UC 404
LBL-28994

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

Materials & Chemical Sciences Division

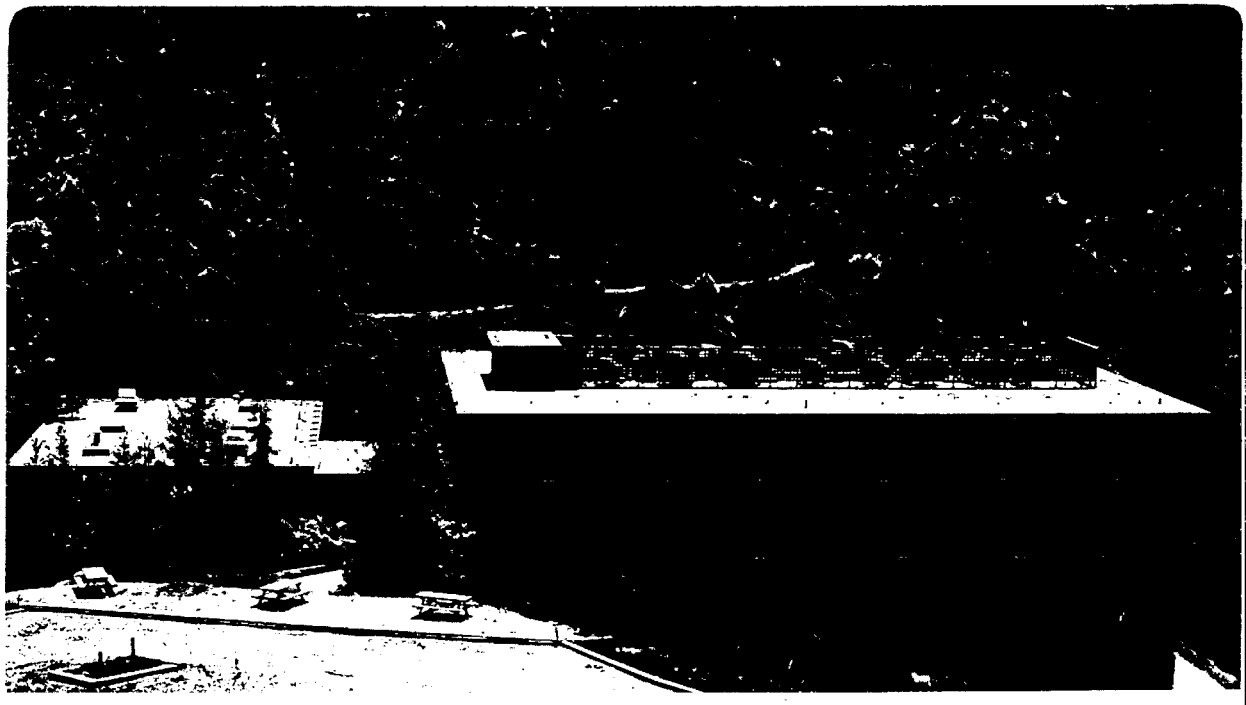
Quantum Monte Carlo for Atoms and Molecules

R.N. Barnett
(Ph.D. Thesis)

November 1989

For Reference

Not to be taken from this room



LBL-28994
Copy 1
Bldg. 50 Library.

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098.

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBL-28994

QUANTUM MONTE CARLO FOR ATOMS AND MOLECULES*

Ph.D. Thesis

Robert Nicholas Barnett

Department of Chemistry

University of California, Berkeley

and

Materials and Chemical Sciences Division

Lawrence Berkeley Laboratory

1 Cyclotron Road

Berkeley, California 94720

November 1989

*This research was supported by the Director, Office of Basic Energy Research Sciences, Chemical Sciences Division of the U. S. Department of Energy under Contract No. DE-AC03-76SF00098.

Quantum Monte Carlo for Atoms and Molecules

by

Robert Nicholas Barnett

Abstract

The diffusion quantum Monte Carlo with fixed nodes (QMC) approach has been employed in studying energy-eigenstates for 1-14 electron systems. Previous work employing the diffusion QMC technique yielded energies of high quality for H_2, LiH, Li_2, and H_2O. Here, the range of calculations with this new approach has been extended to include additional first-row atoms and molecules. In addition, improvements in the previously computed fixed-node energies of LiH, Li_2, and H_2O have been obtained using more accurate trial functions. All computations were performed within, but are not limited to, the Born-Oppenheimer approximation. In our computations, the effects of variation of Monte Carlo parameters on the QMC solution of the Schrödinger equation were studied extensively. These parameters include the time step, renormalization time and nodal structure. These studies have been very useful in determining which choices of such parameters will yield accurate QMC energies most efficiently. Generally, very accurate energies (90-100% of the correlation energy is obtained) have been computed with single-determinant trial functions multiplied by simple correlation functions. Improvements in accuracy should be readily obtained using more complex trial functions.

Further work involved substantial modifications to our QMC algorithm in order to develop new approaches for computing properties other than the energy. Calculating such properties involves the computation of expectation values of coordinate operators, over the exact distribution, which yield which dipole and higher moments. Several algorithms were studied for H, H_2 and LiH. Expectation values and moments are in excellent agreement, $\approx 0.5\%$, with exact theoretical results, H and H_2, or experiment, LiH.

Finally, the scope of diffusion QMC has been further broadened by developing novel methods to compute double-state expectation values, e.g., transition dipole moments. The calculations were motivated by the fact that transition dipole moments are of great interest in chemistry and physics but are difficult to compute using expansion-based techniques. In studying the feasibility of QMC to obtain accurate double-state expectation values, several new QMC approaches for determining several states simultaneously were developed and employed in computing the transition dipole moment for the $1s \rightarrow 2p_x$ transition of the hydrogen atom. The most efficient of these was used to compute the $2^2S \rightarrow 2^2P$ transition dipole moment of lithium atom. The resulting moment and QMC-computed energies of each state yield an oscillator strength and excited state lifetime in excellent agreement with precise experimental measurements. For the oscillator strength the QMC and experimental values are 0.742(7) versus 0.742(1), and for the excited-state lifetime, 27.41(35) $\times 10^{-9}$s versus 27.29(4) $\times 10^{-9}$s. (For QMC values the numbers in parentheses represent one standard deviation in the mean.)

Dedication

To my mother and father

Acknowledgments

I wish to express my gratitude to Professor William A. Lester, Jr., for advice and support during my graduate student years. The benefits I have derived from his guidance will prove invaluable for years to come. I am also thankful for the kindness and understanding he has shown me; it has been a pleasure working with him. I am deeply indebted to Dr. Peter J. Reynolds. I owe much of my scientific development and knowledge of QMC to his advice and teaching. In addition, he has been extremely helpful in introducing and commenting on many of the ideas presented in this thesis.

I would also like to thank the former members of Prof. Lester's research group Drs. Randall Grimes and Christopher Dateo, for much assistance with the HONDO program and useful advice in preparing for my preliminary examinations. I also thank Kent Owen and Dr. Brian Hammond for fruitful discussions on QMC. I would especially like to thank Dr. Hammond for his advice on bureaucratic details.

I am also indebted to Maudie Noyd for her help with manuscripts and other assistance she has provided. Finally, I thank Linda Allcroft and Larry Joseph whose friendship has given me great support.

This research was supported by the Director, Office of Basic Energy Research Sciences, Chemical Sciences Division of the U. S. Department of Energy under Contract No. DE-ACO3-76SF00098.

Table of Contents

Chapter 1. Introduction 1

1.1 Theoretical Background 1

1.2 Approximate Theoretical Approaches 4

1.3 Self-Consistent Field and Configuration Interaction 5

Chapter 2. Quantum Monte Carlo 9

2.1 Introduction 9

2.2 Monte Carlo Background 11

2.3 Quantum Monte Carlo 13

2.4 Fixed-Node Diffusion Quantum Monte Carlo 17

2.4.1 The Schrodinger Equation in Imaginary Time 17

2.4.2 Importance Sampling 20

2.4.3 The Short-Time Green's Function 21

2.4.4 Fermi Statistics: The "Fixed-Node" Approximation 23

2.4.5 Variational Monte Carlo 25

2.5 Quantum Monte Carlo Algorithm 26

Chapter 3. The Computation of QMC Energies 35

3.1 Introduction 35

3.2 Time-Step Bias in Diffusion QMC Calculations 36

3.3 Computations of Energy Differences 44

3.3.1 The Barrier to H + H_2 Exchange 45

3.3.2 The Electron Affinity of Fluorine 53

3.3.3 The Bond Energy of Nitrogen 60

3.4 Energy Computations on Other First-Row Atoms and Molecules 63

3.5 Computing the QMC Energy by Difference 74

Tables and Figures for Chapter 3 84

Chapter 4. Expectation Values of Coordinate Operators 128

4.1 Introduction 128

4.2 Algorithms for the Computation of Pure Expectation Values 131

4.2.1 Pure Expectation Values by a Single QMC Walk 131

4.2.2 Branching Algorithms with Weighting 136

4.2.3 Pure Expectation Values by VMC with QMC "Side Walks" 139

4.3 Results for H and H_2 141

4.4 Properties for LiH and BH 145

4.5 Moments with Odd Powers 155

Tables and Figures for Chapter 4 158

Chapter 5. The Computation of Transition Dipole Moments 170

5.1 Introduction 170

5.2 Approximate Transition Dipole Moments 172

5.3 Exact Transition Dipole Moments: QMC Walks within QMC

Walks (Method 1) 174

5.4 Exact Transition Dipole Moments: VMC Walk with QMC Side Walks (Method 2) 180

5.5 Exact Transition Dipole Moments: Green's Function Approach (Method 3) 188

5.6 Evaluation of Each Method 196

5.7 Calculation of the Li $2^2S \rightarrow 2^2P$ Oscillator Strength 198

Tables and Figures for Chapter 5 205

Chapter 6. Summary and Conclusions 219

Appendix A. A Condition for the Equivalence of Nodal Volumes 222

Appendix B. The Computation of the Trial Function and its Derivatives 224

Appendix C. Antisymmetry Nodes of H_3 227

Appendix D. Variance of Rounded Versus Unrounded Weights 229

References and Footnotes. 231

Chapter 1

Introduction

1.1 Theoretical Background

A fundamental tenet of theoretical chemistry is that quantities which can be measured experimentally may also be computed from quantum mechanics. That is, for a given system, e.g., atom or molecule, the theoretical approach involves finding the state wavefunction, Φ_I, whose modulus squared is the probability density function for the system in state I. Since relativistic effects have little affect on most chemical properties, the wavefunction is obtained by solving, as originally proposed by Schrödinger [1], the equation which bears his name,

$$H\Phi_I = E_I\Phi_I \ . \tag{1.1}$$

In this equation, H is the Hamiltonian operator for the system in question, and E_I is the energy for the system in state I. The wavefunctions, Φ_I, are time-independent or "stationary" states. To obtain a time-dependent wavefunction, $\Phi_I(t)$, the time-dependent Schrödinger equation

$$i\hbar\frac{\partial\Phi(t)}{\partial t} = H\Phi(t) \tag{1.2}$$

must be solved. If A is an operator corresponding to an observable, i.e., a quantity which may in principle be measured, then the expectation value of A

$$<A> \equiv <\Phi_I | A | \Phi_I> \tag{1.3}$$

represents the value of this observable upon measurement. This last equation demonstrates the contributions theory can make; any property of a state can be determined once Φ_I is known.

For a system with N particles the Hamiltonian is given by

$$H = \sum_{i=1}^{N} T_i + V(1, 2, \cdots, N) , \tag{1.4}$$

where $T_i = -\frac{1}{2}\hbar^2 m_i^{-1}\nabla_i^2$ is the kinetic energy operator (in coordinate space) and V is the potential energy. For a system with N_α nuclei and N_e electrons, H, in atomic units ($\hbar$, m_e (electron rest mass), and e (unit charge) = 1) is,

$$H = -\frac{1}{2}\sum_{\alpha=1}^{N_\alpha} m_\alpha^{-1}\nabla_\alpha^2 - \frac{1}{2}\sum_{i=1}^{N_e}\nabla_i^2 - \sum_{i,\alpha}^{N_e, N_\alpha} Z_\alpha r_{i\alpha}^{-1} + \sum_{i<j}^{N_e} r_{ij}^{-1} + \sum_{\alpha<\beta}^{N_\alpha} Z_\alpha Z_\beta r_{\alpha\beta}^{-1}, \tag{1.5}$$

where Greek and Roman subscripts denote nuclei and electrons, respectively. Also, m_α is the mass of nucleus α and Z_α its charge.

The theoretical approach for studying chemical systems is now formulated: solve the Schrödinger equation (1.1) with the Hamiltonian given by Eq. (1.5). Unfortunately, such solutions are only known, in closed analytic form, for one-electron atoms. The first step towards obtaining approximate, albeit accurate, wavefunctions is to write the Hamiltonian as,

$$H = T_N + H_e , \tag{1.6}$$

where the electronic Hamiltonian, H_e, is

$$H_e \equiv -\frac{1}{2}\sum_{i=1}^{N_e}\nabla_i^2 - \sum_{i,\alpha}^{N_e, N_\alpha} Z_\alpha r_{i\alpha}^{-1} + \sum_{i<j}^{N_e} r_{ij}^{-1} + \sum_{\alpha<\beta}^{N_\alpha} Z_\alpha Z_\beta r_{\alpha\beta}^{-1}. \tag{1.7}$$

and T_N is the first term in Eq. (1.5). Eigenfunctions of H_e, Ψ_i, are then obtained by

solving

$$H_e \Psi_i(\underline{r}; \underline{R}) = E_i(\underline{R})\Psi_i(\underline{r}; \underline{R}) , \tag{1.8}$$

where $\underline{r}$ and $\underline{R}$ represent all the electron and nuclear coordinates, respectively. The semi-colon indicates that the dependence on the nuclear coordinates is understood to be parametric. The full solutions to Eq. (1.1) may then be written as

$$\Phi_I(\underline{r}, \underline{R}) = \sum_{i=1}^{\infty} \chi_i^I(\underline{R})\Phi_i(\underline{r}; \underline{R}) \tag{1.9}$$

with

$$[T_N + E_i(\underline{R})]\chi_i^I(\underline{R}) = E_I \chi_i^I(\underline{R}) . \tag{1.10}$$

A further simplification results from the Born-Oppenheimer approximation[2], where the nuclear and electronic motions are assumed to be separable. Since nuclear masses are much greater than the electron mass, this approximation is generally very accurate. In this approximation the total wavefunction is written as a single term of Eq. (1.9),

$$\Phi_I(\underline{r}, \underline{R}) = \chi_i^I(\underline{R})\Phi_i(\underline{r}; \underline{R}) . \tag{1.11}$$

The quantity, $E_i(\underline{R})$ in Eq. (1.8), is referred to as the potential energy surface (PES) for the nuclear motion for the electronic state i.

The problem of solving the Schrödinger equation with the full Hamiltonian is reduced to finding eigenfunctions of H_e, and then solving for the nuclear motion from Eq. (1.10). However, eigenfunctions of H_e are only known for one-electron atoms and one-electron homonuclear diatomics. This is due to the fact that analytic wave functions which exactly describe the instantaneous correlations of two or more electrons are not known. Therefore, approximate, though generally accurate, approaches are employed in theoretical chemistry calculations as outlined in the next section.

1.2 Approximate Theoretical Approaches

The benefit of approximate theoretical techniques is that, since the exact solution can not be determined, an algorithm for obtaining inexact, but hopefully very accurate, wavefunctions is employed. The most widely used of such approaches are those based on expansions in basis functions[3], perturbation theory[4] and density functional theory[5]. In studies of chemical systems expansion methods are the most commonly employed.

For the lowest-energy state of a given symmetry, expansion techniques can take advantage of the variational principle[6] in estimating the accuracy of a computed energy. This principle states that the energy expectation value of a function, Ψ,

$$E[\Psi] = \int d\underline{R}\, \Psi^* H \Psi / \int d\underline{R}\, |\Psi|^2 \tag{1.12}$$

is always above the exact-ground state energy. (From hereon, $\underline{R}$ denotes a 3N-dimensional vector specifying the electronic coordinates.) Therefore, parameters in Ψ are varied to obtain a minimum in $E[\Psi]$. (A similar variational principle for excited states has been formulated by MacDonald[7].)

Expansion techniques are useful because, when Ψ is expanded in a complete set of basis functions, the minimization of $E[\Psi]$ does yield the ground-state energy, E_0. Also, for judicious choices of the basis functions, convergence to E_0 can be rapid. Accurate energies had been obtained as early as the work of Hylleraas[8] in computing the ground-state energy of He. Essentially exact energies have been calculated by Pekeris[9] for He, and by Kolos and Wolniewicz[10] for H_2. Using Hylleraas-type expansions, King and Shoup[11] and Alhenius and Larsson[12] computed highly accu-

rate energies for the ground state and first-excited P state of Li, respectively. For these systems the near-convergence to the exact energy is greatly aided by employing functions which explicitly involve powers of r_{ij}. For other atoms and molecules this explicit description of electron correlation is generally not incorporated into Ψ due, in part, to the reduced usefulness of Hylleraas expansions for more complex systems. In addition, powers of r_{ij}, when used with Slater-type orbital (STO) basis sets, necessitate the computation of the many-center two-electron integrals by numerical quadrature. The result is that large expansions in one-electron coordinates are a more practical alternative, but the accuracies described above in this paragraph are usually not attainable.

Despite this drawback, useful chemistry is derived from expansion techniques because much of chemistry is concerned with energy differences, e.g., bond and excitation energies, and potential energy surfaces. The difference between two computed total energies, which are upper bounds, may be much more accurate than either of these energies individually, due to cancellation of errors. Also, the properties of a PES are unaffected by the addition of a constant. Therefore, a computed PES may yield accurate properties, e.g., spectroscopic constants, by being above, but nearly parallel to, the exact surface. However, such fortuitous cancellation of errors does not always occur. Noted exceptions are the electron affinities of the first-row atoms - despite apparently accurately computed energies for the atom and anion[13] .

1.3 Self-Consistent Field and Configuration Interaction

Since the Self-Consistent Field (SCF) and Configuration Interaction (CI) techniques

are cornerstones of theoretical chemistry, they are outlined below.

The SCF method is based on an orbital description and was developed by Hartree[14], Fock[15] and Slater[16]. (For an excellent description, see Roothaan[17].) In this approach orbitals, ψ_i, are expanded in a set of one-particle basis functions, ϕ_j, and multiplied by a function of spin, $\chi_i(s)$,

$$\psi_i(\underline{x}) = (\sum_j c_{ij}\phi(\underline{r})_j)\chi_i(s) \ , \tag{1.13}$$

where $\underline{x}$ is the spin-space variable $\underline{rs}$, and $\chi_i(\frac{1}{2}) = \alpha$ ("spin-up") and $\chi_i(-\frac{1}{2}) = \beta$ ("spin-down). This orbital description yields an N-electron wavefunction

$$\Psi = \psi_1(1)\psi_2(2)\cdots\psi_N(N) \ . \tag{1.14}$$

To obtain anti-symmetry upon electron interchange, a Slater determinant is formed,

$$\Psi(1, 2, \cdots, N) = \sum(-)^P P\,\psi_1(1)\psi_2(2)\cdots\psi_N(N) \ , \tag{1.15}$$

where $\sum(-)^P P$ is over all possible N-particle permutations.

The optimization of the linear coefficients, $\{c_{ij}\}$, to minimize $E[\Psi]$, requires that the orbitals satisfy

$$H_1\psi_i(\underline{x}_1) + \sum_{j=1}^{N}\int\psi_i(\underline{x}_1)\psi_j(\underline{x}_2)r_{12}^{-1}(1 - P_{12})\psi_i(\underline{x}_1)\psi_j(\underline{x}_2)d\underline{x}_2 = \varepsilon_i\psi_i(\underline{x}_1). \tag{1.16}$$

In Eq. (1.16), H_1 is the one-electron part of the Hamiltonian,

$$H_1 = -\frac{1}{2}\nabla_1^2 - \sum_{\alpha}^{N_\alpha} Z_\alpha r_{1\alpha}^{-1} + \sum_{j\neq i}^{N_e} r_{1j}^{-1} + \sum_{\alpha<\beta}^{N_\alpha} Z_\alpha Z_\beta r_{\alpha\beta}^{-1} \ , \tag{1.17}$$

Note that in Eq. (1.16), each orbital is dependent on the $N - 1$ remaining orbitals so that the N equations must be solved simultaneously. The solutions are therefore obtained, i.e., the values of $\{c_{ij}\}$ are found which satisfy Eq. (1.16) for all orbitals ψ_i, in an iterative process known as the "self-consistent field" (SCF) approach. (See Ref. 17 for more details.) In the limit that a complete 1-particle basis $\{\phi_i\}$ is used, the

lowest energy obtainable from this technique, the Hartree-Fock (HF) energy, results. The approximation inherent in this method is that electrons move in a potential resulting from a static field. That is, the instantaneous correlations of the electronic motion are not described.

The HF energy is used to define a term widely known in the field of theoretical chemistry – the correlation energy. For a given state, the correlation energy is defined as the difference between the eigenvalue of the Hamiltonian employed for this state and its HF energy[18]. Improvements over the HF description are then measured in terms of the amount of correlation energy gained.

The Slater determinant represents an N-particle wavefunction based on an expansion in a 1-particle basis set. To improve on this function, by describing instantaneous electron correlations, an expansion in N-particle functions (Slater determinants) is usually constructed[19]. This is the basis of the CI approach in which the wavefunction is written as

$$\Psi=\sum c_I \Phi_I \ , \tag{1.18}$$

where each Φ_I now represents a Slater determinant. However, as referred to in the last section, such expansions can converge very slowly[20]. This difficulty has lent impetus to the development of new techniques.

In Chapter 2 the quantum Monte Carlo (QMC) approach, with emphasis on fixed-node QMC using a ''short-time'' Green's function, is discussed. In Chapter 3 computations of QMC energies for several systems, 1-14 electrons, are presented and discussed. Chapter 4 discusses and presents new algorithms for the computation of single-state properties other than the energy, and in Chapter 5 these ideas are extended

to the calculation of expectation values across different states. Chapter 6 ends this dissertation with a summary and conclusions.

Chapter 2

Quantum Monte Carlo

2.1 Introduction

As discussed in the previous chapter, expansion techniques can yield a very slowly converging energy, and if cancellation of errors is not complete, inaccurate energy differences result. Additionally, while the energy may be well-converged, other computed properties may still not be accurate. This is because the energy, $E[\Psi]$, is accurate to second order in the error in Ψ, but properties obtained from expectation values of operators which do not commute with the Hamiltonian are accurate only to first order. Since Ψ is chosen to minimize the computed energy, rather than local deviations from the eigenfunction, Φ, regions will exist in which Ψ is a poor approximation to Φ. While this effect may not be significant for the energy, results for properties that are sensitive to the local accuracy of Ψ in such regions may be quite poor. For example, accuracy in Ψ at the periphery of an atom or molecule is not important for energy minimization, but is important for properties such as polarizabilities and permanent moments[21]. A noted example of the difficulty in obtaining accurate results, despite very accurately computed energies, is the computation of transition dipole moments using expansion techniques[22].

The sensitivity of many atomic and molecular properties to the accuracy of Ψ,

rather than of $E[\Psi]$, translates into a marked dependence on the 1-particle expansion (basis set). Therefore, for such properties, large basis sets are required[21], greatly increasing the computational cost. Finally, convergence of a computed property (other than the energy) to the exact value, as the 1- and N-particle bases are increased, is not necessarily monotonic.

An alternative to minimization of the energy is to minimize the expectation value of the variance[23], $U^2[\Psi]$, rather than the energy. Here,

$$U^2[\Psi] \equiv \int (H\Psi - E\Psi)^2 d\underline{R} / \int |\Psi|^2 d\underline{R} \ . \tag{2.1}$$

The integrand in the numerator can be significant even when Ψ is small. Therefore, the quality of the wavefunction should be much more uniform than when energy minimization is employed. Also, since $U^2 \geq 0$, this functional is variational. Unfortunately, minimizing U^2 is quite difficult when expansion techniques are employed, due to the presence of $(H\Psi)^2$ in the integrand. However, such minimizations present no special difficulty for Monte Carlo techniques, and are currently being employed with much success[24,25].

As stated in the discussion to this point, obtaining highly accurate solutions of the Schrödinger equation remains a challenging problem. Therefore, it is of interest to investigate a method which is not predicated upon large 1- and N-particle expansions yet can, in principle, yield exact results. One approach, generally referred to as quantum Monte Carlo (QMC), is employed in the computations described in this thesis. Several variants of QMC exist. A brief discussion of Monte Carlo the and variants of QMC are discussed in the next two sections.

2.2 Monte Carlo Background

Monte Carlo methods employ random numbers to solve the problem at hand. Despite the random nature of the approach, Monte Carlo techniques may be employed to compute quantities which are not inherently statistical.

A simple example is offered by the computation of a definite integral,

$$I = \int_0^1 f(x)dx \ . \tag{2.2}$$

To compute this integral by Monte Carlo, one may select a set of values of x, $\{x_i\}$, such that $x_i = \xi_i$, where ξ_i is a uniform random variate between 0 and 1. This is referred to as "sampling" values of x [and hence of $f(x)$]. A Monte Carlo "estimate" of I, I_N, is then given by

$$I_N = \frac{1}{N}\sum_{i=1}^{N} f(x_i) \ , \tag{2.3}$$

where N is the number of points sampled and $\lim_{N\to\infty} I_N = I$. The statistical nature of the method is now apparent; for finite values of N the average in Eq. (2.3) need not equal I. That is, the "statistical error", generally taken to be one standard deviation in I_N, will cause $I_N \neq I$. From the central limit theorem, for N sufficiently large independent values of I_N may be assumed to be distributed according to a Gaussian distribution[26] for which one standard deviation is referred to as the statistical error in I_N.

It is instructive to consider the statistical error of this example and a well-known method for reducing it. For N points, the statistical error in sampled values of f is most generally measured as

$$\sigma_f(N) = [N^{-1}\sum_{i=1}^{N}(f(x_i) - I)^2]^{\frac{1}{2}} . \tag{2.4}$$

(When the average I is not known it is replaced by Eq. (2.3) and N^{-1} is replaced by $(N-1)^{-1}$.) Generalizing to sampling the entire distribution yields the statistical error in a sampled value of f as

$$\sigma_f = [\int_0^1 f^2(x)dx - (\int_0^1 f(x)dx)^2]^{-\frac{1}{2}} . \tag{2.5}$$

The standard deviation in the computed average of f, I_N, is given by σ_N, $\sigma_N = \sigma_f/\sqrt{N}$ [28], or when σ_f is not known, $\sigma_f(N)$ is employed. The statistical error in I_N may be decreased to any value desired simply by sampling more values of x. However, the dependence of σ_N on N causes the process of error reduction to be one of diminishing returns. That is, to reduce the statistical error by an order of magnitude requires an increase of two orders of magnitude in the sampling.

An efficient way to reduce the statistical error is to employ importance sampling[27]. Importance sampling involves recasting Eq. (2.2) as

$$I = \int_0^1 [\frac{f(x)G}{g(x)}]\frac{g(x)}{G}dx , \tag{2.6}$$

where values of x are now sampled from the probability distribution $g(x)/G\ (\geq 0)$ where $\int_0^1 g(x)dx = G$. Now, Monte Carlo estimates of I can be obtained from

$$I_N = \frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{g(x_i)}G . \tag{2.7}$$

Note that for $g(x) = f(x)$, $(f(x) \geq 0)$, $I_N = I$ and the statistical error is zero for all values of N. Of course, such a selection implies $G = \int_0^1 f(x)dx = I$ is known, which is

assumed not to be the case. Therefore, one chooses a function $g(x) \geq 0$ (with G known) which mimics the behavior of $f(x)$, i.e., $g(x) \approx f(x)$. To the extent that g resembles f, values of x are sampled where f is large, or important, thus "importance sampling". In addition, since $g \approx f$, fluctuations in the ratio $f(x)/g(x)$ will be much less than those in $f(x)$ with the result that $\sigma_{(f/g)} << \sigma_f$, and the statistical error of I_N is reduced.

The simple example above involved solving a non-stochastic problem using random numbers. Naturally, Monte Carlo techniques may be employed in studying random processes as well, such as neutron transport[29] and Brownian motion. The use of Monte Carlo techniques for solving the Schrödinger equation is discussed in the next section.

2.3 Quantum Monte Carlo

Over forty years ago Fermi suggested that the time-dependent Schrödinger equation could lend itself to simulation when written in imaginary time[30]. The first significant QMC calculations were undertaken by Kalos in studying three- and four-body nuclei[31], atomic helium[32], and a Boson fluid interacting through a Lennard-Jones potential[33]. In these calculations the time-independent Schrödinger equation is cast into an integral equation. That is,

$$H\psi = E\psi, \quad H = H_0 - W \tag{2.8}$$

is written in integral form as

$$\psi(\underline{R}') = \int G_0(\underline{R} \rightarrow \underline{R}')[E + W(\underline{R})]\psi(\underline{R})d\underline{R} \ . \tag{2.9}$$

Here G_0 is the inverse operator of H_0 which satisfies $H_0 G_0(\underline{R} \rightarrow \underline{R}') = \delta(\underline{R} \rightarrow \underline{R}')$

and any other boundary conditions inherent in the physics of the problem at hand. The zeroth-order Hamiltonian, H_0, and by implication G_0, is chosen so that values of $\underline{R}'$ can be readily sampled from $G_0(\underline{R} \rightarrow \underline{R}')$. The Green's function, G_0, serves as transition probability density function and W, the residual part of the Hamiltonian, acts as a weighting factor. The procedure for obtaining the ground state is to iterate a set of points sampled from some initial distribution, $\psi^{(0)}$, to a set distributed according to a final distribution, $\psi^{(n)}$. It is easily shown[33] that as n becomes large, $\psi^{(n)}$ becomes the lowest-energy eigenfunction of H not orthogonal to $\psi^{(0)}$. This approach yielded encouraging results for the problems studied. Basically, a differential equation is solved by casting it into an integral equation and sampling a Green's function by Monte Carlo - thus "Green's function Monte Carlo (GFMC)".

A major advance in GFMC was the inclusion of importance sampling in computations on helium-like systems at 0° Kelvin[34]. Here the integral equation is derived using the full Green's function, i.e., $W = 0$, and $\psi_0 = E_0 H^{-1}\psi_0$ is written

$$\psi^{(n+1)}(\underline{R}') = E_R \int G(\underline{R}' \rightarrow \underline{R})\psi^{(n)}(\underline{R})d\underline{R} \tag{2.10}$$

To implement importance sampling using a trial function, Ψ_T, Eq. (2.10) is rewritten in terms of $f^{(n)}(\underline{R}) \equiv \Psi_T(\underline{R})\psi^{(n)}(\underline{R})$ as

$$f^{(n+1)}(\underline{R}') = E_R \int G_I(\underline{R} \rightarrow \underline{R}')f^{(n)}(\underline{R})d\underline{R} \ , \tag{2.11}$$

where $G_I(\underline{R} \rightarrow \underline{R}') \equiv \Psi_T(\underline{R}')G(\underline{R}' \rightarrow \underline{R})/\Psi_T(\underline{R})$. Also, the reference energy, E_R, is chosen as $\approx E_0$, $E_0 > 0$, in order to stabilize the normalization of the solution since $\psi^{(n+1)} = (E_R/E_0)\psi^{(n)}$ at large n. For $\Psi_T = \psi_0$ the energy can be obtained immediately with no statistical error[34]. Therefore, in a vein similar to that described in the first section, Ψ_T is chosen to mimic the behavior of the lowest-energy eigenfunction to

reduce statistical error. Subsequently, variants of the GFMC approach have been developed. For a discussion of QMC techniques see Refs. 26,35-36.

