Green's Function Monte Carlo for Lattice Fermions

E. Manousakis
Florida State University, Tallahassee, FL 32306-3016

January 12, 2012

Outline

Introduction

We are going to discuss the application of the Green's function Monte Carlo (GFMC) Method for lattice Fermions.

- The GFMC method has been successfully applied to lattice bosons [Trivedi and Ceperley, 1989, Trivedi and Ceperley, 1990, Carlson, 1989, Runge, 1992]. The application is straightforward as it does not suffer from the fermion minus-sign problem.
- The GFMC method has been extended to the study of lattice fermions for the case of the few holes in the t-J model [Boninsegni and Manousakis, 1992, Boninsegni and Manousakis, 1993] and for a finite density of fermions [Hellberg and Manousakis, 1997, Hellberg and Manousakis, 2000].

Projection

Projection

- We project a trial wave function $|\Psi\rangle$ onto the ground state by generating a series of increasingly accurate approximants

$$|m\rangle = (H - W)^m |\Psi\rangle.$$

H is the Hamiltonian and W is an appropriate constant.

Projection

- We project a trial wave function $|\Psi\rangle$ onto the ground state by generating a series of increasingly accurate approximants

$$|m\rangle = (H - W)^m |\Psi\rangle.$$

H is the Hamiltonian and W is an appropriate constant.

- Expanding $|\Psi\rangle$ in the basis of the exact eigenstates $|\Phi_n\rangle$

$$|\Psi\rangle = a_0|\Phi_0\rangle + a_1|\Phi_1\rangle + \cdots \qquad a_n = \langle\Phi_n|\Psi\rangle$$

where the a_n's are expansion coefficients.

Projection

- We project a trial wave function $|\Psi\rangle$ onto the ground state by generating a series of increasingly accurate approximants

$$|m\rangle = (H - W)^m |\Psi\rangle.$$

H is the Hamiltonian and W is an appropriate constant.

- Expanding $|\Psi\rangle$ in the basis of the exact eigenstates $|\Phi_n\rangle$

$$|\Psi\rangle = a_0|\Phi_0\rangle + a_1|\Phi_1\rangle + \cdots \qquad a_n = \langle\Phi_n|\Psi\rangle$$

where the a_n's are expansion coefficients.

- With E_n the corresponding exact eigenenergies, we obtain

$$\begin{aligned} |m\rangle &= a_0(E_0 - W)^m|\Phi_0\rangle + a_1(E_1 - W)^m|\Phi_1\rangle + \cdots \\ &\propto \left\{|\Phi_0\rangle + \frac{a_1}{a_0}\left(\frac{E_1 - W}{E_0 - W}\right)^m |\Phi_1\rangle + \cdots\right\} \end{aligned}$$

Projection

- We project a trial wave function $|\Psi\rangle$ onto the ground state by generating a series of increasingly accurate approximants

$$|m\rangle = (H - W)^m |\Psi\rangle.$$

H is the Hamiltonian and W is an appropriate constant.

- Expanding $|\Psi\rangle$ in the basis of the exact eigenstates $|\Phi_n\rangle$

$$|\Psi\rangle = a_0|\Phi_0\rangle + a_1|\Phi_1\rangle + \cdots \qquad a_n = \langle\Phi_n|\Psi\rangle$$

where the a_n's are expansion coefficients.

- With E_n the corresponding exact eigenenergies, we obtain

$$\begin{aligned} |m\rangle &= a_0(E_0 - W)^m|\Phi_0\rangle + a_1(E_1 - W)^m|\Phi_1\rangle + \cdots \\ &\propto \left\{|\Phi_0\rangle + \frac{a_1}{a_0}\left(\frac{E_1 - W}{E_0 - W}\right)^m |\Phi_1\rangle + \cdots\right\} \end{aligned}$$

- For large m $|m\rangle \to |\Phi_0\rangle$ provided that

$$|E_{n>0} - W| < |E_0 - W|$$

Projection (Continued)

Projection (Continued)

- The inequality

$$|E_{n>0} - W| < |E_0 - W|$$

is easy to satisfy for lattice Hamiltonians which are bounded from below *and* above, by choosing W to be at least as large as the largest eigenvalue of H. Continuum problems require a different form for the projection operator. In what follows, we assume the offset constant W is incorporated in the Hamiltonian.

Projection (Continued)

- The inequality

$$|E_{n>0} - W| < |E_0 - W|$$

is easy to satisfy for lattice Hamiltonians which are bounded from below *and* above, by choosing W to be at least as large as the largest eigenvalue of H. Continuum problems require a different form for the projection operator. In what follows, we assume the offset constant W is incorporated in the Hamiltonian.

- The rate of convergence with m

$$|m\rangle \quad \propto \quad \left\{ |\Phi_0\rangle + \frac{a_1}{a_0} \left(\frac{E_1 - W}{E_0 - W} \right)^m |\Phi_1\rangle + \cdots \right\}$$

is governed by the overlap a_0 of the trial state with the ground state and the energy of the lowest excited state overlapping the trial state. We will describe the steps taken to insure fast convergence.

Projection(Continued)

Projection(Continued)

- To calculate ground state expectation values of an arbitrary operator A, we take the large m limit of

$$\frac{\langle \Phi_0 | A | \Phi_0 \rangle}{\langle \Phi_0 | \Phi_0 \rangle} = \lim_{m \to \infty} \frac{\langle m | A | m \rangle}{\langle m | m \rangle}.$$

For large values of m, we cannot evaluate H^m directly. The number of position-space states generated diverges exponentially with the power m, so we calculate H^m by a stochastic method similar to Neumann-Ulam matrix inversion.

Projection(Continued)

- To calculate ground state expectation values of an arbitrary operator A, we take the large m limit of

$$\frac{\langle \Phi_0 | A | \Phi_0 \rangle}{\langle \Phi_0 | \Phi_0 \rangle} = \lim_{m \to \infty} \frac{\langle m | A | m \rangle}{\langle m | m \rangle}.$$

  For large values of m, we cannot evaluate H^m directly. The number of position-space states generated diverges exponentially with the power m, so we calculate H^m by a stochastic method similar to Neumann-Ulam matrix inversion.
- We decompose H into a product of a transition probability $p_{\alpha\beta}$ to make a transition from state α to state β and a residual weight $w_{\alpha\beta}$ as

$$H_{\alpha\beta} = p_{\alpha\beta} w_{\alpha\beta}$$

  where

$$\sum_{\beta} p_{\alpha\beta} = 1, \quad p_{\alpha\beta} \geq 0. \tag{1}$$

Projection (Continued)

To evaluate

$$\langle \alpha_0 | H^m | \alpha_m \rangle = \sum_{\alpha_1,\ldots,\alpha_{m-1}} \langle \alpha_0 | H | \alpha_1 \rangle \langle \alpha_1 | H | \alpha_2 \rangle \ldots \langle \alpha_{m-1} | H | \alpha_m \rangle$$

stochastically, we average over m-step random walks $\alpha_0 \to \alpha_1 \to \cdots \to \alpha_{m-1} \to \alpha_m$, where each α_i is a position space state, giving each walk the accumulated weight

$$W(\alpha_0, \alpha_1, \ldots, \alpha_m) = w_{\alpha_0\alpha_1} w_{\alpha_1\alpha_2} \cdots w_{\alpha_{m-1}\alpha_m}.$$

The probability of the walk $\alpha_0 \to \alpha_1, \cdots \to \alpha_m$ is

$$P(\alpha_0, \alpha_1, \ldots, \alpha_m) = p_{\alpha_0\alpha_1} p_{\alpha_1\alpha_2} \cdots p_{\alpha_{m-1}\alpha_m}. \tag{2}$$

Thus, it follows that

$$\langle \alpha_0 | H^m | \alpha_m \rangle = \sum_{\alpha_1,\ldots,\alpha_{m-1}} W(\alpha_0, \alpha_1, \ldots, \alpha_m) \tag{3}$$

for a large number of walks guided by the probability (2).

