PHYSICAL REVIEW B **77**, 075105 (2008)

Convexity of the self-energy functional in the variational cluster approximation

Andriy H. Nevidomskyy,* David Sénéchal, and A.-M. S. Tremblay
Départment de Physique, Université de Sherbrooke, Sherbrooke, Québec, Canada J1K 2R1

(Received 5 July 2007; revised manuscript received 24 September 2007; published 7 February 2008)

In the variational cluster approximation (VCA) (or variational cluster perturbation theory), widely used to study the Hubbard model, a fundamental problem that renders variational solutions difficult in practice is its known lack of convexity at stationary points, i.e., the physical solutions can be saddle points rather than extrema of the self-energy functional. Here, we suggest two different approaches to construct a convex functional $\Omega[\Sigma]$. In the first approach, one can show analytically that in the approximation where the irreducible particle-hole vertex depends only on center of mass coordinates, the functional is convex away from phase transitions in the corresponding channel. Numerical tests on a tractable version of that functional show that convexity can be a nuisance when looking for instabilities both in the pairing and particle-hole channels. Therefore, an alternative phenomenological functional is proposed. Convexity is explicitly enforced only with respect to a restricted set of variables, such as the cluster chemical potential that is known to be otherwise problematic. Numerical tests show that our functional is convex at the physical solutions of VCA and allows second-order phase transitions in the pairing channel as well. This opens the way to the use of more efficient algorithms to find solutions of the VCA equations.

DOI: 10.1103/PhysRevB.77.075105 PACS number(s): 71.10.Fd, 71.15.Dx, 71.10.−w

I. INTRODUCTION

Effective functionals, such as the Landau-Ginzburg free energy functional in the vicinity of phase transitions, have long been used to study classical and quantum systems. Typically, the effective functional $F[h]$ is obtained in terms of a relevant variable h, such as the order parameter in the Landau-Ginzburg theory or the electron density in the density functional theory.[1] The optimal value of h is then obtained from the requirement that the functional be stationary at the solution, $\delta F/\delta h=0$.

In the formalism of Luttinger-Ward[2] and Baym-Kadanoff,[3] one of the best known functional approaches for correlated electrons,[4] a functional $\Omega[G]$ is stationary and equal to the grand potential at the exact physical value of G. In such a scheme, the functional dependence $\Omega[G]$ is not known exactly. It can be approximated perturbatively by summing a subset of the infinite series of skeleton diagrams that define the Luttinger-Ward functional. On the other hand, one of the most effective methods to obtain nonperturbative results is the dynamical mean-field theory (DMFT).[5] Chitra and Kotliar[6] have shown how this theory can be obtained by modifying the Kadanoff-Baym functional and making the local approximation on the stationarity condition. More recently, a general scheme for generating a wide class of nonperturbative approximations for the Hubbard model[7] from a functional has been proposed by Potthoff.[8] In this method, known as the self-energy-functional approach (SFA), a new functional $\Omega[\Sigma]$ of the self-energy Σ is constructed, which is stationary at the physical solution. The functional itself is unknown explicitly, but Potthoff suggested a particular way of calculating a stationary solution $\delta\Omega/\delta\Sigma=0$ with the help of a reference system, typically a cluster of finite size, which can be solved exactly. This particular implementation of the self-energy functional with no bath (contrary to DMFT) goes under the name of the *variational cluster approximation* (VCA). DMFT and generalizations thereof, (such as cellular dynamical mean-field theory[9]) can be obtained as various special cases of SFA[8,10] corresponding to different choices of reference systems and/or approximations of the stationarity condition. (The functionals of Chitra and Kotliar[6] and of Potthoff[8,10] are, in fact, identical, as shown in Appendix C).

One desirable feature of functional approaches is the variational principle that guarantees that an approximate grand potential is an upper bound to the true grand potential. Such a variational principle is missing for the stationary solutions of both the Baym-Kadanoff functional[3,11] and Potthoff's self-energy functional[8] since stationary solutions are known to be saddle points rather than extrema.

A question thus remains open up to present, whether or not it is possible to construct a functional, be it a functional of the Green function G or of the self-energy Σ, such that its stationary solutions would always be extrema (say, minima) of the functional. If so, this would mean that the functional is convex at the physical solution, which is a first necessary step on the way to prove the variational principle. For an infinite-coordination Bethe lattice, this question was answered positively by Kotliar,[11] who proved that a functional could be constructed such that its extrema occurred at the physical local Green function of the Hubbard model. However, an attempt by Chitra and Kotliar to find its analog for a finite-dimensional lattice was only partially successful.[6]

Another motivation to find convex functionals is a practical one. In VCA, the functional is definitely not convex, for example, when the intracluster chemical potential is varied. The physical solutions are then saddle points. However, most efficient numerical algorithms such as, e.g., the conjugate gradient method, have been designed to find extrema of a functional rather than saddle points. Although one may attempt to use such algorithms to also find saddle points (by minimizing the magnitude square of the gradients of the functional[12]), unphysical solutions may occur.

In this paper, we reexamine the above problem and show that a convex functional can be found. In particular, we construct a new functional of the self-energy $\Omega[\Sigma]$, such that its stationary solutions are minima when the irreducible particle-hole vertex depends only on center of mass coordinates and the system is away from phase transitions in the corresponding channel. Going beyond this approximation involves complicated integrals that cannot be treated analytically; however, numerical tests on a tractable version of the functional suggest that it is indeed always convex at the physical solution. Moreover, we show that such a construction is not unique and that several functionals can be constructed that differ in the higher-order terms of the expansion with respect to the self-energy.

Despite the convexity of the proposed functional, its implementation requires additional approximations and it turns out to be inadequate to detect a second-order phase transition for both Cooper pairing and antiferromagnetic instabilities. There, the functional appears to always be convex at the *paramagnetic* solution, whereas in the case of second-order phase transitions, the paramagnetic solution is rightly expected to be a saddle point, with minima developing instead at a finite value of the symmetry-breaking order parameter. In other words, the sought convexity of the functional "overdoes" its job, imposing too stringent conditions on the resulting physical solution.

Given that the VCA method has originally been developed to study broken-symmetry phases, the aforementioned feature of the proposed functional is especially undesirable. To cure this drawback, we propose a different, this time phenomenological, approach that ensures convexity at the physical solution and, moreover, respects the tendency of the system to develop an instability toward a broken-symmetry phase. This finding opens up different perspectives in the use of the convexity property of the functional, in particular, permitting to apply powerful numerical techniques, such as the conjugate gradient method, which are only guaranteed to work if the solution is known to be an extremum (and not a saddle point) of the functional.

The paper is organized as follows. First, in Sec. II, we review the VCA as proposed by Potthoff,[8] followed by a general discussion of the stability of the stationary solution and criteria for convexity in Sec. III. We then proceed to derive a different functional of self-energy in Sec. IV, with the proof of its convexity given in Appendix A and the recipe for incorporating it into the VCA framework in Appendix B. The second part of Sec. IV is devoted to numerical tests of the proposed functional on a two-dimensional Hubbard model, which show that the functional is convex but that it fails to correctly describe second-order phase transitions into a broken-symmetry phase. Triggered by this negative result, we propose in Sec. V an answer to the convexity problem while correcting the aforementioned failure to describe second-order phase transitions. The adequacy of this different construction is corroborated by numerical tests. More technical aspects of the work are detailed in appendixes.

II. REVIEW OF THE CONVENTIONAL VARIATIONAL CLUSTER APPROXIMATION SCHEME

Consider, following Potthoff,[4,8] a general Hamiltonian $H = H_0(\mathbf{t}) + H_1(\mathbf{U})$ with one-particle hopping parameters $\mathbf{t}$ and two-particle interaction parameters U:

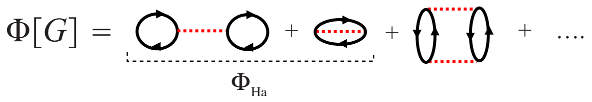

FIG. 1. (Color online) Luttinger-Ward functional $\Phi[G]$ constructed as a sum of renormalized skeleton diagrams with appropriate combinatorial coefficients (not shown). The Hartree approximation consists in considering the first two diagrams only of the series.

$$H = \sum_{ij} t_{ij} c_i^\dagger c_j + \frac{1}{2} \sum_{ijkl} U_{ijkl} c_i^\dagger c_j^\dagger c_l c_k. \tag{1}$$

Here, i, j, k, and l refer to an orthonormal and complete set of one-particle basis states. The equilibrium thermodynamics and elementary one-particle excitations of the system for temperature T and chemical potential μ are fully described by the one-particle Matsubara Green function[13] defined by the imaginary-time ordered product T

$$G_\sigma(\mathbf{r}_1,\tau_1;\mathbf{r}_2,\tau_2) = -\langle \mathrm{T}[c_\sigma(\mathbf{r}_1,\tau_1) c_\sigma^\dagger(\mathbf{r}_2,\tau_2)] \rangle, \tag{2}$$

where the symbol $\langle \ldots \rangle$ denotes thermal average and $G_\sigma(1,2) \equiv G_\sigma(\mathbf{r}_1,\tau_1;\mathbf{r}_2,\tau_2)$ can be seen as $\langle \mathbf{r}_1,\tau_1 | \hat{G}_\sigma | \mathbf{r}_2,\tau_2 \rangle$ or G_{12}, a matrix of two indices, each of which stands for both space and imaginary time.

