Atomic Scale Simulations

Projector Quantum Monte Carlo

Projector Monte Carlo

We now turn to a potentially more powerful method where a function of the Hamiltonian projects out the the ground state, hence the name, projector Monte Carlo. The nomenclature of the various quantum Monte Carlo methods is not at all standardized. Table I shows the operators that have been used as projectors, or Green's functions. For simplicity I will only discuss Diffusion Monte Carlo although most of what I say carries over immediately to the other projectors.

A sequence of trial functions is defined by repeatedly applying the projector, $G(R, R')$: to some initial state $\psi_0(R)$:

$$\psi_{n+1}(R) = e^{-\tau(\mathcal{H}-E_T)}\psi_n(R) = \int dR' G(R, R')\psi_n(R'). \tag{1}$$

The effect on the trial function of the Green's function is seen by expanding the trial function in the set of exact eigenfunctions ϕ_α of the Hamiltonian. The nth iterate is:

$$\psi_n(R) = \sum_\alpha \phi_\alpha(R)\langle\phi_\alpha|\psi_0\rangle e^{-n\tau(E_\alpha - E_T)}. \tag{2}$$

The Green's functions shown in the table will project out the state of lowest energy having a non-zero overlap with the initial trial function:

$$\lim_{n\to\infty} \psi_n(R) = \phi_0(R)\langle\phi_0|\psi_0\rangle e^{-n\tau(E_0 - E_T)}. \tag{3}$$

The role of the *trial energy*, E_T is to keep the overall normalization of ψ_n fixed, which implies $E_T \approx E_0$. The *time step*, τ, controls the rate of convergence to the ground state. It must be taken small to allow us to make accurate approximations to the Green's function.

Since the evaluation of the Green's function involves a $3N$ dimensional integral, if N is larger than two or three one must do the integral with Monte Carlo. The interpretation of Eq. (1) is very similar to the Markov chain we discussed earlier except than the normalization is not fixed. The probability of starting a random walk at R_1 is $\psi_0(R_1)$. For the moment let us discuss bosons where ψ_0 is non-negative. To sample $\psi_1(R)$, we choose moves from R_0 to R_1 from the Green's function $G(R_1, R_0)$. Projector MC is different than a Markov process in that the Green's function is non-normalized as we will see. It does describe a Markov process in ensemble space.

Trotter's theorem[2] says that in the limit of small time step we are allowed to consider the kinetic energy and potential energy terms independently. In the limit that the time step approaches zero, a coordinate space representation of the Green's function is:

$$< R|e^{-\tau(\mathcal{H}-E_T)}|R' >= (4\pi\lambda\tau)^{-3N/2} e^{-\frac{(R-R')^2}{4\lambda\tau}} e^{-\tau(V(R)-E_T)} \; + \; \mathcal{O}(\tau^2), \tag{4}$$

The iteration equation, Eq. (1) , has a simple interpretation in terms of branching random walks since the first factor is the Green's function for diffusion and the second is multiplication of the distribution by a positive scalar. Luckily both are non-negative so a probabilistic interpretation is possible. Such

$Method$		$G(R,R')$	$ref.$
$Diffusion$	DMC	$\exp[-\tau(\mathcal{H}-E_T)]$	[9, 5]
$Green'sFunction$	$GFMC$	$[1+\tau(\mathcal{H}-E_T)]^{-1}$	[22, 4]
$Power$	PMC	$[1-\tau(\mathcal{H}-E_T)]$	[1]

Table 1: The Green's functions for various projection methods. τ is the timestep, and E_T is the trial energy. They have all been normalized to be unity and to have the same derivative for $\tau = 0$.

is not the case for arbitrary Hamiltonians. The branching process makes projector Monte Carlo differ from a Markov process: walks are allowed to split and to die.

