The Variational Cluster Approximation for Hubbard Models : Practical Implementation

David Sénéchal
Département de physique, Université de Sherbrooke and RQMP
Sherbrooke (Québec) Canada, J1K 2R1
david.senechal@usherbrooke.ca

Abstract

The Variational cluster approximation (VCA) is an approximate method for finding properties the Hubbard model, the basic model used in the study of strongly correlated electron systems. Such systems include, for instance, high-temperature and organic superconductors. The method involves solving exactly the Hubbard model on a small cluster of lattice sites, after adding fields that represent the effect of the cluster's environment. We present an outline of the method and describe its implementation. In particular, we show how to efficiently carry frequency-wavevector integrals that are central to the method.

1. Context

The starting point in the microscopic study of a material is typically an *electronic structure calculation*, i.e., a determination of the different electronic energy bands of the material. This is done in the framework of Density Functional Theory (DFT, for a review, see [5]), and many commercial or open source programs are available for this task. However, the very notion of band structure is approximate. It merely provides a basis of quantum states in terms of which the electrons can be optimally expressed as independent (i.e., effectively non interacting) objects. However, in many materials – called *strongly correlated* – residual electron-electron interactions are too strong for the independent electron picture to be satisfactory. The best known examples are high-temperature superconductors, whose parent compounds are predicted to be metallic by band-structure calculations, whereas they are actually antiferromagnetic insulators. A lot of effort has been devoted to understand strongly correlated systems in the condensed matter physics community in the last twenty years, and novel numerical approaches have been developed. One of these approaches, developed in the last five years, is the *variational cluster approximation* (VCA)[9, 3]. The purpose of this work is to outline the numerical implementation of this method, and to point out tricks that make it efficient.

The prototype of physical models used to study strongly correlated systems is the so-called *one-band Hubbard model*, defined in the language of second quantization by the following quantum mechanical Hamiltonian:

$$H = \sum_{\mathbf{r},\mathbf{r}',\sigma} t_{\mathbf{r},\mathbf{r}'} c^\dagger_{\mathbf{r}\sigma} c_{\mathbf{r}'\sigma} + U \sum_{\mathbf{r}} n_{\mathbf{r}\uparrow} n_{\mathbf{r}\downarrow} - \mu \sum_{\mathbf{r},\sigma} n_{\mathbf{r},\sigma} \quad (1)$$

In this expression, $\mathbf{r}$ and $\mathbf{r}'$ denote lattice sites and σ the two possible spin projections of each electron ($\uparrow$ or $\downarrow$). The operator $c_{\mathbf{r}\sigma}$ annihilates an electron of spin σ in a Wannier orbital centered at site $\mathbf{r}$, whereas its Hermitian conjugate $c^\dagger_{\mathbf{r}\sigma}$ creates an electron in the same state. The operator $n_{\mathbf{r}\sigma} = c^\dagger_{\mathbf{r}\sigma} c_{\mathbf{r}\sigma}$ gives the number of electrons of spin σ at site $\mathbf{r}$. $t_{\mathbf{r},\mathbf{r}'}$ is the quantum mechanical amplitude for an electron to hop from site $\mathbf{r}$ to site $\mathbf{r}'$ and is in principle obtained from electronic structure calculations. In addition, U is the electrostatic energy cost for having two electrons of opposite spins occupy the same Wannier orbital. The chemical potential μ controls the electron density.

This model, although very simple to formulate, is impossible to solve analytically except in a few special cases and has been the object of intense effort in the theoretical solid state physics community in the last 20 years. One of the basic questions is whether this simple model can explain common features of high-temperature superconductors, such as the proximity of antiferromagnetism and d-wave superconductivity and the range of these phases as a function of electron concentration, the pseudogap phenomenon, etc. In recent years, the answer to that question has been converging towards a *yes*, in great part thanks to the developement of novel approaches to the Hubbard model, falling under the generic name of *quantum cluster models*, among which the VCA studied in this paper[11, 1].

In order to grasp the difficulty of solving numerically the Hubbard model – e.g. of calculating the lowest eigenstate of the Hamiltonian H, let us consider the size of the quantum-mechanical Hilbert space. On a system containing L sites

1550-5243/08 $25.00 © 2008 IEEE
DOI 10.1109/HPCS.2008.18

and N_σ electrons of spin σ, that dimension is

$$d_H = \frac{L!}{N_\uparrow!(L-N_\uparrow)!}\frac{L!}{N_\downarrow!(L-N_\downarrow)!} \tag{2}$$

For a half-filled, zero spin system ($N_\uparrow = N_\downarrow = L/2$), this translates into $d_H = (L!/(L/2)!^2)^2$, which behaves like $L4^L$ for large L. In other words, the size of the eigenproblem grows exponentially with system size. By contrast, the non-interacting problem can be solved only by concentrating on one-electron states. For this reason, exact diagonalization of the Hamiltonian is restricted to systems of the order of 16 sites or less. But physically, such important characteristics as antiferromagnetism and superconductivity are the result of a spontaneous symmetry breaking that can only occur in the thermodynamic limit $L \to \infty$, and whose signature is very difficult to see in small systems. Quantum cluster methods [6] provide an approximate solution to the Hubbard model by solving exactly, on a small cluster of sites, a modified Hamiltonian containing additional terms that mimick the presence of the infinite lattice around the cluster. These approaches are based on rigorous variational principles or contain some from of self-consistency that fix what the correct additional terms should be.