We start by defining the Luttinger-Ward functional $\Phi[G]$, constructed formally as a sum of all closed, irreducible skeleton diagrams involving fully renormalized ("dressed") Green functions,[2] as illustrated in Fig. 1. We note that this functional is universal, that is, it does not depend on the kinetic part of the Hamiltonian $H_0(\mathbf{t})$, but only on the vertices $\mathbf{U}$ and the Green function itself.

The important property of the Luttinger-Ward functional is that its functional derivative gives the self-energy of the system:

$$\frac{\delta\Phi[G]}{\delta G_\sigma(1,2)} = \Sigma_\sigma(2,1) = \Sigma_\sigma[G]. \tag{3}$$

This equation serves as a definition of the self-energy as a functional of G.[14] We can now introduce the grand potential of the system as the Baym-Kadanoff (BK) functional[3] defined in terms of the fully dressed Green function as follows:

$$\Omega_{BK}[G] = \Phi[G] - \mathrm{Tr}((G_0^{-1} - G^{-1})G) + \mathrm{Tr}\,\ln(G). \tag{4}$$

Inverting Eq. (3) (locally[8]) to obtain the Green function as a functional of Σ, we can now express the Luttinger-Ward functional as a functional of the self-energy $\Phi[G[\Sigma]]$. Substituting this into the expression for the Baym-Kadanoff functional (4) and using Dyson's equation $G^{-1} = G_0^{-1} - \Sigma$, Potthoff proposed the following functional of the self-energy:

$$\Omega[\Sigma] = \Phi[G[\Sigma]] - \mathrm{Tr}(\Sigma G) - \mathrm{Tr}\,\ln(G_0^{-1} - \Sigma). \tag{5}$$

Recognizing that the first two terms on the right-hand side of the Potthoff functional represent the Legendre transform of the Luttinger-Ward functional

$$F[\Sigma] \equiv \Phi[G[\Sigma]] - \mathrm{Tr}(\Sigma G), \tag{6}$$

one can now rewrite the expression for the grand potential as follows:

$$\Omega[\Sigma] = F[\Sigma] - \mathrm{Tr}\ln(G_0^{-1} - \Sigma). \tag{7}$$

Using Eqs. (3) and (6), it is easy to show that the following equation holds:

$$\frac{\delta F[\Sigma]}{\delta \Sigma_\sigma(1,2)} = -G_\sigma(2,1) = -G_{21}[\Sigma]. \tag{8}$$

It can be viewed as the definition of the Green function in terms of the self-energy Σ. Using this, we immediately arrive at the conclusion that the variational derivative of the grand potential in Eq. (7) vanishes at the true physical solution:

$$\left.\frac{\delta\Omega[\Sigma]}{\delta\Sigma}\right|_{sol} = -G + \left.(G_0^{-1} - \Sigma)^{-1}\right|_{sol} = 0 \tag{9}$$

by virtue of the Dyson equation. At the physical solution, $\Omega[\Sigma]=\Omega[G]$ is equal to the true grand potential. Therefore, solving the problem amounts to finding such a function Σ that satisfies the above stationarity condition.

Since the Luttinger-Ward functional $\Phi[G]$ is a universal functional of the interaction $\mathbf{U}$ and of the stationary value of G, its Legendre transform $F[\Sigma]$ is a universal functional of $\mathbf{U}$ and Σ. In other words, the value of this functional should not depend on one-body operators such as the hopping matrix $\mathbf{t}$. This last fact is the crucial point that allows one to define VCA.

One proceeds as follows. In the case of the Hubbard model where the interaction is local, one can modify the hopping matrix elements to subdivide the infinite cluster into disjoint identical clusters. One can then compute the grand potential Ω' and the self-energy Σ' of that problem exactly (for example, by means of exact diagonalization). The one-body part of the cluster Hamiltonian can contain a different hopping matrix, along with site energies, chemical potential, and Weiss fields, all of which are used as variational parameters. Since one can write the grand potential for the cluster problem as

$$\Omega'[\Sigma'] = F'[\Sigma'] - \mathrm{Tr}\ln(G_0'^{-1} - \Sigma'), \tag{10}$$

the universality of the functional $F[\Sigma]$ allows one to find its *exact* value for the solution of the cluster problem:

$$F[\Sigma'] \equiv F'[\Sigma'] = \Omega' + \mathrm{Tr}\ln(G_0'^{-1} - \Sigma'). \tag{11}$$

Using this last result, the functional $\Omega[\Sigma]$ in Eq. (7) can be evaluated exactly when $\Sigma \to \Sigma'$

$$\Omega[\Sigma'] = \{\Omega'[\Sigma'] + \mathrm{Tr}\ln(G_0'^{-1} - \Sigma')\} - \mathrm{Tr}\ln(G_0^{-1} - \Sigma'). \tag{12}$$

Since they are the solutions of the cluster problems, the self-energies Σ' can be varied through the one-body terms of the clusters (which we define to also contain Weiss fields for various order parameters). These are collectively represented by the matrix t'_{ij}. Clearly, the self-energies obtained in this way will span only a small subspace of an infinite-dimensional space of all possible variations, namely, only those that can be represented as the physical self-energies of a cluster $\Sigma'(t')$ parametrized by the matrix t'_{ij}. The corresponding stationary solution is obtained by searching for values of t'_{ij} such that

$$\frac{d\Omega}{dt'_{i,j}} \equiv \frac{\delta\Omega}{\delta\Sigma} \cdot \frac{d\Sigma}{dt'_{i,j}} = 0. \tag{13}$$

The set of Eqs. (12) and (13) forms the essence of the VCA quantum cluster method.

In what follows, the numerical tests are performed on the example of the Hubbard model on a square lattice, for which the cluster Hamiltonian is defined as follows:

$$H' = \sum_{\mathbf{x},\mathbf{x}',\sigma} t_{\mathbf{x}\mathbf{x}'} c^\dagger_{\mathbf{x}\sigma} c_{\mathbf{x}'\sigma} - \sum_{\mathbf{x},\mathbf{x}'} (\Delta^\dagger_{\mathbf{x}\mathbf{x}'} c_{\mathbf{x}\uparrow} c_{\mathbf{x}'\downarrow} + \mathrm{H.c.}) - M \sum_{\mathbf{x},\sigma} e^{i\mathbf{Q}\mathbf{x}} (-1)^\sigma n_{\mathbf{x}\sigma} - \mu \sum_{\mathbf{x}\sigma} n_{\mathbf{x}\sigma} + U \sum_{\mathbf{x}} n_{\mathbf{x}\uparrow} n_{\mathbf{x}\downarrow}. \tag{14}$$

Here, as before, $c^\dagger_{\mathbf{x}\sigma}$ is the electron creation operator at site $\mathbf{x}$ with spin σ, $n_{\mathbf{x}\sigma}$ is the particle number operator, and M and $\Delta_{\mathbf{x}\mathbf{x}'}$ are the Weiss fields corresponding to the antiferromagnetic and superconducting order parameters, respectively. For a singlet superconductivity, we have $\Delta_{\mathbf{x}\mathbf{x}'} = \Delta_{\mathbf{x}'\mathbf{x}}$. In particular, for the $d_{x^2-y^2}$ symmetry, the Weiss field is defined as follows ($\mathbf{e}$ is a lattice vector):

$$\Delta_{\mathbf{x},\mathbf{x}+\mathbf{e}} = \begin{cases} D & \text{for } \mathbf{e} = \pm\hat{x} \\ -D & \text{for } \mathbf{e} = \pm\hat{y}. \end{cases} \tag{15}$$

III. STABILITY OF THE STATIONARY SOLUTION

Let us consider fluctuations around the stationary solution of an arbitrary self-energy functional:

$$\delta\Omega = \Omega[\Sigma + \delta\Sigma] - \Omega[\Sigma] = \int\int \delta\Sigma_\sigma(1',1) \frac{\delta^2\Omega}{\delta\Sigma_\sigma(1',1)\delta\Sigma_{\sigma'}(2',2)} \delta\Sigma_{\sigma'}(2',2). \tag{16}$$

The spin indices have been written explicitly and must be summed over. From now on, we will not write them explicitly to have a lighter notation.

