated from cluster quantities. In addition, since the self-energy is formed from higher-order products of the Green function, e.g., $\Sigma(\mathbf{r}) \sim [G(\mathbf{r})]^3 \sim D^{-3r/2}$ for the second-order contribution in the Hubbard model, in high dimensions it falls faster with increasing r than the Green function itself. Thus, the correction terms coming from $\mathbf{R} \neq 0$ will be smaller for irreducible quantities such as the self-energy than it will be for reducible quantities like the Green function. Since the range of the correlations that are treated increases with the cluster size, away from a transition, the irreducible quantities calculated on the cluster will have converged to acceptable values before their reducible counterparts.

Finally, we note that while in Secs. IV A and IV B we used $1/D$ arguments to justify the approximations made in the DCA, the DCA is not systematic in $1/D$. For example, even for short distances r , which would correspond to low orders in $1/D$, $\bar{G}(\mathbf{r}, \omega)$ contains contributions $G(\mathbf{r} + \mathbf{R}, \omega)$ corresponding to much larger distances and higher orders in $1/D$. Furthermore, since the density of states of the finite-dimensional lattice is used to calculate the host propagator \mathcal{G} , the approximation includes corrections to all orders in $1/D$. In fact, we have shown in this section that the cluster quantities differ from those of the real lattice by terms of order $(\Delta k)^2 = 4\pi^2/N_c^{2/D}$. Thus, the DCA is a systematic approximation in $1/N_c$, not $1/D$.

D. Causality

We can also show that the DCA algorithm is fully causal, i.e., that the spectral weight is conserved and that the imaginary parts of the single-particle retarded Green functions and self-energies are negative definite. Here, since many methods can be used to solve the cluster problem, we will assume that all are causal, i.e., given a causal \mathcal{G} , then the resulting Σ^c and G^c are also ensured to be causal by the method chosen to solve the cluster problem. Furthermore, $\bar{G}(\mathbf{K}, \omega)$ is causal since $\Sigma^c(\mathbf{K}, \omega)$ is causal. Thus, Eq. (3) is the only step in the algorithm where problems with causality could occur. In Ref. 9 we argued using a continued fraction expansion that the $\tilde{\mathbf{k}}$ averaging (coarse graining) of Eq. (2) adds a causal piece to the self-energy of \bar{G} that allows \mathcal{G} to remain causal even after the subtraction of $-\Sigma^c(\mathbf{K}, \omega)$ in Eq. (3). Here, we give a simple *geometrical* argument (which is recast as a formal proof in Appendix C) that causality holds for rather general models, including the HM and the FKM.

There are two steps to the argument: first, we must show that weight is conserved, and second, that the imaginary part of \mathcal{G} is negative semidefinite. The first part follows from the causality of Σ^c and \bar{G} which both fall off inversely with frequency at large ω , and in particular $\bar{G} \sim 1/\omega$. From Eq. (3)

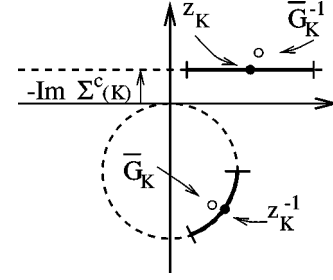


FIG. 5. Illustration of the essential steps of the proof that the DCA is causal (see text).

it is then apparent that $\mathcal{G} \sim 1/\omega$ so that spectral weight is preserved. The second part of the argument is sketched in Fig. 5. The imaginary part of $\mathcal{G}(\mathbf{K}, \omega) = [\bar{G}(\mathbf{K}, \omega)^{-1} + \Sigma^c(\mathbf{K}, \omega)]^{-1}$ is negative provided that $\text{Im}[\bar{G}(\mathbf{K}, \omega)^{-1}] \geq -\text{Im}\Sigma^c(\mathbf{K}, \omega)$. $\bar{G}(\mathbf{K}, \omega)$ can be written as $\bar{G}(\mathbf{K}, \omega) = (N_c/N)\sum_{\tilde{\mathbf{k}}}(z_{\mathbf{K}+\tilde{\mathbf{k}}})^{-1}(\omega)$, where the $z_{\mathbf{K}+\tilde{\mathbf{k}}}(\omega)$ are complex numbers with a positive semidefinite imaginary part $-\text{Im}\Sigma^c(\mathbf{K}, \omega)$. For any \mathbf{K} and ω , the set of points $z_{\mathbf{K}+\tilde{\mathbf{k}}}(\omega)$ are on a segment of the dashed *horizontal* line in the upper half plane due to the fact that the imaginary part is *independent* of $\tilde{\mathbf{k}}$. The mapping $z \rightarrow 1/z$ maps this line segment onto a segment of the dashed circle shown in the lower half plane. $\bar{G}(\mathbf{K}, \omega)$ is obtained by summing the points on the circle segment, yielding the empty dot that must lie *within* the dashed circle. The inverse necessary to take $\bar{G}(\mathbf{K}, \omega)$ to $1/\bar{G}(\mathbf{K}, \omega)$ maps this point onto the empty dot in the upper half plane which must lie *above* the dashed line. Thus, the imaginary part of $\bar{G}(\mathbf{K}, \omega)^{-1}$ is greater than or equal to $-\text{Im}\Sigma^c(\mathbf{K}, \omega)$. This argument may easily be extended for $\mathcal{G}(z)$ for any z in the upper half plane. Thus \mathcal{G} is completely analytic in the upper half plane.

V. DCA FOR THE FALICOV-KIMBALL MODEL

Here we illustrate the power of the DCA with a QMC simulation of the two-dimensional Falicov-Kimball model. The FKM is studied, instead of, for example, the much more complicated Hubbard model (for which there is work in progress¹⁹), for several reasons. First, the FKM is perhaps the simplest model of correlated electrons which retains a complex phase diagram, including a Mott transition and a charge-density-wave (CDW) ordering transition.²⁰ Second, it has been extensively studied by de Vries *et al.* with QMC simulations²¹ of finite-sized systems which may be compared to our results. Third, it is possible²¹ to calculate the real-frequency spectra without the need for computationally expensive numerical analytic continuation. Finally, it is of considerable experimental interest.²²

The FKM can be considered as a simplified Hubbard model in which one spin species is prohibited to hop. In the particle-hole symmetric case the Hamiltonian reads

$$H = -t \sum_{\langle i,j \rangle} d_i^\dagger d_j - \mu \sum_i (n_i^d + n_i^f) + U \sum_i n_i^d n_i^f, \quad (12)$$

with $n_i^d = d_i^\dagger d_i$, $n_i^f = f_i^\dagger f_i$, and $\mu = U/2$. For a two-dimensional (2D) square lattice with nearest-neighbor hop-

ping ($\langle i, j \rangle$) the dispersion is $\epsilon_{\mathbf{k}} = -2t(\cos k_x + \cos k_y)$. We measure energies in units of the hopping element t . Consequently, the bandwidth of the noninteracting system is $W = 8$. For $D \geq 2$ the system has a phase transition from a homogeneous high-temperature phase with $\langle n_i^d \rangle = \langle n_i^f \rangle = 1/2$ to a checkerboard phase [a charge density wave with ordering vector $\mathbf{Q} = (\pi, \pi, \dots)$] with $\langle n_i^d \rangle \neq \langle n_i^f \rangle$ for $0 < U < \infty$.²³

A. Exact enumeration

In contrast to the Hubbard and related models, the DCA for the FKM can be solved without the application of QMC since the f electrons are static, acting as a kind of annealed disorder potential to the dynamic d electrons. Here, we generalize the algorithm of Brandt and Mielsch²⁴ to a finite-size cluster. We first compute the Boltzmann weights w_f of all configurations $\{f\}$ of f electrons on the cluster, given an initial host Green function \mathcal{G}_{ij} of the d electrons via $w_f = w_f^0/Z$, where

$$w_f^0 = 2^{N_c} \prod_{\omega_n} \det \frac{\mathcal{G}_{ij}^{-1}(i\omega_n) - U n_i^f \delta_{ij}}{i\omega_n \delta_{ij}} \quad (13)$$

is the unnormalized weight, and $Z = \sum_{\{f\}} w_f^0$ is the ‘‘partition sum.’’ The determinant is to be taken over the spatial indices. This expression is written such that the product converges at large frequencies. Given the weights, the new d -electron cluster Green function is given by

$$G_{ij}^c(z) = \sum_{\{f\}} w_f [\mathcal{G}_{ij}^{-1}(z) - U n_i^f \delta_{ij}]^{-1} \quad (14)$$

for an arbitrary complex frequency argument z , in particular also for $z = i\omega_n$ (Matsubara) and $z = \omega + i\eta$ (retarded). The self-consistency loop closes by use of the Eqs. (2), (3), and (4).

Because the number of f configurations grows exponentially with the cluster size the exact enumeration method is confined to small clusters (up to 4×4 in the broken symmetry state, see below). We first simultaneously determine the weights and the Matsubara Green function. Then, we use knowledge of the weights to find the retarded Green function. Convergence of the algorithm is fast for Matsubara frequencies, but relatively slow for real frequencies.

B. Quantum Monte Carlo

The FKM is particularly suitable to a QMC evaluation of the configuration sums since the f electrons are themselves like classical Ising spin variables. Following De Raedt and von der Linden,²¹ given a particular configuration, we can propose ‘‘spin flips,’’ corresponding to a change of the f -occupation $n_i^f \rightarrow 1 - n_i^f$ at a single site i . The ratio R of weights w_f' of the proposed configuration to the weight w_f of the original configuration is (at half filling)

$$R = \prod_{\omega_n > 0} [1 - \lambda_i G_{i,i}^c(i\omega_n)] [1 - \lambda_i G_{i,i}^{c*}(i\omega_n)], \quad (15)$$

with $\lambda_i = -Us(i)$ and $s(i) = 2n_i^f - 1$. Note that the ratio R is always real and positive since the Matsubara Green function is Hermitian $G_{i,i}^c(-i\omega_n) = G_{i,i}^{c*}(i\omega_n)$. This holds for any fill-

ing. Consequently, there is no sign problem as there is, e.g., in the Hubbard model away from half filling.

A configuration change is accepted by comparing a random number in the interval (0,1) to $R/(1+R)$ (‘‘heat bath method’’) or to R itself (‘‘Metropolis method’’). Once the change at site i is accepted, the Green function is updated via

$$G_{j,k}^{c'}(i\omega_n) = G_{j,k}^c(i\omega_n) + \frac{\lambda_i G_{j,i}^c(i\omega_n) \otimes G_{i,k}^c(i\omega_n)}{1 - \lambda_i G_{i,i}^c(i\omega_n)}, \quad (16)$$

where \otimes denotes a direct matrix product (no summation). Most of the total CPU time is consumed by this updating step. However, the fact that we can work with frequencies rather than imaginary time drastically reduces the amount of time required. Note that although Eq. (16) is written for Matsubara Green functions an analogous relation holds for the real frequency Green functions which allows us to calculate dynamical properties without the need for analytic continuation. On the other hand, the ratio R is completely determined by the Matsubara Green function. This means that we determine the acceptance from the Matsubara Green function and then update both the Matsubara and the real-frequency retarded Green function ‘‘simultaneously.’’

The measurement of the two-particle properties consumes large amounts of memory and CPU time. Since they are not required for the self-consistency cycle (Fig. 2), they are measured only after convergence of the single-particle properties. In fact, due to the enormous size of the susceptibility matrix it is often worthwhile to separate the single- and two-particle calculations to different computer runs.