2. The Variational Cluster Approximation

2.1. Formal aspects of the method

Let us express the Hubbard Hamiltonian more generally as

$$H = \sum_{\alpha\beta} c_\alpha^\dagger t_{\alpha\beta} c_\beta + \sum_{\alpha\neq\beta} U_{\alpha\beta} n_\alpha n_\beta \tag{3}$$

where α and β are composite indices denoting site ($\mathbf{r}$), spin (σ) and possibly band indices, and where $n_\alpha = c_\alpha^\dagger c_\alpha$ is the number of electrons in the one-particle state α.

A key component in the solution of a many-body Hamiltonian is the one-particle Green function $G_{\alpha\beta}(z)$, which can be defined as follows at zero temperature:

$$\begin{aligned} G_{\alpha\beta}(z) = \langle\Omega| c_\alpha \frac{1}{z - H + E_0} c_\beta^\dagger |\Omega\rangle \\ + \langle\Omega| c_\beta^\dagger \frac{1}{z + H - E_0} c_\alpha |\Omega\rangle \end{aligned} \tag{4}$$

where $|\Omega\rangle$ is the quantum mechanical ground state of the model, with energy E_0. The Green function, if known, gives access to all the one-particle properties of the system, such as the spectral function or the average of any one-particle operator. In the absence of interaction ($U = 0$), the Green function takes the simple (matrix) form $\mathsf{G}(z) \equiv \mathsf{G}_0(z) = (z - \mathsf{t})^{-1}$. In the presence of interaction, it takes rather the general form $\mathsf{G}(z)^{-1} = z - \mathsf{t} - \mathsf{\Sigma}(z)$, where $\mathsf{\Sigma}(z)$ is called the *self-energy* (the above constitutes in fact a definition of $\mathsf{\Sigma}$).

Potthoff has shown [7] that a many-body Hamiltonian $H = T + V$ containing a one-body part T (quadratic in the creation and annihilation operators) and a two-body part V, can be associated a functional $\Omega_{T,V}[\mathsf{\Sigma}]$ of the form

$$\Omega_{T,V}[\mathsf{\Sigma}] = F_V[\mathsf{\Sigma}] - \mathrm{Tr}\,\ln(\mathsf{G}_0^{-1} - \mathsf{\Sigma}) \tag{5}$$

where F_V is a universal functional, i.e., a functional that depends only on the form of the interaction V and not on T. The symbol 'Tr' stands for a functional trace, that translates into an sum over frequencies and sites indices (or, alternately, wavevectors). The functional $\Omega_{T,V}[\mathsf{\Sigma}]$ is stationary at the physical self-energy of the system, meaning that if $\mathsf{\Sigma}$ is the self-energy given by Dyson's equation $\mathsf{\Sigma} = \mathsf{G}_0^{-1} - \mathsf{G}^{-1}$, then any variation $\delta\mathsf{\Sigma}$ of the self-energy from that physical value leaves $\Omega_{T,V}[\mathsf{\Sigma}]$ unchanged at first order in $\delta\mathsf{\Sigma}$. Moreover, the value of $\Omega_{T,V}[\mathsf{\Sigma}]$ at that point is the physical grand potential of the system. The idea is to find an approximate form of $\mathsf{\Sigma}$ by means of this variational principle, and to derive from this an approximate Green function G of the Hubbard model.

The approximation strategy behind the VCA is to evaluate the functional $\Omega_{T,V}[\mathsf{\Sigma}]$ exactly, albeit in a restricted variational space[8]. The variational space is restricted to the physical self-energies of a family of *reference systems* described by a Hamiltonian $H' = T' + V$ which differs from H only by its one-body part, and which can be solved numerically. The universality of $F_V[\mathsf{\Sigma}]$ may be expressed as

$$\begin{aligned} \Omega_{T,V}[\mathsf{\Sigma}] + \mathrm{Tr}\,\ln(\mathsf{G}_0^{-1} - \mathsf{\Sigma}) = \\ \Omega_{T',V}[\mathsf{\Sigma}] + \mathrm{Tr}\,\ln(\mathsf{G}_0'^{-1} - \mathsf{\Sigma}) \end{aligned} \tag{6}$$