For a stationary point to be numerically stable, it must be a minimum. A maximum will also do since it suffices to change the sign. Correspondingly, the functional derivative in Eq. (16) must be negative or positive definite, respectively. It is easy to verify that for Potthoff's self-energy functional [Eq. (7)], the second functional derivative is given by

$$\frac{\delta^2\Omega}{\delta\Sigma_{11'}\delta\Sigma_{22'}} = \Gamma_{11';22'} + G_{1'2}G_{2'1}, \tag{17}$$

where $\Gamma_{11';22'}$ stands for the second functional derivative of the universal functional $F[\Sigma]$:

$$\Gamma_{11';22'} \equiv \frac{\delta^2 F}{\delta\Sigma_{11'}\delta\Sigma_{22'}} \tag{18}$$

and is, thus, a tensor of the fourth rank.

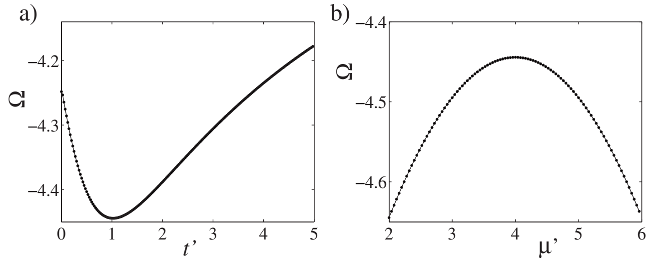

FIG. 2. Functional Ω for the two-dimensional (2D) Hubbard model ($U/t=8$) calculated from the exact-diagonalization solution of a 2×2 cluster using Eq. (12), plotted as a function of (a) cluster nearest-neighbor hopping parameter t' and (b) cluster chemical potential μ'. The variational solution of the problem in the space of cluster parameters ($t'=1$, $\mu'=4$) is a saddle point instead of an extremum.

It follows directly from the expression for the second functional derivative Eq. (17) that the stationary solution is a saddle point. Following the arguments of Ref. 6, this is simplest to illustrate in the absence of two-body interactions in the Hamiltonian, when $\Sigma=0$ and $F[\Sigma]=0$. Then the first term in Eq. (17) vanishes, while in Matsubara-Fourier space the last term leads to

$$\frac{\delta^2\Omega}{\delta\Sigma\,\delta\Sigma} = G(k_1')G(k_2')\,\delta(k_1'-k_2)\,\delta(k_2'-k_1). \qquad (19)$$

In the sector of zero total momentum and total energy ($k_1' = -k_2'$), this quantity decouples in 2×2 blocks with zero diagonal elements and equal off-diagonal components equal to $(\omega_n^2+\varepsilon_{k_1}^2)^{-1}$. The eigenvalues are, thus, both positive and negative, which corresponds to a saddle point. The effect of interactions, $\Gamma_{11';22'}$, cannot cure this problem for all wave vectors and frequencies.

In the case of VCA, the relevant question concerns stability with respect to variations in the cluster parameters. It has been pointed out empirically in the original VCA proposal[8] that variations with respect to site energies or chemical potentials lead to saddle points. This is illustrated in Fig. 2, where the grand potential obtained from a 2×2 cluster solution is plotted as a function of the intracluster nearest-neighbor hopping t' [Fig. 2(a)] and of the cluster chemical potential μ' [Fig. 2(b)]. Clearly, the stationary solution is a minimum in the former case and a maximum in the latter. Yet, varying the cluster chemical potential μ' has been shown[16] to be essential in the VCA scheme for obtaining thermodynamic consistency, which requires that the number $\langle n\rangle$ of electrons in the system has the same value when calculated from the two independent relations:

$$\langle n\rangle = -\frac{\partial\Omega}{\partial\mu} = -\frac{1}{\pi}\int_{-\infty}^{\infty} d\omega f(\omega)\,\mathrm{Im}[\mathrm{Tr}\, G(\omega+i0^+)], \qquad (20)$$

where $f(\omega)=[\exp(\omega/T)+1]^{-1}$ is the Fermi function. Therefore, it is preferable to let μ' vary and somehow deal with the fact that the grand potential is known to be nonconvex in this case.

The fact that the solution is a saddle point rather than an extremum has consequences for practical implementations, since there seems to exist no robust numerical search algorithms for a saddle point, whereas many such algorithms have been developed for extrema[17] searches. It would, therefore, be desirable from this point of view, as well as for the reasons outlined in Sec. I, to find such a functional $\Omega[\Sigma]$ whose stationary solutions would be guaranteed to be extrema. In the next section, we shall demonstrate that such a functional can, indeed, be formally constructed, and its convexity can be proved rigorously given certain approximations.

IV. CONVEX FUNCTIONALS Ω[Σ]

In this section, we first derive a general expression for a convex functional. The drawback of this approach is that in practical implementations, convexity is preserved for order parameters as well. This makes it an undesirable feature in the presence of phase transitions. In the next section, we will restrict ourselves to preserving convexity as a function of the cluster chemical potential, which in practice is the main problem to be solved.

A. Analytical results for convex functionals

The main idea of the current approach is to add a term $\Delta\Omega[\Sigma]$ to the grand potential that is quadratic in Σ and that does not alter the stationary solution of the original potential $\Omega[\Sigma]$, but renders it a positive-definite functional of the self-energy. We first introduce an auxiliary functional $\mathrm{f}[\Sigma]$ that is defined by

$$\mathrm{f}[\Sigma] = \mathrm{f}(1,2) \equiv \frac{\delta\Omega[\Sigma]}{\delta\Sigma(2,1)} = \frac{\delta F[\Sigma]}{\delta\Sigma(2,1)} + (G_0^{-1}-\Sigma)^{-1}(1,2) \qquad (21)$$

and that, by virtue of the stationarity condition Eq. (9), vanishes identically at the stationary solution of $\Omega[\Sigma]$.

It turns out that the simplest term $\Delta\Omega[\Sigma]$ that can be added to the grand potential is of the form

$$\Delta\Omega_1 = -\frac{1}{2}\mathrm{Tr}([(G_0^{-1}-\Sigma)\mathrm{f}]^2) = -\frac{1}{2}\mathrm{Tr}\left(1+(G_0^{-1}-\Sigma)\frac{\delta F}{\delta\Sigma}\right)^2. \qquad (22)$$

Indeed, $\Delta\Omega_1$ and its functional derivative both vanish at the stationary solutions of the grand potential, following the definition of the functional $\mathrm{f}[\Sigma]$ and, therefore, the new functional

$$\Omega_1[\Sigma] = \Omega[\Sigma] + \Delta\Omega_1[\Sigma] \qquad (23)$$

yields the same stationary solution as the original one. The proof of the convexity of this new functional in the special case where the irreducible particle-hole vertex is local is given in Appendix A.

We note in passing that our choice of $\Delta\Omega[\Sigma]$ is not unique. To second order in the quantity $\mathrm{f}[\Sigma]$, the correction

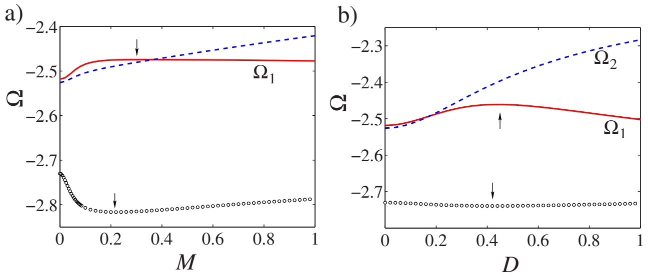

FIG. 3. (Color online) Dependence of the original Potthoff functional Ω (open circles), functional Ω_1 (solid red line), and Ω_2 (broken blue line) on the symmetry-breaking Weiss fields conjugate to (a) antiferromagnetic staggered magnetization and (b) *d*-wave superconductivity. Arrows denote positions of minima of Ω and maxima of Ω_1. Results are for the Hubbard model on a square lattice at half-filling with $U/t=4$. Curiously, both SC and antiferromagnetic solutions are found at this low value of *U*, but the magnetic solution is favored since it has lower total energy.

$\Delta\Omega_1$ given by Eq. (22) is the same as, for example, the following one:

$$\begin{aligned}\Delta\Omega_2 &= \mathrm{Tr}\ln(1-(G_0^{-1}-\Sigma)\mathrm{f}) + \mathrm{Tr}((G_0^{-1}-\Sigma)\mathrm{f})\\ &= \mathrm{Tr}\ln\left((G_0^{-1}-\Sigma)\frac{\delta F}{\delta\Sigma}\right) + \mathrm{Tr}\left(1+(G_0^{-1}-\Sigma)\frac{\delta F}{\delta\Sigma}\right),\end{aligned} \tag{24}$$

meaning that it leaves the stationary point of the original functional unchanged and leads to exactly the same expression for the second functional derivative as Eq. (22). Note also that despite being similar in spirit to the work of Chitra and Kotliar,[6] the above derivation is significantly different, as demonstrated in Appendix C where the connection between the Chitra-Kotliar result and the Potthoff functional is exposed.