VI. RESULTS

In this section we present results from both exact enumeration and QMC simulation of the two-dimensional FKM for a variety of parameters and cluster sizes. There is considerable latitude in the selection of the cluster momenta. This is because (i) the sites on the cluster do not really correspond to the physical lattice, and (ii) because for large clusters any differences due to this choice should vanish. Here, for an $L \times L$ cluster we choose either $K_{\alpha l} = \pi(2l/L - 1)$, or $K_{\alpha l} = \pi(2l/L - 1) - \pi/L$ (where l is an integer $1 \leq l \leq L$, and $\alpha = x$ or y). These choices, respectively, correspond to periodic or antiperiodic boundary conditions for the cluster Green function $G^c(x+L, y, \omega) = \pm G^c(x, y, \omega)$. Unless otherwise noted we use periodic boundary conditions in both spatial directions. Antiperiodic boundary conditions (in both directions) are used only for some data in Fig. 10.

A. Density of states and spectral function

We begin by discussing the (local) density of states (DOS) and the \mathbf{K} -dependent spectral function shown in Figs. 6–8. In Fig. 6 we show the local DOS for various cluster sizes up to 8×8 for the half-filled model and display only the positive frequencies. The full spectrum is symmetric, due to particle-hole symmetry, as shown in the inset. With the exception of a peak which develops at $\omega = \pm U/2$, the spectrum converges quickly as N_c increases. In fact, the convergence to the thermodynamic limit is apparently much faster than that seen in finite-sized lattice simulations,²⁰ where even

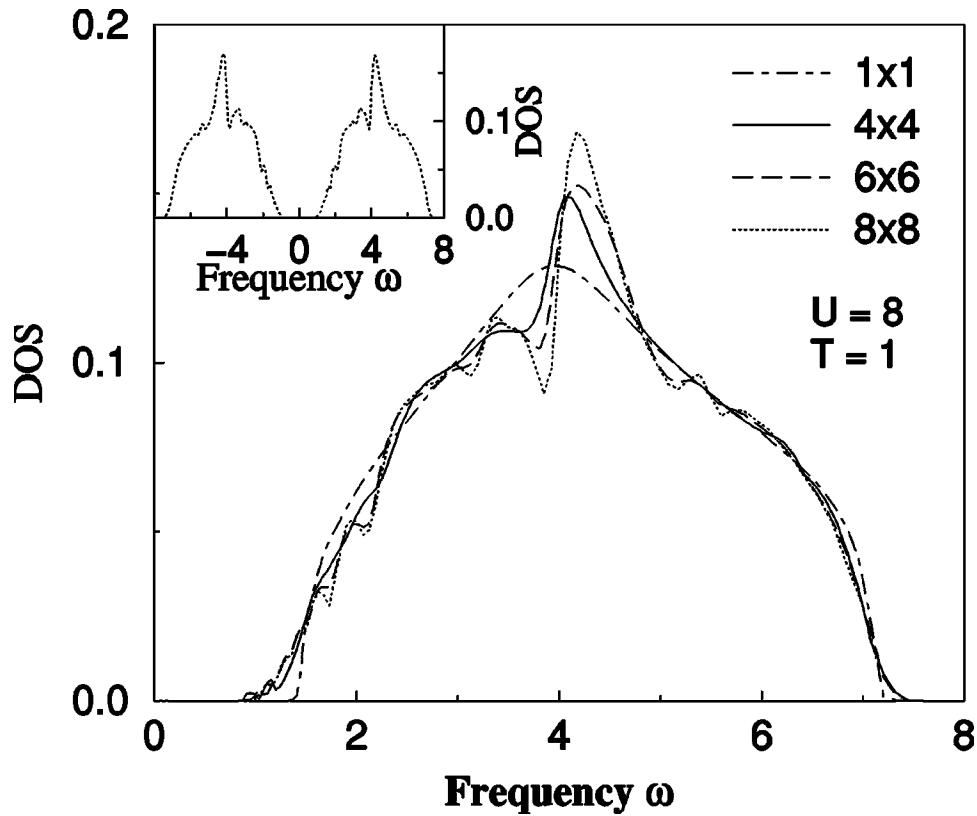


FIG. 6. Local density of states for various cluster sizes. The density of states is essentially converged for the 6×6 cluster, though some fine structure near $\omega = \pm U/2$ continues to emerge for the larger cluster sizes (see discussion in text).

for an 8×8 system, the broadened spectra are often composed of a set of discrete spikes.

Furthermore, the DOS develops three distinct primary features also seen in the finite-size calculations.²⁰ First, as shown in Fig. 6, for large $U \geq U_M$ the DOS develops a Mott gap centered at $\omega = 0$, even though $T \gg T_c$. The value of U_M at this temperature changes slowly with cluster size, with $U_M \approx 5$. Second, as shown in Fig. 7, for $U < U_M$, upon decreasing the temperature the DOS for $N_c > 1$ develops a pseudogap at the Fermi energy associated with charge-

ordering fluctuations. This pseudogap is absent when $N_c = 1$ (as are the charge-ordering fluctuations), and it becomes more pronounced as the cluster size increases. Third, as the charge ordering becomes more pronounced, either by lowering the temperature or increasing the cluster size, a sharp peak begins to develop in the DOS shown in Figs. 6 and 7 at $\omega = \pm U/2$. In the ordered state, each occupied f (d) orbital is surrounded by four occupied d (f) orbitals. Thus, for large U and low T the electrons become highly localized so the spec-

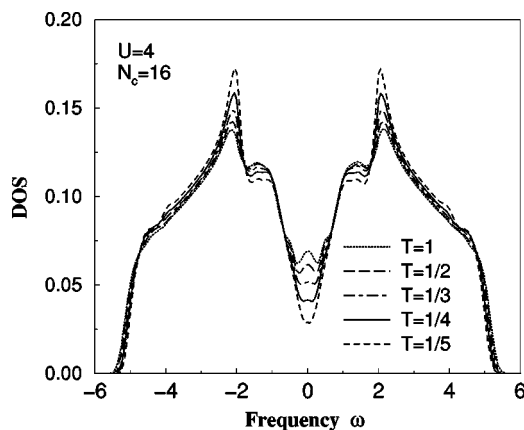


FIG. 7. Local density of states when $U=4$ for a 4×4 cluster at various temperatures. The DOS develops a pseudogap as the temperature approaches $T_c \approx 0.189$. This shows the influence of the nonlocal CDW fluctuations present in the DCA ($N_c > 1$). In the DMFA ($N_c = 1$), there is no T dependence of the DOS above T_c .

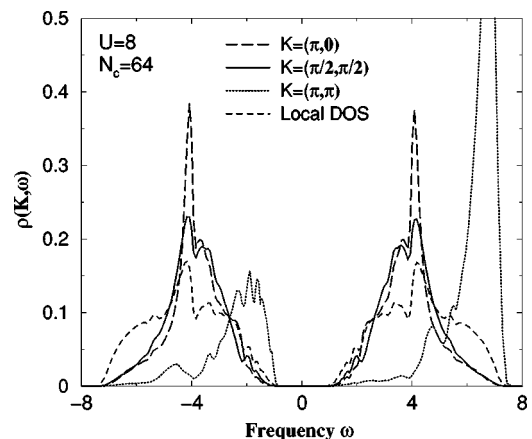


FIG. 8. Spectral function $\rho(\mathbf{K}, \omega)$ for various cluster momenta \mathbf{K} . Note the three peak feature for $\mathbf{K} = (\pi, \pi)$ at the upper edge of the lower band.

trum will develop very narrow “atomic” peaks at $\omega = \pm U/2$.

In addition, there are a surprising number of smaller features which emerge in the DOS. This is true even for the largest cluster, in some sense even more so, as some fine structure in Fig. 6 seems to develop for the 8×8 cluster that was only vaguely present for smaller clusters. This fine structure is more visible in the momentum-resolved spectral function $\rho(\mathbf{K}, \omega) = (1/\pi) \text{Im} G(\mathbf{K}, \omega)$, see Fig. 8. In particular, note the three peak feature at negative frequencies for $\mathbf{K} = (\pi, \pi)$. Of course, we really do not know how the DOS for the infinite lattice is supposed to look like. The extremely smooth form the DMFA provides is mostly due to the lack of associated energy scales. In the DCA we have at least U and $J = t^2/2U$, and, in principle, many other scales can be constructed representing collective excitations of the cluster with charges. That such features emerge as the cluster size is increased can be understood by the following argument. In addition to the self-energy arising from interactions on the cluster the host also provides a self-energy and therefore a broadening. Consequently, features that are in principle present for smaller clusters like 4×4 are washed out by the host’s broadening. Only as the host becomes less important (as cluster size increases) do the smaller energy features emerge from the background.

B. Phase diagram and finite-size scaling

We now discuss the phase diagram and its dependence on cluster size. In Ref. 9 we showed that the transition temperature of the CDW transition was significantly suppressed with respect to the DMFA when nonlocal correlations come into play. We have since extended this analysis in two directions.

In Ref. 9 the result for the 2×2 cluster was computed via the exact enumeration method in the broken symmetry phase. This means we actually simulated two 2×2 clusters forming a bipartite cluster of $2 \times 2 \times 2 = 8$ sites. The extension of the above described exact enumeration method is straightforward and involves Green functions that are now 2×2 matrices with respect to the bipartite cluster (A and B sublattice index). T_c was then obtained by three steps: (1) We apply a staggered field at low enough temperatures (below the expected T_c) to drive the system into the broken-symmetry state with $\langle n_{i \in A}^d \rangle \neq \langle n_{j \in B}^d \rangle$. (2) We remove the staggered field. The system relaxes but stays in the broken if $T < T_c$. (3) We increase T until the system enters the uniform phase with $\langle n_{i \in A}^d \rangle = \langle n_{j \in B}^d \rangle$. This method is very precise, but for larger clusters very time consuming. Using the QMC method in the broken-symmetry phase is possible, but T_c cannot be determined precisely due to critical fluctuations. So the above described method is limited to at most 4×4 clusters, or a total of 32 sites. This also means that a systematic finite-size analysis with this method alone would not be possible.

In order to get T_c for larger clusters we choose a different route. We compute the staggered charge susceptibility $\chi(\mathbf{Q} = (\pi, \pi))$ with the method discussed in Appendix A. Because the host always provides a mean-field environment, the susceptibility diverges as $\chi(\mathbf{Q} = (\pi, \pi)) \propto (T - T_c)^{-\gamma}$ with a mean-field exponent $\gamma = 1$ for T close enough to T_c . (Critical fluctuations cause γ to deviate from the mean-field value

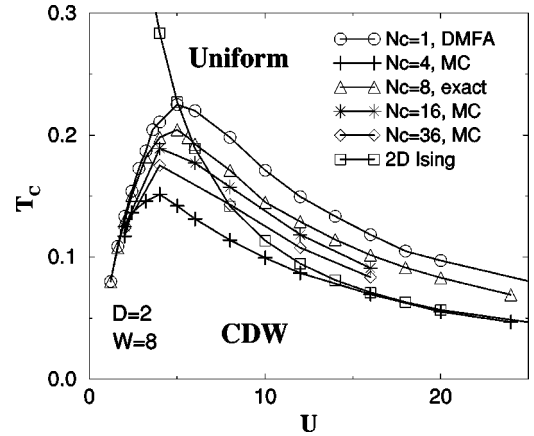


FIG. 9. Phase diagram for various cluster sizes N_c . With the exception of $N_c = 4$ (see text) the T_c monotonically converge with increasing cluster size. At large U the system maps to a 2D Ising model with $J = 1/(2U)$.

for somewhat larger values of $T - T_c$.) This again allows a precise estimate of T_c . The computational drawback here is the enormous memory requirements of the susceptibility matrix needed at intermediate steps of the calculation.

