Title
FIXED-NODE QUANTUM MONTE CARLO FOR MOLECULES

Author
Reynolds, P.J.

Publication Date
1982-04-01

LBL-13356
Preprint

NRCC NATIONAL RESOURCE FOR COMPUTATION IN CHEMISTRY

Submitted to the Journal of Chemical Physics

FIXED-NODE QUANTUM MONTE CARLO FOR MOLECULES

Peter J. Reynolds, David M. Ceperley, Berni J. Alder, and William A. Lester, Jr.

April 1982

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA

Prepared for the U.S. Department of Energy under Contract DE-AC03-76SF00098 and W-7405-ENG-48, and the National Science Foundation under Interagency Agreement CHE-7721305

DISCLAIMER

LBL-13356
Preprint

FIXED-NODE QUANTUM MONTE CARLO FOR MOLECULES

Peter J. Reynolds and David M. Ceperley

National Resource for Computation in Chemistry,
Lawrence Berkeley Laboratory, University of California
Berkeley, California 94720

Berni J. Alder

Lawrence Livermore National Laboratory
University of California, Livermore, California 94550

and

William A. Lester, Jr.

National Resource for Computation in Chemistry
Lawrence Berkeley Laboratory, University of California
Berkeley, California 94720

This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Chemical Sciences Division of the U. S. Department of Energy under Contract Nos. W-7405-ENG-48 and DE-AC03-76SF00098 and by the National Science Foundation under Grant No. CHE-7721305.

FIXED-NODE QUANTUM MONTE CARLO FOR MOLECULES*§

Peter J. Reynolds† and David M. Ceperley¶

National Resource for Computation in Chemistry,
Lawrence Berkeley Laboratory, University of California
Berkeley, California 94720

Berni J. Alder

Lawrence Livermore National Laboratory
University of California, Livermore, California 94550

and

William A. Lester, Jr.§§

National Resource for Computation in Chemistry
Lawrence Berkeley Laboratory, University of California
Berkeley, California 94720

Abstract

The ground-state energies of H_2, LiH, Li_2, and H_2O are calculated by a fixed-node quantum Monte Carlo method, which is presented in detail. For each molecule, relatively simple trial wavefunctions, Ψ_T, are chosen. Each Ψ_T consists of a single Slater determinant of molecular orbitals multiplied by a product of pair-correlation

*This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Chemical Sciences Division of the U. S. Department of Energy under Contract Nos. W-7405-ENG-48 and DE-AC03-76SF00098 and by the National Science Foundation under Grant No. CHE-7721305.

§ A preliminary account of this work was presented at the American Conference on Theoretical Chemistry, Boulder, CO, 22-26 June 1981.

† Permanent address: Department of Physics, Boston University, Boston, MA 02215 and Materials and Molecular Research Division, Lawrence Berkeley Laboratory, University of California, Berkeley, CA 94720.

¶ Permanent address: Lawrence Livermore National Laboratory, Livermore, CA 94550.

§§ Permanent address: Materials and Molecular Research Division, Lawrence Berkeley Laboratory, and Department of Chemistry, University of California, Berkeley, CA 94720.

(Jastrow) functions. These wavefunctions are used as importance functions in a stochastic approach that solves the Schroedinger equation by treating it as a diffusion equation. In this approach, Ψ_T serves as a "guiding function" for a random walk of the electrons through configuration space. In the fixed-node approximation used here, the diffusion process is confined to connected regions of space, bounded by the nodes (zeroes) of Ψ_T. This approximation simplifies the treatment of Fermi statistics, since within each region an electronic probability amplitude is obtained which does not change sign. Within these approximate boundaries, however, the Fermi problem is solved exactly. The energy obtained by this procedure is shown to be an upper bound to the true energy. For the molecular systems treated, at least as much of the correlation energy is accounted for with the relatively simple Ψ_T's used here as by the best configuration interaction calculations presently available.

1. Introduction

Accurate calculations of molecular properties such as binding energies, bond strengths, charge distributions, and potential energy surfaces are an important goal of quantum chemistry. Most approaches in use today for such calculations involve some combination of the multi-configuration self-consistent-field and configuration interaction (CI) methods,[1] or involve many-body perturbation theory.[2] In this paper, however, we use a quantum Monte Carlo scheme which in principle can give exact results.

Configuration interaction wavefunctions, for example, have been able to account typically for about 75% of the correlation energy[3] of a molecule such as water.[4] However, much interesting chemistry occurs on an energy scale of only a fraction of the correlation energy. For example, the O-H bond strength in water is about 50% of the correlation energy. Thus, the ground-state energy computed using large CI wavefunctions differs from the exact (non-relativistic, Born-Oppenheimer) energy by an amount on this same order of magnitude. Furthermore, it can be difficult to improve CI results because convergence to the exact result is slow and can be non-uniform.[5] An additional limitation with CI is that the computational effort increases with somewhere between the fourth and fifth power of the number of electrons in the molecular system, effectively restricting the size of molecules the method can treat.

A direction that shows promise, and avoids the inherent limitations of expansion approaches, is the use of quantum Monte Carlo (QMC) methods.[6-18] These methods were developed and used primarily in the

fields of nuclear and condensed matter physics. Only recently have chemical calculations by QMC methods been carried out.[11,14,17,18] QMC methods are of both the variational type,[8,12,18] in which the Monte Carlo method is used to numerically evaluate expectation values obtained from a given (generally optimized) trial wavefunction Ψ_T, and of the "exact" type in which the Schroedinger equation is solved. In these latter approaches it is not necessary to compute a highly accurate wavefunction in order to determine molecular properties. Instead, these QMC methods use various procedures to stochastically sample the exact wavefunction, $\Phi(\underset{\sim}{R})$, of a molecular system, subject only to statistical errors. Properties of interest are in effect "measured" as the system evolves under the Schroedinger equation. When a stationary state is obtained, averages of the measured quantities provide the desired expectation values.

Recent developments in the exact QMC methods include a reduction in statistical error[10,14-17] by use of importance sampling,[19] and the ability to treat Fermi statistics[11,14-17]. These developments are described here, and have been used in the calculation of the ground-state energies of H_2, LiH, Li_2, and H_2O which we report. We use a "fixed-node" approximation[15-17] to treat Fermi statistics. This approximation may be removed, however, as has been done for the homogeneous electron gas.[15] It is employed here because it simplifies the calculation, and can be argued on physical grounds as representing a very good approximation. A postiori, the fixed-node approximation is justified by the results, since this procedure yields accuracies comparable to or better than the CI method. Furthermore, in this

stochastic approach the computational time rises only with the second power of the number of electrons in the molecule.[20]

The outline of this paper is as follows. In Section 2, we present the stochastic diffusion method for solving the Schroedinger equation. We also discuss importance sampling, the choice of the trial function Ψ_T, and the fixed-node approximation used to treat Fermi statistics. Furthermore, we demonstrate that the fixed-node approximation retains the character of a variational method--i.e., that the calculated energy is an upper bound to the true energy. In Section 3 we outline the algorithm used for the QMC calculations. We present and discuss results for H_2, LiH, Li_2, and H_2O in Section 4. A summary and conclusion comprise Section 5. Finally, the Appendix gives the details of the trial wavefunctions used in this study.

2. Diffusion Monte Carlo

A starting point of this approach is to write the Schroedinger equation in imaginary time as

$$-\frac{\partial \Phi(\underset{\sim}{R},t)}{\partial t} = [\,-D\nabla^2 + V(\underset{\sim}{R}) - E_T]\,\Phi(\underset{\sim}{R},t) \quad . \tag{1}$$

Here $D \equiv \hbar^2/2m_e$, $\underset{\sim}{R}$ is the 3N-dimensional vector specifying the coordinates of the N electrons of the molecule, t is imaginary time measured in units of $\hbar$, and

$$V(\underset{\sim}{R}) = \sum_{i>j} \frac{e^2}{r_{ij}} - \sum_{i,\alpha} \frac{Z_\alpha e^2}{r_{i\alpha}} + \sum_{\alpha>\beta} \frac{Z_\alpha Z_\beta e^2}{r_{\alpha\beta}} \quad , \tag{2}$$

is the potential energy of the molecule. Also, $r_{ab} \equiv |\vec{r}_a - \vec{r}_b|$,

Roman indices label electronic coordinates, Greek indices label nuclear coordinates, e is the electron charge, and Z_α is the atomic number of nucleus α. E_T represents a constant shift in the zero of energy, whose introduction proves useful.