The two major difficulties encountered by Monte Carlo approaches in studying chemical systems are sampling the Green's function, since H^{-1} is generally not known, and obtaining an anti-symmetric solution. The latter difficulty is the most significant[37] as it is the only remaining barrier to the computation of statistically exact ground-state energies.

The lowest-energy Fermi state, described by an antisymmetric eigenfunction ϕ_0, is generally higher in energy than the nodeless Bose ground state. Therefore, the approach described thus far will not yield antisymmetric eigenfunctions unless the Monte Carlo solution is forced to have nodes. If the nodes of the Fermi ground state are known, ϕ_0 may be obtained simply by not allowing these nodes to be crossed. However, the exact location of the nodes of eigenfunctions describing chemical ground states are unknown. In this case the true nodal structure must be estimated by an approximate one, and these nodes *are* crossed. A sign of +1 is assigned to a point if it is arrived at after an even number of node crossings and the sign is -1 otherwise. Since ignoring the signs yields the (symmetric) Bose ground state, the lowest-energy (antisymmetric) Fermi state is obtained as the difference, by using the signs, between points sampled from the Bose ground state. This method becomes unstable because the energy of the Bose state (E_0^B), being lower than the Fermi-state energy (E_0^F), means that the normalization of the Bose state, and hence its statistical error, increases geometrically relative to that of the Fermi state[35,36]. Therefore, the Fermi state, obtained effectively as the difference between two continually growing "+" and "–"

distributions, is made statistically insignificant by the statistical errors in these components. (For a simple example of this effect see Ref. [38]). The time by which the Fermi state is "lost" depends on $|E_0^B - E_0^F|$, and so instabilities occur more quickly as the size of the system increases. Thus far, the use of GFMC to obtain the lowest-energy Fermi state has been limited to atoms and molecules with ten electrons or less.

An alternative to carrying signs is to "fix" the preassigned nodes.[39] That is, moves in $\underline{R}$-space are *not* allowed to cross prescribed nodal boundaries so that the instabilities induced by carrying signs is removed. The Schrödinger equation is now solved within each nodal volume. This approach, discussed in detail in the next section, yields an error depending on the accuracy of the fixed nodes. If the "fixed-node" error is sufficiently small, then the stability obtained by fixing the nodes is justified.

As stated on the previous page, the exact Green's function is generally not available in closed analytic form. In GFMC it is sampled in an iterative (and presumably convergent) process [34,35,40-42]. The complex and iterative nature of this procedure causes the sampling of the exact Green's function to be time consuming. If an accurate approximation to the Green's function, G_a, is known, then a direct sampling of G_a will be more efficient and yield little error. Such an approximation, to be discussed in more detail in the next section, was introduced by Ceperley and Alder to solve the Schrödinger equation in imaginary time[39].

The solution of the Schrödinger equation in imaginary time with fixed nodes, "Fixed-Node diffusion QMC", employing an analytic "short-time" approximation to the Green's function is now discussed.

2.4 Fixed-Node Diffusion Quantum Monte Carlo

2.4.1 The Schrödinger Equation in Imaginary Time

The QMC technique employed here is now described in detail. This description is similar to that of Reynolds *et al.*[43].

The theory begins with the time-dependent Schrödinger equation in imaginary time ($t \rightarrow t/i, \hbar = 1$)

$$-\frac{\partial \Phi(R, t)}{\partial t} = [H - E_R]\Phi(\underline{R}, t) , \tag{2.12}$$

where E_R is a reference energy which only affects the (imaginary) time dependence of $\Phi(\underline{R}, t)$. While the Hamiltonian above is completely general, from hereon the electronic Hamiltonian, H_e, is employed since this is the context of our computations. (Calculations employing the full non-relativistic Hamiltonian have been performed by Ceperley and Alder[44] for solid hydrogen at high pressures.) Writing Eq. (2.12) with H_e (in atomic units) yields

$$-\frac{\partial \Phi(R, t)}{\partial t} = [-\frac{1}{2}\nabla^2 + V - E_R]\Phi(\underline{R}, t) , \tag{2.13}$$

where, for N electrons,

$$\nabla^2 \equiv \sum_{i=1}^{N} \nabla_i^2 , \tag{2.14}$$

and V is the potential energy, [cf. Eq. (1.7)].

To determine the time dependence of Φ, this function is expanded in the complete set of normalized eigenfunctions of H,

$$\Phi(\underline{R}, t) = \sum_{i=0} c_i(t)\phi_i(\underline{R}) . \tag{2.15}$$

Substituting this expansion into Eq. (2.13) and solving for the expansion coefficients

yields

$$c_i(t) = c_i(0)\exp[-t(E_i - E_R)], \quad c_i(0) = \int \phi^*(\underline{R})\Phi(\underline{R}, 0)d\underline{R} \ . \tag{2.16}$$

Therefore, we may write

$$\Phi(\underline{R}, t) = \sum_{i=0} c_i(0)\exp[-t(E_i - E_R)]\phi_i(\underline{R}) \ , \tag{2.17}$$

so that the asymptotic form of Φ is

$$\Phi(\underline{R}, t) = c_0\exp[-t(E_0 - E_R)]\phi_0(\underline{R}) \ , \tag{2.18}$$

with $c_0 \equiv c_0(0)$. The subscript 0 denotes the lowest-energy eigenstate not orthogonal to $\Phi(\underline{R}, 0)$. The excited states decay, relative to ϕ_0, according to $\exp[-t(E_i - E_0]$, and E_R is chosen to be $\approx E_0$ to minimize the time-dependence of the asymptotic form of Φ.

The ground state $\phi_0(\underline{R})$ is then obtained by propagating an initial function sufficiently far in imaginary time via Monte Carlo simulation. Eq. (2.13) corresponds to a combination of diffusion ($\partial\Phi/\partial t = \frac{1}{2}\nabla^2\Phi$) and a first-order rate process ($\partial\Phi/\partial t = -(V - E_R)\Phi$). The QMC approach simulates these two processes by sampling the Green's function of the imaginary-time Schrödinger equation (2.13). This Green's function is

$$G(t) = \exp[-t(T + V - E_R)] \ , \tag{2.19}$$

and, in coordinate space, the time evolution of Φ is

$$\Phi(\underline{R}', t+\tau) = \int G(\underline{R} \rightarrow \underline{R}', \tau)\Phi(\underline{R}, t)d\underline{R} \ . \tag{2.20}$$

The Green's function is a solution to Eq. (2.13) in the coordinate $\underline{R}'$ with the boundary condition $G(\underline{R} \rightarrow \underline{R}', 0) = \delta(\underline{R} - \underline{R}')$.

The diffusion and first-order rate process (branching) may be simulated by employing a "short-time" approximation to the Green's function, G_a. That is, as $t \rightarrow 0$,

$$G(\tau) = G_a(\tau) \equiv \exp(-\tau T)\exp[-\tau(V - E_R)] \ , \tag{2.21}$$

where τ signifies small values of t, and G_a is accurate to order τ^2. The first factor corresponds to diffusion and the second to birth ($V < E_R$) or death ($V > E_R$). The form of G_a in coordinate space is easily derived, and the QMC simulation is readily performed.

This simulation involves propagating a set of points, initially distributed as $\Phi(\underline{R}, 0)$, until ϕ_0 is obtained as determined by the convergence of the energy or some other quantity. Subsequently, more points are sampled and the ground state energy, E_0, is computed from the average of the potential energy;

$$\begin{aligned} <V> \equiv \lim_{M\to\infty} \frac{1}{M}\sum_{i=1}^{M} V(\underline{R}_i) &= \int\phi_0(\underline{R})V(\underline{R})d\underline{R} \ / \int\phi_0(\underline{R})d\underline{R} \\ &= \int\phi_0(\underline{R})H\cdot 1 d\underline{R} \ / \int\phi_0(\underline{R})\cdot 1 d\underline{R} = E_0 \ . \end{aligned} \tag{2.22}$$

The second line is obtained from the first since (in coordinate space) the kinetic energy operator annihilates the constant function, 1, and the last equality is obtained from the hermiticity of H. Since error arises from G_a for non-zero τ (time step), either small time steps must be employed, or results must be extrapolated to $\tau = 0$.

Simulations of the type described above have been carried out by Anderson for one- to four-electron systems[45-47]. Large statistical errors in the energy resulted, e.g., about 30 kcal/mol for Be, because the energy is obtained from the average of the potential energy, cf. Eq. (2.22), a quantity which possesses large fluctuations. Despite the fact that computers would be one to two orders of magnitude faster in the near future, these calculations demonstrated that importance sampling would be required for precise QMC computations.

2.4.2 Importance Sampling

A particularly elegant way of incorporating importance sampling into the QMC approach was introduced by Kalos *et al*.[34] as described in section 2.3. In diffusion QMC, importance sampling involves defining a new density function $f(\underline{R}, t) \equiv \Psi_T(\underline{R})\Phi(\underline{R}, t)$. Rewriting Eq. (2.13) in terms of f yields

$$-\frac{\partial f}{\partial t} = -\frac{1}{2}\nabla^2 f + [E_L(\underline{R}) - E_R]f + \frac{1}{2}\nabla\cdot[f F_Q(\underline{R})] , \tag{2.23}$$

where $E_L(\underline{R}) \equiv \Psi_T^{-1}(\underline{R}) H \Psi_T(\underline{R})$ and $F_Q(\underline{R}) \equiv \nabla \ln|\Psi_T(\underline{R})|^2$ are the "local energy" and "quantum force", respectively.

Once again a diffusion equation is obtained. Here, however, the diffusive motion occurs in the presence of an external force, F_Q. The role of Ψ_T, in guiding the random walk, can be seen from the form of F_Q. As $|\Psi_T|^2$ decreases, $|F_Q|$ increases and the walk is guided away from regions where $|\Psi_T|^2$ is small. Another important difference, from the case of no importance sampling, is that the branching term is now given by $E_L(\underline{R}) - E_R$. For judicious choices of the trial function, e.g., to remove or minimize the Coulomb singularities, the branching is now much better behaved than $V(\underline{R}) - E_R$. Note that as $\Psi_T \rightarrow \phi_0$, the branching term becomes a constant, $E_0 - E_R$, since $E_L = \phi_0^{-1} H \phi_0 = E_0$. Thus, Ψ_T is chosen to approximate ϕ_0.

The asymptotic form of $\Phi(\underline{R}, t)$, cf. Eq. (2.18), yields the large t form of f,

$$f(\underline{R}, t) = c_0 \exp[-t(E_0 - E_R]\Psi_T(\underline{R})\phi_0(\underline{R}) . \tag{2.24}$$

As the convergence is exponential in t, the "asymptotic" form of $f(\underline{R}, t)$ is readily obtained. The QMC energy is now given by the average of the local energy over points sampled from $f(\underline{R}, t)$. That is,

$$<E_L> \equiv \lim_{M\to\infty} \frac{1}{M}\sum_{i=1}^{M} E_L(\underline{R}_i) = \int f(\underline{R}, t) E_L(\underline{R}) d\underline{R} \;/\; \int f(\underline{R}, t) d\underline{R}$$

$$= \int \phi_0(\underline{R}) H \Psi_T(\underline{R}) d\underline{R} \;/\; \int \phi_0(\underline{R}) \Psi_T(\underline{R}) d\underline{R} = E_0 \,, \quad (2.25)$$

where the last equality is deduced from the hermiticity of H, and the cancellation of the time-dependence requires that E_R not depend on $\underline{R}$. Since the statistical error in E_L vanishes when $\Psi_T = \phi_0$, the statistical error in averages of the local energy is greatly reduced (an order of magnitude or more), in comparison to those of the potential energy, for reasonable choices of Ψ_T.

2.4.3 The Short-Time Green's Function

The time evolution of the probability density function, $f(\underline{R}, t)$, is

$$f(\underline{R}', t+\tau) = \int d\underline{R} f(\underline{R}, t) G(\underline{R} \to \underline{R}', \tau) \,. \quad (2.26)$$

Here, as in Sec. 4.2.1, the Green's function defines a move from $\underline{R}$ to $\underline{R}'$ in time τ. This function is a solution of the imaginary-time Schrödinger equation (with importance sampling), Eq. (2.23), with the boundary condition $G(\underline{R} \to \underline{R}', 0) = \delta(\underline{R}' - \underline{R})$.

A "short-time" approximation to G is employed to simulate the diffusion, drift and branching of Eq. (2.23). It is obtained by assuming that the local energy and quantum force are constant during the course of a move from $\underline{R}$ to $\underline{R}'$. With these assumptions, valid at $\tau = 0$, this Green's function is given by

$$G_a(\underline{R} \to \underline{R}', \tau) = (2\pi\tau)^{-3N/2} \exp\left[-\left[\underline{R}' - \underline{R} - \frac{\tau}{2} F_Q(\underline{R})\right]^2 / 2\tau\right]$$

$$\times \exp\left[-\tau\left[(E_L(\underline{R}') + E_L(\underline{R}))/2 - E_R\right]\right] \,. \quad (2.27)$$

The first factor is the transition probability of moving from $\underline{R}$ to $\underline{R}'$ in time τ, and the

second arises from the branching and gives the weight, relative to that at $\underline{R}$, to be assigned to $\underline{R}'$. This approximate Green's function becomes exact as $\tau \rightarrow 0$[47-49] (except, perhaps, at the nuclei). For non-zero values of the time step, τ, the asymptotic form of $f(\underline{R}, t)$ approximates $\Psi_T \phi_0$, so that computed results will differ from the $\tau = 0$ limit. The difference between $\tau \neq 0$ and $\tau = 0$ values is referred to as "time-step bias". Therefore, computations must be performed at values of τ at which the bias is masked by the statistical error, i.e, is not significant, or results must be extrapolated to $\tau = 0$.

The Green's function in Eq. (2.27) is modified slightly in order to maintain detailed balance.[43] In this context, detailed balance means that the probability of moving from $\underline{R}$ to $\underline{R}'$ equals the probability of the reverse move. The modification involves employing a diffusion QMC Green's function, G_D,

$$G_D(\underline{R} \rightarrow \underline{R}', \tau) \equiv G_a(\underline{R} \rightarrow \underline{R}', \tau) A(\underline{R} \rightarrow \underline{R}', \tau) , \tag{2.28}$$

where

$$A(\underline{R} \rightarrow \underline{R}', \tau) \equiv \min(1, W(\underline{R} \rightarrow \underline{R}', \tau)) \tag{2.29}$$

with

$$W(\underline{R} \rightarrow \underline{R}', \tau) \equiv \frac{|\Psi_T(\underline{R}')|^2 G_a(\underline{R}' \rightarrow \underline{R}, \tau)}{|\Psi_T(\underline{R})|^2 G_a(\underline{R} \rightarrow \underline{R}', \tau)} . \tag{2.30}$$

The factor, A, incorporates an acceptance/rejection step into the algorithm. That is, after a move is completed, it is accepted with probability A. As $\tau \rightarrow 0, A \rightarrow 1$ and G_a becomes exact implying that G_D is also exact at $\tau = 0$.

Finally, when $\Psi_T = \phi_0$, G_D yields $f = |\Psi_T|^2$ for any value of τ. This results from the fact that G_D satisfies detailed balance for any value of τ. i.e.,

$$G_D(\underline{R} \rightarrow \underline{R}',\tau)|\Psi_T(\underline{R})|^2 = |\Psi_T(\underline{R}')|^2 G_D(\underline{R}' \rightarrow \underline{R},\tau) \ . \tag{2.31}$$

As illustrated in Ref. 26, when detailed balance is satisfied between a transition probability function, K, and a given probability distribution function, g, sampling from K (G_D for any τ) will yield convergence to g, ($|\Psi_T(\underline{R})|^2$). Generally, i.e., when Ψ_T only approximates ϕ_0, maintaining detailed balance reduces the time step bias. The use of G_D is discussed in the presentation of the diffusion QMC algorithm in Sec. 2.5.

2.4.4 Fermi Statistics: The "Fixed-Node" Approximation

As is well known, the eigenfunctions of chemical systems must change sign upon the interchange of electron coordinates since electrons are spin one-half (Fermi) particles. The QMC formalism described thus far, however, requires $f(\underline{R}, t) = \Psi_T(\underline{R})\Phi(\underline{R}, t)$ to be non-negative. If Ψ_T possesses the same nodes as does ϕ_0, then f will remain non-negative and the Fermi system may be treated immediately and exactly. For example, to obtain the exact energy of the $2p_x$ state of H atom simply requires Ψ_T to have a node at $x = 0$. In general, the symmetry of an atom or molecule does not completely specify the location of the nodes.[51] While techniques do exist for exactly sampling any Fermi state[35-36,40-42], they suffer from the instabilities discussed in the previous section.

The technique employed here for treating a Fermi system fixes the nodes of $\Phi(\underline{R}, t)$, and, therefore of ϕ_0, to be those of the trial function. That is, moves which cross these nodes are not allowed so that the simulations are confined within the nodal volumes of Ψ_T. Therefore, the Schrödinger equation is solved with the nodes of Ψ_T as boundary conditions, and statistical error does not diverge since signs are not carried.

When the assigned nodes are not those of the lowest-energy eigenfunction, the QMC solution is approximate. This approximation, however, has been found to be quite accurate in many QMC computations[42-44,52-58].

It is now shown that the fixed-node solution (still referred to here as ϕ_0) yields an upper bound to the ground-state energy. First, let the trial function $\Psi_T(\underline{R}, \underline{s})$ be antisymmetric in the electron coordinates $(\underline{r}, s)$. Let v_α be a nodal volume of Ψ_T, i.e., a volume enclosed by a nodal surface, and $\phi_\alpha(\underline{R}, \underline{s})$ the lowest-energy eigenfunction, with eigenvalue ε_α, obtained from fixed-node diffusion QMC (QMC from hereon). Then

$$H\phi_\alpha = \varepsilon_\alpha \phi_\alpha, \qquad \underline{R} \in v_\alpha$$

$$\phi_\alpha \Psi_T \geq 0 \,, \tag{2.32a}$$

and

$$\phi_\alpha = 0, \qquad \underline{R} \notin v_\alpha \,. \tag{2.32b}$$

The full anti-symmetric eigenfunction, $\hat{\phi}_\alpha(\underline{R}, \underline{s})$ is then obtained by summing over all permutations P,

$$\hat{\phi}_\alpha(\underline{R}, \underline{s}) = \sum_P (-)^P \hat{\phi}_\alpha(P\underline{R}, P\underline{s}) \,, \tag{2.33}$$

and from the variational principle

$$\frac{\int d\underline{R}\, \hat{\phi}_\alpha^* H \hat{\phi}_\alpha}{\int d\underline{R}\, \hat{\phi}_\alpha^* \hat{\phi}_\alpha} = \varepsilon_\alpha \geq E_0 \,. \tag{2.34}$$

As seen from above, different nodal volumes may yield different eigenvalues. Following the discussion which led to the asymptotic form of $f(\underline{R}, t)$, cf. Eq. (2.24), the same reasoning shows that for t sufficiently large, there results

$$f(\underline{R}, t) = c_0^{(m)} \exp[-t(\varepsilon_m - E_R)] \Psi_T \phi_0^{(m)} , \tag{2.35}$$

where m labels the nodal volume in which the QMC energy is a minimum. If the energy differences, $|\varepsilon_m - \varepsilon_\alpha|$, are sufficiently small, they may be masked by statistical error or lack of full convergence to the form of f given by Eq. (2.35). However, if all nodal volumes may be related by a symmetry operation which does not change the Hamiltonian, then the QMC energy obtained in each is the same, see Appendix A. Note that if Ψ_T possesses only the required anti-symmetry nodes then the nodal volumes are all related to each other by a permutation, P, of electron coordinates. Since $PH = H$, all the ε_α will be identical. This result may be generalized to operators other than H which are symmetric upon permutations of electron coordinates.

2.4.5 Variational Monte Carlo

In this section we describe the sampling of $|\Psi_T|^2$ using a guided Metropolis walk. This approach is generally referred to as variational Monte Carlo (VMC). (The reference to variational results from the fact that the variational quantity $E[\Psi_T]$ may be computed. However, the QMC energy is also variational for ground states, cf. Sec. 2.4.4.) As discussed in Sec. 2.4.4, the fixed-node energy is determined by the nodes of Ψ_T and the statistical error by fluctuations in the local energy, $\Psi_T^{-1} H \Psi_T$. This central role of the trial function makes it highly desirable to compute the expectation values

$$A_T \equiv \int d\underline{R}\, |\Psi_T|^2\, \Psi_T^{-1} A\, \Psi_T \;/\; \int d\underline{R}\, |\Psi_T|^2 . \tag{2.36}$$

(Specifically, $A = H$ yields $E[\Psi_T]$.)

The QMC simulations require only the gradients and Laplacians of the trial func-

tion. Therefore, Ψ_T may possess features, such as functions of electron-electron separation, which are desirable but prevent the analytic computation of A_T. However, in these instances expectation values may be obtained from averages over points sampled from $|\Psi_T|^2$, i.e.,

$$\lim_{m\to\infty} \frac{1}{M} \sum_{i=1}^{M} \Psi_T^{-1}(\underline{R}_i) A \Psi_T(\underline{R}_i) = A_T \tag{2.37}$$

for $\underline{R}_i$ sampled from $|\Psi_T|^2$. A useful method of sampling points from $|\Psi_T|^2$ is obtained by noting that $f = |\Psi_T|^2$ is an exact solution to the imaginary-time Schrödinger equation, (2.23), when the branching term is suppressed. Therefore, employing the Green's function of Eq. (2.28), but without the branching factor, allows the distribution $|\Psi_T|^2$ to be sampled correctly for arbitrary values of τ, cf. Sec. 2.4.4. Generally, τ is chosen such that the average probability of accepting a move is approximately 0.5.

2.5 QMC Algorithm

In this section, the QMC algorithm for converging to $\Psi_T \phi_0$ and computing the energy is discussed. The algorithm for performing the VMC computations is identical to that employed for QMC with the exception that branching is not simulated, cf. Sec. 2.4.

Prior to the Monte Carlo simulations, a trial function is chosen for importance sampling. As stated in the previous section, the ideal choice is $\Psi_T = \phi_0$ which, of course, is generally not possible. In practice, a trial function which accurately approximates the ground-state eigenfunction yet is relatively simple, so that its gradients and Lapla-

cians can be evaluated quickly, is desirable. The optimum mix of simplicity and accuracy is difficult to determine because QMC energies can not be predicted on the basis of a trial function's nodal structure. Instead different trial functions are employed and the resulting QMC energies yield, *a postieriori*, their quality.

Below the form of the trial function derived from the product of a single Slater determinant and correlation functions is presented. The extension to the case of several determinants readily follows.

A Slater determinant depends on both the spatial and spin coordinates of the electrons. For one with n_α spin-up and n_β spin-down electrons, a function which depends only on the spatial variables is obtained by assigning electrons 1 through n_α a spin of $\frac{1}{2}$ (up) and electrons $n_\alpha + 1$ through $n_\alpha + n_\beta$ a spin of $-\frac{1}{2}$ (down). The result is that the Slater determinant becomes a product of two determinants – one for the spin-up electrons and one for those with spin-down. This assignment of spins introduces no error for a spin-independent Hamiltonian.

The trial function, for N electrons, is now written as

$$\Psi_T(1, 2, \cdots, N) = \det[\psi_1^\alpha(1)\ \psi_2^\alpha(2)\ \cdots\ \psi_{n_\alpha}^\alpha(n_\alpha)]\cdot\det[\psi_1^\beta(n_\alpha+1)\ \psi_2^\beta(n_\alpha+2)\ \cdots\ \psi_{n_\beta}^\beta(n_\alpha+n_\beta)] \times F(1, 2, \cdots, N)\ . \tag{2.38}$$

An orbital subscript (α or β) denotes that the orbital is occupied by an electron of that spin. The correlation function F is introduced to give an improvement over the independent particle approximation. For the calculations performed in this thesis, F is given by

$$F(\underline{R}) = G(\{r_{ij}\})H(\{r_{i\alpha}\})\ . \tag{2.39}$$

The "electron-electron" correlation function, G, is usually chosen to have the form

$$G(\{r_{ij}\}) = \exp[\sum_{i<j} U_1(r_{ij})] \tag{2.40}$$

where several different forms of U_1 have been studied. The exponential form is useful because it is without nodes. The parameters in U_1 are generally chosen to minimize $E[\Psi_T]$, and to remove the Coulomb singularity in the local energy at $r_{ij} = 0$, i.e., to satisfy the electron-electron cusp condition. The "electron-nuclear" correlation function employed here is,

$$H(\{r_{i\alpha}\}) = \exp[\sum_{i,\alpha} U_2(r_{i\alpha})] \ . \tag{2.41}$$

The form of U_2 studied thus far is,

$$U_2(r_{i\alpha}) = -\lambda_\alpha r_{i\alpha}/(1 + \nu_\alpha r_{i\alpha}) \ . \tag{2.42}$$

The parameters, λ_α and ν_α are chosen to minimize $E[\Psi_T]$ and/or reduce or remove the singularity in the local energy at $r_{i\alpha} = 0$, the electron-nuclear cusp condition. As can seen from the equations above, the correlation functions are always positive, so that the nodes of Ψ_T are completely specified by the determinants.

In performing QMC (or VMC) computations, trial function ratios, the quantum force F_Q, and the local energy E_L must be evaluated. To do this efficiently, the inverses of the matrices corresponding to the determinants in Eq. (2.38) are computed. The result is that the quantities above can be readily obtained as scalar products [37]. Details are reserved for Appendix B.

The QMC algorithm employed here is very similar to that of Reynolds *et al*.[43] and is as follows:

(1) The initial probability density, $f(\underline{R}, 0)$, is chosen as $|\Psi_T(\underline{R})|^2$. This is accomplished by first randomly choosing a set of $N_0 \approx 100 - 300$ points, $\{\underline{R}_i\}$, and using

VMC to obtain N_0 new points sampled from $|\Psi_T(\underline{R})|^2$. (During this step "trial" expectation values, A_T, are computed.) Convergence to $|\Psi_T|^2$ is rapid since large time steps, τ, can be employed.

It is very important, however, that none of the initial random points be "too" close to a node of the trial function. This is because the acceptance probability, A, cf. Eqs. (2.29) and (2.30), becomes zero for a point located on a node of Ψ_T. The acceptance/rejection step is useful because it yields detailed balance in the simulation and guarantees convergence to $|\Psi_T|^2$ for arbitrary time steps. However, since the probability of moving to a node of Ψ_T is zero, then, by detailed balance, the probability of moving off a node is also zero. Therefore, each point (walker) is checked to see if it is moving through $\underline{R}$-space and discarded if it is not.

For an ensemble of points, $\{\underline{R}_i\}$, distributed according to $|\Psi_T|^2$, the next steps concern propagating these walkers in imaginary time, by employing G_D, cf. Eq. (2.28), until a new ensemble distributed as $\Psi_T\phi_0$ is obtained.

(2) Each point $\underline{R}_m$ in the ensemble corresponds, for N electrons, to a 3N-dimensional vector specifying the positions of the electrons, that is,

$$\underline{R}_m = (\underline{r}_1^{(m)}, \underline{r}_2^{(m)}, \cdots, \underline{r}_N^{(m)}) . \tag{2.43}$$

To move the first electron, a new set of coordinates, $\underline{r}_1'^{(m)}$, is chosen from the diffusion + drift factor of the Green's function,

$$\underline{r}_1'^{(m)} = \underline{r}_1^{(m)} + \sqrt{\tau}\underline{\chi} + \frac{1}{2}\tau F_{Q_1}(\underline{r}_1^{(m)}, \underline{r}_2^{(m)}, \cdots, \underline{r}_N^{(m)}) , \tag{2.44}$$

where $F_{Q_1} = 2\Psi_T^{-1}\nabla_1\Psi_T$. The second term of Eq. (2.44) simulates the diffusion, that is, $\underline{\chi}$ is a three-dimensional Gaussian random variable (obtained by the Box-Mueller

method[60]), and the last term simulates the drift. If a node is crossed, i.e., $\Psi_T(\underline{r}_1^{(m)}, \cdots)\,\Psi_T(\underline{r}_1'^{(m)}, \cdots) \leq 0$, then the nodal boundary conditions are enforced by either deleting this walker or rejecting ($\underline{r}_1'^{(m)} = \underline{r}_1^{(m)}$) this move. (The relative merits of "rejection" versus "deletion" are discussed in Chapter 3.)

(3) To incorporate detailed balance, the acceptance probability, A, is computed. If $A = 1$, the move is accepted, if $A < 1$, then the move is accepted with probability A. Algorithmically, this is implemented by obtaining a number I_A,

$$I_A \equiv \mathrm{int}(A + \xi) \tag{2.45}$$

where ξ is a uniform random variate between 0 and 1. The move is accepted if $I_A = 1$ and rejected if $I_A = 0$. It is easily seen, by integrating $I_A\, d\xi$ from 0 to 1, that the average value of I_A is A.

The quantity, I_A is stored in order to compute the "acceptance ratio" of the calculation. The acceptance ratio is defined as the number of moves accepted divided by the total number of moves attempted. This ratio gives an estimate of how close the simulation is to the $\tau = 0$ limit. That is, for the exact Green's function, G, it is readily deduced from the eigenfunction expansion of G, that W and, therefore, A are equal to one for all moves, cf., Eqs. (2.29) and (2.30). Since the short-time Green's function is exact at $\tau = 0$, then the acceptance ratio will converge to unity as the time step is made small.

(4) Steps (2) and (3) are performed for each of the remaining $N - 1$ electrons to complete the move $\underline{R}_m \rightarrow \underline{R}'_m$.

(5) The "multiplicity", M_m, is computed in order to simulate the branching where,

$$M_m \equiv \exp[-\tau_a \{[E_L(\underline{R}_m) + E_L(\underline{R}'_m)]/2 - E_R\}] \ . \tag{2.46}$$

The diffusion time in Eq. (2.46) is τ_a rather than τ because some of the electron moves may have been rejected. The mean-squared distance the electrons would diffuse in time τ, without rejection, corresponds to that of simple Brownian motion,

$$<r^2_{total}> = 3\tau \ , \tag{2.47}$$

and, in keeping with this correspondence, rejections yield an actual diffusion time given by

$$<r^2_{accepted}> = 3\tau_a \ . \tag{2.48}$$

Therefore, we have

$$\tau_a = \tau <r^2_{accepted}>/<r^2_{total}> \ . \tag{2.49}$$

Since each electron is moved individually, the local energies are computed as the sum of one-electron contributions. That is,

$$E_L(\underline{R}_m) = \sum_{i=1}^{N} E_{L_i}(\underline{r}_1'^{(m)}, \underline{r}_2'^{(m)}, \ \cdots \ , \underline{r}_{i-1}'^{(m)}, \underline{r}_i^{(m)} \ , \ \cdots \ , \underline{r}_N^{(m)}) \tag{2.50a}$$

and

$$E_L(\underline{R}'_m) = \sum_{i=1}^{N} E_{L_i}(\underline{r}_1'^{(m)}, \underline{r}_2'^{(m)}, \ \cdots \ , \underline{r}_i'^{(m)}, \underline{r}_{i+1}^{(m)} \ , \ \cdots \ , \underline{r}_N^{(m)}) \ . \tag{2.50b}$$

Also, for $\underline{R} \equiv (\underline{r}_1, \underline{r}_2, \ \cdots \ , \underline{r}_N)$, $E_{L_i}(\underline{R})$ is defined as

$$E_{L_i}(\underline{R}) \equiv -\frac{1}{2}\Psi_T^{-1}(\underline{R})\nabla_i^2\Psi_T(\underline{R}) - \sum_{\alpha} r_{i\alpha}^{-1} + \frac{1}{2}\sum_{j\neq i} r_{ij}^{-1} + N^{-1}\sum_{\alpha,\beta} Z_\alpha Z_\beta r_{\alpha\beta}^{-1} \ . \tag{2.51}$$

The multiplicity is computed from the local energies and the actual diffusion time. The branching may then be simulated in either of two ways, or as a combination of both. The first, referred to as "integer rounding", signifies that I_m copies of walker m are created at $\underline{R}'_m$ where

$$I_m \equiv \text{int}[M_m + \xi], \tag{2.52}$$

and ξ is once again a uniform random variate between 0 and 1. The second approach, "weighting", simply uses M_m to define a weight, $w(\underline{R}'_m)$. For the walker at $\underline{R}'_m$,

$$w(\underline{R}'_m) \equiv M_m w(\underline{R}_m) \ . \tag{2.53}$$

The benefit of weighting is that it is always exact rather than exact only on average (with integer rounding). However, if only weighting is employed, the sampling will become inefficient because the weights diverge towards 0 or ∞ as the simulation proceeds. This problem is not present with integer rounding because each walker has a weight of unity (though more than one walker may be present at a point in $\underline{R}$-space). A useful combination of weighting and integer rounding is to employ weighting with the constraint that the weights remain between an upper ($w_{\max}$) and lower ($w_{\min}$) bound. That is, if $w(\underline{R}) \geq w_{\max}$ at a point $\underline{R}$, $I_w \equiv \text{int}[w(\underline{R}) + \xi]$ walkers are created at $\underline{R}$ and each is assigned new weight of $w(\underline{R})/I_w$. If $w(\underline{R}) \leq w_{\min} < 1$, integer rounding is employed by assigning this walker a new weight of $\text{int}[w + \xi]$. The usefulness of this method of simulating the branching is discussed in subsequent chapters.

