Green's function Monte Carlo method with exact imaginary-time propagation

K. E. Schmidt, Parhat Niyaz, and A. Vaught
Department of Physics and Astronomy, Arizona State University, Tempe, Arizona 85287, USA

Michael A. Lee
Department of Physics, Kent State University, Kent, Ohio 44242, USA

(Received 19 December 2001; revised manuscript received 16 October 2003; published 24 January 2005)

We present a general formulation of the Green's function Monte Carlo method in imaginary-time quantum Monte Carlo which employs exact propagators. This algorithm has no time-step errors and is obtained by minimal modifications of the time-independent Green's function Monte Carlo method. We describe how the method can be applied to the many-body Schrödinger equation, lattice Hamiltonians, and simple field theories. Our modification of the Green's function Monte Carlo algorithm is applied to the ground state of liquid ^{4}He. We calculate the zero-temperature imaginary-time diffusion constant and relate that to the effective mass of a mass-four ''impurity'' atom in liquid ^{4}He.

DOI: 10.1103/PhysRevE.71.016707 PACS number(s): 02.70.Ss, 05.30.−d, 67.20.+k

I. INTRODUCTION

The first effective quantum Monte Carlo method, the Green's function Monte Carlo (GFMC) method as developed by Kalos [1,2], can calculate the exact ground state of a many-body system. It has been applied to a wide class of many-body problems that include many-boson systems, quantum lattice Hamiltonians, and simple field theories [3]. It is generally applicable to many-body problems where the Green's function is positive definite and can be used as a probability density that produces the random walks in the simulation.

Since Kalos' original work, several alternative quantum simulation methods have been developed. Here we will compare with one popular alternative, the imaginary-time Green's function Monte Carlo method as developed by Anderson, Kalos, Ceperley, and others [4–6]. For calculating response functions [8–12] it is convenient, but by no means necessary, to use the imaginary-time propagator, $\exp(-Ht)$, where H is the Hamiltonian and t is the imaginary time. The imaginary-time GFMC is able to carry out ground state simulations that solve the time-dependent Schrödinger equation in imaginary time. The imaginary-time GFMC method has the advantage of being more intuitive but it is generally implemented approximately. These approximations have been addressed by many authors, but, while exact [13,14] and nearly exact [15–19] variations have been developed, they have not enjoyed wide use because of their increased complexity.

This paper presents an exact algorithm for imaginary-time propagation that is theoretically and computationally a simple extension of the Kalos time-independent GFMC method [2]. This algorithm retains the advantages of the time-independent formulation in that it does not suffer the ills of the approximate time evolution propagators, e.g., time-step errors, node crossing, and singular behavior with coulomb and other divergent potentials. This paper will focus on the formal relationship between time-dependent and time-independent exact propagators and the construction and implementation of our propagator.

In Sec. II, we give an overview of Green's function Monte Carlo methods and then discuss the general formulation that circumvents approximations. We show how the exact imaginary-time propagator can be implemented by a slight modification of the method described by Kalos, Levesque, and Verlet (KLV) [2] for the exact time-independent propagator. We show how the method can be applied to a lattice model in Sec. III and compare our implementation to quantum simulations of the spin-$\frac{1}{2}$ Heisenberg antiferromagnet [21]. In Sec. IV we formulate the exact imaginary-time method for a simple scalar field theory and relate it to the method of Carlson [22]. We describe a coordinate space implementation appropriate to the many-body Schrödinger equation in Sec. V. Finally, in Sec. VI we use our method to calculate the zero-temperature imaginary-time diffusion constant of liquid helium and relate this to the effective mass of mass-four ''impurity'' atoms.

II. THE GREEN'S FUNCTION MONTE CARLO METHOD

Here we review the Green's function Monte Carlo method, first showing the formal relationships between the various propagators and then describing our implementation. We include in our discussion the class of Monte Carlo implementations of iterative schemes to calculate the ground state for a Hamiltonian, particularly time-dependent, imaginary-time, and pseudotime formulations.

A. Formal relationships among propagators

The Schrödinger equation in imaginary time is

$$H\psi(t) = -\frac{\partial \psi(t)}{\partial t}. \tag{1}$$

(Here, and in all equations that follow, we have assumed time units in which Planck's constant $\hbar$ is unity.) The absence of the usual $\sqrt{-1}$ in Eq. (1) makes it a diffusion equation in a many-dimensional space. In the diffusion analogy, ψ is a time-dependent density that evolves in time as pre-

scribed by the imaginary-time Green's function or propagator. The time evolution is governed by the imaginary-time propagator

$$G(t)=\exp[-(H-E_T)t], \tag{2}$$

where E_T is a trial energy, eventually set to the calculated ground state energy.

The imaginary-time Green's function Monte Carlo method numerically iterates the equation

$$\psi(t+\Delta t)=e^{-(H-E_T)\Delta t}\psi(t), \tag{3}$$

to obtain $\psi(t\rightarrow\infty)$ which converges to the lowest energy state contained in $\psi(t=0)$ (normally the ground state). Formally, convergence occurs because the amplitude of any excited state component of energy E in the initial wave function is diminished by the factor

$$e^{-(E-E_T)\Delta t} \tag{4}$$

after each time step Δt.

Propagators other than Eq. (2) have been used to obtain the ground state. For small Δt, the linear term in the expansion of the propagator in Eq. (2),

$$1-(H-E_T)\Delta t, \tag{5}$$

can be employed. This propagator is only useful if the spectrum of H is bounded above as well as below, i.e., for a finite basis such as those encountered in finite lattice problems [23–26]. To ensure the positive definite character of the propagator, the off diagonal elements of the Hamiltonian must be negative and one must choose Δt small enough so that the second term in Eq. (5) is less than unity. As in Eq. (4), application of this operator means that the amplitude of any excited state component of energy E in the initial wave function is diminished by the factor $1-(E-E_T)\Delta t$ after each time step Δt. This method is equivalent to the well-known "power method" [27] used to find the largest eigenvalue/eigenvector of a matrix. The constant E_T and the factor Δt serve to shift and scale the eigenvalue spectrum so that the ground state of the system has the largest magnitude eigenvalue of Eq. (5).

An operator which avoids the problem of the small time step needed by Eq. (5) is

$$1/[1+(H-E_T)\Delta t]. \tag{6}$$

This propagator and the propagator of Eq. (5) are equal to the imaginary-time Green's function to lowest order in Δt. The iteration of all three propagators will converge to the true ground state, but only for the full imaginary-time Green's function does the variable "t" have the significance of (imaginary) time.

The KLV method [2] solves the time-independent Schrödinger equation

$$H\psi=E\psi, \tag{7}$$

and is also formally analogous to the power method. This requires that a constant E_C be added to H to make the eigenvalue spectrum positive definite. The largest eigenvalue of the Green's function (the inverse of the shifted Hamiltonian) is the inverse of the shifted ground state energy, and the power method applies. An eigenstate of H of energy E then satisfies the Green's function equation,

$$\frac{1}{E_C+E}\psi=\frac{1}{E_C+H}\psi\equiv G_{KLV}\psi. \tag{8}$$

The rightmost equality in this equation defines the KLV Green's function.

From G_{KLV} one forms the KLV propagator equation to be iteratively applied to an approximate ground state,

$$\psi=(E_C+E_T)G_{KLV}\psi=\frac{E_C+E_T}{E_C+H}\psi, \tag{9}$$

with a trial energy E_T adjusted to equal the true ground state energy E_0. With each iteration, the amplitude of an excited state component of energy E is diminishes by a factor $(E_C+E_T)/(E_C+E)$.

One can compare the KLV propagator, Eq. (9), and the short-time approximation to the imaginary-time propagator, Eq. (6). The KLV propagator can be written as

$$\frac{E_C+E_T}{E_C+H}=[1+(H-E_T)\Delta t]^{-1}. \tag{10}$$

The equality determines the relationship between the constant energy E_C and the "time step," $E_C+E_T=1/\Delta t$.

The KLV propagator, Eq. (9), can also be written in terms of the full imaginary-time propagator, Eq. (2), as

$$\begin{aligned}(E_C+E_T)G_{KLV}&=[1+(H-E_T)\Delta t]^{-1}\\&=\frac{1}{\Delta t}\int_0^\infty dt e^{-(H-E_T)t}e^{-t/\Delta t}.\end{aligned} \tag{11}$$

The KLV method uses this equation to sample the propagator G_{KLV} of Eq. (8). The right-hand side of Eq. (11) is applied to a wave function and the integral over time is performed by sampling a time. The propagation of the single point is then carried out for that sampled time and, when repeated for many points, provides a correct statistical average of the integral over all times, i.e., the left-hand side of Eq. (11).

The above analysis then reduces to the simple statement that when the KLV method makes a Monte Carlo move, it is made with the imaginary-time propagator for a time randomly sampled from $\exp(-t/\Delta t)/\Delta t$. We have glossed over the details (see KLV [2] and Sec. V), but this observation points us to the minor modification of the KLV method required to exactly sample the imaginary-time propagator. We need only enforce a specific time step in each propagation to change the algorithm to provide an exact sampling of the imaginary-time propagator.

B. Iterative equation for the imaginary-time propagator

For most problems, the exact imaginary-time propagator is not known in an analytic form, but a procedure for exactly sampling it can be developed. An integral equation for the full imaginary-time Green's function can be formed if a Green's function $G^{(U)}(t)$ is known for a Hamiltonian $H^{(U)}(t)$ [13,14]. ($H^{(U)}$ is often taken independent of t.) The known Green's function satisfies

$$(H^{(U)}(t)-E_T)G^{(U)}(t)=-\frac{\partial G^{(U)}(t)}{\partial t}, \tag{12}$$

and will be the starting point for sampling the exact Green's function. While one usually chooses $H^{(U)}$ to be an approximation to the Hamiltonian in question and solves Eq. (12) for $G^{(U)}$, in principle, one could choose $G^{(U)}$ and use Eq. (12) as defining $H^{(U)}$. An approach of this sort has been followed by Ceperley [20], producing a generalization of the KLV algorithm.

The construction of the exact Green's function begins with the identity

$$e^{-(H-E_T)t}=G(t)=G^{(U)}(t)-\int_0^t dt'\frac{\partial}{\partial t'}[G(t-t')G^{(U)}(t')], \tag{13}$$

where we have used the condition that $G(0)=G^{(U)}(0)=I$, the identity operator. Substituting the known time derivatives from Eqs. (2) and (12) into Eq. (13) yields

$$G(t)=G^{(U)}(t)+\int_0^t dt' G(t-t')[H^{(U)}(t')-H]G^{(U)}(t'). \tag{14}$$

This is the formal representation of the imaginary-time Green's function which will be the basis of the procedures we develop in the remainder of this paper. This equation is well known in real-time scattering theory [28] and was previously employed in a formulation of a generalized imaginary-time diffusion Monte Carlo method [13,14]. It is the exact implementation of the relation in imaginary-time GFMC with a fixed time step that is the substance of this paper.