If $\mathsf{\Sigma}$ is the physical self-energy of H', then $\Omega_{T',V}[\mathsf{\Sigma}] = \Omega'$, $\mathsf{\Sigma} = \mathsf{G}_0'^{-1} - \mathsf{G}'^{-1}$ and thus

$$\Omega_{T,V}[\mathsf{\Sigma}] = \Omega' - \mathrm{Tr}\,\ln\left[1 + (\mathsf{G}_0^{-1} - \mathsf{G}_0'^{-1})\mathsf{G}'\right] \tag{7}$$

In VCA, H' is chosen to represent a finite cluster of sites: The infinite lattice γ (e.g. the square lattice) is now regarded as a superlattice Γ of identical clusters (Fig. 1). Since the interaction U is local, H' differs from H only by the one-body part of the Hamiltonian: The new hopping matrix t' does not link sites belonging to different clusters. In addition, t' may contain new terms, called *Weiss fields*, that are local to each cluster and represent the effect of the rest of the lattice on each cluster, in some way compensating for the inter-cluster hopping terms that have been suppressed. The form of these Weiss fields is set by physical insight and their value determined by the variational principle. The crucial point, of course, is that H' can now be solved numerically because the clusters are small and independent of each other.

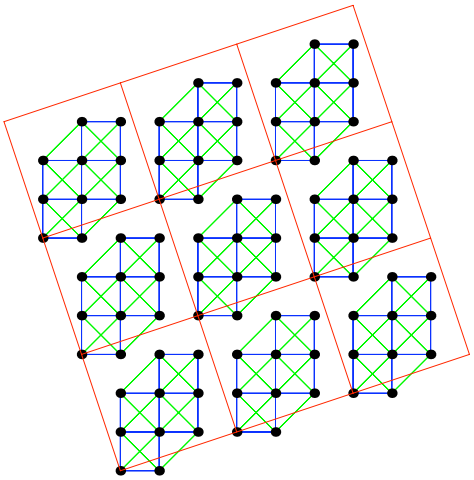

Figure 1. Tiling the square lattice with 10-site clusters (hopping terms in blue and green).

The hopping matrix $\mathtt{t}$ can be represented in a number of bases. Because of the translation invariance of the original lattice, it is diagonal in a basis of plane waves:

$$t_{\mathbf{r},\mathbf{r}'} = V_\gamma \int_{\mathrm{BZ}} \mathrm{e}^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \tilde{t}(\mathbf{k}) \tag{8}$$

where the integral is carried over the Brillouin zone of the reciprocal lattice and V_γ represents the volume of the unit cell. $\tilde{t}(\mathbf{k})$ is the dispersion relation of electrons roaming across the lattice in the absence of interaction U. For instance, in the case of a one-dimensional system with nearest-neighbor hopping $-t$, we find $\tilde{t}(k) = -2t\cos(k) - \mu$. Alternatively, $\mathtt{t}$ can be expressed in a basis of plane waves propagating in the superlattice only, with wavevectors restricted to the *reduced Brillouin zone* (RBZ), i.e., the Brillouin zone of the superlattice. In this case, the index $\mathbf{r}$ is limited to the cluster only and, if $\tilde{\mathbf{r}}$ denotes the base position of a cluster (a site of the superlattice), one may write

$$t_{\mathbf{r}+\tilde{\mathbf{r}},\mathbf{r}'+\tilde{\mathbf{r}}'} = V_\Gamma \int_{\mathrm{RBZ}} \mathrm{e}^{i\mathbf{k}\cdot(\tilde{\mathbf{r}}-\tilde{\mathbf{r}}')} t_{\mathbf{r},\mathbf{r}'}(\mathbf{k}) \tag{9}$$

and the same hopping matrix as above would now be expressed, on a two-site cluster, as

$$\mathtt{t}(k) = \begin{pmatrix} -\mu & -t(1+\mathrm{e}^{ik}) \\ -t(1+\mathrm{e}^{-ik}) & -\mu \end{pmatrix} \tag{10}$$

In this representation, a more explicit expression for the functional (7) is

$$\Omega(\mathtt{t}') = \Omega' - \int_0^\infty \frac{\mathrm{d}\omega}{\pi} \int_{\mathrm{RBZ}} \ln \left| \det \left[1 + (\mathtt{t}(\mathbf{k}) - \mathtt{t}')\mathsf{G}'(i\omega) \right] \right| \tag{11}$$

where (i) Ω' is the exact ground state energy of the cluster Hamiltonian H' (chemical potential included); (ii) $\mathsf{G}'(z)$ the associated Green function; (iii) $\mathtt{t}(\mathbf{k})$ is the mixed representation of the hopping matrix $\mathtt{t}$ of the original Hamiltonian; (iv) $\mathtt{t}'$ is the hopping matrix of the cluster Hamiltonian H'; (v) The summation is taken over wavevectors in the reduced Brillouin zone and (vi) The integral is carried over the positive imaginary frequency axis (at zero temperature). $\mathtt{t}$ and G are now matrices of order L (the number of sites), or $2L$ if the two spin species are not conserved independently.

