PHYSICAL REVIEW A VOLUME 40, NUMBER 3 AUGUST 1, 1989

Treating nodes and singularities in quantum simulation of electronic structure: A sampling technique using multiple time steps

Young Hee Lee* and Michael A. Lee
Department of Physics, Kent State University, Kent, Ohio 44242

(Received 25 August 1988)

Quantum simulation of the electronic structure of atoms and molecules provides a rigorous upper bound to the true ground-state energy in these systems. Two approximations that are normally invoked, the short-time approximation for the Green's function and the fixed-node approximation for the true ground-state wave function, are investigated. It is found that certain approximations to the short-time Green's function introduce significant errors in the wave function near the nodes and Coulomb singularities. These errors can cause the rigorous upper bound to be violated. Introducing a multiple-time-step method we will show how this can be treated properly.

I. INTRODUCTION

The Green's-function Monte Carlo (GFMC) method developed by Ceperly and Kalos[1] was the first quantum-simulation method applicable to the full many-body Schrödinger equation. Its greatest success was initially in obtaining exact ground-state energies for boson quantum fluids, liquid helium in particular.[2] To date, a practical formulation of the GFMC or other simulation methods for the many-fermion problem has eluded researchers, although a few special cases have been solved.[3]

The quantum simulation of many-fermion problems has thus progressed through the development of several approximation schemes. In applications involving the electronic structure of atoms and molecules, most of these approximation schemes are some variation of a method which has become known as the GFMC short-time approximation.[4] In Sec. II we will derive the short-time approximation to the GFMC and specifically identify the sources of error which occur. We wish to explicitly distinguish between errors that result from the short-time nature of Green's function and those which are inherent in dealing with fermion problems.

The principal results of this study were that we were able to separate the sources of error which occurred as a result of the finite size of the time step, the attractive Coulomb singularity, and the nodes of the wave function. With use of improvement which we call the multiple-time-step method near the nucleus and nodes, several significant sources of error were removed. The only sources of error that remained came from the finite size of the time step and the fixed-node approximation, which will be explained in Sec. III. The calculation of the correlation energy of lithiumlike atoms will be shown as a realistic example demonstrating the effectiveness of these improvements.

II. GREEN'S-FUNCTION MONTE CARLO METHOD

A. The diffusion analogy

Kalos and co-workers developed the GFMC method by noting the analogy between the N-body Schrödinger equation with the Hamiltonian H,

$$\left[-\frac{\hbar^2}{2m}\nabla^2+V(R)\right]\psi_S(R,t)=i\hbar\frac{\partial\psi_S(R,t)}{\partial t}$$

or (1)

$$H\Psi_S=i\hbar\frac{\partial\Psi_S}{\partial t}\ ,$$

and the diffusion equation with an absorption term $A(R)$,

$$-\tfrac{1}{2}D\nabla^2\psi_D(R,\tau)+A(R)\psi_D(R,\tau)=-\frac{\partial\psi_D(R,\tau)}{\partial\tau}\ , \tag{2}$$

where R is a three-dimensional position vector, $R=(r_1,r_2,\ldots,r_N)$ and hence ∇^2 is a three-dimensional Laplacian, and D is a diffusion coefficient.

If the absorption term $A(R)$ is equal to $V(R)$, and, setting $\hbar=m=D=1$ and $\tau=it$, the two equations are exactly equivalent. With this notation the Schrödinger equation becomes

$$-\tfrac{1}{2}\nabla^2\psi(R,\tau)+V(R)\psi(R,\tau)=-\frac{\partial\psi(R,\tau)}{\partial\tau}\ . \tag{3}$$

Thus the solutions of the diffusion equation $\psi_D(R,\tau)$ are equivalent to the solutions of the Schrödinger equation $\psi_S(R,t)$ in imaginary time. It is easy to demonstrate that Eq. (3) can be used directly to find the ground state energy of the Schrödinger equation. The formal quantum-mechanical expression for the time evolution of the $\Psi(R,\tau)$ from some initial state $\Psi(R,0)$ is given by

$$\psi(R,\tau)=\sum_n a_n\phi_n(R)\exp(-E_n\tau)\ , \tag{4}$$

where

$$a_n=\int dR\ \psi(R,0)\phi_n(R)\ .$$

Equation (4) decays with time. Actually, the decay process occurs with only positive E_n, but this can always be achieved by subtracting constant energy E_T from the Hamiltonian. It is necessary to choose a trial ground-state energy E_T which is nearly equal to the true energy

E_0, so that at large time, all the excited states decay out and the dominant component in the wave function (or density) is the one with the lowest eigenvalue. Thus, assuming that one has a good guess for the ground-state energy, the asymptotic behavior of Eq. (4) is

$$\psi(R,\tau)\rightarrow a_0\phi_0(R)\exp[-(E_0-E_T)\tau]\simeq a_0\phi_0(R)\ , \qquad (5)$$

where $n=0$ is referred to the ground state of the wave function.

The value of E_T is refined during the course of the simulation, so that the density ψ is independent of τ for large τ.