While implementing Eqs. (22) and (24) in practice, a problem arises since Eq. (21) for the auxiliary functional f[Σ] contains a functional derivative $\delta F[\Sigma]/\delta\Sigma$ whose value is not readily available in practical calculations. Therefore, an approximation must be made in order to implement the new functional into the VCA scheme. One such elegant approximation is given in Appendix B, and the conclusions of the numerical verification of its convexity is the subject of the following section.

B. Drawback of this type of functionals

While numerical tests show convincingly (see Appendix B for details) that the functionals proposed in Eqs. (22) and (B5) are, indeed, convex with respect to all cluster variational parameters, including the cluster chemical potential μ', there is a major drawback. The numerical tests show one undesirable property, namely, its failure to correctly predict second-order phase transitions, at least for the approximation to f[Σ] proposed in Appendix B.

Consider Fig. 3(a) that shows the dependence of Ω and of the new functionals given by Eqs. (B4) and (24) on the magnitude of the antiferromagnetic Weiss field *M*, which is defined by Eq. (14). Unlike the original functional, which has a minimum at a finite value of *M*, both new functionals have a single minimum at $M=0$, thereby favoring the paramagnetic (PM) solution. The behavior as a function of the superconducting *d*-wave Weiss field, *D* defined in Eq. (15), follows the same pattern, as shown in Fig. 3(b). Only the original functional Ω exhibits a minimum at nonzero value of *D*, whereas both functionals proposed in this work have no other minima except at $D=0$. Clearly, such behavior of the new functionals is unphysical since the existence of antiferromagnetism in the Hubbard model, for example, around $U=4$ at half-filling, is solidly established.

As discussed in Appendix A, we would have expected that at least in the particle-hole channel, instabilities (such as antiferromagnetism) could have been detected by the present functional. This might well have been the case had we been able to evaluate the proposed functionals given by Eqs. (22)–(24) exactly. However, the key quantity $\delta F[\Sigma]/\delta\Sigma$ entering Eq. (21) remains unknown and had to be approximated in terms of known cluster quantities as explained in Appendix B. As a result, even though, by construction, the functionals Ω_1 and Ω_2 should have exactly the same value at the stationary solution as the original Potthoff functional, the numerical results in Fig. 3 show unequivocally that the approximation we used for $f[\Sigma]$, as given by Eq. (B2), does not satisfy this requirement. As a consequence, the plots of Ω in Fig. 3 suggest that, within this unavoidable approximation, the requirement of convexity in some sense overdoes its job by rendering the paramagnetic solution always a minimum and, hence, ignoring a possible instability toward a broken-symmetry phase.

Unfortunately, we have found no other satisfactory way of implementing the functional $f[\Sigma]$ within the VCA framework. Nevertheless, the proposed functionals (22)–(24) are still useful in practical calculations, given their very desirable feature of being convex at the solution. Perhaps in the future, a better approximation could be invented for evaluating these functionals in practice.

V. CURING THE PROBLEM OF CONVEXITY ARISING FROM A RESTRICTED SET OF VARIABLES

In practical implementations of VCA, the cluster chemical potential μ' leads to a first-order saddle point of the grand potential $\Omega(\mu')$ even in the paramagnetic state. Although there exist some numerical techniques that can search for such saddle-point solutions in the multidimensional space of cluster parameters,[18] it would still be highly desirable, in the context of VCA, to come up with a practical way to transform the saddle point coming from variation of the cluster chemical potential μ' into a minimum, without affecting other saddle points that have physical meaning, such as those that occur in first-order transitions. It is very important to achieve this goal while also preserving the "right" of the system to develop a second-order phase transition, thereby avoiding the problem discussed above in Sec. IV B and illustrated in Fig. 3. In this section, we show that such a construction is, indeed, possible and it seems to work well even in cases where the analytically obtained approximations given in Eqs. (B4) and (B5) fail.

Let us recall that the general form of the convex correction developed in Sec. IV was

$$\Delta\Omega \sim \mathrm{Tr}\left(\frac{\delta\Omega}{\delta\Sigma}\right)^2. \tag{25}$$

Typically, we know from experience about the existence of a certain variational parameter h (here, the cluster chemical potential μ') such that the Ω lacks convexity at the stationary solution $h=h_0$. Let us modify the correction in Eq. (25) as follows:

$$\Delta\Omega \sim \mathrm{Tr}\left(\frac{\delta\Omega}{\delta\Sigma}\cdot\frac{\partial\Sigma}{\partial h}\right)^2 = \mathrm{Tr}\left(\frac{\partial\Omega}{\partial h}\right)^2, \tag{26}$$

where we have effectively substituted the unknown functional derivative $\delta\Omega/\delta\Sigma$ by a much simpler derivative with respect to the variational parameter h, which can be easily calculated numerically.

We, thus, postulate the following functional:

$$\Omega_\lambda = \Omega + \frac{\lambda}{2}\left(\frac{\partial\Omega}{\partial h}\right)^2, \tag{27}$$

where λ is some empirical coefficient that should be chosen such that the resulting potential Ω_λ is a convex function of h. By construction, the additional term vanishes at the stationary solution $h=h_0$, where $\partial\Omega/\partial h=0$, and the value of the new potential coincides with the old one $\Omega(h_0)$. For the second derivative at h_0, we have

$$\left.\frac{\partial^2\Omega_\lambda}{\partial h^2}\right|_{h=h_0} = \frac{\partial^2\Omega}{\partial h^2}\left(1+\lambda\frac{\partial^2\Omega}{\partial h^2}\right). \tag{28}$$

If the original functional is already convex, $\partial^2\Omega/\partial h^2>0$, the new functional will be convex too. If, however, Ω has a maximum and not a minimum at h_0, the coefficient λ has to satisfy the following inequality to ensure that the new functional is convex:

$$\lambda > \lambda_0(h) \equiv \left|\left(\frac{\partial^2\Omega}{\partial h^2}\right)^{-1}\right|. \tag{29}$$

In particular, note that choosing $\lambda=2\lambda_0(h)$ will result in the new functional being convex everywhere and having the same absolute value of curvature as the original one: $\partial^2\Omega_\lambda/\partial h^2=-\partial^2\Omega/\partial h^2$. Chosen in such a way, the behavior of the newly constructed functional Ω_λ as a function of the cluster chemical potential $h\equiv\mu'$ is plotted in Fig. 4.

As we see, Ω_λ can clearly be made convex with a suitable choice of the coefficient λ. The lowest allowed value λ_0 that yields a convex functional is itself a function of parameters of the model, such as Hubbard U, as is clear from comparison of Figs. 4(a) and 4(b). In particular, we find $\lambda_0=5.8$ for $U=4t$ and $\lambda_0=9.9$ for $U=8t$.

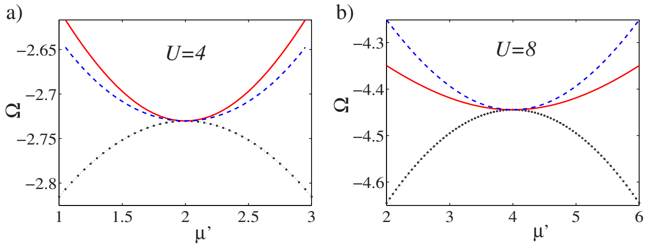

FIG. 4. (Color online) Dependence on the cluster chemical potential μ' of Potthoff's Ω (black dots), and the functional Ω_λ in Eq. (27) for two choices of the coefficient λ: the constant $\lambda=15$ (solid red line) and the variable $\lambda=2\lambda_0(\mu')$ (dashed blue line). Results for the half-filled Hubbard model with (a) $U/t=4$ and (b) $U/t=8$ are shown.

Note that by taking derivatives with respect to cluster parameters in Eq. (26) instead of the more general Eq. (25), we have substituted a strong requirement of convexity with respect to all variations in Σ with a much weaker one, which only requires that the functional be convex with respect to a particular parameter (or set of parameters) h. Let us then see what effect this has on the dependence of Ω_λ on a symmetry-breaking parameter, such as the superconducting Weiss field D. The first derivative with respect to D reads

$$\frac{\partial\Omega_\lambda}{\partial D} = \frac{\partial\Omega}{\partial D} + \lambda\frac{\partial^2\Omega}{\partial\mu'\partial D}\frac{\partial\Omega}{\partial\mu'}. \tag{30}$$

All the stationary points of Ω (where derivatives with respect to both μ' and D vanish) are also stationary points of the new functional Ω_λ. That functional can also have additional stationary points that are not stationary points of Ω, but with the procedure suggested below to choose λ, we found this not to be an issue.

In numerical calculations, we searched for minima of Ω_λ and found that, indeed, one recovers as minima the same solutions (μ'_0,D_0) that were saddle points of Ω. To illustrate how one can choose λ in practice, we consider several cases. First, in Fig. 5(a), μ' is fixed at the known superconducting solution μ'_0, and λ is taken either as constant or defined by the procedure in Eq. (29) that guarantees convexity. All curves have their minimum at the same value $D=D_0$ as expected.