Importance sampling

The Monte Carlo sum (3) is evaluated most efficiently using importance sampling. We cannot use the trial state as a guiding function for the random walk: the guiding function must be positive for all allowed states. We let,

$$p_{\alpha\beta} = \frac{1}{z_\alpha} \frac{\Psi^G_\beta}{\Psi^G_\alpha} H_{\alpha\beta} \tag{4}$$

where Ψ^G is our guiding function and

$$z_\alpha = \sum_\beta \frac{\Psi^G_\beta}{\Psi^G_\alpha} H_{\alpha\beta}. \tag{5}$$

Defined in this way, (4) satisfies (1), and the residual weight is

$$w_{\alpha\beta} = z_\alpha \frac{\Psi^G_\alpha}{\Psi^G_\beta}, \tag{6}$$

and the accumulated weight for the m-step walk given by Eq. (2).

Importance sampling (Continued)

For an antisymmetric trial wave function Ψ^T, the standard algorithm to evaluate

$$\langle \Psi^T | H^m | \Psi^T \rangle = \sum_{\alpha\beta} \Psi^{T*}_{\alpha} \langle \alpha | H^m | \beta \rangle \Psi^T_{\beta}, \tag{7}$$

where $\Psi^{T*}_{\alpha} = \langle \Psi^T | \alpha \rangle$, is to generate a set of M initial states $\{\alpha_i\}$ with probabilities proportional to $Q_{\alpha_i} \propto |\Psi^T_{\alpha_i}|^2$ using Metropolis sampling as in Variational Monte Carlo. At each initial state $|\alpha_i\rangle$, we start an m-step random walk, ending in the state $|\beta_i\rangle$. For large M,

$$\langle \Psi^T | H^m | \Psi^T \rangle \rightarrow \frac{1}{M} \sum_{i=1}^{M} \frac{\Psi^{T*}_{\alpha_i} W(\alpha_i, ..., \beta_i) \Psi^T_{\beta_i}}{Q_{\alpha_i}}. \tag{8}$$

An efficient approach

An efficient approach

- The standard algorithm is inefficient since a random walk in configuration space of length m must be generated for each term in the sum (8). The details of the intermediate states are thrown away.

An efficient approach

- The standard algorithm is inefficient since a random walk in configuration space of length m must be generated for each term in the sum (8). The details of the intermediate states are thrown away.
- The expectation value $\langle \Psi | H^m | \Psi \rangle$ can be evaluated more efficiently if the generation of the initial states $\{\alpha_i\}$ is combined with the generation of the random walks.

An efficient approach

- The standard algorithm is inefficient since a random walk in configuration space of length m must be generated for each term in the sum (8). The details of the intermediate states are thrown away.
- The expectation value $\langle\Psi|H^m|\Psi\rangle$ can be evaluated more efficiently if the generation of the initial states $\{\alpha_i\}$ is combined with the generation of the random walks.
- After a large number of random walk steps new states are distributed with probability

$$Q_\alpha \propto z_\alpha[\Psi_\alpha^G]^2 \tag{9}$$

which is derived by solving the "detailed balance" condition

$$Q_\alpha p_{\alpha\beta} = Q_\beta p_{\beta\alpha}, \tag{10}$$

where $p_{\alpha\beta}$ is the probability to make a transition from configuration α to β, and Q_α is the probability to visit a state $|\alpha\rangle$.

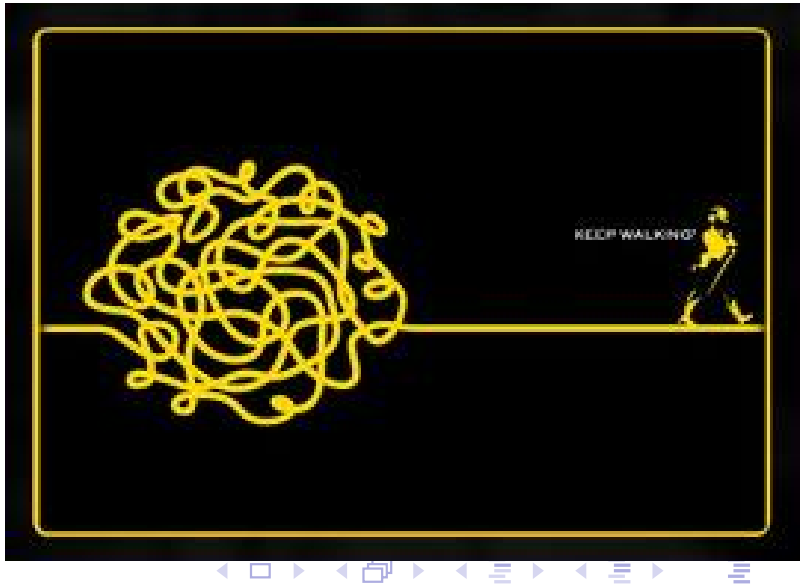
An efficient approach (Continued)

An efficient approach (Continued)

- Thus we may use states generated in the m-step random walk as initial states for new m-step random walks. For maximum efficiency, we use every state generated as the starting point for a new walk, so at each step we calculate different stages of m-walks simultaneously. We simply generate one very long random walk using the probability (4).

An efficient approach (Continued)

- Thus we may use states generated in the m-step random walk as initial states for new m-step random walks. For maximum efficiency, we use every state generated as the starting point for a new walk, so at each step we calculate different stages of m-walks simultaneously. We simply generate one very long random walk using the probability (4).
- Keep walking



An efficient approach (Continued)

An efficient approach (Continued)

- At each step in the walk, we look m steps into the past to evaluate an element of (8). The computer time needed to calculate a given number of observations of $\langle H^m \rangle$ is independent of m. An additional advantage is that since only one long random walk is generated, we may calculate all different powers m in parallel. The fundamental observation becomes

$$\langle \Psi^T | H^m | \Psi^T \rangle = \frac{1}{M} \sum_{i=1}^{M} \frac{\Psi^{T*}_{\alpha_{i-m}} W(\alpha_{i-m}, ..., \alpha_i) \Psi^T_{\alpha_i}}{z_{\alpha_{i-m}} |\Psi^G_{\alpha_{i-m}}|^2}. \quad (11)$$

An efficient approach (Continued)

An efficient approach (Continued)

- The method is easily generalized to evaluate the expectation value $A_m \equiv \langle \Psi | H^m A H^m | \Psi \rangle$ for any *diagonal* operator A, such as the density or spin structure factors, $n_i n_j$ and $S_i^z S_j^z$.

An efficient approach (Continued)

- The method is easily generalized to evaluate the expectation value $A_m \equiv \langle \Psi | H^m A H^m | \Psi \rangle$ for any *diagonal* operator A, such as the density or spin structure factors, $n_i n_j$ and $S_i^z S_j^z$.
- At each stage in the walk, we look m steps into the past to obtain the expectation value of $\langle A \rangle$ and $2m$ steps into the past to calculate the accumulated weight. By summing M observations from a walk, we find

$$A_m = \frac{1}{M} \sum_{i=1}^{M} \frac{\Psi_{\alpha_k}^{T*} W(\alpha_k, ..., \alpha_i) \Psi_{\alpha_i}^{T}}{z_{\alpha_k} |\Psi_{\alpha_k}^{G}|^2} \langle \alpha_j | A | \alpha_j \rangle. \qquad (12)$$

where $j = i - m$ and $k = i - 2m$.