The GFMC method does not attempt to simulate the Schrödinger equation directly in real time, but substitutes the diffusion process. The density evolves with time until it approaches the ground-state wave function. One may say that the GFMC method simulates the Schrödinger equation in imaginary time. These formal manipulations establish the equivalence of the Schrödinger equation in imaginary time with the diffusion equation in real time.

In the simulation of quantum systems, the Green's function for the Schrödinger equation plays a key role. The Green's function is defined as

$$-\tfrac{1}{2}\nabla^2_{R_2}G(R_2,\tau_2,R_1,\tau_1)+V(R-E_T)G(R_2,\tau_2,R_1,\tau_1) = -\frac{\partial G(R_2,\tau_2,R_1,\tau_1)}{\partial\tau_2}\ , \qquad (6)$$

with the boundary condition that in the limit $\tau_2\rightarrow\tau_1$, $G(R_2,\tau_1,R_1,\tau_1)=\delta(R_2-R_1)$. Here the trial energy E_T is subtracted from the both sides of the equation. The time evolution of the density (or wave function) can be done with this propagator G by the integral equation

$$\psi(R_2,\tau_2)=\int dR_1 G(R_2,\tau_2,R_1,\tau_1)\psi(R_1,\tau_1)\ . \qquad (7)$$

Given the ability to calculate G, Eq. (7) provides one with the ability to perform the simulation.

The physical interpretation of this equation is that, for given time τ_1 at position R_1, the probability of moving a particle to a position R_2 at τ_2 is proportional to the Green's function, i.e., a configuration at position R_2 at τ_2 is sampled from the propagator $G(R_2,\tau_2,R_1,\tau_1)$.

B. Short-time approximate Green's function

Since the exact analytic form of the Green's function is not available, one needs to construct certain approximations to it. In this section the short-time approximation to the GFMC will be explained. Once the Green's function with the short-time approximation is supplied, Eq. (7) is iterated for large times.

For sufficiently short time,

$$\tau=\tau_2-\tau_1\ ,$$

such that τ is very small, the configuration does not propagate very far. Since the displacement of a configuration from an initial position R_1 is small, one can assume that the potential energy is approximately constant in the vicinity of R_1,

$$U=V(R_2)-E_T\simeq\text{const}\ .$$

If the potential were a constant, then Eq. (6) would become

$$-\tfrac{1}{2}\nabla^2 G_u(R_2,R_1,\tau)+UG_u(R_2,R_1,\tau) = -\frac{\partial G_u(R_2,R_1,\tau)}{\partial\tau}\ , \qquad (8)$$

with the boundary condition $G(R_2,R_1,0)=\delta(R_2-R_1)$. The solution for this equation is well known for a constant potential U,

$$G_u(R_2,R_1,\tau) = (2\pi\tau)^{-3N/2}\exp[-(R_2-R_1)^2/2\tau-U\tau]\ . \qquad (9)$$

This approximate equation is employed in the short-time GFMC formulation. Knowing the short-time approximate Green's function, the integration of Eq. (7) can be performed by stochastic techniques. In the GFMC method the wave function is supplied as a set of $3N$-dimensional position vectors of electrons, which is called configuration. Given a configuration R_1 at τ_1, the position of R_2 is sampled from the normalized Green's function

$$(2\pi\tau)^{-3N/2}\exp[-(R_2-R_1)^2/2\tau]$$

and the second factor is a measure of the weight which the new position contributes to the integral. One accounts for this weight as follows. First evaluate the weight $W(R_2)=\exp[-U(R_2)\tau]$. If $W(R_2)$ is less than 1, then with probability $W(R_2)$, the configuration at R_2 is retained as a member of the ensemble at time τ. If $W(R_2)$ is greater than 1 but less than 2, then R_2 is accepted once and a second identical configuration is accepted with the probability $W(R_2)-1$ as a representation of $\psi(R_2,\tau_2)$. For $W(R_2)>2$ this process is repeated. Thus the configurations in the ensemble are continually being created or destroyed depending on the weights.

C. Importance sampling

Because the iteration of Eq. (7) is performed stochastically by moving a finite number of configurations forward in time, the number of configurations in the ensemble necessarily fluctuates during the course of the simulation. To reduce these purely statistical variations in the population one needs another variance reduction method. This can be achieved with the introduction of an approximate trial wave function. This trial wave function must be known analytically, and may be obtained from any reasonably accurate approximate method. If one simply multiplies the trial wave function $\psi_T(R_2)$ on both sides of Eq. (7), one may rewrite the propagator equation as

$$\psi_T(R_2)\psi(R_2,\tau_2) = \int dR_1[\psi_T(R_2)G(R_2,R_1,\tau)/\psi_T(R_1)]\psi_T(R_1) \times\psi(R_1,\tau_1)\ , \qquad (10)$$

defining a new density, $f(R,\tau)$, and importance kernel,

$k(R_2,R_1,\tau)$,

$$f(R,\tau)=\psi_T(R)\psi(R,\tau)\ ,$$
$$K(R_2,R_1,\tau)=\psi_T(R_2)G(R_2,R_1,\tau)/\psi_T(R_1)\ . \tag{11}$$