The objective is to obtain the solution to the time-independent Schroedinger equation. As is apparent, this may be obtained from a steady-state solution to Eq.(1). Let us solve Eq.(1) by expanding $\Phi(\underset{\sim}{R},t)$ in a complete set of eigenfunctions $\phi_i(\underset{\sim}{R})$ of the Hamiltonian, and substituting this expansion into Eq. (1). One finds

$$\Phi(\underset{\sim}{R},t) = \sum_i N_i e^{-(E_i-E_T)t} \phi_i(\underset{\sim}{R}) \quad , \tag{3}$$

where E_i are the energy eigenvalues corresponding to the $\phi_i(\underset{\sim}{R})$. The coefficients N_i depend on the initial conditions. At sufficiently long times, only the term with the lowest energy survives in Eq. (3). Hence if $N_0 \neq 0$, the asymptotic solution to Eq. (1) is

$$\Phi(\underset{\sim}{R},t) = N_0 e^{-(E_0-E_T)t} \phi_0(\underset{\sim}{R}) \quad . \tag{4}$$

If E_T is adjusted to be the true ground-state energy, E_0, the asymptotic solution <u>is</u> a steady-state solution, corresponding to the ground-state eigenfunction ϕ_0. If, however, we require explicitly that Φ be orthogonal to ϕ_0, then $N_0=0$ in Eq.(3) and the asymptotic solution gives the first excited state.

Equation (1) is a diffusion equation in a 3-N dimensional space, and as such may be readily simulated.[11] Here $\Phi(\underset{\sim}{R},t)$ plays the role

of the density of diffusing particles. If the $[E_T - V(\underset{\sim}{R})]\Phi$ term were absent, Eq.(1) would be the usual diffusion equation, with a diffusion constant D--the coefficient of the Laplacian. This simple equation can be simulated by a random walk of particles through configuration space--the well-known "drunkard's walk". On the other hand, if the term $[E_T - V(\underset{\sim}{R})]\Phi$ were present alone on the right-hand side, Eq. (1) would be a rate equation, describing branching processes such as radioactive decay or exponential birth and death processes in a population. Thus, the entire equation can be simulated as a combination of a diffusion and a branching process, in which the number of diffusers increases or decreases at a given point proportional to the density of diffusers already there. This branching serves to decrease the probability density in regions where $V(\underset{\sim}{R})$ is large, and enhance it in regions of favorable potential energy.

For the diffusion interpretation to be valid, however, Φ must always be positive, since it is a population density. Φ may also be everywhere negative, since the overall phase of the wavefunction is arbitrary. Thus, at first glance, it seems that the process is restricted to functions $\phi_0(\underset{\sim}{R})$ that have no nodes, such as for Bose systems in their ground state. If, however, ϕ_0 does have nodes, and hence changes sign as in a Fermi system, the apparent limitation of the diffusion analogy can be dealt with by treating positive and negative regions separately. If we do not allow diffusion between these regions, we have the fixed-node approximation. Releasing this constraint will not be pursued here, but will be the subject of a subsequent paper.

Importance Sampling

Solving Eq. (1) by a random-walk process with branching is inefficient, because the branching rate--which is proportional to the Coulomb potential $V(\underset{\sim}{R})$--can diverge to $\pm\infty$. This leads to large fluctuations in the number of diffusers, and to slow convergence when calculating averages such as $\langle V(\underset{\sim}{R})\rangle$ and hence E_0. However, the fluctuations, and hence the statistical uncertainties, can be greatly reduced[10,14-17] by the technique of importance sampling[19].

Importance sampling enables one to work with a probability distribution other than $\Phi(\underset{\sim}{R})$, to obtain the same averages. Anderson[14] has explored several ways of using known properties of the ground state to reduce statistical fluctuations. Here, however, we use a simple method introduced earlier by Kalos for many-Boson systems[10], since this method is most readily generalized to many-electron systems, and provides a convenient way of dealing with Fermi statistics. One simply multiplies Eq.(1) by a known trial function, $\Psi_T(\underset{\sim}{R})$, and rewrites it in terms of a new probability distribution $f(\underset{\sim}{R},t)$ given by

$$f(\underset{\sim}{R},t) \equiv \Phi(\underset{\sim}{R},t)\ \Psi_T(\underset{\sim}{R}) \quad . \tag{5}$$

Rearranging terms in the resultant equation leads to

$$-\frac{\partial f(\underset{\sim}{R},t)}{\partial t} = -D\nabla^2 f + (E_L(\underset{\sim}{R}) - E_T)\ f + D\nabla\cdot(fF_Q(\underset{\sim}{R})) \quad . \tag{6}$$

Here $E_L(\underset{\sim}{R}) \equiv H\Psi_T/\Psi_T$ is the local energy obtained from the trial function, and $F_Q(\underset{\sim}{R}) \equiv \nabla \ell n|\ \Psi_T(\underset{\sim}{R})\ |^2 = 2\nabla\Psi_T(\underset{\sim}{R})/\Psi_T(\underset{\sim}{R})$. The quantity F_Q

plays the role of a "quantum force," as may be seen from a classical analogue. If we equate the quantum mechanical probability distribution $|\Psi_T|^2$ with a Boltzmann probability $e^{-\beta U}$, (thereby defining U), then F_Q is proportional to the force due to the potential U. In fact, even in the quantum case, Ψ_T is often written as an exponential of a sum over pseudopotentials.[21]

Equation (6), which incorporates importance sampling through Ψ_T, is a diffusion equation for a density function $f(\underset{\sim}{R},t)$. The branching term is now proportional to the "excess local energy" $(E_L(\underset{\sim}{R}) - E_T)$, which, unlike the original branching term, with a good choice of Ψ_T need not become singular when $V(\underset{\sim}{R})$ does. Thus, to control branching we need to choose Ψ_T such that E_L is everywhere as smooth as possible.[22] In particular, Ψ_T should have the correct cusp behavior as any two particles approach each other.[23] As Ψ_T better approximates the correct wavefunction, $E_L(\underset{\sim}{R})$ will tend to approach E_0 throughout configuration space. As a consequence, the excess local energy becomes independent of $\underset{\sim}{R}$, and branching can be greatly reduced by a proper choice of E_T.

Also, an additional term $D\nabla\cdot(fF_Q)$ now appears in Eq.(6). This new term acts to impose a directed drift velocity on the diffusion, just as a similar term in the Smoluchowski equation[24] gives the correction to Brownian diffusion in an external potential. In regions of low probability--where $\Psi_T(\underset{\sim}{R})$ is small--one can see that $F_Q(\underset{\sim}{R})$ is large, and hence any diffusers reaching such a region are driven away. Thus, the advantage of Eq. (6) over Eq. (1) is that the diffusion process for $f(\underset{\sim}{R},t)$ is guided by Ψ_T (through the force F_Q), so that sampling is

performed preferentially in regions where Ψ_T is large. Hence, it is evident that importance sampling will be most useful if Ψ_T is a good approximation to ϕ_0. In fact, as $\Psi_T \to \phi_0$, $E_L(\underset{\sim}{R}) \to E_0$ independent of $\underset{\sim}{R}$. Thus, in this case, the variance and hence the statistical uncertainty of $\langle E_L \rangle$ will vanish. In practice, importance sampling with a good approximate trial function containing whatever information is known about the exact wavefunction--such as the cusp conditions--yields averages with much lower statistical uncertainties than can be obtained without importance sampling.

Note that from Eqs. (4) and (5), the asymptotic solution to Eq.(6) is

$$f(\underset{\sim}{R},t) = \Psi_T(\underset{\sim}{R})\ \phi_0(\underset{\sim}{R})\ e^{-(E_0-E_T)t} \quad . \tag{7}$$

By adjusting E_T one may achieve a steady-state solution where, on the average, the branching leads to no net change in the population. The value of E_T obtained in this way is the energy E_0. Actually, as we shall see later, the average $\langle E_L \rangle$, taken as the electrons diffuse, will also yield E_0--even when Ψ_T is approximate. Were it not for the branching term, however, this average would be equal to the expectation value $\langle \Psi_T | H | \Psi_T \rangle$, since the solution to Eq.(6) would then be $f = |\Psi_T|^2$. In other words, without branching we would obtain the variational energy of Ψ_T, rather than E_0.

Sampling from the distribution $f = |\Psi_T|^2$ is the basis of the variational QMC method.[8,12,18] Although, with good choices of Ψ_T, the variational QMC procedure can yield accurate results, there is little justification in using that method when a minor modification in the

stochastic procedure (i.e. adding the branching term) can eliminate the variational approximation. Thus, in the procedure used in the present computations it is possible to obtain higher accuracy than the variational method since Ψ_T is only the starting point, and branching serves to correct the distribution $|\Psi_T|^2$ in regions where it is poor. We discuss the details of the fixed-node QMC solution to Eq. (6) and the calculation of E_0 in Section 3. First, we discuss our choice of Ψ_T, and the problem associated with Fermi statistics.