Variations of the above formalism have been implemented for both time-independent [20] and time-dependent GFMC [1,29], but it can be made far more efficient when statistical fluctuations in the simulation are reduced by importance sampling. Importance sampling is a general technique for reducing statistical fluctuations [30], described for quantum simulation by KLV. Importance sampling and the technicalities involved will be discussed in detail in the following section and Sec. V.

C. Importance sampling

Importance sampling utilizes an approximate form of the ground state wave function to significantly improve the efficiency of the Monte Carlo calculation. In liquid ^{4}He, for example, we have a very good understanding of the two- and three-body correlations in the ground state [31–34].

In this section we formulate importance sampling and the implementation of the exact imaginary-time algorithm on a problem with a finite basis. While a finite basis is usually encountered in lattice problems, this representation in no way restricts the generality of what we will present. Section V contains the continuous space formulation.

In the finite basis, the matrix elements of the Hamiltonian and Green's functions are finite matrices, H_{ij}, $G_{ij}(t)$, and $G^{(U)}_{ij}(t)$, and the wave function is a finite vector, $\psi_i(t)$. The fundamental equations, Eqs. (3) and (14), become

$$G_{ij}(t)=G^{(U)}_{ij}(t)+\int_0^t dt'\sum_{kl} G_{ik}(t-t')\times[H^{(U)}_{kl}(t')-H_{kl}]G^{(U)}_{lj}(t'), \tag{15}$$

$$\psi_n(t+\Delta t)=\sum_m G_{nm}(\Delta t)\psi_m(t). \tag{16}$$

Given some analytic approximate ground state wave function, with a real amplitude, $\psi_T(m)$, on the mth basis state, we use it to define a new set of propagators and densities in the usual way by writing

$$\tilde{\psi}_m(t)=\psi_T(m)\psi_m(t),$$

$$\tilde{H}_{nm}=\frac{\psi_T(n)H_{nm}}{\psi_T(m)},$$

$$\tilde{G}^{(U)}_{nm}(t)=\frac{\psi_T(n)G_{nm}(t)}{\psi_T(m)},$$

$$\tilde{G}_{nm}(t)=\frac{\psi_T(n)G_{nm}(t)}{\psi_T(m)}. \tag{17}$$

The importance sampled propagator $\tilde{G}$ defines a new random walk with a new density, $\tilde{\psi}_m(t)=\psi_T(m)\psi_m(t)$. Because the new density converges to the product of the trial wave function and the exact wave function, it is nearly equal to the true quantum probability density. This is only one of the advantages of importance sampling. Substituting these definitions into our fundamental equations, Eqs. (15) and (16), shows that the importance sampled equations are simply given by writing a tilde over all quantities. These equations,

$$\tilde{G}_{ij}(t)=\tilde{G}^{(U)}_{ij}(t)+\int_0^t dt'\sum_{kl}\tilde{G}_{ik}(t-t')\times[\tilde{H}^{(U)}_{kl}(t')-\tilde{H}_{kl}]\tilde{G}^{(U)}_{lj}(t'), \tag{18}$$

$$\tilde{\psi}_n(t+\Delta t)=\sum_m \tilde{G}_{nm}(\Delta t)\tilde{\psi}_m(t), \tag{19}$$

are used in the numerical implementation in the following sections.

Carrying out an iteration based on Eq. (19) is simple if we suppose for a moment that $\tilde{G}_{nm}$ is known. Initially (and at every iteration of the calculation) one has an ensemble of points called ''walkers'' distributed over the basis states with a probability proportional to $\tilde{\psi}_m$. Each walker is moved from a state m to a state n with a probability proportional to $\tilde{G}_{nm}$. While $\tilde{G}_{nm}$ is positive definite by construction, it is not normalized in the sense that summing over all possible final positions does not necessarily yield unity,

$$0 \leqslant \sum_n \tilde{G}_{nm} \not\equiv 1. \tag{20}$$

As a result, unlike classical Monte Carlo, the number of walkers in the ensemble may change during the iteration.

For a walker moving from a state m, if the sum in Eq. (20) is less than unity, then the sum is the probability that the walker will not be eliminated from the simulation. If the sum is greater than one, the algorithm must produce, on an average, more than one walker. The exact number of resulting walkers at the new positions may vary from one iteration to the next, but must, on average, equal the magnitude of the sum in Eq. (20). Only after many iterations, when the ensemble representing $\tilde{\psi}(t)$ for large t has converged and E_T has been adjusted to be the ground state energy, will the total number of walkers in the ensemble be constant on the average.

Improvement in the efficiency of the calculation with importance sampling results from an enormous decrease in statistical fluctuations of the number of walkers or in the density representing the ground state accompanied by a large decrease in the statistical fluctuations of expectation values. In practice, the calculation may be quite close to the ideal case where there are no population fluctuations. The algorithm is designed to make use of the information in the importance function and we therefore begin by considering the case of perfect importance sampling, when the exact ground state is used as the trial wave function.

1. Ideal importance sampling

With perfect importance sampling one would have $\psi_T = \psi_0$ and $E_T = E_0$ where ψ_0 and E_0 are the exact ground state wave function and energy. While, we do not know these quantities, we design the algorithm so that there are no population fluctuations in this limit.

In the process iterating Eq. (16), the GFMC repeatedly moves an ensemble consisting of a list of walkers (coordinates) that are sampled from $\tilde{\psi}(t)$ to new positions sampled from $\tilde{\psi}(t+\Delta t)$. The expected number of walkers resulting from an initial walker at state m is given by the sum in Eq. (20) which, with ψ_T replaced by ψ_0 in Eq. (17), becomes unity,

$$\sum_n \tilde{G}_{nm}(t) = \sum_n \frac{\psi_0(n) G_{nm}(t)}{\psi_0(m)} = 1. \tag{21}$$

This follows from the definition of $\tilde{G}$, the hermiticity of G_{nm}, and the fact that the time-independent ψ_0 satisfies Eq. (16). Unit normalization means each input walker will produce exactly one output walker when perfect importance sampling is used.

To sample $\tilde{G}_{ij}(t)$, we mean that we have a walker at site j and we wish to select a site i from all possible final sites so that the probability of selecting a particular i is proportional to $\tilde{G}_{ij}(t)$. We use Eq. (13) to expand $\tilde{G}(t)$ in terms of the known $\tilde{G}^{(U)}$,

$$\tilde{G}_{ij}(t) = \tilde{G}^{(U)}_{ij}(t) + \int_0^t dt' \left(-\frac{\partial}{\partial t'} \left[\sum_k \tilde{G}_{ik}(t-t') \tilde{G}^{(U)}_{kj}(t') \right] \right). \tag{22}$$

The unit normalization of perfect importance sampling from Eq. (21) tells us that for any site j, the two terms on the right of Eq. (22) have normalizations that add to unity. We may thus sample the site i from either term if we first select that term with a probability proportional to its norm. The probability of selecting the first term is $\Sigma_i \tilde{G}^{(U)}_{ij}(t)$, and the probability of selecting the second term is just one minus this.

If the first term is chosen to be sampled, then a site i is sampled from $\tilde{G}^{(U)}_{ij}(t)$ and the process of moving one walker one time step t has been completed. If the second term is selected, then we proceed by recognizing that the integrand is positive definite and can be normalized to form a joint probability distribution for the variables i and $t' \in [0,t]$. Since it is a joint probability distribution function in two variables, we can write it as the product of a marginal probability distribution function for t' and the conditional probability distribution function for i, given t'. We shall construct these probability distributions explicitly in the following section.

The probability of sampling the second term in Eq. (22) is also the sum over final states, i. The expected number of walkers, using perfect importance sampling, produced by the second term for a value of t' in an interval dt' is

$$dt' \left(-\frac{\partial}{\partial t'} \sum_k \tilde{G}^{(U)}_{kj}(t') \right), \tag{23}$$

since the sum over i in the second term of Eq. (22) eliminates the $\tilde{G}$ by using Eq. (21).

If we sample a time t' from the probability distribution of Eq. (23) over the interval $0 < t' < \infty$, the probability that t is shorter than t' is just the probability of picking the first term. The probability that t is longer than t' is the probability of picking the second term. In other words, we may choose to have the walker on site j make a time step t to a site i selected from the first term or the second term of Eq. (21) depending on whether a time t' sampled from Eq. (23) is shorter or longer than the time step t.

This structure has a simple physical interpretation. The time t is the time the walker must propagate to finish the time step. We sample a time t' which is the time for the walker to

"scatter" corresponding to the difference between $\tilde{G}$ and $\tilde{G}^{(U)}$. If the time to finish the time step is smaller, then the walker does not scatter and the walk is completed by sampling from $\tilde{G}^{(U)}$. If, however, the time t' is smallest, then the scattering takes place before the walker completes its time step. The second term must then be sampled. For this to work we require that Eq. (23) defines a normalized probability density on $0<t'<\infty$, i.e.,

$$G^{(U)}(t'\to\infty)=0. \tag{24}$$

For a time-independent $H^{(U)}$, this means that the ground state energy of $H^{(U)}$ is higher than E_T. Once a term is selected, we sample the states of the walkers with probability again given by the expected contribution of that state to the number of walkers. We note that the asymptotic requirement in Eq. (24) is just a matter of convenience to simplify the sampling process because the value $\tilde{G}^{(U)}(t')$ for $t'>t$ is never actually needed.

2. Nonideal importance sampling

In actual implementations, one must deal with a trial wave function that approximates the ground state. Having chosen a trial function, the random walk and resulting density are given by exactly implementing Eq. (18). Minimizing the variance, for a given trial function, means minimizing the fluctuations in the population of walkers and this characterizes the optimal random walk algorithm.

We will describe two implementations of the finite time-step algorithm, both of which are exact in the sense that they sample the exact Green's function correctly on the average. We will compare these methods to related implementations of the imaginary-time and time-independent algorithms, and discuss why one or the other may be considered more nearly "optimal" for nonideal importance sampling.

We wish to sample a new location i for a walker at j from $\tilde{G}_{ij}(t)$ using Eq. (22), or equivalently Eq. (16). We assume that we have constructed the time independent ψ_T and $\tilde{H}^{(U)}$, and are able to sample a position i from $\tilde{G}^{(U)}_{ij}(t')$ for any time $t'\leqslant t$. These assumptions are common to both algorithms.

With nonideal importance sampling, we lose the unit normalization of $\tilde{G}_{ij}(t)$ that was the basis of Eq. (21) and the sampling method described in Sec. II C 1. This is compensated for by weighted sampling techniques (described below) which produce an average number of walkers equal to the correct normalization.