With these definitions, Eq. (10) becomes

$$f(R_2,\tau_2)=\int K(R_2,R_1,\tau)f(R_1,\tau_1)dR_1\ , \tag{12}$$

with the initial condition $K(R_2,R_1,0)=\delta(R_2-R_1)$. An intuitive appreciation of the effect of importance sampling can be obtained by considering Eq. (11). It is the ratio of the trial wave function $\psi_T(R_2)/\psi_T(R_1)$ that gives the biased random walk (biased means that the random walk is drawn to the favored region). For example, when the ratio is less than 1, the probability of moving R_1 to R_2 is decreased. Thus this ratio gives a biased random walk toward higher probability densities and the configurations spend more time in regions favored by the trial wave function. The diffusion analogies with this new propagator equation are described in other places.[5]

What is mostly used is not Eq. (11) but the one with further approximations. Equation (11) can be rewritten as

$$\begin{aligned}K(R_2,R_1,\tau)&=[\psi_T(R_2)/\psi_T(R_1)]G(R_2,R_1,\tau)\\&=G(R_2,R_1,\tau)\{\exp[\ln\psi_T(R_2)-\ln\psi_T(R_1)]\}.\end{aligned} \tag{13}$$

Expanding $\ln\psi_T(R_2)$ in a Taylor series about R_1 and neglecting third-order terms, and also neglecting τ^2 terms, the short-time approximate kernel becomes, in an efficient form,

$$\begin{aligned}K_u(R_2,R_1,\tau)&=(2\pi\tau^*)^{-1/2}\exp[-(R_2-R_1-F\tau^*)^2/2\tau^*]\\&\quad\times\exp\{-[H\psi_T(R_1)/\psi_T(R_1)-E_T]\tau\}\ ,\end{aligned} \tag{14}$$

where

$$\tau^*=\tau/[1-\nabla^2\ln\psi_T(R_1)\tau]\ ,$$
$$F=\nabla\ln\psi_T(R_1)\ . \tag{15}$$

Here we have also neglected some second-order terms of the form

$$\tfrac{1}{2}(R_{2i}-R_{1i})(R_{2j}-R_{1j})\frac{\partial^2}{\partial R_{1i}\partial R_{1j}}\ln\psi_T\ ,$$

where i is not equal to j, and i,j represent x,y,z coordinates.

Inspection of this approximation shows that the effect of importance sampling is to shift the position of the center of the Gaussian toward a favorable region of the trial wave function by an amount $F\tau^*$ and to change the width by τ^*. The weight is

$$\exp\{-[H\psi_T(R_1)/\psi_T(R_1)-E_T]\tau\}$$

and this term is always close to 1 if the trial energy is chosen such that $H\psi/\psi\simeq E_T$. The most important success of this equation is that the weight is almost constant. This is because the total energy does not change much, even though the potential energy changes rapidly. As a result, the number of configurations generated by Eq. (12) is almost constant. Thus, statistical fluctuation is reduced in this way.

III. DIFFICULTIES ASSOCIATED WITH FERMION SYSTEMS

In characterizing the errors introduced by the short-time approximation in many-electron atoms, several complications arise. As before, one does not have an analytic expression for the correct Green's function for the full many-body Schrödinger equation and only the short-time approximate Green's function is available. Furthermore, one does not know the position of the nodes of the true ground-state wave function. Finally, the potential is more complex and in particular it is singular.

Introducing a one-dimensional system, authors have shown[6] that the error in the energy is linear in the size of the time step when the importance kernel was used and the error in the Green's function mostly occurred near the nodes and singularities. In addressing these problems, we have focused our attention on the smallest atomic system that has an antisymmetric ground-state lithium atom.

Lithium has three electrons. The spin of electron 1 can be assigned to spin β (spin down); and electrons 2 and 3, to spin α (spin up), respectively. Since the spin of electrons 2 and 3 is the same, one must choose the spatial wave function to be an antisymmetric function under interchange of r_2 and r_3. One can express the spatial wave function as

$$\psi_D(1,2,3)=\psi_1(r_1)|\psi_1(r_2)\psi_2(r_3)|\ , \tag{16}$$

where ψ_i are the atomic spatial orbitals and $|\,|$ is the Slater determinant. This will be part of our trial wave function in the GFMC calculation. The total wave function is Eq. (16) times the Jastrow factor, which will be discussed later in this section.

In practice, there are several difficulties associated with the simulation of atomic electronic structure. These problems are interrelated, but we have tried to separate them in a way that they can be discussed separately: (a) finite time step, (b) fixed-node approximation, (c) nodal surface, (d) Coulomb repulsion singularity between electrons, (e) Coulomb attractive singularity near nucleus, and (f) infinite variance in the energy.

As explained before, the finite size of the time step is introduced to derive Eqs. (14) and (15). In fact, this causes many difficulties. When this short-time approximate Green's function is derived, it is assumed that the change in potential is almost constant during the finite time step. However, the change in the Coulomb potential near the nucleus is not small. As a result, the Green's function is inaccurate. Therefore, the energy calculated from sampled configurations is inaccurate and a rigorous bound to the true ground-state energy is not guaranteed. Thus the Green's function must be corrected; otherwise, this method is no longer reliable. In fact, the finite size of the time step is interrelated with other problems. This will be expanded upon when each problem is discussed later.