The computer algorithm is quite simple: an ensemble of configurations is constructed with a Metropolis sampling procedure for $\psi_0(R)$. This is the *zeroth generation*, *i.e.* $n = 0$. The number of configurations is the *population* of the zeroth generation, P_0. Points in the next generation are constructed by sampling the Gaussian distribution in Eq. (4) and then branching. The number of copies of R' in the next generation is the integer part of

$$m = u + \exp[-\tau(V(R) - E_T)] \tag{5}$$

where u is a uniform random number in $(0, 1)$. If we average over u, we see that the average density of points so sampled is precisely equal to the Green's function. If the potential energy is less than the ground state energy, duplicate copies of the configuration may be generated. In succeeding generations, these walks propagate independently of each other. In places of high potential energy, random walks are terminated.

This procedure is a Markov process where the state of the walk in the nth generation is given by $\{P_n; R_1, R_2, \ldots R_{P_n}\}$. Hence it has a unique stationary distribution, constructed to be the ground state wave function. As a result the population (the number of walkers) fluctuates from step to step. It executes a undirected random walk and if uncontrolled, will either reach zero or go off to infinity. The trial energy, E_T, must be adjusted to keep the population within computationally acceptable limits. This is done by adjusting the trial energy. The smoothest way of doing this feedback is:

$$E_T = E_0 + \kappa \ln(P^*/P), \tag{6}$$

where P is the current population, P^* is the desired population, E_0 is the best guess of the ground state energy, and κ is a feedback parameter adjusted to be small as possible while achieving the goal of stabilizing the population around the target, P^*. If κ is too large, one can bias the distribution.

Importance Sampling

The algorithm as described above was first suggested by Fermi, and actually tried out in the first days of computing some forty years ago [3]. But it fails on many-body systems because the potential is unbounded. For example, a coulomb potential can go to both positive and negative infinity as two charges approach each other. In the case that like charges come close together branching will kill the entire configuration. In the case of unlike charges, the configuration can branch into an unlimited number of copies and those copies will continue branching. Even with a bounded potential the method becomes very inefficient as the number of particles increases since the branching factor grows with the number of particles. Mathematically it is correct, but computationally unstable and inefficient.

There is a simple cure discovered by Kalos [4] for GFMC, and extended by Ceperley and Alder to Diffusion Monte Carlo and fermion systems[5]. *Importance sampling* multiplies the underlying probability distribution by a known, approximate solution called the *trial* or *guiding* function, $\Psi(R)$. Multiply Eq. (1) by Ψ, the trial function, and define $f_n(R) = \Psi(R)\psi_n(R)$. Then:

$$f_{n+1} = \Psi e^{-\tau(\mathcal{H}-E_T)}\psi_n = \int dR' \tilde{G}(R,R') f_n(R') \tag{7}$$

where $\tilde{G}(R,R') = \Psi^{-1}e^{-\tau(\mathcal{H}-E_T)}\Psi$ is the importance-sampled Green's function and the initial conditions are $f_0(R) = \Psi(R)\psi_0(R)$. It is easily shown by differentiating $\tilde{G}$ with respect to τ that it satisfies the evolution equation:

$$-\frac{\partial \tilde{G}(R,R_0;\tau)}{\partial \tau} = -\sum_i \lambda_i \nabla_i [\nabla_i \tilde{G} + 2\tilde{G}\nabla_i \ln(\Psi(R))] + [E_L(R) - E_T]\tilde{G}, \tag{8}$$

where $E_L(R)$ is the local-energy defined in the VMC section. As we discussed earlier, Trotter's theorem[2] says that for short enough time steps each term on the right-hand side can be considered as an independent process in the random walk. The three terms on the right-hand side correspond to diffusion, drifting and branching. We have already discussed diffusion and branching. We just have to add drift to the previous algorithm.