The two algorithms we will consider differ primarily in the manner in which a new position i is to be sampled from $\tilde{G}^{(U)}_{ij}$ or the correction term in Eq. (16). The first algorithm adopts the approach that, for a good importance function, the relative weights of these two choices should be approximately given by the same condition of whether t' sampled from Eq. (23) is less than or greater than t. Since this method of selecting the first or second term in the $\tilde{G}_{ij}(t)$ equation is not exact, we must assign compensating weights to each term so that the expected number of walkers sampled for any site i starting from j is on the average correct, i.e., equal to the value of $\tilde{G}_{ij}(t)$.

The weighted sampling techniques that we employ are elucidated in texts on probability theory [30]. We will provide here only the importance sampled equation and the algorithm to sample the position of the walker that is propagated at time t.

To describe the first algorithm, we write the importance sampled Green's function equation as

$$\tilde{G}_{ij}(t)=\left[\frac{\tilde{G}^{(U)}_{ij}(t)}{\sum_n \tilde{G}^{(U)}_{nj}(t)}\right]_2\left[\sum_n \tilde{G}^{(U)}_{nj}(t)\right]_1+\int_0^t\sum_{kl}[\tilde{G}_{ik}(t-t')]_7\left[\frac{\sum_{mn}[\tilde{H}^{(U)}_{mn}(t')-\tilde{H}_{mn}]\tilde{G}^{(U)}_{nj}(t')}{\sum_{mn}[\tilde{H}^{(U)}_{mn}(t')-E_T\delta_{mn}]\tilde{G}^{(U)}_{nj}(t')}\right]_6$$
$$\times\left[\frac{\tilde{H}^{(U)}_{kl}(t')-\tilde{H}_{kl}}{\sum_m[\tilde{H}^{(U)}_{ml}(t')-\tilde{H}_{ml}]}\right]_5\left[\frac{\sum_m[\tilde{H}^{(U)}_{ml}(t')-\tilde{H}_{ml}]\tilde{G}^{(U)}_{lj}(t')}{\sum_{mn}[\tilde{H}^{(U)}_{mn}(t')-\tilde{H}_{mn}]\tilde{G}^{(U)}_{mn}(t')}\right]_4\left[-\frac{\partial}{\partial t'}\sum_n\tilde{G}^{(U)}_{nj}(t')dt'\right]_3, \tag{25}$$

where the subscripts on the brackets are simply labels that we can refer to in the text. The first thing to note is that Eq. (25) is identical to Eq. (22). Term 1 cancels the denominator of term 2. Term 3 cancels the denominator of term 6, where we use Eq. (12). The Hamiltonian part of the numerator of term 4 cancels with the denominator of term 5. The denominator of term 4 cancels with the numerator of term 6. This leaves the numerator of term 2 for the first term, and term 7 along with the numerator of term 6 and the $\tilde{G}^{(U)}$ in the numerator of term 4.

The algorithm to sample a new walker position from Eq. (25) is the following.

(1) For each walker at position j, which needs to be propagated, the time left to propagate is t. (Initially, t is the full time step.)

(2) Sample a time t' $(0<t'<\infty)$ from the probability

distribution of term 3. The probability that the time t' is greater than t is term 1; if the sampled $t'>t$, then sample a new state i from the normalized probability distribution given by term 2. The walker has now walked for a time t, and the weight is one. This walker has finished its walk and been propagated a full time step.

(3) If the time in step 2 is less than t, then we have sampled a time t' from term 3. We now sample the intermediate state l from the normalized probability distribution of term 4, using the time t' already sampled, to evaluate $\tilde{G}^{(U)}(t')$.

(4) Once l is sampled, the intermediate state k is sampled from term 5. The walker is now in state k, and we have taken care of all but terms 6 and 7. Term 6 is taken to be the weight of the walker. If perfect importance sampling is used, this weight is one since in that case the local energy defined by

$$E_n^{local}=\sum_m \tilde{H}_{mn} \tag{26}$$

is equal to the ground state energy E_0. However, with realistic importance functions, this term will not be identically one. This weight can be included as a branching/absorption factor, or it can be carried along with the walker.

(5) Only term 7 is left. This is the full Green's function for walking from k to some state i in time $t-t'$. To sample term 7, we simply reset the time left $t\rightarrow t-t'$ and repeat these steps. Eventually, the sampled time t' in step 2 will be greater than the time left and the walk terminates.

The above algorithm produces exactly one output walker when perfect importance sampling is used. Otherwise, the weight, term 6, compensates for the assumptions made when sampling the terms so that the exact equation is reproduced even when the importance sampling is not perfect.

Many other algorithms that sample exactly like this one in the perfect importance sampling limit will sample differently when the importance sampling is not perfect. Our second algorithm is an example of such a modification which includes a term $\exp[(E_j^{local}-E_T)t]$ in the time sampling. With perfect importance sampling, $E_j^{local}=E_T=E_0$ being independent of j, nothing is changed in that limit. In the nonideal case, the sampling process must include the appropriate compensating weights so that the algorithm remains exact. The Green's function equation can be written in the same way as before:

$$\begin{aligned}\tilde{G}_{ij}(t)=&[e^{-(E_i^{local}-E_T)t}]_3\left[\frac{\tilde{G}_{ij}^{(U)}(t)\exp[(E_i^{local}-E_T)t]}{\sum_n \tilde{G}_{nj}^{(U)}(t)\exp[(E_n^{local}-E_T)t]}\right]_2\left[\sum_n \tilde{G}_{nj}^{(U)}(t)e^{(E_n^{local}-E_T)t}\right]_1\\&+\int_0^t dt'\sum_{kl}[\tilde{G}_{ik}(t-t')]_8[e^{-(E_l^{local}-E_T)t'}]_7\left[\frac{\tilde{H}_{kl}^{(U)}(t')-\tilde{H}_{kl}}{\sum_m[\tilde{H}_{ml}^{(U)}(t')-\tilde{H}_{ml}]}\right]_6\\&\times\left[\frac{\sum_m[\tilde{H}_{ml}^{(U)}(t')-\tilde{H}_{ml}]\tilde{G}_{lj}^{(U)}(t')\exp[(E_l^{local}-E_T)t]}{\sum_{mn}[\tilde{H}_{mn}^{(U)}(t')-\tilde{H}_{mn}]\tilde{G}_{nj}^{(U)}(t')\exp[(E_n^{local}-E_T)t]}\right]_5\left[-\frac{\partial}{\partial t'}\sum_n \tilde{G}_{nj}^{(U)}(t')e^{(E_n^{local}-E_T)t'}\right]_4.\end{aligned} \tag{27}$$

The additional factors again cancel, leaving this equation analytically identical to Eq. (22).

For this second algorithm, the time t' is sampled from the probability density in term 4. We require term 4 to go to zero as $t'\rightarrow\infty$, i.e., the ground state energy of $G_{nj}^U(t')$ must be larger than $E_n^{local}-E_T$ for all n. If t' is greater than t, state i is sampled from term 2, and the walk is terminated with a walker of weight given by term 3. If t' is less than t, states l and k are sampled from terms 5 and 6 respectively, and the walk continues with a walker that needs to be propagated from k to some i with weight given by term 7.

The advantage of this formalism is that the weight term is always proportional to $\exp[-(E_i^{local}-E_T)t]$ where t is the time for i sampled from $\tilde{G}_{ij}^{(U)}$ either in the term 5 or in the term 2. In the limit that the time step goes to zero, this is the Feynman-Kac formula [35]. In both algorithms, large fluctuations in E^{local} will give a large variance. Since the algorithms are based on the assumption that a good importance function is available, both formalisms can suffer large variances if the trial function is poor.

In practice, we should select the algorithm that gives the lowest variance for a given amount of computer time. Exactly sampling terms 5 and 6 of Eq. (27) can be difficult; often approximations to these terms are used and the ratio of the correct term to the approximation is included in the weight. These approximations will increase the variance of the results even in the perfect importance sampling limit. They will not change the exact character of the algorithm.

III. A LATTICE HAMILTONIAN EXAMPLE

The spin-$\frac{1}{2}$ quantum antiferromagnet ground state has been solved using Monte Carlo methods by several groups [21,36,12]. We will add nothing to their results, however, the Hamiltonian is simple enough so that we can solve the imaginary-time equations easily. We will then relate this solution to the method using Eq. (5).

Carlson [21] and others [36] exploited the fact that the Hamiltonian can be written in terms of boson variables and the basis set can be chosen, for a square lattice, so that the ground state is positive. The standard spin Hamiltonian is

$$H = -J \sum_{\langle ij \rangle} S_i S_j , \tag{28}$$

where the sums $\langle ij \rangle$ are over nearest neighbor states and S is the spin vector of the spin-$\frac{1}{2}$ particles at each site. This is transformed [21] to a boson system on a two-dimensional lattice,

$$H = -\frac{J}{2} \sum_{\langle ij \rangle} (b_i^\dagger b_j + b_j^\dagger b_i) + J \sum_{\langle ij \rangle} n_i n_j . \tag{29}$$

This Hamiltonian describes hard-core bosons on a square lattice. The energy is the energy above the classical perfect antiferromagnet ground state energy. Boson operators $b_i^\dagger$ (b_i) create (destroy) a boson at site i. The bosons have a hard-core interaction because the number operators n_i, generated by the transformation, only take on the values 0,1, i.e., no double occupancy. The basis states are chosen to be those which have total S_z zero, since only these states are important in the ground state of the antiferromagnet. This corresponds to a half filled lattice of bosons. Different fillings of bosons correspond to different S_z.

We begin by constructing the trial wave function and importance sampled matrices needed for the first of our Green's function algorithms, Eq. (25). Several choices are available that were used in earlier simulations [21,26,36] but for the algorithm discussion we need not make a specific choice of a trial wave function.

The simplest choice for $G^{(U)}$ is to take $H^{(U)}$ to be the diagonal part of H. Then, $H_{ij}^{(U)} = H_{jj}\delta_{ij}$, and $G_{ij}^{(U)}(t) = \exp[-(H_{jj} - E_T)t]\delta_{ij}$. Since the true ground state energy is known to be negative, and H_{jj} is positive, then $G_{ij}^{(U)}(t)$ is a decaying exponential and can be normalized and used as a probability distribution in time. (For other fillings, where the exponential may not be negative, a time could be sampled from a decaying exponential and appropriate weighting factors included.)