An efficient approach (Continued)

- The method is easily generalized to evaluate the expectation value $A_m \equiv \langle \Psi | H^m A H^m | \Psi \rangle$ for any *diagonal* operator A, such as the density or spin structure factors, $n_i n_j$ and $S_i^z S_j^z$.
- At each stage in the walk, we look m steps into the past to obtain the expectation value of $\langle A \rangle$ and $2m$ steps into the past to calculate the accumulated weight. By summing M observations from a walk, we find

$$A_m = \frac{1}{M} \sum_{i=1}^{M} \frac{\Psi_{\alpha_k}^{T*} W(\alpha_k, ..., \alpha_i) \Psi_{\alpha_i}^{T}}{z_{\alpha_k} |\Psi_{\alpha_k}^{G}|^2} \langle \alpha_j | A | \alpha_j \rangle. \qquad (12)$$

  where $j = i - m$ and $k = i - 2m$.
- The speed of convergence of the procedure with power m is determined by the ratio $R = |E_1 - W|/|E_0 - W| < 1$. Since this gap is caused by the finite size of the system, we generally calculated powers of the Hamiltonian up to several times the linear system size.

Trial wave functions

Trial wave functions

- For high efficiency, care is needed to choose a trial state with maximal overlap with the true ground state. We restrict ourselves to total spin singlet states with zero momentum and try to write a very arbitrary form yielding a good initial guess throughout the phase diagram.

Trial wave functions

- For high efficiency, care is needed to choose a trial state with maximal overlap with the true ground state. We restrict ourselves to total spin singlet states with zero momentum and try to write a very arbitrary form yielding a good initial guess throughout the phase diagram.
- We use a Jastrow resonating-valence-bond wave function for the trial-state, written as

$$\begin{aligned} \Psi^T &= \prod_{i<j,\sigma,\sigma'} f(\mathbf{r}_{i,\sigma} - \mathbf{r}_{j,\sigma'}) |\mathrm{RVB}\rangle \qquad (13) \\ &= \prod_{i<j,\sigma,\sigma'} f(\mathbf{r}_{i,\sigma} - \mathbf{r}_{j,\sigma'}) P_N \prod_{\mathbf{k}} (u_{\mathbf{k}} + v_{\mathbf{k}} c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow}) |0\rangle, \end{aligned}$$

where $c^\dagger_{\mathbf{k}\sigma}$ is the usual Fermion creation operator and P_N projects the state onto the subspace with the number of particles fixed to be N.

Trial wave functions (Continued)

Trial wave functions (Continued)

- It is important that the Jastrow factor f correlate *all* pairs of particles independent of spin, yielding a correlated state that is still a total spin singlet. If we allow different Jastrow factors for like and unlike spins, the resultant state would be a superposition of many spin states and in general would overlap excited states with non-zero spin *closer* in energy to the ground state than the lowest spin zero excited state, resulting in *slower* projection than a singlet trial state with higher variational energy.

Trial wave functions (Continued)

- It is important that the Jastrow factor f correlate *all* pairs of particles independent of spin, yielding a correlated state that is still a total spin singlet. If we allow different Jastrow factors for like and unlike spins, the resultant state would be a superposition of many spin states and in general would overlap excited states with non-zero spin *closer* in energy to the ground state than the lowest spin zero excited state, resulting in *slower* projection than a singlet trial state with higher variational energy.
- The ratio $a_{\mathbf{k}} \equiv v_{\mathbf{k}}/u_{\mathbf{k}}$ is the physical quantity, and, assuming $a_{\mathbf{k}} = a_{-\mathbf{k}}$, we define $a(\mathbf{r})$ as its Fourier transform, then:

$$|\mathrm{RVB}\rangle = \left(\sum_{\mathbf{r}_{i\uparrow}, \mathbf{r}_{j\downarrow}} a(\mathbf{r}_{i\uparrow} - \mathbf{r}_{j\downarrow}) c^{\dagger}_{\mathbf{r}_{i\uparrow}} c^{\dagger}_{\mathbf{r}_{j\downarrow}} \right)^{N/2} |0\rangle \quad (14)$$

$$a(\mathbf{r}) = \sum_{\mathbf{k}} a_{\mathbf{k}} \cos(\mathbf{k} \cdot \mathbf{r}). \quad (15)$$

Trial wave functions (Continued)

Therefore, $|RVB\rangle$ can be written as the $\frac{N}{2} \times \frac{N}{2}$ determinant

$$|\mathbf{D}| = \begin{vmatrix} a(\mathbf{r}_{1\uparrow} - \mathbf{r}_{1\downarrow}) & a(\mathbf{r}_{1\uparrow} - \mathbf{r}_{2\downarrow}) & \cdots & a(\mathbf{r}_{1\uparrow} - \mathbf{r}_{\frac{N}{2}\downarrow}) \\ a(\mathbf{r}_{2\uparrow} - \mathbf{r}_{1\downarrow}) & a(\mathbf{r}_{2\uparrow} - \mathbf{r}_{2\downarrow}) & \cdots & a(\mathbf{r}_{2\uparrow} - \mathbf{r}_{\frac{N}{2}\downarrow}) \\ \vdots & \vdots & \ddots & \vdots \\ a(\mathbf{r}_{\frac{N}{2}\uparrow} - \mathbf{r}_{1\downarrow}) & a(\mathbf{r}_{\frac{N}{2}\uparrow} - \mathbf{r}_{2\downarrow}) & \cdots & a(\mathbf{r}_{\frac{N}{2}\uparrow} - \mathbf{r}_{\frac{N}{2}\downarrow}) \end{vmatrix} \tag{16}$$

in position space.

In this form, $|\mathrm{RVB}\rangle$ spans a broad class of Fermion wave functions.

A Fermi liquid state corresponds to

$$a_{\mathbf{k}} = \begin{cases} 1 & \mathbf{k} \in \text{Fermi sea} \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

while by allowing other choices for $a_{\mathbf{k}}$, the wave function can describe a pairing state, which may be s-wave, d-wave, or something more general.

Guiding function

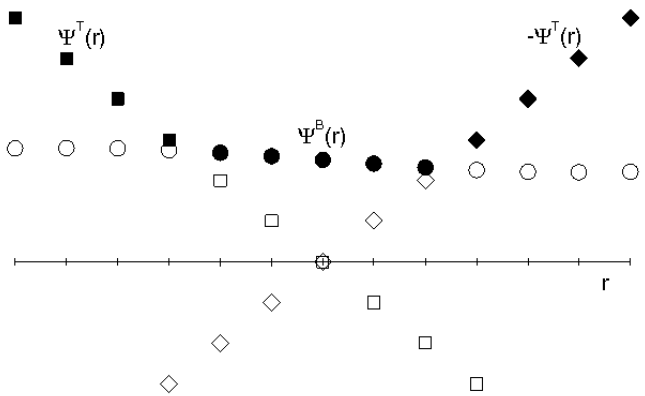
We are tempted to use the magnitude of the trial state as our guiding wave function but this would be a serious mistake. By construction, the sites of a periodic lattice lie at high symmetry points, and the nodes of a Fermion wave function also respect these symmetries. One finds $\langle \Psi_\alpha^T | \alpha \rangle = 0$ for a significant fraction of states in the Hilbert space not violating the Pauli exclusion principle.
We choose as guiding function one that reduces the functuations in $|\Psi^T|/|\Psi^G|$. We define

$$\Psi^G \equiv \max\left\{|\Psi^T|, c\Psi^B\right\}, \tag{18}$$

where Ψ^B is a positive function, typically a good variational state of the bosonic Hamiltonian. We take Ψ^B to be a spin-dependent Jastrow function. This is similar to a choice used in continuum problems, but on a discrete lattice, it is not necessary to match the first derivatives as in the continuum.