Choice of Ψ_T

As discussed above, the role of $\Psi_T(\underset{\sim}{R})$ in the fixed-node QMC method is that of a guiding function for importance sampling. As such, its role is primarily in variance reduction--i.e. a better Ψ_T leads to smaller statistical errors for the same amount of sampling. Thus, in attempting to reduce the statistical error there is a tradeoff between using a more elaborate form for Ψ_T (which generally takes longer to compute) and using a simple Ψ_T (which must be sampled more times). In the fixed-node approximation, however, Ψ_T also determines the location of the nodal surfaces, where the approximate $\phi_0(\underset{\sim}{R})$ must vanish. How well the nodes are represented will determine how close to E_0 one can ultimately come. Once the nodes are established, however, the choice of Ψ_T affects only the variance and not the expectation value of the energy. In this paper the optimum mix of simplicity and accuracy in choosing Ψ_T has not been investigated. Instead, several simple forms for Ψ_T are used to minimize the time necessary for computing the trial function and its derivatives. The choices are specified in the Appendix.

In the cases treated here, molecular orbitals ψ_k are formed from linear combinations of atomic Slater-type orbitals (LCAO), or are Gaussian-like orbitals localized to the vicinity of a single nucleus. These in turn are used to form a Slater determinant with the symmetry of the ground state. To allow explicitly for electron correlation in the wavefunction, the determinant is multiplied by a Jastrow pair-correlation factor[21] of the form $\exp(\sum_{ij} U_{ij})$ where $U_{ij} = ar_{ij}/(1 + br_{ij})$. This Padé form is the simplest function having the desired properties, that U_{ij} be linear in r_{ij} at small r_{ij} to satisfy the cusp condition[23], and that U_{ij} approach a constant + $O(1/r_{ij})$ so that the wavefunction factors at large r_{ij}. Thus,

$$\Psi_T(\underset{\sim}{R}) = \det|D^{\alpha}_{k\ell}|\ \det|D^{\beta}_{k\ell}|\exp(\sum_{i>j} U_{ij}) \quad , \tag{8}$$

where[25] $D^{s}_{k\ell} = \psi_k(\vec{r}_\ell;s)$, s is the spin state, and ψ_k is the $k^{\underline{th}}$ molecular orbital. Trial wavefunctions of the form (8) have been shown to be quite reasonable in variational calculations.[18,26]

The cusp condition on the wavefunction, necessary for cancelling the singularity in $V(\underset{\sim}{R})$ when two particles approach the same position, fixes the value of a in the Jastrow factor[23]; the variable b, and the parameters in the Slater determinant, may be adjusted by a variational QMC procedure, to achieve the lowest energy. More sophisticated Padé approximants for U_{ij} may also be used. However, since the Jastrow factor is always positive, any such adjustments will not change the nodes of Ψ_T. Thus, adjustment of the Jastrow factor alone can only affect the variance of the energies obtained by the method presented here.

In addition to trial functions constructed from Slater-type orbitals, we have also tried a significantly different trial function which, for some molecules, offers the possibility of higher accuracy and greater computational speed. The molecular orbitals, instead of being of the LCAO form--where each atomic orbital may have a different center--are instead each localized about one center. Different molecular orbitals may, however, have different centers. These molecular orbitals have the form

$$\psi_k(\vec{r}) = \exp\left(\frac{-(\vec{r}-\vec{c}_k)^2}{w_k^2 + v_k|\vec{r}-\vec{c}_k|}\right) , \tag{9a}$$

where $\vec{c}_k$, w_k, and v_k are variational parameters. A Slater determinant is then formed from these orbitals. The full trial function, in addition to having the product of the Slater determinant with an electron-electron correlation factor, as in Eq.(8), also has an electron-nuclear Jastrow factor to satisfy the cusp condition when $r_{i\alpha}\to 0$. Explicitly, this trial function has the form

$$\Psi_T(\underset{\sim}{R}) = \det|D^{\alpha}_{k\ell}|\,\det|D^{\beta}_{k\ell}|\,\exp\Big[\sum_{i>j}\frac{a_{ij}r_{ij}}{1+br_{ij}} - \sum_{i,\alpha}\frac{Z_\alpha a_{i\alpha}r_{i\alpha}}{1+br_{i\alpha}}\Big] , \tag{9b}$$

where $a_{ij} = \begin{cases} e^2/8D & \text{if ij are like spins,} \\ e^2/4D & \text{if ij are unlike spins,} \end{cases}$

and $a_{i\alpha} = e^2/2D$. The parameter b in the three cases may be written in terms of a single parameter β as $b = (a/\beta)^{1/2}$, where a is either a_{ij} or $a_{i\alpha}$. The values of β, $\vec{c}_k$, w_k, and v_k which we have used are given in Table AIV in the Appendix.

A trial function of the form of Eq.(9) has built into it a number of desirable properties:

a) It contains the correct cusp conditions for both like and unlike electron spins. It is not generally realized that the cusp conditions for these two cases are different because of the presence of the determinant.

b) It also contains the correct cusp behavior as the electron-nuclear separation $r_{i\alpha}$ becomes small. Since $\psi_k(\vec{r})$ is quadratic in $\vec{r}$ at the origin, the determinant does not affect the cusp behavior. Therefore, if the nuclear positions change by a small amount it does not become immediately necessary to re-optimize the determinant, since the trial function will change in such a way that the electron-nuclear cusp is preserved.

c) For two separated, closed-shell molecules (A and B), the wavefunction will naturally have the form

$$\Psi_A(\underset{\sim}{R}_A)\ \Psi_B(\underset{\sim}{R}_B)\ \exp[-\beta\Delta V(\underset{\sim}{R})] \quad , \tag{10}$$

where $\underset{\sim}{R}_A$ and $\underset{\sim}{R}_B$ are the electronic coordinates of molecules A and B, $\Delta V(\underset{\sim}{R})$ is the potential energy of interaction of molecule A with B, and $\beta=a/b^2$. This form is reminiscent of that obtained in the Hylleraas-Hasse variational treatment of intermolecular forces.[27]

d) If the Jastrow factor is dropped and v_k is set to zero, these orbitals become floating spherical Gaussians (FSGO). However, the complete trial function is superior to one constructed from FSGO's, primarily because of the correlation factors.

e) This wavefunction can be applied more efficiently in QMC approaches to large systems, e.g., containing up to 250 electrons[28]. In such large systems, trial functions made with "localized" orbitals, such as those given by Eq. (9), will be computationally more efficient

than ones formed from "delocalized" orbitals, such as LCAO, since in the former case sparse matrix algorithms may be used to calculate the Slater determinant and its inverse matrix. Matrix manipulation for large systems is the most time-consuming step.

f) Finally, the parameters in this trial function are much simpler to interpret physically. Thus, one should be able to understand how they will change as a molecule is distorted. At large separations, β can be related to the polarizability, and v*, the maximum value of v_k, is $1/\sqrt{2I}$, where I is the first ionization potential.

Fermi Statistics

The diffusion equation formulation requires that the density of diffusers be non-negative. In Eq.(1) this required that $\Phi(\underset{\sim}{R},t)$, and hence $\phi_0(\underset{\sim}{R})$, either had no nodes--leading to a Bose ground state--or that we could treat the positive and negative regions of ϕ_0 separately. In Eq. (6), on the other hand, it is $f(\underset{\sim}{R},t) = \Psi_T(\underset{\sim}{R})\,\Phi(\underset{\sim}{R},t)$ that must not change sign. Thus, if Ψ_T were to have the exact nodes of the ground state, one could treat the Fermion system immediately and exactly, since f would never change sign. Unfortunately, very little is known about the exact location of the nodes in molecular systems.[29] From symmetry one can only determine points on the nodal surface. Nevertheless, an exact simulation of Eq. (6) by QMC methods is possible.[15-17] However, in this paper we deal with Fermi systems by employing the simpler fixed-node approximation: we solve Eq. (6) in each nodally bounded volume of Ψ_T separately, with the boundary condition that $\Phi(\underset{\sim}{R},t) = 0$ at, and only

at, the boundary of this volume.[30] In terms of the description of diffusion with branching, this means that when a configuration diffuses to the boundary, it is killed. The diffusion process thus leads to the lowest energy solution with no internal nodes (i.e., the Bose ground state in such a bounded system) for each nodally bounded region of Ψ_T. The approximate ground state is then taken as the antisymmetrized $\Phi(\underset{\sim}{R})$, obtained by permutation "reflections" of the asymptotic $\Phi(\underset{\sim}{R})$ in the nodal region having the lowest energy. We show now that the energy expectation value, calculated with the approximate density $f(\underset{\sim}{R}) = \Psi_T(\underset{\sim}{R})\,\Phi\,(\underset{\sim}{R})$ is an upper bound to the true ground-state energy.