We now discuss how the fixed-node approximation is introduced in simulating the nodal problems. In applying the GFMC method to any quantum problem, the wave function $\psi(R,\tau)$ is interpreted as a positive probability density. However, since the fermion eigenfunction of the Hamiltonian must be antisymmetrized in space and spin variables, $\psi(R,\tau)$ is not, in general, positive definite everywhere. The subspace of particle positions at which a wave function is zero is called the nodal surface of the wave function. Because of the antisymmetric property of the wave function, there are two equivalent regions of space in the wave function which are positive or negative. These regions are separated by a nodal surface. These two equivalent regions are related by the interchange of the coordinates, as required by fermion statistics. Hence, if one knows the function $\psi(R,\tau)$ in one such region, its value in the other regions can be obtained by suitable interchanges of the coordinates. If these nodal surfaces were known exactly in advance, one could use the GFMC to solve the problem by imposing the boundary condition that the solution vanish at the nodes. Unfortunately, the nodal surfaces are not known in advance but come as a part of the solution. However, even though one does not know the nodal surface in the true wave function ψ, one can use an analytic trial wave function $\psi_T(R)$ which "guesses" the true nodal surfaces. Then the true wave function follows the same nodal surface in the whole process as the trial wave function's guesses. Thus this requires that the GFMC solution vanish at the nodal surface prescribed by the trial wave function. This is known as the fixed-node approximation. There are some benefits of using this approximation. First, since the trial wave function gives a rigorous upper bound to the energy with a prescribed nodal structure, an upper bound is still obtained in the GFMC method as long as the Green's function is correct. Second, the trial wave function can be used as an importance function to reduce variance. Third, it is a good approximation resulting in errors in the energy of 1 part in 10^4. If the Green's function satisfies Eq. (6) with the boundary condition $G(R_2,\tau_1,R_1,\tau_1)=\delta(R_2-R_1)$ and one additional boundary condition that the Green's function goes to zero at the nodal surface, then it gives a rigorous upper bound to the energy. Since the exact Green's function is not available in the N-fermion problem, one may use the short-time approximation to the Green's function in practice. Then there is no guarantee that this will give a rigorous upper-bound condition. However, if the approximation satisfies Eq. (6) in the zero-time-step limit with all boundary conditions satisfied, then it gives a rigorous upper bound in the zero-time-step limit. Moskowitz *et al.* proved[4] that the energy obtained by the GFMC method was always a rigorous upper bound to the true ground-state energy for fermions, when the Green's function satisfied the condition that it was zero at the nodes of the trial wave function and satisfied the correct cusp condition near the nucleus. In the same paper, they calculated an energy for LiH that violated this bound; as they noted, the Green's function did not satisfy the correct cusp condition near the nucleus. A possible way of correcting the procedure by using a multiple time step will be discussed in a later section. Thus far, the fixed-node approximation has been discussed. The fixed-node approximation requires that the GFMC solution vanish at the nodal surface. In principle, one can investigate current flow on the nodal surface, by the diffusion analogy, to see what actually happens on the nodal surface. Consider the diffusion-equation analogue of the Schrödinger equation with source term $S(r)$

$$\frac{\partial\rho}{\partial t}+\nabla\cdot\mathbf{J}=S\ ,$$

where $S=(E_T-V)\rho$. For an equilibrium state (stationary state), such as the ground state, $\partial\rho/\partial t=0$ and integrating the above equation over a volume consisting of a region bounded by nodal surfaces gives

$$\int d^3R\ \nabla\cdot\mathbf{J}=\int\mathbf{J}\cdot d\mathbf{a}=-\int d^3R\,(E_T-V)\rho\ ,$$

where $J=-\frac{1}{2}\nabla\rho$. Again,

$$\int\mathbf{J}\cdot d\mathbf{a}=-\int d^3R\,(E_T-V)\rho\ .$$

The interpretation of the equation is as follows. The left-hand side of the equation is the net current flow across the nodal surface and the right-hand side is a small but nonzero quantity because the potential fluctuates as a function of configurations. Thus the current flow across the nodal surface, which is represented by the number of configurations, is not zero since there is a nonzero contribution from the integration of $(E_T-V)\rho$. This is expected because the fermion ground-state wave function is linear in the spatial coordinates and so its derivative evaluated at the nodal surface gives some constant. This phenomenon is altered when importance sampling is introduced. In this case, the current is represented by $J=-\nabla f=-\nabla(\psi\psi_T)$, which is equal to zero because f is proportional to the square of the spatial coordinates, and its derivative, which is evaluated at the nodal surface, is zero. When the short-time approximate Green's function is introduced, the condition that the solution vanish at the nodal surface is violated. In other words, since the short-time approximate Green's function, which we introduced from Eqs. (14) and (15), does not satisfy the antisymmetric condition, some configurations are still sampled from the nodal surface. As a result, the current flow across the nodal surface is not zero. Fortunately this number is very small compared to the total number of configurations during the process. It was noted that, when a bigger time step was used, this number increased. In addition, this number can be used to check how reliable this method can be.