To gain more insight into the convergence process, we can also check whether the solution for D is stable in cases where the value of the cluster chemical potential μ' is slightly off the true solution μ'_0. Such a situation is illustrated in Fig. 5(b), where the chemical potential μ' is fixed at the PM value $\mu'_{\mathrm{pm}}=2.45$ instead of the true superconducting (SC) solution $\mu'_0=2.28$. As expected from Eq. (27), both the old and the new functionals coincide at the paramagnetic solution $D=0$ since, by construction, $\partial\Omega/\partial\mu'=0$ there. The situation is different, however, away from $D=0$. Fixing $\lambda=\mathrm{const}$ [solid line in Fig. 5(b)] yields a minimum at an incorrect value of $D=0.84$ instead of $D_0=0.56$. By contrast, a variable coefficient $\lambda=2\lambda_0(\mu')$, where λ_0 is determined at each point from Eq. (29), gives a minimum of $\Omega_\lambda(D)$ [dashed line in Fig. 5(b)] that is already very close to the true solution. Varying μ' as well eventually leads to the correct solution in all cases, as mentioned above. An empirical case-by-case analysis has shown that the choice $\lambda=2\lambda_0(\mu')$ gives consistently reliable results and is preferred over fixing λ to a constant value. Besides, such a choice avoids the ambiguity in the value of λ and ensures that the functional $\Omega_\lambda[\mu']$ is always a convex one, according to the inequality in Eq. (29).

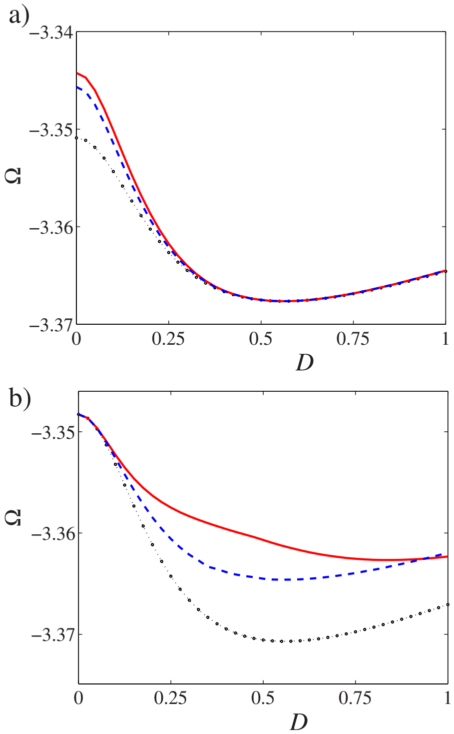

FIG. 5. (Color online) Dependence on the symmetry-breaking superconducting Weiss field D of Potthoff's grand potential (black dots), and of the functional in Eq. (27) constructed for two choices of the coefficient λ: the constant $\lambda=15$ (solid red curve) and the variable $\lambda=2\lambda_0(\mu')$ (dashed blue curve). The calculations are for the electron-doped Hubbard model (2.2% doping) with $U=4t$ and with next-nearest-neighbor hopping $t_2=0.3t$. The cluster chemical potential μ' was fixed at the value corresponding to (a) the true SC solution ($D=0.56$) and (b) the PM solution ($D=0$).

The example of $\Omega(D)$ shown above in Fig. 5 corresponds to a second-order phase transition with the only minimum of the potential Ω occurring at a nonzero value of the symmetry-breaking order parameter (D in this case). An interesting question arises as to the performance of the proposed functional in the case of a first-order phase transition. There, two minima occur, corresponding to paramagnetic ($D=0$) and ordered ($D=D_0$) phases, separated by an unstable maximum (or, generally, saddle point) of the functional. First of all, it is important to realize that the construction Eq. (27) is only aimed at removing the saddle point with respect to the cluster chemical potential μ', and *not* the physical order parameter D. To see the effect of our construction on $\Omega(D)$, let us take further derivative of Eq. (30) with respect to D at the stationary point where $\partial\Omega/\partial\mu_0'=0$:

$$\left.\frac{\partial^2\Omega_\lambda}{\partial D^2}\right|_{\mu'=\mu_0'}=\left.\frac{\partial^2\Omega}{\partial D^2}\right|_{\mu_0'}+\lambda\left(\left.\frac{\partial^2\Omega}{\partial\mu'\partial D}\right|_{\mu_0'}\right)^2. \quad (31)$$

Since the second term above is always positive, a minimum of the old functional would necessarily be a minimum of the new functional Ω_λ. As for the unstable phase characterized by a maximum of $\Omega(D)$, one would expect it to remain so provided the mixed derivative $\frac{\partial^2\Omega}{\partial\mu'\partial D}$ is sufficiently small, which we have found to be the case in practice. Therefore, no qualitative changes in the behavior of Ω as a function of a symmetry-breaking order parameter are expected in the current scheme. As for the choice between the two minima occurring at $D=0$ and $D=D_0$, the values of the potential at these two points have to be compared to ensure that the true physical solution corresponds to the global minimum.

A technical remark regarding computational efficiency is due. The knowledge of the derivatives $\partial\Omega/\partial h$ and $\partial^2\Omega/\partial h^2$ is required in order to implement the convex correction (27) with $\lambda=2\lambda_0(h)$. Numerically, this requires the knowledge of the functional Ω at three points for the finite-differencing algorithm to work. For this reason, the computational cost of evaluating Ω at a given point in parameter space is three times that of the original functional Ω. To counterbalance, the minima of the constructed *convex* functional may now be searched more efficiently using powerful numerical methods designed for extrema search, such as the conjugate gradient method.[17]

We note in passing that if, for the sake of the argument, we do not vary the cluster chemical potential μ' but instead identify it with the true potential μ, then we could rewrite Eq. (27) as

$$\Omega_\lambda \sim \Omega+\lambda\langle n\rangle^2, \quad (32)$$

where we have used the fact that $\langle n\rangle=-\partial\Omega/\partial\mu$. The above Eq. (32) is equivalent to defining a new chemical potential $\mu_\lambda=\mu-\lambda\langle n\rangle$. In this respect, we note a similarity in spirit of our approach to a modified self-consistent scheme proposed in the framework of DMFT by Tong *et al.*,[19] although in the latter case the goal was not the convexity of the potential $\Omega(\mu)$ but a useful description of a first-order phase transition in terms of a single-valued function $\mu_\lambda(\langle n\rangle)$ instead of a multivalued $\mu(\langle n\rangle)$, typical of a first-order phase transition. Despite the conceptual differences between these two approaches, it is instructive that, as in our case, the authors of Ref. 19 have found that the final results do not depend on the choice of the empirical parameter λ, provided it was large enough.

VI. CONCLUSIONS

It is known that physical solutions obtained with VCA are, in general, saddle points rather than extrema of the functional $\Omega[\Sigma]$. This is so, in particular, in the important practical case where the cluster chemical potential is varied. Saddle points are notoriously more difficult to find numerically than extrema. We have, thus, constructed a different functional that we proved to be a convex functional of the self-energy at the stationary solution, at least in the case where the irreducible particle-hole vertex depends only on the center of mass coordinates (Appendix A). It turns out, however, that implementing the proposed functional in practice is far from simple, since it involves, as can be seen in Eq. (22), an unknown functional derivative $\delta F/\delta\Sigma$ of the Legendre transform $F[\Sigma]$ of the Luttinger-Ward functional. An approximation, therefore, had to be made to express this functional derivative in terms of cluster-defined quantities, as detailed in Appendix B. The corresponding numerical results in Sec. IV B show that the functional is, indeed, convex even in the general nonperturbative case. However attractive this may be, the proposed functional changes qualitatively the behavior of Ω with respect to symmetry-breaking order parameters, such as magnetism or superconductivity, rendering instead the paramagnetic solution stable.

To cure this problem, we proposed a functional Eq. (27) and a (nonunique) recipe to find the coefficient λ that guarantees convexity with respect to cluster chemical potential μ' only. This approach removes the saddle point normally associated with μ', has the same physical solutions as the original problem, and leaves unchanged the minimum or maximum character of the functional with respect to variations of symmetry-breaking order parameters, such as magnetism or superconductivity.

Despite the nonuniqueness of the proposed convex functional and the admittedly phenomenological basis of its derivation, we argue that this is an important result. In particular, our findings open a way to the use of powerful numerical algorithms in solving for minima, such as, e.g., conjugate gradients, that only work provided the functional is convex at the physical solution. We consider the use of such efficient algorithms as highly desirable for the VCA approach. The question of the existence of a functional that would give a bound for the true grand potential and would thereby implement the variational principle is left open.

ACKNOWLEDGMENTS

We thank G. Kotliar for useful discussions and his comments on the manuscript, and J. P. Dussault for pointing out Ref. 18. Computations were performed on the Elix and the Ms RQCHP clusters. The work was supported by FQRNT (Québec), NSERC (Canada), CFI (Canada), CIAR, and the Tier I Canada Research Chair Program (A.-M.S.T.).