Since $G^{(U)}$ is diagonal we can simplify Eq. (25) and the algorithm for implementing it. A diagonal $H^{(U)}$ means $\tilde{H}_{ij}^{(U)} = H_{ij}^{(U)}$ and $\tilde{G}_{ij}^{(U)} = G_{ij}^{(U)}$. This makes term 4 and term 2 in Eq. (25) simply δ_{lj}. We can rewrite Eq. (25) with these simplifications as

$$\tilde{G}_{ij}(t) = \delta_{ij}[\exp\{-(H_{jj} - E_T)t\}]_1 + \int_0^t dt' \sum_k [\tilde{G}_{ik}(t-t')]_7 \times \left[\frac{H_{jj} - E_j^{local}}{H_{jj} - E_T}\right]_6 \left[\frac{-\tilde{H}_{kj}(1-\delta_{kj})}{\sum_m (-\tilde{H}_{mj})(1-\delta_{mj})}\right]_5 \times [(H_{jj} - E_T)\exp\{-(H_{jj} - E_T)t'\}]_3 . \tag{30}$$

We have retained the labels of the terms to match those in Eq. (25). The absence of term 4 means that step 3 in the algorithm for Eq. (25) is omitted, i.e., l is set equal to j.

We may restate the algorithm as it applies to the above equation and the problem at hand. Sample a time to hop, t', from the normalized probability distribution $(H_{jj} - E_T)\exp[-(H_{jj} - E_T)t']$. If $t' > t$, then the walker has finished and remains at site j. Otherwise, sample a boson hop from site j to a site k according to the relative value of the importance function at the hopped position, i.e., sample k from the probability distribution function of term 5. The walker at site k is assigned a weight of term 6, $(H_{jj} - E_j^{local})/(H_{jj} - E_T)$. As mentioned before, the weight can be kept or included in a branching/absorption process. Finally, subtract t' from the time left for the walker to complete the time step, and repeat the process until the walk terminates.

This resulting algorithm can be related to that of Trivedi and Ceperley [36] with only minor modifications. They used the propagator of Eq. (5). As the system size becomes large, the largest eigenvalue also gets large. To keep the propagator positive, the ''time step'' in Eq. (5) must get small. In this case, the straightforward algorithm to sample Eq. (5) becomes inefficient. The straightforward algorithm simply samples the move for a walker from state j to i by sampling i from the relative value of $\delta_{ij} - (\tilde{H}_{ij} - E_T\delta_{ij})\Delta t$. As Δt becomes small, almost all of the steps will be sampled from the diagonal term. Trivedi and Ceperley therefore modified the algorithm so that they sampled the number of steps until the walker hops. If we take the limit of their time step, $\Delta t \to 0$, replace the number of steps n by the corresponding time $n\Delta t$, and propagate all the walkers to the same time t, we recover our algorithm above.

One advantage of our formalism is that it is easy to see how to expand G in terms of another Green's function $G^{(U)}$. For example, solving the spin-$\frac{1}{2}$ quantum antiferromagnet for high values of S_z corresponds to a low density of bosons. If the bosons are well separated, a $G^{(U)}$ given by solving for the propagation of each boson in a domain around its current position would be more efficient. The method would then be very much similar to that given in Sec. V for the many-particle Schrödinger equation. The simple diagonal $G^{(U)}$ can, of course, be viewed as a domain Green's function (see Sec. V) with the domain consisting of a single site. Extending the domain of $G^{(U)}$ should yield a more computationally efficient simulation.

IV. A SIMPLE FIELD THEORY

Scalar field theories can be solved using these GFMC methods. Here we show how the method developed by Carl-

son and used by Carlson and Schmidt [22] to solve the polaron and other scalar field theories can be cast in the exact imaginary-time GFMC form. The method is quite similar to that used for the lattice Hamiltonian. The main difference is that the degrees of freedom are continuous. We will present a lattice formulation in momentum space. The continuous formulation [22] is obtained in the usual way by replacing the sums over momenta by integrals.

The Hamiltonian for the scalar field theory is

$$H=\sum_{\vec{k}}\ \omega_f(k)a^\dagger(\vec{k})a(\vec{k})+\sum_{\vec{k}}\ \omega_b(k)b^\dagger(\vec{k})b(\vec{k})-\sum_{\vec{k},\vec{q}}\ \beta(q) \times[a^\dagger(\vec{k}-\vec{q})a(\vec{k})b^\dagger(\vec{q})+a^\dagger(\vec{k})a(\vec{k}-\vec{q})b(\vec{q})], \tag{31}$$

where $a(\vec{k})$ $[a^\dagger(\vec{k})]$ is the fermion destruction (creation) operator and $b(\vec{k})$ $[b^\dagger(\vec{k})]$ is the boson destruction (creation) operator for momentum $\vec{k}$. The corresponding free particle energies are $\omega_f(k)$ and $\omega_b(k)$ and the interaction, which can create or destroy bosons of momentum $\vec{q}$, has a strength given by $\beta(q)$. This interaction term corresponds to the hopping term in the lattice model, and to the kinetic energy in the real-space Schrödinger Hamiltonian. Carlson and Schmidt discuss several different forms for the free particle spectrums, $\omega_f(k)$ and $\omega_b(k)$, and the interaction $\beta(q)$ with applications to a polaron model and a model with a nucleon and scalar mesons. We will leave the particular choices of these quantities unspecified and refer the interested reader to the original work [22] for specific results.

A basis state $|\vec{K}_n\rangle\equiv|\vec{k}_1,\vec{k}_2,\ldots,\vec{k}_n,\vec{k}_f\rangle$ will specify the number of bosons, n, and the momenta of each boson, $\vec{k}_i$, and the fermion, $\vec{k}_f$. Carlson and Schmidt chose the trial wave function for one fermion and total momentum $\vec{Q}$ to be

$$|\Psi_T\rangle=P(\vec{Q})\sum_{\vec{k}_f}\ g(k_f)a^\dagger(\vec{k}_f)\exp\left[\sum_{\vec{k}}\ f(k)b^\dagger(\vec{k})\right]|0\rangle, \tag{32}$$

where $P(\vec{Q})$ projects the state onto total momentum $\vec{Q}$. They describe the variational Monte Carlo procedure used to optimize a trial wave function using standard methods to sample the density

$$|\Psi_T(\vec{K}_n)|^2=g^2(k_f)\prod_{i=1}^{n}\ f^2(k_i)/n!. \tag{33}$$

To match the algorithm described in Ref. [22], we use our second algorithm described by Eq. (27). $H^{(U)}$ must first be chosen. The simplest choice, following Carlson, is similar to the simplest lattice choice. We take $H^{(U)}$ to be diagonal in k space and boson number,

$$H^{(U)}=\sum_{\vec{k}}\ \omega_f(k)a^\dagger(\vec{k})a(\vec{k})+\sum_{\vec{k}}\ \omega_b(k)b^\dagger(\vec{k})b(\vec{k}). \tag{34}$$

The algorithm described by Eq. (27) requires the local energy, Eq. (26), $E_j^{local}=\Sigma_m\tilde{H}_{mj}$. The index j now refers to a particular basis state which we denoted above as $|K_n\rangle$ and the sum is over all other basis states of the same total momentum. We can separate E_j^{local} into three parts:

$$E_j^{diag}=\omega_f(k_f)+\sum_i\ \omega_b(k_i),$$

$$E_j^{crea}=\frac{\sum_{\vec{q}}\ \langle\Psi_T(\vec{K}_{n+1})|\beta(q)b^\dagger(\vec{q})a^\dagger(\vec{k}-\vec{q})a(\vec{k})|\Psi_T(\vec{K}_n)\rangle}{\langle\Psi_T(\vec{K}_n)|\Psi_T(\vec{K}_n)\rangle},$$

$$E_j^{dest}=\frac{\sum_{i}\ \langle\Psi_T(\vec{K}_{n-1})|\beta(q)b(\vec{q})a^\dagger(\vec{k}+\vec{q})a(\vec{k})|\Psi_T(\vec{K}_n)\rangle}{\langle\Psi_T(\vec{K}_n)|\Psi_T(\vec{K}_n)\rangle}. \tag{35}$$

The first component of E^{local} is the energy of the diagonal Hamiltonian $H^{(U)}$ for the particular walker and is simply the sum over boson and fermion single particle energies. The creation energy term, E^{crea}, requires the sum (integral) over all possible values of the created boson's momentum. The matrix element conserves momentum, so the vector $\vec{K}_{n+1}$ has the same boson momenta of the n-boson state and one additional boson of momentum q which is subtracted from the fermion momentum. When making a move to a state with one fewer boson, the sum in E^{dest} covers the momentum values of the n possible destroyed bosons. (Note that both destruction and creation energies are negative.) This can be generalized to include more than one fermion [22].

We will now construct the propagation equation based on our second algorithm, Eq. (27), and show that it is a fixed time-step version of the algorithm use in Ref. [22]. The particular choice for $H^{(U)}$, Eq. (34), means $G^{(U)}$ is diagonal in the basis set chosen. Substituting $\tilde{G}_{ij}^{(U)}=\delta_{ij}G_{jj}^{(U)}$ allows the simplification, which we need, of Eq. (27),

$$\begin{aligned}\tilde{G}_{ij}(t)=&[\exp(-(E_j^{local}-E_T)t)]_3[\delta_{ij}]_2[G_{jj}^{(U)}(t)\\ &\times\exp((E_j^{local}-E_T)t)]_1+\int_0^t\sum_k\ [\tilde{G}_{ik}(t-t')]_8\\ &\times[\exp\{-(E_j^{local}-E_T)t'\}]_7\\ &\times\left[\frac{H_{kj}^{(U)}(t')-\tilde{H}_{kj}}{\sum_m\ (H_{mj}^{(U)}(t')-\tilde{H}_{mj})}\right]_6\\ &\times\left[-\frac{\partial}{\partial t'}\{G_{jj}^{(U)}(t')\exp[(E_j^{local}-E_T)t']\}dt'\right]_4.\end{aligned} \tag{36}$$

We have retained the structure of the propagation equation, Eq. (12), and the labeling of each of the terms, so that we can refer to the steps in the general sampling procedure. Term 5

equaled δ_{lj} and was removed with the sum over l. We have dropped the unnecessary tilde from the diagonal $H^{(U)}$.

Implementing the walk proceeds as before. For a walker at an initial location j, we evaluate the three components of E_j^{local} using Eq. (35). We must sample a time from term 4 above. Because $G_{jj}^{(U)}$ contains the diagonal energy E_j^{diag}, the time t' is simply sampled from the normalized exponential,

$$p_j(t)=\frac{\exp[(E_j^{crea}+E_j^{dest})t]}{-(E_j^{crea}+E_j^{dest})}. \tag{37}$$

If the time sampled is greater than the specified time step t, no ''scattering'' has occurred, so set $t'=t$ and sample the first term which means stay in the same place and complete the walk by assigning a weight equal to term 3. If t' is less than t, then sample a new location k from the normalized distribution of term 6. The new ''location'' k means a state with either one more or one less boson since the numerator of term 6 contains only the off diagonal part of the Hamiltonian and it connects only states with a different number of bosons. We can *a priori* decide whether to create or destroy a boson by selecting the option with the relative weight of the two corresponding contributions in term 6. Then, either a creation or a destruction move must be made with the probabilities

$$P_j^{crea}=\frac{E_j^{crea}}{E_j^{crea}+E_j^{dest}},$$

$$P_j^{dest}=\frac{E_j^{dest}}{E_j^{crea}+E_j^{dest}}. \tag{38}$$

If a destruction move is chosen, then one of the n bosons is deleted and its momentum transferred to the fermion. The selection of the boson is done with a probability proportional to the matrix element $H_{kj}^{(U)}(t')-\tilde{H}_{kj}$. Alternatively, this is equal to sampling one of the n terms in E_j^{dest} given in the sum in Eq. (35) with a probability equal to the value of the term divided by E_j^{dest}. If a creation move is selected, exactly the same procedure is followed of sampling a term from the sum in E_j^{crea} in Eq. (35) except that now there are an infinite number of such possible final states. There are a number of techniques available for sampling from such a distribution [30]. Finally, the new location k has been selected and it is assigned a weight $\exp[-(E_j^{local}-E_T)]$ and the walk continues with a new time $t-t'$ as before.