Expectation values are obtained from the following averages,

$$\bar{A}_M = \sum_{m=1}^{M} I_m A(\underline{R}_m) / \sum_{i=1}^{M} I_m \qquad \text{(integer rounding)} \tag{2.54a}$$

or

$$\bar{A}_M = \sum_{m=1}^{M} w(\underline{R}_m) A(\underline{R}_m) / \sum_{i=1}^{M} w(\underline{R}_m) \ , \quad \text{(weighting)} \tag{2.54b}$$

where A is a coordinate operator. (For example, to compute the energy, $A(\underline{R}_m) = E_L(\underline{R}_m)$.)

(6) Repeat steps (2)-(5) for all walkers in the ensemble.

(7) Repeat steps (2)-(6) for N_τ times, where $t_b \equiv N_\tau \tau$, and the "block" time (t_b) is generally chosen to be several inverse hartrees. At this point a block is completed and averages given by Eq. (2.54a or b) are computed.

Block averages are useful because, given the small size of τ, there exists a high degree of correlation between successive moves. Therefore, the QMC calculation is divided into blocks and final averages are obtained from block values. If the blocks are sufficiently large, i.e. t_b and/or N_0 are large, then the correlation between block values is small and the statistical error from their average will be accurate.

(8) If the reference energy, E_R, is a poor approximation to E_0, it may be updated, to minimize the dependence of $f(\underline{R}, t)$ on t, according to

$$E_R(\text{new}) = \frac{1}{2}[E_R(\text{old}) + E_b] \,, \tag{2.55}$$

(or some other combination) where E_b, a "block energy", corresponds to the average in Eq. (2.54) with $A = E_L$, and M is the number of points sampled in the block.

Note from Eq. (2.55) that updating E_R in this way introduces a weak dependence on the values of $\underline{R}$ sampled in the block, since $E_b = E_b(\{\underline{R}_m\}, t)$. It is assumed, however, when obtaining E_0 from the average of the local energy, that the reference energy does not depend on $\underline{R}$. In the limit that every point in $\underline{R}$-space is sampled in the block, E_b and E_R depend only on t. Therefore, for sufficiently large blocks it may be assumed that E_R depends only on the imaginary time. In practice, the reference energy is <u>not</u> updated when the QMC energy, for the value of τ employed, can be estimated to within $\simeq 0.01h$ as is often the case.

(9) Before proceeding to the next block, the number of walkers, which may now be

different than N_0 due to branching, is "renormalized" to N_0. This step is necessary because the fluctuations in the ensemble size increase as the simulation proceeds. Renormalization introduces a bias[40] since adding or removing a walker at $\underline{R}_m \in \{ \underline{R}_i \}$, to obtain N_0 walkers, is not arbitrary for blocks of finite length. Therefore, block lengths (t_b), are chosen as large as possible, to minimize this (generally not visible) bias, yet small enough so that the ensemble size will not fluctuate to zero or beyond the allocated memory.

(10) Repeat steps (2)-(9) until the block energies have converged, implying that $f(\underline{R}, t) \propto \Psi_T(\underline{R})\phi_0(\underline{R})$.

(11) Reset all averages to zero. Then repeat steps (2)-(9) until enough blocks, N_b, are obtained so that statistical errors in each computed quantity, $\bar{A}$,

$$\bar{A} \equiv \frac{1}{N_b}\sum_{i=1}^{N_b} A_b \ , \tag{2.56}$$

are small.

Chapter 3

The Computation of QMC Energies

3.1 Introduction

In this chapter computations of QMC energies are presented and discussed for a variety of atoms and molecules. These computations allow several important facets of fixed-node diffusion QMC to be studied.

As mentioned in Chapter 2, a short-time approximation to the Green's function is employed. In principle this approximation is easily removed by computing over a range of time steps which yields a reliable estimate of $\tau = 0$ results. The required range of time steps depends on the behavior of the computed quantity versus the time step. If time-step bias is sufficiently large, or is not consistent, then obtaining unbiased estimates, from extrapolation or otherwise, will necessitate computations at small time steps. However, efficiency is generally inversely proportional to τ, due to the increase in correlation between moves, so that small–τ computations are costly. Therefore, it is of interest to gain an understanding of this bias. Towards this end, time-step, as well as other, effects are studied in detail for several well-understood systems in the next section.

It is also of interest to examine the error introduced by forcing the nodes of the QMC solution to be those of the trial function – the fixed-node approximation. As

proven in Chapter 2, this approximation yields a variationally bounded energy for lowest-energy states of a given symmetry. Also, as the nodes become exact and time-step bias is removed, QMC results do as well. Beyond these facts, however, the parameterization of the fixed-node error versus nodal structure is unknown. This is due, in part, to the complex nature of the nodal (hyper-)surfaces of anti-symmetric functions. Therefore, the accuracy of a *trial* function's nodal structure, in terms of the quality of the QMC energy or other computed quantities, must be ascertained directly by computation if previous experience is lacking.

In studying the quality of fixed-node energies, several energy computations have been performed and are reported in the remaining sections. For trial functions with nodal surfaces simple enough to be drawn, nodes have been plotted and compared with computed energies. When this is not the case, conclusions of a general nature concerning the effects of one- and many-particle basis sets may still be made.

3.2 Time-Step Bias in Diffusion QMC Calculations

The magnitude of the time-step bias is considered for all the atoms and molecules for which we have performed QMC computations. Since the energy generally possesses the smallest relative statistical error, time-step bias is often most visible in this quantity. For few-electron atoms and molecules, time-step and other biases may be studied in greatest detail because of the smaller computational cost incurred. The results of QMC computations on several small systems are presented and discussed below.

The first effect studied here concerns the acceptance/rejection (AR) step. As

discussed in Chap. 2, this step maintains detailed balance and is expected to reduce time-step bias. The model system studied is the ground state of the H atom with $\Psi_T = \exp(-\zeta r)$, $\zeta = 1.05$. The results presented in Figs. 3.1 and 3.2 demonstrate the importance of AR in reducing time-step bias. Time-step bias is increased by an order of magnitude in the energy (cf. Fig. 3.1), and in r_M and r^2_M (cf. Fig. 3.2, where $A_M = \langle\Psi_T|A|\phi\rangle/\langle\Psi_T|A|\phi\rangle$), for most values of the time step (τ) when AR is omitted. Consequently, the value of τ at which accurate results may be obtained is approximately an order of magnitude smaller, increasing computational cost by this amount, without the AR step. Given the size of the effect found here, the use of AR is highly desirable in computing diffusion QMC energies of Coulombic systems.

When the trial function is exact, no time-step bias is present, and therefore, increasing trial function accuracy should reduce this bias. This effect is demonstrated in Figs. 3.3 and 3.4 in which QMC energies for the ground state of H are plotted for several values of ζ, $\Psi_T = \exp(-\zeta r)$. While the amount of bias is roughly independent of the sign of the error in ζ, compare the $\zeta = 0.95$ and $\zeta = 1.05$ (exact $\zeta = 1$) energies in Fig. 3.3, decreasing ζ to 0.90 leads to a noticeable increase in the size of the time-step bias. Finally, Fig. 3.4 shows that very poor choices of Ψ_T lead to great difficulty in computing the energy accurately. Near-statistical agreement with the exact energy is obtained only for the very small time step of 0.0005 h^{-1} when $\zeta = 0.50$. Furthermore, the practically monotonic behavior of the QMC energies plotted in Fig. 3.3, $\zeta = 0.90$, 0.95 and 1.05, is contrasted by the fluctuating $\zeta = 0.50$ energies. Such fluctuations greatly increase the uncertainty in extrapolations to $\tau = 0$.

In addition to the trial function, another consideration is the enforcement of the

nodal boundaries. Several aspects of this are studied for the $2p_x$ state of H; Figs. 3.5-3.7 present results.

The first effect examined here is that arising from the finite probability of a walker crossing and recrossing a node during the diffusion step of a move. To remove this source of bias a walker should be deleted (or the move rejected) with the probability, which is non-zero for $\tau > 0$, that a node cross-recross (CRC) occurred. For the $2p_x$ state, this probability may be computed exactly[46] since the node is known to be at $x = 0$, ("NO CRC" in Figs. 3.5-3.7). Alternately, and more generally applicable, the location of the node may be approximated using the gradient of the trial function, and, assuming the node to be a plane perpendicular to the gradient, the CRC probability may be estimated, "∇ CRC".

Fig. 3.5 presents energies computed with and without CRC bias and its approximate treatment using $\nabla\Psi_T$, $\Psi_T = x\exp(-0.55r)$ (the exact exponent is 0.5). A node crossing is treated by deleting the walker. It is notable that removal of CRC bias actually *increases* time-step bias, as seen by comparing the NO CRC with the standard walk (CRC) results. Apparently the CRC bias cancels another so that its removal worsens computed energies. This is discussed later when rejection of moves crossing a node is considered. Approximately removing CRC bias with the ∇ CRC technique is found to slightly increase bias with respect to the standard walk energies, probably to the extent that cross-recrossings are removed.

The remaining set of calculations in Fig. 3.5 neglect CRC bias. Instead the "ITERATIVE" technique of Anderson[59] is employed to reduce time-step bias. This approach improves the short-time Green's function through the use of a more

appropriate quantum force, F_Q in Eq. (2.23), for each move. After sampling a point $\underline{R}'$ from $G_a(\underline{R} \rightarrow \underline{R}', \tau)$, $\underline{R}$ is moved to $\underline{R}''$ with a quantum force of

$$F_Q = \frac{1}{2}[F_Q(\underline{R}) + F_Q(\underline{R}')] \ .$$

The extra computation resulting from the increased complexity of the moves will be justified if time-step bias is reduced by about a factor of two. The energies in Fig. 3.5 show that this iterative technique does reduce bias, in comparison to the standard walk, but by no more, if not less, than a factor of two.

The next consideration is whether to establish nodal boundary conditions by either deleting a walker or rejecting a move upon crossing a node. For non-zero time steps the two algorithms will yield different densities near the nodes and, therefore, different results. A good example demonstrating the differences between deletion and rejection is that of a particle in an infinitely deep well. With rejection, choosing the trial function and reference energy exactly for the ground state yields a stable algorithm and correct results. However, if walkers are deleted upon crossing the boundaries of the well, a population decaying in time and a biased growth energy are obtained. Other quantities, such as the density of the walkers near the boundaries, could now be inaccurate as well.

Energies computed with the deletion approaches possessing the least amount of time-step bias, the CRC and iterative walks in Fig. 3.5, are contrasted in Fig. 3.6 with rejection (upon crossing $x = 0$) walks which do and do not possess CRC bias. The trial function is that of Fig. 3.5. Comparing standard diffusion QMC (CRC) deletion and rejection energies shows a large reduction in time-step bias when rejection is employed. Energies resulting from rejecting moves yields improvement over the

iterative (deletion) approach as well. Also, removal of CRC has a much less pronounced effect with rejection.

The worsening of the energy by removing CRC bias when deleting walkers is now readily explained. The QMC energy is computed from the average of the local energy, E_L, over $\Psi_T \phi$ (cf. Eq. 2.25), where

$$E_L(\underline{r}) = -\frac{1}{2}\zeta^2 + \frac{2\zeta-1}{r} . \tag{3.1}$$

If $\zeta > 0.5$, the local energy tends to be higher near the node, i.e., smaller values of $|x|$ tend to give smaller values of r and higher local energies. Therefore, too low a density near $x = 0$ yields an energy biased below the correct result – as occurs with deletion. Removing CRC bias, and thereby deleting more walkers near $x = 0$, further increases time-step bias as observed in Fig. 3.5. Since rejection energies are much more accurate, the description of the density near the node is better with rejection versus deletion. A further verification of this fact is found by comparing rejection and deletion results for averages of $3x^2-r^2$ over $\Psi_T \phi$. For $0.05 \leq \tau \leq 1.00\ h^{-1}$, rejecting moves gives accurate values while deleting walkers yields values which are far too high.

The effects discussed above are finally considered for $\zeta = 0.45$, i.e., ζ below the exact, and results are presented in Fig. 3.7. As implied by Eq. (3.1), deletion energies are now biased above the exact, and once again rejection yields far better energies than deletion.

Due to the results presented thus far, rejection is employed in QMC calculations subsequent to the discovery of the reduced time-step bias discussed above. Also, given the unsatisfactory results of the ∇ CRC and iterative techniques, they are not

implemented in other QMC computations.

The last, and perhaps most important, idea for reducing time-step bias concerns the choice of the trial function, which is used for importance sampling and setting the nodal boundary conditions.

As stated in Chap. 2, the trial function is a product of an antisymmetric function and an everywhere positive correlation function, F, cf. Eq. (2.38). That is,

$$\Psi_T(\underline{R}) = \mathrm{Det}^{\alpha}\{\psi_i(\underline{R}_i)\}\mathrm{Det}^{\beta}\{\psi_j(\underline{R}_j)\}F(\underline{R}) , \tag{3.2}$$

where $\{\psi_i\}$ and $\{\psi_j\}$ are the molecular orbitals (MOs) occupied by spin-up (α) and spin-down (β) electrons, respectively. For example, a single-determinant trial function describing the ground state of Be takes the form

$$\Psi_T(\underline{R}) = \begin{vmatrix} \psi_1(\underline{r}_1) & \psi_1(\underline{r}_2) \\ \psi_2(\underline{r}_1) & \psi_2(\underline{r}_2) \end{vmatrix} \begin{vmatrix} \psi_1(\underline{r}_3) & \psi_1(\underline{r}_4) \\ \psi_2(\underline{r}_3) & \psi_2(\underline{r}_4) \end{vmatrix} F(\underline{R}) , \tag{3.3}$$

where $\underline{R} = (\underline{r}_1, \underline{r}_2, \underline{r}_3, \underline{r}_4)$. A multi-determinant trial function is, by analogy with Eq. (3.2), a sum of products of "spin-up" multiplied by "spin-down" determinants.

The correlation functions employed here are given by

$$\begin{aligned} F(\underline{R}) &= G(\{r_{ij}\})H(\{r_{i\alpha}\}) \\ &= \exp\Big[\sum_{i<j} U_1(r_{ij})\Big]\exp\Big[\sum_{i,\alpha} U_2(r_{i\alpha})\Big] , \end{aligned} \tag{3.4}$$

where r_{ij} and $r_{i\alpha}$ denote electron-electron (E-E) and electron-nuclear (E-N) separations, respectively. Parameters in U_1, for the forms we have considered, may readily be chosen to satisfy the electron-electron cusp conditions. The electron-nuclear function, U_2,

$$U_2(r_{i\alpha}) = -\lambda_\alpha/(1 + \nu_\alpha r_{i\alpha}) , \tag{3.5}$$

combines with the antisymmetric component of Ψ_T in determining the behavior of the

local energy as an electron approaches a nucleus. For a nucleus with charge, Z_α, the $r_{i\alpha}^{-1}$ singularity in E_L is removed when

$$[r_{i\alpha}E_L(\underline{R})]_{r_{i\alpha}=0} = [-\frac{1}{2}r_{i\alpha}\Psi_T(\underline{R})\nabla_i^2\Psi_T(\underline{R})]_{r_{i\alpha}=0} - Z_\alpha = 0. \tag{3.6}$$

Abbreviating $Det^s\{\psi_i(\underline{R}_i)\}$ as D^s, and employing Ψ_T and U_2 as given by Eqs. (3.2) and (3.5), respectively, yields for the E-N cusp condition,

$$-\frac{1}{2}[r_{i\alpha}(D^s)^{-1}\nabla_i^2 D^s]_{r_{i\alpha}=0} + \lambda_\alpha = Z_\alpha \,. \tag{3.7}$$

For a Slater-type orbital (STO) basis set, the first term in Eq. (3.7) is generally non-zero and depends on the positions of the spin-s electrons not on the nucleus. Therefore, Eq. (3.7) may only be solved approximately if two or more electrons possess the same spin.

We have found that satisfying the E-N cusp condition, exactly when possible or approximately otherwise, reduces time-step bias. The importance of satisfying the E-N cusp condition, for purposes of reducing time-step bias, derives from the Green's function, G_D [cf. Eqs. (2.28)-(2.30)], we employ in diffusion QMC. Since this Green's function yields no time-step bias when Ψ_T is exact, and E_L is a constant, reducing fluctuations in the local energy should lead to a decrease in bias. Insofar as removing singularities in E_L is an important step towards achieving a constant local energy, satisfying the E-N cusp condition should yield decreases in bias in comparison to energies computed otherwise.

For systems with two or more electrons of a given spin, it is useful to determine the behavior of the first term in Eq. (3.7), T_1, in order to choose the value of λ_α which best satisfies the E-N cusp condition. Here, this is done by sampling several thousand values of $\underline{R}$ from $|\Psi_T|^2$. For a set of coordinates $\underline{R}$, if a value of $r_{i\alpha}$ is sufficiently

small, T_1 is computed and a value of λ_α is tabulated according to $\lambda_\alpha = Z_\alpha - T_1$. We have found, for several atoms and molecules described by single-determinant trial functions with STO basis sets, that the distribution of these values of λ_α is tightly clustered about its mode, "λ_α(cusp)". An example is given in Fig. 3.8. Therefore, with $\lambda_\alpha = \lambda_\alpha$(cusp) we expect that the E-N cusp condition(s) are approximately, yet optimally, satisfied.

In Figs. 3.9-3.11 QMC energies are computed with and without satisfying the E-N cusp conditions for He, H_2, and LiH, respectively. The two trial functions employed for each system possess the same nodal structures and approximately identical energy expectation values, $E[\Psi_T] = \langle\Psi_T|H|\Psi_T\rangle$. For the first two systems, with only one electron of each spin, the E-N cusp conditions may be satisfied exactly with the appropriate choice of λ. For LiH, this condition is obeyed approximately employing computed values of λ_H(cusp) and λ_{Li}(cusp).

Figs. 3.9-3.11 demonstrate the dramatic reductions of time-step bias in the QMC energy obtained with values of λ_α optimized to remove the E-N singularities, "NO CUSP", in comparison to energies computed with values of λ_α hand-optimized to minimize $E[\Psi_T]$, "CUSP". The statistical error in each QMC energy is 0.00085 h (0.5 kcal/mol) or less. For each system NO CUSP energies are in statistical agreement with the exact energy, or for LiH the extrapolated to $\tau = 0$ value, at a time step an order of magnitude larger than when CUSP energies obtain similar agreement. As with the effects of rejection versus deletion upon crossing a node, the significance of satisfying the E-N cusp condition was discovered after many of the calculations reported below.

3.3 Computations of Energy Differences

The fixed-node diffusion QMC approach is a relatively new *ab initio* technique. Initially, energies of good or excellent quality were obtained for H_2, LiH, Li_2, and H_2O employing simple, i.e., single-determinant, small basis set, trial functions.[43] This success raises the question of the capability of QMC to compute accurately energy differences such as classical barrier heights, binding energies, level splittings, electron affinities, and so on.

The first such application was the study of the singlet-triplet energy difference, T_0, of methylene.[52] This system was of interest due to previous experimental discrepancies, 8-10[61] versus 19.5[62] kcal/mol, which have since been resolved, $T_0 \sim 9$ kcal/mol.[63] The QMC energy of each state, computed with single-determinant, double zeta (DZ) basis set, trial functions, are below any previous *ab initio* values and also very accurate, 98% and 96% of the correlation energy (CE) obtained for the singlet and triplet state, respectively. Furthermore, the QMC-computed T_0, 9.4(2.2) kcal/mol, is in good accord with accepted experimental results. (The large statistical error arises from the fact that these early QMC calculations were performed on a VAX 780.)

The work described in this section continues the computation of energy differences. The systems studied provide a proving ground of the QMC approach. In Sec. 3.3.1 the classical barrier height to the H + H_2 exchange reaction is discussed. Sec. 3.3.2 presents QMC energies of F and F^- from which an accurate electron affinity is obtained. Sec. 3.3.3 concludes Sec. 3.3 with a calculation of the binding energy of

N_2.

3.3.1 The Barrier to H + H_2 Exchange

The computation of the classical barrier to the H + $H_2 \rightarrow H_2$ + H exchange reaction is often used as a test case for theoretical methods of determining potential energy surfaces (PESs). Indeed, there have been many semi-empirical and *ab initio* calculations of the PES of three hydrogen atoms, and, most importantly, the barrier to H-atom exchange. A review of such efforts is given by Truhlar and Wyatt.[64] The large configuration interaction (CI) calculations of Liu[65] and Siegbahn and Liu[66] have made the H_3 PES the most accurately determined surface to date. Most recently Liu[65] has reported an improved saddle-point energy and a barrier estimated to be accurate to 0.1 kcal/mol. Finally, a released-node Green's function Monte Carlo (GFMC) calculation by Ceperley and Alder[41] has yielded a barrier height which is exact to within statistical error (0.08 kcal/mol). Given the high accuracy of previous work and the relative simplicity of a three electron system, the barrier to this H-atom exchange reaction serves as an excellent test case for diffusion QMC with fixed nodes.

All our QMC computations on H_3 employ deletion upon crossing a node. Since single-determinant trial functions yielded accurate energies in previous QMC calculations[43,52], such trial functions are also used here. Several basis sets, however, are utilized to examine their effect on the QMC energy. These basis sets are given in Table 3.1, and range in size from single zeta (SZ) to 4Z with four polarization functions. A basis set with a function between each nuclei is also employed. The MO coefficients (not shown) are readily obtained using standard molecular codes, e.g.,

HONDO.[68]

In the MO picture, the ground state of H_3 consists of two electrons in the lowest-energy MO, σ_g symmetry, and one in the next MO, σ_u symmetry. The QMC approach, in assigning a spin to each electron, reduces the full 3×3 Slater determinant to a spin-free product of a 2×2 determinant and a single MO. Including the correlation functions yields,

$$\Psi_T(\underline{r}_1, \underline{r}_2, \underline{r}_3) = \begin{vmatrix} \psi_1(\underline{r}_1) & \psi_1(\underline{r}_2) \\ \psi_2(\underline{r}_1) & \psi_2(\underline{r}_2) \end{vmatrix} \psi_1(\underline{r}_3)\exp[\sum_{i<j} U_1(r_{ij})]\exp[\sum_{i\alpha} U_2 r_{i\alpha}] \ . \quad (3.8)$$

For the electron-electron (E-E) correlation function,

$$U_1(r) = (a_1 r + a_2 r^2)/(1 + b_1 r + b_2 r^2). \quad (3.9)$$

The value of a_1 is chosen to satisfy the E-E cusp condition for electrons of opposite spin. (The like-spin cusp condition, satisfied when $a_1 = 0.25$[43], is not important since Ψ_T vanishes when such electrons come together.) The remaining parameters are optimized to yield a minimum in the VMC energy. Here, this optimization was performed by hand, i.e., by comparing VMC energies between several sets of parameter values, and is therefore not complete.

The E-E correlation function is monotonically increasing (for the parameters chosen here, cf. Table 3.2) which pushes electrons further apart thereby expanding the density. Thus the usefulness of describing instantaneous electron correlations is partially negated by creating a density more diffuse than the SCF density which is qualitatively correct on average. The E-N correlation function pulls the electrons back towards the nuclei to improve the density, and also further lowers the VMC energy. The function U_2 is given by Eq. (3.5). The parameters λ and ν are also partially optimized. All

correlation function parameters are presented in Table 3.2.

Classical barrier heights, computed with the trial functions described in Tables 3.1 and 3.2, are presented in Table 3.3. These barriers, in kcal/mol, are computed as the difference between the H_3 saddle-point energy and the exact ground state energies of H and H_2.[10] The exact energies of H and H_2 are employed because the QMC energies are exact for these nodeless systems. The QMC barriers, therefore, are statistical upper bounds to the exact. The saddle-point geometry, collinear with adjacent atoms separated by 1.757 bohr, is taken from Liu's large CI calculation.[65]

The basis sets range from minimum size (SZ), Ψ_1, to near-Hartree Fock (HF) quality, Ψ_6. A spin-unrestricted HF ("UHF") trial function, Ψ_2, is also included and yields an SCF barrier height below that of the HF limit. The addition of the correlation functions is seen to significantly improve trial function quality. In general, 40-65% of the CE is gained and the accuracy in the barrier height is improved by about a factor of two in comparison to the SCF results, see E_B (VMC) in Table 3.3.

In computing the QMC energy, the effect of the QMC parameters selected for the simulation, time-step, number of points in the ensemble, block times, etc., must be considered. (The block time is defined as the number of steps each walker is moved in each block multiplied by the time step.) Besides time-step bias, other effects may be present. As discussed in Chap. 2, the number of points in the ensemble is reset to its initial value at the end of each block by deleting or adding points (renormalization). Renormalization introduces a small bias which decreases as the frequency of renormalization decreases.[40] In addition to this bias, the updating of the reference energy, E_R, has a small effect. This effect is due to the fact that updating E_R by averaging in

the current block energy, cf. Eq. (2.55), yields a correlation between E_R and the sampling of configuration space. This correlation decreases as the ensemble size and block time increase.

In Table 3.3, an initial set of QMC calculations (fourth column) is compared with a final set (last column) demonstrating the effect of QMC parameters on the computed energy. In the initial QMC calculations the time steps employed were quite small, 0.00125 to 0.00500 h^{-1}, and no time-step dependence is observed. The QMC barrier heights, however, are significantly above the estimated exact value of 9.6 kcal/mol. This poor accuracy was found to be due to short block times. In the first three computations in column 4 of Table 3.3, the block time is 0.5 h^{-1} and 1.0 h^{-1} for the last. Also the reference energy is updated at the end of every block. For the last set of QMC computations, all block times are increased to 5 h^{-1} and E_R is not updated but fixed throughout the entire simulation. This reduction of renormalization error, and elimination of correlation of between E_R and block energies, yields a significant improvement in the QMC-computed barrier heights. Therefore, a general practice of setting block times to several inverse hartrees and using a constant reference energy is strongly recommended.

The final QMC barriers, computed at $\tau = 0.005\ h^{-1}$, demonstrate the improvements over the VMC energy obtained by the QMC approach. Even the very simple SZ trial function recovers over 90% of the CE. Trial functions $\Psi_3 - \Psi_6$ yield QMC energies which obtain over 99%, resulting in barriers within 0.1 to 0.3 kcal/mol of Liu's best (unbounded) estimate and Ceperley and Alder's released-node result which possesses only statistical error. Finally, note that our most accurate and precise barrier, obtained

from Ψ_5, is in excellent agreement with the essentially exact values quoted above, cf. Table 3.4.

The basis-set dependence of the QMC energy, in contrast to more standard *ab initio* methods, is very weak. All QMC barriers are within 2.5 kcal/mol of each other. Eliminating Ψ_1 as too simplistic reduces differences in QMC barrier heights to only 0.5 kcal/mol for basis sets ranging in size from two functions on each atom to eight. Variation in the computed barriers is further reduced to only 0.2 kcal/mol (roughly the size of the statistical errors) by removing the UHF trial function energy.

The UHF trial function is now considered as a special case. The SCF barrier height of this trial function, Ψ_2, is below those of the remaining spin-restricted ("RHF") trial functions. The QMC barriers are much more similar, and that of Ψ_2 is actually the least accurate with the exception of the Ψ_1 barrier height. This phenomenon relates to the fact that the distinction between UHF and RHF trial functions is less significant in the context of QMC calculations on the three-electron system studied here.[69] The QMC energy is determined by the nodal structure which, ignoring the positive correlation functions, is given for Ψ_2 by Φ,

$$\Phi(\underline{r}_1, \underline{r}_2, \underline{r}_3) = \psi_1(\underline{r}_3) \begin{vmatrix} \psi_1'(\underline{r}_1) & \psi_1'(\underline{r}_2) \\ \psi_2(\underline{r}_1) & \psi_2(\underline{r}_2) \end{vmatrix} . \tag{3.10}$$

For the ground state of H_3, ψ_1 and ψ'_1 correspond to the lowest-energy MO and therefore should be without nodes. Consequently, an RHF analog of Ψ_2 which possesses the same nodal structure, and thus identical QMC energy, is readily obtained by replacing ψ_1 by ψ_1' in Eq. (3.11). To the extent that the linear coefficients in the 2×2 determinant of the RHF analog are not fully optimized, in comparison to those result-

ing from a spin-restricted SCF calculation, a UHF trial function may actually yield a higher QMC energy. For Ψ_2, ψ_1 was also found to possess an ellipitically shaped node far from the nuclei, cf. Fig. 3.15, which *will* increase the QMC energy. Therefore, it is not surprising that the QMC barrier of Ψ_2 is among the least accurate – despite having the lowest SCF energy.

As noted in the discussion of the QMC barriers in Table 3.3, the QMC energies resulting from $\Psi_3 - \Psi_6$ show little variation. Since the QMC energy is determined by the nodal structure of the trial function employed, the most direct explanation of the QMC barrier heights lies in an examination of the trial function nodal structures.

In each of Figs. 3.12-3.14, curves representing nodes of the 2×2 determinant are plotted for Ψ_2, Ψ_4 and Ψ_5, respectively. Full three-dimensional nodal surfaces are obtained by rotating the curves about the internuclear axis, and the trial function vanishes when the two like-spin electrons lie anywhere on it (see Appendix C). As seen from the plots, the highly curved nodal surfaces of Ψ_1 are unique, while those of Ψ_4 and Ψ_5, as well as those of the remaining trial functions (not shown), are practically indistinguishable. The dissimilarity of the SZ surface and the similarity among the nodal surfaces of $\Psi_2 - \Psi_6$ thus explains the resulting QMC barriers.