The Diffusion Monte Carlo (DMC) algorithm is:

1. The ensemble is initialized by sampling from $\Psi^2(R)$ using VMC.

2. The points in the configuration are advanced in time as:

$$R_{n+1} = R_n + \chi + \lambda\tau\nabla \ln(\Psi(R_n)^2), \tag{9}$$

   where χ is a normally distributed $3N$ dimensional random vector with variance $2\lambda\tau$ and zero mean. The last term is the drift.

3. The number of copies of each configuration is the integer part of

$$\exp(-\tau(E_L(R_n) - E_T)) + u, \tag{10}$$

   where u is a uniformly distributed random number in $(0,1)$ and E_T is the current trial energy. As the trial function approaches the exact eigenfunction, the branching factor approaches unity; thus a sufficiently good trial function can control the branching.

4. The energy is calculated as the average value of the local energy: $E_0 = \langle E_L(R_n)\rangle$.

5. The trial energy is periodically adjusted to keep the population stable using Eq. (6).

6. To obtain ground state expectations of quantities other than the energy, one must correct the average over the DMC walk using the *mixed estimator*, $V_{mix} = \langle\phi_0|V|\Psi\rangle$, and the variational estimator [6]. For example the potential energy is calculated as:

$$\langle\phi_0|V|\phi_0\rangle \approx 2\langle\phi_0|V|\Psi\rangle - \langle\Psi|V|\Psi\rangle + \mathcal{O}([\phi_0 - \Psi]^2). \tag{11}$$

   The first term on the LHS is the mixed estimator produced by the projector Monte Carlo, the second term the variational estimate. If the mixed estimator equals the variational estimator then the trial function has maximum overlap with the ground state. For a strictly positive quantity like the density or pair correlation function we can use another estimator:

$$\langle\phi_0|\rho|\phi_0\rangle \approx \frac{[\langle\phi_0|\rho|\Psi\rangle]^2}{\langle\Psi|\rho|\Psi\rangle} + \mathcal{O}([\phi_0 - \Psi]^2) \tag{12}$$

to keep the extrapolated density positive.

Note that repeated use of step 2 alone would generate a probability density proportional to Ψ^2, if we turn off the branching we recover VMC. One can substantially improve the DMC algorithm given above by using rejections to put in the exact detailed balance property as detailed in ref. [7]. Further recent improvements are described in ref[8].

In the GFMC algorithm introduced by Kalos there is no error resulting from taking a finite timestep which makes it very useful for performing precise energy calculations. Its essence is identical to the above algorithm. The new algorithmic features of GFMC are the introduction of intermediate points and the sampling of the value of the timestep. But the other features are very similar.

The Fixed-Node and Fixed-Phase Method

We have not discussed at all the problem posed by fermi statistics or complex-valued wavefunctions to the projector Monte Carlo method. First let us consider the difficulty in implementing the non-importance sampled algorithm. The initial condition $\phi_0(R)$ is not a probability distribution since a fermion trial function will have an equal volume of positive and negative regions (assuming it can be made real at all.) We must use the initial sign of the wave function as a weight for the random walk. That leads to an exact but slowly converging algorithm that we will discuss in the next subsection.

Importance sampling cures this defect of the initial condition since the initial distribution $|\Psi(R)|^2$ is positive, but the importanced sampled Green's function, $\tilde{G}(R, R')$ can be negative, if in a step from R to R' we have $\Psi(R)\Psi(R') < 0$. If we sample the absolute value of the Green's function, then sign can be used as a weight. The walk will count negatively towards averages until such a time as it recrosses the node. This leads to a growing statistical variance for all matrix elements. There is a simple way to avoid the sign: forbid moves in which the sign of the trial function changes. This is the fixed-node (FN) approximation[9].