The algorithm implemented by Carlson and Schmidt differs from ours in that they always moved the walker at each step. They accomplish this by sampling the time until a creation/destruction move is made using the probability distribution of term 4, i.e., Eq. (37). This, in our language, effectively sets $t=\infty$, so term 2 is never sampled. We use the term ''effectively'' because this is not a rigorous analogy. In their implementation, the simulation was divided into several large time intervals and when a walker had passed one of the interval boundaries, its propagation was suspended until all other walkers had passed the same boundary. These intervals have, in a practical sense, the same effect as making the time steps in the exact-time algorithm equal to the large time interval.

V. THE MANY-BODY SCHRÖDINGER EQUATION

In this section we will discuss in some detail the exact imaginary-time method for a system of N equal mass bosons in three-dimensional space interacting with central pairwise potentials. We are adopting units where $\hbar^2/2m$ and $\hbar$ are unity. We shall ultimately apply this method to the ground state of liquid helium, but generalization to unequal masses and more general potentials is straight forward. Generalization to fermions is problematic [7].

The many-body Hamiltonian we are considering is simply

$$H=-\nabla^2+V(R), \tag{39}$$

where $R=(\vec{r}_1,\vec{r}_2,\ldots,\vec{r}_N)$ is the $3N$ dimensional vector of N particles with positions $\vec{r}_i$ and ∇^2 is the sum of laplacians with respect to the individual particle coordinates, $\Sigma\nabla_i^2$.

While real-space GFMC is applicable to a wide class of problems, it carries with it a few technical aspects that did not arise in the theoretical formulations of the lattice and field theory problems of Secs. III and IV. In particular, we must deal with infinitely repulsive potentials and additional sampling approximations that are related to issues of efficiency but which retain the exact character of the method. We will not discuss the issues of the attractive coulomb potential or nodal surfaces of wave functions because the usual methods that have been applied [14,37] can be carried over directly to the exact imaginary-time formulation.

A. Coordinate space algorithm

In Sec. II A, Eq. (11), we showed that the KLV Green's function for the time-independent Schrödinger equation was formally related to the time integral of the imaginary-time propagator. Although not explicitly formulated this way, the KLV algorithm can be understood as sampling a time step from the exponential probability distribution function of Eq. (11),

$$p(t)=\frac{1}{\Delta t}e^{-t/\Delta t}, \tag{40}$$

to select the time step for the Monte Carlo move. Thus, our modification merely changes the randomly chosen time step to a prescribed time step. We proceed by following the basic ideas of the Kalos, Levesque, Verlet algorithm [2] and note where differences exist.

The imaginary-time Green's function, Eq. (2), in a real-space basis, is the solution to the $3N$ coupled differential equations,

$$[-\nabla^2+V(R)-E_T]G(R,R',t)=-\frac{\partial G(R,R',t)}{\partial t}. \tag{41}$$

We will employ our first algorithm, Eq. (25) of Sec. II B, to express this exact Green's function as a known Green's function plus a correction term.

The simplest approach takes a Hamiltonian $H^{(U)}$ with a constant potential U if the constant potential can be chosen to be an upper bound for $V(R)$. In helium and many other systems, $V(R)$ is often not bounded above. To circumvent this problem, as well as to make the method more efficient in the general case, the domain of $H^{(U)}$ is restricted. Equivalently, outside a finite domain D the potential in $H^{(U)}$ is taken to be infinite.

The Green's function for a constant potential U in the domain D is

$$G^{(U)}(R,R',t)=\langle R|e^{-(H^{(U)}-E_T)t}|R'\rangle, \quad R,R'\epsilon D, \tag{42}$$

and satisfies

$$(H^{(U)}-E_T)G^{(U)}(R,R',t)=(-\nabla^2+U-E_T)G^{(U)}(R,R',t) = -\frac{\partial G^{(U)}(R,R',t)}{\partial t}, \tag{43}$$

with the condition $G^{(U)}=0$ for R or R' outside the domain D and $G^{(U)}(R,R',0)=\delta(R-R')$.

Because $G^{(U)}$ is defined on a finite domain, the derivation of the real-space propagator equation, equivalent to Eq. (14), is modified by the addition of a surface term. We begin with Eq. (13), in coordinate space,

$$\langle R|e^{-(H-E_T)t}|R'\rangle=G(R,R',t) =G^{(U)}(R,R',t)-\int_D dR''\int_0^t dt'\frac{\partial}{\partial t'} \times[G(R,R'',t-t')G^{(U)}(R'',R',t')]. \tag{44}$$

Carrying out the time derivatives to get H and $H^{(U)}$ as in Eq. (14), yields

$$G(R,R',t)=G^{(U)}(R,R',t)+\int_0^t dt'\int_D dR''G(R,R'',t-t') \times H^{(U)}(R'')G^{(U)}(R'',R',t') -G^{(U)}(R'',R',t')H(R'')G(R,R'',t-t'), \tag{45}$$

where the argument R'' for $H(R'')$ and $H^{(U)}(R'')$ identifies the coordinates for ∇^2 and $V(R'')$.

Applying Green's theorem to $G\nabla^2G^{(U)}-G^{(U)}\nabla^2G$ yields the fundamental equation of the coordinate space algorithm, giving the Green's function G, in a form that can be iterated, as

$$G(R,R',t)=G^{(U)}(R,R',t)+\int_0^t dt'\int_D dR''G(R,R'',t-t') \times[U-V(R'')]G^{(U)}(R'',R',t') +\int_0^t dt'\int_{\partial D} dS''G(R,R'',t-t') \times[-\hat{n}''\cdot\vec{\nabla}''G^{(U)}(R'',R',t')]. \tag{46}$$

The last term on the right-hand side is integrated over the bounding surface, ∂D, of the domain of $G^{(U)}$ and the second term is integrated over the domain. The term $[-\hat{n}''\cdot\vec{\nabla}''G^{(U)}(R'',R',t')]$ is minus the normal derivative at the surface. This new surface term is a feature that can be naturally fit into the diffusion interpretation of the simulation.

Both the bound U and the domain D are chosen differently for each value of R'. The only restriction on the domains is that it must eventually be possible to propagate from any valid part of configuration space to any other. The value of $U(R')$ is selected to be an upper bound to $V(R'')$ for R'' inside the domain so that all terms in Eq. (46) are positive.

Having constructed the coordinate space integral equation for $G(R,R',t)$, we can write down the importance sampled coordinate space equation for $\tilde{G}$ and the integral equation to be iterated to obtain the importance sampled density $\tilde{\psi}=\psi_T\psi$ corresponding to Eqs. (18) and (19) in Sec. II C. Following the definitions of Eq. (17), Eq. (46) becomes

$$\tilde{G}(R,R',t)\equiv\frac{\psi_T(R)}{\psi_T(R')}G(R,R',t)=\frac{\psi_T(R)}{\psi_T(R')}G^{(U)}(R,R',t) +\int_D dR''\int_0^t dt'\frac{\psi_T(R)}{\psi_T(R'')}G(R,R'',t-t') \times[U-V(R'')]\frac{\psi_T(R'')}{\psi_T(R')}G^{(U)}(R'',R',t') +\int_S dR''\int_0^t dt'\frac{\psi_T(R)}{\psi_T(R'')}G(R,R'',t-t') \times\frac{\psi_T(R'')}{\psi_T(R')}[-\hat{n}''\cdot\vec{\nabla}''G^{(U)}(R'',R',t')] \tag{47}$$

and

$$\tilde{\psi}(R,t+\Delta t)=\int dR'\tilde{G}(R,R',\Delta t)\tilde{\psi}(R',t). \tag{48}$$

Unlike our general case, Eq. (18), Eq. (47) with importance sampling cannot be transcribed from the original form, Eq. (46), by simply placing tildes over each function. The gradient at the surface prevents this elegance.

B. Ideal sampling in coordinate space

Sampling the Green's function equation, Eq. (47), proceeds along the lines described for the lattice and field theory examples. The task is more complicated because of the surface term and because $G^{(U)}(R,R',t)$ is not diagonal in the coordinate space basis. As a consequence, the sampling of both the time and positions from the importance sampled functions will have to be done in a nonoptimal fashion. Our substitute procedure will remain exact because the sampling will include compensating weighting factors. Further, we will argue that the increase in variance will be relatively small and acceptable. The main source of variance will be the approximate nature of the trial wave function.

As before, we will develop a procedure for sampling which introduces as little variance as possible and thus is as near zero variance as possible when the importance function approaches the exact ground state. It is thus useful to write down this ideal case and integrate Eq. (47) assuming $\psi_T = \psi_0$. Then

$$\int dR\tilde{G}(R,R',t)=1 \tag{49}$$

on the left side of Eq. (47) and inside both integrals of the right side. This leaves the identity,

$$\begin{aligned}1=&\int dR\tilde{G}^{(U)}(R,R',t)\\&+\int_0^t dt'\left[\int_D dR''[U-V(R'')]\frac{\psi_T(R'')}{\psi_T(R')}G^{(U)}(R'',R',t')\right.\\&\left.+\int_S dR''\frac{\psi_T(R'')}{\psi_T(R')}[-\hat{n}''\cdot\vec{\nabla}''G^{(U)}(R'',R',t')]\right].\end{aligned} \tag{50}$$

From the above equation, we can read off the procedure to be used for exact importance sampling, in analogy to Eq. (22). The term in large brackets is a normalized probability distribution in the t' variable. As before (e.g., Sec. II C 1) we would sample t' from this distribution and if $t'>t$ then the $\tilde{G}^{(U)}(R,R',t)$ term is to be sampled for a new location R of the walker and the walk terminates. If $t'<t$, then we must choose a time sampled from the distribution of times for moves to the surface and the distribution of times for interior (volume) moves with a probability proportional to the two integrals over R''. Finally, having selected the surface or volume time, the location R'' is to be sampled from the corresponding integrand scaled by its normalization. This R'' is assigned a new time $t-t'$ to complete the time step and the sampling of $\tilde{G}(R,R'',t-t')$ is done to continue the walk.