The remaining question concerns the variation in the QMC barriers of $\Psi_2 - \Psi_6$. The differences among the QMC results of $\Psi_3 - \Psi_6$ are small enough to be caused by statistical error, especially for Ψ_3–Ψ_5, or the slight differences in the nodes of the 2×2 determinant. It is important to note, however, that the ground-state MO of Ψ_2, Ψ_5 and Ψ_6 possess a node far from the nuclei, as shown in Fig. 3.15. (For Ψ_2 this node is in the MO of the 1×1 determinant only.) Therefore, Ψ_5 and Ψ_6 possess, in addition to the

exchange nodes, a node when like-spin electrons are on a node of ψ_1 or when the unlike-spin electron is on this node. The additional node in Ψ_2 concerns only the unlike-spin electron. The additional node in Ψ_5 remains over 20 bohr from the nuclei and should have little, if any, effect on the energy. On the other hand, the node in ψ_1 is much closer in for Ψ_2 and Ψ_6 which may account for barrier heights above Liu's estimate by more than statistical error.

The additional node and slightly higher barrier height of Ψ_6 show that nodal structure does not necessarily improve simply because a HF-limit basis set is approached. This is reasonable since SCF calculations preferentially optimize the wave function where it is large, while QMC calculations require accuracy in the nodes where the exact solution is zero. Thus the accuracy of the SCF-determined nodal structure may be non-monotonic or relatively constant, as seen here, as the basis set is increased.

In addition to an accurate energy, a trial function is desired which does not incur large computational cost in obtaining a sufficiently small statistical error. While a more complex trial function may yield a smaller statistical error for a given number of sampled points, it may still be less efficient than a simpler trial function due to the increased effort of computing quantities necessary for the simulation, e.g., F_Q and E_L.

In Table 3.5 the relative efficiencies of the trial functions are shown. The relative efficiency of Ψ_i, is a measure of how quickly, in comparison to Ψ_1, the statistical error in the QMC energy is reduced to a given value. The efficiency of the UHF trial function is seen to be the lowest, further obviating its usefulness. Most notable is that the efficiency of Ψ_6 is only slightly higher than that of Ψ_1, and lower than those of $\Psi_3 - \Psi_5$. Therefore, expansion of the basis set to near-HF quality actually reduces

efficiency (and, as discussed above, also possibly the accuracy).

Thus far a detailed discussion of time-step bias has been omitted. As stated about the "short block time" barrier heights, τ–dependence of the QMC energies was not observed for $0.00125 \leq \tau \leq 0.00500 \ h^{-1}$. We now extend the time-step range to $[0.00125, 0.10] \ h^{-1}$ in computing QMC energies with Ψ_4; results are plotted in Fig. 3.16. The barrier heights plotted in Fig. 3.16 are all in statistical agreement with each other demonstrating that even at $\tau = 0.1 \ h^{-1}$ time-step bias is beneath the statistical error of 0.25 kcal/mol. Furthermore, the barrier computed at $0.005 \ h^{-1}$, reported in Table 3.3, is in agreement with that obtained by extrapolation, 9.86 (0.20) versus 9.86 (0.22) kcal/mol. This agreement implies that the barriers in Table 3.3 of the remaining trial functions, computed at $\tau = 0.005 \ h^{-1}$, also possess a negligible amount of time-step bias.

The success in accurately calculating the H_3 saddle-point energy encouraged the computation of energies at two other points along the reaction coordinate. The reaction-coordinate geometries are also taken from Ref. 65. For these points Ψ_3 is employed. This choice is dictated by the desire to use a small basis set beyond SZ which treats all the hydrogen atoms equivalently. In Fig. 3.17, the QMC reaction coordinate energies, and a spline fit to Liu's, are compared. The QMC reaction coordinate energies, which are upper bounds since the QMC asymptotic energies are exact, agree with Liu's results to within the statistical error, 0.25 kcal/mol.

In conclusion, single-determinant trial functions with small basis sets yield essentially exact energies for the parts of the H_3 PES studied. As a result, the QMC-computed classical barrier height to the $H + H_2 \rightarrow H_2 + H$ exchange reaction, 9.71

(0.13) kcal/mol, is in excellent agreement with the exact released-node value of 9.65 (0.08) kcal/mol and Liu's CI estimate of 9.59 ± 0.06 kcal/mol. The computational cost of these calculations was somewhat high; approximately five hours of Cray-1S time to obtain a statistical error of 0.25 kcal/mol using Ψ_3. Since this time, vectorization and other enhancements have been incorporated into our QMC codes. A more complete QMC study of the H_3 PES, for purposes of investigating the possibilities of resonances, has recently been undertaken by Maria Soto. These calculations on H_3 now yield the above statistical error in 30 minutes or less on a Cray/XMP.

3.3.2. The Electron Affinity of Fluorine

The accurate determination of electron affinities (EAs) remains a challenging theoretical task because the total energies of an atom or molecule and its anion must be computed to the same accuracy. Since an accurate determination of an EA requires an accurate computation of the difference between two species with much different electronic structures, difficulty arises using standard *ab initio* approaches because a given level of theory generally describes the atom and anion to different levels of accuracy. Attempts to circumvent this difficulty leads to the use of very large basis sets and CI expansions. However, though accurate energies may result, the energy difference, EA, may still be of poor quality. This shortcoming of standard methods makes desirable an exploration of alternative theoretical techniques for obtaining EAs.

As seen from our computations on H_3, simple trial functions, single-determinant with small basis sets, yielded an energy difference accurate to 0.1 kcal/mol. Such accuracy is a strong motivation to use QMC in approaching the difficult task of

calculating an accurate EA of F. Once again simple trial functions are employed to obtain the desired energy difference.

The history of EA calculations begins effectively with the semiempirical work of Clementi[70] and Clementi and Mclean[71] who tabulated non-relativistic energies for several atoms and ions. These energies were obtained by subtracting semiempirical estimates of the relativistic corrections[72] from the experimental energies.[73] The energies so obtained for F and F^- led to a non-relativistic EA which was within 1% of experiment. The point to note here is that the relativistic corrections to the energy of F and F^- are almost identical – leading to an effective cancellation in determining the EA of fluorine. This is expected since the added electron goes into the valence shell where relativistic effects are small. Subsequent semiempirical estimates of the EA of F and other atoms were performed in the late 1960's.[74,75] In those studies, correlation energies for atoms and positively charged ions were estimated by configuration interaction (CI) methods. Non-relativistic "experimental" correlation energies[76] (also for the atoms and the positive ions) were combined with these CI energies to infer correlation energies for the negative ions. Electron affinities were thus deduced from experimental results and CI estimates of the correlation energies of atoms and *cations*.

Because it is desirable to compute physical quantities without reliance on experimental data, several *ab initio* studies of EAs were undertaken in the early 1970's.[77,78] To compute reliable *ab initio* EAs requires the accurate treatment of electron correlation. This is readily demonstrated by the poor quality of Hartree-Fock (HF) EAs.[79] Among the first-row atoms, for example, HF theory correctly predicts

the *sign* of the EA for only two cases. Even for these systems, the HF EAs are only 40% of the experimental values.

To incorporate correlation effects, Weiss[77] and Moser and Nesbet[78] employed the atomic Bethe-Goldstone method to calculate total energies for first-row atoms and ions. The technique relies on two assumptions: (a) that the total energy can be written as

$$E = E_0 + \sum_n \varepsilon_n + \sum_{mn} \varepsilon_{mn} \ , \tag{3.11}$$

where E_0 is the HF energy and ε_n and ε_{mn} are one- and two-particle correlation energies, respectively; and (b) that these correlations can be calculated independently using a superposition of configurations wave function. That is, the correlation energy of the mn^{th} pair, ε_{mn}, is taken to be the energy lowering (relative to HF) upon including *only* double excitations of the mn pair in a CI expansion. The Moser-Nesbet and the Weiss EAs were found to be in good accord with experiment, having an error of only 2% for F.

Sasaki and Yoshimine[13] (SY), however, showed that summing one- and two-particle correlations, as implied in Eq. (3.11), would over-estimate the EA of F by about 6% (0.20 eV) if a complete basis set were used. The SY study involved extended basis sets and large CI expansions in computing total energies of the first-row atoms and their singly-charged anions. The basis sets contained up to i functions (l=6), while the configuration list consisted of all single and double excitations from the lowest-energy configuration, as well as selected triples and quadruples. Though the SY total energies for F and F^- are the lowest obtained to date from variational calculations – having achieved over 95% of the correlation energy – the EA is nevertheless

0.28 eV (8%) less than the observed value. Thus SY's study again demonstrates the difficulty of determining the EA of F (and other atoms) by variational techniques.

Recently, several new theoretical estimates of the EA of F have been reported.[80-83] The techniques employed include many-body perturbation theory[78,79] as well as multiconfiguration self-consistent field (MCSCF) plus CI methods.[82,83] These sophisticated treatments still result in errors of 4-10%. Most recently, Moeller-Plesset perturbation theory[84] and coupled-cluster (CC)[85] computations have yielded very accurate EAs of 3.47 and 3.37 eV., respectively. The CC calculations also yielded 95% of the (CE) for F and F^-.

It is worth noting that the experimental determination of EAs is also difficult, with experimental values for the EA of F ranging from 3.4-3.5 eV.[86-88] Presently, however, the recommended EA of F is 3.399 eV.[88] For a further discussion of the experimental work, see Ref. 82.

The accuracy of the H_3 energy resulting from single-determinant trial functions with a DZ basis set led to the use of such trial functions for F and F^-. The F atom exponents and linear coefficients are those of Clementi and Roetti[89], while those of F^- resulted from a hand-optimization for a minimum in the SCF energy using HONDO.[68] The basis sets for the atom and anion are presented in Table 3.6.

For purposes of comparison, the SCF energy of the F-atom basis set is within 0.006 h of the HF limit. The SCF energy for F^-, and therefore agreement with HF, is more difficult to analyze because SCF energies resulting from HONDO are too high in comparison to the true SCF energy when STO basis sets are employed. The error, which is more noticeable for higher-Z systems, results from using (a maximum of six)

Gaussian functions to describe each STO in the basis set. For example, a HONDO SCF calculation on F with its DZ basis set, though giving accurate linear coefficients, yielded an energy 0.0336 h above Clementi and Roetti's value. Subtracting this difference from the HONDO SCF energy gives an estimate of the correct SCF energy of F^- lying within 0.021 h of the HF limit. The poorer agreement, compared to F, of the F^- SCF energy with the HF is not surprising given the difficulty of describing the diffuse covalent electron density of an anion with a small basis set. Given the H_3 results, however, agreement or disagreement of the SCF with the HF energy was not a major concern in considering nodal structure accuracy.

The correlation functions used here have the same form as those employed for H_3, cf. Eqs. (3.4), (3.5) and (3.9). In selecting correlation function parameters, optimization algorithms, besides hand-optimization, were not available. Unfortunately, the difficulty of hand-optimization, in which parameter values are selected and the corresponding VMC energies are then compared, greatly increases with the number of electrons and the total energy. This is because the statistical error in the VMC energy, for a given amount of computation time, generally rises rapidly with the size of the system. Therefore, while a_1 is optimally chosen to satisfy the like-spin E-E cusp condition, the degree of optimization of the remaining parameters is very crude. Despite these difficulties, about 40% of the CE was obtained in the VMC energies of the atom and anion. Correlation function parameters are reported in Table 3.6.

In Table 3.7 our QMC results are compared with previous studies, see also Ref. 54. The QMC total energies are quite good, achieving over 90% of the correlation energy for both F and F^-. These energies are lower than those of all previous variational

calculations, with the sole exception of those of SY.[13] It should be pointed out, however, that it is quite difficult to perform such large basis set, extended CI calculations, as done in Ref. 13, on most chemical systems. Moreover, the QMC *electron affinity* gives much better agreement with experiment[88] than that of SY, as well as giving better agreement than all other variational calculations. Moreover, the present QMC EA is also as good as or better than non-variationally obtained EAs.

Our QMC energies for several time steps, τ, are presented in Table 3.8, and shown graphically in Figure 3.18. In the figure one sees that the fixed-node energies of F and F^- increase with decreasing τ. Furthermore, the data strongly suggests that the energy is linear in τ for small τ. As can be seen in Fig. 3.18, it appears that the F^- energy has begun to deviate from linearity by $\tau = 0.005\ h^{-1}$, while the F energy remains linear in this region. In other words, the observed linear domain is smaller for F^- than for F.

The statistical error in the QMC EA, albeit reasonable, is larger than preferred. For example, the mean differs from experiment by only 1.5% but the statistical error yields a small, but not negligible, probability (about 15%) that the QMC EA is inaccurate by 5% or more. To reduce the statistical error in the EA requires a reduction in the uncertainties of the QMC energies. These energies are obtained from extrapolation to $\tau = 0$ as indicated in Table 3.8 and shown graphically in Fig. 3.18. As seen from Table 3.8, time-step bias is now much larger than was the case for the low-Z system, H_3. For H_3 a time step as large as $0.1\ h^{-1}$ was acceptable, while here $\tau = 0.00125\ h^{-1}$ still shows some bias in the F-atom energy and the very small time step of $0.0005\ h^{-1}$ yields a bias for F^-. The result is that reliable extrapolations to $\tau = 0$ require the

computation of QMC energies at small time steps which greatly decreases efficiency. As an example of the computing required, at $\tau = 0.005\ h^{-1}$, roughly 3 hours of Cray/XMP time were needed to calculate the energy for F, and similarly for F^-, to the accuracy shown in Table 3.7. For the time-steps used in these calculations, halving τ doubles the computer time needed to achieve a given statistical accuracy. Thus the overall computational cost was quite high. (The effect of this cost, however, was significantly reduced by running our small-memory code at low priorities.) Therefore, obtaining smaller statistical errors, by simply continuing the calculations as discussed above, are not practical.

After the EA calculation, however, we discovered the effect of the E-N cusp upon the time-step bias, cf. Sec. 3.2 and Figs. 3.9-3.11. Therefore, it is of interest to compare the values of λ employed above with those which optimally satisfy the E-N cusp condition, λ(cusp), see Table 3.9. For F and F^-, the distributions of λ values, each computed to satisfy the cusp condition for a given set of coordinates, are sharply peaked, and each resembles a Gaussian with a longer tail at larger values, as in Fig. 3.8. For F, 68% of the computed λ values lie within 0.006 of λ(cusp) = 0.064, while 67% differ from λ(cusp) = 0.0165 by only $\pm$ 0.0015 for F^-.

As seen in Table 3.9, the discrepancy between the hand-optimized value of λ used in the QMC calculation and λ(cusp) is far greater, as is the time-step bias, for F^- than for F. A subsequent calculation of the F^- energy at $\tau = 0.005\ h^{-1}$ with $\lambda = 0.0165$ yielded an energy of -99.8393(37) h. This energy possesses far less bias than the previously computed value at this time step, -99.9141(28)h, and is in agreement with that computed at $\tau = 0.0005\ h^{-1}$, -99.8375(31) h. Thus the same amount of bias is

obtained at an order of magnitude larger time step implying, though not proving since only one time step is compared, that a similar increase in efficiency can be obtained. Resetting λ to λ(cusp) for F should also yield a (less dramatic) reduction of time-step bias.

In summary, the main difficulty encountered in variational calculations of EAs is that of describing the atom and the anion to the same accuracy. Generally, the atom, being easier to describe than the anion, is treated more accurately at the same level of theory. This leads to an underestimation of the EA. Even for a large basis set, large CI calculation, such as that of SY, this effect persists. For example, SY underestimate the EA of F by 0.28 eV. On the other hand, with relatively simple importance functions, i.e., DZ basis-set, single determinant, QMC appears to treat F and F^- to virtually the same accuracy (the EA is in error by 0.05 ± 0.11 eV). Apparently, the quality of the *nodal structures* obtained in the SCF procedure does not vary significantly from F to F^-, and so the fixed-node error in the total energy of each species is almost identical: 99.9691(21) and 99.9699(45)% of the estimated total energy is gained for F and F^-, respectively.

3.3.3. The Bond Energy of Nitrogen

The accurate determination of the bond energy of N_2 is a challenging task for theoretical chemists. Theoretical studies of N_2, to list a few, have included expansion[92-96], perturbation theory[97-100], coupled-cluster[101], and local density approximation[102,103] approaches.

The difficulty faced by expansion methods is best exemplified by Siegbahn's

calculation[92] which yielded a dissociation energy, D_e, 0.45 eV below the experimental value of 9.91 eV[104] despite the large basis set employed. Using bond-centered functions, Wright and Buenker were able to obtain a D_e of 9.96 eV. This accurate value, however, was obtained by a balancing between the relatively small numbers of atom-centered and bond-centered functions employed. As found in their calculations, changing this balance destroys the close agreement with experiment. Full CI calculations with basis sets large enough to give the desired accuracy, 0.1 eV or better, are not currently possible due to computational cost. This difficulty has led to the development of schemes[92,96] which approximate the full CI limit without excessive CPU and storage requirements.

Computations based on Moeller-Plesset perturbation theory are more accurate. Values of 9.70[98,99] and 9.80 eV[100] have been reported.

As an interesting and difficult problem, the computation of the dissociation energy of N_2 serves as an excellent benchmark for further investigating the capabilities of fixed-node diffusion QMC. For H_3, the fixed-node approximation and the time-step bias are of little consequence, and a near-exact energy is obtained. For F and F^-, on the other hand, the fixed-node approximation resulted in extrapolated energies noticeably above estimates of the exact Born-Oppenheimer energy, and the large time-step bias led to high computational costs. The fact that the total energies of the atom and anion are obtained to practically identical accuracy, however, yielded a very accurate EA. If the similarity in the N and N_2 QMC energies is that of those for F and F^-, then the QMC dissociation energy will also be of high quality.

Continuing in the same vein as the F-F^- calculations, the nodes are given by a

single-determinant constructed from a double-zeta basis set. In addition, the forms of the E-E and E-N correlation functions are the same as those employed in Secs. 3.3.1 and 3.3.2 but with $a_2 = b_2 = 0$. These parameters are considered the least important (they do not affect the E-E cusp condition) and are, therefore, not employed since their usefulness was furthered limited by the unavailabilty of an efficient optimization algorithm. The basis-set exponents are taken from Clementi and Roetti.[89] For the atom, the linear coefficients are also given by Ref. 89, and those of the molecule were computed with HONDO.[68] The exponents and hand-chosen correlation function parameters are reported in Table 3.10. These simple trial functions yielded 40 and 18% of the CE for N and N_2, respectively. The lower percentage of the CE obtained for N_2 is due, in part, to the greater degree of basis-set incompleteness for the molecule.

The QMC energies of N and N_2 were computed at several time steps, and later recomputed, after the importance of satisfying the E-N cusp condition was discovered, with $\lambda = \lambda(\text{cusp}) \simeq 0.05$. These energies are presented in Table 3.11 and plotted in Figs. 3.19 and 3.20. The QMC energies, for which the E-N cusp condition was not considered, once again possess a large amount of time-step bias. The energies computed at the smallest values of τ, 0.00060 h^{-1} and 0.00125 h^{-1} for N and N_2, respectively, are still biased above the extrapolated result. Linear least-squares fits, which pass within the statistical error of each point, yield energies corresponding to 90.5% of the CE for both N and N_2. While accuracy in the CE is the same for both the atom and the molecule, this is not the case for the total energy. The result is that the QMC dissociation energy of 9.3(1) eV is significantly different from the experimental value of 9.91 eV, i.e., an error of 6%.

In addition, there appears to be some discrepancy between the extrapolated energies obtained from computations with and without the E-N cusp condition satisfied. Since changing λ does not affect the nodal structure, the energy at $\tau = 0$ should not depend on this parameter. The QMC energies of N and N_2 with $\lambda = \lambda(\text{cusp})$, as expected, have much less bias, the slope of E versus τ is reduced by about a factor of six, and are therefore more reliable. Thus the bias in the N and N_2 QMC energies is problematic to the extent that time steps even smaller than the ones employed are required to obtain a reliable extrapolation to $\tau = 0$ when $\lambda \neq \lambda(\text{cusp})$.

The dissociation energy obtained with $\lambda(\text{cusp})$, however, is not significantly affected; $D_e = 9.0(3)$ eV. Also, the amount of correlation energy obtained for the atom is once again the same as that obtained for the molecule, now 87%. The fact that the accuracy in the total energy differs between the two species, as opposed to the case of F and F^-, is not necessarily unreasonable. Unlike F and F^-, there is a large difference in the total energy between the two systems studied. Therefore, if the accuracy obtainable by a given trial function form depends on the total energy, and thus complexity, of an atom or molecule, then electron affinities are more easily computed than bond energies. Another dissimilarity, not present for F and F^-, is the different symmetry between N and N_2. Further experience is necessary to predict the amount of cancellation of error obtained with a given type of trial function when computing energy differences.

3.4 Energy Computations on Other First-Row Atoms and Molecules

As seen from the calculations presented in Secs. 3.2-3.3, the fixed-node error of

single-determinant trial functions is very small for H_3 and increases to about 10% of the CE for larger systems. Beyond this, little else is discernible concerning the fixed-node error of a given type of trial function. This section presents QMC energies for a variety of low-Z atoms and molecules in order to gain an understanding of sources of fixed-node error in QMC calculations.

The exact ground-state energies of two-electron atoms and molecules are readily obtained by QMC since these systems do not possess spatial nodes. This fact is demonstrated by calculations on the ground state of H_2 and He, cf. Figs. 3.9 and 3.10.

Turning now to two-electron systems which possess nodes, we consider the first three excited singlet S states of He. As these states have the same symmetry as the ground state, their QMC energies are not variational. A simple example of $E_{\mathrm{QMC}} < E_{exact}$ is given by considering the $2s$ state of H with the node displaced from its exact location at $r = 2$, e.g., $r > 2$. In this case the nodal volume containing the nucleus is too large and the QMC energy obtained by sampling in this region will be lower than the exact $2s$-state energy, i.e., as the node is moved out to infinity the ground-state energy is recovered. Since the lowest-energy nodal volume dominates after sufficiently long times, the asymptotic QMC energy will be too low. However, if the nodes of the trial function describing the excited state in question are sufficiently accurate, the computed QMC energy will be of high quality.

For the excited ^1S states of He the two electrons are much farther apart than in the ground state. Thus their effect on each other, i.e., the degree to which their movement is correlated, is reduced. The result is that compact trial functions yield accurate VMC energies. Such trial functions[107] are employed here with the hope that their nodal

structures are very good. These trial functions are given by

$$\Psi_T(\underline{r}_1, \underline{r}_2) = \psi_1(r_1)\psi_2(r_2) + \psi_1(r_2)\psi_2(r_1). \tag{3.12}$$

In Eq. (3.12), ψ_1 is simply a Slater $1s$ orbital and ψ_2 is a Slater $1s$ orbital multiplied by a function which is a linear combination of the first four Laguerre polynomials.[107] The accuracies of $E_T = \langle\Psi_T|H|\Psi_T\rangle$ are very good, the errors are only 3, 0.9, and 0.4 *milli-hartree* (*mh*) for the first three excited states, respectively, cf. Table 3.12.

The QMC energies, computed employing deletion upon crossing a node, are compared with exact values in Table 3.12. The time-step bias, which is visible because of the small statistical errors, is not large, and reliable extrapolations are easily obtained. Differences between rejection and deletion are probably not observable at the time steps employed.

For ground states, the nodal volumes, which arise from exchange antisymmetry, are equivalent. This is not the case for most excited states for which additional nodes generally result in nodal volumes that are not equivalent, in terms of the QMC energy, unless the trial function's nodal structure is correct. This effect was studied for the $1s3s\ ^1S$ and $1s4s\ ^1S$ states by sampling points only in regions which possess the same sign[Ψ_T]. (It should be pointed out, however, that, given the number of nodal volumes in the trial functions describing these states, there is probably more than one distinct nodal volume for a given value of sign[Ψ_T].) As seen in Table 3.12, the QMC energy does show a significant dependence on the nodal volume(s) sampled. Since this dependence vanishes as the nodes become exact, the differing QMC energies are indicative of the error in the nodal structure of the trial function. Therefore, despite

the high accuracies of $E[\Psi_T]$ for each state, the location of the nodes is far from correct.

The incorrect placement of the nodes is further reflected by the QMC energies obtained. We see that the QMC energies do not recover a large percentage of the remaining energy, $E_T - E_{exact}$, cf. Table 3.13. For the first-excited state, only 59% of the remaining energy is recovered (for H_3 over 99% is recovered). For the second-excited state, sampling only where $\Psi_T > 0$ significantly improves the energy over naive sampling, cf. Table 3.12, yet the amount of energy recovered is still modest, 75%. Interestingly, the QMC energies for the $1s4s\,^1S$ state show no improvement over E_T when $\Psi_T > 0$ and the QMC energy is actually above E_T when $\Psi_T < 0$. Theoretically $E_{\mathrm{QMC}} \leq E_T$, but only for each nodal volume. Therefore, these counterintuitive results for the last excited state must be caused by sampling in nodal volumes which possess relatively high energies – despite varying sign$[\Psi_T]$. Though higher-energy nodal volumes will decay away after sufficient imaginary time, cf. Sec. 2.4, the energy differences found here are too small, $\approx 1\ mh$, for this decay to be significant for the simulation times we employed, $\approx 1000\ h^{-1}$.

Finding *all* the distinct nodal volumes of a trial function may be difficult when studying excited states of systems more complex than He. However, the differences in QMC nodal volume energies found here are less than 1 *mh* (0.6 kcal/mol) which is generally an acceptable accuracy when computing total energies. If the differences in nodal volume energies are greater, the populations of the higher-energy volumes will decay away more quickly allowing convergence to the lowest-energy volume to be more easily obtained. Therefore, the presence of distinct nodal volumes may not be

significant for computations of excited-state energies of other systems.

Other Monte Carlo algorithms for computing excited state energies exist. [109-110] These approaches are variational, can compute several energies simultaneously, and do not have nodal-volume dependences. They do, however, present other difficulties and have not yet been employed in studying electronic states so comparison is difficult.

QMC energies for other systems which we studied are compared with estimates of the exact electronic energy in Table 3.14. In all cases, the electronic state is the lowest in energy for the symmetry considered so QMC energies are variational to the extent that bias and statistical error are eliminated. The QMC energies are generally obtained from single-determinant trial functions and are discussed on a case by case basis.

We now consider the QMC energies computed for the ground state (2^2S) and first-excited P state (2^2P) of Li. These two states are among the simplest for which fixed-node error is a concern. As such, they afford an excellent starting point for assessing the accuracy obtained from a single-determinant trial function. The near-HF basis set and linear coefficients are taken from Weiss.[111] The E-N cusp condition is well satisfied by the Slater determinant so only the E-E correlation function described in Sec. 3.3 (with $a_2 = b_2 = 0$) is employed. The nodal boundary conditions are enforced by rejection.

No statistically significant time-step dependence in the QMC energy of the excited state is found at $\tau = 0.005$ and $\tau = 0.010\ h^{-1}$, the statistical error in the difference between the two energies being over three times larger than the difference itself. Therefore, the reported excited-state energy is the weighted (by σ^{-2}) average of the

energies obtained at the time steps above. Given the imperceptible amount of bias in the P-state energy, the reported ground-state energy is simply obtained from a calculation at $\tau = 0.010h^{-1}$. The QMC energies agree with the exact values to well within their statistical errors (0.24 *mh* or 0.15 kcal/mol). Though some fixed-node error must result from the trial functions, it is found to be very small for these simple systems, as is the case for H_3. Actually, the largest errors appear to be statistical, e.g., the P-state energy is below, but within less than one standard deviation of, the exact energy.

Turning to more complex electronic structures, we now consider LiH and Li_2, at their experimental equilibrium internuclear separations. With single-determinant trial functions, the QMC energies reported in Table 3.14 are seen to be very accurate. The LiH QMC energy, -8.06908(43) h, is as low as that of the largest CI calculation to date[112], 132,000 configuration state functions and $E = -8.06904\ h$, and only slightly above the energy (which is not variational) obtained in a recent coupled-cluster calculation,[86] $E = -8.06951\ h$. The Li_2 energy differs from the exact by only 2.2 *mh* and gains 98% of the CE. This energy is more accurate than any other *ab initio* result. (However, most *ab initio* calculations emphasize accuracy in the potential energy curve rather than the total energy.)

In spite of the high accuracy in the QMC energies, fixed-node error is now noticeable. Therefore, it is of interest to consider deficiencies in the trial functions which may have led to the errors observed. For LiH, the basis set, given in Ref. 43, is a truncation of Cade and Huo's[113] with the linear coefficients re-optimized by the SCF technique. Using one-, two-, and four-determinant trial functions, Harrison and Handy

(HH)[56] obtained QMC energies of -8.0696(7), -8.0701(4), and -8.0697(3) h, respectively. The basis set employed in these calculations is identical to ours with the addition of a $2s$ STO on H. These calculations were all performed at $\tau = 0.010\ h^{-1}$. Given that SCF and MCSCF wave functions generally satisfy the E-N cusp condition very well (HH did not employ an E-N correlation function), and our experience with LiH, cf. Fig. 3.11, the time-step bias in the three energies presented above should be insignificant. Most noteworthy are the facts that the HH single-determinant energy appears to be better than our own, -8.0696(7) versus -8.0691(4) h, very similar to those of the multi-determinant trial functions, and in near-agreement with the exact, -8.0702 h. These facts imply that a single-determinant trial function should yield an almost exact QMC energy. Very recently, we discovered a node in the lowest-energy MO of our trial function. Removing this node, by slightly changing two of the linear coefficients (which hardly changed the exchange nodes), and recomputing the QMC energy yielded a much improved energy, $E = -8.0702(6)\ h$. This experience with LiH, and that with H_3, demonstrates the importance of checking for spurious nodes in the MOs used to construct the Slater determinant.

The Li_2 energy, though very accurate, also possesses a noticeable fixed-node error. It is well-known from MCSCF calculations, however, that the near $s-p$ degeneracy of the Li atom can be important in describing this molecule. Also, the wave function should properly describe dissociation, e.g., a single-determinant corresponds to a combination of ionized and neutral atoms as the internuclear separation becomes large. The importance of $s-p$ degeneracy and proper dissociation is reflected in the large mixing coefficients of the determinants describing these effects in a CI or MCSCF

expansion, for an example see Ref. 114. For the QMC calculations, the trial function's nodes are those given by a single determinant[25]. Therefore, it is reasonable to expect that adding determinants which account for $s-p$ degeneracy and allow for proper dissociation will yield a trial function giving an improved QMC energy.

Most recently, we have performed a QMC calculation with a four-determinant trial function which describes the effects discussed above. The QMC energies are computed with $\lambda = \lambda(\text{cusp})$ for $0.005 \leq \tau \leq 0.100\ h^{-1}$, and time-step dependence is no longer perceptible by $\tau = 0.035\ h^{-1}$. The QMC energy is now found to be -14.9945(4) h, in excellent agreement with the exact value, -14.9945 h. This result demonstrates that the QMC energy may be computed to near-exact quality by considering effects that are important in expansion techniques.

Perhaps the most instructive case of how a single-determinant trial function can yield a poor QMC energy is the ground state of Be. In Table 3.14, we see that only 89% of the CE is recovered. At first this result is surprising given the accuracy of the QMC energy of LiH and Li_2, systems with as many or more electrons and of lower symmetry. However, the $s-p$ degeneracy effect in this atom is known to be especially strong. Since treating this degeneracy is most important in obtaining an accurate wavefunction in standard *ab initio* approaches, it is of interest to determine whether an improvement in nodal structure is also obtained. This question is considered by the calculations of HH who employed, in addition to the determinant arising from the $|1s^2 2s^2|$ occupancy, another corresponding to $|1s^2 2p^2|$. (A $|2s'^2 2s^2|$ determinant, to describe radial correlation, was also included, but it had a much smaller mixing coefficient, 0.02 versus 0.30 for the $|1s^2 2p^2|$ determinant, indicating that this effect is

much less important.) Once again, improving the trial function to describe well-known effects yields a much more accurate QMC energy; HH obtained 98% of the CE.