In a diffusion process, forbidding node crossings puts a zero boundary condition on the evolution equation for the probability. This solves the wave equation with the boundary conditions that the solution vanish wherever the trial function vanishes. The calculated energy will be an upper bound to the exact ground state energy[10], in fact the best possible upper bound with the given boundary conditions. One is exactly solving the wave equation inside of all the nodal regions, but there is a mismatch of the derivative of the solution across the boundary. With the FN method, we do not necessarily have the exact fermion energy, but the results are much superior to those of VMC. No longer do we have to optimize two-body correlation factors, three-body terms etc., since the nodes of the trial function are unchanged by those terms. The fixed-node method gets rid of the basis set problem to a large extent. One generally finds that the systematic error in the FN calculation is three to ten times smaller than it would be for a well-optimized VMC energy.

The nodes play a very important role since, as we have seen, if the nodes were exactly known, the many-fermion system could be treated by Monte Carlo methods without approximation. Let me briefly recap a few basic facts about nodal surfaces. First note that the ground state wave function can be chosen real in the absence of magnetic fields; the nodes are the set of points where $\phi(R) = 0$. Since this is a single equation, the nodes are in general a $3N-1$ dimensional hypersurface. (A common confusion is between these many-body nodes and those of the spin-orbits which are 2D surfaces in a 3D space.) When any two particles with the same spin are at the same location the wave function vanishes. These coincident planes, with $\mathbf{r}_i = \mathbf{r}_j$ are $3N-3$ dimensional hypersurfaces. In 3D space they do not exhaust the nodes, but are a sort of scaffolding. The situation is very different in 1D where the set of nodes is usually equal to the set of coincident hyperplanes. Fermions in 1D are equivalent to 1D bosons with a no-exchange rule.

Nodal volumes of ground state wave functions possess a tiling property[11]. To define this property

first pick a point, R_0, which does not lie on the nodes. Consider the set of points which can be reached from R_0 by a continuous path with $\phi(R) \neq 0$. This is the volume in phase space accessible to a fixed-node random walk starting at R_0. Now consider mapping this volume with the permutation operator (only permute like spins), *i. e.* relabel the particles. The tiling theorem says that this procedure completely fills phase space, except, of course, for the nodes. Thus one does not have to worry about where the random walk started; all starting places are equivalent. This theorem applies for any fermion wave function which is the ground state for some local Hamiltonian. Excited states, ground states of non-local Hamiltonians, or arbitrary antisymmetric functions need not have the tiling property. More extensive discussion of fermion nodes and some pictures of cross-sections of free particle nodes are given in ref. [11].

Let us now consider systems for which the wavefunction is necessarily complex-valued. Two examples are when there is a strong magnetic field and when one wants to work in a state of fixed non-zero linear or angular momentum such as a vortex. The generalization of the VMC method is straightforward in principle: one simply samples the square of the modulus of the wavefunction. See, for example ref. [12]. One complication is in finding good wavefunctions, particularly in periodic boundary conditions since now the phase of the wavefunction can be periodic or more generally quasi-periodic. For the projector MC methods, the fixed-node method can be generalized to the *fixed-phase* method. Here the phase is specified by a variational wavefunction such as HF and the modulus is exactly solved for using the Diffusion Monte Carlo method. All that needs to be changed (over the zero field situation) is to add an additional term to the potential energy equal to:

$$V_{eff}(R) = V(R) + \lambda \sum_{i=1}^{n} [\nabla_i \phi(R) + \mathbf{A}(\mathbf{r}_i)]^2 \tag{13}$$

where $\phi(R)$ is the phase and $\mathbf{A}$ the vector potential. If the phase is exact, the exact energy is obtained even if the trial modulus was not exact. Otherwise, the best upper bound over all functions with that phase is found. Applications to quantum Hall systems are discussed in ref. [13]. An application to a vortex in superfluid helium is discussed in ref. [14]

Exact Fermion and Excited State Methods

As accurate as the FN method might be, it is still unsatisfactory since one does not know how the assumed nodal structure will affect the final result. One might guess that long-range properties, such as the existence or non-existence of a fermi surface will be determined by the assumed nodes. The FN algorithm only improves the bosonic correlations of the trial function, and may not change the genuine fermion features. There are some fairly simple ways of improving on the FN method, but their use is limited to small systems, though by small it may be possible to do rather accurate "exact calculations" of fifty or more particles. We explain this limitation below.