After these preliminaries we now discuss the results of these calculations in Figs. 9 and 10. Figure 9 shows the phase diagram for various cluster sizes, all of them equipped with periodic boundary conditions (PBC). In addition, we show the T_c of the 2D Ising model given by $T_c^{\text{Ising}} = 2.268J$ with a coupling $J = 1/(2U)$. We show the Ising result because the half-filled FKM reduces to an Ising model with such a coupling in the limit of large $U \gg W$. The FKM data are all obtained from the evaluation of the susceptibility with the MC method except for the $N_c = 8$ data which are obtained by the exact enumeration method in the broken symmetry (two 2×2 clusters). For the DMFA the two methods give identical results (within 1% accuracy). The phase boundary has always the same general shape for the FKM data, with a slightly cluster size dependent maximum at about half the bandwidth W .

The results from the MC method converge monotonically with cluster size with one notable exception: The 2×2 cluster ($N_c = 4$) has the lowest T_c of all, and even seems to fall below the Ising results for all U . The reason for this exceptional behavior is not entirely clear to us. At first one might consider a double counting of neighbors and a resulting doubling of the energy scale common in standard lattice methods to be the reason. But clearly, the T_c 's of all clusters agree well at small U where only local correlations are important. This rules out a simple doubling of the energy scale. A likely reason for this unusual behavior lies in the particular way the BZ is sampled in the 2×2 cluster, see Fig. 2. The only points on the Fermi surface are $\mathbf{K} = (\pi, 0), (0, \pi)$. These, however, are also the points responsible for the van Hove singularity of the noninteracting system. In comparison to other momenta on the Fermi surface these points have extraordinary large scattering rates, making them unfavorable for the formation of CDW fluctuations driving the transition. As a consequence, the T_c for this cluster is exceptionally low.

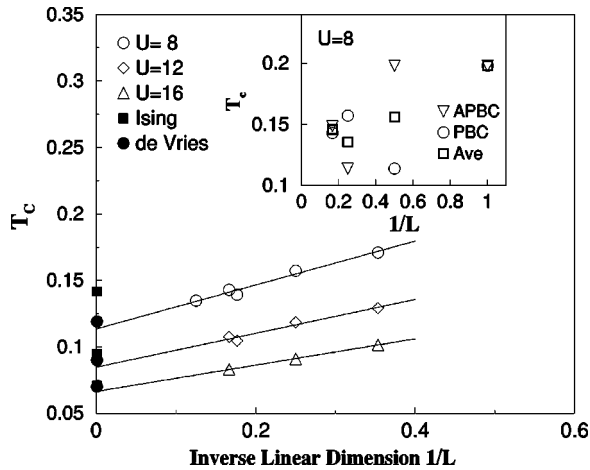


FIG. 10. T_c as a function of inverse linear cluster dimension for the larger clusters and various U . The Ising limit, and de Vries *et al.*²⁰ estimates of T_c from simulations of finite-sized clusters are shown for comparison. The extrapolated T_c 's generally fall below the finite-size estimates as well as the Ising limit (which should serve as an upper bound and become exact for large U). The inset shows the influence of the cluster boundary conditions on T_c . The effect of boundary conditions becomes smaller with increasing cluster size.

Although the T_c results from a given method are monotonically decreasing (with the one exception noted above) it is not obvious how to scale the data as a function of cluster size; for, to our knowledge, a rigorous finite-size scaling theory for a *quantum-dynamical* cluster coupled to a *quantum-dynamical* host does not exist. However, such questions have been addressed in the context of systematic self-consistent cluster approximations for *classical* statistical systems, in particular, the 2D Ising model,²⁵ which should be relevant to our problem, at least for large U . Furthermore, on general grounds one expects that for critical phenomena at *finite temperatures* the asymptotic scaling properties even of a *quantum* system will be determined by the same universality class as for the corresponding *classical* system (i.e., with the same order-parameter symmetry and the same spatial dimensionality). Hence, one expects²⁵ that our results for $T_c(L) - T_c(\infty)$ should scale asymptotically as $L^{-1/\nu}$, i.e., as $1/L$, since $\nu=1$ for the 2D Ising Model. In Fig. 10 we therefore plot the T_c data as a function of $1/L$ (or $1/\sqrt{N_c}$ for the broken-symmetry results). In the main part of the plot we show the results for large clusters with PBC which scale approximately linearly with $1/L$. The $N_c=32$ result (broken symmetry) for $U=8$ and $U=12$ is a bit lower than the T_c for $N_c=36$ (MC). This shows that the two methods are not easy to combine, but the difference seems small enough not to disrupt the predominant linear scaling with $1/L$.

For $U=16$ the cluster T_c 's scale well and the extrapolation to the infinite system comes very close to the Ising limit (or the results of de Vries *et al.*). For smaller U the Ising model is not appropriate, and it shows, as the Ising T_c is much higher than the extrapolated T_c of the clusters. However, the extrapolated cluster results are very close to the results obtained from finite-sized lattice simulations. The fact that the cluster estimates of T_c consistently fall below de Vries results is likely due to finite-sized effects (de Vries *et al.* simulated lattices of up to 64 sites). We also note that

the T_c 's of the 2×2 cluster (not shown in Fig. 10) are in excellent agreement with the cluster extrapolated values and the Ising result for large U . We have currently no explanation for this phenomenon. Though probably pure coincidence, the fact remains: the T_c of the 2×2 cluster seems to provide a good estimate of the T_c of the $D=2$ FKM.

The inset shows the same T_c 's as in the main plot (all determined via MC) for $U=8$ of various cluster sizes, and in addition the T_c 's for the same clusters equipped with antiperiodic boundary conditions (APBC). As noted before, the DCA does not intrinsically determine the choice of cluster momenta. But different choice of cluster momenta will also in general affect T_c and other quantities. As PBC and APBC seem to span the entire range it is interesting to see by how much the T_c 's differ. As illustrated in the inset it matters quite a bit for very small clusters, but not much once we consider clusters of the 6×6 size.²⁶ The difference for 2×2 clusters is extreme for the following reason: we noted above that the 2×2 cluster with PBC has the lowest T_c of all clusters with PBC. The 2×2 cluster with APBC, on the other hand, is identical to the single site cluster (which has the maximum T_c) by virtue of the symmetry of the square lattice.²⁷ Similarly, the 4×4 cluster with APBC is by symmetry identical to the 2×2 cluster with PBC. But once we go to cluster sizes beyond this such identifications are no longer possible. Concurrently, the T_c 's of the clusters also depend less and less on the boundary conditions (of course, boundary conditions are irrelevant in the thermodynamic limit). For 6×6 clusters the difference is down to about 5%.

C. Energy, entropy, and specific heat

The DCA differs from the DMFA through the introduction of nonlocal dynamical correlations. For example, in the FKM, the DCA exhibits fluctuations associated with charge ordering that are absent in the DMFA. To illustrate this, we calculated specific heat divided by the temperature shown in Fig. 11, using a recently developed maximum-entropy method.²⁸ The DMFA ($N_c=1$) result displays a single peak in C/T associated with the suppression of local charge fluctuations and the formation of the Mott gap in the single-particle density of states (Fig. 6). As shown in the inset to Fig. 11, the integrated weight in the peak is $0.69 \approx \ln(2)$; however, the infinite temperature entropy $\int_0^\infty (C/T) dT = 2 \ln(2)$ for the half-filled model. Thus, only half of the entropy is quenched, with the remainder associated with the disorder in n_f ; i.e., $n_f=0$ or $n_f=1$ with equal probability on each site when $N_c=1$, regardless of the configurations of neighboring sites. However, when $N_c=4$, C/T displays an additional lower-temperature peak slightly below $T=T_c$. We believe this peak is due to critical fluctuations associated with charge ordering.

To test the identification of the two peaks seen in the DCA specific heat, we plot $C(T)$ for a variety of values of U when $N_c=4$ in Fig. 12. The location of the upper peak increases monotonically with U , consistent with the association of this peak with local charge fluctuations. However, the location of the lower peak does not depend monotonically on U , but rather changes in rough proportion to the CDW ordering temperature shown in Fig. 9. Similar results have been obtained in Ref. 20, though we want to point out that in

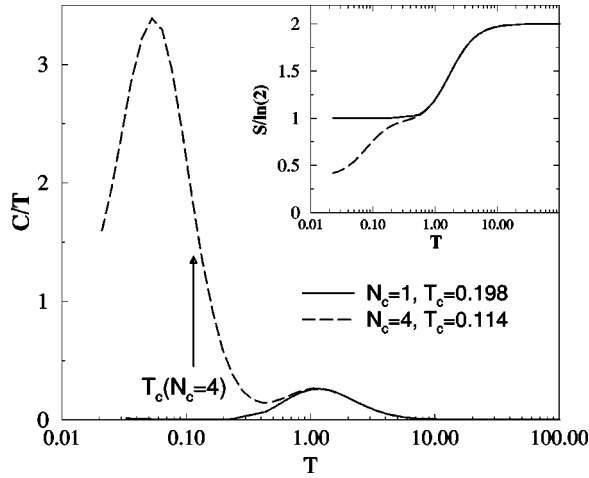


FIG. 11. Specific heat versus temperature for one- and four-site clusters calculated with exact enumeration when $U=8$. For $N_c=1$, there is a single peak with integrated weight $\ln(2)$ associated with the suppression of local charge fluctuations. For $N_c=4$, there is an additional peak at lower temperatures associated with critical fluctuations near the charge ordering transition temperature. T_c for $N_c=4$ is indicated by an arrow. The entropy $S(T') = \int_0^{T'} dT [C(T)/T]$ is shown in the inset divided by $\ln(2)$.

our case the position of the lower peak is below T_c for the given parameters. The rise of this lower peak with U for low U (below the maximum T_c and the opening of the Mott gap) is similar to the half-filled Hubbard model.²⁹

The total entropy in these lower peaks can be substantial. For example, when $U=8$, the entropy $S(T') = \int_0^{T'} dT [C(T)/T]$ in the lower peak is 0.41 whereas that in the upper peak is $0.69 \approx \ln 2$. Thus, the fluctuations associated with charge ordering quench most of the entropy needed to form a proper ground state with $S=0$.

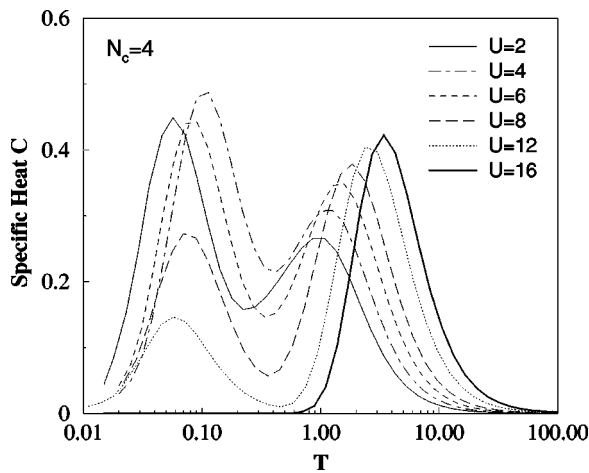


FIG. 12. Specific heat versus temperature for four-site clusters calculated with exact enumeration. The position and height of the lower peak, associated with charge ordering, is nonmonotonic in U . For small U the peak rises and moves to higher temperatures, for large U the trend is opposite. This tracks the behavior of T_c with U . The upper peak, associated with local (Mott) charge fluctuations, moves higher temperatures and becomes more pronounced as U increases.