In practice, just a few variational parameters are used (typically less than 4). For instance, a reference Hamiltonian containing Weiss fields allowing for Néel antiferromagnetism and superconductivity (denoted respectively M and $\Delta_{\mathbf{r},\mathbf{r}'}$) would be

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

where $\mathbf{Q} = (\pi, \pi)$ is the antiferromagnetic wavevector and, for instance,

$$\Delta_{\mathbf{r},\mathbf{r}'} = \begin{cases} \Delta & \text{if} \quad \mathbf{r}' - \mathbf{r} = \hat{\mathbf{x}} \\ -\Delta & \text{if} \quad \mathbf{r}' - \mathbf{r} = \hat{\mathbf{y}} \end{cases} \tag{12}$$

for $d_{x^2-y^2}$ superconductivity. The functional (11) then becomes a function $\Omega(M, \Delta, \mu')$ whose stationary points must be found. These correspond to approximate physical solutions and the corresponding self energies $\Sigma = \mathsf{G}_0'^{-1} - \mathsf{G}'^{-1}$ are then used to construct the approximate Green function of the original Hubbard model via the relation

$$\begin{aligned} \mathsf{G}^{-1}(\mathbf{k}, z) &= \mathsf{G}_0^{-1}(\mathbf{k}, z) - \Sigma(z) \\ &= \mathsf{G}'^{-1}(z) - \mathsf{V}(\mathbf{k}) \end{aligned} \tag{13}$$

where $\mathsf{V}(\mathbf{k}) = \mathsf{G}_0'^{-1} - \mathsf{G}_0^{-1}(\mathbf{k})$ is the frequency-independent difference between the hopping matrix in the lattice model (H) and that of the cluster model (H').

Note that anomalous terms of the form $c_{\mathbf{r}\uparrow} c_{\mathbf{r}'\downarrow}$ do not take the form (3) at first sight. But they do if we use the so-called Nambu representation:

$$c_\alpha = (c_{\mathbf{r}_1,\uparrow}, \ldots, c_{\mathbf{r}_L,\uparrow}, c^\dagger_{\mathbf{r}_1,\downarrow}, \ldots, c^\dagger_{\mathbf{r}_L,\downarrow}) \tag{14}$$

where $\alpha = 1, \ldots, 2L$. It is however impossible to express in the form (3) a Hamiltonian that contains both anomalous terms and terms that do not conserve spin (spin-flip terms).

2.2. Implementation layout

Implementing the VCA requires (i) calculating the function (11) for any given set of variational parameters and (ii) finding the stationary points of that function. The latter must be done keeping in mind that (i) function evaluation

is costly and (ii) we are looking not for minima only, but for saddle points in general. Let us collectively denote the variational parameters as x. Calculating $\Omega(\mathsf{x})$ involves the following components:
(i) Calculating the ground state $|\Omega\rangle$ of Hamiltonian H'.
(ii) Calculating the Green function G' of the cluster.
(iii) Performing the frequency and wavevector sums (11).
Except for very small clusters ($L \leq 6$ sites), the exact diagonalization (steps (i) and (ii)) is dominating computation time and memory if the frequency-wavevector integral is done efficiently (see Sect. 4.2). The rest of the paper will be devoted to outlining the exact diagonalization procedure, the frequency-wavevector sum and the saddle-point search.

3. Exact diagonalization of the cluster

3.1. Coding of the states

The exact diagonalization proceeds by constructing an explicit matrix representation of the Hamiltonian H' in the quantum Hilbert space. The first step is to define a coding scheme for the basis states of that space. We use the occupation number basis, in which each state is specified by the occupation number n_α ($= 0$ or 1) ($\alpha = 1, \ldots, 2L$, since we have L sites and two spin projections). Thus the string of n_α's forms the binary representation of a 32-bit unsigned integer (if $L \leq 16$) and represents the state

$$(c_1^\dagger)^{n_1} \cdots (c_{2L}^\dagger)^{n_{2L}} |0\rangle \tag{15}$$

where the order in which the creation operators are applied on the vacuum state $|0\rangle$ is a matter of convention, but important.

Not all the states (15) belong to the relevant Hilbert space, however, because of particle number conservation. Specifically, in the absence of superconductivity or spin-flip terms, the total number of electrons of each spin ($N_\uparrow$ and $N_\downarrow$) is conserved by the cluster Hamiltonian H' and therefore the ground state will belong to the tensor product $V_{N_\uparrow} \otimes V_{N_\downarrow}$, where V_n is the Hilbert space of n electrons of a given spin. In the presence of a singlet superconducting Weiss field, the total particle number is not conserved, but the total spin is, so that the ground state lives in the direct sum

$$\sum_{n=1}^{L} V_n \otimes V_n \ . \tag{16}$$

Other Weiss fields are possible that conserve the total number of particles, without conserving the spin projection, and in that case another combination of tensor products is used. In any case, for adressing purposes, the necessary basis states must be labelled by consecutive integers and a look-up table between these labels and the binary representations of basis states is needed. This is manageable from the memory viewpoint because of the tensor product structure (16).