The above description is ideal and, with a realistic importance function, we must carry out these steps in a way that includes weights. There will be some additional population fluctuation resulting from the fact that we include weights that are not unity. Our first step in implementing a nonideal algorithm is constructing the domain Green's function $G^{(U)}$. We can then sample a time from this function and the positions R'' on either the interior or the surface of the domain. These sampling procedures do not include the importance function or the $U-V(R'')$ factor, so we will modify the sampling procedures in an approximate fashion and construct the appropriate weighting factors to make the algorithm exact.

1. Constructing $G^{(U)}(t)$

Before we can develop a sampling procedure for the various terms in Eq. (47), it is necessary to construct the domain Green's function $G^{(U)}$. KLV arrived at the time-independent form by first solving for the time dependent $G^{(U)}$ which we seek. A hyperspherical domain in 3N dimensions was used by Kalos in his early investigations [1,38] of few-body problems. An improved Green's function, suited to many-body systems, was developed by Kalos, Levesque, and Verlet [2]. The domain D is a set of nonoverlapping spheres, centered about each particle. The radius of each sphere and the upper bound U are adjusted at every step of the walk. As long as the bound is rigorous and the spheres nonoverlapping, any choice for domain size and U is legal, but very large domains usually result in large values of the bound U which limits the distance walkers move in a single step. The optimization of these is easily done by trial and error, so we will omit further discussion.

Since the spheres in the domain D are nonoverlapping, $G^{(U)}$ can be factored into a product of N one-body functions,

$$\begin{aligned}G^{(U)}(R,R';t)&=\langle R|\exp[-(H^{(U)}-E_T)t]|R'\rangle\\&=\exp(-(U-E_T)t)\prod_{i=1}^{N}\frac{1}{a_i^3}g\left(\frac{|\vec{r}_i-\vec{r}_i'|}{a_i},\frac{t}{a_i^2}\right).\end{aligned} \tag{51}$$

The sphere radius for the ith particle is a_i, and $\vec{r}_i$ and $\vec{r}_i'$ are the three-dimensional positions of particle i. The analytic forms of the $g(\rho,t)$ are well known from classical diffusion theory. They describe the time evolution from a unit point source at the center of a unit sphere with a unit diffusion constant, as it diffuses toward the perfectly absorbing boundary of the surface of the sphere. The function g, as defined by the above equation, satisfies

$$-\nabla_\rho^2 g(\rho,t)=-\frac{\partial g(\rho,t)}{\partial t}\quad(0<\rho<1) \tag{52}$$

with the boundary conditions $g(\rho,t\to 0)=\delta^3(\vec{\rho})$ and $g(\rho=1,t)=0$. The time-independent $G_{KLV}^{(U)}$ of KLV is defined, in analogy with G_{KLV} in Eq. (11), as the integral over time with the constant E_C to make the eigenvalue spectrum positive, i.e.,

$$G_{KLV}^{(U)}(R,R')=\int_0^\infty dt e^{-(U+E_C)t}\prod_{i=1}^{N}\frac{1}{a_i^3}g\left(\frac{|\vec{r}_i-\vec{r}_i'|}{a_i},\frac{t}{a_i^2}\right). \tag{53}$$

It should be noted that the second argument in g is really $\hbar t/2m_i a_i^2$, where m_i is the mass of particle i, but we have

assumed $\hbar=1$ and $\hbar^2/2m=1$ with all the masses the same. A change of the scaling factor is the only change necessary for unequal mass particles.

The function $g(\rho,t)$ is easily calculated from the eigenfunction expansion of Eq. (52) to be

$$g(\rho,t)=\frac{1}{2\rho}\sum_{n=1}^{\infty} n\sin(n\pi\rho)\exp(-\pi^2n^2t)$$
$$=\frac{1}{\rho\sqrt{(4\pi t)^3}}\exp\left(-\frac{\rho^2}{4t}\right)+\frac{1}{\rho\sqrt{(4\pi t)^3}}\sum_{n=1}^{\infty}(\rho+2n)$$
$$\times\exp\left(-\frac{(\rho+2n)^2}{4t}\right)-(\rho-2n)\exp\left(-\frac{(\rho-2n)^2}{4t}\right). \tag{54}$$

The second equality is the result of using the Poisson summation formula on the eigenfunction expansion. The result of the Poisson formula can be recognized as the contributions for a unit source at the center of the sphere and a set of image sources outside the sphere. Only five terms of the series are required to obtain convergence machine precision on 64-bit computers for small t ($t<0.11$) using the second form and for large t ($t>0.11$) using the first.

The integral over $d^3\rho$ of $g(\rho,t)$ is necessary to sample new positions from $G^{(U)}$. To machine precision, again only a few terms are needed,

$$h(t)=\int_0^1 d^3\rho g(\rho,t)$$
$$=\begin{cases}-2\sum_{n=1}^{4}(-1)^n\exp(-\pi^2n^2t), & t>0.16\\ 1-\frac{2}{\sqrt{\pi t}}\sum_{n=1}^{2}\exp\left(-\frac{(2n-1)^2}{4t}\right), & t<0.16\end{cases}. \tag{55}$$

The surface term, $-\hat{n}\cdot\vec{\nabla}G^{(U)}(R,R',t)$, is calculated using Eq. (51) as

$$-\hat{n}\cdot\vec{\nabla}G^{(U)}(R,R',t)|_{\partial D}=e^{-(U-E_T)t}\sum_{i=1}^{N}\frac{-1}{a_i^4}g'\left(1,\frac{t}{a_i^2}\right)$$
$$\times\prod_{j\neq i}\frac{1}{a_j^3}g\left(\frac{|\vec{r}_j-\vec{r}_j'|}{a_j},\frac{t}{a_j^2}\right). \tag{56}$$

Since g is spherically symmetric, the normal derivative is a constant. Its normalization is

$$-\frac{d\,h(t)}{dt}=\int d^3\rho\left[-\frac{\partial g(\rho,t)}{\partial t}\right]=\int d^3\rho[-\nabla_\rho^2 g(\rho,t)]$$
$$=-4\pi\hat{n}\cdot\vec{\nabla}_\rho g(\rho,t)|_{\rho=1}. \tag{57}$$

This is just a consequence of current conservation in the analogous diffusion problem. The flux through the boundary is just the negative of the density change in the domain. With these results, the normalizations of $G^{(U)}$ and $-\hat{n}\cdot\vec{\nabla}G^{(U)}$ are known and each of these distributions can be sampled once the time t is chosen.

All of the sampling techniques we have described in this section are essentially a restatement of the constructions employed by KLV. Their method of choosing the time sampled it from the exponential in Eq. (53). Our formulation, thus far, has been toward implementing our first algorithm as embodied in Eq. (25) and its coordinate space manifestation, Eq. (47). That algorithm requires that we sample a time from $-\partial\Sigma_n\tilde{G}_{nj}^{(U)}(t')/\partial t'$. The sum over final states in this expression includes both diffusion that terminates inside the spheres and diffusion to the surface. We will consider both processes in the following section.

2. Sampling the domain Green's function $G^{(U)}$

The time derivative of the domain Green's function is a probability density in the two variables R and t, since it is positive definite and normalized,

$$\int_0^\infty dt'\int_D dR\left(-\frac{\partial}{\partial t'}G^{(U)}(R,R',t')\right)=1. \tag{58}$$

This follows from the conditions that $G^{(U)}$ is $\delta(R-R')$ at $t'=0$, and zero at $t'=\infty$. Having constructed $G^{(U)}$ analytically in the preceding section, we wish to sample R and t from the probability distribution function, in the form we need it. Since it is a joint probability distribution function in two variables, we can write it as the product of a marginal probability distribution function for t and the condition probability distribution function for R given t,

$$-\frac{\partial}{\partial t'}G^{(U)}(R,R',t')=\left[\frac{-\frac{\partial}{\partial t'}G^{(U)}(R,R',t')}{\int_D dR-\frac{\partial}{\partial t'}G^{(U)}(R,R',t')}\right]$$
$$\times\int_D dR\left[-\frac{\partial}{\partial t'}G^{(U)}(R,R',t')\right]. \tag{59}$$

The rightmost factor in this equation is the marginal probability distribution function, the probability of sampling a time t independent of the value of R sampled. The conditional probability distribution function in brackets is the probability that an R is sampled, given that t was sampled.

We can transform the probability distribution function into a form that can be used for our purposes because $G^{(U)}$ satisfies the Schrödinger equation, Eq. (43). If we integrate both sides of Eq. (43) over the domain D and over the time domain, we find that the unit normalization can be expressed as three terms:

$$1=\int dR\, G^{(U)}(R,R',t) + \int_0^t dt' \int_D dR''(U-E_T)G^{(U)}(R'',R',t') + \int_0^t dt' \int_S dR''[-\hat{n}''\cdot\vec{\nabla}''G^{(U)}(R'',R',t')]. \quad (60)$$

The first term on the right is the integral from $t'=t$ to $t'=\infty$. The last two terms are the integral up to t. The last two terms are obtain by substituting $H^{(U)}-E_T$ for the time derivative, then transforming the Laplacian to a surface integral.

There is a one-to-one correspondence between the terms in Eq. (60) and the normalization condition for the case of ideal sampling given by Eq. (50). As we will show below, we can sample all the terms in Eq. (60). Then we will use this as a step towards our ultimate goal of sampling the terms in the full Green's function equation, Eq. (47), when a nonideal importance function is used. We will use the correspondence between Eqs. (60) and (50) to construct the approximate probability distribution functions that we sample, and then modify them by the appropriate corrective weights.

The first term in Eq. (60) equals the probability that t', sampled from the marginal probability distribution function, will be greater than t. The second and third terms are the integrals of the joint probability distribution function over D (including S) and up to time t and have the value equal to the probability of sampling a time t' less than t.

From this equation, we can see that if we want to sample the marginal probability distribution function in t' over the interval $0<t'<\infty$, then we may sample from

$$\int_D dR''(U-E_T)G^{(U)}(R'',R',t') + \int_S dR''[-\hat{n}''\cdot\vec{\nabla}''G^{(U)}(R'',R',t')], \quad (61)$$

and this is the marginal probability distribution function in Eq. (59).

In our algorithm, if the time sampled is less than t, then we need to sample the conditional probability distribution function in Eq. (59) for a position R. The marginal probability distribution function for t' in Eq. (61) is the sum of two distributions. If t' is sampled from the first, then R is sampled from the interior point of D and the conditional probability distribution function for that distribution is the integrand of the first term. If t' is sampled from the second distribution, then likewise a point on the surface is to be sampled from the integrand.