APPENDIX A: CONVEXITY OF THE ANALYTICAL FUNCTIONAL

We aim to prove the convexity of the proposed functional $\Omega_i[\Sigma]$ given by either Eq. (22) ($i=1$) or (24) ($i=2$). At the true solution where $\delta\Omega_i/\delta\Sigma = \mathrm{f}[\Sigma]=0$, we have that

$$\left.\frac{\delta^2\Omega_i}{\delta\Sigma_{11'}\delta\Sigma_{22'}}\right|_{\mathrm{f}=0} = \frac{\delta^2\Omega}{\delta\Sigma_{11'}\delta\Sigma_{22'}} - \mathrm{Tr}\left[G^{-1}\frac{\delta \mathrm{f}}{\delta\Sigma_{11'}}G^{-1}\frac{\delta \mathrm{f}}{\delta\Sigma_{22'}}\right]\Bigg|_{\mathrm{f}=0}, \tag{A1}$$

with $G^{-1}=(G_0^{-1}-\Sigma)$ and matrix multiplication implied for the indices not explicitly written. We then use the definitions

$$\frac{\delta \mathrm{f}_{12}}{\delta\Sigma_{34}} = \frac{\delta^2 F}{\delta\Sigma_{21}\delta\Sigma_{34}} + \frac{\delta G_{12}}{\delta\Sigma_{34}} = \Gamma_{21;34} + G_{13}G_{42}, \tag{A2}$$

where $\Gamma_{21;34}$ has the symmetry $\Gamma_{21;34}=\Gamma_{34;21}$, and obtain

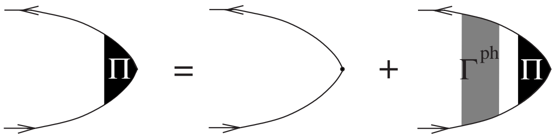

FIG. 6. Bethe-Salpeter equation for the particle-hole vertex Π in the language of Feynman diagrams.

$$\frac{\delta^2\Omega_i}{\delta\Sigma_{11'}\delta\Sigma_{22'}} = -\Gamma_{11';22'} - (G_0^{-1}-\Sigma)_{33'}\cdot\Gamma_{43';11'}\cdot(G_0^{-1} - \Sigma)_{44'}\cdot\Gamma_{34';22'}, \tag{A3}$$

with the summation implied over the repeated indices. The last term may be rewritten in the following short-hand matrix notation:

$$(G_0^{-1}-\Sigma)_{33'}\cdot\Gamma_{43';11'}\cdot(G_0^{-1}-\Sigma)_{44'}\cdot\Gamma_{34';22'} = \mathrm{Tr}[G^{-1}\Gamma^T_{;11'}G^{-1}\Gamma^T_{;22'}], \tag{A4}$$

where the transpose applies only to the indices involved in the matrix multiplication.

Irreducible vertices are usually easier to approximate. So it is useful to work with the second functional derivative of the Luttinger-Ward functional $\Phi[G]$ with respect to the single-particle Green function:

$$\Gamma^{\mathrm{ph}}{}_{11';22'} \equiv \frac{\delta^2\Phi[G]}{\delta G_{1'1}\delta G_{2'2}} = \frac{\delta\Sigma_{11'}}{\delta G_{2',2}}, \tag{A5}$$

where we have used the property Eq. (3) of the Luttinger-Ward functional. The superscript in the notation Γ^{ph} indicates that this symmetric matrix ($\Gamma^{\mathrm{ph}}_{12;34}=\Gamma^{\mathrm{ph}}_{34;12}$) is the *irreducible particle-hole vertex*, i.e.,

$$\Pi^{-1}_{12;34} = \delta_{13}\delta_{24} - G_{22'}\Gamma^{\mathrm{ph}}_{2'3';34}G_{3'1}, \tag{A6}$$

where Π is the full (reducible) particle-hole vertex. In the language of Feynman diagrams, the latter can be cast as shown in Fig. 6.

It is required on physical grounds that the reducible particle-hole vertex Π should be positive definite at zero frequency for stability in the particle-hole channel and, hence, it follows that its inverse obeys the same property. If there is an instability that affects the particle-hole channel directly or indirectly, it should still be visible in Π_i^{-1}.

Using the identity $\delta F/\delta\Sigma = -G$ and the definition of Γ given by Eq. (18), we obtain (with summation over repeated indices implied) that

$$\Gamma_{12;1'2'}\Gamma^{\mathrm{ph}}_{1'2';34} = \frac{\delta(-G)_{21}}{\delta\Sigma_{1'2'}}\cdot\frac{\delta\Sigma_{1'2'}}{\delta G_{43}} = -\delta_{2,4}\delta_{1,3} = -I, \tag{A7}$$

which again has the structure of matrix multiplication if the pairs of indices on either side of the semicolon are flattened (combined as one). Using this result, we obtain

$$\mathrm{Tr}(G^{-1}\Gamma^{T}_{;11'}G^{-1}\Gamma^{T}_{;22'})\mathrm{Tr}(G\Gamma^{\mathrm{ph}}_{22';}G\Gamma^{\mathrm{ph}}_{33';}) = \delta_{1,3}\delta_{1',3'} = I. \tag{A8}$$

Using this last identity and Eq. (A3) for $\delta^2\Omega_i/\delta\Sigma^2$, we find

$$-\frac{\delta^2\Omega}{\delta\Sigma_{11'}\delta\Sigma_{22'}}\mathrm{Tr}(G\Gamma^{\mathrm{ph}}_{22';}G\Gamma^{\mathrm{ph}}_{33';}) = \delta_{1,3}\delta_{1',3'} - G_{14}\Gamma^{\mathrm{ph}}_{44';33'}G_{4'1'}. \tag{A9}$$

As has been pointed out above, the right-hand side of this equation is positive definite at zero frequency, following from the requirement that the kernel of the particle-hole vertex equation be positive.

Let us now concentrate on $L_{22';33'}[G]=\mathrm{Tr}(G\Gamma^{\mathrm{ph}}_{22';}G\Gamma^{\mathrm{ph}}_{33';})$ appearing on the left-hand side of Eq. (A9). We assume that the irreducible particle-hole vertex is local (like in the random phase approximation), i.e.,

$$\Gamma^{\mathrm{ph}}_{11';22'} = \delta_{11'}\delta_{22'}\tilde{\Gamma}^{\mathrm{ph}}_{12}. \tag{A10}$$

With this approximation, we can write for the left-hand side of the stability equation (A9):

$$\begin{aligned}-\frac{\delta^2\Omega}{\delta\Sigma_{11}\delta\Sigma_{22}}G_{6'5'}\Gamma^{\mathrm{ph}}_{22;5'5'}G_{5'6'}\Gamma^{\mathrm{ph}}_{6'6';33}\\ = -\frac{\delta^2\Omega}{\delta\Sigma_{11}\delta\Sigma_{22}}\tilde{\Gamma}^{\mathrm{ph}}_{25'}G_{6'5'}G_{5'6'}\tilde{\Gamma}^{\mathrm{ph}}_{6'3} = -\frac{\delta^2\Omega}{\delta\Sigma_{11}\delta\Sigma_{22}}\tilde{L}_{23}.\end{aligned} \tag{A11}$$

Taking the Fourier-Matsubara transform, we find

$$\tilde{L}(\mathbf{Q},i\omega_\nu) = \tilde{\Gamma}^{\mathrm{ph}}(\mathbf{Q},i\omega_\nu)\left[\int_{\mathbf{k}}\sum_{i\omega_n}G(\mathbf{k},i\omega_n)G(\mathbf{k}+\mathbf{Q},i\omega_n+i\omega_\nu)\right]\tilde{\Gamma}^{\mathrm{ph}}(\mathbf{Q},i\omega_\nu) = -\tilde{\Gamma}^{\mathrm{ph}}(\mathbf{Q},i\omega_\nu)\chi(\mathbf{Q},i\omega_\nu)\tilde{\Gamma}^{\mathrm{ph}}(\mathbf{Q},i\omega_\nu), \tag{A12}$$

where the dressed Lindhard function $\chi(\mathbf{Q},i\omega_\nu)$ (no vertex correction) denotes the value of the integral (defined with the minus sign) in the last equation. Using the spectral representation, one can show that $\chi(\mathbf{Q},i\omega_\nu)$ for a system at equilibrium is always positive for all Matsubara frequencies. Also, $\tilde{\Gamma}^{\mathrm{ph}}(\mathbf{Q},i\omega_\nu)$ is real, following its spectral properties, so $(\tilde{\Gamma}^{\mathrm{ph}}(\mathbf{Q},i\omega_\nu))^2$ is positive. It is, thus, clear that the quantity $\tilde{L}(\mathbf{Q},i\omega_\nu)$ must be negative definite, with possible exception of phase transitions that, in this simple approximation, could appear in the particle-hole channel. Comparing this with the identity given by Eq. (A9):

$$-\frac{\delta^2\Omega_i}{\delta\Sigma\,\delta\Sigma}\cdot L[G] = 1 - G\Gamma^{\mathrm{ph}}G, \tag{A13}$$

$$-\frac{\delta^2\Omega_i}{\delta\Sigma\,\delta\Sigma}(\mathbf{Q},i\omega_\nu)\cdot\tilde{L}(\mathbf{Q},i\omega_\nu) = 1 + \chi(\mathbf{Q},i\omega_\nu)\tilde{\Gamma}^{\mathrm{ph}}(\mathbf{Q},i\omega_\nu) \tag{A14}$$

and remembering that its right-hand side is positive definite, we arrive at the conclusion that, at least in the above approximation for the particle-hole vertex, the object $\delta^2\Omega_i/\delta\Sigma^2$ must also be positive definite at zero frequency. This, in turn, means that our proposed grand potential $\Omega_i[\Sigma]$, defined by Eq. (23) and either expression (22) or (24), is a convex functional of the self-energy within this approximation.