Once the coding scheme is set up, the Hamiltonian matrix needs to be built. That matrix is sparse, and so can be stored in memory. Many terms in the Hamiltonian (like the hopping terms) have a tensor product form and so just storing the tensor factors is needed and the memory required is modest.

3.2. Finding the ground state

The ground state $|\Omega\rangle$ is found by the Lanczos algorithm. This involves selecting a (random) initial state $|\phi_0\rangle$ and constructing a projection $\mathscr{H}$ of the Hamiltonian matrix H' in the Krylov subspace spanned by the vectors

$$\{|\phi_0\rangle, H'|\phi_0\rangle, H'^2|\phi_0\rangle, \ldots, H'^R|\phi_0\rangle\} \tag{17}$$

where R is the number of iterations that is controlled by the convergence of the lowest eigenvalue of $\mathscr{H}$. Convergence is estimated by looking at the relative changes in the ground state energy at each iteration and stopping at, say, 10^{-12}. A description of the application of this algorithm to strongly correlated systems can be found in Refs. [2, 10]. The algorithm converges rapidly (R between 100 and 200 for a 12-site cluster) whenever the ground state is well separated from the first excited state.

3.3. Computing the Green function

Once the ground state is found, the Green function $\mathsf{G}(\omega)$ can be calculated from its definition (4). The matrix $\mathsf{G}(\omega)$ is of order L_G, with $L_G = L$ (the number of cluster sites) if the model conserves spin, and $L_G = 2L$ if spin-flip or anomalous terms are present in the Hamiltonian.

There are two ways to accomplish this: The first (I) is to apply a Lanczos procedure on the excited state $|\phi_0\rangle = (c_\alpha^\dagger + c_\beta^\dagger)|\Omega\rangle$ for each pair (α, β) of indices, as described in [10]. One then obtains a numerical representation of $G_{\alpha\beta}$ in the form of a truncated continued fraction.

Another way (II) is to apply the *band Lanczos* procedure[4]: The states $|\phi_\alpha\rangle = c_\alpha^\dagger|\Omega\rangle$ are first constructed, and then one builds the projection $\mathscr{H}$ of H' Krylov subspace spanned by

$$\begin{aligned}\{|\phi_1\rangle, \ldots, |\phi_{L_G}\rangle, & H'|\phi_1\rangle, \ldots, H'|\phi_{L_G}\rangle, \ldots, \\ & (H')^{\mathscr{M}}|\phi_1\rangle, \ldots, (H')^{\mathscr{M}}|\phi_{L_G}\rangle\}\end{aligned} \tag{18}$$

Again one may control the number of iterations $\mathscr{M}$ by the convergence of the lowest eigenvalue of $\mathscr{H}$, and one must be careful about potential redundant basis vectors in the Krylov subspace, that must be properly deflated[4]. The number of states R in the Krylov subspace at convergence is typically between 100 and 300, depending on system size.

The $R \times R$ matrix $\mathscr{H}$, which has a tridiagonal structure in the ordinary Lanczos method, now has a band structure made of $2L_G$ diagonals around the central diagonal. One must then calculate the projection $Q_{\alpha r}$ of $|\phi_\alpha\rangle$ on the eigenstates of $\mathscr{H}$. The Green function can then be expressed in a Lehman representation:

$$G'_{\alpha\beta}(\omega) = \sum_{r=1}^{R} \frac{Q_{\alpha r} Q^*_{\beta r}}{\omega - \omega'_r} = \left(\mathsf{Q} \frac{1}{\omega - \mathsf{\Lambda}} \mathsf{Q}^\dagger \right)_{\alpha\beta} \qquad (19)$$

where the poles ω_r are a combination of the eigenvalues of $\mathscr{H}$ and of the ground state energy E_0; these poles constitute the diagonal matrix $\mathsf{\Lambda}$. The band Lanczos method (II) requires more memory than the usual Lanczos method (I), but is faster since all pairs (α, β) are covered in a single procedure. Another advantage is that it provides a Lehman representation of the Green function.

3.4. Parallelization

The exact diagonalization of H' can be easily parallelized since the Lanczos and band Lanczos methods rely on matrix-vector multiplications. Parallelization is necessary, in particular if the band Lanczos is to be used, for clusters of more than 14 sites. It is simplest in this case to store the (sparse) Hamiltonian matrix in memory, since the memory needed to store the band Lanczos vectors is typically larger or at least of the same order as the memory needed to store the Hamiltonian matrix. A parallel implementation accross $\mathscr{N}$ nodes proceeds by splitting all vectors into $\mathscr{N}$ equal segments, and the Hamiltonian matrix into $\mathscr{N}^2$ blocks (labelled H'_{IJ}, with $I, J = 1, \ldots, \mathscr{N}$. A single matrix-vector multiplication $|y\rangle = H'|x\rangle$ then proceeds, for each process, by $\mathscr{N}$ successive operations $|y\rangle_I = H_{IJ}|x\rangle_J$, J labelling the different processes and I the successive operations. After each operation the resulting vectors $|y\rangle_I$ must be sent to process I to be summed in a single segment.