Guiding function (Continued)

We rescale c so the effective number of configurations contributing to the norm is approximately $N \approx 1/L$ for an $L \times L$ system.
This guiding function is shown schematically below:



Guiding function (Continued)

For the guiding function, we use the Jastrow-pairing function

$$\Psi^B = \prod_{i<j} f(\mathbf{r}_{i\uparrow} - \mathbf{r}_{j\uparrow}) \prod_{i<j} f(\mathbf{r}_{i\downarrow} - \mathbf{r}_{j\downarrow}) \prod_{i,j} g(\mathbf{r}_{i\uparrow} - \mathbf{r}_{j\downarrow}) \qquad (19)$$

of Bose spin $\frac{1}{2}$ particles (i.e., two kinds of bosons, up bosons and down bosons). We have chosen this function to mimic the physics of the Fermion state as closely as possible without having nodes. Since it is not important to guide with a spin singlet function, we use a more arbitrary spin-dependent Jastrow factor where like spin particles are correlated differently than opposite spin particles.

Walking through the nodes

To evaluate the determinant in the guiding function we used the so called "inverse update" first applied to condensed matter systems by Ceperley, Chester and Kalos. One calculates the determinant and the inverve of the matrix (16) together at the start of each run, an operation taking $O(N^3)$ steps for a $N \times N$ determinant. Then with each single particle move in the random walk, we update the determinant in $O(N)$ steps and the inverse in $O(N^2)$ steps. Starting with the matrix $\mathbf{D}$ and its inverse $\mathbf{I}$, suppose we change row l of the matrix to $D_{lj} \rightarrow r_j$. Since the inverse is the transpose of the matrix of cofactors normalized by the determinant,

$$I_{ij} = \mathrm{cof}_{ji}(\mathbf{D})/|\mathbf{D}|, \tag{20}$$

the ratio of the determinant of $\mathbf{D}$ before and after the change is

$$q \equiv \frac{|\mathbf{D}'|}{|\mathbf{D}|} = \sum_j r_j I_{jl}. \tag{21}$$

Walking through the nodes (Continued)

Walking through the nodes (Continued)

- The new inverse matrix is given by

$$I'_{ij} = I_{ij}\left(1 + \frac{1}{q}\delta_{lj}\right) - \frac{1}{q}I_{il}\sum_k r_k I_{kj}, \tag{22}$$

and one can easily confirm $\sum_j D'_{ij} I'_{jk} = \delta_{ik}$. Changing one column of the matrix results in a similar update for the inverse.

Walking through the nodes (Continued)

- The new inverse matrix is given by

$$I'_{ij} = I_{ij}\left(1 + \frac{1}{q}\delta_{lj}\right) - \frac{1}{q} I_{il} \sum_k r_k I_{kj}, \tag{22}$$

  and one can easily confirm $\sum_j D'_{ij} I'_{jk} = \delta_{ik}$. Changing one column of the matrix results in a similar update for the inverse.
- The algorithm is straightforward, and has been used in many GFMC studies in the continuum and in Variational Monte Carlo on a lattice. However, it cannot be used directly with GFMC on a lattice since the random walk steps directly on nodes for a significant fraction of steps. When the matrix becomes singular, its inverse is undefined.

Walking through the nodes (Continued)

- The new inverse matrix is given by

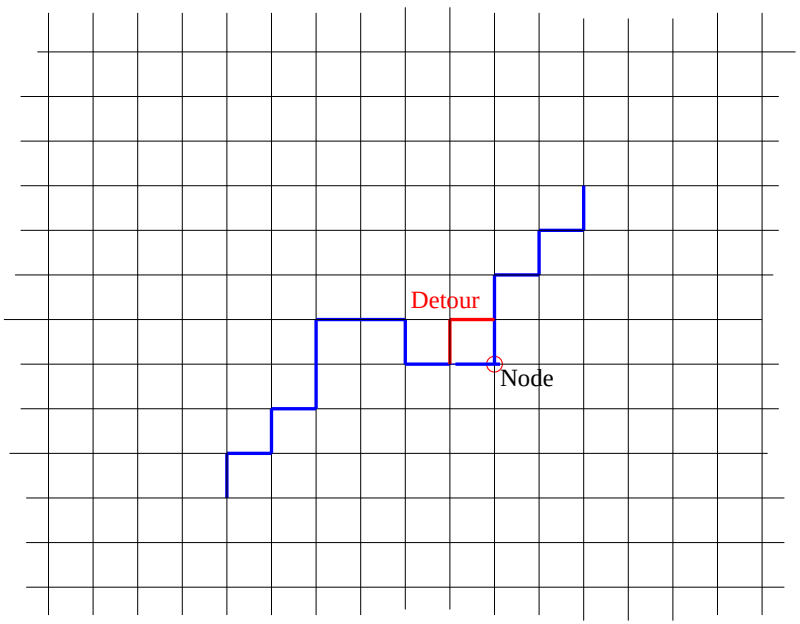
$$I'_{ij} = I_{ij}\left(1 + \frac{1}{q}\delta_{lj}\right) - \frac{1}{q}I_{il}\sum_k r_k I_{kj}, \tag{22}$$

  and one can easily confirm $\sum_j D'_{ij}I'_{jk} = \delta_{ik}$. Changing one column of the matrix results in a similar update for the inverse.
- The algorithm is straightforward, and has been used in many GFMC studies in the continuum and in Variational Monte Carlo on a lattice. However, it cannot be used directly with GFMC on a lattice since the random walk steps directly on nodes for a significant fraction of steps. When the matrix becomes singular, its inverse is undefined.
- One way around this problem is to recalculate the determinant and inverse after walking through a node. However, a large fraction of steps (we find as many as 1 out of 3) will land on nodes, and the running time will scale as $O(N^3)$.

Walking through the nodes-Detour Walk

Detour Walk. We developed a new $O(N^2)$ technique to hop over nodes without recalculation of the determinant or inverse.

The essence of the method is this:
When the random walk generated by the guiding function hits a state or series of states where the determinant of the trial function vanishes, we generate a "detour" walk around the region where the matrix is singular, rejoining the guiding walk when the determinant is non-zero again.



The real random walk goes though the node, the detour walk is a fictitious one which is used only to calculate the determinant and its inverse. It serves only as a calculational tool for the inverse update.

Walking through the nodes (Continued)

Since we often land on a node, where $\Psi_\alpha^T = 0$, and the inverse of the matrix (16) is not defined, it is difficult to calculate the probabilities $p_{\alpha\beta}$ (Eqs. 4 and 5) for the random walk we defined earlier. When we are on a node where $\Psi_\alpha^T = 0$ and we need to choose the next step of the walk from the various possible β states we need to calculate z_α using Eq. (5). Since we are on the node to calculate each Ψ_β^T we need to make a detour walk which in itself takes N^2 steps. Thus, the calculation of z_α is an $O(N^3)$ process, whereas when $\Psi_\alpha^T \neq 0$ it is an $O(N^2)$ process. Remember that the calculation of the determinant without the inverse update, is an $O(N^3)$ problem.

We make the following adjustments so that each step is an $O(N^2)$ process.