The $s-p$ degeneracy effect discussed in the preceding paragraphs, also appears to be significant for the ground state of BH. The result is that the single-determinant QMC energy gains only 88% of the CE. Therefore, a multi-determinant trial function is also recommended for BH as well as other systems which possess degeneracy effects.

In addition to the QMC energy, trial function optimization and the effect of rejection versus deletion were studied for BH. In the calculations described in Sec. 3.3, correlation function parameters are hand-optimized. This method is reasonable for small systems, two and three electrons, but impractical otherwise. At the time of the BH calculations a method for machine-optimizing parameters became available. This method uses correlated sampling and a steepest-descent approach during a VMC walk to obtained optimized parameters.[115] In a calculation on H_2, the algorithm worked very well, an optimized nine-parameter trial function gave a VMC energy which obtained 91% of the CE.

The initial BH trial function, Ψ_1, is given in Table 3.15. The basis set is a truncation of Cade and Huo's[113] with some of the exponents re-optimized in SCF calculations. The correlation function parameters are hand-optimized, and the VMC energy gains only 26% of the CE. The optimized BH trial function, Ψ_2, is given in Table 3.16, where U_1 in the E-E correlation function is now given by,

$$U_1(r) = \ln[1 - b_1 \exp(-r/a_1 - r^2/a_2)] \,, \tag{3.13}$$

with $a_1 = 2b_1/(1 - b_1)$ to satisfy the like-spin E-E cusp condition and $b_1 < 1$. The

trial function, Ψ_2, was obtained by optimizing the exponents, linear coefficients, and most of the correlation function parameters. The parameter a_1 is fixed to satisfy the like-spin E-E cusp condition, and λ_H and λ_B are also fixed since varying these parameters during optimization led to poor results. Generally, attempting to optimize, by minimizing the energy, parameters which have a large effect on cusp conditions leads to an overemphasis in regions near the cusp and a poor VMC energy. The MOs are now optimized in the presence of an explicit description of electron correlation in order to examine the effect on the fixed-node error. The optimized trial function yielded a substantial, though less than hoped, improvement in the VMC energy, 45% of the CE obtained.

Table 3.17 compares the QMC energies of Ψ_1 and Ψ_2 at $\tau = 0.01\ h^{-1}$ The effect of enforcing nodal boundaries by rejection rather than deletion is also studied. The energies in Table 3.17 show that while optimization improves the VMC energy, in this case no improvement in the QMC energy is observed. These results indicate that the limitation of a single determinant in Ψ_T is the overriding factor in determining the accuracy of the QMC energy. Also in Table 3.17, the difference between deletion and rejection is seen to be negligible. This is in contrast with the H ($2p_x$) atom calculations discussed in Sec. 3.2. This is not unreasonable since the treatment of the nodal boundaries should be more important for excited states which possess a greater density of nodes than do ground states. [However, a difference between rejection versus deletion results was noticed in the computation of properties other than the energy (discussed in Chap. 4).]

The last system to be discussed is the often-studied H_2O molecule. The trial

function parameters are obtained from an efficient optimization algorithm[25] and presented in Ref. 25. The E-N correlation function parameters λ_H and λ_O, however, have been changed slightly to optimally satisfy, as described in Sec. 3.2, the E-N cusp conditions. The function U_1 is now given by

$$U_1(r) = -b\exp(-ar/b)\ , \tag{3.14}$$

where $a = 0.5$ to satisfy the like-spin E-E cusp condition. The resulting QMC energies are presented in Table 3.18 as well as the $\tau = 0$ estimate (also in Table 3.14). The amount of CE gained is similar to that of the other "high"-Z systems studied; N, N_2, F and F^-.

The H_2O calculations also emphasize the usefulness of choosing the E-N parameters λ_α so that the E-N cusp condition is best satisfied. Note that observable time-step dependence in the QMC energy has vanished at $\tau \leq 0.0025\ h^{-1}$ allowing a reliable estimate of the $\tau = 0$ energy to be obtained simply from the $\tau = 0.0025$ and $\tau = 0.0010\ h^{-1}$ energies. This behavior contrasts sharply with the F-F^- time-step dependence where biases were still observable at $\tau \approx 0.001\ h^{-1}$.

As seen from the QMC energies presented in this and the preceeding section, useful knowledge, as well as accurate energies, has been gained concerning the accuracy of nodal structures. The most important fact is that obtaining accurate nodal structures requires consideration. For many of the systems studied, a single determinant yields only 90% of the CE is obtained in contrast to the QMC energies of some simple systems which are nearly exact. In addition, it is found, by us for Li_2 and by Harrison and Handy[56] for Be, that improving the trial function to account for well-known effects, e.g., $s-p$ degeneracy, also leads to a significant improvement in the QMC

energy. It is therefore of interest to ascertain whether describing this degeneracy, and other effects, will also yield accurate energies for systems such as BH, N and N_2, where a single determinant is not satisfactory. Similar considerations will also be relevant for other atoms and molecules, leading to an increase in the accuracy and competitiveness of the QMC approach employed here.

3.5 Computing the QMC Energy by Difference

As found in the calculation of the EA of F, the statistical error in QMC computations can be problematic. A major contribution to the statistical error of the QMC energy arises from the magnitude of the total energy. For example, 1 kcal/mol statistical accuracy is obtained for H_2 ($|E| \approx 1.17$) when the relative statistical error in the energy is 0.14%. For F atom, on the other hand, this error in the energy must be *two* orders of magnitude lower to obtain the same precision in the total energy. In contrast, the correlation energy (CE) only increases by *one* order of magnitude in going from H_2 to F, 0.04 h versus 0.33 h, and is generally two orders of magnitude smaller than the total energy. This implies that if the CE and the total energy can be computed to the same relative statistical error after equal computation times, then the statistical error in the CE will be much smaller than that of the total energy. Therefore, QMC energies (E), given as the sum of the HF and CE energies, or alternatively as $E_{\mathrm{VMC}} + (E - E_{\mathrm{VMC}})$ if correlation functions are employed, may be computed more efficiently by calculating energy differences. In this section an idea for directly computing the energy difference, $\Delta \equiv E - E_{\mathrm{VMC}}$, is presented and explored.

The trial functions may or may not contain correlation functions. In the former

case, the statistical error of the QMC energy arises from both Δ and E_{VMC}, while in the latter the statistical error is due solely to that of Δ. The relative merits of including or not including correlation functions in Ψ_T is discussed below.

To describe the method by which Δ can be computed during a single walk, we begin by noting that, as stated in Sec. 2.5, the branching of the QMC walk may be implemented through removing or creating walkers ("integer rounding"), by carrying weights, or a combination of both. Note that when weights are carried and moves are rejected, the VMC and QMC walks are isomorphic. That is, omitting the weights, which suppresses the branching, yields $|\Psi_T|^2$ while including them gives $\Psi_T\phi_0$. (Essentially, the weight is a statistical estimate of ϕ_0/Ψ_T.) Therefore, VMC and QMC expectation values, and more importantly the differences between them, may be obtained in a correlated fashion during a VMC walk in which weights are computed.

To see how the weights are computed, consider a point, $\underline{R}_0$, which is moved to $\underline{R}_1$ after one time step (τ), $\underline{R}_2$ after the second move, and so on. Repeated use of Eq. (2.53) yields (now with slightly different notation for the points sampled)

$$w(\underline{R}_n) = M(\underline{R}_n, \underline{R}_{n-1})M(\underline{R}_{n-1}, \underline{R}_{n-2}) \cdots M(\underline{R}_1, \underline{R}_0)w(\underline{R}_0) \ , \tag{3.15}$$

where $w(\underline{R}_0) \equiv 1$ since points are sampled from $|\Psi_T|^2$. In Eq. (3.15), M is the branching factor of the short-time Green's function, cf. Eq. (2.44), giving

$$w(\underline{R}_n) = \exp\Big[-\tau\Big(\frac{E_L(\underline{R}_0)}{2} + \sum_{i=1}^{n-1} E_L(\underline{R}_i) + \frac{E_L(\underline{R}_n)}{2} - nE_R\Big)\Big] \ . \tag{3.16}$$

Above, $\tau_a = \tau$ is employed because weights may be computed more efficiently with a constant value of the time step. Since small time steps are required to reduce time-step bias, the acceptance ratio is very nearly equal to one yielding $\tau_a \approx \tau$ on average.

Therefore, not adjusting τ to account for rejection of moves should have little effect (which vanishes as $\tau \to 0$). The weight of the next point, $\underline{R}_{n+1}$, is then given by

$$w(\underline{R}_{n+1}) = \exp\Big[-\tau\Big(\frac{E_L(\underline{R}_1)}{2} + \sum_{i=2}^{n} E_L(\underline{R}_i) + \frac{E_L(\underline{R}_{n+1})}{2} - nE_R\Big)\Big] , \tag{3.17}$$

and the weights of subsequent points are computed in a similar fashion. The number of intermediate steps, n, must be chosen large enough so that the convergence time, $t = n\tau$, yields convergence to $\Psi_T \phi_0$ from $|\Psi_T|^2$.

With the weights computed as described by Eqs. (3.15)-(3.17), statistical estimates of the E_{VMC}, E, and Δ are, for a total of N (weighted and unweighted) points

$$E_{\mathrm{VMC}} = N^{-1} \sum_{i=1}^{N} E_L(\underline{R}_i) , \tag{3.18a}$$

$$E = \sum_{i=1}^{N} w(\underline{R}_i) E_L(\underline{R}_i) \sum_{i=1}^{N} w(\underline{R}_i) , \tag{3.18b}$$

and

$$\Delta = \sum_{i=1}^{N} w(\underline{R}_i) E_L(\underline{R}_i) \sum_{i=1}^{N} w(\underline{R}_i) - N^{-1} \sum_{i=1}^{N} E_L(\underline{R}_i) . \tag{3.18c}$$

Since there is a positive correlation between the above VMC and QMC energies, the difference will possess a statistical smaller error than either of those in the VMC or QMC energy.

From Eq. (3.18c), the amount of correlation between the VMC and QMC energies decreases as fluctuations in the weights increase. In turn, these fluctuations depend on the behavior of the local energy and the time, t, required to converge to $\Psi_T \phi_0$. That is,

$$\sigma_w / w = w^{-1} |\partial w / \partial E| \sigma_E ,$$

and from Eq. (3.16) there results,

$$\sigma_w / w = t\sigma_E \ , \qquad t = n\tau \ . \tag{3.19}$$

Possible advantages of employing correlation functions are implied by Eq. (3.19). An SCF trial function yields no statistical error in the VMC energy since this energy is simply the analytically computed SCF value. However, an SCF trial function generally yields a value of σ_E 2-3 times greater, primarily because the like-spin E-E cusp condition is not satisfied, than that obtained with correlation functions. In addition, to the extent that correlation functions improve the accuracy of the trial function, they should reduce the time required to converge to $\Psi_T \phi_0$. For these reasons, improvement over the SCF description will reduce, by reducing σ_E and t, the statistical error in w and, therefore, Δ. If the statistical error in Δ is reduced enough to more than compensate for the statistical error in the VMC energy, the use of correlation functions will be advantageous.

This method of directly computing energy differences was first implemented for H_2O, where Ψ_T is simply an SCF wave function.[124] To ascertain convergence to $\Psi_T \phi_0$, QMC energies and values of Δ are computed for several convergence times in a single calculation. Results are presented in Table 3.19 and plotted in Fig. 3.21; the time step is 0.0025 h^{-1}. The curve shown in the figure is obtained from a fit of the energies at $t = 0$ and $t \geq 0.25\ h^{-1}$ to

$$E(t) = E_0 + b\exp(-\gamma t) \ . \tag{3.20}$$

(The energy at $t = 0$ is included so that the fit yields a reasonable energy at this intercept. Including this point is found to have little effect on the accuracy of the fit at large t.) This equation arises because at large t,

$$\Phi(t) = \phi_0 + \langle\Psi_T|\phi_1\rangle\exp[-t(E_1 - E_0)]\phi_1 \ , \tag{3.21}$$

which gives for the energy, to first order in $|\langle\Psi_T|\phi_1\rangle|^2$,

$$E(t) \sim E_0 + (E_1 - E_0)|\langle\Psi_T|\phi_1\rangle|^2\exp[-2t(E_1 - E_0)] \ . \tag{3.22}$$

In agreement with Eq. (3.22), the data at large values of t, here large is assumed to be $\geq 0.25h^{-1}$, is fit reasonably well by a decaying exponential. More importantly for determining the QMC energy, convergence is obtained at approximately 1 h^{-1}. The converged QMC energies are also seen to be in excellent agreement with the previously computed QMC value of -76.377(7) h.[43] The statistical errors in the QMC energy and Δ generally increase as t increases, as implied by Eq. (3.19). The notable exception occurs at $t = 0.0625\ h^{-1}$ which is most likely due to the influence of excited states which have not decayed away. Calculations at smaller values of t, not shown here, reproduce this effect, and statistical error actually reaches a minimum between $t = 0.125$ and $0.250\ h^{-1}$.

The most significant fact revealed by the data is that the statistical errors do *not* scale with the magnitude of the quantity computed. Though the energy differences are over 100 times smaller in magnitude than the total energy, the statistical errors in E and Δ are of the same order of magnitude. Therefore, large increases in efficiency with this "difference" technique, versus the standard diffusion QMC approach, are not obtained. For example, the converged energy with the smallest statistical error is -76.3818(42) h, and the uncorrelated value is -76.3815(63) h. Apparently an increase in efficiency of a factor of 2.5 is gained. However, since including correlation functions generally reduces statistical error by at least a factor of 2-3, a QMC energy may be readily computed in the standard fashion, i.e., as described in Sec. 2.5, which

possesses a statistical error as small as that obtained in Δ. Therefore, for this H_2O test case, not using correlation functions does not give an increase in efficiency from the difference approach. Hopefully, including correlation functions in Ψ_T will improve the efficiency of this method.

The question of employing correlation functions is studied in detail for H_2. The MO consists of a Slater $1s$ orbital, $\zeta = 1.19$, on each atom and at the center of the bond. The linear coefficients are obtained in an SCF calculation. Computations are performed with and without E-E and E-N correlation functions and with and without satisfying the E-N cusp condition.

Results for the statistical errors, in micro-hartrees (*mih*), of E and $E(\Delta) = E_{\mathrm{VMC}} + \Delta$ are presented in Table 3.20. Several time steps are employed and the convergence time, 4 h^{-1}, has been ascertained to be sufficiently large to yield $\Psi_T \phi_0$. The statistical errors in Table 3.20 are (arbitrarily) normalized to correspond to one hour (Cray/XMP) of computation time as follows. For the SCF trial function, the statistical errors are straightforwardly normalized by $\sqrt{T_\Delta/60}$, where T_Δ is the computation time. For the correlated functions, the statistical error in the QMC energy is also normalized by the factor given above, i.e., it is assumed that only the QMC energy is computed for 60 minutes. However, the reported statistical error in the energy obtained from Δ, $E(\Delta) = E_{\mathrm{VMC}} + \Delta$, is found by considering how the combined statistical error of E_{VMC} and Δ is obtained most efficiently. That is, if T_{VMC} and T_Δ are the computation times yielding the statistical errors σ_{VMC} and σ_Δ, respectively, T'_{VMC} and T'_Δ ($T'_{\mathrm{VMC}} + T'_\Delta = 60$) are desired such that the resulting combined statistical error is a minimum. Therefore, the optimum computation times are obtained by

maximizing the efficiency of computing $E(\Delta)$. This efficiency is given by, cf. Sec. 3.3.1,

$$Eff = (T'_{\mathrm{VMC}} + T'_{\Delta})[\sigma'_{\mathrm{VMC}}{}^2 + \sigma'_{\Delta}{}^2] \ . \tag{3.23}$$

The optimum value of T'_{VMC} is thus obtained as,

$$T'_{\mathrm{VMC}} = 60\left[\frac{\sigma_{\Delta}}{\sigma_{\mathrm{VMC}}}\left[\frac{T_{\Delta}}{T_{\mathrm{VMC}}}\right]^{\frac{1}{2}} + 1\right]^{-1} , \tag{3.24}$$

and $T'_{\Delta} = 60 - T'_{\mathrm{VMC}}$. The resulting statistical errors in E_{VMC} and Δ are $\sigma'_{\mathrm{VMC}} = (T_{\mathrm{VMC}}/T'_{\mathrm{VMC}})^{\frac{1}{2}}\sigma_{\mathrm{VMC}}$ and $\sigma'_{\Delta} = (T_{\Delta}/T'_{\Delta})^{\frac{1}{2}}\sigma_{\Delta}$, respectively, from which the combined statistical error, $\sigma_E(\Delta)$, is obtained. In practice, such optimization is not difficult and may be accomplished by estimating σ_{VMC} and σ_{Δ} in small calculations before proceeding to the computation of precise results. For this reason, $\sigma_E(\Delta)$ is optimized when comparing with σ_E.

In comparing the correlated trial functions, we see that satisfying the E-N cusp condition most often yields larger statistical errors. This result is caused primarily by the use of a bond function which necessitates a value of λ(cusp) for this "nucleus" which is unphysical – large in magnitude and of negative sign. Thus, a trial function of rather poor overall quality is obtained, though the time-step bias is reduced, and statistical errors are large.

The relative efficiency of employing correlation functions in computing $E(\Delta)$ is examined by comparing $\sigma_E(\Delta)$ between the correlated (with the E-N cusp condition not satisfied) and the SCF trial function. Table 3.20 demonstrates that at the larger time steps, 0.04 and 0.05 h^{-1}, not including correlation is actually more efficient, while at $\tau = 0.02$ and 0.01 h^{-1}, the reverse is true. At the larger time steps, the QMC walk

is most efficient and so the cost of a VMC calculation is relatively high. Therefore, the reduction of σ_Δ is more than offset by the added cost of the computing the VMC energy when employing correlation functions. However, at smaller time steps the QMC walk is less efficient and the relative cost of a VMC calculation is small, so that reducing σ_Δ using correlation functions is advantageous.

The question of whether computing the energy by difference yields smaller statistical errors than with the standard approach is answered by comparing values of $\sigma_E(\Delta)$ to σ_E. For all the trial functions used here, and at all values of τ, the statistical error is reduced when the energy is computed as $E(\Delta) = E_{\mathrm{VMC}} + \Delta$. (The exception at $\tau = 0.04\ h^{-1}$ is not considered here since it appears to be statistically anomalous, i.e., $\sigma_E = 55\ mih$ is less than σ_E obtained with the same trial function at a larger time step.) However, the reductions are once again not large considering the small magnitude of Δ. Furthermore, the best single approach is not evident. While computing the energy by difference is most efficient at $\tau = 0.04$ and $0.05\ h^{-1}$ without correlation functions, this method becomes less efficient than a standard calculation [with correlation functions and $\lambda \neq \lambda(\mathrm{cusp})$] at smaller time steps. Employing correlation functions to compute Δ does yield $\sigma_E(\Delta) < \sigma_E$ at small time steps, but by $\tau = 0.05h^{-1}$ the statistical errors are roughly the same. Therefore, in the general case, computing the QMC energy as $E_{\mathrm{VMC}} + \Delta$ may be less efficient, depending on the value of τ and the use of correlation functions.

The final system we have studied using the difference approach is the nitrogen atom. A DZ and HF basis set[89] are employed to determine how much of the fixed-node error results from using a small (DZ) basis set, cf. Sec. 3.3.3. Since the goal is

to compute an unbiased energy most efficiently, a small time step (0.0025 h^{-1}) is chosen and statistical errors are compared at convergence times for which $\Psi_T \phi_0$ is obtained, 0.80 and 0.96 h^{-1} for the DZ and HF trial functions, respectively. Also, the results for H_2 indicate, since τ is small, that the greatest efficiency will be obtained when correlation functions are employed. Finally, $\lambda = \lambda(\text{cusp})$ so that the time-step bias actually is small at $\tau = 0.0025\ h^{-1}$. Correlation function parameters are reported in Table 3.21.

The value of λ(cusp) is obtained somewhat differently than as described in Sec. 3.1. In Sec. 3.1, λ(cusp) is chosen as the mode of the distribution of values which satisfy Eq. (3.7). Below, λ(cusp) is the average value, over $|\Psi_T|^2$, of λ which satisfies the E-N cusp condition.[125] That is,

$$\lambda(\text{cusp}) = -\frac{1}{n}\int d\underline{R}\left[\Psi_T(\underline{R})\sum_{i=1}^{n}\frac{\partial\Psi_T(\underline{R})}{\partial r_i}\right]_{r_i=0} - Z\ , \tag{3.25}$$

where n is the number of electrons and Ψ_T is assumed to be normalized. So that the integration may be performed analytically, the correlation functions of Ψ_T are omitted in Eq. (3.22) yielding

$$\lambda(\text{cusp}) = -\frac{1}{2}\left[\rho^{-1}(\underline{r})\frac{\partial\rho(r)}{\partial r}\right]_{r=0} - Z\ , \tag{3.26}$$

where ρ is the one-electron density function of the Slater determinant. This approach is slightly more convenient since λ(cusp) is obtained directly from the Slater determinant. The difference between the mode value and the average value of λ(cusp) is small in this case, 0.00874, cf. Tables 3.11 and 3.21. The QMC energies computed with the two values of λ(cusp) are compared in Table 3.22, and the mode value is seen to yield a slightly smaller time-step bias.

The statistical errors σ_E and $\sigma_E(\Delta)$ are reported for each trial function in Table 3.23. These errors are normalized to 60 minutes of computation time on a Cray/XMP and obtained in the same way as those for H_2. As seen from Table 3.23, the difference approach fails to yield a significant improvement in comparison to standard QMC. The statistical error in the weights is such that, by the time convergence to $\Psi_T \phi_0$ is obtained, the correlation between E_{VMC} and E is reduced to the point that computing the energy by difference is not useful.

As the final point of interest, it is noted that these calculations gave very similar QMC energies implying that the fixed-node error is due primarily to the use of a single determinant.

Given the results of H_2, N, and H_2O, computing energies by difference does not yield significant increases in efficiency for the trial functions studied. Also, when reductions in statistical error are observed, they are dependent on the time step and the use of correlation functions. It still remains to be determined, however, whether different classes of trial functions (i.e., more highly optimized) will be more useful in the difference approach. Finally, when computing $E(\Delta)$ using correlation functions in Ψ_T, the relative amount of computation of E_{VMC} and Δ should be optimized as described in this section.

Table 3.1. Trial function basis sets.[a]

Importance Function	Outer H's STO	Outer H's ζ	Central H STO	Central H ζ	Bond STO	Bond ζ
Ψ_1	1*s*	1.040	1*s*	1.275	---	---
Ψ_2 (UHF)	1*s* 1*s*′	0.925 1.250	1*s*	1.120	---	---
Ψ_3	1*s* 1*s*′	0.925 1.275	1*s* 1*s*′	0.925 1.275	---	---
Ψ_4	1*s* 1*s*′	0.925 1.275	1*s*	1.120	---	---
Ψ_5[b]	1*s* 1*s*′	0.925 1.250	1*s*	1.120	1*s*	1.175
Ψ_6[c]	1*s* 1*s*′ 2*s* 2*s*′ 2*p* 2*p*′ 3*p* 3*d*	1.000 2.200 1.000 2.200 1.700 2.900 2.900 2.700	1*s* 1*s*′ 2*s* 2*s*′ 2*p* 2*p*′ 3*p* 3*d*	1.000 2.200 1.000 2.200 1.700 2.900 2.900 2.700	---	---

[a] For $\Psi_1 - \Psi_5$, the exponents ζ are hand-optimized using HONDO to obtain a minimum in the SCF energy. For Ψ_3 exponents of the STO's on all the H atoms are constrained to be equal.

[b] A "bond" function is placed at ±0.85 bohr from the central H.

[c] Ref. 65.

Table 3.2. Correlation function parameters.[a]

Correlation Function	Parameters	Ψ_3	Other Ψ_i's
U_1	a_1	0.50	0.50
	b_1	1.00	2.00
	a_2	1.12	0.75
	b_2	1.13	1.00
U_2	λ	0.10	0.10
	ν	0.50	0.50

[a] The parameter a_1 is chosen to satisfy the electron-electron cusp condition for electrons of like spin. The remaining parameters are hand-optimized to minimize the VMC energy.

Table 3.3. SCF, VMC and QMC barriers.[a]

Importance Function	E_B (SCF)	E_B (VMC)	E_B (QMC)[b]	E_B (QMC)[c]
Ψ_1	60.9	36.0 (0.5)	13.40 (0.40)	12.2 (0.50)
Ψ_2	47.0	27.4 (0.2)	11.16 (0.48)	10.20 (0.26)
Ψ_3	53.5	27.8 (0.3)	---	9.70 (0.22)
Ψ_4	53.4	25.9 (0.2)	11.67 (0.38)	9.86 (0.20)
Ψ_5	51.6	24.4 (0.2)	10.39 (0.40)	9.70 (0.13)
Ψ_6	50.3	23.6 (0.4)	---	9.90 (0.24)
HF Limit[d]	50.1	---	---	---

[a] All barriers are obtained by subtracting the exact H + H_2 energy from the computed H_3 saddle-point energy. In Tables 3.3 and 3.4 energy differences are in kcal/mol, and, unless stated otherwise, numbers in parentheses are one standard deviation in the average (statistical error).

[b] All block times are $0.5h^{-1}$, except for Ψ_5 ($1h^{-1}$).

[c] All block times are $5h^{-1}$.

[d] Derived from Ref. 65.

Table 3.4. Comparison of best QMC barrier height with estimates of the exact.

Method	E_B
Best CI[a]	< 9.86 $\approx 9.59 \pm 0.06$
QMC (Ψ_5)	< 9.70 (0.13)
Exact[b]	9.65 (0.08)

[a] Ref. 67.

[b] Ref. 41.

Table 3.5. Trial function properties.

Importance Function	Computation Time (CPU)[a]	QMC statistical error (σ)	Relative Efficiency[b]
Ψ_1	35	0.38	1.00
Ψ_2	40	0.37	0.92
Ψ_3	44	0.31	1.20
Ψ_4	39	0.31	1.35
Ψ_5	50	0.25	1.62
Ψ_6	77	0.25	1.05

[a] The time (in scaled units) required to obtain the statistical error, third column, in the QMC energy.

[b] The relative efficiency of Ψ_i is given by $CPU_1\sigma_1^2 / CPU_i\sigma_i^2$.

Table 3.6 Trial function parameters for F and F^-.

		F	F^-
Exponents[a]	$1s$	10.42450	10.200
	$1s'$	7.66585	7.250
	$2s$	3.13578	3.100
	$2s'$	1.94456	1.700
	$2p$	4.18389	4.125
	$2p'$	1.85602	1.650
Correlation function parameters[b]	a_1	0.50	0.5000
	a_2	0.00	11.5000
	b_1	3.00	2.8750
	b_2	0.00	33.0625
	λ	0.10	0.1500
	ν	0.10	0.2000

[a] The exponents and linear coefficients are from Ref. 89. For F^- the exponents were hand-optimized for a minimum in the SCF energy computed using HONDO[68], the linear coefficients are from the SCF calculation.

[b] The parameter a_1 is chosen to satisfy the like-spin E-E cusp condition. The remaining parameters are hand-optimized for a minimum in the VMC energy.

Table 3.7. Theoretical and experimental energies of F and F^-, and the electron affinity.

	METHOD	ENERGY (h)		EA (eV)
		F	F^-	
Ab initio non-variational	Atomic Bethe-Goldstone	-99.7131[a]	-99.8411[a]	3.47[a]
				3.37[b]
	Perturbation Theory	-99.7299[c]	-99.8595[c]	3.53[c]
		-99.6750[d]	-99.8025[d]	3.47[d]
	Coupled-Cluster[e]	-99.7128	-99.8368	3.37
ab initio variational	HF[f]	-99.4093	-99.4594	1.36
	CI[g]	-99.7166	-99.8312	3.12
	MCSCF+CI	-99.6202[h]	-99.7369[h]	3.18[h]
		-99.62350[i]	-99.73722[i]	3.09[i]
	QMC[j]	-99.7005(21)	-99.8273(34)	3.45(11)
	Empirical	-99.7313[k]	-99.8574(30)[l]	3.43(8)[m]
	Experimental	-99.8059[n]	---	3.399(3)[o]

[a] Ref. 77.
[b] Ref. 78.
[c] Ref. 80.
[d] Ref. 84.
[e] Ref. 85.
[f] Ref. 79.
[g] Ref. 13.
[h] Ref. 82.
[i] Ref. 83.
[j] This work.
[k] From experimental results corrected for relativistic effects in Ref. 76. Ref. 13 corrects an error in the sign of the Lamb shift, resulting in the energy given here.
[l] Ref. 71.
[m] This number is the difference of the empirical non-relativistic energies for F and F^-.
[n] Ref. 86.
[o] Recommended experimental value of Ref. 88.

Table 3.8. QMC energies of F and F^- versus time step.

Time-step (h^{-1})	ENERGY (h)	
	F	F^-
0.00500	-99.7196(17)	-99.9141(28)
0.00250	-99.7106(20)	-99.8769(28)
0.00125	-99.7050(17)	-99.8515(27)
0.00050	---	-99.8375(31)
0.0[a]	-99.7005(21)	-99.8273(34)

[a] Obtained from a linear fit of the energies computed at the three smallest time steps.

Table 3.9. Values of λ and λ(cusp) for F and F^-.

Species	λ	λ(cusp)[a]	$\lambda - \lambda$(cusp)
F	0.1000	0.0640	0.0360
F^-	0.1500	0.0165	0.1335

[a] Computed as described in Sec. 3.2.

Table 3.10 Trial function parameters for N and N_2.

	Orbital	N	N_2
Exponents[a]	$1s$	8.49597	8.49597
	$1s'$	5.98644	5.98644
	$2s$	2.26086	2.26086
	$2s'$	1.42457	1.42457
	$2p$	3.24933	3.24933
	$2p'$	1.49924	1.49924
Correlation function parameters	a_1	0.5	0.5
	b_1	3.0	3.0
	λ	0.2	0.2
	ν	0.3	0.3

[a] The exponents and linear coefficients are from Ref. 89.

Table 3.11 QMC energies for N and N_2 with and without $\lambda = \lambda(\text{cusp})$.[a]

	N		N_2	
τ	E	$E(\lambda(\text{cusp}))$	E	$E(\lambda(\text{cusp}))$
0.01000	-54.6317(18)	-54.5743(15)	-109.5575(79)	-109.4769(69)
0.00750	---	---	-109.5441(63)	-109.4695(59)
0.00500	-54.6030(23)	-54.5702(9)	-109.5231(39)	-109.4769(69)
0.00250	---	-54.5668(13)	-109.5026(35)	---
0.00125	-54.5794(12)	---	-109.4916(36)	---
0.00060	-54.5745(13)	---	---	---
0.0[b]	-54.5716(13)	-54.5648(16)	-109.4835(37)	-109.462(14)
Hartree-Fock	-54.4009[c]		-108.9939[d]	
Exact	-54.5895[e]		-109.535[f]	

[a] For N, $\lambda(\text{cusp}) = 0.0538$, and for N_2, $\lambda(\text{cusp}) = 0.0505$.

[b] Obtained from a linear fit.

[c] Ref. 89.

[d] Ref. 105.

[e] Esimated from experiment in Ref. 13.

[f] Esimated from experiment in Ref. 106.

Table 3.12 Excited-state energies of He.[a] For the $1s\,3s$ and $1s\,4s$ states, calculations were performed only in regions of a given sign[Ψ_T], e.g., $\Psi_T > 0$, restricting the number of nodal volumes sampled.