Even though this is readily implemented, the code scales very poorly and must be used with as few processes possible to fit in memory.

4. Computing frequency-wavevector sums

Once the cluster Green function is known, calculating the functional (11) requires an integral over frequencies and wavevectors of an expression that requires a few linear-algebraic operations to evaluate. Two different methods have been used to compute these sums, described in what follows. One of the goals of this paper is to show that the second method, entirely numerical, is much faster than the first one, which is partly analytic, a result that may seem paradoxical.

4.1. Method I : Exact frequency integration

The integral over frequencies in (11) may be done analytically, with the result [1]

$$\Omega(h) = \Omega'(h) - \sum_{\omega'_r < 0} \omega'_r + \int_{\mathrm{RBZ}} \sum_{\omega_r(\mathbf{k}) < 0} \omega_r(\mathbf{k}) \qquad (20)$$

where the ω'_r are the poles of the Green function G' in the Lehmann representation (19) and the $\omega_r(\mathbf{k})$ are the poles of the VCA Green function $(\mathsf{G}_0^{-1}(\mathbf{k}) - \mathsf{\Sigma})^{-1}$. The latter may be expressed as [1]

$$\mathsf{Q} \frac{1}{\omega - \mathsf{L}(\mathbf{k})} \mathsf{Q}^\dagger \qquad (21)$$

where $\mathsf{L}(\mathbf{k}) = \mathsf{\Lambda} + \mathsf{Q}^\dagger \mathsf{V}(\mathbf{k}) \mathsf{Q}$. The poles $\omega_r(\mathbf{k})$ are nothing but the eigenvalues of the $R \times R$ matrix $\mathsf{L}(\mathbf{k})$.

In practice, the first sum in (20) is readily calculated. The second sum demands an integration over wavevectors. For each wavevector $\mathbf{k}$, one must calculate $\mathsf{L}(\mathbf{k})$ and find its eigenvalues, a process of order R^3. Other linear-algebraic manipulations leading to the diagonalization of $\mathsf{L}(\mathbf{k})$ are typically less time-consuming than the diagonalization itself. The computation time therefore goes like $N_k R^3$, where N_k is the number of points in a mesh covering the reduced Brillouin zone (in fact half of the reduced Brillouin zone, since inversion symmetry is assumed).

4.2. Method II : Numerical frequency integration

An alternate method of computing the integrals in (11) is to perform them in the reverse order, i.e., to first compute the wavevector sum for a fixed frequency ω, and then integrating over frequencies numerically. The method used to sum over wavevectors is exactly the same as in Method I above : a wavevector mesh is set up in the reduced Brillouin zone. This mesh is either a fixed, regular grid, or an adaptive mesh that is refined recursively as needed by comparing a two- and three-points Gauss-Legendre evaluations within each cell (more accurately, the number of function evaluations in each cell is 2^d and 3^d, d being the dimension of space).

The frequency integral is done by dividing the positive imaginary frequency axis in three segments: $[0, \Lambda_1]$, $[\Lambda_1, \Lambda_2]$ and $[\Lambda_2, \infty)$. The constant Λ_1 is a low-energy scale in the problem, like the lowest eigenvalue ω'_r (up to some minimum), whereas Λ_2 is a high-energy scale, like the largest eigenvalue ω'_r (up to some maximum). On each segment a 20-point Gauss-Legendre integration is used. The last segment is in fact treated like an integral over $u = 1/\omega$ from 0 to Λ_2^{-1}. The frequency integral $\int \mathrm{d}\omega\, f(\omega)$ thus takes the form of a weighted sum $\sum_n p_n f(\omega_n)$, where $f(\omega)$ is the wavevector sum conducted at frequency ω. The required

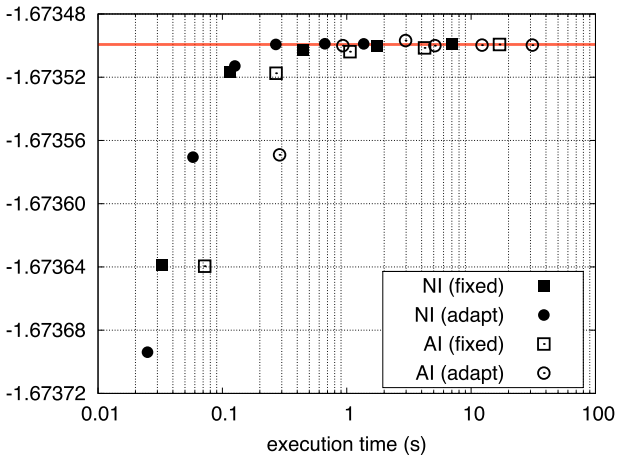

Figure 2. Comparisons of the different integration methods in execution time and accuracy for a 4-site cluster. See text for details.

accuracy of the wavevector sum, which sets the number of points of the wavevector mesh, is conditioned by the weight p_n (i.e. it does not have to be so large when p_n is small). In addition, the integrand of (11) may have sharp structures – thus requiring a fine wavevector mesh – at frequencies close to the real axis (the poles of G are all real) but is increasingly smooth as one moves away from the real axis. Thus the wavevector mesh may be redefined from time to time as one progresses along the imaginary frequency axis, making the mesh coarser and coarser without cost in accuracy.