The transient estimate (TE) method calculates the ratio:

$$E_{TE}(t) = \frac{\int \Psi \mathcal{H} e^{-t(\mathcal{H}-E_T)} \Psi}{\int \Psi e^{-t(\mathcal{H}-E_T)} \Psi} \tag{14}$$

where $\mathcal{H}$ is the exact Hamiltonian (not the fixed-node Hamiltonian) and Ψ is an antisymmetric trial function. Clearly the variational theorem applies so that $E_{TE}(t) \geq E_0$. In fact you can show that the energy converges exponentially fast and monotonically to the exact energy as a function of projection time:

$$\lim_{t\to\infty} E_{TE}(t) = E_0 + \mathcal{O}(e^{-tE_g}). \tag{15}$$

Here E_g is the gap to the next excited state with the same quantum numbers as the fermion ground state. In a Fermi liquid, this is the gap to the state with the same momentum, parity and spin obtained by making 2 particle-hole excitations.

Since we do not know precisely where the nodes are, the walks must be able to go everywhere in configuration space and so the drift term in Eq. (8) must not diverge at the nodes. Hence we must distinguish between the antisymmetric trial function that is used to calculate the energy, $\Psi(R)$, (this is always assumed to be our best variational function) and a strictly positive guide function, $\Psi_G(R)$, used to guide the walks. The guide function appears in the drift and branching terms of Eq. (8) and will be assumed to be a reasonable boson ground state trial function, while the trial function appears in Eq. (14). The Ψ_G importance-sampled Green's function is:

$$\tilde{G}(R,R';t) = \Psi_G(R) < R|e^{-t(\mathcal{H}-E_T)}|R' > \Psi_G^{-1}(R'), \tag{16}$$

and we can rewrite Eq. (14) as:

$$E_{TE}(t) = \frac{\int \sigma(R)E_{LT}(R)\tilde{G}(R,R';t)\sigma(R')\Psi_G^2(R')}{\int \sigma(R)\tilde{G}(R,R';t)\sigma(R')\Psi_G^2(R')} \tag{17}$$

where $\sigma(R) = \Psi(R)/\Psi_G(R)$ and $E_{LT}(R)$ is the local energy of Ψ. In the limit, $\Psi_G \to |\Psi|$, $\sigma(R)$ equals the sign of the trial function at the point R.

The transient estimate algorithm is:

1. Sample configuration R' from the square of the guide function with VMC. That corresponds to the right most factor in Eq. (17).

2. Record the initial weight of the walk, $\sigma(R')$.

3. Propagate the walk forward an amount of time, t with the Green's function, $\tilde{G}(R,R';t)$ by taking many sufficiently many time steps. If a branch occurs, each branch will count separately.

4. The total contribution of the walk arriving at R is $\sigma(R)\sigma(R')$. The energy at projection time t is:

$$E_{TE}(t) = \frac{< [E_{LT}(R) + E_{LT}(R')]\sigma(R)\sigma(R') >}{2 < \sigma(R)\sigma(R') >}, \tag{18}$$

   where the averages are over all random walks generated by this process.

We see that the contribution of the walk is positive if the walk crosses an even number of nodes (or does not cross at all) and is negative if it crosses once or an odd number of times. This is how the nodes of the true wave function can differ from those of the trial function because of an unequal diffusion of walks from the negative to positive regions at particular points on the nodes.

The release node (RN) algorithm[10, 15] is an improvement on this TE method for ground state fermion calculations. Instead of starting the projection from the trial function, one begins the projection from the fixed-node solution. There are several advantages. First, boson correlation within the fixed-nodes is already optimized, thus the projection time is only determined by the time to adjust the position of the nodes. Second, one can directly calculate the difference between the exact result and the fixed-node solution. It turns out that this is given by the local energy of walks as they cross the nodes. Thus the difference is obtained with more statistical accuracy than either energy alone which allows the convergence to be carefully monitored. Finally, the release node method can be conveniently integrated into a fixed-node program. The only modifications are to introduce a guide function, and to keep track of the energy as a function of time since nodal crossing.