Variational principle for the fixed-node process[31]

In this section we discuss the relationship between the fixed-node energy and the Fermion ground-state energy E_0. Let the trial function $\Psi_T(\underset{\sim}{R},s)$ be antisymmetric in both the spatial variables $\vec{r}$ and the spin variables s. Further, let v_α be the connected volumes in 3N-dimensional space bounded by the nodes of Ψ_T, obtained for some arrangement of spins s_0.[32] In each of these volumes there is a unique ground-state eigenfunction $\phi_\alpha(\underset{\sim}{R},s)$ with eigenvalue ε_α, which satisfies the equations

$$\begin{cases} H\phi_\alpha(\underset{\sim}{R},s) = \varepsilon_\alpha\phi_\alpha(\underset{\sim}{R},s) & \underset{\sim}{R} \in v_\alpha \ , \\ \phi_\alpha(\underset{\sim}{R},s)\ \ \Psi_T(\underset{\sim}{R},s) > 0 \end{cases} \tag{11a}$$

and

$$\phi_\alpha(\underset{\sim}{R},s) = 0 \qquad \underset{\sim}{R} \notin v_\alpha \ . \tag{11b}$$

The fixed-node procedure solves this problem exactly in each volume element [i.e., $\Phi(\underset{\sim}{R},t) \xrightarrow{t\to\infty} \phi_\alpha(\underset{\sim}{R},s_0)$]. However, each of the eigenvalues ϵ_α is an upper bound to the Fermion energy E_0, since for each α one can define an antisymmetric function

$$\hat{\phi}_\alpha(\underset{\sim}{R},s) = \sum_P (-)^P \phi_\alpha(P\underset{\sim}{R},Ps) \quad , \tag{12}$$

whose variational energy is

$$\frac{\int d\underset{\sim}{R}\ \hat{\phi}_\alpha^*\, H\hat{\phi}_\alpha}{\int d\underset{\sim}{R}\ \hat{\phi}_\alpha^*\,\hat{\phi}_\alpha} = \epsilon_\alpha \geq E_0 \quad . \tag{13}$$

Here P represents a permutation of the electrons. The approximation to ϕ_0 in Eq. (12) has the nodal structure of Ψ_T. The probability density $f_\alpha(\underset{\sim}{R}) = \hat{\phi}_\alpha(\underset{\sim}{R})\,\Psi_T(\underset{\sim}{R})$, is thus always non-negative, as desired.

There are two important points to note in this proof. First, $\hat{\phi}_\alpha$ is not identically zero. This can be seen because any permutation which maps $\underset{\sim}{R}\epsilon v_\alpha$ back into v_α must be even (otherwise v_α would contain positive and negative regions of Ψ_T), and hence the terms in the sum in Eq. (12) must be all of the same sign. Thus, $\hat{\phi}_\alpha$ cannot be zero inside v_α. Second, $\hat{\phi}_\alpha$ may have a discontinuous gradient at the node, and consequently the Laplacian in the Hamiltonian of Eq. (13) could contain a delta function there. However, such a delta function would not contribute to the integral since $\hat{\phi}_\alpha$ is zero at the boundary.

In the fixed-node process, we attempt to populate as many of the volumes v_α as possible. By Eq. (3), the trial energy necessary to hold the population of walks asymptotically constant will be given by $\epsilon_m \equiv \min_\alpha (\epsilon_\alpha)$, where α ranges over those volumes initially populated.

This energy is clearly also an upper bound to E_0. Thus, at large times, the average value of the local energy will equal ϵ_m:

$$\langle E_L \rangle \equiv \frac{\int f_\infty(\underset{\sim}{R})[H\Psi_T/\Psi_T]\, d\underset{\sim}{R}}{\int f_\infty(\underset{\sim}{R})\, d\underset{\sim}{R}} = \epsilon_m \quad , \tag{14}$$

where

$$f_\infty(\underset{\sim}{R}) \equiv f(\underset{\sim}{R}, t \to \infty) = \sum_\alpha c_\alpha \phi_\alpha(\underset{\sim}{R}, s)\, \Psi_T(\underset{\sim}{R}, s) \tag{15}$$

and c_α is a non-negative constant proportional to the initial population of v_α. The best upper bound (for a given Ψ_T) will be obtained if all of the volumes v_α have been populated. For the true nodes, all these volumes will have the same energy, and it will be irrelevant how we choose the initial ensemble. In fact, no dependence on the initial ensemble has been discovered for the molecular systems treated here.

3. Monte Carlo Solution

To obtain the asymptotic distribution $f_\infty(\underset{\sim}{R})$, which is a solution to Eq. (6), we begin with an arbitrary initial distribution $f(\underset{\sim}{R},0)$--for example, one randomly generated, or one given by $|\Psi_T(\underset{\sim}{R})|^2$ from an earlier variational QMC simulation. The time evolution of $f(\underset{\sim}{R},t)$ is given by

$$f(\underset{\sim}{R}', t+\tau) = \int d\underset{\sim}{R}\, f(\underset{\sim}{R},t)\, G(\underset{\sim}{R} \to \underset{\sim}{R}', \tau) \quad , \tag{16}$$

where the Green's function[33] $G(\underset{\sim}{R} \to \underset{\sim}{R}', \tau)$ is a transition probability for moving the set of coordinates from $\underset{\sim}{R}$ to $\underset{\sim}{R}'$ in time τ. Thus G is a solution to the same differential equation, Eq. (6), as f, but with the boundary condition $G(\underset{\sim}{R} \to \underset{\sim}{R}', 0) = \delta(\underset{\sim}{R}' - \underset{\sim}{R})$.

For short times τ we may assume that both the local energy and the quantum force are constant, independent of $\underset{\sim}{R}$. Then an approximate Green's function solution to equation (6) is

$$G(\underset{\sim}{R} \to \underset{\sim}{R}', \tau) = (4\pi D\tau)^{-3N/2} \, e^{-\tau\{[E_L(\underset{\sim}{R})+E_L(\underset{\sim}{R}')]/2-E_T\}} \, e^{-[\underset{\sim}{R}'-\underset{\sim}{R}-D\tau F_Q(\underset{\sim}{R})]^2/4D\tau} \quad (17)$$

This Gaussian probability distribution has a mean which drifts with a velocity DF_Q, and a width which spreads with time as $\sqrt{\tau}$. This distribution is used to move the electrons. The exponential prefactor of the Gaussian grows or diminishes depending on the relative magnitude of E_L and E_T. This change in normalization results from the branching term in Eq. (6). It is carried out by creating or destroying entire electronic configurations with a probability such that the average number of configurations in the next time step is $\exp(-\tau\{[E_L(\underset{\sim}{R}) + E_L(\underset{\sim}{R}')]/2 - E_T\})$.

Thus to obtain a Monte Carlo solution to Eq. (6)--that is, to find the asymptotic distribution $f(\underset{\sim}{R},t)$--one needs only to apply Eqs. (16) and (17) repeatedly for small τ, until t is sufficiently large. Once the equilibrium distribution for f is obtained, one may take any desired average over the configurations. The explicit algorithm follows:

(0) Before beginning the computation, one must choose $\Psi_T(\underset{\sim}{R})$. As discussed earlier, a trial function which is compact and concise, yet relatively accurate, is ideal [cf. Eqs. (8) and (9)]. Optimization of the parameters in Ψ_T may have been carried out in previous <u>variational</u> QMC runs, or in a self-consistent-field or a Hartree-Fock calculation.

The initial probability density $f(\underset{\sim}{R},0)$ must also be chosen. To increase the speed with which $f(\underset{\sim}{R},t)$ approaches its asymptotic, steady-state solution, we usually choose $f(\underset{\sim}{R},0) = |\Psi_T(\underset{\sim}{R})|^2$. However, any initial choice of f is acceptable, as long as the overlap of $\Phi(\underset{\sim}{R},0) \equiv f(\underset{\sim}{R},0)/\Psi_T(\underset{\sim}{R})$ [cf. Eq. (5)] with the ground-state ϕ_0 is non-zero.

In order to compute Ψ_T, F_Q, and E_T efficiently, the inverse of the Slater matrix is initially computed.[12] Later in the algorithm, as electrons are moved, this inverse is updated. The first and second derivatives of Ψ_T, needed to evaluate F_Q and E_T, can then be obtained readily as scalar products.

(1) Initialize a set of N_c configurations of coordinates $\underset{\sim}{R}$, (the "list"), with the electrons in each configuration distributed with a probability density $f(\underset{\sim}{R},0)$. Typically $N_c \approx 100\text{-}500$. If $f(\underset{\sim}{R},0) = |\Psi_T(\underset{\sim}{R})|^2$, the initial list may be generated by choosing configurations produced in a variational QMC simulation.