The Jastrow factor has been introduced for the correlations of electrons. Electrons will be in favorable positions when they are far away from each other because of their repulsive forces. The Jastrow factor[7] is defined as

$$\Psi_J=\exp[U(r_{ij})]\ ,$$

where

$$U(r_{ij})=\sum\frac{C_1 r_{ij}}{1+C_2 r_{ij}}\ ,$$

r_{ij} is an interparticle distance, and C_1 and C_2 are variational parameters. This gives a higher probability when electrons are far away from each other and, therefore, reduces statistical fluctuations. Even if one were to introduce the Jastrow factor, there still would exist finite probability since the wave function is finite (Ψ_T is not zero) when electrons are close to each other. When electrons are close to each other the potential is divergent. Even though the potential is divergent the total energy is not divergent because the wave function is finite and its second derivative, which is related to the kinetic energy, is divergent in such a way that those two divergences cancel each other. This is called the cusp condition. A main advantage of using the Jastrow factor is that this can reduce the variance in the calculation, since the random walks are biased by the factor. The total trial wave function is now expressed as a product of the Jastrow factor and the Slater determinant,

$$\Psi=\Psi_D\Psi_J\ ,$$

where Ψ_D is the Slater-determinant wave function.

Item *e* identifies the difficulty which arises in treating the attractive Coulomb singularity between the electron and the nucleus. In the derivation of the short-time approximate Green's function it is assumed that the change in potential during a time interval τ is small. If this condition is satisfied, then Eq. (9) is still valid, and in the limit of zero time step a rigorous upper bound will be obtained. Since there is a Coulomb singularity near the nucleus, the potential changes are large such that the short-time approximate Green's function does not satisfy Eq. (9). Also, the importance kernel does not satisfy the correct cusp condition. In fact, part of this second derivative term, which is related to the kinetic energy, is neglected in the derivation of the short-time approximate importance kernel, Eq. (14). It seems that this effect would be small if the correct Green's function were used. Since the short-time approximate Green's function is used near the nucleus, there is no guarantee that this effect is small. Unless this problem is corrected, this method cannot be reliable. Moskowitz *et al.* introduced a square well for the divergent Coulomb potential such that the depth of the well V_0 was chosen to be

$$V_0=3/4R_w^3\int\left[\frac{1}{r}\right]d\mathbf{r}\ ,$$

where the integral limit is from zero to R_w. In the same paper they also stated that in the limit $\tau\to 0$ and $R_w\to 0$ calculation would give the exact result. A better solution to this problem is to use a method which generates the correct Green's function for this singularity. We will introduce the multiple time step in this work to correct this problem.

There is another serious problem in the derivation of Eq. (14), which is related to the infinite variance in the energy. It is assumed that the force ($F=\nabla\ln\psi_T$) and second derivative ($\nabla^2\ln\psi_T$) are small enough so that $\nabla\ln\psi_T\ll 1$ and $\nabla^2\ln\psi_T\ll 1$. In fact, these conditions are violated near the nucleus and nodal surface. It can be shown as follows. The force term is expressed as

$$F=\nabla\ln\psi=\nabla\psi/\psi\ .$$

Since ψ is linear in the spatial coordinates and finite at the nucleus, its derivative is some constant, and so the force term is constant, not divergent. But at the nodal surface, ψ is zero, thus the force term is not constant but divergent. The second derivative term is expressed by two terms,

$$\nabla^2\ln\psi=\nabla^2\psi/\psi-|\nabla\psi|^2/\psi^2\ .$$

At the nucleus, even if the first derivative of the wave function is finite (because of the "cusp"), its second derivative is divergent. At the nodal surface, because ψ, which goes to zero, occurs in the denominator of each term, while the numerator of each term is finite, the second derivative of $\ln\psi$ is divergent. In summary, the first derivative is finite but the second derivative is divergent at the nucleus, and both the first- and second-derivative terms are divergent at the nodal surface. Thus the short-time approximate importance kernel is not correct near the nucleus and the nodal surface. In particular, since this kernel is not correct near the nucleus, a correct cusp condition is not satisfied. As a result the energy is still finite (due to a wrong cusp condition) but shows an infinite variance. In Sec. IV we will show how the infinite variance can be removed.

IV. MULTIPLE-TIME-STEP SAMPLING TECHNIQUE

In Sec. III we argued that a number of approximations are present in the usual implementation of the short-time approximation which cause it to be untenable as a reliable tool in investigating many-fermion problems. In this section we will present a resolution of these problems which is consistent with the original philosophy of the short-time GFMC formalism. In other words, we will present a method of implementing these ideas which has the virtue that the errors due to the finite size of the time step do indeed go to zero as the time step goes to zero.

Our goal is to fix the short-time approximate Green's function by using a multiple-time-step sampling technique and reduce all other errors at least by the same order of magnitude as the size of the time step. We will introduce a multiple-time-step sampling technique in this section and argue that this solves the problem at the nucleus and the nodal surface.