However, there are serious problems with both the TE and RN method. Let us examine how the statistical error of Eq. (14) depends on the projection time. It is not hard to see that the value of both the numerator and denominator are asymptotically proportional to $\exp(-t(E_F - E_T))$. Thus to keep the normalization fixed the trial energy must be equal to E_F. But, because the guide function allows the walks to cross the nodes, the population will increase as $\exp(-t(E_B - E_T))$ where E_B is the boson energy. From this, one can demonstrate that the signal-to-noise ratio vanishes exponentially fast. This is a general result. In any fermion scheme, as soon as negative weights are introduced the statistical error will grow as:

$$\epsilon_{stat} \propto e^{-t(E_F - E_B)}. \tag{19}$$

The behavior is physically easy to understand. Our estimator depends on finding differences between random walks crossing an even or an odd number of times. As soon as there is substantial mixing, the difference becomes harder and harder to see. Note that the exponential growth rate depends on a total energy difference. This implies that the transient estimate algorithm is guaranteed to fail if N is sufficiently large; the statistical errors will be too large. Nonetheless reliable results have been obtained for systems of 54 fermions.

The convergence problem is actually a bit more subtle since the projection time, t, can be optimized. The projection time should be chosen to give approximately equal statistical errors and systematic errors coming from non-convergence of the projection. Taking these errors from eqs. (15,19) we find the total error will decrease as:

$$\epsilon \quad \propto \quad P^{-\eta} \qquad \eta = \frac{E_g}{2(E_F - E_B + E_g)}. \tag{20}$$

where P is the total number of steps in the random walk. Only for bosons will $\eta = 1/2$. Any excited state will converge at a slower rate. Note that $\eta \propto 1/N$ for a fermion system. Inverting this relation, we find that the computer time needed to achieve a given error will increase exponentially with N.

One possibility for improving this convergence is to use all of the information given in the function, $E_{TE}(t)$, rather then just the value of the energy at the largest time. Crudely speaking, we can fit this function with a sum of exponentials and thereby try to extract the asymptotic limit. This "inverse Laplace transform" problem is well-known to be numerically unstable. It has been suggested[21] in the context of Quantum Monte Carlo for lattice models that the proper way to perform such a function fit is with the maximum entropy statistical method, wherein a model of the expected density of states is used to bias the result, thereby regularizing the fitting problem. We[16] have applied these ideas to the TE and RN methods on simple problems and shown that they do indeed reduce the statistical and systematic errors.

There have been many attempts to "solve" the fermion sign problem. For example, one can try to pair positive and negative random walks in the TE method. This is difficult in many dimensions simply because the volume of phase space is so large than random walks rarely approach each other and no such schemes have yet succeeded for more than a few particles.

There is some confusion about the nature of the "fermion" or "sign" problem. Both the TE and RN methods do converge to the exact fermion energy. A proper statement of the fermion sign problem is in terms of complexity: how long does it take to achieve a given error estimate and how does the computer time scale with the number of fermions? In the TE method, the computer time to reach a given precision grows exponentially with the number of fermions. I would say that a solution of the fermion problem would be an approximation free algorithm which scales as some low power of the number of fermions.