Walking through the nodes (Continued)

We define

$$p_{\alpha\beta} = \frac{1}{z_\alpha} f_{\alpha\beta} H_{\alpha\beta} \qquad z_\alpha = \sum_\beta f_{\alpha\beta} H_{\alpha\beta} \tag{23}$$

where if $\Psi_\alpha^T \neq 0$

$$f_{\alpha\beta} = \begin{cases} \Psi_\beta^G / \Psi_\alpha^G & \text{if } \Psi_\beta^T \neq 0 \\ c^2 \Psi_\alpha^B \Psi_\beta^B / |\Psi_\alpha^G|^2 & \text{if } \Psi_\beta^T = 0 \end{cases} \tag{24}$$

and if $\Psi_\alpha^T = 0$,

$$f_{\alpha\beta} = \frac{\Psi_\beta^B}{\Psi_\alpha^B} \tag{25}$$

It is easy to show that detailed balance is obeyed by these definitions. The advantage of using these probabilities is that if $\Psi_\alpha^T = 0$, then calculating Ψ_β^T for all β with $H_{\alpha\beta} \neq 0$ is not required. Since we don't have the inverse matrix of the determinant in Ψ_α^T, such a calculation would be expensive.

Fitting projection output

The Green's Function Monte Carlo procedure takes a trial state and projects it onto the exact ground state. Its output consists of the observables for the energy

$$E^{(n)} = \langle \Psi | H^n | \Psi \rangle \tag{26}$$

where the trial state $|\Psi\rangle$ has been normalized. For any operator A which does not commute with the Hamiltonian, using the present Monte Carlo method we calculate

$$a_{mn} = \langle \Psi | H^m A H^n | \Psi \rangle \tag{27}$$

as functions of powers of the Hamiltonian m and n.
These values converge to their ground state values, except for a normalization factor, for large powers n and m. However, their statistical errors increase exponentially with increasing power due to the Fermion sign problem.

Fitting projection output (Continued)

To extract the most information on the ground state, we use the calculated observable for all powers less than some maximum power. The highly converged small powers provide much more accurate ground state properties than the large powers.
Ground state energy. We define the spectral function $c(\mathcal{E})$ with respect to the trial state $|\Psi\rangle$ as

$$c(\mathcal{E}) \equiv \frac{1}{\pi} \lim_{\eta \to 0^+} Im\langle\Psi|\frac{1}{\hat{H} - \mathcal{E} + i\eta}|\Psi\rangle = \sum_i |\langle\Psi|\Phi_i\rangle|^2 \delta(\mathcal{E} - E_i)$$

We expand the trial state in the exact eigenstates $|\Phi_i\rangle$ of $\hat{H}$ as

$$|\Psi\rangle = \sum_{i=0} a_i |\Phi_i\rangle.$$

The poles of $c(\mathcal{E})$ and of the exact spectral function are at the same energy values for those eigenstates $|\Phi_i\rangle$ which have non-zero overlap with the trial state $|\Psi\rangle$.

Fitting projection output (Continued)

In order to proceed we discretize the energy interval using a fine mesh with ΔE, thus, the energy takes discrete values ${E^*}_m$, $m = 1, 2, ..., M$ and the spectral function $c^*(E)$ is written as

$$c^*(E) = \sum_{m=1}^{M} c^*_m \delta(E - {E^*}_m) \tag{28}$$

where $c^*_m \geq 0$ are non-negative real numbers. Thus, the spectral function is thought of as a histogram, where in each fine slice of the histogram the value of the integral of $c(E)$ multiplied by any function $f(E)$ is simply $c^*_m f({E^*}_m)$. We have used a mesh interval ΔE smaller than the finite-size gap between the lowest and the first excited state. Thus, the contribution of the ground state to c^* is accurately represented as a single delta-function peak. Namely, up to some $m = m_0$, $c^*_{m<m_0} = 0$, while $c^*_{m_0} > 0$ and $c^*_{m_0<m<m_1} = 0$ and $c^*_{m_1} > 0$, etc.

Fitting projection output (Continued)

- The moments of the spectral function $c^*(E)$ can be calculated as:

$$\langle \Psi | H^n | \Psi \rangle = \sum_{m=1}^{M} c_m^* E^{*n}_m \tag{29}$$

where $n = 0, 1, ..., p_{max}$. Since $p_{max} < M$, because typically $p_{max} = 40 - 60$ and $M = 200 - 500$, we have more unknowns (the c_m^*'s) than equations. However, the solution needs to satisfy the contraint $c_m^* \geq 0$ which limits the possible solutions. The optimal way to find the most likely solution is to minimize the χ^2.

- We gain very large computational savings by calculating all powers of H in parallel. However, this results in statistical correlations between results of different powers which must be treated accordingly [Toussaint, 1989].

Fitting projection output (Continued)

We divide the measurements into M bins. The covariance matrix is defined

$$C_{ij} = \frac{1}{M-1}\left(\frac{1}{M}\sum_{k=1}^{M}(E_k^{(i)} - \bar{E}^{(i)})(E_k^{(j)} - \bar{E}^{(j)})\right) \tag{30}$$

where $E_k^{(i)}$ is the average of the i'th power in the k'th bin. For uncorrelated output, C is diagonal. With correlations, χ^2 is defined

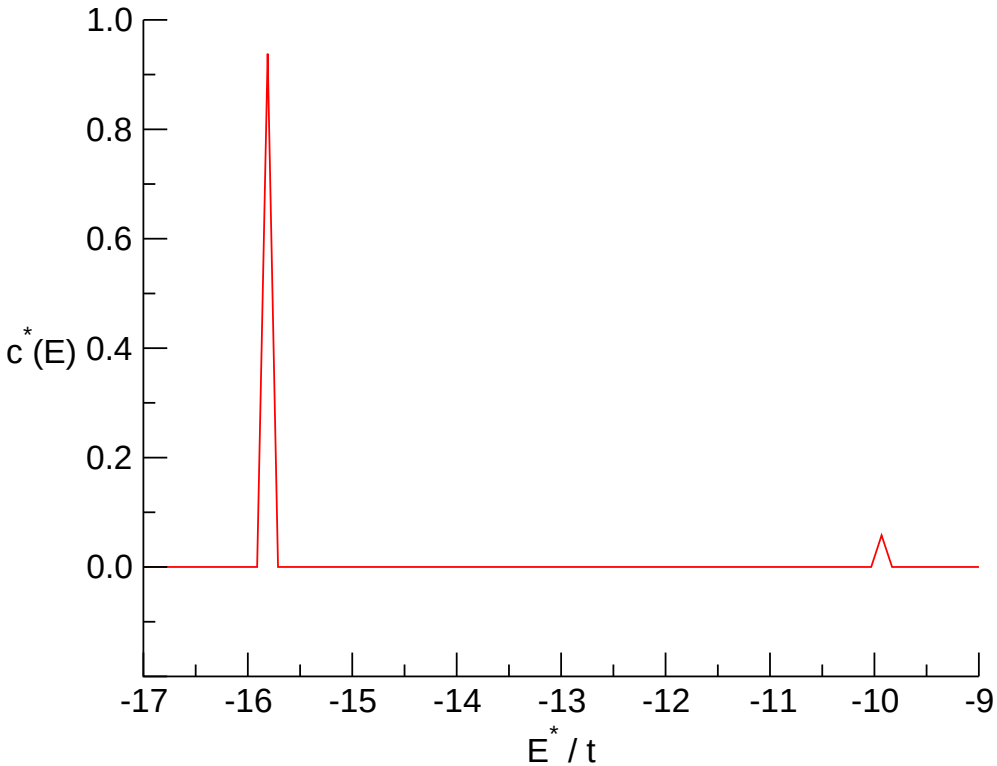
$$\chi^2 = \sum_{ij}(\bar{E}^{(i)} - E^{(i)*})\mathrm{C}_{ij}^{-1}(\bar{E}^{(j)} - E^{(j)*}), \tag{31}$$

where $E^{(i)*}$ is the fitting function given in terms of the coefficients C_m (which are to be determined by this minimization) by means of Eq. (29).