(2) Pick a configuration from the list--assume the next one is the $m^{\underline{th}}$ one. The electrons in this configuration, will be moved, one at a time, by letting them diffuse independently for a time τ, according to the Gaussian part of the transition probability $G(\underset{\sim}{R} \rightarrow \underset{\sim}{R}', \tau)$. If the current electron is the $j^{\underline{th}}$ electron in configuration m, it is moved to [cf. Eq. (17)]

$$\vec{r}_j'(m) = \vec{r}_j(m) + D\tau\vec{F}_Q(\vec{r}_j(m)) + \vec{\chi} \quad , \tag{18}$$

where $\vec{r}_j(m)$, the three-dimensional coordinate of the electron being moved, is the $j^{\underline{th}}$ component of $\underset{\sim}{R}$; $\vec{\chi}$ is a three-dimensional

Gaussian random variable with a mean of zero and a variance of $2D\tau$; and $D = \hbar^2/2m_e$ is the diffusion constant. The time step τ is chosen empirically by determining when decreasing τ no longer affects the results within resolution of the statistical uncertainty. To correct for the effect of the finite time step systematically, one should extrapolate the results obtained for a sequence of decreasing values of τ. We have performed detailed studies of the time-step dependence of the energy only for the H_2 molecule. In Fig. (1) we show the results obtained. The two curves are $E = E_o + A_i \tau^{i/2}$, for i=1, 2. Here E_o is the exact H_2 energy, and the A_i are determined by least squares fits to the data. The coefficients A_i will vanish as Ψ_T approaches the exact ground-state wavefunction. These forms for the τ dependence of E were chosen because the corrections to Eq. (17) go as powers of $\tau^{1/2}$. In both cases, a χ^2 test shows that the Monte Carlo results are consistent with an extrapolation of E to E_o. In practice, we have found that choosing τ small enough that the rejection ratio, in step (3) below, is less than 1% makes the systematic error due to the time step smaller than our statistical errors (cf. Fig. 1). Typically, we chose $\tau \approx 0.003$ h^{-1} for the water molecule, and $\tau \approx 0.005$-0.015 h^{-1} for the other molecules. The time-step error may also be eliminated by use of Green's function Monte Carlo method.[13]

If electron j crosses a node (e.g. as a result of the finite τ in the simulation) the entire configuration is eliminated from the list. (This forces $\phi_\alpha(\underset{\sim}{R})$ to vanish on the other side of the node, satisfying the boundary condition (11b); thus the Schroedinger

equation is solved in each nodal volume separately.) If this is the case, go back to the beginning of step (2) and pick the next configuration. Otherwise, go on to Step 3.

(3) After electron j is "moved" according to Eq. (18), accept the move with a probability

$$A(\underset{\sim}{R} \to \underset{\sim}{R}', \tau) \equiv \min\,(1,\, W(\underset{\sim}{R}', \underset{\sim}{R})), \tag{19}$$

where

$$W(\underset{\sim}{R}', \underset{\sim}{R}) \equiv \frac{\Psi_T^2(\underset{\sim}{R}')G(\underset{\sim}{R}' \to \underset{\sim}{R}, \tau)}{\Psi_T^2(\underset{\sim}{R})\; G(\underset{\sim}{R} \to \underset{\sim}{R}', \tau)} \quad , \tag{20}$$

and G is given by Eq. (17). This step would be unnecessary if G were the exact Green's function, since $W(\underset{\sim}{R}', \underset{\sim}{R})$ would be unity. This can be seen by writing the exact Green's function in its eigenvalue expansion. The approximate Green's function of Eq. (17) becomes exact as $\tau \to 0$, except perhaps at $r_{i\alpha} = 0$. For finite τ, however, when G is only approximate, the acceptance/rejection procedure of this step nevertheless guarantees detailed balance in our operational Monte Carlo transition probability $G(\underset{\sim}{R} \to \underset{\sim}{R}', \tau)\; A(\underset{\sim}{R} \to \underset{\sim}{R}', \tau)$: if this expression replaces G in Eq. (20), W is again unity.[34] This step guarantees that as $\Psi_T \to \phi_0$, the correct distribution, $|\phi_0|^2$, will be sampled for any τ, no matter how large.

(4) After all N electrons in the current configuration, m, have been moved once, advance the time associated with this new configuration $\underset{\sim}{R}'$ by τ. Calculate $E_L(\underset{\sim}{R}')$ and other quantities of interest.

(5) Calculate the multiplicity M_m (the branching probability) for configuration m, from the exponential prefactor of the Gaussian in Eq. (17). Explicitly,

$$M_m = \exp \{ - \tau_a([(E_L(\underset{\sim}{R}) + E_L(\underset{\sim}{R}'))]/2 - E_T)\} \quad . \tag{21}$$

Note that the actual time that the electrons have been drifting and diffusing is slightly shorter than τ due to rejections in step (3). The mean-squared distance the electrons would diffuse in time τ is

$$\langle r^2 \rangle = 6D\tau \quad . \tag{22}$$

However, due to rejections, they only diffuse by[24]

$$\langle r^2_{\text{accepted}} \rangle = 6D\tau_a \quad . \tag{23}$$

This equation defines τ_a used in Eq. (21). Combining Eqs. (22) and (23),

$$\tau_a = \frac{\langle r^2_{\text{accepted}} \rangle}{\langle r^2_{\text{total}} \rangle} \quad . \tag{24}$$

After computing (21), place M_m copies of the new configuration $\underset{\sim}{R}'^{(m)}$, back into the list of molecular configurations. If M_m is not an integer, treat the remainder as a probability: choose a random number ξ between 0 and 1; if the remainder of M_m is greater than ξ, round M_m up. This rounding can be achieved simply by using the integer

$$\bar{M}_m = \text{int } (M_m + \xi) \tag{25}$$

instead of M_m. It is easy to see that $\langle \bar{M}_m \rangle = M_m$, and hence the density of random walks at the point $\underset{\sim}{R}$ will be given by Eq. (17). Note that if $\bar{M}_m = 1$, the total number of configurations in the list is unchanged, while if $\bar{M}_m = 0$ the $m^{\underline{th}}$ configuration is not returned to the list.

(6) If $\bar{M}_m \neq 0$, weight $E_L(\underset{\sim}{R}')$ and other quantities of interest by M_m.

(7) Repeat (2) through (6) until all extant configurations have reached a target time t_{target}. We generally choose t_{target} to be on the order of a few atomic units (inverse hartrees).

(8) Calculate the weighted mean $\langle E_L(\underset{\sim}{R}') \rangle$ as an estimator of the "ground-state" energy ϵ_m [cf. Eq.(14)]. This average is the expectation value of $E_L = H\Psi_T/\Psi_T$ sampled from the distribution $f_\infty(\underset{\sim}{R})$. Also calculate other desired averages.

(9) Use the cummulative estimate of $\langle E_L \rangle$ to update the trial energy E_T. For better convergence, we <u>mix</u> this estimate with the old E_T, so that $(E_T)_{new} = [(E_T)_{old} + \langle E_L \rangle]/2$. If $E_T = \epsilon_m$, asymptotically the number of configurations in the ensemble remains constant [cf. Eq. (4) noting Ref. 30], and thus $\langle M_m \rangle = 1$.

(10) "Renormalize" the number of configurations to the initial list size, N_c, by either randomly elimininating or copying from the existing list of configurations. Reset the time counter to zero in all configurations.

This step completes one "block". The size of a block is determined by N_c, and the target time t_{target}. On the one hand, one should choose both of these as large as possible to avoid introducing a bias in the renormalization procedure, and to minimize the statistical correlation between blocks. Yet, on the other hand, one wants these quantities small enough that in the available computing time there are a large number of blocks to average over, and with which to compute variances.

(11) Repeat steps (2) through (10) until there is no systematic trend reflected in the single block and grouped averages of $\langle E_L \rangle$ and other quantities. At this point steady-state has been reached, and all traces of the initial conditions are gone.

(12) Reset all cumulative averages. Repeat (2) through (10) until the variance in $\langle E_L \rangle$ has reached the desired level.

4. Results and Discussion

We have used the QMC algorithm described in the previous section to calculate the ground-state energies of H_2, LiH, Li_2, and H_2O. The results presented here are obtained for three different types of importance function Ψ_T: A Jastrow electron-electron pair-correlation function multiplied by a Slater determinant of molecular orbitals constructed from (I) a minimal basis set of Slater-type atomic orbitals, (II) a somewhat enhanced basis set and/or an optimized version of (I), and (III) localized Gaussian-like orbitals. These importance functions were described in more detail in Sec. 2 [cf. Eqs. (8) and (9)]. Importance functions of type III contain a spin-dependent

electron-electron Jastrow factor, and are multiplied by an additional electron-nuclear Jastrow factor.

The values of the parameters used for Ψ_T are given in the Appendix. The total energies corresponding to these choices are presented in Table I. For each molecule, we compare the fixed-node QMC ground-state energy with the Hartree-Fock energy, the best CI calculation, and the "exact," clamped (i.e. fixed-nuclei or Born-Oppenheimer), non-relativistic result. All our numbers are presented in atomic units,[45] in which $e^2 = \hbar^2/m_e = 1$. In addition, to gain an appreciation for the quality of each Ψ_T by itself (without the fixed-node procedure) we give the results obtained with the same Ψ_T's used as variational trial wavefunctions rather than as guiding functions. For this comparison, the variational energies have been calculated with the same set of parameters as used in the fixed-node QMC calculation. For most of these wavefunctions, the parameters have only been partially optimized, both to save computer time, and to demonstrate the strength of the fixed-node procedure.