Properties of classical systems can be simulated in time $\mathcal{O}(N)$. Simulations of equilibrium properties of quantum bosons at zero or non-zero temperature are also $\mathcal{O}(N)$. A Heisenberg model on a bipartite lattice, or any 1D fermion system is $\mathcal{O}(N)$. VMC calculations of fermion systems are $\mathcal{O}(N^3)$ in general, but the exponent would be smaller if localized spin-orbits are used. The Hubbard model at half filling on

a bipartite lattice[17] is $\mathcal{O}(N^3)$ using the projection Monte Carlo method and auxiliary field techniques. This is the only non-trivial fermion problem solved. Known algorithms for general fermion systems are $\mathcal{O}(e^{\kappa N})$. Barring a breakthrough, one can still reduce the rate of exponential growth, κ, or use the TE or RN methods to gain confidence in FN and VMC calculations of much larger systems.

Excited States: CFMC

We have discussed to this point only calculations of ground state properties, or more correctly ground states with a given symmetry. Those are the states where one can use VMC and the FNA. If we try to apply either VMC or FN-DMC to excited states, one must always keep the wavefunction orthogonal to lower states. In the case of MC calculations, the lower states may not be explicitly known. However the MacDonald theorem provides a way both to minimize the energy and to keep all states orthogonal. It is a generalization of the TE method that is capable of calculating a spectrum of excited state properties from a single random walk.

The Correlation Function Quantum Monte Carlo method (CFMC) introduced by Ceperley and Bernu [18] starts with a basis of m trial functions, hopefully having a strong overlap with the first m exact eigenfunctions. MacDonald's theorem says that in any finite basis, if both the Hamiltonian and the overlap matrices are simultaneously diagonalized the resulting nth eigenvalue is an upper bound to the nth exact state. This is commonly used in SCF method in chemistry, for example. In QMC we use the Green's function $exp(-t\mathcal{H})$ to make the original basis much better.

$$\tilde{f_i}(t) = exp(-t\mathcal{H})f_i \tag{21}$$

One then uses a single DMC trajectory to calculate the Hamiltonian and overlap matrices of the projected basis as a function of projection time. The name arises because all of the excited state properties are determined from the correlation function of the basis sets along the imaginary time trajectory. At zero projection time what enters are equal time correlation; the method is a generalization of the variational Monte Carlo method.

Having obtained these matrices, one solves the generalized eigenvalue problem for these matrices. It can be shown that the resulting eigenvalues approach the exact eigenvalues from above, monotonically and exponentially fast. However the statistical noise is also growing exponentially fast, just as it would for the TE method (of which CFMC is a generalization.) The method has been applied to small molecules [19] and the excitations of the two dimensional electron gas[20]. It is possible to use the excited state CFMC method to get exact results for systems in strong magnetic fields or for when the wavefunction is complex-valued.

General Features of the Projection Method

Let me briefly summarize some of the strengths and weakness of the projection method. The fixed-node result is guaranteed to be closer to the exact answer than the starting variational trial function. Since the FN algorithm automatically includes bosonic correlation, the results are much less likely to have a systematic bias than does VMC results. There is also the possibility of new physics coming out of the simulation. For example, one may observe a particular type of correlation completely absent from the trial function. It is always good to pay close attention to correlation functions computed by DMC since they are a good way of monitoring the quality of a trial function. But DMC is slower than VMC because the timestep needs to be smaller to have a good approximation to the Green's function. The cost in computer time is typically a factor of 2 to 10.

Although the projected probability distribution converges to the exact answer mathematically, in practice, this does not always occur in a finite length simulation of a many-body systems. The situation

is similar to that of a classical simulation near a phase boundary. Metastable states exist and can have a very long lifetime. In addition the importance sampling biases the result. If the trial function describes a localized solid, even after complete convergence, the correlation functions will show solid-like behavior. Careful observation will reveal liquid-like fluctuations indicating the presence of the other state. The ability to perform simulations in a metastable state is useful but the results must be interpreted with caution. Although the fixed-node approximation dramatically improves energies, other properties, such as the momentum distribution may not be improved. To explore the metal-insulator phase transition with FN-DMC, one must come up with a sequence of nodes spanning the transition and use the upper bound property of the fixed-node approximation. In both VMC and DMC there is a premium for good trial functions; that is the most straightforward way of making progress in solving the many-fermion problem.