VII. CONCLUSIONS

We described in detail the recently introduced⁹ dynamical cluster approximation (DCA) and explained its assumptions and approximations. The DCA systematically introduces nonlocal corrections to the DMFA. The DMFA is recovered by taking the cluster to be a single site, whereas the exact result is obtained when the cluster becomes large. We have shown explicitly that the DCA is causal, systematic, and Φ derivable. Furthermore, as the cluster size increases, it systematically restores momentum conservation neglected in the DMFA. Consequently, the DCA becomes conserving in the thermodynamic limit. We have applied it to an exact enumeration and quantum Monte Carlo study of the two-dimensional Falicov-Kimball model and discussed the density of states and the spectral function, including their causality and cluster size dependence. A pseudogap opens in the density of states at intermediate interactions as the temperature is lowered, a single-particle precursor of the CDW transition at lower temperature. The phase diagram converges monotonically with cluster size, with the notable exception of the 2×2 cluster. The CDW transition temperature scales linearly in the inverse linear dimension of the cluster, as expected for a system in the 2D Ising model universality class. The specific heat clearly displays the critical fluctuations associated with the phase transition, in contrast to the dynamical mean-field theory where such nonlocal fluctuations are absent.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge discussions with N. E. Bickers, P. G. J. van Dongen, J. K. Freericks, D. W. Hess, J. E. Gubernatis, M. Ma, Th. Maier, Th. Pruschke, A. Schiller, V. Sudhinda, and F.-C. Zhang. This work was supported by NSF Grant Nos. DMR-9704021 and DMR-9357199, the U.S. Department of Energy Contract No. W-31-109-ENG-38, the Ohio Board of Regents Research Challenge Award (H.R.K.), and the University Grants Commission, India (H.R.K.). Computer support was provided by the Ohio Supercomputer Center.

APPENDIX A: TWO-PARTICLE PROPERTIES

Here we discuss the calculation of the lattice two-particle properties, such as spin and charge susceptibilities, in terms of the two-particle quantities on the cluster. This is a subtle issue which requires some formal discussion of what quantities from the cluster and lattice should and should not be employed. We will show using the ‘‘Baym-Kadanoff’’ formalism that there is a unique construction for which the susceptibilities correspond to the second derivatives of the corresponding extremal free energy with respect to external fields. This optimal choice corresponds to employing only the irreducible quantities from the cluster when constructing these susceptibilities.

1. Lattice quantities and matrix notation

As discussed in standard texts on quantum many-body theory, the charge and spin susceptibilities at wave vectors \mathbf{q} and frequency $i\nu$ can be calculated from the two-particle Green functions χ as

$$\begin{pmatrix} \tilde{\chi}_{ch}(\mathbf{q}, i\nu) \\ \tilde{\chi}_{sp}(\mathbf{q}, i\nu) \end{pmatrix} = \frac{(k_B T)^2}{N^2} \sum_{\mathbf{kk}'nn', \sigma\sigma'} \chi_{\mathbf{q}, i\nu, \sigma\sigma'}(\mathbf{k}i\omega_n; \mathbf{k}'i\omega_n') \times \begin{pmatrix} 1 \\ \sigma\sigma' \end{pmatrix}, \quad (\text{A1})$$

where χ is the appropriate Matsubara frequency Fourier component of $\langle T_\tau c_{\mathbf{k}+\mathbf{q}\sigma}^\dagger(\tau) c_{\mathbf{k}\sigma}(\tau') c_{\mathbf{k}'-\mathbf{q}\sigma'}^\dagger(\tau'') c_{\mathbf{k}'\sigma'}(\tau''') \rangle$. In diagrammatic perturbation theory, χ gets related to the one-particle irreducible vertex function $\mathbf{T}^{(2)}$ or the particle-hole irreducible vertex function $\mathbf{\Gamma}$ in the standard way as

$$\chi_{\mathbf{q}, i\nu} = \chi_{\mathbf{q}, i\nu}^0 + \chi_{\mathbf{q}, i\nu}^0 \mathbf{T}_{\mathbf{q}, i\nu}^{(2)} \chi_{\mathbf{q}, i\nu}^0 \quad (\text{A2})$$

$$= \chi_{\mathbf{q}, i\nu}^0 + \chi_{\mathbf{q}, i\nu}^0 \mathbf{\Gamma}_{\mathbf{q}, i\nu} \chi_{\mathbf{q}, i\nu}^0. \quad (\text{A3})$$

Here, a matrix notation, regarding $\chi_{\mathbf{q}, i\nu}$, $\mathbf{T}_{\mathbf{q}, i\nu}^{(2)}$, and $\mathbf{\Gamma}_{\mathbf{q}, i\nu}$ as matrices with row and column indices labeled by $(\mathbf{k}i\omega_n\sigma)$ and $(\mathbf{k}'i\omega_n'\sigma')$, respectively, has been used to compactify the equations. $(\mathbf{q}i\nu)$ constitute passive, parametric labels for these matrices. The bare two-particle Green function $\chi_{\mathbf{q}, i\nu}^0$ is the diagonal matrix given by

$$\chi_{\mathbf{q}, i\nu, \sigma\sigma'}^0(\mathbf{k}i\omega_n; \mathbf{k}'i\omega_n') = N \delta_{\sigma\sigma'} \delta_{nn'} \delta_{\mathbf{kk}'} G_\sigma(\mathbf{k}, i\omega_n) G_\sigma(\mathbf{k}+\mathbf{q}, i\omega_n+i\nu). \quad (\text{A4})$$

From the above it follows that

$$[\chi_{\mathbf{q}, i\nu}]^{-1} = [\chi_{\mathbf{q}, i\nu}^0]^{-1} - \mathbf{\Gamma}_{\mathbf{q}, i\nu}, \quad (\text{A5})$$

$$[\mathbf{T}_{\mathbf{q}, i\nu}^{(2)}]^{-1} = [\mathbf{\Gamma}_{\mathbf{q}, i\nu}]^{-1} - \chi_{\mathbf{q}, i\nu}^0. \quad (\text{A6})$$

For completeness, these equations may be diagonalized in the spin label to yield the more familiar forms

$$[\chi_{\alpha, \mathbf{q}, i\nu}]^{-1} = [\chi_{\mathbf{q}, i\nu}^0]^{-1} - \mathbf{\Gamma}_{\alpha, \mathbf{q}, i\nu}, \quad (\text{A7})$$

$$[\mathbf{T}_{\alpha, \mathbf{q}, i\nu}^{(2)}]^{-1} = [\mathbf{\Gamma}_{\alpha, \mathbf{q}, i\nu}]^{-1} - \chi_{\mathbf{q}, i\nu}^0, \quad (\text{A8})$$

where α denotes either the spin or charge channel (sp or ch), and $\Gamma_{sp} = \Gamma_{\sigma, -\sigma} - \Gamma_{\sigma, \sigma}$ and $\Gamma_{ch} = \Gamma_{\sigma, -\sigma} + \Gamma_{\sigma, \sigma}$.

2. Cluster quantities

On the cluster, the two-particle Green functions and vertex functions are calculated at the cluster momenta $\mathbf{Q}, \mathbf{K}, \mathbf{K}'$; which we denote by $\chi_{\mathbf{Q}, i\nu}^c, \chi_{\mathbf{Q}, i\nu}^{0c}, \mathbf{T}_{\mathbf{Q}, i\nu}^{(2)c}$, and $\mathbf{\Gamma}_{\mathbf{Q}, i\nu}^c$, where now the matrix labels correspond to $(\mathbf{K}, i\omega_n, \sigma)$ and $(\mathbf{K}', i\omega_n', \sigma')$ (momenta confined to the cluster momenta). These are then related to each other by the same equations as Eqs. (A5) and (A6), except that the lattice momenta \mathbf{q} are replaced by the cluster momenta \mathbf{Q} . In a diagrammatic perturbation theory treatment of the cluster problem, $\mathbf{\Gamma}_{\mathbf{Q}, i\nu}^c$ is calculated approximately as a function of the cluster propagators. In other treatments of the cluster, such as QMC, one calculates $\chi_{\mathbf{Q}, i\nu}^{0c}$ and $\chi_{\mathbf{Q}, i\nu}^c$ and infers $\mathbf{\Gamma}_{\mathbf{Q}, i\nu}^c$ by using the analog of Eq. (A5) in reverse as

$$\mathbf{\Gamma}_{\mathbf{Q}, i\nu}^c = [\chi_{\mathbf{Q}, i\nu}^{0c}]^{-1} - [\chi_{\mathbf{Q}, i\nu}^c]^{-1}, \quad (\text{A9})$$

and then $\mathbf{T}_{\mathbf{Q}, i\nu}^{(2)c}$ using the analog of Eq. (A6). Both lattice and cluster quantities are now uniquely defined.

3. Coarse-grained quantities

We now define coarse-grained two-particle Green function $\bar{\chi}$, the equivalent of \bar{G} for the single-particle Green function. For this purpose, we write $\mathbf{q} = \mathbf{Q} + \tilde{\mathbf{q}}$, $\mathbf{k} = \mathbf{K} + \tilde{\mathbf{k}}$, $\mathbf{k}' = \mathbf{K}' + \tilde{\mathbf{k}'}$, etc., where $\mathbf{Q}, \mathbf{K}, \mathbf{K}'$ are cluster momenta and $\tilde{\mathbf{q}}, \tilde{\mathbf{k}}, \tilde{\mathbf{k}'}$ are inside the corresponding momentum cells. $\bar{\chi}$ is then given by

$$\begin{aligned} \bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}}, i\nu} &\equiv \bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}}, i\nu, \sigma\sigma'}(\mathbf{K}, i\omega_n; \mathbf{K}', i\omega_n') \\ &= \frac{N_c^2}{N^2} \sum_{\mathbf{kk}'} \chi_{\mathbf{Q}+\tilde{\mathbf{q}}, i\nu, \sigma\sigma'}(\mathbf{K}+\tilde{\mathbf{k}}, i\omega_n; \mathbf{K}'+\tilde{\mathbf{k}'}, i\omega_n'), \end{aligned} \quad (\text{A10})$$

where the first equation again shows the matrix notation. Similarly $\bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}}, i\nu}^0$ is the diagonal matrix with entries given by

$$\begin{aligned} \bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}}, i\nu, \sigma\sigma'}^0(\mathbf{K}, i\omega_n; \mathbf{K}', i\omega_n') \\ = N_c \delta_{\sigma\sigma'} \delta_{\mathbf{KK}'} \delta_{nn'} \left[\frac{N_c}{N} \sum_{\mathbf{k}} G_\sigma(\mathbf{K}+\tilde{\mathbf{k}}, i\omega_n) \right. \\ \left. \times G_\sigma(\mathbf{K}+\tilde{\mathbf{k}}+\mathbf{Q}+\tilde{\mathbf{q}}, i\omega_n+i\nu) \right]. \end{aligned} \quad (\text{A11})$$

For the purposes of calculating $\tilde{\chi}_{ch}(\mathbf{Q}+\tilde{\mathbf{q}}, i\nu)$ and $\tilde{\chi}_{sp}(\mathbf{Q}+\tilde{\mathbf{q}}, i\nu)$, it is enough to compute $\bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}}, i\nu}$, since

$$\begin{pmatrix} \tilde{\chi}_{ch}(\mathbf{Q}+\tilde{\mathbf{q}}, i\nu) \\ \tilde{\chi}_{sp}(\mathbf{Q}+\tilde{\mathbf{q}}, i\nu) \end{pmatrix} = \frac{(k_B T)^2}{N_c^2} \sum_{\mathbf{KK}'nn', \sigma\sigma'} \bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}}, i\nu, \sigma\sigma'} \times (\mathbf{K}i\omega_n; \mathbf{K}'i\omega_n') \begin{pmatrix} 1 \\ \sigma\sigma' \end{pmatrix}. \quad (\text{A12})$$

For the single-particle Green function we had $\bar{\mathbf{G}} = \mathbf{G}^c$, since in that case the coarse graining is done with the *external* momentum. For the two-particle case, the above-defined coarse-grained quantities are *not* identical with $\chi_{\mathbf{Q}, i\nu}^c$ and $\chi_{\mathbf{Q}, i\nu}^{0c}$. The coarse-grained quantities are defined for all external lattice momenta \mathbf{q} , not just the cluster momenta \mathbf{Q} . However, the matrix size is determined by the number of cluster momenta rather than the (infinite) number of lattice momenta. As we will see below, this is a significant numerical simplification, since the calculation of the susceptibilities can be reduced to the solution of a set of linear equations defined on the cluster momenta instead of the momenta of the infinite lattice.