The multiple-time-step method is constructed by noting that the Green's function can be propagated forward in time by itself,

$$G(R_3,\tau_3,R_1,\tau_1)=\int G(R_3,\tau_3,R_2,\tau_2)G(R_2,\tau_2,R_1,\tau_1)dR_2\ . \qquad (17)$$

This says that the Green's function at a large time τ_3 can be obtained from Green's function at an earlier time τ_2 by propagating it forward in time using $G(R_3,\tau_3,R_2,\tau_2)$. In principle, the above process may be repeated indefinitely. Then Eq. (17) becomes

$$G(R_f,\tau_f,R_i,\tau_i)=\int\cdots\int dR_n\cdots dR_1 G(R_f,\tau_f,R_n,\tau_n)\cdots G(R_1,\tau_1,R_i,\tau_i)\ . \tag{18}$$

The Green's function which is propagated from the initial time τ_i to final time τ_f can be expressed as the product of many intermediate Green's functions G_j with a smaller time step $\Delta\tau_j<\Delta\tau_0$, where the original time step is $\Delta\tau_0=\tau_f-\tau_i$. It is also noted that $\Delta\tau_n$ does not have to be the same in the process. One can write down a similar expression for an importance kernel,

$$K(R_f,\tau_f,R_i,\tau_i)=\int\cdots\int dR_1 dR_2\cdots dR_n K(R_f,\tau_f,R_n,\tau_n)K(R_n,\tau_n,R_{n-1},\tau_{n-1})\cdots K(R_1,\tau_1,R_i,\tau_i)\ . \tag{19}$$

In this expression each kernel has a weight. If one separates the weight from each kernel, it may be rewritten as

$$\begin{aligned}WK'(R_f,\tau_f,R_i,\tau_i)=W_1\cdots W_n\int\cdots\int dR_1 dR_2\cdots dR_n K'(R_f,\tau_f,R_n,\tau_n)\\ \times K'(R_n,\tau_n,R_{n-1},\tau_{n-1})\cdots K'(R_1,\tau_1,R_i,\tau_i)\ ,\end{aligned} \tag{20}$$

where primed coordinates represent the kernel without its weight factor. Thus the total weight during the propagation of this kernel from τ_i to τ_f can be expressed by

$$W=\prod_i W_i\ , \tag{21}$$

where W_i is the ith weight of the kernel in each step.

The way to decide the intermediate time step τ_j is shown as follows. As explained earlier, the force and the second derivative of the trial wave function must be small so that the importance kernel, Eq. (14), is valid. Fortunately the force and the second derivative term can be monitored to determine exactly when and where these terms are big. Whenever they become large, that configuration will use the multiple time steps of Eq. (19). Instead of a large-time propagation $\Delta\tau_0=\tau_f-\tau_i$, many small substeps will be chosen. The question of how many steps should be taken is decided in the following way. Let $\Delta\tau_0$ be the initial time step and define two more time steps such that

$$\begin{aligned}\Delta\tau_a&=A/f_1\ ,\\ \Delta\tau_b&=B/f_2\ ,\end{aligned} \tag{22}$$

where f_1 is the force $(=\nabla\ln\psi_T)$,f_2 is the second derivative term $(=\nabla^2\ln\psi_T)$, and A and B are some arbitrary constants. Wherever f_1 or f_2 becomes large, which violates the validity of the approximate kernel, a small enough time step should be used so that those derivatives times the step are small enough to recover the validity of the approximation. Among these two time steps, the smaller will be chosen because the approximation is more seriously violated for the smaller time step. So the next time step may be chosen as

$$\Delta\tau_j=\min(\Delta\tau_0,\Delta\tau_a,\Delta\tau_b)\ , \tag{23}$$

where min means the minimum of those variables. For example, if $\Delta\tau_0$ is chosen, then the error will be the same order of magnitude as the size of the time step. In this case it becomes a normal process, i.e., Eq. (19) is not taken. If $\Delta\tau_a$ or $\Delta\tau_b$ are chosen, it is a sign that the force or second derivative are big. Then the multiple-time-step Green's function Eq. (19) is taken. This process will be finished when $\Delta\tau'$, which is defined by

$$\Delta\tau'=\Delta\tau_0-\Delta\tau_j\ ,$$

becomes zero so that it achieves the zero time limit near the nucleus and nodal surface. Since each subprocess in Eq. (18) satisfies the approximation, Eq. (8), the resulting Green's function is regarded as the correct one. Similarly, since the subprocess in Eq. (19) satisfies the short-time approximation conditions, the resulting kernel is taken as the correct one. After this process the resulting configuration will be taken with the total weight expressed in Eq. (21).

What we have done here is to sample an accurate Green's function near the nucleus and nodal surface. It can be explained as follows how these two problems are treated properly with the multiple-time-step sampling technique.

First, because of the small time step used near the nucleus, the change in potential is small. Thus the approximate Green's function, Eq. (8), is still valid. The error from this process is reduced approximately to the same order of magnitude as the size of the time step. Also, the cusp condition is satisfied, since the Green's function and the approximate kernel are fixed. As a result, the infinite variance is removed and the error in the energy will be the same order as the size of the time step.