Importance sampling is only a partial cure to the unbounded fluctuations of the branching method. As N increases, sooner or later the branching becomes uncontrollable. Most projector Monte Carlo calculations have fewer than several hundred fermions. The finite temperature Path Integral Monte Carlo based on the Metropolis method does not suffer from the problem of uncontrolled branching. Release node calculations only improve the nodes locally. If t is the release node projection time, then we can move the nodes a distance of at most $\sqrt{6N\lambda t}$. One expects that global properties of the nodes will take a much longer time to be fixed by the projection operator.

The projector methods can only calculate energies exactly. For all other properties one must extrapolate out the effect of the importance sampling. This is a real problem if one is interested in obtaining asymptotic behavior of correlation functions. There are ways of getting around some of these problems but none are totally satisfactory. The Path Integral finite temperature methods are much superior to Projector Monte Carlo for calculating correlation functions.

References

[1] N. Trivedi and D. M. Ceperley, Phys. Rev. **B41**, 4552 (1990).

[2] Trotter, H. F., 1959, Proc. Am. Math. Soc. **10**, 545.

[3] M. D. Donsker and M. Kac, J. Res. Natl. Bur. Stan. **44**, 551(1950).

[4] M. H. Kalos, D. Levesque and L. Verlet, Phys. Rev. A **9**, 2178 (1974).

[5] D. M. Ceperley, B. J. Alder, Phys. Rev. Letts. **45**, (1980).

[6] D. M. Ceperley and M. H. Kalos, *Monte Carlo Methods in Statistical Physics*, ed. K. Binder, Springer-Verlag (1979).

[7] Reynolds, P. J., Ceperley, D. M., Alder, B. J. and Lester, W. A., J. Chem. Phys. **77**, 5593 (1982).

[8] C.J. Umrigar, M.P. Nightingale, and K.J. Runge, J. Chem. Phys. **99**, 2865 (1993)

[9] J. B. Anderson, J. Chem. Phys. **63**, 1499 (1975); J. Chem. Phys. **65**, 4122 (1976).

[10] D. M. Ceperley, *Recent Progress in Many-Body Theories*, ed. J. Zabolitzky, Springer-Verlag (1981).

[11] D. M. Ceperley, J. Stat. Phys. **63**, 1237 (1991).

[12] Zhu, X. and S. G. Louie, Phys. Rev. Lett. **70**, 335 (1993).

[13] Ortiz, G., Ceperley, D. M. and Martin R. M., "2-D Fermions in a Magnetic Field:The Fixed-Phase Method" Phys. Rev. Letts. **71**, 2777 (1993).

[14] Ortiz, G. and D. M. Ceperley, Phys. Rev. Letts. **75**, 4642 (1995).

[15] D. M. Ceperley and B. J. Alder, J. Chem. Phys. **81**,5833 (1984).

[16] M. Caffarel and D. Ceperley, J. Chem. Phys. **97**, 8415 (1992).

[17] J. Hirsch, Phys. Rev. **B28**, 4059 (1983); Phys. Rev. Letts. **51**, 1900 (1983); Phys. Rev. **B31**, 4403 (1985).

[18] Ceperley, D. M. and Bernu, B., J. Chem. Phys. **89**, 6316 (1988).

[19] B. Bernu, D. M. Ceperley and W. A. Lester Jr., J. Chem. Phys. **93**, 552 (1990).

[20] Kwon, Y., Ceperley, D. M. and Martin, R. M., Phys. Rev. **B53**, 7376 (1996).

[21] R. N. Silver, D. S. Sivia, and J. E. Gubernatis, Phys. Rev. **B41**, 2380 (1990).

[22] M. H. Kalos, Phys. Rev. **128**, 1791 (1962); J. Comp. Phys. **2**, 257 (1967).

©D. M. Ceperley, 2000.