4. Two prescriptions

Two different prescriptions for computing $\bar{\chi}$ out of cluster quantities suggest themselves (a third possibility, approximating $\bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}}, i\nu}$ by $\chi_{\mathbf{Q}, i\nu}^c$, is obviously too crude to be discussed further). The first one corresponds to replacing $\mathbf{T}_{\mathbf{Q}+\tilde{\mathbf{q}}, i\nu, \sigma\sigma'}^{(2)c}(\mathbf{K}+\tilde{\mathbf{k}}, i\omega_n; \mathbf{K}'+\tilde{\mathbf{k}'}, i\omega_n')$ by $\mathbf{T}_{\mathbf{Q}, i\nu, \sigma\sigma'}^{(2)c}(\mathbf{K}, i\omega_n; \mathbf{K}', i\omega_n')$ in the expression for $\bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}}, i\nu}$ derived from Eq. (A2). We then get the equation

$$\bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}},i\nu} \cong \bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}},i\nu}^0 + \bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}},i\nu}^0 \mathbf{T}_{\mathbf{Q},i\nu}^{(2)c} \bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}},i\nu}^0. \quad (\text{A13})$$

This means we have identified the *reducible* two-particle vertex $\mathbf{T}^{(2)}$ of the cluster and the lattice at the cluster momenta.

The second prescription, that we argue below is the *correct* prescription, is to replace $\Gamma_{\mathbf{Q}+\tilde{\mathbf{q}},i\nu,\sigma\sigma'}^{(2)}(\mathbf{K}+\tilde{\mathbf{k}},i\omega_n;\mathbf{K}'+\tilde{\mathbf{k}}',i\omega_n')$ by $\Gamma_{\mathbf{Q},i\nu,\sigma\sigma'}^{(2)c}(\mathbf{K},i\omega_n;\mathbf{K}',i\omega_n')$ in the integral equation for $\bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}},i\nu}$ derived from Eq. (A3). This leads to the equation

$$\bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}},i\nu} \cong \bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}},i\nu}^0 + \bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}},i\nu}^0 \mathbf{\Gamma}_{\mathbf{Q},i\nu}^c \bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}},i\nu}, \quad (\text{A14})$$

whence

$$\bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}},i\nu} = ([\bar{\chi}_{\mathbf{Q}+\tilde{\mathbf{q}},i\nu}^0]^{-1} - \mathbf{\Gamma}_{\mathbf{Q},i\nu}^c)^{-1}. \quad (\text{A15})$$

Here, we have identified the *irreducible* two-particle vertex $\mathbf{\Gamma}$ of the cluster and the lattice at the cluster momenta. Either Eqs. (A13) or (A15) can then be used in Eq. (A12) to compute $\tilde{\chi}_{ch}$ and $\tilde{\chi}_{sp}$. At this stage it is not clear which prescription is better or whether both could be feasible approximations. We will now show that internal consistency and Φ derivability in the Baym-Kadanoff sense do single out the second prescription, Eq. (A15).

5. Relation to Φ derivability

The Baym-Kadanoff¹¹ Φ functional is diagrammatically defined as

$$\Phi(\mathbf{G}) = \sum_l p_l \text{tr}[\mathbf{\Sigma}_\sigma^l \mathbf{G}_\sigma]. \quad (\text{A16})$$

The trace indicates summation over frequency, momentum, and spin. Here, $\mathbf{\Sigma}_\sigma^l$ is the set of irreducible self-energy diagrams of l th order in the interaction, \mathbf{G}_σ is the dressed Green function related to $\mathbf{\Sigma}_\sigma$ and the bare lattice Green function \mathbf{G}_σ^0 via the Dyson equation $\mathbf{G}_\sigma^{-1} = \mathbf{G}_\sigma^{0-1} - \mathbf{\Sigma}_\sigma$, and p_l is a counting factor equal to the number of occurrences of \mathbf{G}_σ in each term (for Hubbard-like models, $p_l = 1/l$). The free energy \mathcal{J} can be expressed in terms of the ‘‘linked cluster expansion’’ W as $\mathcal{J} = -k_B T W$ with

$$W = \Phi(\mathbf{G}) - \text{tr}[\mathbf{\Sigma}_\sigma \mathbf{G}_\sigma] - \text{tr} \ln[-\mathbf{G}_\sigma]. \quad (\text{A17})$$

With the above definitions it holds that $\mathbf{\Sigma}_\sigma = \delta\Phi/\delta\mathbf{G}_\sigma$, as required for a ‘‘ Φ -derivable’’ theory, and the free energy is stationary under variations of \mathbf{G} . In addition, the irreducible vertex function is obtained by a second variation of Φ , $\mathbf{\Gamma}_{\sigma,\sigma'} = \delta^2\Phi/(\delta\mathbf{G}_\sigma \delta\mathbf{G}_{\sigma'}) = \delta\mathbf{\Sigma}_\sigma/\delta\mathbf{G}_{\sigma'}$.

The DCA can be microscopically motivated by our choice of the Laue function Δ_{DCA} in Eq. (9). The effect of the chosen Laue function is the replacement of the $\mathbf{\Sigma}_\sigma$ and $\mathbf{\Gamma}_{\sigma,\sigma'}$ by the corresponding coarse-grained quantities (indicated by the bars). For example, consider the relation $\mathbf{\Sigma} = T^{(2)}\mathbf{G}$ (order by order in the diagrammatic series). The vertices connecting the Green function to $T^{(2)}$ do not preserve momentum within the cells about the cluster momentum due to the DCA Laue function. Consequently, the lattice Green function \mathbf{G}_σ is replaced by the coarse-grained Green function $\bar{\mathbf{G}}_\sigma$. The external momentum label (\mathbf{k}) of the self-energy is

in principle still a lattice momentum; however, the self-energy will only depend through the function $\mathbf{M}(\mathbf{k})$ on \mathbf{k} . If we use this self-energy in, e.g., the calculation of its contribution to the Φ functional, the Laue function on the vertices will ‘‘reduce’’ both the self-energy as well as the diagram closing Green function to their corresponding coarse-grained expressions. Consequently, the DCA Φ functional reads

$$\Phi_{DCA}(\mathbf{G}) = \sum_l p_l \text{tr}[\bar{\mathbf{\Sigma}}_\sigma^l \bar{\mathbf{G}}_\sigma]. \quad (\text{A18})$$

In correspondence to the lattice system,

$$\frac{\delta\Phi_{DCA}}{\delta\bar{\mathbf{G}}_\sigma} = \bar{\mathbf{\Sigma}} = \frac{\delta\Phi_{DCA}}{\delta\mathbf{G}_\sigma}, \quad (\text{A19})$$

where the second equality follows since the variation $\delta/\delta\mathbf{G}_\sigma$ corresponds to cutting a Green-function line, so that $\delta\bar{\mathbf{G}}_{\sigma\mathbf{K}}/\delta\mathbf{G}_{\sigma'\mathbf{K}'} = \delta_{\mathbf{K},\mathbf{M}(\mathbf{k}')}\delta_{\sigma,\sigma'}$. It follows that the DCA estimate of the lattice free energy is $\mathcal{J}_{DCA} = -k_B T W_{DCA}$, where

$$W_{DCA} = \Phi_{DCA} - \text{tr}[\bar{\mathbf{\Sigma}}_\sigma \bar{\mathbf{G}}_\sigma] - \text{tr} \ln[-\bar{\mathbf{G}}_\sigma]. \quad (\text{A20})$$

Now W_{DCA} is stationary with respect to $\bar{\mathbf{G}}_\sigma$,

$$\delta\mathcal{J}_{DCA}/\delta\bar{\mathbf{G}}_\sigma = -\bar{\mathbf{\Sigma}}_\sigma + \mathbf{\Sigma}_\sigma = 0, \quad (\text{A21})$$

which means that $\bar{\mathbf{\Sigma}}_\sigma$ is the proper approximation for the lattice self-energy corresponding to Φ_{DCA} .

The susceptibilities are thermodynamically defined as second derivatives of the free energy with respect to external fields. Now $\Phi_{DCA}(\mathbf{G})$ and $\bar{\mathbf{\Sigma}}_\sigma$, and hence \mathcal{J}_{DCA} depend on these fields only through $\bar{\mathbf{G}}_\sigma$ and \mathbf{G}_σ^0 . Following Baym¹¹ it is easy to verify that, the prescription (A12)+(A15), with

$$\mathbf{\Gamma}_{\sigma,\sigma'} \approx \bar{\mathbf{\Gamma}}_{\sigma,\sigma'} \equiv \delta\bar{\mathbf{\Sigma}}_\sigma/\delta\mathbf{G}_{\sigma'}, \quad (\text{A22})$$

yields the same estimate that would be obtained from the second derivative of W_{DCA} with respect to the applied field. For example, the first derivative of the partition function W_{DCA} with respect to a spatially homogeneous external magnetic field h is the magnetization,

$$m = \text{tr}[\sigma \mathbf{G}_\sigma]. \quad (\text{A23})$$

The susceptibility is given by the second derivative,

$$\frac{\partial m}{\partial h} = \text{tr} \left[\sigma \frac{\partial \mathbf{G}_\sigma}{\partial h} \right]. \quad (\text{A24})$$

We substitute $\mathbf{G}_\sigma = (\mathbf{G}_\sigma^{0-1} - \bar{\mathbf{\Sigma}}_\sigma)^{-1}$, and evaluate the derivative,

$$\frac{\partial m}{\partial h} = \text{tr} \left[\sigma \frac{\partial \mathbf{G}_\sigma}{\partial h} \right] = \text{tr} \left[\mathbf{G}_\sigma^2 \left(1 + \sigma \frac{\partial \bar{\mathbf{\Sigma}}_\sigma}{\partial \mathbf{G}_{\sigma'}} \frac{\partial \mathbf{G}_{\sigma'}}{\partial h} \right) \right], \quad (\text{A25})$$

where $\partial m/\partial h = \tilde{\chi}_{sp}(\mathbf{q}=0, i\nu=0)$. If we identify $\chi_{\sigma,\sigma'} = \sigma(\partial\mathbf{G}_\sigma/\partial h)$ and $\chi_\sigma^0 = \mathbf{G}_\sigma^2$, collect all of the terms within both traces, and sum over the cell momenta $\tilde{\mathbf{k}}$, we obtain the two-particle Dyson’s equation

$$2(\bar{\chi}_{\sigma,\sigma} - \bar{\chi}_{\sigma,-\sigma}) = 2\bar{\chi}_{\sigma}^0 + 2\bar{\chi}_{\sigma}^0(\bar{\Gamma}_{\sigma,\sigma} - \bar{\Gamma}_{\sigma,-\sigma})(\bar{\chi}_{\sigma,\sigma} - \bar{\chi}_{\sigma,-\sigma}) \quad (\text{A26})$$

which is equivalent to Eq. (A15). We see that indeed it is the irreducible quantity, i.e., the vertex function, for which cluster and lattice correspond.

In summary, the choice of the Laue function and the requirement of a Φ -derivable theory ultimately determine the way lattice properties are constructed out of cluster properties. The usefulness of the DCA lies in the fact that both the single- and the two-particle irreducible properties ($\bar{\Sigma}$ and $\bar{\Gamma}$) can be determined from the cluster problem, i.e., $\bar{\Sigma} = \Sigma^c$ and $\bar{\Gamma} = \Gamma^c$. Note that although this construction is unique and Φ derivable, because of the partial violation of momentum conservation at each internal vertex described by Δ_{DCA} certain Ward identities will be violated in any dimension, even for the single-site cluster (DMFA) appropriate in $D = \infty$. This will be discussed in Appendix D.

APPENDIX B: DCA FOR PROBLEMS WITH EXTENDED RANGE OR ELECTRON-PHONON INTERACTIONS

In this appendix we present an extension of the DCA to problems with extended range interactions, such as in the extended Hubbard model.