The difference between the variational QMC results and the exact results is due, of course, to the inexactness of the trial wavefunction Ψ_T. The statistical error bars have nothing to do with this difference, but only measure how accurately the variational energy of Ψ_T has been obtained. This statistical uncertainty (standard deviation) is inversely proportional to the square root of the number of independent averages, and thus may be reduced by making more or longer computer runs. The difference between the fixed-node QMC results and the exact results is, on the other hand, due solely to the

approximation made in forcing the ground state to have the approximate nodes of Ψ_T. Again, the statistical error bars measure only the uncertainty in the measured quantity--this time the ground-state energy of the functions ϕ_α (cf. Sec.2). Numerically, the fixed-node approximation is quite good. For a given Ψ_T, a fixed-node QMC calculation is much superior to a variational QMC calculation, gaining approximately 90% of the energy missing in the variational treatment. In fact, the fixed-node procedure may be thought of as a stochastic method of "correcting" the variational wavefunction Ψ_T through the branching process described by Eq. (6). Thus, the variational result obtained with Ψ_T is only the starting point for the fixed-node calculation.

For H_2, where the ground state has no nodes--since there is only one electron in each spin state--the fixed-node QMC results are exact, except for the time step error (cf. Fig. 1) which can be eliminated. The only remaining uncertainty is statistical. On the other hand, the same trial function (e.g. Ψ_I) used variationally rather than as a guiding function, although better than Hartree-Fock, gives only 50% of the correlation energy. For LiH, the three trial functions, used as variational wavefunctions, range from considerably worse than Hartree-Fock to considerably better (achieving approximately 60% of the correlation energy). Note, however, that in each case--regardless of the quality of the starting Ψ_T--the fixed-node calculation brings the result almost 90% of the way from the variational result to the exact result. Similarly, for Li_2, Ψ_I starts off worse than Hartree-Fock, while Ψ_{II} and Ψ_{III} already achieve variationally 22%

and 62% of the correlation energy respectively. Just as for LiH, for all three trial functions, the fixed-node QMC achieves an additional 90% of the energy difference between the variational starting point and the exact result. For H_2O, however, the quality of neither of the two trial functions used is especially good. Variationally, Ψ_I is worse than Hartree-Fock, and Ψ_{II} is only about 17% better than Hartree-Fock. Nevertheless, using the fixed-node QMC with these trial functions, almost 80% of the energy difference between the variational and exact energies is obtained. We have, however, been unable to effectively optimize Ψ_{III} for water.

Considering the simplicity of our trial functions, it is perhaps remarkable that variationally we obtain with Ψ_{III} about 2/3 of the correlation energy for H_2, LiH, and Li_2. Nevertheless, applying the fixed-node procedure with Ψ_{III} is better still, and yields between 95% and 100% of the correlation energy, although presently with a statistical error of about 2% of the correlation energy. For these molecules, the results obtained with Ψ_{III} appear to be somewhat superior to these obtained with the LCAO-type functions Ψ_I and Ψ_{II}. The chief drawbacks appear to be the difficulty in optimizing the parameters in Ψ_{III} for larger molecules, such as H_2O, and in using spherical orbitals to represent directional bonds.

We note that for all the molecules treated here, the fixed-node procedure appears to obtain a fairly constant fraction of the energy that the variational wavefunction misses. Thus even the simplest trial function, Ψ_I, is able to achieve from 50-90% of the

correlation energy. Furthermore, the fixed-node procedure with our best choice of Ψ_T does at least as well as CI presently does. As is clear from Table I, the results improve by choosing better Ψ_T's, since in general these will more accurately represent the nodes of the true wavefunction. Thus there is potential for still higher absolute accuracy, in addition to the reduction in statistical error which may be achieved by running longer. Finally, by release of the fixed-node constraint it should be possible to achieve 100% of the correlation energy, even for the simplest trial functions. This will be the subject of a future paper.

Thus far we have used only a few hours of computer time per molecule on a CDC 7600. There is, however, no fixed amount of computer time that is necessary; more or less can be used. The effect of a longer run is to increase the precision with which the computed averages, such as the energy, are known. The error bars obtained are the standard deviations of the mean. By running as long as we have, we have reduced these error bars sufficiently that for each Ψ_T the effect of the fixed-node approximation becomes visible. The approximation manifests itself when the statistical error bars do not encompass the exact answer. For H_2, however, where there is no approximation, reduction of the statistical error will give the exact answer more and more precisely.

It is also of interest to inquire whether this approach is applicable to nuclear separations away from the equilibrium configuration. The QMC procedure applies equally well in this case, and in Table II we present some results for the the ground-state

energy of Li_2 at a few different nuclear separations. Note that we use the same importance function--with the same parameters--for all nuclear separations. Although this choice is not optimal, pointwise agreement with the exact results is nevertheless quite good. However, the estimates for the energy are statistically independent and thus have separate error bars. This is troublesome for an accurate calculation of potential energy surfaces. However, work is in progress on a differential QMC scheme, which would eliminate this problem and give more accurate relative energies than is possible from separate calculations of the absolute energies.

It is also worth noting that the Born-Oppenheimer approximation, used throughout, can also be relaxed. This is achieved by allowing the nuclei, as well as the electrons, to diffuse. The diffusion constant for each nucleus is then $\hbar^2/2M_{nuc}$, where M_{nuc} is the mass of the nucleus. Thus the nuclei diffuse considerably slower than the electrons. This, however, may make the calculation much longer.

5. Summary and Conclusion

We have presented the theory and an algorithm for obtaining a stochastic solution to the Schroedinger equation by treating it as a diffusion equation, and applied it to H_2, LiH, Li_2, and H_2O. The procedure described here made use of the fixed-node approximation, in which the positive and negative regions of the importance function Ψ_T are treated separately, and electrons from one region are prevented from diffusing across the nodes of Ψ_T, into another region. Using relatively simple forms for Ψ_T, and only modest computational effort,

we were able to obtain at least as much of the correlation energy as the CI method does for the molecules treated. A further increase in accuracy seems also readily achievable. Increasing the length of the runs reduces the statistical uncertainty (standard deviation of the mean) associated with the calculated averages. Absolute error can also be reduced, by choosing importance functions which better approximate the nodes of the true wavefunction, or by relaxing the fixed-node constraint.[15,16] In this paper it was also demonstrated that the approximate energies calculated within the fixed-node scheme are upper bounds to the true ground-state energy.

Given the relative ease of computation, and the potential for high accuracy, this method holds exceptional promise for wide application in quantum chemistry. Future work should include (1) use of more accurate compact trial wavefunctions, (2) development of adaptive (self-improving) Monte Carlo schemes, (3) work on stable methods for eliminating the fixed-node constraint, (4) use of differential methods to obtain potential surfaces, (5) excited-state calculations, (6) evaluation of expectation values other than the energy, (7) elimination of the Born-Oppenheimer approximation, and (8) relativistic calculations.

Note Added: After this paper was completed we received a preprint from Moskowitz et. al.[47] of a paper in which a similar approach to ours is used to study the molecule LiH. Their results suggest that, due to an approximation in their Green's function, the electron-nuclear singularity can cause the fixed-node energy to be less than the true energy, even at very small time steps. When they remove this singularity, their energy (like ours) is an upper bound to the true energy. Their best bounded value for the total ground-state energy of LiH, obtained using a generalized valence bond trial function, is (within statistical errors) the same as our energy for LiH obtained with trial function Ψ_{III}.

6. Acknowledgment

We would like to thank Dr. Michel Dupuis for making helpful comments on the manuscript.

Appendix

We present here the details of the trial wavefunctions Ψ_T used in the present calculations. In Table AI we give the equilibrium nuclear geometry used for each molecule. The nuclear coordinates are held fixed in this work (Born-Oppenheimer approximation). The electronic coordinates are generated by using three different trial wavefunctions as importance functions (cf. Sect. 2).

Wavefunctions of type Ψ_I and Ψ_{II} are described by Eq. (8). For these wavefunctions, the linear combination of atomic orbitals used for each molecular orbital, together with the orbital exponents ζ and the coefficients a and b in the Jastrow factor, are given in Tables AII and AIII. The molecular orbitals for H_2O, not listed in these Tables, are from Aung et. al.[46]

Wavefunctions of type Ψ_{III} are described by Eq. (9). Each molecular orbital is localized on <u>one</u> center, and there is no linear combination formed. The parameters β, w_k, v_k, and $\vec{c}_k$ are given in Table AIV.