Fitting projection output (Continued)

- When C is diagonal, its inverse is trivial. For more general C, small errors in its components can result in large errors in its inverse, so it is important to calculate C accurately.
- Increasing the number of bins decreases the statistical error in C but increases the systematic error due to autocorrelations. To balance these two sourses of error, we choose the number of measurements in each bin to be $n = Mp_{\max}$, where $p_{\max}$ is the maximum power of the Hamiltonian [Toussaint, 1989].
- We calculate statistical errors with the bootstrap method.

Fitting projection output (Continued)



The lowest value of E^* where we have a delta function peak gives the lowest eigenstate of H which is not orthogonal to the trial state. The value of the peak gives the square of the ovelap of the lowest energy state to the trial state.

Fitting projection output (Continued)

For an arbitrary operator A, we have

$$\begin{aligned}\langle\Psi|H^n A H^n|\Psi\rangle &= \sum_{ij}(E_i E_j)^P \langle\Psi|\Phi_i\rangle\langle\Phi_i|A|\Phi_j\rangle\langle\Phi_j|\Psi\rangle \\ &= \int dE E^{2p} a(E) \end{aligned} \tag{32}$$

where the operator overlap function $a(E)$ is given by

$$a(E) = \sum_{ij} \delta(E + \sqrt{E_i E_j})\langle\Psi|\Phi_i\rangle\langle\Phi_i|A|\Phi_j\rangle\langle\Phi_j|\Psi\rangle. \tag{33}$$

Following the approach for the energy spectral function, we can define a discrete overlap function

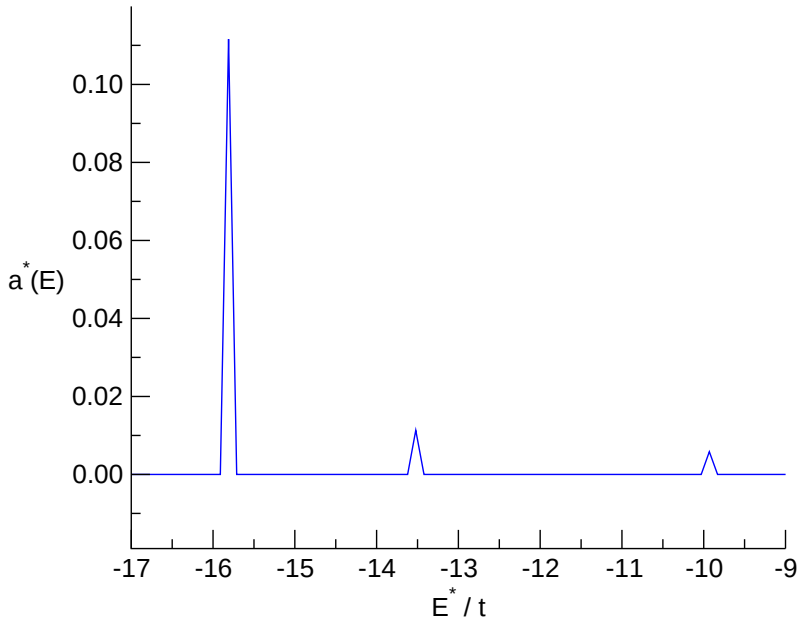
$$a^*(E) = \sum_{i=0}^{M} \delta(E - E_i^*) a_i^*. \tag{34}$$

Fitting projection output (Continued)

Here the values of E where the spectral function $a^*(E)$ attains peaks are all possible geometric means $\sqrt{E_i E_j}$ of all the eigenenergies which correspond to eigenstates which have non-zero overlap with $|\Psi\rangle$ and they give non-zero matrix element of A. The lowest energy peak corresponds to the geometric mean of the ground state energy with itself, i.e $E_0 = \sqrt{E_0 E_0}$ thus, if the ground state is not degenerate it is uniquely specified. Here we also need to solve for all the a_i^* given that they obey the following $p_{\max}$ equations:

$$\begin{aligned} \langle\Psi|A|\Psi\rangle &= \sum_{i=0}^{M} a_i^* \\ \langle\Psi|HAH|\Psi\rangle &= \sum_{i=0}^{M} (E_i^*)^2 a_i^* \\ &\cdot \\ &\cdot \end{aligned}$$

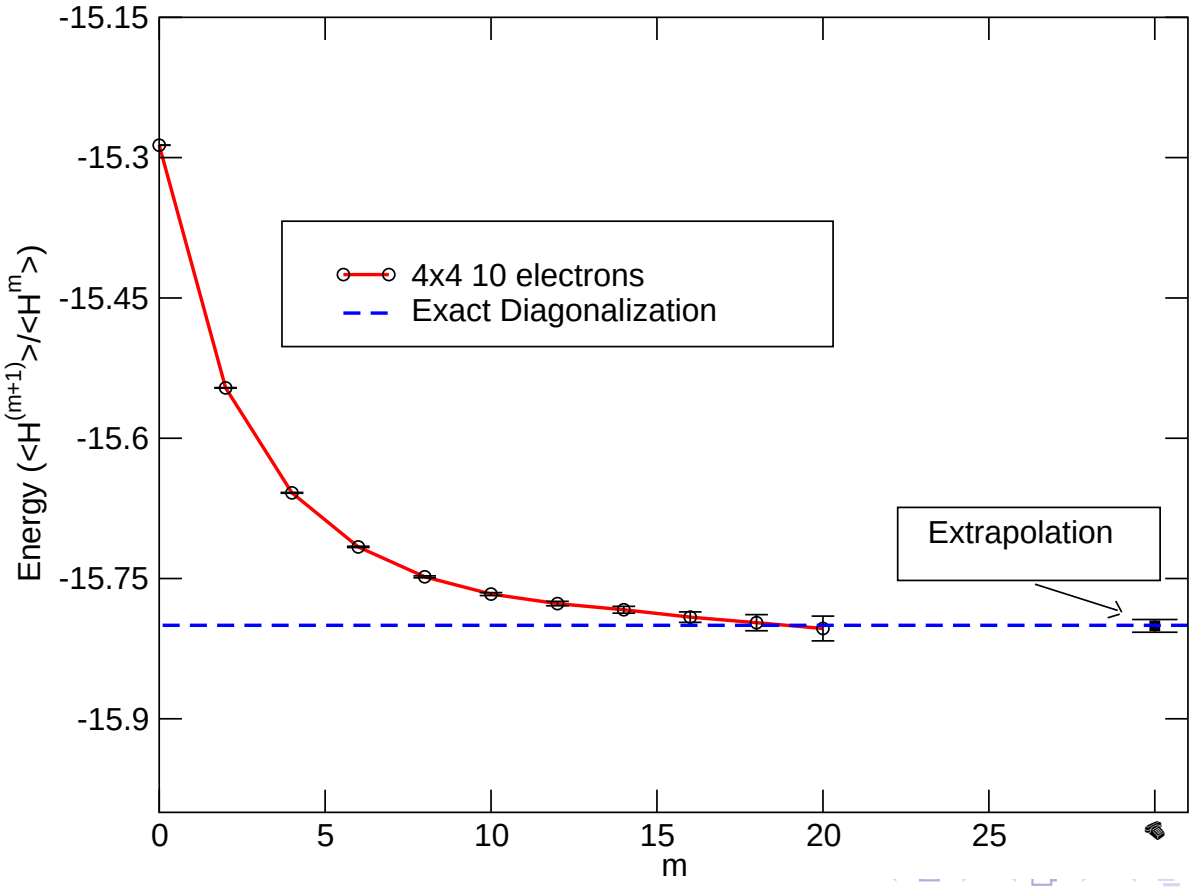
Fitting projection output (Continued)