Consider the partition function for such a model written in terms of Fermionic functional integrals:

$$\begin{aligned} \mathcal{Z} = & \int_{cc^\dagger} \exp - \int_0^\beta d\tau \left[\sum_{ij} c_i^\dagger(\tau) \{ (\partial_\tau - \mu) \delta_{ij} - t_{ij} \} c_j(\tau) \right. \\ & \left. + U \sum_i \hat{n}_{i\uparrow}(\tau) \hat{n}_{i\downarrow}(\tau) + \frac{1}{2} \sum_{i \neq j} \sum_{\sigma\sigma'} V_{ij} \hat{n}_{i\sigma}(\tau) \hat{n}_{j\sigma'}(\tau) \right]. \end{aligned} \quad (\text{B1})$$

By introducing a real, continuous Hubbard-Stratonovich field $\phi_i(\tau)$ which couples to the local charge density $\hat{n}_i \equiv \sum_\sigma \hat{n}_{i\sigma}$, we can write

$$\begin{aligned} \mathcal{Z} = & \int_{cc^\dagger} \int_\phi \exp - \int_0^\beta d\tau \left[\sum_{ij} c_i^\dagger(\tau) \{ (\partial_\tau - \tilde{\mu}) \delta_{ij} - t_{ij} \} c_j(\tau) \right. \\ & + \tilde{U} \sum_i n_{i\uparrow}(\tau) n_{i\downarrow}(\tau) + \frac{\tilde{V}_o^2}{2} \sum_{ij} \phi_i(\tau) (\tilde{V})^{-1}_{ij} \phi_j(\tau) \\ & \left. + \tilde{V}_o \sum_i \phi_i(\tau) \hat{n}_i(\tau) \right]. \end{aligned} \quad (\text{B2})$$

Here, $\tilde{V}_{ij} = \tilde{V}_o \delta_{ij} - V_{ij}$ with \tilde{V}_o so chosen as to make \tilde{V} positive definite (and hence invertible), $\tilde{U} = (U + \tilde{V}_o)$, and $\tilde{\mu} = \mu - \frac{1}{2} \tilde{V}_o$. For example, for the extended Hubbard model with nearest-neighbor interaction of strength V , $\tilde{V}_o = zV$, where z is the coordination number of the lattice.

Now it is straightforward to devise the DCA for this coupled Fermion-boson problem. The cluster problem we need to solve corresponds to the functional integral given by

$$\begin{aligned} \mathcal{Z}_c = & \int_{cc^\dagger} \int_\phi \exp - \left[\int_0^\beta d\tau \int_0^\beta d\tau' \sum_{ij} \{ c_i^\dagger(\tau) \mathcal{G}_{ij}^{-1}(\tau - \tau') \right. \\ & \times c_j(\tau') + \phi_i(\tau) \mathcal{D}_{ij}^{-1}(\tau - \tau') \phi_j(\tau') \} \\ & \left. + \int_0^\beta d\tau \sum_i \{ \tilde{U} \hat{n}_{i\uparrow}(\tau) \hat{n}_{i\downarrow}(\tau) + \tilde{V}_o \phi_i(\tau) \hat{n}_i(\tau) \} \right]. \end{aligned} \quad (\text{B3})$$

The cluster problem is to be treated by some technique to obtain the cluster propagators and self energies: $G^c(\mathbf{K})$, $\Sigma^c(\mathbf{K})$ for the electrons and $D^c(\mathbf{Q})$, $\Pi^c(\mathbf{Q})$ for the field ϕ , at cluster momenta \mathbf{K} and \mathbf{Q} . One has the Dyson equations

$$[G^c(\mathbf{K})]^{-1} = \mathcal{G}^{-1}(\mathbf{K}) - \Sigma^c(\mathbf{K}), \quad (\text{B4})$$

$$[D^c(\mathbf{Q})]^{-1} = \mathcal{D}^{-1}(\mathbf{Q}) - \Pi^c(\mathbf{Q}), \quad (\text{B5})$$

where the frequency arguments have been suppressed for convenience.

The self-consistent embedding of the above cluster in the effective medium defined by the rest of the sites of the original lattice is obtained by assuming that $\Sigma^c(\mathbf{K})$, and $\Pi^c(\mathbf{Q})$ represent good approximations to the (coarse-grained averages of the) lattice self-energies, and that $G^c(\mathbf{K})$ and $D^c(\mathbf{Q})$ must equal the coarse-grained averages of the corresponding lattice Green functions

$$G^c(\mathbf{K}) = \bar{G}(\mathbf{K}) \equiv \sum_{\mathbf{k}} \frac{1}{i\omega_n + \tilde{\mu} - \epsilon_{\mathbf{k}+\mathbf{K}} - \Sigma^c(\mathbf{K})}, \quad (\text{B6})$$

$$D^c(\mathbf{Q}) = \bar{D}(\mathbf{Q}) \equiv \sum_{\mathbf{q}} \frac{1}{\tilde{V}_{\mathbf{Q}+\mathbf{q}}^{-1} - \Pi^c(\mathbf{Q})}. \quad (\text{B7})$$

Thus, the self-consistency loop is closed by recalculating $\mathcal{G}_{\mathbf{K}}^{-1}$ and $\mathcal{D}_{\mathbf{Q}}^{-1}$ using the Dyson equations backwards as

$$\mathcal{G}^{-1}(\mathbf{K}) = \bar{G}^{-1}(\mathbf{K}) + \Sigma^c(\mathbf{K}), \quad (\text{B8})$$

$$\mathcal{D}^{-1}(\mathbf{Q}) = \bar{D}^{-1}(\mathbf{Q}) + \Pi^c(\mathbf{Q}). \quad (\text{B9})$$

We note that for the one-site cluster, the resulting DMFA does not correspond to the approximation resulting from scaling V as V^*/d (whence in the $D \rightarrow \infty$ limit only the Hartree contribution to Σ survives), but is a rather different approximation which includes local dynamical charge fluctuations and local screening effects.⁸ It is formally similar to the problem obtained in the DMFA of the Holstein-Hubbard model. Correspondingly, the DCA for this latter model can be formulated analogously to the above.

APPENDIX C: PROOF OF CAUSALITY

In this appendix we prove that the DCA formally preserves the condition of positive semidefiniteness of the single-particle spectral functions. The proof requires that the cluster problem is solved by methods that preserve causality (exact enumeration, QMC, etc.). For simplicity of notation the proof is explicitly given for Hubbard-like models, but it can be easily generalized to the PAM, multiband models and models with nonlocal interactions.

Most steps of the DCA algorithm are easily seen to preserve the causality property. We assume a causal \mathcal{G} , so that $-\text{Im } \mathcal{G} > 0$, as a starting point of the iteration. If the method to solve the cluster problem preserves causality the resulting cluster Green function G^c will also be causal. With Dyson's equation we obtain a causal cluster self-energy. This self-energy is also assumed to be the lattice self-energy of the infinite lattice at the cluster momenta. Therefore, the lattice Green function [the summand of Eq. (2)] is also causal. As the coarse-grained Green function \bar{G} is obtained by an average of causal Green functions it must be causal, too.

The nontrivial step is to show that Eq. (3) does not lead to an acausal \mathcal{G} for the next iteration. The spectral function of \mathcal{G} will be positive semidefinite if

$$\text{Im}[\bar{G}(\mathbf{K}, \omega)^{-1}] \geq -\text{Im } \Sigma^c(\mathbf{K}, \omega). \quad (\text{C1})$$

We write $\bar{G}(\mathbf{K}, \omega)$ as $\bar{G}(\mathbf{K}, \omega) = (N_c/N) \sum_{\tilde{\mathbf{k}}} [z_{\mathbf{K}+\tilde{\mathbf{k}}}(\omega)]^{-1}$ with $z_{\mathbf{K}+\tilde{\mathbf{k}}}(\omega) = x_{\tilde{\mathbf{k}}}(\mathbf{K}, \omega) + ia(\mathbf{K}, \omega)$. Now $z_{\mathbf{K}+\tilde{\mathbf{k}}}(\omega)$ is the inverse of our estimate of the Green function of the infinite lattice with a real part $x_{\tilde{\mathbf{k}}}(\mathbf{K}, \omega) = \omega - \epsilon_{\mathbf{K}+\tilde{\mathbf{k}}} - \text{Re } \Sigma^c(\mathbf{K}, \omega)$ and an imaginary part $a(\mathbf{K}, \omega) = -\text{Im } \Sigma^c(\mathbf{K}, \omega)$, with $a(\mathbf{K}, \omega)$ a positive semidefinite function of \mathbf{K} and ω but independent of $\tilde{\mathbf{k}}$. Graphically, the proof of Eq. (C1) is illustrated in Fig. 5.

We now proceed to show the validity of Eq. (C1) in a rigorous fashion. To simplify notation we will suppress the common indices \mathbf{K} and ω . We also specify to the retarded Green functions with $\omega \rightarrow \omega + i\eta$ with positive infinitesimal η . The sum over $\tilde{\mathbf{k}}$ in the definition of \bar{G} runs over $n = N/N_c$ terms. Each term is a complex number with a positive definite imaginary part a that is *independent* of the summation index. Equation (C1) is now cast into the following proposition.

Proposition. For $j=1, \dots, n$, let $z_j \in \mathbf{C}$, where \mathbf{C} is the set of complex numbers, and $\text{Im}(z_j) = a > 0$. If

$$\bar{G} := \frac{1}{n} \sum_{j=1}^n \frac{1}{z_j} \text{ then } \text{Im}(\bar{G}^{-1}) \geq a,$$

with equality if and only if $z_1 = \dots = z_j = \dots = z_n$.

Proof. If $w = u + iv = 1/z$ with $z = x + iy$, then the line $\text{Im } z = a$ in the extended z plane, given by

$$\text{Im}(z) = y = a = \frac{-v}{u^2 + v^2},$$

is mapped in a one-to-one fashion onto the circle

$$u^2 + \left(v + \frac{1}{2a}\right)^2 = \left(\frac{1}{2a}\right)^2$$

in the extended w plane, with center $-i/2a$ and a radius of $r = 1/2a$. It follows that \bar{G} lies on or inside this circle,

$$\left| \bar{G} - \left(\frac{-i}{2a}\right) \right| = \frac{1}{n} \left| \sum_{j=1}^n \left(\frac{1}{z_j} + \frac{i}{2a}\right) \right| \leq \frac{1}{n} \sum_{j=1}^n \left| \frac{1}{z_j} + \frac{i}{2a} \right| = \frac{1}{2a}, \quad (\text{C2})$$

where we have used the triangular inequality. The bijective function $z = 1/w$ maps a point w strictly inside the circle to a point z with $\text{Im}(z) > a$ (and conversely)

$$\text{Im } z = \frac{-v}{u^2 + v^2} > a$$

if and only if

$$u^2 + \left(v + \frac{1}{2a}\right)^2 < \left(\frac{1}{2a}\right)^2.$$

Hence, $\text{Im}(\bar{G}^{-1}) \geq a$, where equality holds if and only if $z_1 = \dots = z_j = \dots = z_n$.

Because of the infinitesimal η we had $a > 0$ for the above proof. However, if $\text{Im } \Sigma^c(\mathbf{K}, \omega) = 0$, the resulting imaginary part of \mathcal{G} is proportional to $-\eta$. This is the case, e.g., for frequencies larger than the bandwidth. Hence, the bandwidth of \mathcal{G} is identical to the bandwidth of \bar{G} and G^c , i.e., there is no band broadening induced by the coarse-graining procedure.