	$1s\,2s\ ^1S$
$\langle\Psi_T\|H\|\Psi_T\rangle$[b]	-2.14307
QMC(τ)	
0.050	-2.14490(6)
0.025	-2.14485(7)
0.010	-2.14480(5)
0.000[c]	-2.14478(6)
Exact[d]	-2.14598

	$1s\,3s\ ^1S$	
$\langle\Psi_T\|H\|\Psi_T\rangle$[b]	-2.06036	
QMC(τ)		$\Psi_T > 0$
0.050	-2.06026(11)	-2.06144(8)
0.020		-2.06122(7)
0.005		-2.06108(17)
0.000[c]		-2.06105(17)
Exact[d]		-2.06128

	$1s\,4s\ ^1S$	
$\langle\Psi_T\|H\|\Psi_T\rangle$[b]	-2.03320	
QMC(τ)	$\Psi_T > 0$	$\Psi_T < 0$
0.050	-2.03026(11)	-2.06144(8)
0.025		-2.03122(7)
0.000[c]		-2.03105(17)
Exact[d]		-2.03358

[a] All time steps are in h^{-1} and energies in h.

[b] Ref. 107.

[c] Obtained from a linear fit.

[d] Ref. 108.

Table 3.13. Percent energy recovered by QMC.

State	$\%E_{rec}$
$1s\,2s\ ^1S$	58.8(2)
$1s\,3s\ ^1S$	75(18)
$1s\,4s\ ^1S$	2(3)

[a] $\%E_{rec} \equiv (E_{\mathrm{QMC}} - E_{exact})/(E_T - E_{exact})$, $E_T = \langle\Psi_T|H|\Psi_T\rangle$.

Table 3.14. Single-determinant QMC energies for selected first-row atoms and molecules.[a]

Species	E_{HF}	E_{exact}	E_{QMC}	%CE
Li	-7.43273[b]	-7.47807[c]	-7.47809(24)	100.0(5)
Li(2^2P)	-7.36507[d]	-7.41016[e]	-7.41031(22)	100.3(5)
LiH	-7.98735[f]	-8.07019[g]	-8.06908(43)	98.7(5)
Li_2	-14.872[h]	-14.9945[i]	-14.9923(8)	98.2(8)
Be	-14.57302[b]	-14.6673[j]	-14.6566(7)	88.7(7)
BH	-25.1314[k]	-25.2829[l]	-25.2650(11)	88.2(7)
H_2O	-76.0675[m]	-76.4376[m]	-76.4017(15)	90.3(4)

[a] All energies are in h and computed at the experimental geometry. Electronic ground states are computed unless stated otherwise.

[b] Ref. 89.

[c] Ref. 116.

[d] Ref. 111.

[e] Ref. 12.

[f] Ref. 117.

[g] Computed here by adding D_e = 0.09243 h [102] to the finite mass energies of Li and H. The resulting finite mass energy of LiH is then converted to the infinite mass (Born-Oppenheimer) value reported above.

[h] Ref. 118.

[i] Obtained by the same method as the LiH energy reported above with D_e = 0.03881 h [119].

[j] Ref. 120.

[k] Ref. 121.

[l] Ref. 122.

[m] Ref. 123.

Table 3.15. BH, trial function parameters of Ψ_1.

Orbital	ζ	MO Linear Coefficients		
		ψ_1	ψ_2	ψ_3
$1s$(B)	4.00000	-0.62188	0.15929	-0.14387
$1s'$(B)	5.75000	-0.39038	0.02268	-0.01033
$2s$(B)	1.35000	0.00355	-0.57974	0.76663
$2p_x$(B)	0.91077	0.00439	0.01103	-0.28741
$2p_x'$(B)	1.50000	-0.00506	-0.21690	-0.33636
$1s$(H)	1.18274	-0.00244	-0.50324	-0.37319
$1s'$(H)	2.90014	0.00044	0.00636	0.01699
$2p_x$(H)	1.70000	-0.00051	0.04439	0.01163

Correlation function parameters

a_1	0.5	λ_B	0.15
b_1	2.0	ν_B	2.00
		λ_H	0.02
		ν_H	0.20

Table 3.16. BH, trial function parameters of Ψ_2.

Atom	ψ_1		ψ_2		ψ_3	
Orbital	ζ	c	ζ	c	ζ	c
$1s$ (B)	4.02586	-0.55410	4.00922	0.09996	4.00035	-0.13592
$1s'$(B)	5.81396	-0.49116	5.75058	0.02154	5.75015	-0.01402
$2s$ (B)	1.35016	0.00455	1.47658	-0.61542	1.28784	0.76316
$2p_x$ (B)	0.91076	0.00324	0.91062	0.01059	0.89497	-0.29426
$2p_x'$ (B)	1.50008	-0.00975	1.50666	-0.26620	1.54825	-0.36945
$1s$ (H)	---	---	1.17272	-0.46348	1.17524	-0.35145
$1s'$(H)	---	---	2.90011	0.00653	2.90030	0.01977
$2p_x$ (H)	1.69990	-0.00517	1.69911	0.03565	1.69986	0.01068

Correlation function parameters

a_1	0.66667	λ_B	0.15
b_1	0.25000	ν_B	2.00
a_2	1.00300	λ_H	0.02
		ν_H	0.20

Table 3.17. QMC energies of Ψ_1 and Ψ_2.[a]

Trial function	E (deletion)	E (rejection)
Ψ_1	-25.2663(11)	-25.2667(11)
Ψ_2	---	-25.2651(10)

[a] Energies are in h. Calculations were performed at several times steps yielding energies within the statistical errors shown, but direct comparisons were made only at $\tau = 0.01\ h^{-1}$.

Table 3.18. QMC energies of H_2O versus time step.

τ (h^{-1})	E (h)
0.0100	-76.4276(29)
0.0075	-76.4215(18)
0.0050	-76.4116(19)
0.0025	-76.4015(27)
0.0010	-76.4018(18)
0.0000[a]	-76.4017(15)

[a] Since the τ–dependence of the QMC energy is not observable at $\tau \leq 0.0025\ h^{-1}$, the $\tau = 0$ estimate is obtained from a weighted (by σ^{-2}) average of the $\tau = 0.0025$ and $0.0010\ h^{-1}$ energies.

Table 3.19. QMC energy of H_2O by the difference method versus convergence time.

t (h^{-1})	E (h)	$E(\Delta)$ $(h)^a$
0.0000	-75.9693[b]	-75.9693
0.0625	-76.1655(64)	-76.1653(39)
0.1250	-76.2245(53)	-76.2244(25)
0.2500	-76.2912(51)	-76.2911(28)
0.5000	-76.3482(57)	-76.3485(40)
1.1250	-76.3815(63)	-76.3818(42)
1.5000	-76.3852(87)	-76.3857(68)
2.2500	-76.3763(80)	-76.3763(78)

[a] Small differences between the standard estimates of the QMC energy, E, and these results arise from the sampled VMC energy being slightly different than -75.9693 due to statistical error.

[b] Ref. 124.

Table 3.20. Statistical errors of H_2 QMC energies.

	SCF trial function		E-N cusp condition satisfied		E-N cusp condition not satisfied	
τ	σ_E	$\sigma_E(\Delta)$	σ_E	$\sigma_E(\Delta)$	σ_E	$\sigma_E(\Delta)$
0.05	106	50	73	73	72	68
0.04	125	62	55	80	84	66
0.03	124	83	117	89	93	83
0.02	176	121	115	99	102	77
0.01	261	119	134	128	183	94

[a] Statistical errors are in micro-hartrees and time steps are in h^{-1}.

Table 3.21. Correlation function parameters of N.

Basis set	λ	ν	a_1	b_1
DZ	0.06254	0.07500	0.5	3.0
HF	-0.22099	0.15000	0.5	3.5

Table 3.22. QMC energies of N with different values of λ(cusp)[a].

τ	E [λ(cusp) mode][b]	E [λ(cusp) average][b]
0.0100	-54.5743(15)	-54.5828(17)
0.0050	-54.5702(9)	-54.5765(17)
0.0025	-54.5668(13)	-54.5682(19)

[a] The units for the energy and time step are h and h^{-1}, respectively.

[b] The mode value is 0.05380 and the average value is 0.06254.

Table 3.23. Statistical errors in the QMC energies of N.[a]

Basis set	σ_E	$\sigma_E(\Delta)$
DZ	1.5	1.6
HF	2.9	2.3

[a] Statistical errors, in milli-hartrees, are obtained as described in the text and correspond to 1 hour of Cray/XMP computer time.

Figure 3.1. QMC energies with and without acceptance/rejection. Since time-step bias is eliminated at $\tau = 0$ for both sets of calculations, energies should converge to the exact (solid line) value of -0.5 h. For Figs. 1 and 2, $\Psi_T = \exp(-1.05r)$ and statistical errors, representing one standard deviation in the mean of a Gaussian distribution, are no larger than the symbol sizes.

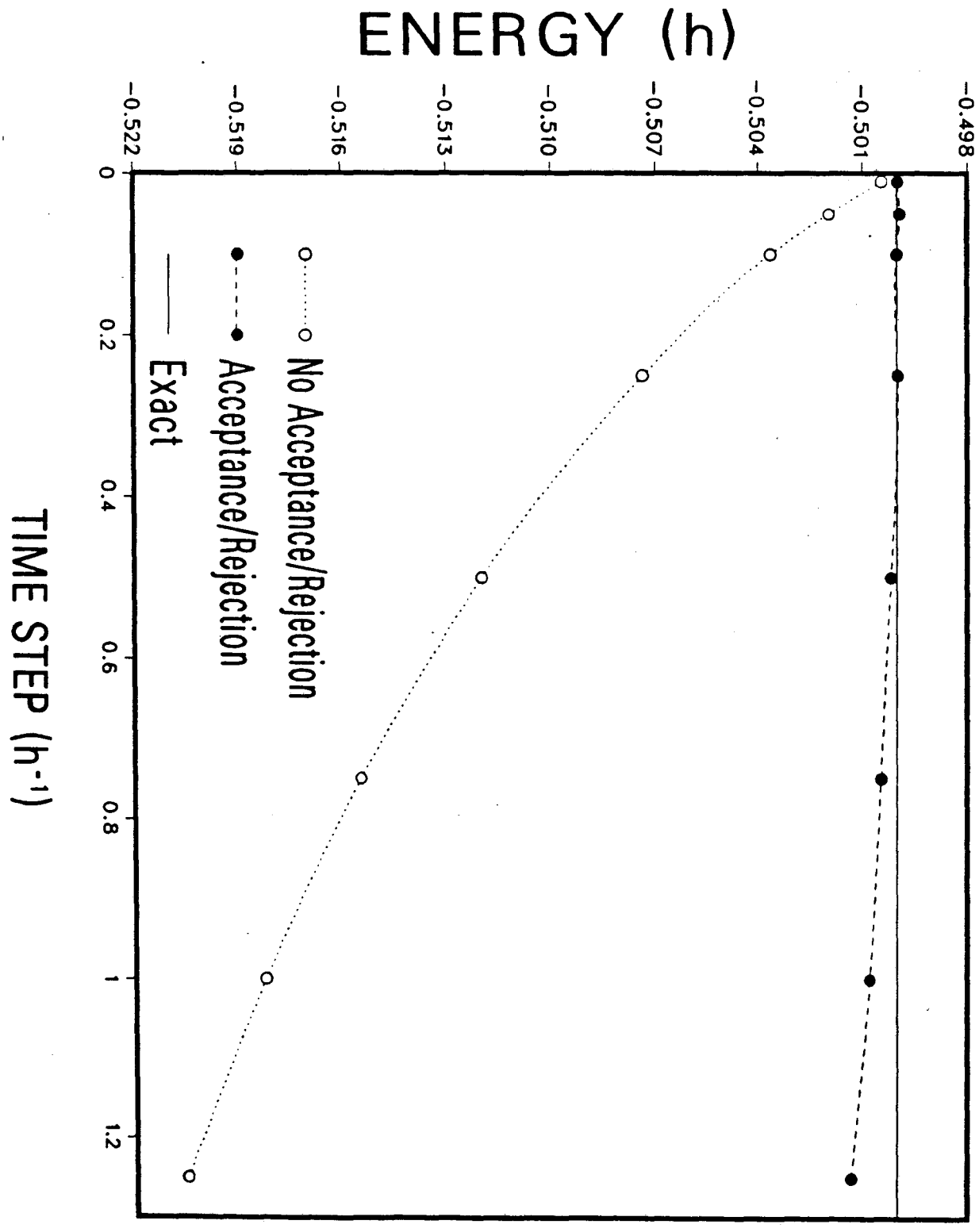


Figure 3.2. Comparison of r_M and r^2_M with and without acceptance/rejection. Results are plotted for mixed expectation values, $A_M \equiv \langle\Psi_T|A|\phi\rangle/\langle\Psi_T|\phi\rangle$.

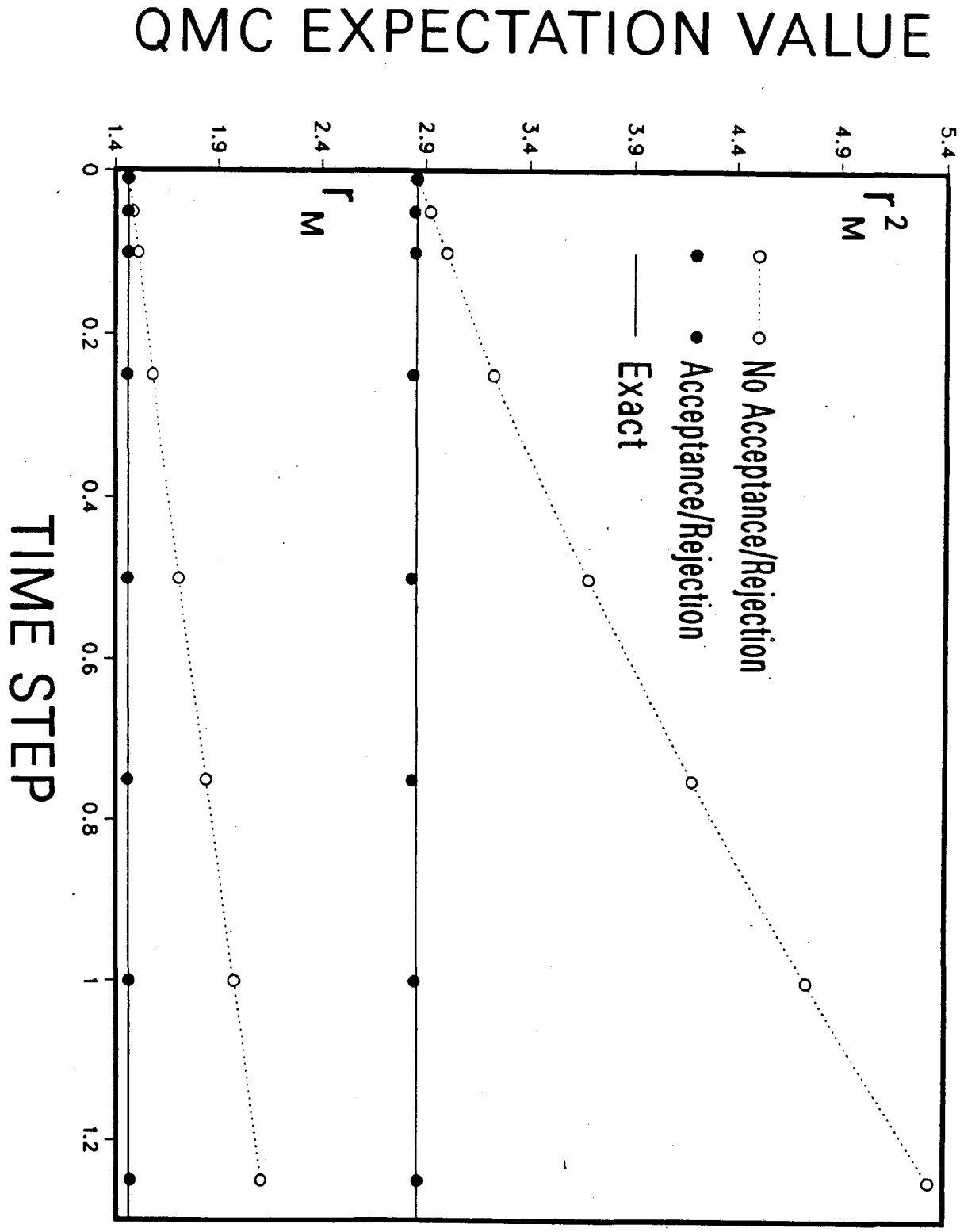


Figure 3.3. Time-step behavior of H atom energies. QMC energies are computed over a range of time steps for several exponents, ζ, in the Slater 1*s* orbital employed as a trial function. Statistical errors, when not shown explicitly, are no larger than the point sizes.

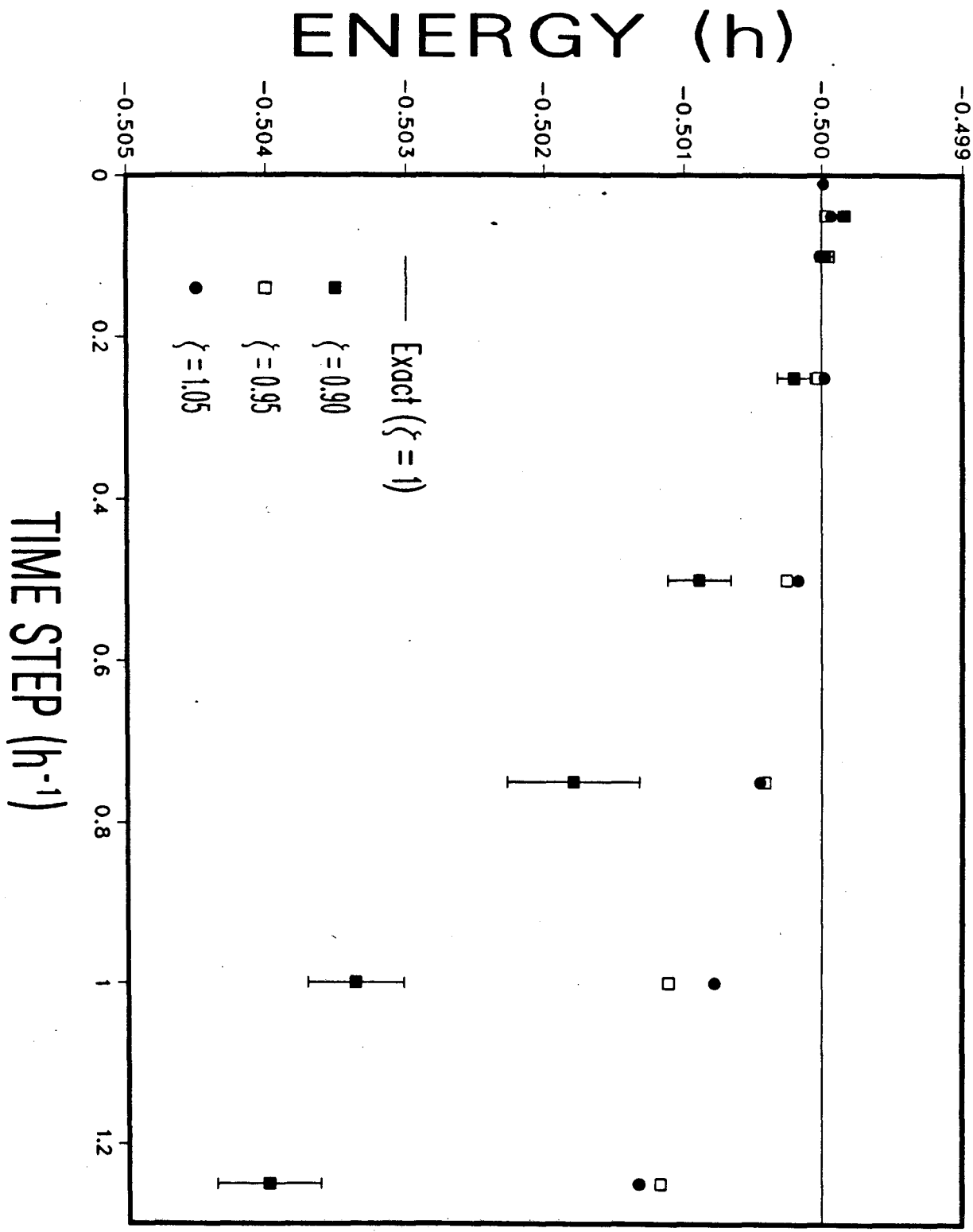


Figure 3.4. Time-step behavior of H-atom energies using poor trial functions.

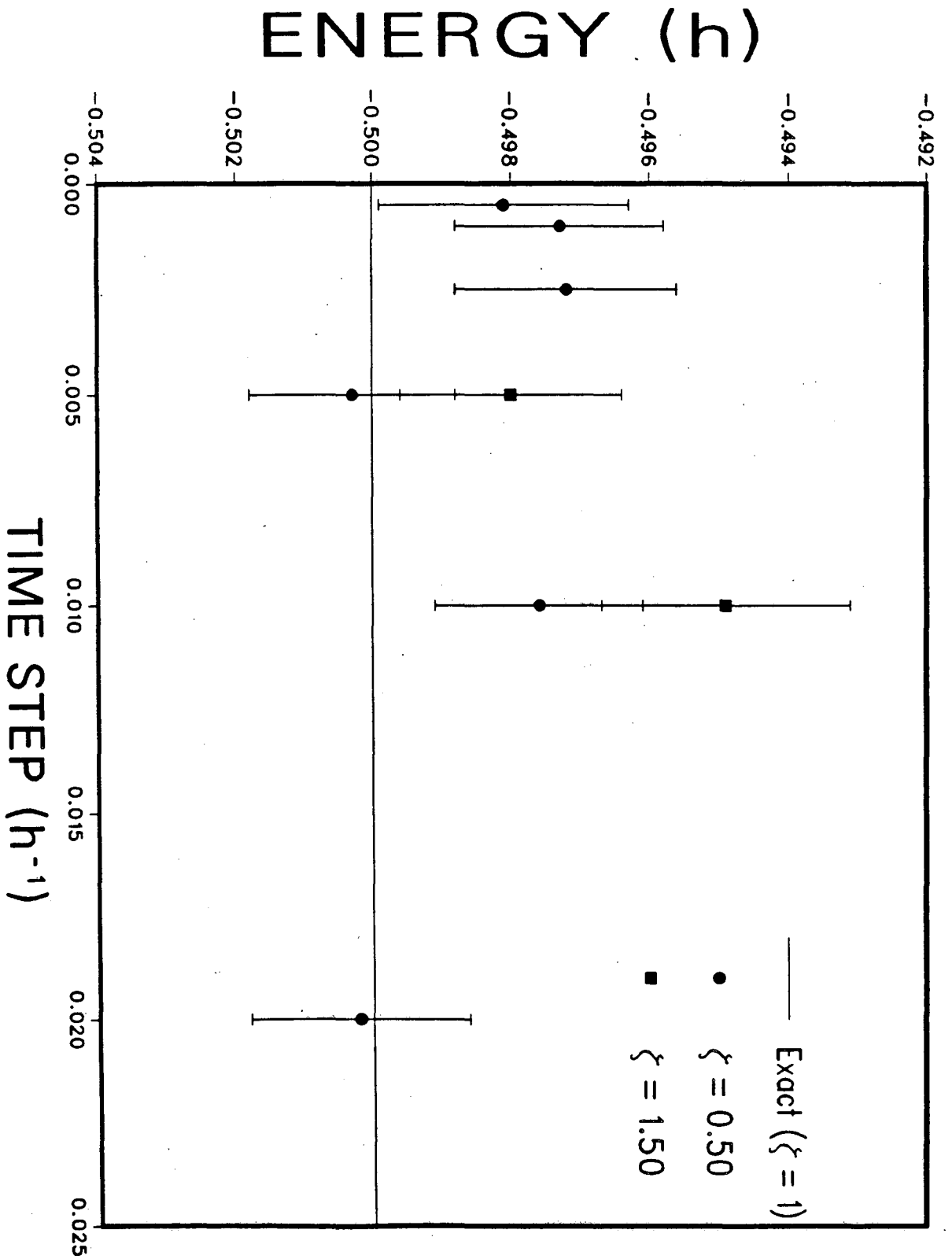


Figure 3.5. QMC energies of the $2p_x$ state of H using different algorithms. For all calculations the nodal boundary condition is enforced by deleting any walker which attempts to cross a node. In Figs. 5 and 6 the trial function is a $2p_x$ Slater orbital with $\zeta = 0.55$. For Figs. 5-7 "CRC" refers to not considering the possibility of walkers crossing and recrossing a node, "ITERATIVE" refers to Anderson's method of reducing time-step bias, "∇ CRC" indicates that the gradient of the trial function is used to approximately locate the node in computing the cross-recross probability, and "NO CRC" means that the cross-recross probability is computed exactly.

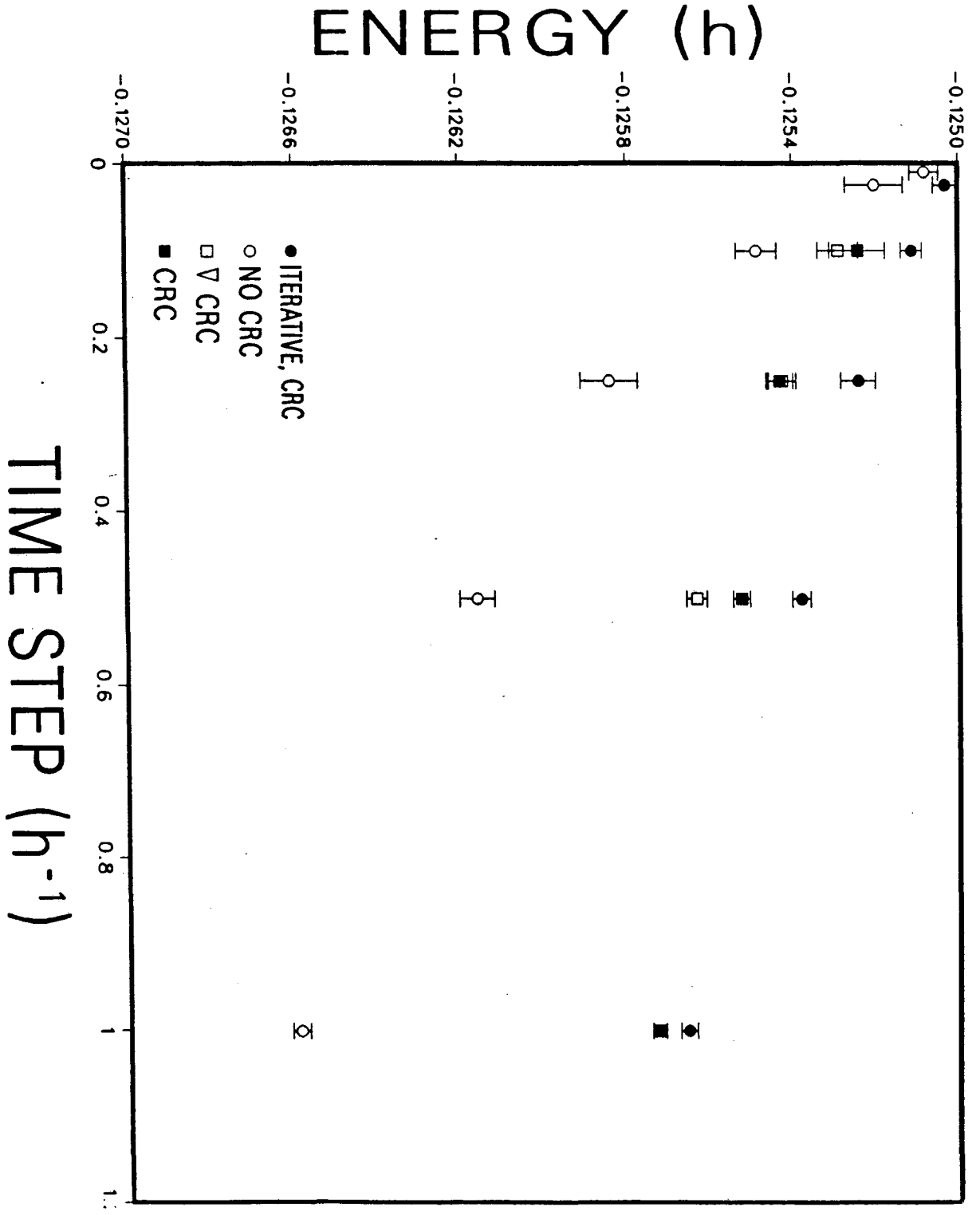
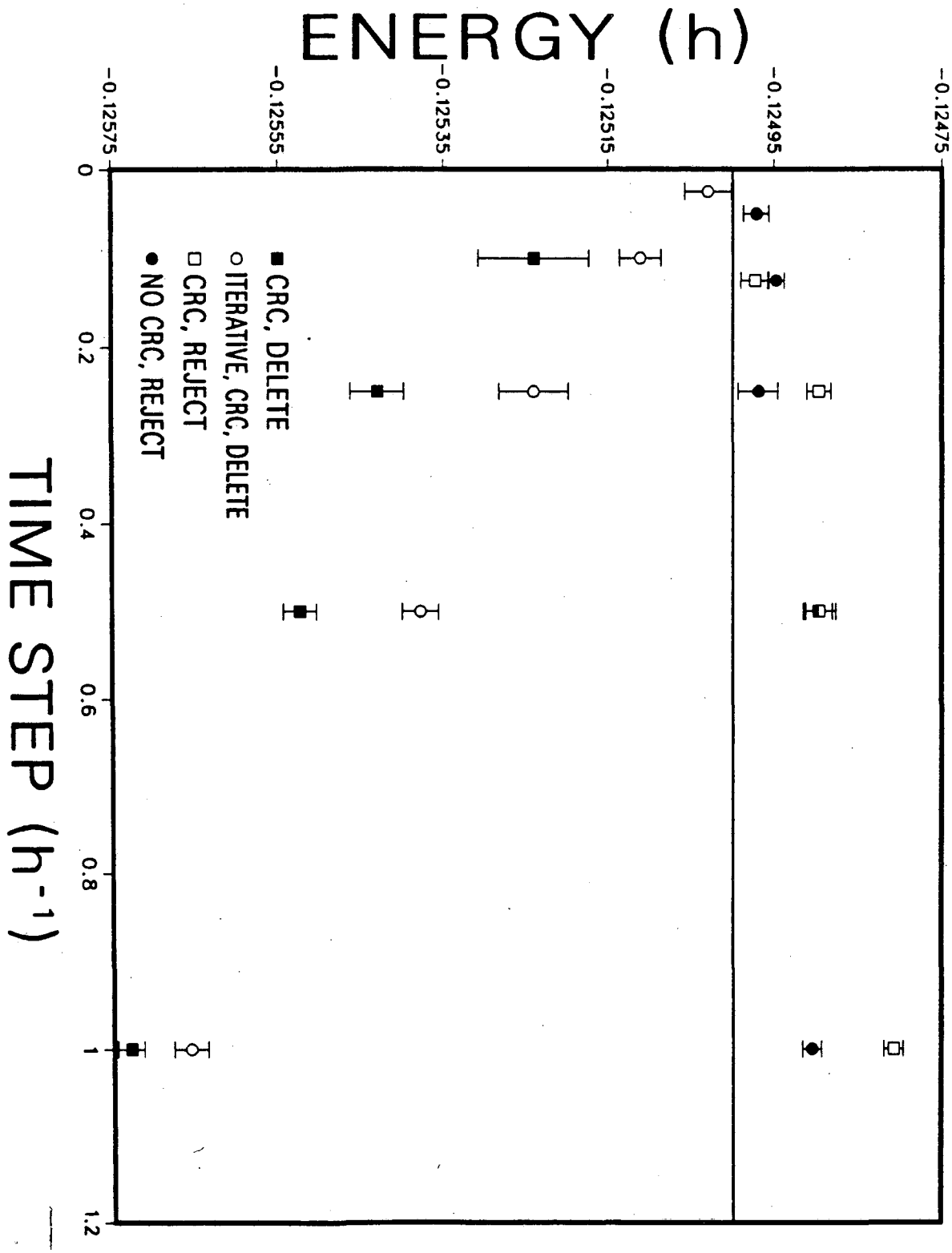


Figure 3.6. Rejection versus deletion for the $2p_x$ state. QMC energies are compared for the cases of rejecting a move ("REJECT") and deleting a walker ("DELETE") upon an attemtped crossing of the node.