Second, the importance kernel was not correct near the nodal surface. With the introduction of the multiple time step, the Green's function was improved and, furthermore, the importance kernel was corrected such that fewer configurations were sampled near the nodal surface. Since the antisymmetric condition is still not satisfied in the importance kernel, the boundary condition that the solution vanish at the nodal surfaces is still violated, but with the multiple-time-step sampling technique, fewer configurations are sampled near the nodal surface. By doing this, the error in the importance kernel is reduced by approximately the same order of magnitude as the size of the time step which is consistent with the initial assumption when it was derived.

The only remaining source of errors is from the finite size of the time step and the fixed-node approximation. These two problems are inherent in the method. The finite size of the time step can be removed by taking the limit to zero time step. The energy for the zero time step can be obtained from extrapolation. This will be shown in Sec. V.

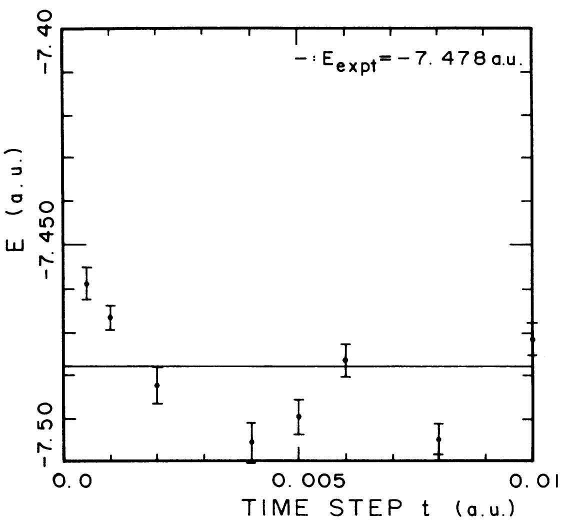

FIG. 1. Calculations of the ground-state energy of the lithium atom in terms of time steps from the old method.

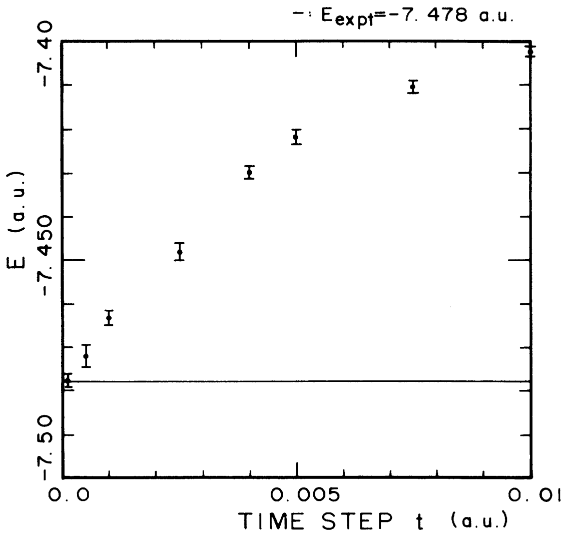

FIG. 2. Calculations of the ground-state energy of the lithium atom in terms of time steps from the multiple-time-step sampling technique.

V. COMPARISON AND RESULTS

We will provide numerical results in this section which show the conventional short-time approximate GFMC fails and how the new method introduced in this study works. We chose to study the lithium atom since it has an antisymmetric wave function (i.e., helium has a symmetric wave function). Because it is highly desirable to supply the best initial information to the configurations, a variational Monte Carlo (VMC) calculation was done. Another purpose of doing this is to compare its results with the GFMC results and see the improvement.

Figure 1 shows the behavior of the old method. Illustrated is the ground-state energy of the lithium atom for various sizes of the time step. Total run time was 40 a.u. before and after the equilibrium. When $\tau=0.001$ was used, the run time was 10 h with the IBM 3081D. The parallel line in the figure is the experimental value for a lithium atom. As explained, since the short-time approximate importance kernel is inaccurate near the nodal surface and the nucleus, it turns out to be totally unreliable. The scatter of energy for different sizes of the time step is definite evidence of the infinite variance mentioned earlier. Furthermore, it is clear that the exact energy would not be obtained from the extrapolation to the zero time limit.

Shown in Fig. 2 are the results of the new method introduced in this study. Since the Green's function is corrected properly with the introduction of the multiple-time-step sampling technique, the problems near the nodal surface are reduced and the cusp condition is satisfied. As a result, the infinite variance is removed and the error generated from the nodes and Coulomb singularities is approximated to an order of the same magnitude as the finite size of the time step. In other words, most of the errors result from the inherent finite size of the time step. Clearly, the errors in the energy are larger for larger time steps. This behavior is very predictable because the exact ground-state energy can be obtained from the extrapolation to zero time step. Also notice that for small time steps the curve is linear in the step which is consistent with the one-dimensional case. The extrapolation gives very accurate results. For example, the ground-state energy for the lithium atom is 7.478 a.u. experimentally and our answer from the extrapolation gives 7.479±0.0022 a.u. This variance would be reduced

TABLE I. Comparison of HF, VMC, and GFMC methods for various nuclei (all in a.u.). The numbers in parentheses are uncertainty figures.