Having outlined the steps needed to sample t and R from Eq. (59) we can now specify the particulars needed to sample each of the terms in Eq. (60). Our first step is to sample a time from the distribution in Eq. (61). This can be done by sampling a time t' from the probability density

$$P(t')=-\frac{d}{dt'}\left[\prod_{i=1}^{N} h\left(\frac{t'}{a_i^2}\right)e^{-\tilde{U}t'}\right]. \quad (62)$$

The constant $\tilde{U}$ is $U-E_T$, but later we will want to generalize the procedure so we will treat it as a an arbitrary positive constant. Equation (62) can be sampled by taking the minimum of the times sampled from each $-dh(t'/a_i^2)/dt'$ and $\tilde{U}\exp(-\tilde{U}t')$. (This is an example of a combinatorial sampling method discussed in Ref. [30].) The former can be sampled by direct application of the transformation method [30],

$$t_i=a_i^2 h^{-1}(\xi_i), \quad (63)$$

and the latter from

$$t_0=-\frac{\ln(\xi_0)}{\tilde{U}}, \quad (64)$$

where ξ_0 through ξ_N are $N+1$ independent random numbers on the interval $(0,1)$.

By taking the smallest time of the $N+1$ values sampled above, we have sampled a time from the probability distribution function of Eq. (61). Because we know which t_i was the smallest, we also know the term in Eq. (61) from which t' was sampled. The probability that the t' sampled from $-dh(t'/a_i^2)/dt'$ is the smallest is the probability of sampling t between t and $t+dt$ times the probability of all the other $t's$ being larger, which is

$$-\frac{d}{dt'}\left[h\left(\frac{t'}{a_i^2}\right)dt'\prod_{j\neq i} h\left(\frac{t'}{a_j^2}\right)e^{-\tilde{U}t'}\right]. \quad (65)$$

Equation (65) is the normalization of the $-\hat{n}\cdot\vec{\nabla}_i''G^{(U)}(R'',R';t')$ term if $\tilde{U}=U-E_T$. Similarly, if the t' sampled from $\tilde{U}e^{-\tilde{U}t'}$ is the smallest, the probability is

$$\tilde{U}e^{-\tilde{U}t'}dt'\prod_i h\left(\frac{t'}{a_i^2}\right), \quad (66)$$

which is the normalization of $\tilde{U}G^{(U)}(R'',R',t')$.

Knowing t' and knowing the term from which it was sampled allows us to sample the new position R''. If the t' from $\tilde{U}e^{-\tilde{U}t'}$ is smallest, sample an R'' from $G^{(U)}(R'',R',t')/\int dR''G^{(U)}(R'',R',t')$. This latter sampling is simply done by sampling a ρ for each particle from $g(\rho,t'/a_i^2)/h(t'/a_i^2)$ and letting $|\vec{r}_i-\vec{r}_i'|=a_i\rho$. Sampling from the surface term is very similar. If the smallest t' was sampled from $-dh(t'/a_i^2)/dt'$, then to sample the surface derivative of $G^{(U)}$, the position of the ith particle must be sampled at the surface of its sphere, and the other particles sampled as before. Since the normal derivative of $g(\rho,t')$ is independent of position on the surface, this can be easily accomplished by sampling ρ values as above and simply promoting particle i randomly to the surface of its sphere.

Routines to sample ρ from $g(\rho,t)d^3p$ and t from $-d\,h(t)/dt$ are available in the literature [39].

The above procedure sampled time and position from Eq. (61). To complete the desired sampling of Eq. (60), we need only to check if the smallest time sampled was greater than the time step t. If so, then set $t'=t$ and sample the new location from the interior of the domain as described above. That location has then been sampled from the integrand of the first term of Eq. (60) and we have sampled $G^{(U)}$.

C. Importance sampling the full Green's function in coordinate space

The above section presents a way of sampling time and sampling positions for a probability distribution function which has the same domain, $t' \in (0,t)$ and $R'' \in D$ and $R'' \in S$, as the terms in the full Green's function equation, Eq. (47). Our analysis of the ideal case told us that the sampling of position and time should be done from Eq. (50), even if the importance function is not exact. Since exact sampling of Eq. (50) is not possible, we must introduce weights. We could proceed by simply taking the three terms of Eq. (60) and multiplying and dividing each term in the full Green's function equation, but that defeats the whole effect of importance sampling. Instead, we modify the sampling procedure for $G^{(U)}$ to make it more closely resemble $\tilde{G}^{(U)}$, and then multiply and divide corresponding terms of the full Green's function equation.

We begin by changing $\tilde{U}$ from $U-E_T$ to

$$\tilde{U}=[U-V(R')]. \tag{67}$$

This choice is not unique. Whatever change is made to $\tilde{U}$ must be compensated for with appropriate factors included in the weight terms to correct for this difference. This change tends to increase the number of volume steps selected which corresponds to picking the first term in Eq. (47).

The inclusion of some effect of the importance function, $\psi_T(R')/\psi_T(R)$, in $\tilde{G}^{(U)}$ is done by using its gradient to select the direction of $\vec{r}_i'-\vec{r}_i$. This is accomplished by expanding

$$\frac{\psi_T(R')}{\psi_T(R)} \approx \prod_i \left(1+|\vec{\nabla}_i \ln \psi(R)||\vec{r}_i'-\vec{r}_i|\cos\theta_i\right)$$
$$\approx \prod_i \left(1+\min(|\vec{\nabla}_i \ln \psi(R)||\vec{r}_i'-\vec{r}_i|,1)\cos\theta_i\right), \tag{68}$$

where θ_i is the angle between the gradient and $\vec{r}_i'-\vec{r}_i$. The min function constrains the expression to be positive definite so that is can be used as a probability density. The right-hand side of Eq. (68) can be recognized as a sum of a uniform and linear distribution in $\cos\theta$ since the solid angle element is $d\cos\theta\, d\phi$. The ϕ is sampled uniformly. After sampling from Eq. (68) we must, as usual, include an additional correction factor in the weight for the walker which is simply

$$W_\psi(R,R')=\frac{\psi_T(R')}{\psi_T(R)}\prod_i \left[1+\min(|\vec{\nabla}_i \ln \psi(R)| \times|\vec{r}_i'-\vec{r}_i|,1)\cos\theta_i\right]^{-1}. \tag{69}$$

The choice of $\tilde{U}$, E_C, and the a_i's should be done to maximize approximately the average time t' per step. Typically, there is a large range of values that give similar results. The complete algorithm for a $3N$ dimensional code is the following.

(1) Select a walker R_i from the previous time step. The iteration of the equation is complete when no previous walkers remain. Calculate its $U, V(R_i)$, the a_i's, and the trial wave function value and gradient.

(2) Sample the smallest t' from the N surface terms $-dh(t'/a_i^2)/dt'$ and from $\tilde{U}\exp(-\tilde{U}t')$ with $\tilde{U}=[U-V(R_i)]$. If $t'>t$, set $t'=t$.

(3) Sample distances $|\vec{r}_m-\vec{r}_m'|=\rho a_m$ from $g(\rho,t'/a_m^2)$ with t' the smallest from step 2.

(4) If a surface time of particle k is the smallest (and less than t) make its $|\vec{r}_k'-\vec{r}_k|=a_k$, then (i) sample the set of angles θ_i from right-hand side of Eq. (68), and calculate weight W_ψ, and (ii) set $t=t-t'$, assign the weight W_ψ to the walker, and continue the walk step 1.

(5) If the volume time sampled from $\tilde{U}\exp(-\tilde{U}t)$ is the smallest (but less than t), then assign angles as above and calculate the weight $W=[U-V(R_i')]/[U-V(R_i)]W_\psi\exp([E_T-V(R_i)]t)$. Move the walker to R_i'. Set $t=t-t'$, assign the walker the weight W and restart the walk at step 1.

If the volume time sampled is greater than t, then set it to t and sample the angles as above and calculate the weight, $W=W_\psi\exp([E_T-V(R_i)]t)$. Move the walker to R_i', assign the walker the weight W and the walk terminates.

The weight factors above were ''assigned'' to the walker. In the most common implementations of GFMC, this means that when a walker is selected from the previous generation it is assigned a weight of unity. All subsequent assignments multiply this weight. If the walk does not terminate, but t is reset to $t-t'$, the walker retains its weight when it reinitiates the walk for the remainder of the time step and the weight continues to be assigned. In the final step above, when the walk terminates, one takes the integer part of the weight plus a random number, $W+\xi$, and if this is greater than one, one replicates the walker and that number of new walkers begins the next time step. If the integer is zero, then the walker is eliminated from the population in the next time step.

The process is repeated for as many time steps as required for the simulation to converge and enough statistics to be calculated to accurately determine the desired averages.

VI. APPLICATION TO THE EFFECTIVE MASS PROBLEM

An advantage to using the imaginary-time formalism is that it allows one to calculate imaginary-time correlation functions directly. A straightforward application would be to calculate the effective mass of impurity atoms (such as a ^{3}He atom) in liquid ^{4}He. Here we will apply the method to a

distinguishable ^{4}He atom in liquid ^{4}He.

The effective mass of the impurity is given by calculating the diffusion constant of the impurity in the imaginary-time simulation. This is easily understood by looking at the two-time density-density distribution function of the impurity. If we were looking at a free particle, this distribution function is simply the three-dimensional Gaussian Green's function, with width $\sigma=(\hbar^2 t/m)^{1/2}$, so that its diffusion constant is $\hbar^2/m$. The impurity's density-density distribution function is not in general a Gaussian. However, at long times, the positions of the background liquid ^{4}He atoms become uncorrelated, and the distance the impurity moves in a time t will be given by a sum of uncorrelated random variates. So for times much longer than this correlation time, the distribution of positions will again be Gaussian. The width of the Gaussian, $\sigma=(\hbar^2 t/m^*)^{1/2}$, defines the effective mass.

The standard method to define the effective mass of an impurity atom is to look at the long wavelength limit

$$\frac{1}{m^*}=\lim_{p\to 0}\frac{1}{p}\frac{\partial E(p)}{\partial p}, \tag{70}$$

where $E(p)$ is the energy of the system with the single particle excitation, corresponding to the impurity, having momentum p.

This is equivalent to our diffusion constant expression. One way to verify this is to imagine a direct calculation of $E(p)$ using Green's function Monte Carlo. We add to the Hamiltonian an external potential that is zero when impurity has z coordinate $0<z<L$, and is positive infinity otherwise. We assume that the system without the external potential is solved using Green's function Monte Carlo. For convenience, we also assume that the time step Δt for the Green's function Monte Carlo calculation is small. In that case, the simulation with the external potential could done as in an exact imaginary-time formulation for a single particle. At the end of a step we simply ask if the impurity has left the interval $0<z<L$. If so the walker is removed. If not, the walker is given a slightly larger weight $\exp[E(p)\Delta t]$. $E(p)$, the growth estimate for the excitation energy, is chosen to keep the average number of walkers constant. The dynamics for large L is identical to that for a single particle. If the energy of this system is calculated using Green's function Monte Carlo, the energy of the excitation will be set by the time it takes the impurity to diffuse between the two limits

A. Numerical results

The time-step algorithm of Sec. V was carried out for a system of 64 ^{4}He atoms interacting through the HFDHE2 two-body potential given by Aziz [40]. GFMC calculations with this potential have been shown to produce the ground state properties of liquid and solid helium accurately [41]. Periodic boundary conditions were employed. The importance function included two- and three-body correlation terms [42].