It should be noted that being a convex functional of the self-energy is not necessarily the same as being a convex functional of a given cluster parameter h'. Indeed, differentiating again the first derivative $d\Omega/dh'$ given by Eq. (13), one obtains

$$\frac{d^2\Omega_i}{dh'^2} = \int_1\int_2\frac{\delta^2\Omega_i}{\delta\Sigma(1)\delta\Sigma(2)}\frac{d\Sigma(2)}{dh'}\frac{d\Sigma(1)}{dh'} + \int_1\frac{\delta\Omega_i}{\delta\Sigma(1)}\frac{d^2\Sigma(1)}{dh'^2}. \tag{A15}$$

Note that in the above, we have assumed translational invariance, and, hence, self-energy can be written as a function of one momentum and frequency variable only. Since the functional derivative in the first term of Eq. (A15) is almost certainly positive definite at the stationary solution, as suggested above, the first term is guaranteed to be positive too. The situation with the second term is more complicated. Naively, it may seem that this term vanishes since, at the stationary solution, $\delta\Omega/\delta\Sigma$ is zero by definition. This would be true if the variational space of the cluster parameter h' was sufficient to describe the complete variational space of the functional $\Omega[\Sigma(h')]$. Unfortunately, and this is the main approximation entering the VCA method, this may not always be true. In fact, one may only vary $\delta\Sigma(h')$ in a subspace that can be parametrized by the cluster parameter h', which is a small part of the whole infinite-dimensional space of variation $\delta\Sigma$. Therefore, the solution obtained by requiring that the derivative $\partial\Omega/\partial h'$ vanishes is not necessarily the same as the true solution where $\delta\Omega/\delta\Sigma=0$. Although we expect this contribution to be very small, it is difficult to judge the convexity of the given grand potential with respect to some cluster parameter, even if the potential is known to be a convex functional of Σ. Since varying h' is the best one can achieve within the VCA approach, we stress that numerical tests are always desirable to verify the convexity of a given functional $\Omega[\Sigma(h')]$.

APPENDIX B: IMPLEMENTATION OF THE CONVEX FUNCTIONAL IN VARIATIONAL CLUSTER APPROXIMATION

In Sec. IV, we proposed a correction $\Delta\Omega_{1,2}$ to the original grand potential $\Omega[\Sigma]$, such that it is rendered a convex functional of the self-energy. The issue that we address here is how to implement this correction in the framework of the VCA quantum cluster method.

The key element entering the expressions (22) and (24) for $\Delta\Omega$ is the auxiliary functional $\mathrm{f}[\Sigma]$ defined by Eq. (21) as a functional derivative $\mathrm{f}[\Sigma]\equiv\delta\Omega/\delta\Sigma$. Unfortunately, it is not possible to evaluate this functional derivative even numerically since we have no direct way of varying the self-energy Σ (however, we develop a variant of this idea further in Sec. V). The unknown element in $\mathrm{f}[\Sigma]$ Eq. (21) is the functional derivative $\delta F[\Sigma]/\delta\Sigma$, where $F[\Sigma]$ is a universal functional of the self-energy, as explained in Sec. II. This functional derivative can be evaluated at the *ground state solution of the cluster* by virtue of Eq. (8):

$$\frac{\delta F[\Sigma]}{\delta\Sigma}\approx\left.\frac{\delta F[\Sigma']}{\delta\Sigma'}\right|_{sol'}=-G', \tag{B1}$$

where, as before, we denote the quantities belonging to the cluster by prime. Now that both terms on the right-hand side of Eq. (21) are known, we can write

$$\mathrm{f}[\Sigma]\approxeq -G'+\frac{1}{G_0^{-1}-\Sigma}\equiv -G'+G. \tag{B2}$$

We stress that this last equation is approximate. This is because the self-energy Σ' of the cluster solution, at which the derivative in Eq. (B1) is evaluated, is not, generally speaking, equal to the *true lattice* self-energy. Another way to see this is to note that, by construction, $\mathrm{f}[\Sigma]$ must vanish identically at the true stationary solution *of the lattice*, whereas it is clear that the right-hand side of Eq. (B2) can only vanish if the two Green functions are equal to each other at the ground state *of the cluster tiling*. Rigorously speaking, this can only be the case in the limit of infinitely large cluster, where $G'\to G$.

Using this approximation Eq. (B2), we can now rewrite Eq. (22) for $\Delta\Omega_1$ as

$$\Delta\Omega_1[\Sigma]=-\frac{1}{2}\mathrm{Tr}[1-(G_0^{-1}-\Sigma)G']^2, \tag{B3}$$

so that, by virtue of Eq. (12), the new convex functional $\Omega_1=\Omega+\Delta\Omega_1$ finally becomes

$$\Omega_1[\Sigma]\approxeq\Omega'-\mathrm{Tr}\ln[(G_0^{-1}-\Sigma)G'] -\frac{1}{2}\mathrm{Tr}[1-(G_0^{-1}-\Sigma)G']^2. \tag{B4}$$

Similarly, for the correction $\Delta\Omega_2$ expressed by Eq. (24), we obtain an alternative expression for the new functional

$$\Omega_2[\Sigma]\approxeq\Omega'+\mathrm{Tr}[1-(G_0^{-1}-\Sigma)G']. \tag{B5}$$

The proposed functionals Eqs. (B4) and (B5) have been implemented into the VCA scheme using the Lanczos algorithm of exact diagonalization (ED) to solve the cluster problem. The Hubbard model on a square lattice with nearest-neighbor hopping t has been studied, and a cluster of 2×2 sites was used in the ED scheme. While the hopping parameter t and the chemical potential μ of the lattice model remain fixed, we have the freedom of varying the corresponding parameters of the cluster t' and μ', whose optimal values should be found by solving the stationarity equation (13).

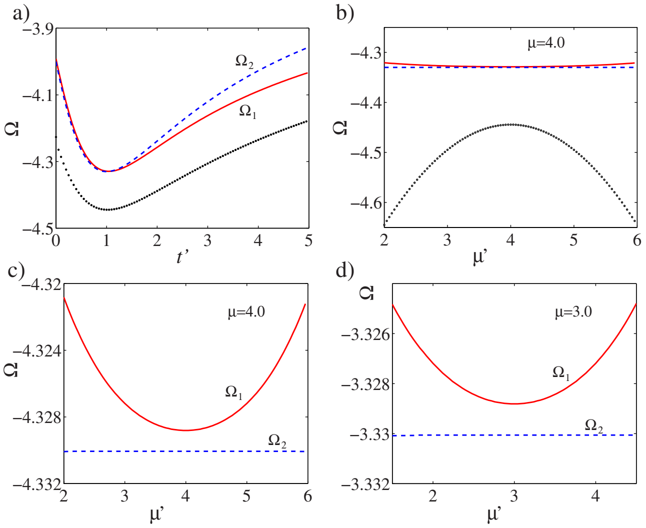

FIG. 7. (Color online) Comparisons of three different functionals for the 2D Hubbard model ($U/t_{\mathrm{lat}}=8$): the original functional from Eq. (12) (black dots), and the two functionals proposed in this work, given respectively by Eq. (B4) (solid red curve) and Eq. (B5) (dashed blue curve). The panels show (a) dependence of all three functionals on the nearest-neighbor hopping parameter t' of the cluster and (b) the same for the dependence on the chemical potential μ' of the cluster. Comparison between the two functionals proposed in this work as functions of the cluster chemical potential μ' for (c) half-filled case, $\mu_{\mathrm{lat}}=4t$, and (d) away from half-filling, $\mu_{\mathrm{lat}}=3t$ (corresponds to 0.4% electron doping).

Figure 7(a) shows the comparison between the original grand potential Ω (dotted line) and the two potentials derived in Appendix B, as functions of the variational cluster hopping parameter t'. All three functionals appear to have minimum at the same value of $t'=t=1$, and the proposed new functionals are convex functions of t' near the solution, as is the original grand potential Ω.