Tests have been conducted in order to compare the speed and accuracies of the two methods: the analytic frequency integration (AI) described in Sect. 4.1, and the numerical frequency integration (NI) described in this section. In each case, a fixed mesh and an adaptive mesh have been used. The results are displayed in Figs. 2 and 3.

Fig. 2 shows the relation between execution time (in seconds) on a 2.16 MHz Intel Core 2 Duo processor and the value of the functional obtained with four different methods: (i) Numerical integration (NI) with a fixed wavevector grid, (ii) NI with an adaptive wavevector grid, (iii) Analytic integration (AI) with a fixed wavevector grid and (iv) AI with an adaptive wavevector grid. For each method, different points correspond to different required accuracies of the wavevector integral. The system used in that case was the two-dimensional Hubbard model with nearest-neighbor hopping $t = 1$, Coulomb repulsion $U = 8$ and chemical potential $\mu = 1$, on a 4-site cluster. The time to perform the exact diagonalization is negligible here. The value of the chemical potential used makes the system metallic, i.e., there are poles of the VCA Green function at zero frequencies, which is more demanding on the integral because of sharp features of the integrand. Deep within ordered phases there are generally gaps in the physical spectrum that make the integrals converge faster, but the location of phase boundaries, where these gaps disappear, is generally of great interest. One sees from Fig. 2 that the NI method with adaptive mesh converges fastest, about 3 times faster than the AI method. The two horizontal red lines represent the two converged values (for NI and AI). They differ by less than 10^{-6}, a difference due to the finite mesh used in the frequency integral. This accuracy is more than adequate for applications.

Fig. 3 shows the same, this time for a 12-site cluster. The exact diagonalization is done beforehand, and is the same for all integration methods shown. In that case the number R of poles was 312, instead of $R = 32$ for the 4-site cluster used in Fig. 2. Here the gain in using NI is greater even, as the method is over 500 times faster than AI with a adaptive mesh. Indeed, it is prohibitively expensive to use the AI method for anything but very small clusters. Method I (AI) is of order $N_k R^3$, whereas method II (NI) involves linear-algebraic operations on the Green function, of order $N_k N_\omega L_G^3$ (say, for a fixed wavevector grid and N_ω frequencies). The ratio of execution times between the two method should roughly be

$$\frac{\text{time(AI)}}{\text{time(NI)}} \sim \left(\frac{R}{L_G}\right)^3 \frac{1}{N_\omega} \tag{22}$$

In the case of Fig. 2, $R = 32$, $L_G = 4$ and $N_\omega = 60$, which gives a ratio of ~ 8. In the case of Fig. 3, $R = 312$, $L_G = 12$ and $N_\omega = 60$, which gives a ratio of ~ 290. In both cases this overestimates by a factor 2 to 3 the measured advantage of NI, which is nevertheless enormous.

5. Searching for stationary points

Let x_i be the n different variational parameters used in VCA. Once the function $\Omega(\mathsf{x})$ may be efficiently calculated,

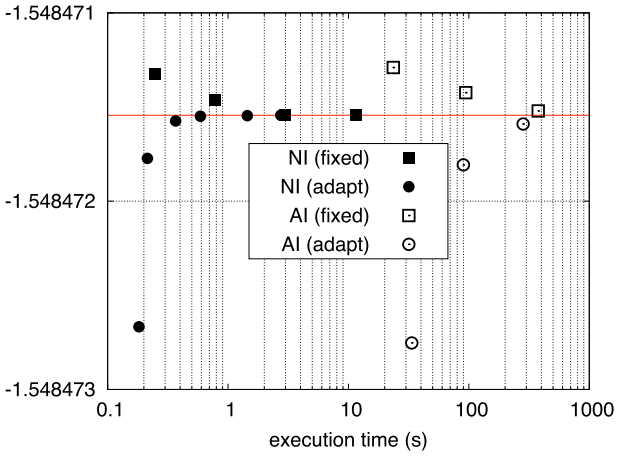

Figure 3. Same as Fig. 2 for a 12-site cluster.

it remains to find a stationary point of that function. This point is not necessarily a minimum in all directions. Indeed, experience has shown that Ω is a maximum as a function of the cluster chemical potential μ', while it is generally a minimum as a function of symmetry-breaking Weiss fields like M or Δ.