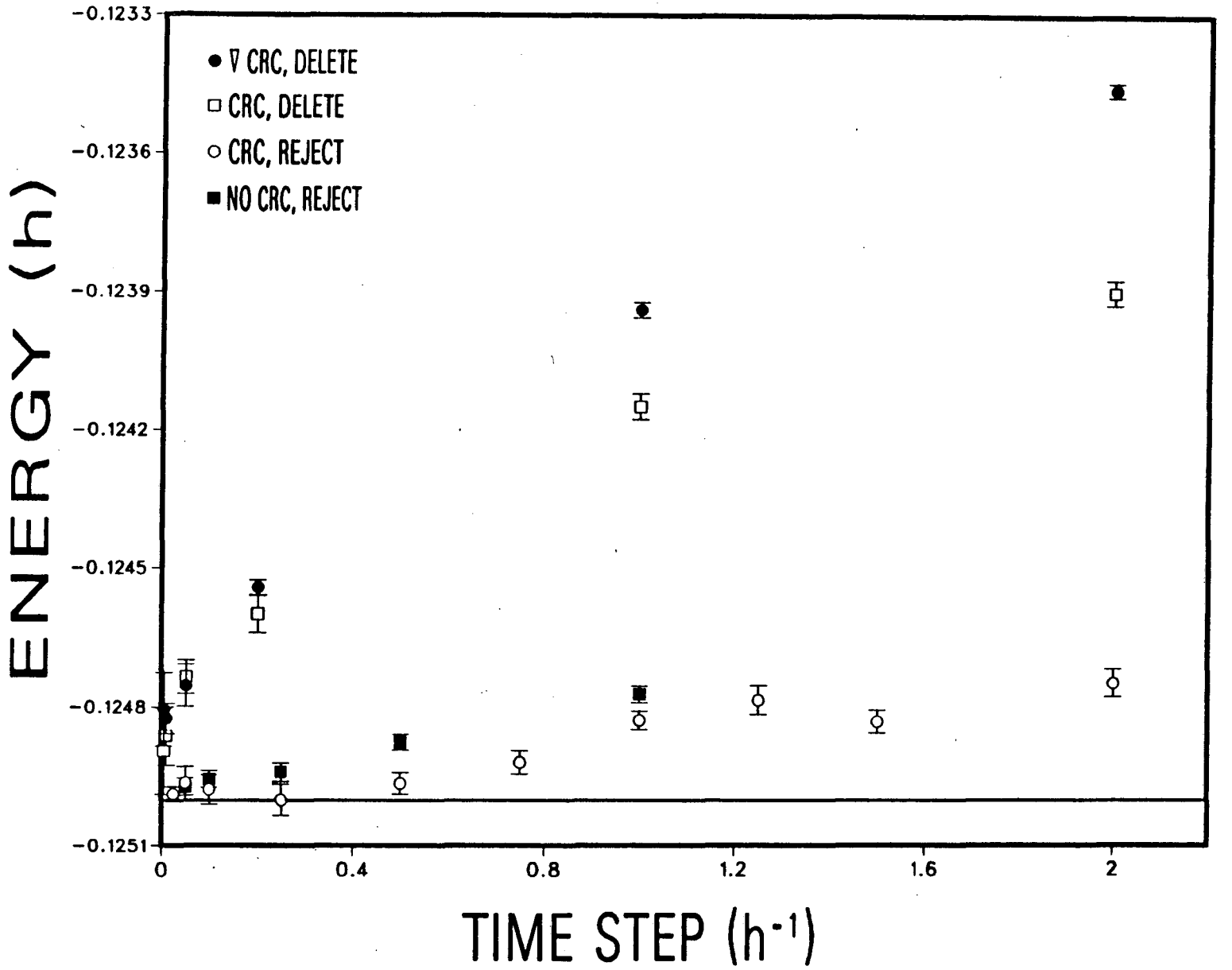


Figure 3.7. Rejection versus deletion for the $2p_x$ state, $\zeta = 0.45$.

Figure 3.8. Values of λ which satisfy the electron-nuclear (E-N) cusp condition for LiH. For a point in configuration space, the value of λ_α which satisfies the E-N cusp condition is given by, $\lambda_\alpha = (\frac{1}{2} r_{i\alpha} \psi^{-1} \nabla_i^2 \psi)_{r_{i\alpha}=0} + Z_\alpha$, see text. Here, ψ represents the determinant constructed from the MOs occupied by electrons of a given spin. Namely, for LiH, $\psi = [\psi_1(1)\psi_2(2) - \psi_1(2)\psi_2(1)]$ if $i = 1, 2$ or $\psi = [\psi_1(3)\psi_2(4) - \psi_1(3)\psi_2(4)]$ if $i = 3, 4$. Values of λ_α are computed at points sampled from $|\Psi_T|^2$ with $r_{i\alpha} < r_{min}$. The cutoff, r_{min}, has little effect, and is chosen large enough to yield several hundred values of λ_α. The distributions below indicate that the E-N cusp condition is best satisfied on average for $\lambda_{Li} = 0.031$ and $\lambda_H = 0.061$.

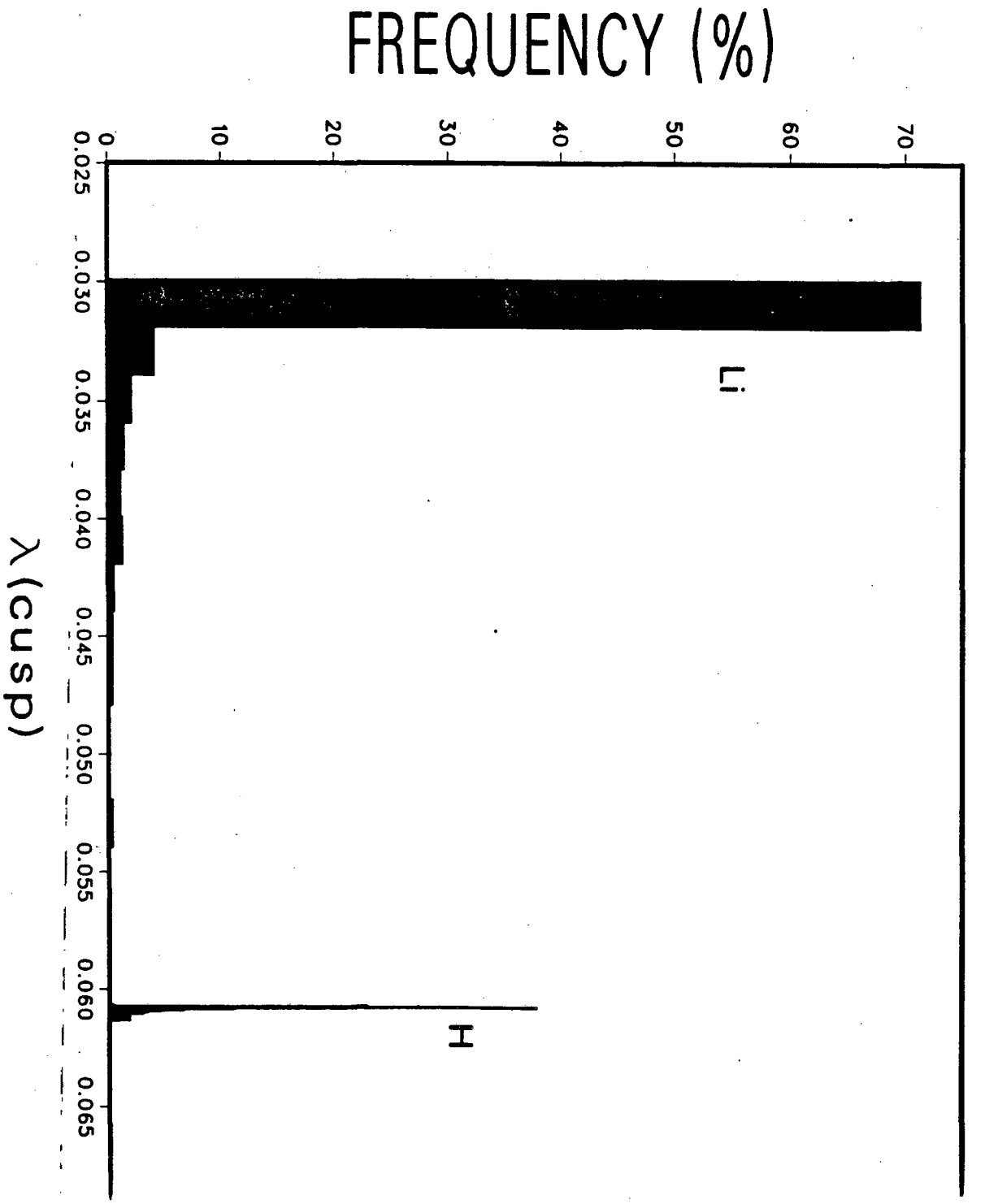


Figure 3.9. Effect of satisfying the electron-nuclear (E-N) cusp condition for He. In Figs. 3.9 and 3.10 open circles signify that the E-N cusp condition is satisfied, and closed circles indicate that it is not. The straight line represents the exact energy.

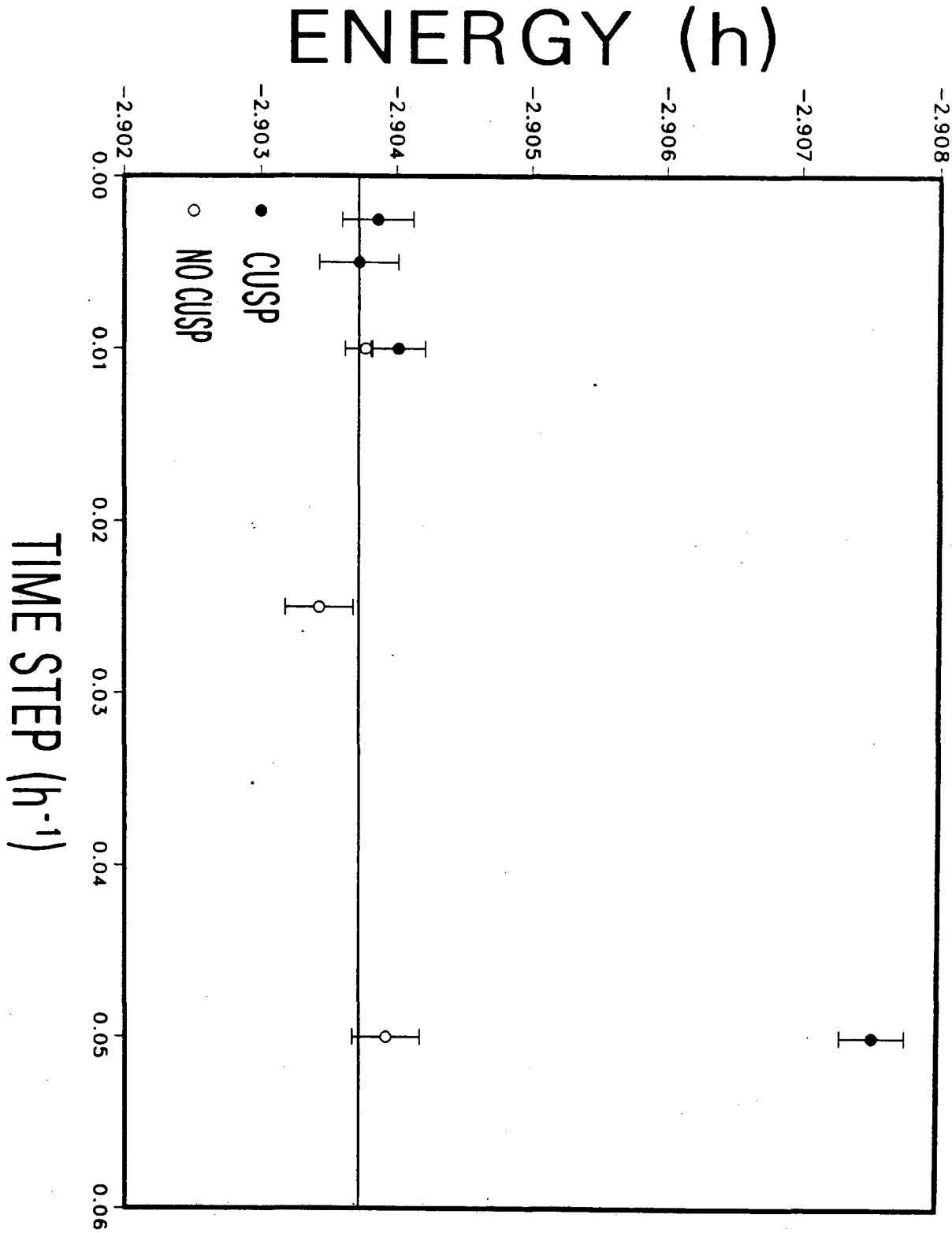


Figure 3.10. Effect of satisfying the E-N cusp condition for H_2. The straight line represents the exact energy.

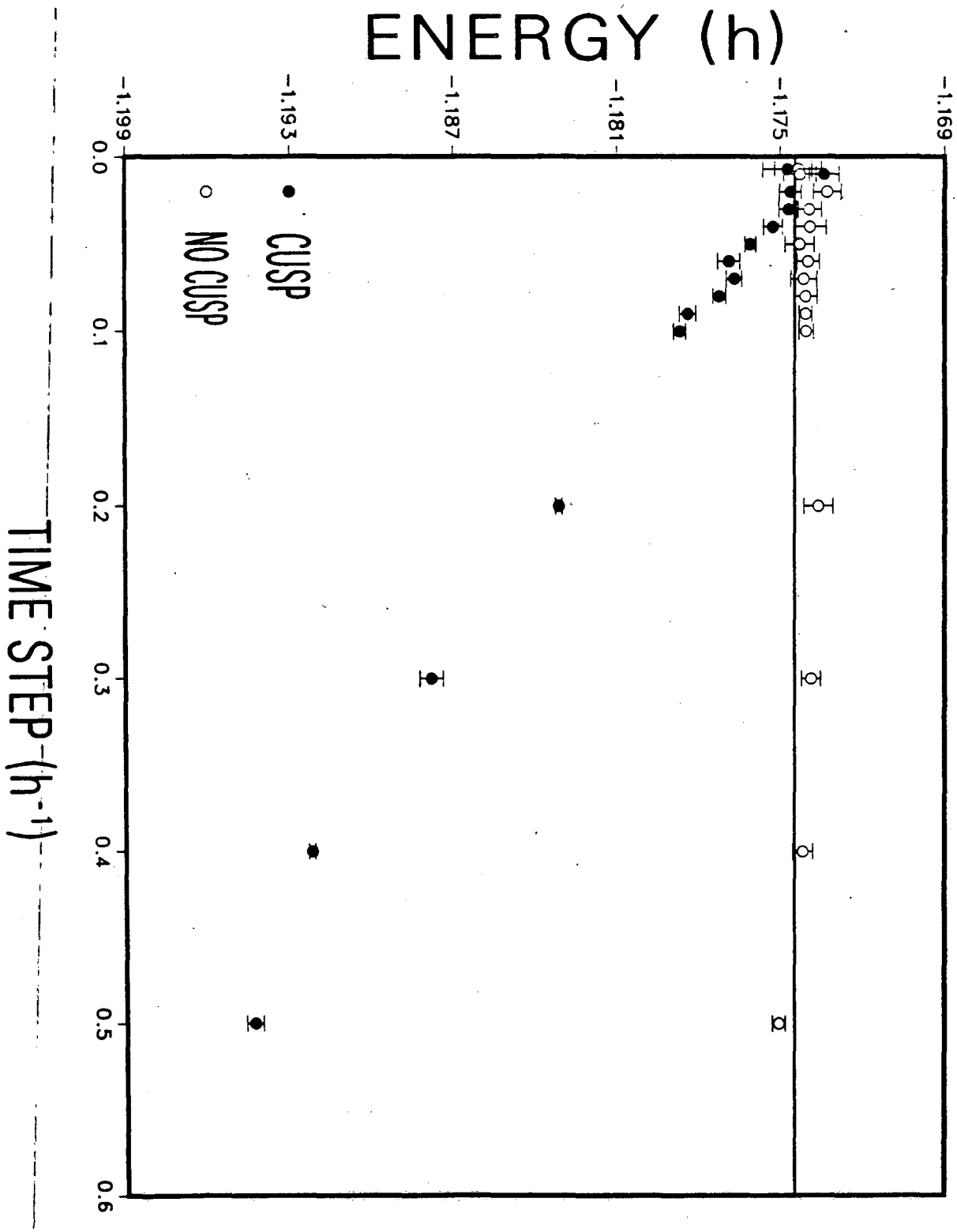


Figure 3.11. Effect of satisfying the E-N cusp condition for LiH. The open circles denote computations of the energy with λ_H and λ_{Li} chosen to satisfy the E-N cusp condition on average. The straight line represents the energy at $\tau = 0$ obtained from a second-order fit of the open-circled points.

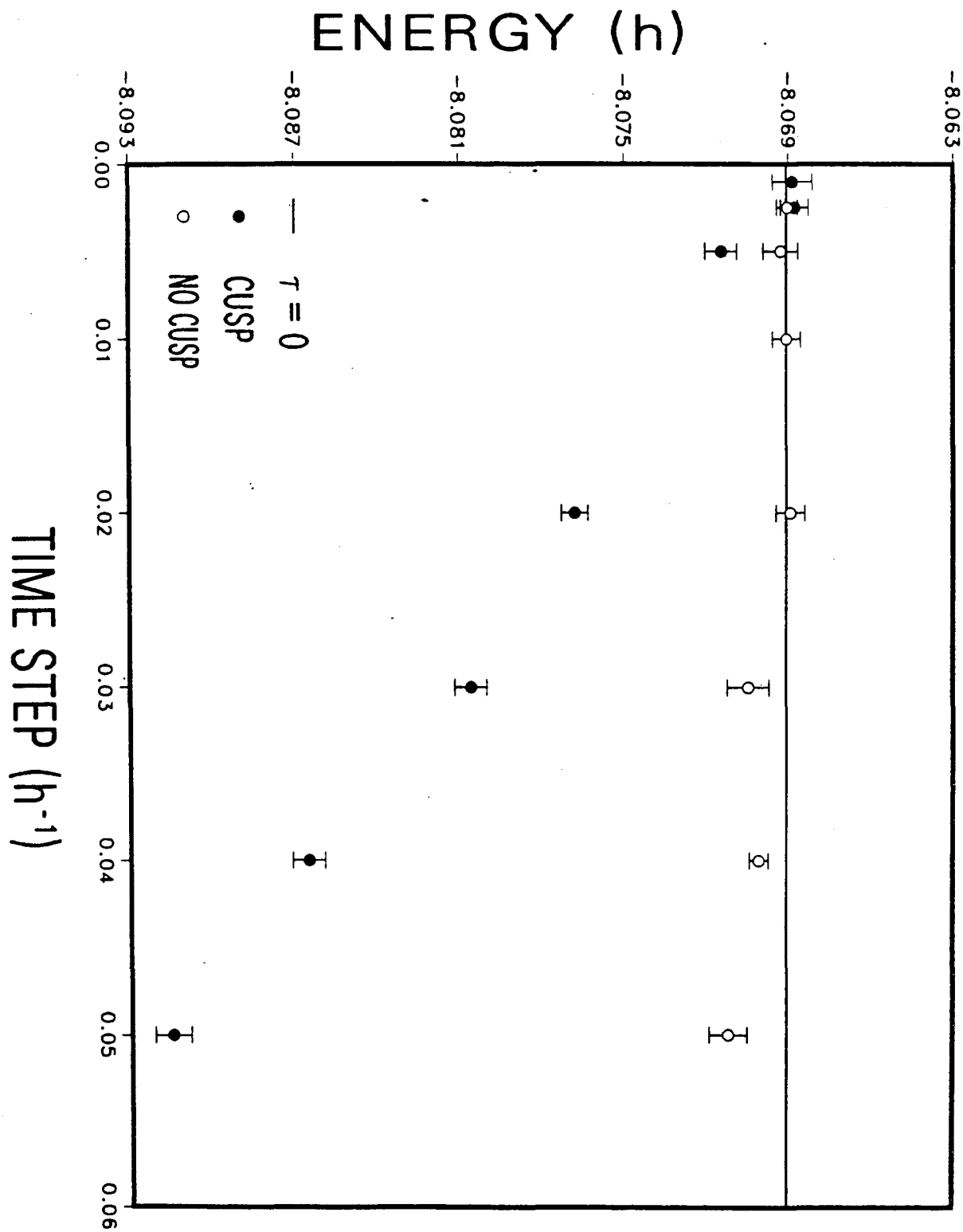


Figure 3.12 Exchange nodes of Ψ_1. The curves are cross sections through a selection of nodal surfaces arising from the exchange antisymmetry. Full nodal surfaces are obtained by rotating the curves about the internuclear axis. Each surface is obtained by fixing the position of one electron on it and finding the locus of points for the other like-spin electron at which $\Psi_1 = 0$. It can be easily shown, cf. Appendix 3.1, that Ψ_1 is zero whenever both like-spin electrons are *anywhere* on a surface. In Figs. 3.12-3.15, distances are in bohr and the circles represent the hydrogen nuclei.

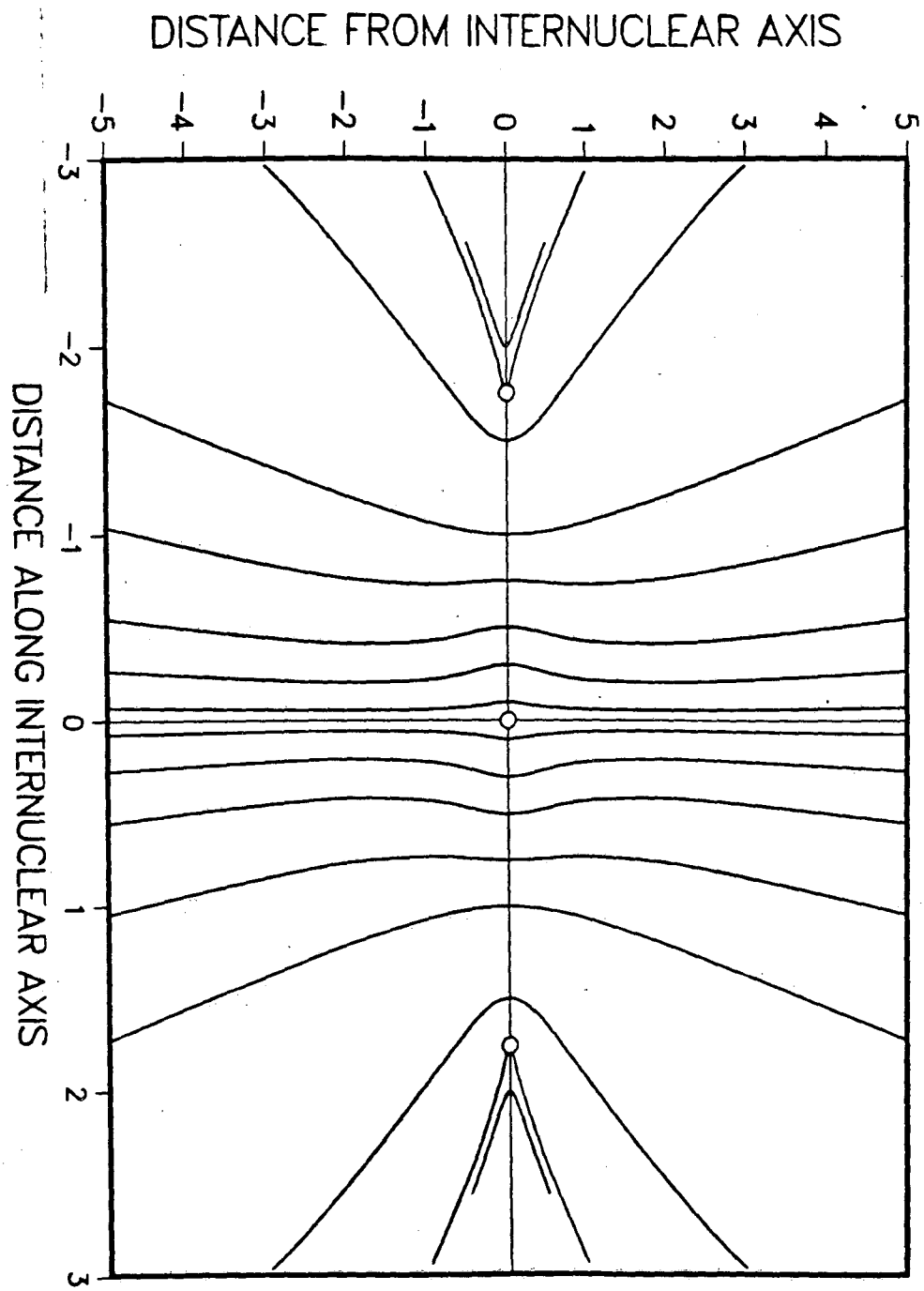


Figure 3.13 Exchange nodes of Ψ_4. See Fig. 3.12 for further explanation. Note how different these nodes are from those of Ψ_1, shown in Fig. 3.12. This is consistent with the difference in the QMC energies obtained with Ψ_1 and Ψ_4.

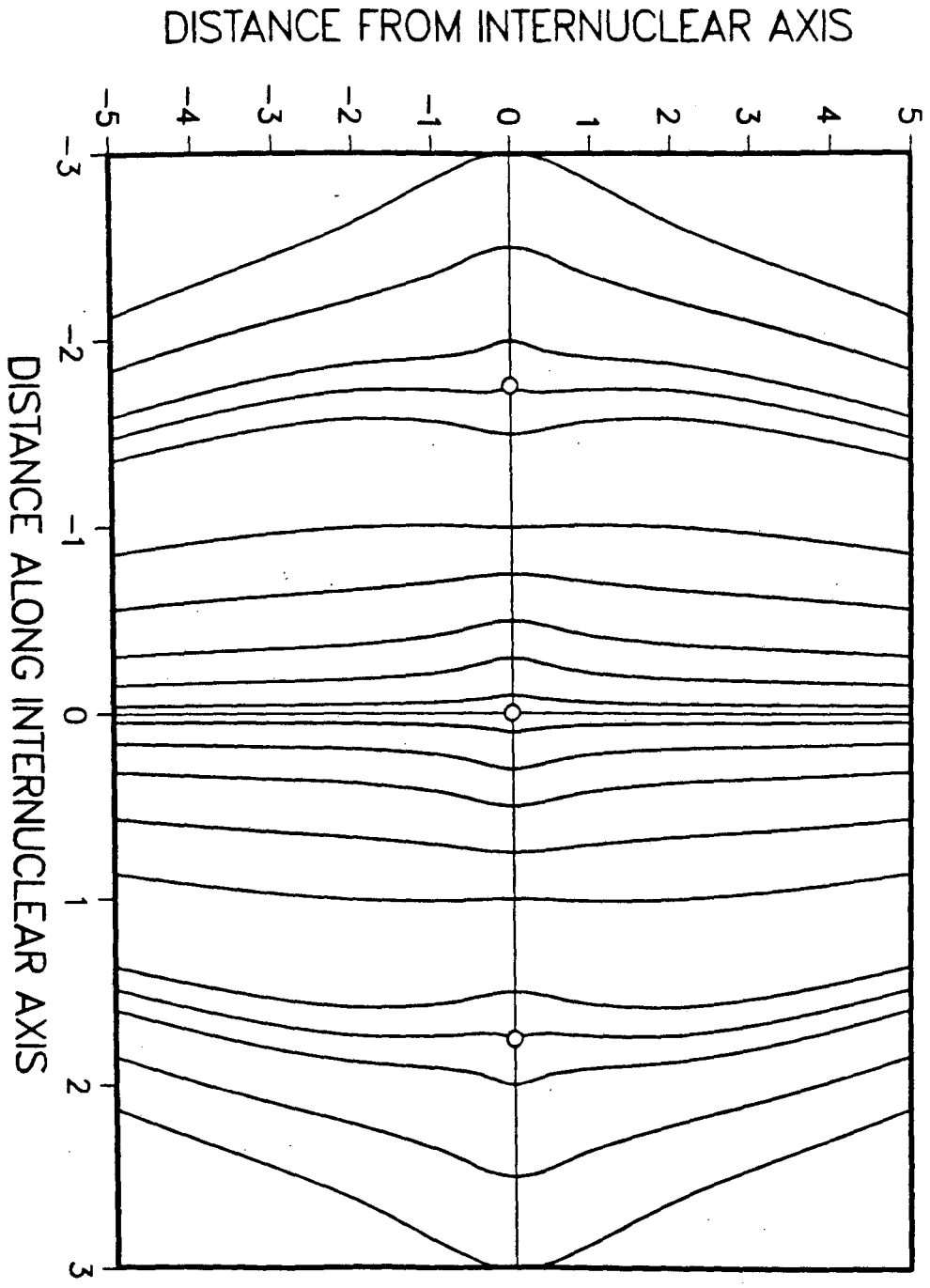


Figure 3.14 Exchange nodes of Ψ_5. See Fig. 3.12 for further explanation. Note the similarity between these nodes and the exchange nodes of Ψ_4 (Fig. 3.13), and the dissimilarity between these nodes and the nodes of Ψ_1 (Fig. 3.12). The exchange nodes of Ψ_2, Ψ_3, and Ψ_6 (not shown) are qualitatively the same as those shown here. This is consistent with the close similarity in the QMC energies obtained with Ψ_2–Ψ_6.