Generalization to multiband models such as the PAM is straightforward. Without going into the details of the model we note that there are two species of fermions which are coupled by on-site hybridization. The d electrons are itinerant and noninteracting, whereas the f -electrons are localized (no bare hopping) and have a Hubbard interaction. The f -electron Green function has two self-energies, from the Hubbard interaction and the hybridization, respectively. Both self-energies are causal (negative semidefinite and decaying like $1/\omega$). In contrast to the Hubbard self-energy the self-energy due to the hybridization is known explicitly and does depend on all the lattice momenta, therefore also on the $\tilde{\mathbf{k}}$ momenta in the cells about the cluster momenta. For a given \mathbf{K} and ω the imaginary part of this self-energy is bounded from above by some value $-b_{\min}(\mathbf{K}, \omega)$. Consequently, we can prove in analogous fashion that

$$\text{Im}[\bar{G}_f(\mathbf{K}, \omega)^{-1}] \geq a(\mathbf{K}, \omega) + b_{\min}(\mathbf{K}, \omega),$$

where $-a(\mathbf{K}, \omega)$ is the self-energy due to the Hubbard interaction of the f electrons.

A last remark on the possibility of self-energy interpolation is in order here. At first glance one might try to improve the calculation by employing an interpolation of the cluster self-energy between the cluster momenta in the coarse-graining step, Eq. (2), rather than using the ‘‘rectangular’’ approximation for the lattice self-energy $\Sigma(\mathbf{K} + \tilde{\mathbf{k}}, \omega) \approx \Sigma^c(\mathbf{K}, \omega)$. However, as one can easily convince oneself given the above proof, *any* interpolation scheme will violate causality if $\text{Im } \Sigma^c(\mathbf{K}, \omega)$ has a minimum somewhere in the BZ. This will generally be so except in the case of the single site cluster, in which there is nothing to interpolate. This further limits the freedom of the coarse-graining procedure.

APPENDIX D: CONSERVATION OF THE DMFA AND DCA

An approximation which satisfies the various Ward identities is identified as a ‘‘conserving approximation’’ since the Ward identities are derived from conservation laws. Baym

and Kadanoff^{30,11} showed that a sufficient condition to guarantee that an approximation is conserving is for it to be Φ derivable and self-consistent. Energy, particle number, and momentum are also assumed to be conserved at each internal vertex, which may be assured by properly constructing the diagrams from the lattice propagator $G_{\mathbf{k}}$ using well-known Feynman rules. Specifically, the functional $\Phi(G(\mathbf{k}, \omega), U)$ is a set of closed graphs formed from the lattice propagators $G(\mathbf{k}, \omega)$ and interactions U . The one- and two-particle self-energies are calculated from functional derivatives of $\Phi(G(\mathbf{k}, \omega), U)$, $\Sigma(\mathbf{k}, \omega) = \delta\Phi/\delta G(\mathbf{k}, \omega)$, $\Gamma_{\sigma, \sigma'} = \delta^2\Phi/\delta G_{\sigma} \delta G_{\sigma'}$. The equation $\Sigma(\mathbf{k}, \omega) = \delta\Phi/\delta G(\mathbf{k}, \omega)$ must be solved self-consistently until $G(\mathbf{k}, \omega)$ converges. As an additional consequence, Baym showed that quantities calculated within such an approximation were unique.

In the infinite-dimensional formalism of Metzner and Vollhardt momentum conservation is violated at internal vertices. Consequently, Φ is a functional of the local propagator $G_{ii}(\omega)$ rather than the lattice propagator $G(\mathbf{k}, \omega)$, and the corresponding self-energies are obtained from functional derivatives of $\Phi(G_{ii}(\omega), U)$ and are therefore also local. However, we may also expect violations of some conservation laws. If a proper $\Phi(G(\mathbf{k}, \omega), U)$ is taken, all non-local diagrams which are higher order in $1/D$ vanish, so that $\Phi(G(\mathbf{k}, \omega), U) = \Phi(G_{ii}(\omega), U) + \mathcal{O}(1/D)$. Each functional derivative with respect to the Green function breaks an internal line and so reduces the order of the approximation by \sqrt{D} .¹⁵ It follows then that the self-energy is also local $\delta\Phi(G(\mathbf{k}, \omega), U)/\delta G(\mathbf{k}, \omega) = \Sigma(G(\mathbf{k}, \omega), U) = \Sigma(G_{ii}(\omega), U) + \mathcal{O}(1/\sqrt{D})$. However, a problem emerges at the two-particle, or higher, level since $\Gamma(G(\mathbf{k}, \omega), U) = \Gamma(G_{ii}(\omega), U) + \mathcal{O}(1)$ for any D , with the difference due to needed nonlocal corrections. Equivalently, if Φ is evaluated in the limit $D \rightarrow \infty$ before the functional derivatives are evaluated, then $\Gamma(G(\mathbf{k}, \omega), U) = \Gamma(G_{ii}(\omega), U)$; however, if the order is reversed, then corrections of order unity are required.³¹ Thus, due to the lack of momentum conservation, the DMFA does not provide a unique prescription for the calculation of two-particle properties and thus it need not be conserving.

For example, the equation of continuity, $\nabla \cdot \mathbf{J} - \partial\rho/\partial t = 0$, which describes charge conservation by electric currents, yields the original Ward¹⁰ identity

$$i\nu_{\alpha}\Lambda_0 - \mathbf{q} \cdot \mathbf{\Lambda} = \Sigma(\mathbf{k} + \mathbf{q}, i\nu_{\alpha} + i\omega_n) - \Sigma(\mathbf{k}, i\omega_n), \quad (\text{D1})$$

where Λ_0 and $\mathbf{\Lambda}$ are the scalar and vector components of the dressed vertex function such that

$$\Lambda_0(\mathbf{k}, \mathbf{q}, i\omega_n, i\nu_{\alpha}) = \frac{T}{N} \sum_{\mathbf{k}', n'} G(\mathbf{k}', i\omega'_n) G(\mathbf{k}' + \mathbf{q}, i\omega'_n + i\nu_{\alpha}) \times \mathbf{T}_{\mathbf{q}, i\nu_{\alpha}}^{(2)}(\mathbf{k}, i\omega_n; \mathbf{k}', i\omega'_n) \quad (\text{D2})$$

and

$$\mathbf{q} \cdot \mathbf{\Lambda}(\mathbf{k}, \mathbf{q}, i\omega_n, i\nu_{\alpha}) = \frac{T}{N} \sum_{\mathbf{k}', n'} (\epsilon_{\mathbf{k}'+\mathbf{q}} - \epsilon_{\mathbf{k}'}) G(\mathbf{k}', i\omega'_n) \times G(\mathbf{k}' + \mathbf{q}, i\omega'_n + i\nu_{\alpha}) \times \mathbf{T}_{\mathbf{q}, i\nu_{\alpha}}^{(2)}(\mathbf{k}, i\omega_n; \mathbf{k}', i\omega'_n). \quad (\text{D3})$$

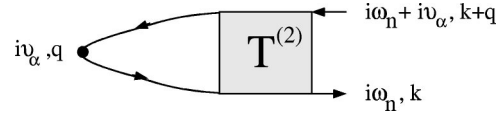


FIG. 13. Definition of $i\nu_{\alpha}\Lambda_0$ and $\mathbf{q} \cdot \mathbf{\Lambda}$. Here, each solid line is a full lattice propagator $G(k, \omega)$, the filled box is the full particle-hole reducible two-particle T -matrix, and the filled circle \bullet is $i\nu_{\alpha}$ or $\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}$ for $i\nu_{\alpha}\Lambda_0$ or $\mathbf{q} \cdot \mathbf{\Lambda}$, respectively.

Here $\mathbf{T}^{(2)}$ is the corresponding particle-hole reducible two-particle T matrix

$$\mathbf{T}_{\mathbf{q}, i\nu_{\alpha}}^{(2)} = \mathbf{\Gamma}_{\mathbf{q}, i\nu_{\alpha}}^{ph} (1 - \chi_{\mathbf{q}, i\nu_{\alpha}}^0 \mathbf{\Gamma}_{\mathbf{q}, i\nu_{\alpha}}^{ph})^{-1},$$

and $\mathbf{\Gamma}_{ph} = \mathbf{\Gamma}_{\sigma, \sigma} + \mathbf{\Gamma}_{\sigma, -\sigma}$ is the particle-hole irreducible two-particle self-energy, with $(\mathbf{k}, i\omega_n)$ and $(\mathbf{k}', i\omega'_n)$ as the matrix indices, and χ^0 is the diagonal matrix with entries $\chi_{\mathbf{q}, i\nu_{\alpha}}^0(i\omega_n, i\omega'_n) \equiv N \delta_{nn'} \delta_{\mathbf{k}\mathbf{k}'} G(\mathbf{k}, i\omega_n) G(\mathbf{k} + \mathbf{q}, i\omega_n + i\nu_{\alpha})$, and $\epsilon_{\mathbf{k}}$ the bare electronic dispersion. The corresponding diagrams are illustrated in Fig. 13.

When this formalism is applied as the DMFA in finite dimensions, the conservation of Ward identities does not follow from the arguments of Baym and Kadanoff. If we write down a proper $\Phi(G(\mathbf{k}, \omega), U)$, the only way to obtain the local generating function $\Phi(G_{ii}(\omega), U)$ used in the DMFA is to ignore momentum conservation within each graph and sum over each internal momenta independently. This clearly violates the requirement for a conserving approximation that momentum be conserved at each internal vertex,¹¹ so the conserving property of the theory is lost.

This can be seen from a direct examination of Ward's original identity, i.e., the Ward identity, Eq. (D1), is not satisfied for a general \mathbf{q} except when $i\nu_{\alpha}$ is zero. To see this, note that from Eqs. (D2) and (D3) and some simple algebra one can write

$$i\nu_{\alpha}\Lambda_0 - \mathbf{q} \cdot \mathbf{\Lambda} = \frac{T}{N} \sum_{\mathbf{k}', n'} [\{G(\mathbf{k}', i\omega'_n) - G(\mathbf{k}' + \mathbf{q}, i\omega'_n + i\nu_{\alpha})\} + \{\Sigma(\mathbf{k}' + \mathbf{q}, i\nu_{\alpha} + i\omega_n) - \Sigma(\mathbf{k}', i\omega_n)\}] \times G(\mathbf{k}', i\omega'_n) G(\mathbf{k}' + \mathbf{q}, i\omega'_n + i\nu_{\alpha}) \times \mathbf{T}_{\mathbf{q}, i\nu_{\alpha}}^{(2)}(\mathbf{k}, i\omega_n; \mathbf{k}', i\omega'_n). \quad (\text{D4})$$

Specializing now to the DMFA, the required Ward identity can be written as

$$\Sigma(i\nu_{\alpha} + i\omega_n) - \Sigma(i\omega_n) = \frac{T}{N} \sum_{j, n'} [\{G_{ii}(i\omega'_n) - G_{ii}(i\omega'_n + i\nu_{\alpha})\} \delta_{ij} + \{\Sigma(i\nu_{\alpha} + i\omega_n) - \Sigma(i\omega_n)\}] \times \exp(i\mathbf{q} \cdot \mathbf{r}_{ij}) G_{ij}(i\omega'_n) \times G_{ji}(i\omega'_n + i\nu_{\alpha}) \mathbf{T}_{\mathbf{q}, i\nu_{\alpha}}^{(2)}(i\omega_n, i\omega'_n), \quad (\text{D5})$$

where we have used the DMFA in the second step and assumed that Σ and $\mathbf{\Gamma}^{ph}$ are momentum independent, so $\mathbf{T}_{\mathbf{q}, i\nu_{\alpha}}^{(2)} = \mathbf{\Gamma}_{\mathbf{q}, i\nu_{\alpha}}^{ph} (1 - \chi_{\mathbf{q}, i\nu_{\alpha}}^0 \mathbf{\Gamma}_{\mathbf{q}, i\nu_{\alpha}}^{ph})^{-1}$ has only the momentum dependence it inherits from $\chi_{\mathbf{q}, i\nu_{\alpha}}^0$. Clearly, when $i\nu_{\alpha}$ is zero,

the right-hand side (RHS) vanishes for arbitrary \mathbf{q} and the Ward identity is satisfied. But when $i\nu_\alpha$ is nonzero, the second term on the right-hand side has a nontrivial \mathbf{q} dependence in general, and the Ward identity is violated since the left-hand side (LHS) of Eq. (D5) is \mathbf{q} independent.