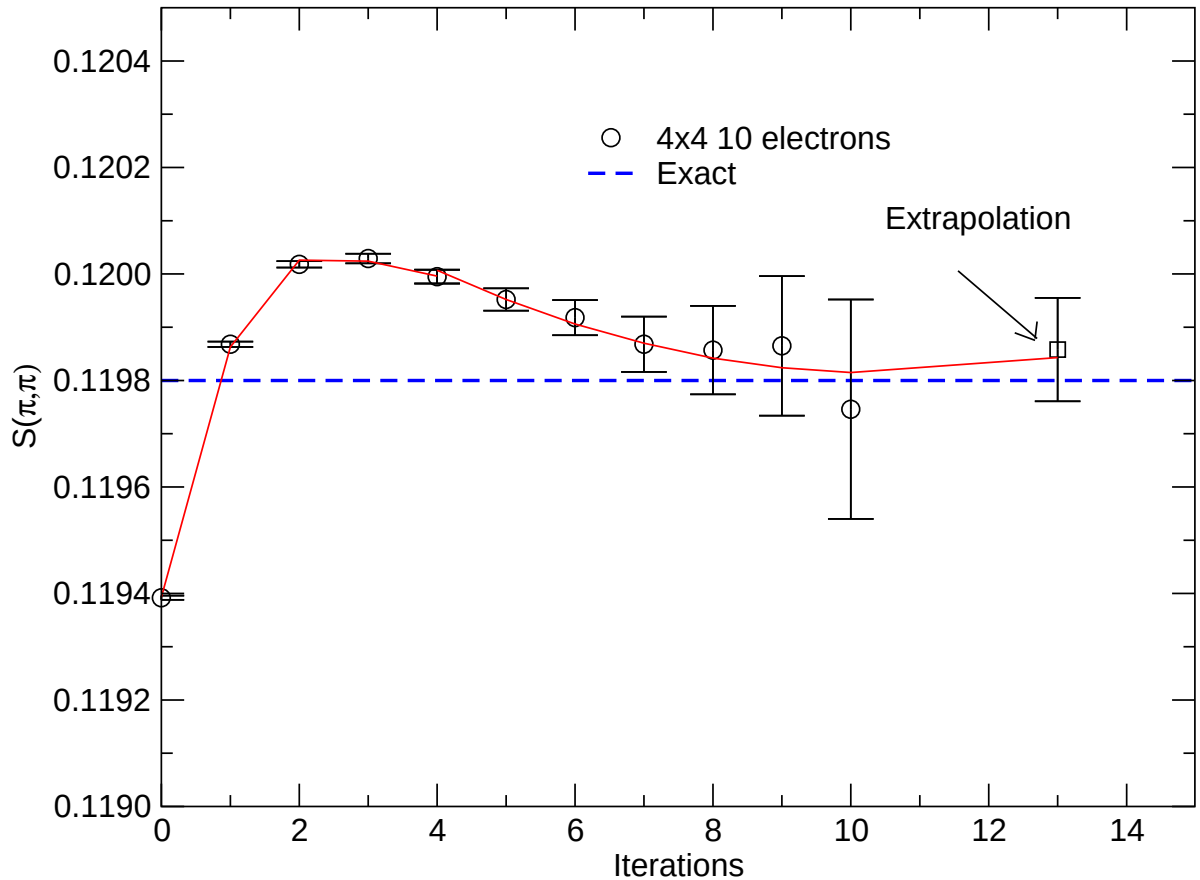
In principle, we can also extract information about the excited states along with the ground state. This possibility is indicated by the fact that we can see higher energy peaks in these spectral functions.

Test of convergence

We have tested our method by comparing our results with exact results for the 4×4 size lattice with several electrons. In Fig. 44 we show the results for the energy as a function of the iteration for the case of 10 electrons.



Test of convergence (Continued)



In Fig. 45 we compare our results for the spin-spin correlation function with that obtained by exact diagonalization of the $4x4$ size system with 10 electrons.

Conclusions

Conclusions

- We have developed an efficient Quantum Monte Carlo method which resembles the GFMC method where without eliminating the minus-sign the fluctuations are controlled with the aid of appropriately constructed guiding functions up to a certain high power of the Green's function.

Conclusions

- We have developed an efficient Quantum Monte Carlo method which resembles the GFMC method where without eliminating the minus-sign the fluctuations are controlled with the aid of appropriately constructed guiding functions up to a certain high power of the Green's function.
- Starting from a good initial state allows us to achieve convergence before the statistical errors become too large.

Conclusions

- We have developed an efficient Quantum Monte Carlo method which resembles the GFMC method where without eliminating the minus-sign the fluctuations are controlled with the aid of appropriately constructed guiding functions up to a certain high power of the Green's function.
- Starting from a good initial state allows us to achieve convergence before the statistical errors become too large.
- We developed a powerful technique which uses all the calculated powers of the Hamiltonian to extrapolate to infinite power.

Conclusions

- We have developed an efficient Quantum Monte Carlo method which resembles the GFMC method where without eliminating the minus-sign the fluctuations are controlled with the aid of appropriately constructed guiding functions up to a certain high power of the Green's function.
- Starting from a good initial state allows us to achieve convergence before the statistical errors become too large.
- We developed a powerful technique which uses all the calculated powers of the Hamiltonian to extrapolate to infinite power.
- This technique comes also with solutions to a number of other technical problems such as:

Conclusions

- We have developed an efficient Quantum Monte Carlo method which resembles the GFMC method where without eliminating the minus-sign the fluctuations are controlled with the aid of appropriately constructed guiding functions up to a certain high power of the Green's function.
- Starting from a good initial state allows us to achieve convergence before the statistical errors become too large.
- We developed a powerful technique which uses all the calculated powers of the Hamiltonian to extrapolate to infinite power.
- This technique comes also with solutions to a number of other technical problems such as:
- a) enabling the guided random walk to walk through the nodes with an $O(N^2)$ algorithm using the idea of "detour walk". This does not restrict our calculation within the fixed node approximation.

Conclusions (Cont'd)

Conclusions (Cont'd)

- b) instead of using multiple walkers the introduction of a single walker whose walk is very long allows us to compute all the desired powers of H^m $m = 0, 1, ..., p_{max}$ in parallel by looking back in the past p_{max} steps of the walk.

Inverse update through nodes

To evaluate the determinant in the trial state, we use the usual "Inverse Update" trick first applied to condensed matter systems by Ceperley, Chester, and Kalos [Ceperley et al., 1977]. We calculate the determinant and inverse of the matrix (16) together at the start of each run, an operation taking $O(N^3)$ steps for a $N \times N$ determinant. Then with each single particle move in the random walk, we update the determinant in $O(N)$ steps and the inverse in $O(N^2)$ steps.

Starting with the matrix $\mathbf{D}$ and its inverse $\mathbf{I}$, suppose we change row l of the matrix to $D_{lj} \rightarrow r_j$. Since the inverse is the transpose of the matrix of cofactors normalized by the determinant,

$$I_{ij} = \mathrm{cof}_{ji}(\mathbf{D})/|\mathbf{D}|, \tag{36}$$

the ratio of the determinant of $\mathbf{D}$ before and after the change is

$$q \equiv \frac{|\mathbf{D}'|}{|\mathbf{D}|} = \sum_j r_j I_{jl}. \tag{37}$$

Inverse update through nodes (Continued)

The algorithm is straight forward, and has been used in many GFMC studies in the continuum and in Variational Monte Carlo on a lattice. However, it cannot be used directly with GFMC on a lattice since the random walk steps directly on nodes for a significant fraction of steps. When the matrix becomes singular, its inverse is undefined, and the algorithm breaks down.