5.1. Method I : The Newton-Raphson algorithm

The Newton-Raphson algorithm allows one to find stationary points with a small number of function evaluations. One starts with a trial point x_0 and an initial step h. Let e_i denote the unit vector in the direction of axis i of the variational space. The function Ω is then calculated at as many points as necessary to fit a quadractic form in the neighborhood of x_0. This requires $(n+1)(n+2)/2$ evaluations, at points like x_0, $\mathsf{x}_0 \pm h\mathsf{e}_i$, and a few of $\mathsf{x}_0 + h(\mathsf{e}_i + \mathsf{e}_j)$. The stationary point x_1 of that quadratic form is then used as a new starting point, the step h is reduced to a fraction of the difference $|\mathsf{x}_1 - \mathsf{x}_0|$, and the process is iterated until convergence on $|\mathsf{x}_i - \mathsf{x}_{i-1}|$ is achieved. A variant of this method, the quasi-Newton algorithm, may also be used, in which the full Hessian matrix of second derivatives is not calculated. It requires in general more iterations, but fewer function evaluations at each step.

The advantage of the Newton-Raphson method lies in its economy of function evaluations. Its disadvantage is a lack of robustness. One has to be relatively close to the solution in order to converge to it. But one typically runs parametric studies in which an external (i.e. non variational) parameter of the model is varied, such as the chemical potential μ or the interaction strength U. In this context, the solution associated with the current value of the external parameter may be used as the starting point for the next value, and in this fashion, by proximity, one may conduct rather robust calculations.

5.2. Method II : The conjugate-gradient algorithm

A more robust method, albeit more time consuming, is the conjugate-gradient algorithm, which we will not explain here as it is amply documented and fairly common. However, this algorithm finds minima (or maxima), not saddle points in general. We must therefore take the extrinsic step of identifying parameters (like μ' above) that are expected to drive maxima of Ω, and a complementary set of parameters (like M and Δ above) that drive minima of Ω. One then, iteratively, finds maxima and minima with the two sets of parameters in succession, and stops when convergence on $|\mathsf{x}_i - \mathsf{x}_{i-1}|$ has been achieved. This method is suitable to find a first solution when the Newton-Raphson method fails to deliver one. It may however converge to minima that are in fact singularities of Ω, i.e., points where the derivatives are not defined. Such points may occur as the result of energy-level crossings in clusters and are an artifact of the finite-cluster size.

6. Conclusion

We have outlined the numerical implentation of the Variational Cluster Approximation, a quantum cluster method well suited to study the ordered phases of the Hubbard model, used in the field of strongly correlated electrons. The implementation combines various tools such as: the Lanczos and band Lanczos algorithms for very large matrices, adaptive mesh for wavevector integration and finally Optimization tools like the Newton-Raphson and the conjugate gradient methods. In particular, we have shown that the fastest method to perform the frequency-wavevector integral (11) is to first sum over wavevectors and then over frequencies, despite the possibility of using an analytic form for the frequency integral when the latter is performed first.

References

[1] M. Aichhorn, E. Arrigoni, M. Potthoff, and W. Hanke. Variational cluster approach to the hubbard model: Phase-separation tendency and finite-size effects. *Phys. Rev. B*, 74(23):235117, 2006.

[2] E. Dagotto. Correlated electrons in high-temperature superconductors. *Rev. Mod. Phys.*, 66:763, 1994.

[3] C. Dahnken, M. Aichhorn, W. Hanke, E. Arrigoni, and M. Potthoff. Variational cluster approach to spontaneous symmetry breaking: The itinerant antiferromagnet in two dimensions. *Phys. Rev. B*, 70:245110, 2004.

[4] R. Freund. *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, chapter 4.6. SIAM, 2000.

[5] W. Kohn. Nobel lecture: Electronic structure of matter - wave functions and density functionals. *Rev. Mod. Phys.*, 71:1253, 1999.

[6] T. Maier, M. Jarrell, T. Pruschke, and M. H. Hettler. Quantum cluster theories. *Rev. Mod. Phys.*, 77:1027, 2005.

[7] M. Potthoff. Self-energy-functional approach to systems of correlated electrons. *Eur. Phys. J. B (France)*, 32(4):429 – 36, 2003.

[8] M. Potthoff. Dynamical variational principles for strongly correlated electron systems. *Adv. Solid State Phys.*, 45:135, 2005.

[9] M. Potthoff, M. Aichhorn, and C. Dahnken. Variational cluster approach to correlated electron systems in low dimensions. *Phys. Rev. Lett.*, 91:206402, 2003.

[10] D. Sénéchal. A cluster method for spectral properties of correlated electrons. In D. Sénéchal, editor, *High Performance Computing Systems and Applications and OSCAR symposium*, 2003.

[11] D. Sénéchal, P.-L. Lavertu, M.-A. Marois, and A.-M. S. Tremblay. Competition between antiferromagnetism and superconductivity in high-T_c cuprates. *Phys. Rev. Lett.*, 94:156404, 2005.