The algorithm was initially used to calculate the equilibrium ground state energy by carrying out 1000 time steps of size 0.001/K. The calculation required a simple modification

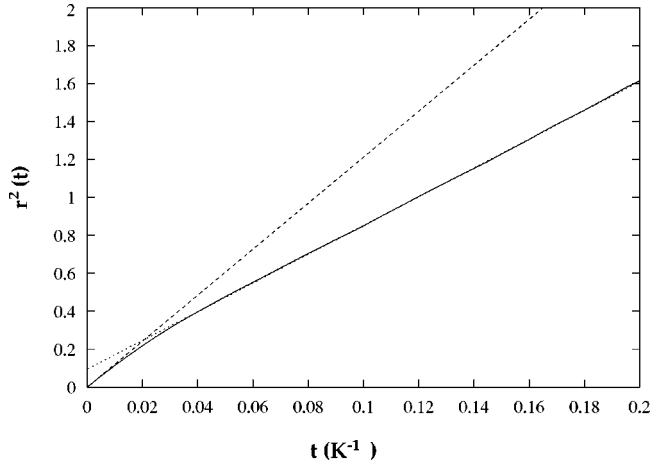

FIG. 1. Mean square diffusion distance vs imaginary time for a distinguishable mass-four helium atom. Short time slope indicate bare mass and long time slope indicates effective mass.

of the usual GFMC time-independent algorithm. We observed no differences in the convergence of the two methods either in CPU time or in variance.

An equilibrated population of 1000 walkers was the starting point of the calculation of the imaginary-time diffusion constant. Each walker carried a unique tag, so that if a branching process occurred, progeny would be given a new unique tag and would retained the tag of the parent walker. The calculation was carried out so that after each time step, the entire population with identification information was stored. With the identification information, the parent walkers could be traced back through every step of the simulation. Each atom in each walker was thus distinguishable. For any atom at a time t, the quantity $|\vec{r}(t)-\vec{r}(0)|$ could be calculated.

Figure 1 shows the average value of the mean square displacement calculated versus imaginary time. As expected from our discussion above, the initial slope reflects the bare mass of the helium atom and is 1.0 based on a simply observing that the $m^*=1$ line in Fig. 1 matches the slope near the origin. The linear asymptotic form is achieved quickly, after no less than 0.1/K in imaginary time. In spatial terms, this is about 1 Å, or about the average inter-particle spacing. It is somewhat unexpected, that in less than 1 Å, the particle which moves in imaginary time are essentially uncorrelated with the original position. Fitting this form over the last half of the graph yields a slope of $m^*=1.60\pm.01m$, for an effective mass four impurity.

We calculate the diffusion of a mass-four distinguishable atom with interactions identical to all other helium atoms because the process of exchange in a real helium fluid causes the effective mass to lose physical significance. The definition in terms of the excitation spectrum fails because there are no excitations with the quadratic form required in Eq. (70). Alternatively, one can understand this in terms of the fact that the simulation, as we have described it, does not include the symmetry requirement that the ground state be symmetric under exchange of particle coordinates. The boson ground state simulation remains exact without the symmetry requirement as long as we calculate expectation values

averaged over all possible $N!$ exchanges of particle coordinates. Expectation values relating correlations of quantities at different imaginary times must also average over all exchanges to be correct. Such an average would yield an infinite diffusion constant consistent with the zero effective mass calculated from Eq. (70).

GFMC imaginary-time methods can be applied to obtain other dynamical physical properties with appropriate use of transforms implemented with, for instance, maximum entropy techniques [8,9,43,44,12]. Another application is the determination of the effective mass of truly distinguishable impurity atoms such as ^{3}He [45]. Such calculations require significantly more computational time. A few hours of time on a 100 MFlop computer is sufficient to produce the results in Fig. 1. This is due to the fact that all helium atoms could be tracked in the diffusion process, not just one atom of a walker.

VII. CONCLUSION

We have presented an exact form of a GFMC calculation that iterates the imaginary-time propagator with no time-step errors. Our description includes implementations for lattice Hamiltonians, continuous-space N-body problems, and simple field theories. We have demonstrated that in these classes of problems there is no increased computational or algorithmic complexity; indeed, simple alterations to existing computer codes are all that is required in each of the examples. Our calculation of the pseudoeffective mass of a helium added to liquid helium was included as a tangible example of the implementation of the coordinate space form where we could calculate all of the usual ground state properties and extend this to an imaginary-time quantity previously inaccessible to time-independent and pseudotime algorithms. We have observed that popular implementations of imaginary-time propagator algorithms are approximate and require significant additional programming and tests for time-step errors. An algorithm with fewer heuristic modifications and a guaranteed time-step independence is offered as a replacement.

ACKNOWLEDGMENTS

This work was supported by the NSF, Grant Nos. CHE 9407309, NSF-ASC 9720221, and NSF-ACI 0081324. Calculations were performed at the Ohio Supercomputer Center under Grant Nos. PGS178 and PGS030.

[1] M.H. Kalos, Phys. Rev. **128**, 1791 (1962).
[2] M.H. Kalos, D. Levesque, and L. Verlet, Phys. Rev. A **9**, 2178 (1974).
[3] M.P. Nightingale and C.J. Umrigar, *Quantum Monte Carlo Methods in Physics and Chemistry* (Kluwer Academic, Boston, 1999).
[4] J.W. Moskowitz, K.E. Schmidt, M.A. Lee, and M.H. Kalos, J. Chem. Phys. **77**, 349 (1982).
[5] P.J. Reynolds, D.M. Ceperley, B.J. Alder, and W.A. Lester, J. Chem. Phys. **77**, 5593 (1982).
[6] J.B. Anderson, J. Chem. Phys. **65**, 4122 (1976).
[7] F. Pederiva and M.H. Kalos, Comput. Phys. Commun. **122**, 440 (1999).
[8] R.N. Silver, D.S. Sivia, and J.E. Gubernatis, Phys. Rev. B **41**, 2380 (1990).
[9] J.E. Gubernatis, M. Jarrell, R.N. Silver, and D.S. Sivia, Phys. Rev. B **44**, 6011 (1991).
[10] M. Harris and P. Ballone, Solid State Commun. **105**, 725 (1995).
[11] S. Baroni and S. Moroni, Phys. Rev. Lett. **82**, 4745 (1999).
[12] P. Henelius, A.W. Sandvik, C. Timm, and S.M. Girvin, Phys. Rev. B **61**, 364 (2000).
[13] R.K. Kalia, P. Vashishta, and M.A. Lee, Solid State Commun. **52**, 873 (1984).
[14] D.W. Skinner, J.W. Moskowitz, M.A. Lee, P.A. Whitlock, and K.E. Schmidt, J. Chem. Phys. **83**, 4668 (1985).
[15] Y.H. Lee and M.A. Lee, Phys. Rev. A **40**, 1207 (1989).
[16] C.J. Umrigar, M.P. Nightingale, and K.J. Runge, J. Chem. Phys. **99**, 2865 (1993).
[17] M. Mella, G. Morosi, and D. Bressanini, Phys. Rev. E **61**, 2050 (2000).
[18] H.A. Forbert and S.A. Chin, Phys. Rev. B **63**, 144518 (2001).
[19] S. Marten and A. Luchow, J. Chem. Phys. **115**, 5362 (2001).
[20] D. Ceperley, J. Comp. Physiol. **51**, 404 (1983).
[21] J. Carlson, Phys. Rev. B **40**, 846 (1989).
[22] J. Carlson and K.E. Schmidt, *Monte Carlo Approaches to Effective Field Theories* (Plenum, New York, 1992), Vol. 3.
[23] D. Weaire and A.R. Williams, J. Phys. C **10**, 1239 (1977).
[24] D. Weaire and V. Srivastava, J. Phys. C **10**, 4309 (1977).
[25] C.S. Hellberg and E. Manousakis, Phys. Rev. Lett. **78**, 4609 (1997).
[26] C.S. Hellberg and E. Manousakis, Phys. Rev. B **61**, 11787 (2000).
[27] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Quantum Mechanics* (Cambridge University Press, Cambridge, 1992).
[28] L. Schiff, *Numerical Recipes in Fortran* (McGraw-Hill, New York, 1968).
[29] M.A. Lee, P. Vashishta, and R.K. Kalia, Phys. Rev. Lett. **51**, 2422 (1983).
[30] M.H. Kalos and P.A. Whitlock, *Monte Carlo Methods Volume I: Basics* (John Wiley, New York, 1986).
[31] K.E. Schmidt, M.A. Lee, M.H. Kalos, and G.V. Chester, Phys. Rev. Lett. **47**, 807 (1981).
[32] Q.N. Usmani, S. Fantoni, and V.R. Pandharipande, Phys. Rev. B **26**, 6123 (1982).
[33] P.A. Whitlock and G.V. Chester, Phys. Rev. B **35**, 4719 (1987).
[34] S. Moroni, S. Fantoni, and G. Senatore, Phys. Rev. B **52**, 13 547 (1995).
[35] R.P. Feynman, *Statistical Mechanics* (Addison-Wesley, Reading, 1972).
[36] N. Trivedi and D.M. Ceperley, Phys. Rev. B **41**, 4552 (1990).

[37] J.W. Moskowitz and K.E. Schmidt, J. Chem. Phys. **85**, 2868 (1986).
[38] M.H. Kalos, Nucl. Phys. A **126**, 609 (1969).
[39] K.E. Schmidt, *Variational and Green's Function Monte Carlo of Few-Body Systems*, Lecture Notes in Physics, Vol. 273 (Springer, Berlin, Heidelberg, New York, 1987).
[40] R. Aziz, V. Nain, J. Carley, W. Taylor, and G. McConville, J. Chem. Phys. **70**, 4330 (1979).
[41] M.H. Kalos, M.A. Lee, P.A. Whitlock, and G.V. Chester, Phys. Rev. B **24**, 115 (1981).
[42] K. Schmidt, M.H. Kalos, M.A. Lee, and G.V. Chester, Phys. Rev. Lett. **45**, 573 (1980).
[43] E.S.G. Krilov and B.J. Berne, J. Chem. Phys. **114**, 1075 (2001).
[44] T.H.M. Asakawa and Y. Nakahara, Prog. Part. Nucl. Phys. **46**, 459 (2001).
[45] D.M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).