REFERENCES

1. H.F. Schaefer, Electronic Structure of Atoms and Molecules (Addison-Wesley, Reading, MA) 1972.

2. R.J. Bartlett, Ann. Rev. Phys. Chem. 32, 359 (1981) and references therein.

3. The correlation energy is defined as the difference between the Hartree-Fock energy and the exact, non-relativistic, Born-Oppenheimer energy. It is so named since it accounts for the electronic correlation missing in the Hartree-Fock model.

4. B. J. Rosenberg and I. Shavitt, J. Chem. Phys. 63, 2162 (1975); W. Meyer, Int. J. Q. Chem., Symp. No. 5, 341 (1971).

5. Nevertheless, very good predictions for a variety of molecular properties have been achieved (see e.g. Ref. 1). For example, relatively accurate values for bond strengths are obtained with the CI method by subtracting the energy obtained for the molecule from the sum of the energies obtained for the two fragments which result when the bond is broken. Significantly, however, the large mathematical uncertainty inherent in this procedure--e.g., due to the nature of the convergence, and hence to the mathematically unknown degree of cancellation in the separate errors--is generally not systematically improvable.

6. N. Metropolis and S.M. Ulam, J. Amer. Statist. Assoc. 44, 247 (1949).

7. M.H. Kalos, Phys. Rev. 128, 1791 (1962); J. Comp. Phys. 2, 257 (1967).

8. W. L. McMillan, Phys. Rev. A138, 442 (1965).

9. R.C. Grimm and R.G. Storer, J. Comp. Phys. 7, 134 (1971).

10. M.H. Kalos, D. Levesque and L. Verlet, Phys. Rev. A 9, 2178 (1974).

11. J.B. Anderson, J. Chem. Phys. 63, 1499 (1975); ibid 65, 4121 (1976).

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

13. D. M. Ceperley and M. H. Kalos, in Monte Carlo Methods in Statistical Physics, K. Binder ed. (Springer Verlag, Berlin, 1979), pp. 145-97.

14. J.B. Anderson, J. Chem. Phys. 73, 3897 (1980); F. Mentch and J. B. Anderson, J. Chem. Phys. 74, 6307 (1981).

15. D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. 45, 566 (1980).

16. D.M. Ceperley, "The Stochastic Solution of the Many-Body Schroedinger Equation for Fermions" in Recent Progress in Many-Body Theories, J.G. Zabolitzky, M. de Llano, M. Fortes, and J.W. Clark eds., (Springer-Verlag, 1981) pp. 262-9.

17. B.J. Alder, D.M. Ceperley and P.J. Reynolds, J. Phys. Chem. (to be published).

18. J. W. Moskowitz and M. H. Kalos, Int. J. Q. Chem. 20, 1107 (1981) and J.W. Moskowitz, K.E. Schmidt, M.A. Lee, and M.H. Kalos, J. Chem. Phys. 76, 1064 (1982).

19. J.M. Hammersley and D.C. Handscomb, Monte Carlo Methods, (Chapman and Hall, 1964) pp. 57-9.

20. The computational complexity of our code goes as $AN^3 + BN^2$, where the cubic term comes from inverting the Slater matrix order N times (once for each electron moved), and the quadratic term comes from computing pairwise interactions. Generally $A << B$. For N sufficiently large, the N^3 term would ultimately dominate, although the algorithm is effectively N^2 in the range of N we have treated. In large systems, where N^3 would begin to dominate, suitable modifications can be made, by use of sparse matrix algorithms, to eliminate this term. The computational complexity then goes as $B'N^2$. However, $B' > B$, making this modification costly at small N.

21. R. Jastrow, Phys. Rev. 98, 1479 (1955); R. B. Dingle, Phil. Mag. 40, 573 (1949).

22. In fact, the constancy of the local energy can be used as a quantitative measure of the accuracy of any proposed Ψ_T.

23. The cusp condition is a requirement on a wavefunction Ψ that the leading singularity in $V(\underset{\sim}{R})$, when two particles come together, cancels when evaluating the energy $H\Psi/\Psi$. This leads to the conditions that, for two electrons

$$\frac{1}{\Psi}\frac{\partial \Psi}{\partial r_{ij}}\bigg|_{r_{ij}=0} = \begin{cases} e^2/8D & \text{i,j like spins,} \\ e^2/4D & \text{i,j unlike spins,} \end{cases} \quad \text{and} \quad \frac{1}{\Psi}\frac{\partial \Psi}{\partial r_{j\alpha}}\bigg|_{r_{j\alpha}=0} = -Z_\alpha e^2/2D,$$

for an electron and a nucleus. Thus, e.g., for opposite spins at small r_{ij}, $\Psi \propto \exp(e^2 r_{ij}/4D)$, implying that the coefficient a in U_{ij} equals $e^2/4D$.

24. See e.g. D. M. Ceperley, M. H. Kalos, and J. L. Lebowitz, Macromolecules 14, 1472 (1981).

25. If the trial function contains only a single Slater determinant, the full Slater matrix can always be block diagonalized into spin up and spin down submatrices by relabeling the coordinates.

26. Good results have also been obtained with other forms for the pair-correlation function. See e.g. C.C.J. Roothaan and A.W. Weiss, Rev. Mod. Phys. 32, 194 (1960); W. Kolos and C.C.J. Roothaan, Rev. Mod. Phys. 32, 205 (1960); and W.A. Lester Jr. and M. Krauss J. Chem. Phys 41, 1407 (1964); ibid 44, 207 (1966).

27. E.A. Hylleraas, Z. Phys. 65, 209 (1930); H. R. Hasse, Proc. Camb. Phil. Soc. 26, 542 (1930); J. O. Hischfelder, C. F. Curtiss and R. B. Bird, Molecular Theory of Gases and Liquids, (Wiley, 1954), pp. 942-6.

28. D. M. Ceperley and B. J. Alder, Physica 108B, 875 (1981).

29. R.J. White and F.H. Stillinger, Phys. Rev. A 3, 1521 (1971); D.J. Klein and H.M. Pickett, J. Chem. Phys. 64, 4811 (1976).

30. Because of the boundary condition imposed on Φ by this approximation, the expansion of Eq. (3) must be in terms of eigenfunctions of H within the separate volume elements. Thus, the spectrum of eigenvalues E_i will not be exactly that of the true Fermion problem unless the nodes are correct. In particular, E_0 of Eqs.(4,7) is replaced by ε_α in volume v_α.

31. The proof given here is an expanded version of the proof given in Ref. 16.

32. For any given total spin, the particular spin configuration s_0 is unimportant since the electrons can be simply relabeled.

33. The use of a Green's function here is not to be confused with the "Green's function Monte Carlo" (GFMC) method of Kalos described in Ref. 13.

34. W. K. Hastings, Biometrika 57, 97 (1970).

35. W. Kolos and C.C.J. Roothaan, Rev. Mod. Phys. 32, 219 (1960).

36. P.E. Cade and W. M. Huo, J. Chem. Phys. 47, 614 (1967).

37. G. Das and A.C. Wahl, J. Chem. Phys. 44, 87 (1966).

38. S. Hagstrum and H. Shull, Rev. Mod. Phys. 35, 624 (1963).

39. W. Kolos and L. Wolniewicz, J. Chem. Phys. 41, 3663 (1964); ibid 43, 2429 (1965); ibid 49, 404 (1968).

40. C.F. Bender and E.R. Davidson, J. Phys. Chem. 70, 2675 (1966).

41. G. Das, J. Chem. Phys. 46, 1568 (1967).

42. D.D. Konowalow and M.L. Olson, J. Chem. Phys. 71, 450 (1979).

43. G.C. Lie and E. Clementi, J. Chem. Phys. 60, 1275 (1974).

44. G.C. Lie and E. Clementi, J. Chem. Phys. 60, 1288 (1974).

45. Thus, energy is in hartrees, length in bohr, charge in units of e, and the diffusion constant D = 1/2.

46. S. Aung, R.M. Pitzer, and S.I. Chan, J. Chem. Phys. 49, 2071 (1968).

47. J.W. Moskowitz, K.E. Schmidt, M.A. Lee, and M.H. Kalos, preprint (1982).

Table I. Comparison of the total ground-state energy obtained with the fixed-node QMC procedure, versus the estimated Hartree-Fock limit, CI, and "exact" energies. Except as noted, "exact" means the non-relativistic, Born-Oppenheimer energy, derived from experiment. The "quality" of each of the three importance functions (Ψ_I, Ψ_{II}, and Ψ_{III}) is also indicated, by giving the energy obtained from them in a variational calculation. Energies are in hartrees.