Even in the $D \rightarrow \infty$ limit the Ward identity is not always satisfied. From the form of Eq. (D5) it is clear that the Ward identity is only satisfied when

$$\begin{aligned}\chi_q^0(i\omega_n, i\nu_\alpha) &\equiv (1/N) \sum_{\mathbf{k}} G(\mathbf{k}, i\omega_n) G(\mathbf{k} + \mathbf{q}, i\omega_n + i\nu_\alpha) \\ &= \chi_{ii}^0(i\omega_n, i\nu_\alpha) \equiv G_{ii}(i\omega_n) G_{ii}(i\omega_n + i\nu_\alpha).\end{aligned}$$

This is true for a generic \mathbf{q} where $X(\mathbf{q}) = (1/D) \sum_l \cos q_l = 0$.¹⁵ Then, the nonlocal parts of the second term in the RHS of Eq. (D5) can be neglected, and the Ward identity, which now involves only the local Σ , Γ and G is exactly satisfied, as can be directly shown from the effective single site problem using equations of motion. However,

there is a set of \mathbf{q} of measure zero within the Brillouin zone, which unfortunately includes the values $\mathbf{q} = 0$ and $\mathbf{q} = (\pi, \pi, \dots)$, for which $\mathbf{X}(\mathbf{q})$ is finite and $\chi_q^0(i\omega_n, i\nu_\alpha) \neq \chi_{ii}^0(i\omega_n, i\nu_\alpha)$, with corrections of order unity. For these values of \mathbf{q} the nonlocal parts in the second term can no longer be discarded, and the Ward identity is again violated. Consistent with this observation one may show to all orders in perturbation theory that nonlocal corrections to the $D = \infty$ two-particle self-energy remain finite for a set of measure zero points in the Brillouin zone. Apparently, for these points, the nonlocal corrections to the two-particle self-energy are needed to satisfy the Ward identity, or, equivalently, the theory is only conserving if the limit as $D \rightarrow \infty$ is evaluated only after the functional derivatives of Φ (e.g., $\Gamma_{\sigma, \sigma'} = \delta^2 \Phi / \delta \mathbf{G}_\sigma \delta \mathbf{G}_{\sigma'}$) are evaluated.³²

In a similar way, one may explore violation of the Ward identities by the DCA. The required Ward identity in this case can be written as

$$\begin{aligned}\Sigma_c(\mathbf{K} + \mathbf{Q}, i\nu_\alpha + i\omega_n) - \Sigma_c(\mathbf{K}, i\omega_n) &= \frac{T}{N} \sum_{\mathbf{K}', \tilde{\mathbf{k}}, \mathbf{k}, \mathbf{n}'} \{ [G(\mathbf{K}' + \tilde{\mathbf{k}}, i\omega_n') - G(\mathbf{K}' + \tilde{\mathbf{k}} + \mathbf{Q} + \tilde{\mathbf{q}}, i\omega_n' + i\nu_\alpha)] \\ &+ [\Sigma_c(\mathbf{K}' + \mathbf{Q}, i\nu_\alpha + i\omega_n) - \Sigma_c(\mathbf{K}', i\omega_n)] G(\mathbf{K}' + \tilde{\mathbf{k}}, i\omega_n') G(\mathbf{K}' + \tilde{\mathbf{k}} + \mathbf{Q} + \tilde{\mathbf{q}}, i\omega_n' + i\nu_\alpha) \} \\ &\times \mathbf{T}_{\mathbf{Q} + \tilde{\mathbf{q}}, i\nu_\alpha}^{(2)c}(\mathbf{K}, i\omega_n; \mathbf{K}', i\omega_n'),\end{aligned}\quad (\text{D6})$$

where we have used the DCA in assuming that Σ and Γ are dependent only on the cluster momenta, and $\mathbf{T}^{(2)c}$ is defined in Appendix A. Now it is clear that, to the extent that the RHS depends on \mathbf{q} , the Ward identity will not be satisfied, even in the static case.

However, the DCA will be conserving in the limit of large cluster size, since momentum conservation at the internal vertices is restored (with corrections of order Δk). Here, we assume that the method used to solve the cluster is exact, or that if an approximate methods used, that the corresponding self-energy diagrams are formed from derivatives of a generating functional and employ fully dressed propagators [i.e., $\bar{G}(\mathbf{k}, \omega)$, not $\mathcal{G}(\mathbf{k}, \omega)$] so that we approximate $\Phi(G(\mathbf{k}, \omega)) \approx \Phi(\bar{G}(\mathbf{k}, \omega))$. Then, the DCA is conserving to the extent that $\Gamma_{\mathbf{q}}(\mathbf{k}, \mathbf{k}')$ and $\Sigma_{\mathbf{k}}$ are well approximated by the cluster quantities. Since $\Gamma = \Gamma^c + \mathcal{O}(\Delta k^2)$ and $\Sigma = \Sigma^c + \mathcal{O}(\Delta k^2)$, the DCA is able to restore the conservation properties lost in the

DMFA when $\Delta k = \pi/L \rightarrow 0$ with corrections of order $\mathcal{O}(\Delta k^2)$.

In this appendix we have shown that due to violations of momentum conservation, the DMFA is not a conserving approximation in any dimension D . Violations of Ward's original identity also emerge for the DMFA even when $D \rightarrow \infty$ for a vanishingly small set of momenta \mathbf{q} which includes $\mathbf{q} = 0$, but not for general momenta \mathbf{q} . There are concomitant requisite nonlocal corrections to the infinite-dimensional irreducible vertex functions for a set of measure zero points in the infinite-dimensional Brillouin zone which are necessary to restore the Ward identity for all \mathbf{q} . In finite dimensions, the DMFA violates conservation in a finite fraction of the Brillouin zone due to the lack of momentum conservation in the internal vertices of the generating functional. Momentum conservation is restored by DCA systematically as the cluster size increases, and so the DCA restores the conserving nature of the approximation.

*Present address: Forschungszentrum Karlsruhe, Postfach 3640, D-76021 Karlsruhe, Germany.

†Present address: Department of Physics, Miami University, Oxford, OH 45056.

¹W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989).

²M. Jarrell, Phys. Rev. Lett. **69**, 168 (1992).

³A. Georges and G. Kotliar, Phys. Rev. B **45**, 6479 (1992).

⁴T. Pruschke, M. Jarrell, and J.K. Freericks, Adv. Phys. **42**, 187 (1995).

⁵A. Georges, G. Kotliar, W. Krauth, and M. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).

⁶P.G.J. van Dongen, Phys. Rev. B **50**, 14 016 (1994). It follows immediately from Eq. (A9) therein, that the spectral weight has an acausal part of order $\exp(-2/\sqrt{d})$ [P.G.J. van Dongen (private communication)].

⁷A. Schiller and K. Ingersent, Phys. Rev. Lett. **75**, 113 (1995).

⁸Q. Si and J. L. Smith, Phys. Rev. Lett. **77**, 3391 (1996); J. L. Smith and Q. Si, Phys. Rev. B **61**, 5184 (2000).

⁹M. H. Hettler, A. N. Tahvildar-Zadeh, M. Jarrell, T. Pruschke, and H. R. Krishnamurthy, Phys. Rev. B **58**, 7475 (1998).

¹⁰J.C. Ward, Phys. Rev. **78**, 182 (1950).

¹¹G. Baym, Phys. Rev. **127**, 1391 (1962).

- ¹²C.M. Varma, P.B. Littlewood, S. Schmitt-Rink, E. Abrahams, and A.E. Ruckenstein, *Phys. Rev. Lett.* **63**, 1986 (1989).
- ¹³Y. Zha, V. Barzykin, and D. Pines, *Phys. Rev. B* **54**, 7561 (1996).
- ¹⁴N.E. Bickers and D.J. Scalapino, *Ann. Phys. (N.Y.)* **193**, 206 (1993).
- ¹⁵E. Müller-Hartmann, *Z. Phys. B: Condens. Matter* **74**, 507 (1989).
- ¹⁶A. Gonis, *Green Functions for Ordered and Disordered Systems* (North-Holland, Amsterdam, 1992), p. 204.
- ¹⁷This is the only choice for the cluster \mathbf{K} values that are evenly spaced in the D -dimensional Brillouin zone and form a group under addition. However, one may also choose \mathbf{K} values that correspond to a cluster with lower dimensionality than D . For example, one could choose just two cluster \mathbf{K} values, corresponding to a two-site cluster. However, care must be taken to properly symmetrize the corresponding cluster self energy, before using it in Eqs. (2) and (3).
- ¹⁸D.F. Elliot and K.R. Rao, *Fast Transforms: Algorithms, Analyses, Applications* (Academic Press, New York, 1982).
- ¹⁹H. Akhlaghpour *et al.* (unpublished).
- ²⁰P. de Vries, K. Michelsen, and H. De Raedt, *Phys. Rev. Lett.* **70**, 2463 (1993); *Z. Phys. B: Condens. Matter* **92**, 353 (1993); **95**, 475 (1994).
- ²¹H. DeRaedt and W. von der Linden, in *The Monte Carlo Method in Condensed Matter Physics*, edited by K. Binder (Springer, Berlin, 1992).
- ²²M.P. Pasternak, R.D. Taylor, A. Chen, C. Meade, L.M. Falicov, A. Gieseckus, R. Jeanloz, and P.Y. Yu, *Phys. Rev. Lett.* **65**, 790 (1990); A.L. Chen, P.Y. Yu, and R.D. Taylor, *Phys. Rev. Lett.* **71**, 4011 (1993).
- ²³U. Brandt and R. Schmidt, *Z. Phys. B: Condens. Matter* **63**, 45 (1986); **67**, 43 (1987).
- ²⁴U. Brandt and C. Mielsch, *Z. Phys. B: Condens. Matter* **75**, 365 (1989).
- ²⁵M. Suzuki, *J. Phys. Soc. Jpn.* **55**, 4205 (1986).
- ²⁶In situations where only small clusters are numerically accessible it might be advisable to average over “all” possible choices of the cluster momenta, thus reducing the effect of “special” points in the Brillouin zone.
- ²⁷This can be best seen in momentum space. For APBC the number of independent \mathbf{K} points, e.g., for a 4×4 cluster, is the same as the number for a 2×2 cluster with PBC. After a shift and a rescaling (doubling of the lattice constant) the equivalence of the two clusters is easily established.
- ²⁸C. Huscroft, R. Gass, and M. Jarrell, cond-mat/9906155 (unpublished).
- ²⁹D. Duffy and A. Moreo, *Phys. Rev. B* **55**, 12 918 (1997).
- ³⁰G. Baym and L.P. Kadanoff, *Phys. Rev.* **124**, 287 (1961).
- ³¹We thank P.G.J. van Dongen for pointing out that different results are obtained depending upon the order in which limit as $D \rightarrow \infty$ or the functional derivatives of Φ are evaluated. See also P.G.J. van Dongen, G.S. Uhrig, and E. Müller-Hartmann, *Phys. Rev. Lett.* **83**, 2865 (1999).
- ³²The fact that the DMFA is not conserving when $D \rightarrow \infty$ does not indicate that it is not exact in this limit. In fact, the theory remains exact, since the requisite nonlocal corrections to the irreducible vertex Γ occur only for a set of points of measure zero in the Brillouin zone. Thus, where Γ is used in an integral equation to calculate a measurable quantity, such as a susceptibility, or even the dc conductivity, the contribution of these nonlocal corrections is negligible.