Z	HF	VMC	GFMC	E_{corr}
3	−7.4327	−7.450(7)	−7.4790(22)	0.0463
4	−14.2774	−14.287(9)	−14.3149(72)	0.0462
5	−23.3760	−23.399(7)	−23.4234(35)	0.0475
6	−34.7261	−34.752(12)	−34.7743(38)	0.0482
7	−48.3268	−48.324(17)	−48.3537(42)	0.0407
8	−64.1780	−64.164(19)	−64.2156(58)	0.0376

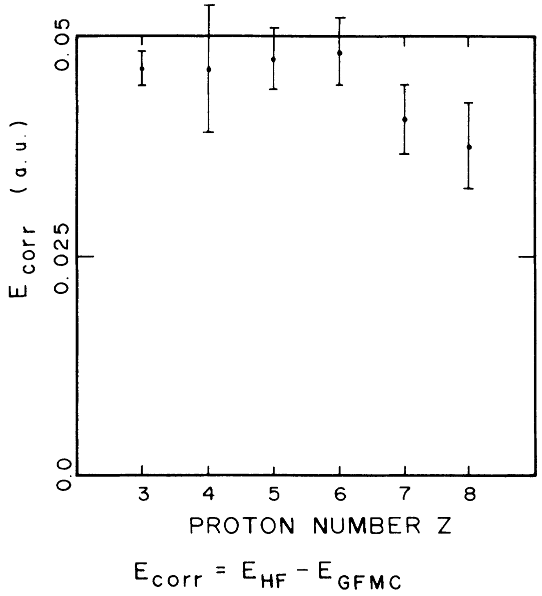

FIG. 3. Correlation energies for various nuclei.

if longer runs were taken. In addition to this calculation of the lithium atom we did some more calculations for lithium-like atoms (arbitrary z with three electrons) using the multiple-time-step sampling technique. Table I shows the results of the variational Monte Carlo method, which we calculated, the Hartree-Fock (HF) method,[8] and our new method. The VMC method gives a little better result than the HF method, since a better trial wave function (with the Jastrow factor) was used. Table I also shows that this improvement is not noticeable for heavier atoms because of the statistical uncertainty. The main goal of this calculation is the GFMC results. It will be convenient to define a correlation energy as the result of the HF method minus that of the GFMC method. Consistently the GFMC method gives better ground-state energy than the HF method. Illustrated in Fig. 3 are the correlation energies for various nuclei. For heavier nuclei, it is observed that a smaller time step is required to achieve the linear fit for the extrapolation. The increase of the correlation energies for heavier nuclei is not noticeable. This is the evidence that the HF approximation, which basically comes from the independent electron model, improves for heavier nuclei, since the electrons become more independent because of smaller screening effects for heavier nuclei.

VI. CONCLUSION

It has been shown that the error in the Green's function mostly occurred near the nodes and Coulomb singularities. We were able to separate the sources of errors from the finite size of the time step and those from nodes and Coulomb singularities and the fixed-node approximation. As a result, a correct treatment near the nodes and Coulomb singularities using a multiple-time-step sampling technique recovered most of the experimental values. Our calculation gave 7.479±0.0022 a.u. for the lithium atom, while the experimental value was 7.478 a.u. What it means is that the fixed-node approximation, which is inherently introduced, is a very good approximation.

---

*Present address: Department of Physics, Jeonbug National University, Jeonju, Jeonbug 560-756, Korea.

[1] D. M. Ceperly and M. H. Kalos, in *Monte Carlo Method in Statistical Physics*, edited by K. Binder (Springer, Berlin, 1979), Chap. 4.

[2] M. H. Kalos, M. A. Lee, P. A. Whitlock, and G. V. Chester, Phys. Rev. B **24**, 115 (1981).

[3] D. M. Arnow, M. H. Kalos, M. A. Lee, and K. E. Schmidt, J. Chem. Phys. **77**, 5562 (1982); M. A. Lee, K. A. Motakabbir, and K. E. Schmidt, Phys. Rev. Lett. **53**, 1191 (1984); M. A. Lee, K. E. Schmidt, M. H. Kalos, and G. V. Chester, Phys. Lett. **46**, 728 (1981); J. B. Anderson, J. Chem. Phys. **63**, 1499 (1975); R. K. Kalia, P. Vashishta, and M. A. Lee, Solid State Commun. **52**, 873 (1984).

[4] J. W. Moskowitz, K. E. Schmidt, M. A. Lee, and M. H. Kalos, J. Chem. Phys. **77**, 349 (1982).

[5] Y. H. Lee, Ph.D thesis, Kent State University, 1986 (unpublished) (microfilm available through the University of Michigan).

[6] Y. H. Lee and M. A. Lee, J. Phys. Soc. Jpn. **58**, 91 (1989).

[7] K. E. Schmidt, M. A. Lee, and M. H. Kalos, J. Chem. Phys. **77**, 5562 (1982), and references therein.

[8] C. S. Sharma, Proc. Phys. Soc., London **80**, 839 (1962).