One way around this problem is to recalculate the determinant and inverse after walking through a node. However, in a reasonably dense system, a large faction of steps will land on nodes, and the running time will scale as $O(N^3)$.

We developed a new $O(N^2)$ technique to hop over nodes without recalculation of the determinant or inverse. The essence of the method is this: Let us suppose that the random walk visits a node; namely, the particles were in a configuration $\vec{R} = (\vec{r}_{1\uparrow}, \vec{r}_{2\uparrow}, ..., \vec{r}_{N/2\uparrow}, \vec{r}_{1\downarrow}, \vec{r}_{2\downarrow}, ..., \vec{r}_{N/2\downarrow})$ and by moving a particle, say the first up-spin particle from position $\vec{r}_{1\uparrow}$ to $\vec{r}'_{1\uparrow}$ the

Inverse update through nodes (Continued)

When the random walk generated by the guiding function hits a state or series of states where the determinant of the trial function vanishes, we generate a "detour" walk around the region where the matrix is singular, rejoining the guiding walk when the determinant is non-zero again.

To choose the detour walk, we simply delay any move causing the determinant to vanish and place the particle number and its future site at the beginning of a list of moves to make. For any subsequent move of a particle of the same spin, we try to move the first particle in the list to that site. If that move yields a non-zero determinant, we accept it and attempt to move the next particle in the list in the same manner. We repeat the process until either all moves give zero determinant or the list is empty, in which case the true determinant is not zero.

Obviously, the procedure will not produce a non-zero determinant when the true determinant is zero. However, it is important to

Inverse update through nodes (Continued)

We represent the rows of the matrix (16) by $\mathbf{D} = \{|\mathbf{r}_1), |\mathbf{r}_2), \ldots, |\mathbf{r}_n)\}$ where $|\vec{r}_i)$ represents the row

$$|\vec{r}_i) = (a(\vec{r}_{i\uparrow} - \vec{r}_{1\downarrow}), a(\vec{r}_{i\uparrow} - \vec{r}_{2\downarrow}), ..., a(\vec{r}_{i\uparrow} - \vec{r}_{N/2\downarrow})) \tag{39}$$

which is labeled by $\vec{r}_i$. Suppose moving the first up particle to a new site, changing the first row label to $\mathbf{r}_1 \rightarrow \mathbf{s}$, yields a zero determinant. Then

$$|\mathbf{s}) = \alpha_2|\mathbf{r}_2) + \alpha_3|\mathbf{r}_3) + \cdots + \alpha_n|\mathbf{r}_n) \tag{40}$$

for some coefficients $\alpha_2, \alpha_3, \ldots, \alpha_n$.

Inverse update through nodes (Continued)

Let the next random walk step move the second particle, changing row $\mathbf{r}_2 \rightarrow \mathbf{t}$. Simply by checking if the matrix $\mathbf{D}' = \{|\mathbf{t}), |\mathbf{r}_2), \ldots, |\mathbf{r}_n)\}$ has zero determinant, we can determine if the true matrix $\mathbf{D}'' = \{|\mathbf{s}), |\mathbf{t}), |\mathbf{r}_3), \ldots, |\mathbf{r}_n)\}$ is singular. If $|\mathbf{D}'| \neq 0$, we accept the move, swap the particles, and try to move the first particle again. If $|\mathbf{D}'| = 0$,

$$|\mathbf{t}) = \beta_2|\mathbf{r}_2) + \beta_3|\mathbf{r}_3) + \cdots + \beta_n|\mathbf{r}_n) \tag{41}$$

for certain coefficients $\beta_2, \beta_3, \ldots, \beta_n$. Combining (40) and (41) to eliminate $\mathbf{r}_2$, we see that $|\mathbf{D}'| = 0$ implies $|\mathbf{D}''| = 0$. Thus by simply checking single particle moves, we can verify that the determinant of the matrix two steps away is zero. The argument is easily generalized to any number of delayed moves.

Inverse update through nodes (Continued)

For a Fermi liquid state, (16) may be expanded into the product of two Slater determinants, and this algorithm suffices as it stands. However with a pairing trial state, this decomposition is not possible, we must consider moves of opposite spin electrons causing the determinant to vanish.
Suppose we find moving either the first up particle, changing the first row to $D_{1j} \rightarrow r_j$, or the first down particle, changing the first column to $D_{i1} \rightarrow c_i$, results in a zero determinant for matrix (16). If $I_{11} \neq 0$ we move the first row with the first element shifted by $1/I_{11}$, noting

$$|\mathbf{D}'| = \begin{vmatrix} r_1 + \frac{1}{I_{11}} & r_2 & \cdots & r_n \\ D_{21} & D_{22} & \cdots & D_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ D_{n1} & D_{n2} & \cdots & D_{nn} \end{vmatrix} = |\mathbf{D}|. \qquad (42)$$

We then try to change the first column in the standard manner.

Inverse update through nodes (Continued)

If $I_{11} = 0$, this modified step is no longer possible, and we need to prove that the true determinant,

$$|\mathbf{D}''| = \begin{vmatrix} x & r_2 & \cdots & r_n \\ c_2 & D_{22} & \cdots & D_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & D_{n2} & \cdots & D_{nn} \end{vmatrix} = 0, \tag{43}$$

vanishes. Here x is the upper left element after both moves. We know there exist coefficients $\alpha_2, \alpha_3, \ldots, \alpha_n$ such that $r_j = \sum_{i \geq 2} \alpha_i d_{ij}$ for all j. Since the inverse is related to the matrix of cofactors by (36), $\mathrm{cof}_{11}(\mathbf{D}) = 0$, and there are other coefficients $\beta_2, \beta_3, \ldots, \beta_n$ such that $0 = \sum_{i \geq 2} \beta_i d_{ij}$ for all $j \geq 2$. If $\sum_{i \geq 2} \beta_i c_i = 0$, then $|\mathbf{D}''| = 0$ trivially. Otherwise let $\gamma_i = \alpha_i + \lambda \beta_i$ where

$$\lambda = \left(x - \sum_{i \geq 2} \alpha_i c_i \right) \Big/ \sum_{i \geq 2} \beta_i c_i. \tag{44}$$


Boninsegni, M. and Manousakis, E. (1992).
Phys. Rev. B, 46:560.


Boninsegni, M. and Manousakis, E. (1993).
Phys. Rev. B, 47:11897.


Carlson, J. (1989).
Phys. Rev. B, 40:846.


Ceperley, D. M., Chester, G. V., and Kalos, M. H. (1977).
Phys. Rev. B, 16:3081.


Hellberg, C. S. and Manousakis, E. (1997).
Gfmc for latticle fermions.
Phys. Rev. Lett., 78:4609.


Hellberg, C. S. and Manousakis, E. (2000).
Green's function monte carlo for lattice fermions: Application to the t-j model.
Phys. Rev. B, 61:11787.


Runge, K. J. (1992).
Quantum monte carlo calculation of the long-range order in the heisenberg antiferromagnet.
Phys. Rev. B, 45:12292.


Toussaint, D. (1989).
World Scientific, Singapore.


Trivedi, N. and Ceperley, D. M. (1989).
Green-function monte carlo study of quantum antiferromagnets.
Phys. Rev. B, 40:2737.


Trivedi, N. and Ceperley, D. M. (1990).
Ground-state correlations of quantum antiferromagnets: A green-function monte carlo study.
Phys. Rev. B, 41:4552.