		H_2	LiH	Li_2	H_2O
Hartree-Fock		-1.1336 [a]	-7.987 [b]	-14.872 [c]	-76.0675 [d]
Ψ_I	variational	-1.1507±0.0009	-7.91±0.01	-14.85 ±0.03	-75.69 ± 0.03
	fixed-node	-1.1745±0.0008	-8.047±0.005	-14.985±0.005	-76.23 ± 0.02
Ψ_{II}	variational		-7.975±0.005	-14.900±0.004 [e]	-76.13 ±0.07
	fixed-node		-8.059±0.004	-14.991±0.007	-76.377±0.007
Ψ_{III}	variational	-1.162±0.001	-8.041±0.008	-14.95 ± 0.01	
	fixed-node	-1.174±0.001	-8.067±0.002	-14.990±0.002	
Best CI		-1.1731 [f]	-8.0606 [g]	-14.903 [h]	-76.3683 [i]
"Exact"		-1.17447... [j]	-8.0699 [k,m]	-14.9967 [l,m]	-76.4376 [d]

(a) Obtained with a nine term expansion in Ref. 35.
(b) Ref. 36.
(c) Ref. 37.
(d) Rosenberg and Shavitt in Ref. 4.
(e) Variational energy from Moskowitz and Kalos in Ref. 18.
(f) Ref. 38. Of course, better correlated wavefunctions than CI exist for H_2. For example, Ref. 35 obtains E=-1.1744 from a 40 term expansion which includes r_{ij} explicitly, and the "exact" result of Ref. 39 also uses this method.
(g) Ref. 40.
(h) Refs. 41, 42.
(i) Meyer in Ref. 4.
(j) Ref. 39. This value is not derived from experiment, but directly from theory.
(k) Ref. 43.
(l) Ref. 44.
(m) Here the zero-point energy has not been subtracted; also the relativistic correction is assumed independent of $r_{\alpha\beta}$, and the Lamb shift has not been included.

Table II. Ground-state energies at selected nuclear separations for Li_2. Results of the fixed-node QMC calculation, obtained using the importance function Ψ_{II}, are compared with Hartree-Fock and "exact" energies (in hartrees). Typical statistical uncertainty in the fixed-node results is 0.005 a.u.

R (Bohr)	E_{H-F} [a]	E_{F-N}	$E_{"exact"}$ [b]
3	-14.786	-14.905	-14.915
4	-14.853	-14.968	-14.983
5.05	-14.872	-14.991	-14.997
6	-14.869	-14.985	-14.992
7	-14.859	-14.976	-14.982

(a) Refs. 37, 41, 42.

(b) Ref. 44.

Table AI. Nuclear geometry used for each molecule.

Molecule	Bond length (Bohr)
H_2	1.401
LiH	3.015
Li_2	5.05
H_2O	O-H: 1.8111 (Θ_{H-O-H} = 104° 27')

Table AII. Trial wavefunctions Ψ_I. The form of Ψ_I is given by Eq. (8). The Slater determinant of molecular orbitals (MO's) is constructed from the linear combination of atomic Slater type orbitals (STO's) shown here. The orbital exponents (ζ) and the parameters (a, b) in the Jastrow correlation factor are also listed. Atomic units ($bohr^{-1}$) are used for a and b.

Molecule	(a, b)	STO	(ζ)	MO Coefficients ψ_1	ψ_2	ψ_3
H_2	(0.28,0.05)	$1s_a$	(1.285)	1		
		$1s_b$	(1.285)	1		
LiH	(0.5, 0.5)	$1s_{Li}$	(2.8)	1	0	
		$2pz_{Li}$	(1.2)	0	1.1	
		$1s_H$	(1.27)	0	1.0	
Li_2	(0.5, 0.5)	$1s_a$	(2.8)	1	0	0
		$1s_b$	(2.8)	0	1	0
		[a] $2s_a$	(1.2)	0	0	1
		[a] $2s_b$	(1.2)	0	0	1
H2O	(0.5, 3.5)	See wavefunction I of Aung et. al.[b]				

(a) The 2s atomic orbitals used here for Li_2 are hydrogenic 2s orbitals rather than STO's.

(b) Ref. 46.

Table AIII. Trial wavefunctions Ψ_{II}. Notation as in Table AII.

Molecule	(a, b)	STO	(ζ)	MO Coefficients ψ_1	ψ_2	ψ_3
LiH	(0.5, 0.5)	$1s_{1,Li}$	(2.521)	0.894	-0.128	
		$1s_{2,Li}$	(4.699)	0.103	-0.004	
		$2s_{Li}$	(0.797)	-0.003	0.346	
		$2pz_{1,Li}$	(0.737)	-0.001	0.176	
		$2pz_{2,Li}$	(1.2)	-0.004	0.046	
		$1s_{1,H}$	(0.888)	0.007	0.601	
		$1s_{2,H}$	(1.566)	0	0.1	
		$2pz_H$	(1.376)	0.002	0.017	
[a] Li_2	(0.5, 1.0)	$1s_a$	(2.69)	1	1	0
		$1s_b$	(2.69)	1	-1	0
		$2s_a$	(0.694)	0	0	1
		$2s_b$	(0.694)	0	0	1
H_2O	(0.5, 3.5)	See wavefunction II of Aung et. al.[b]				

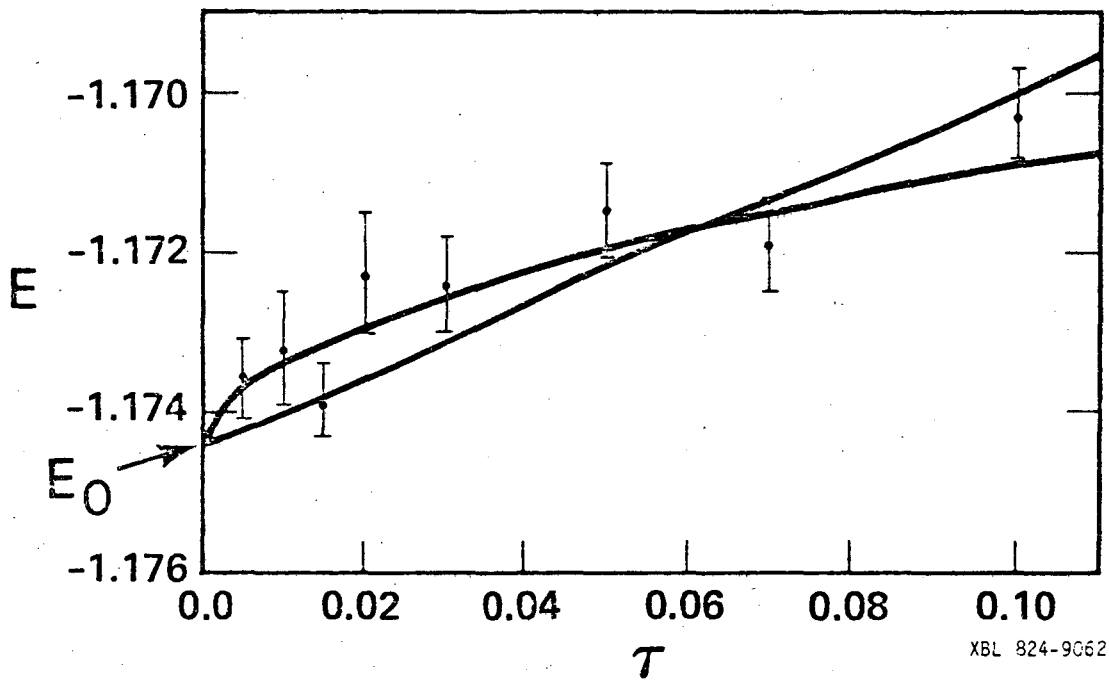
(a) From Moskowitz and Kalos in Ref. 18.

(b) Ref. 46.

Table AIV. Parameters used in the localized trial wavefunctions Ψ_{III} which are of the form of Eq.(9). For each molecule, the separate rows give the parameters for one molecular orbital (MO). The molecules are aligned along the x-axis with the center of the bond at the origin. The parameters have been optimized for a linear combination of the lowest energy and lowest variance of the energy. Numbers are in atomic units.

Molecule	β	MO	w_k	v_k	$\vec{c}_k \cdot \hat{x}$
H_2	9.913	ψ_1	2.74	0.0	0
LiH	1.0358	ψ_1	0.568	0.4143	-1.490
		ψ_2	0.8937	1.463	1.513
Li_2	0.5766	ψ_1	0.509	0.41	2.525
		ψ_2	0.509	0.41	-2.525
		ψ_3	3.33	1.49	0

Figure 1. The total energy of the H_2 molecule computed by the Monte Carlo fixed-node algorithm versus the time step τ. The two fits to the data shown are $E_o + A_i \tau^{i/2}$ where $i = 1, 2$. A χ^2 test shows that for either power law the Monte Carlo results are consistent with an extrapolation to the exact ground-state energy E_o. The trial function used in this calculation was Ψ_{III}.


-1.170
-1.172
-1.174
-1.176
E
E_0
0.0
0.02
0.04
0.06
0.08
0.10
τ
XBL 824-9062

TECHNICAL INFORMATION DEPARTMENT
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720