The dependence of the grand potential on the cluster chemical potential μ' is plotted in Fig. 7(b) for the half-filled case. Functional Ω_1, shown by the solid line, clearly is a convex function near the stationary solution $\mu_0'=4t$, unlike the original grand potential Ω (dotted line) that develops a maximum at this point. The details of the behavior of the new functionals near μ_0' appear more clearly in the blown up version of the same figure shown in Fig. 7(c). The functional Ω_2 (dashed line) has vanishing second derivative within machine precision around the stationary solution, and is therefore less suitable for practical calculations compared to Ω_1 close to half-filling. The same conclusion equally holds away from half-filling, as illustrated in Fig. 7(d).

The above illustrations as well as many tests performed for different values of the parameters U, t, and μ, all show that the proposed functional Ω_1 given by Eq. (B4) develops a minimum at the stationary solution with respect to both variational parameters μ' and t'.

Despite these encouraging results, note that the actual values of $\Omega_1(h)$ and $\Omega_2(h)$ at the solution $h=h_0$, while being very close to each other, deviate appreciably from the original grand potential at the same point $\Omega(h_0)$, where by h we denote the manifold of variational parameters of the cluster, i.e., $h=\{t',\mu',\ldots\}$. First of all, this is not surprising in light of the approximation made in Eq. (B2) and the discussion that follows. Indeed, $f(\Sigma(h))$ can only vanish at the *true solution*, where we expect

$$\left.\frac{\delta F[\Sigma]}{\delta \Sigma}\right|_{h=h_0}(\mathbf{k},\omega) + \frac{1}{G_0^{-1}(\mathbf{k},\omega)-\Sigma_{h=h_0}(\mathbf{k},\omega)} = 0. \tag{B6}$$

Instead, we have made an approximation for $f(\Sigma)$ which reads (with $\mathbf{k}$ and ω dependence written explicitly)

$$f(\Sigma_h(\omega)) \approxeq -G_h'(\omega) + \sum_k \frac{1}{G_0^{-1}(\mathbf{k},\omega)-\Sigma_h(\mathbf{k},\omega)}, \tag{B7}$$

where $G_h'(\omega)$ is the cluster Green's function obtained by, e.g., exact diagonalization, and we have dropped the cluster site indices in $G_{\alpha\beta}$ and $\Sigma_{\alpha\beta}$ for simplicity.

The approximation (B7) essentially tells us to consider not all possible variations $\delta\Sigma$ in the calculation of the variational derivative $\delta F[\Sigma]/\delta\Sigma$, but only a *limited subset* of them, namely, such Σ_h that can be parametrized by a set of cluster variational parameters h. Naturally, there is no guarantee that the expected equality (B6) will hold at the approximate solution and, generally speaking, $f(h_0)$ is nonzero. Despite this drawback, the proposed functionals can still be useful in practical calculations, although care has to be taken in cases where a second-order phase transition is expected, as explained in Sec. IV B.

APPENDIX C: CONNECTION TO THE CHITRA-KOTLIAR FUNCTIONAL

The purpose of this appendix is to identify the connection that exists between Potthoff's original functional,[8] the convex functional $\Omega[\Sigma]$ proposed in this work (Sec. IV), and an earlier attempt to construct a convex functional undertaken by Chitra and Kotliar (CK) in Ref. 6. According to the latter study, one can construct an improved version of the Baym-Kadanoff functional $\Omega_{BK}[G]$ given by Eq. (4) as follows:

$$\Omega_{\mathrm{CK}}[G] = \Omega_{BK}[G] - \mathrm{Tr}\ln(1+JG) + \mathrm{Tr}(JG), \tag{C1}$$

where J is the external source field coupled to the electron's Green function

$$-J = \frac{\delta\Omega_{BK}}{\delta G} = G^{-1} - G_0^{-1} + \frac{\delta\Phi}{\delta G}, \tag{C2}$$

where $\Phi[G]$ is the Luttinger-Ward functional. Since $J=0$ when the Dyson equation is satisfied, both functionals are equal to the grand potential at the stationary solution. Chitra and Kotliar have shown[6] that their new functional has a different stability criterion for its stationary solution from that of the Baym-Kadanoff functional; however, they have been unable to prove its convexity at the stationary point. In fact, they have shown explicitly that, in the Hartree approximation, their proposed functional Ω_{CK} is *unstable* for repulsive interactions. This prohibits the use of the Chitra-Kotliar functional as a variational free energy and leaves open the question of convexity.

Is there a way to relate the Chitra-Kotliar functional Eq. (C1) to the functionals of self-energy $\Omega_1[\Sigma]$ and $\Omega_2[\Sigma]$ proposed in Sec. IV? From the previous two equations, it is easy to show that [see Eq. (12) in Ref. 6]

$$\Omega_{\mathrm{CK}}[G] = \Phi[G] - \mathrm{Tr}\left(G\frac{\delta\Phi}{\delta G}\right) - \mathrm{Tr}\ln\left[G_0^{-1} - \frac{\delta\Phi}{\delta G}\right]. \tag{C3}$$

Using the relation (3), $\delta\Phi/\delta G=\Sigma[G]$, and the Legendre transform of the Luttinger-Ward functional[8] as in Eq. (6),

$$F[\Sigma] \equiv \Phi[G[\Sigma]] - \mathrm{Tr}(\Sigma G), \tag{C4}$$

Ω_{CK} in Eq. (C3) becomes

$$\Omega_{\mathrm{CK}}[G(\Sigma)] = F[\Sigma] - \mathrm{Tr}\ln[G_0^{-1} - \Sigma]. \tag{C5}$$

This last expression is nothing else but Potthoff's functional $\Omega[\Sigma]$ as given in Eq. (7). In other words, we see that the Chitra-Kotliar functional of G, if expressed in terms of the self-energy Σ, is equivalent to Potthoff's $\Omega[\Sigma]$. Although both groups begin with the Baym-Kadanoff functional to propose their own functional, they use the Dyson equation in apparently different but, in fact, equivalent ways.

As regards the different functionals of the self-energy introduced in this work (see Sec. IV), their derivation, despite being similar in spirit to that of Chitra and Kotliar,[6] has required additional ingredients that are contained neither in the Potthoff grand potential $\Omega[\Sigma]$ nor in the Chitra-Kotliar functional.

---

*Present address: Department of Physics and Astronomy, Rutgers University, 136 Frelinghuysen Road, Piscataway, NJ 08854, USA; nevidomskyy@cantab.net

[1] R. Dreizler and E. Gross, *Density Functional Theory: An Approach to the Quantum Many-body Problem* (Springer-Verlag, Berlin, 1990).

[2] J. Luttinger and J. Ward, Phys. Rev. **118**, 1417 (1960).

[3] G. Baym and L. Kadanoff, Phys. Rev. **124**, 287 (1961).

[4] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, Rev. Mod. Phys. **78**, 865 (2006).

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

[6] R. Chitra and G. Kotliar, Phys. Rev. B **63**, 115110 (2001).
[7] J. Hubbard, Proc. R. Soc. London, Ser. A **276**, 238 (1963).
[8] M. Potthoff, Eur. Phys. J. B **32**, 429 (2003).
[9] G. Kotliar, S. Y. Savrasov, G. Palsson, and G. Biroli, Phys. Rev. Lett. **87**, 186401 (2001).
[10] M. Potthoff, M. Aichhorn, and C. Dahnken, Phys. Rev. Lett. **91**, 206402 (2003).
[11] G. Kotliar, Eur. Phys. J. B **11**, 27 (1999).
[12] M. Potthoff and M. Aichhorn (private communication).
[13] S. Doniach and E. H. Sondheimer, *Green's Functions for Solid State Physicists* (Benjamin, New York, 1974), Chap. 4.
[14] In Eq. (3), the Luttinger-Ward functional $\Phi[G]$ was taken as a basis for the derivation of the VCA method; however, this choice is not unique and depends on the Hamiltonian at hand. For instance, recently, a generalized Luttinger-Ward functional has been suggested in Ref. 15 to include the possibility of nonlocal Coulomb repulsion terms not accounted for in the Hubbard Hamiltonian (1).
[15] N.-H. Tong, Phys. Rev. B **72**, 115104 (2005).
[16] M. Aichhorn, E. Arrigoni, M. Potthoff, and W. Hanke, Phys. Rev. B **74**, 024508 (2006).
[17] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C++: The Art of Scientific Computing* (Cambridge University Press, Cambridge, England, 2002), Chap. 10.
[18] P. Culot, G. Dive, V. Nguyen, and J. Ghuysen, Theor. Chim. Acta **82**, 189 (1992).
[19] N.-H. Tong and F.-C. Pu, Phys. Rev. B **62**, 9425 (2000); N.-H. Tong, S.-Q. Shen, and R. Bulla, *ibid.* **70**, 085118 (2004).