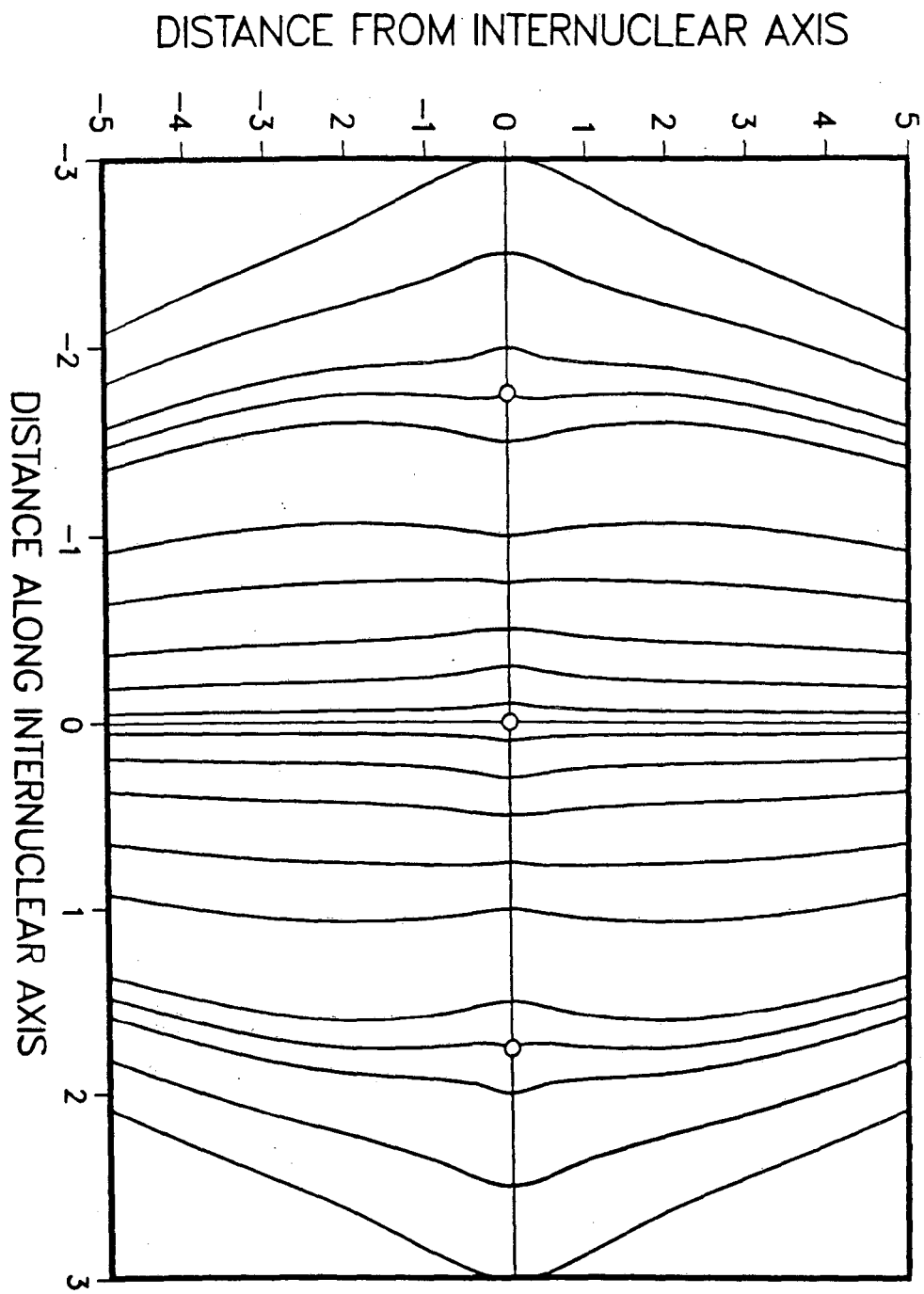


Figure 3.15. Nodes of the σ_g MO's of Ψ_2, Ψ_5, and Ψ_6. The curves shown are cross sections through the nodal surfaces. The trial functions Ψ_5 and Ψ_6 are zero if the unlike-spin or *both* like-spin electrons are on the corresponding nodal surface. Trial function Ψ_2 is zero if the *unlike*-spin electron is on the pictured nodal surface.

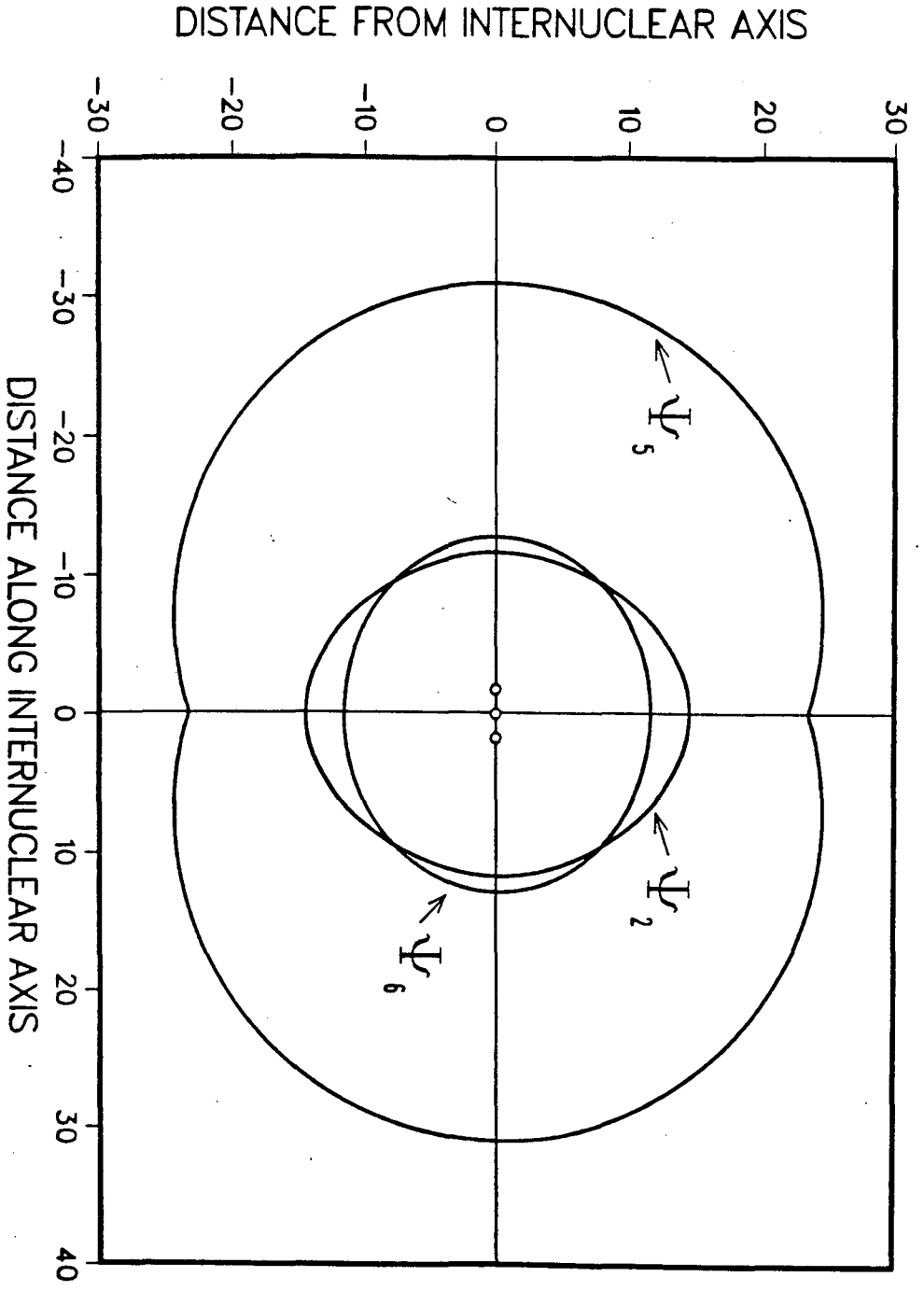


Figure 3.16. QMC barrier height dependence on time step for Ψ_4. The intercept gives an unbiased estimate of the QMC barrier height as $E_B = 9.86(0.22)$ kcal/mol. The curve is the second-order least squares fit $E(\tau) = 9.86 + 5.61\tau - 73.2\tau^2$.

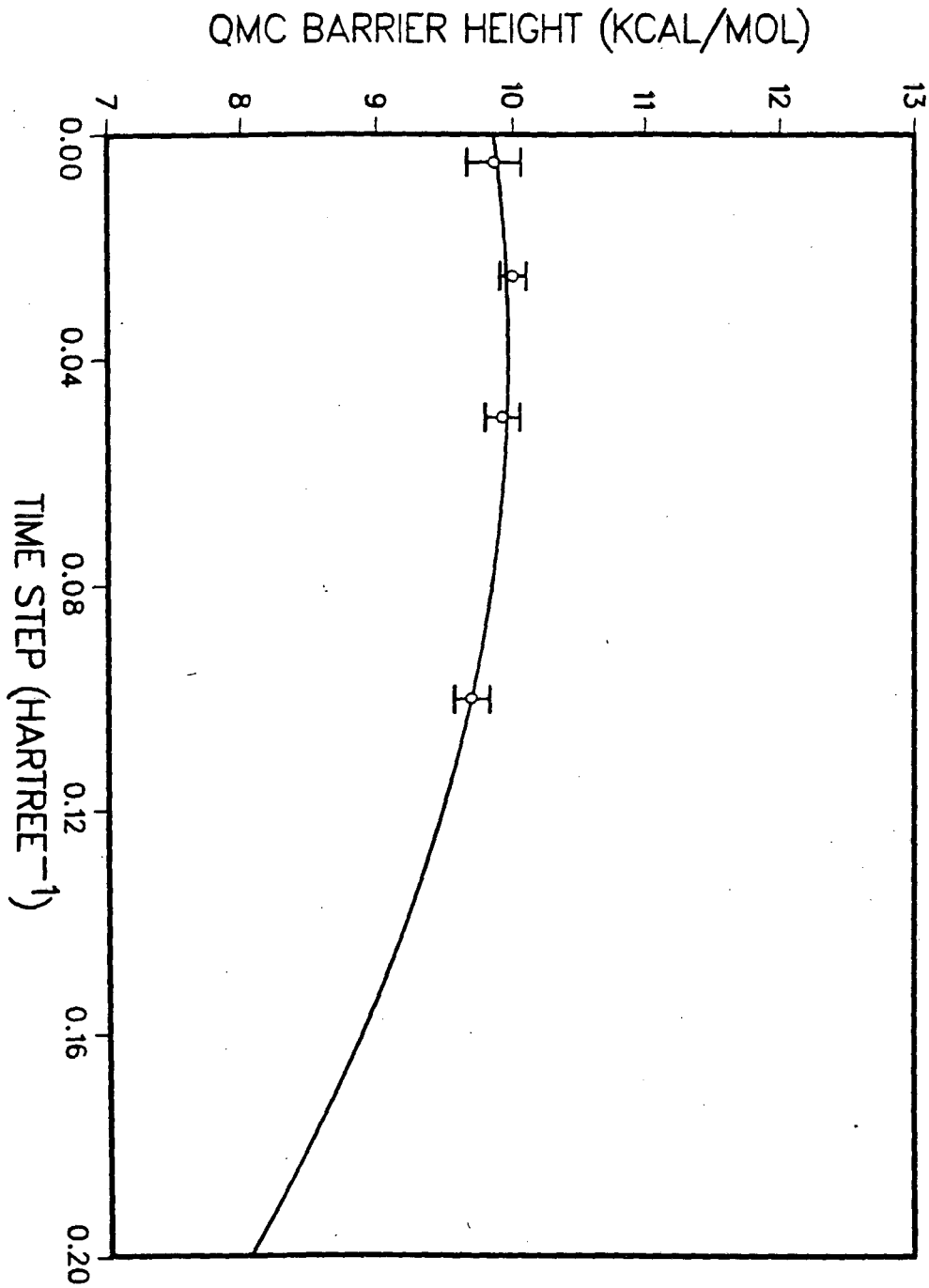


Figure 3.17. Reaction path comparison of QMC energies with a fit of the best CI results. The reaction coordinate r_c is the distance of the point (r_1, r_2) from the from the saddle point. The geometries for the Monte Carlo points are $r_1 = r_2 = 1.757$ bohr ($r_c = 0$); $r_1 = 2.084$, $r_2 = 1.550$ bohr ($r_c = 0.387$); and $r_1 = 2.572$, $r_2 = 1.448$ bohr ($r_c = 0.872$). The solid curve is a spline fit to Liu's reaction path data.

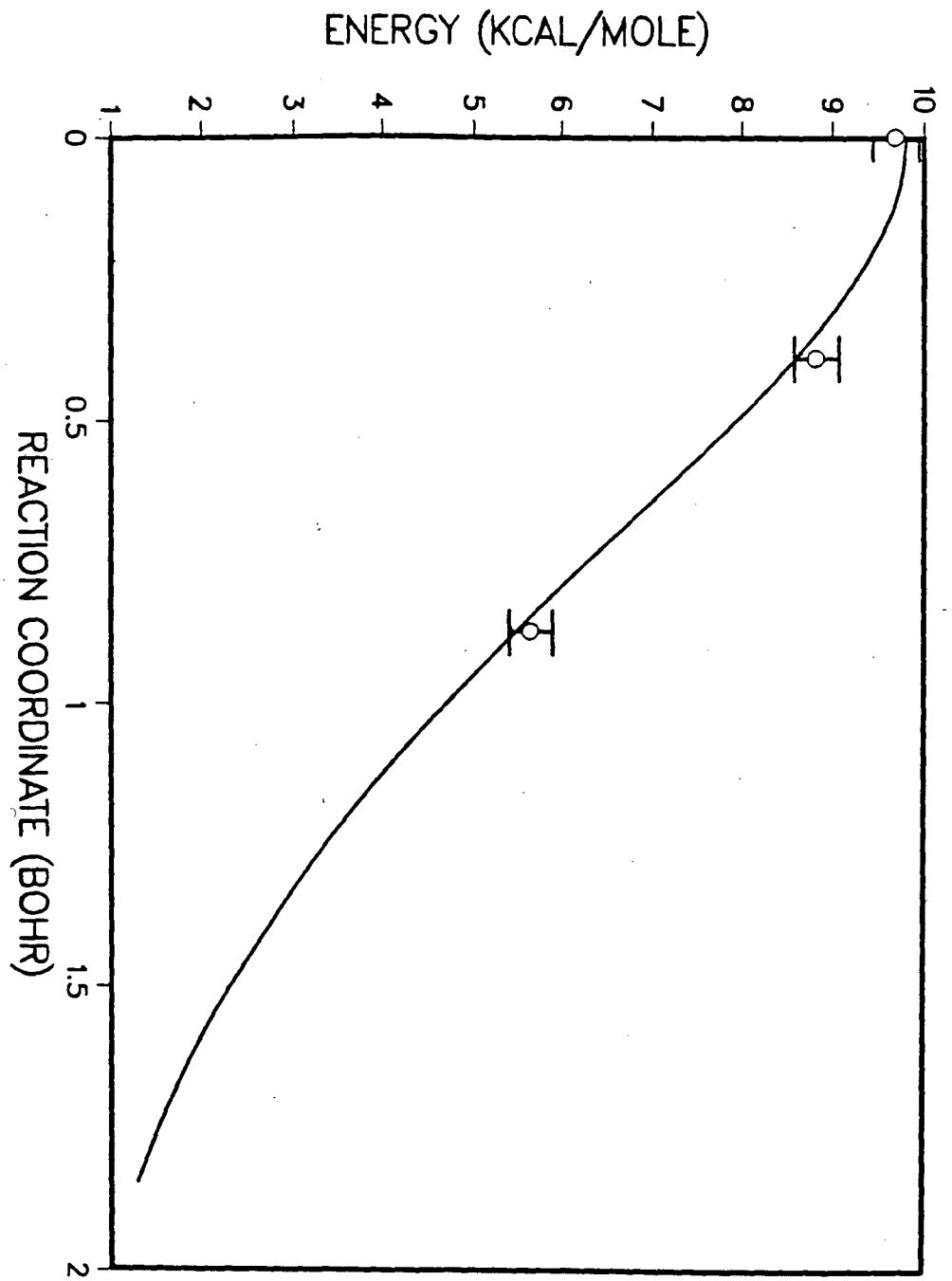
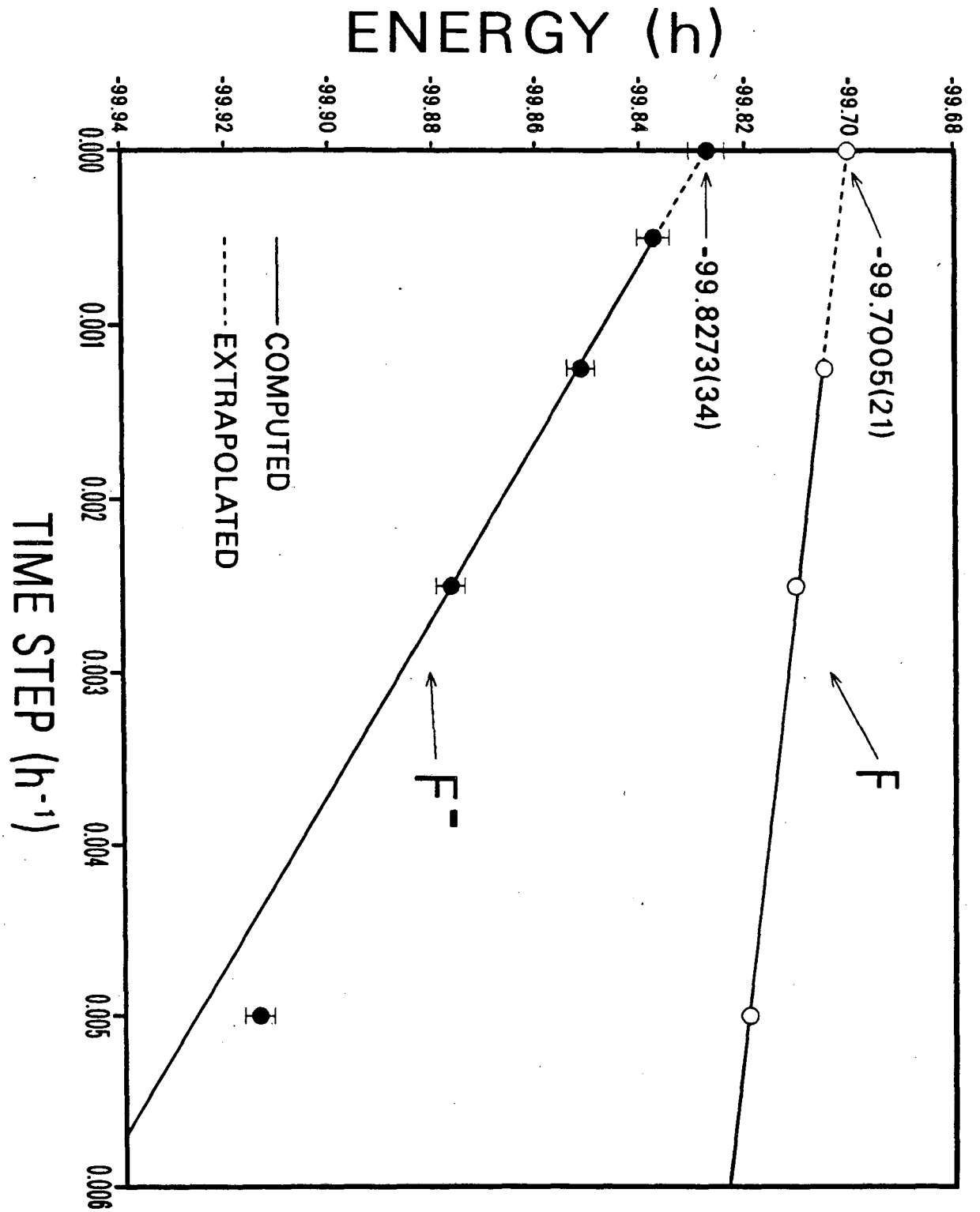


Figure 3.18. QMC energy versus time step for F and F^-. For F, the statistical errors are contained within the points.



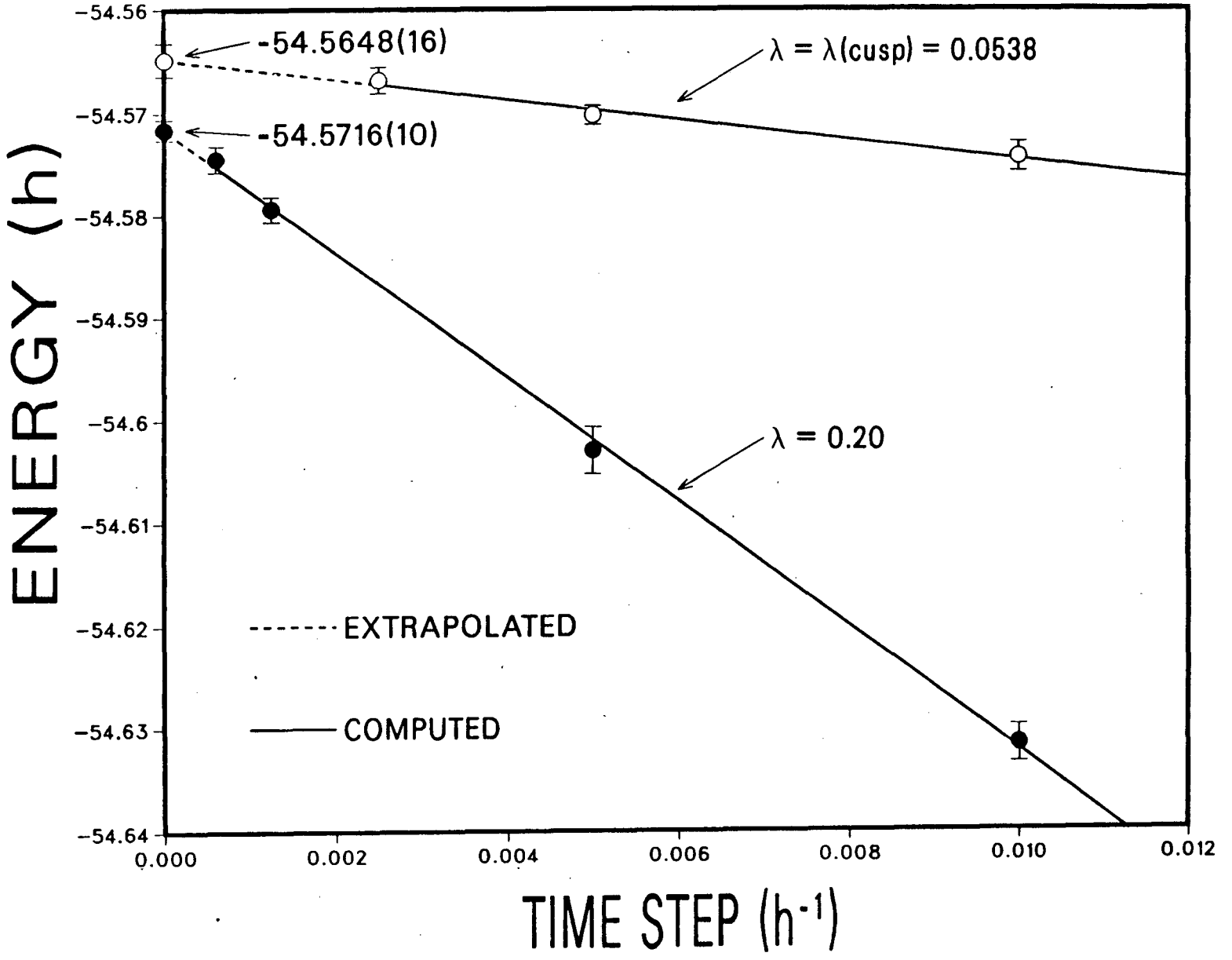


Figure 3.19. QMC energy versus time step for N with and without $\lambda = \lambda(\text{cusp})$.

Figure 3.20. QMC energy versus time step for N_2 with and without $\lambda = \lambda(\text{cusp})$.

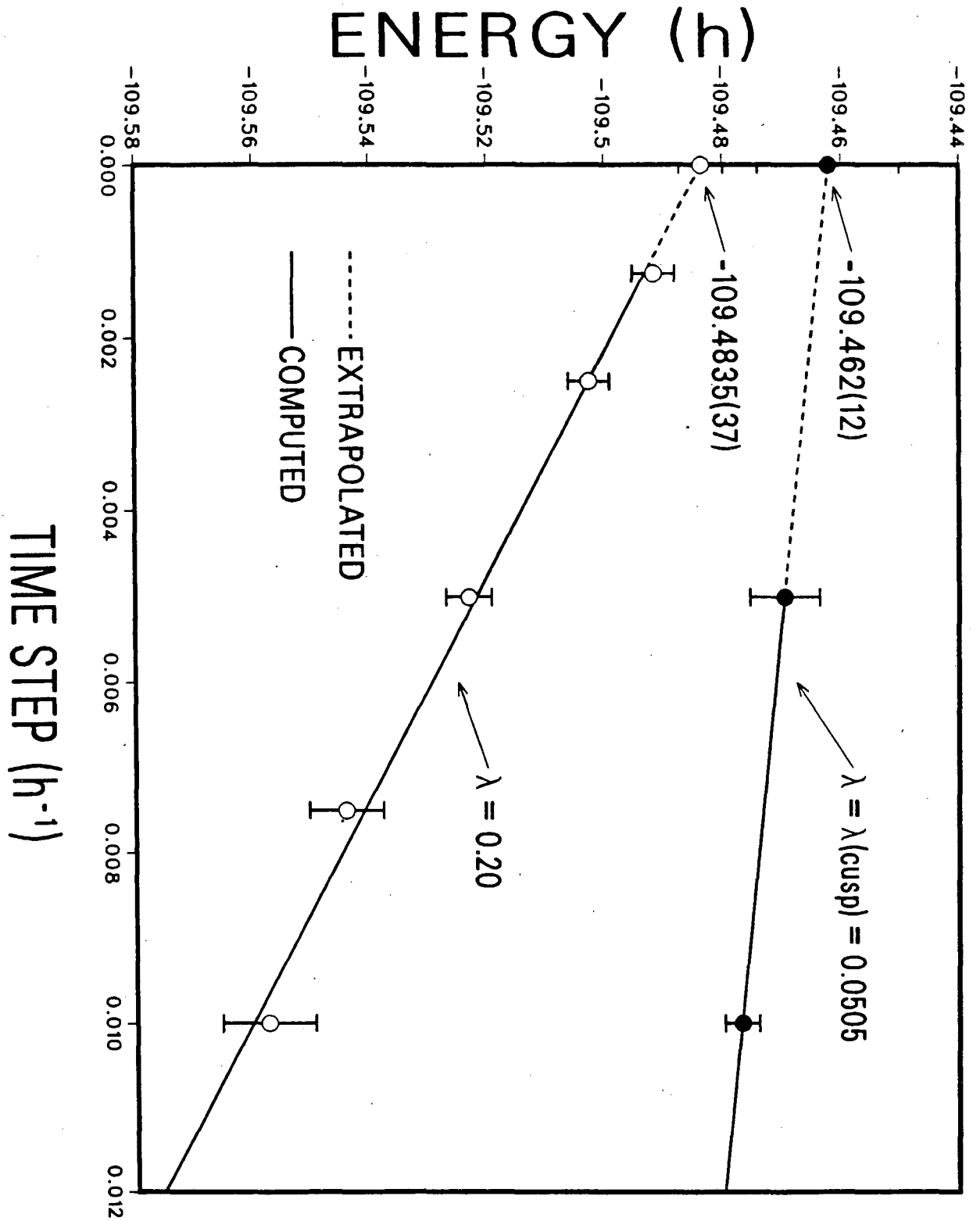
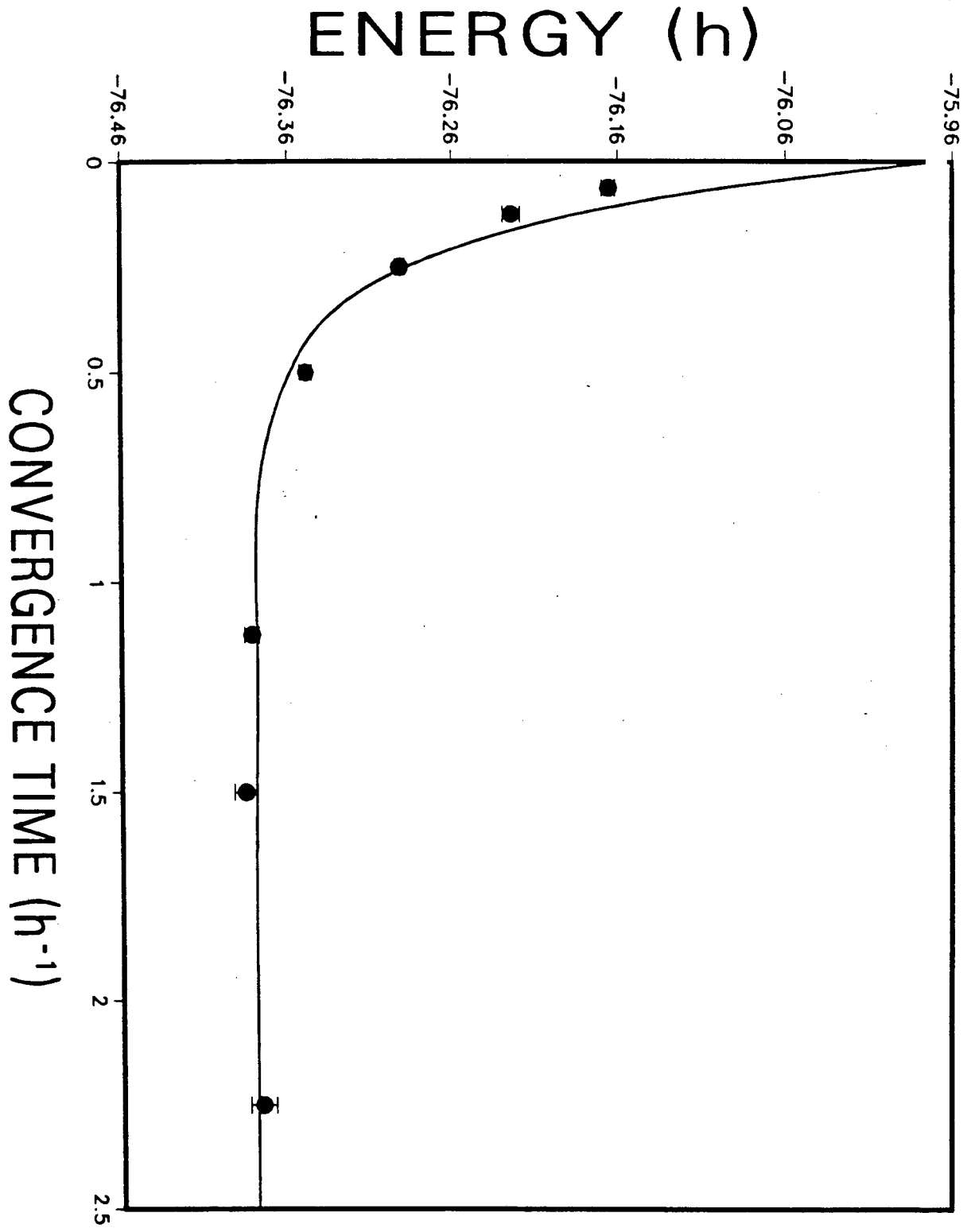


Figure 3.21. QMC energy versus convergence time for H_2O. The QMC energy is computed by difference and is given by $E(t) = \Delta(t) + E_{SCF}$, cf. Sec. 3.5. The computed energies are fitted to $a + b\exp(-\gamma t)$ at $t = 0$ and $t \geq 0.25\ h^{-1}$.



Chapter 4

Expectation Values of Coordinate Operators

4.1 Introduction

Over the last ten to fifteen years Monte Carlo techniques have been increasingly applied to quantum mechanical problems.[126] Several of these QMC methods have been employed in order to obtain stochastic solutions of the Schrödinger equation for atomic and molecular systems. The focus of most of these approaches has been the accurate computation of the total electronic energy of small atoms and molecules. Since energy is only one of many important properties, it is desirable to evaluate expectation values of operators other than the Hamiltonian. For properties whose operators do not commute with H, such as functions of coordinates from which static moments of the charge distribution may be obtained, the usual evaluation of QMC averages as the so-called mixed expectation values is not exact.[127] However, as is well known, these moments must be accurately determined to describe the interaction of a molecule with an electric field or with another species at long range.

As seen in Chap. 3, the fixed-node diffusion QMC approach yields accurate energies for a variety of first-row atoms and molecules. For the larger systems, approximately 90% of the CE is obtained, and computations on 2-4 electron systems yield 98-100%. Given this accuracy, it is of interest to broaden the scope of QMC so that

properties other than the energy may be studied.

The QMC approach described in Chap. 2 samples the "mixed" distribution $\Psi_T \phi_0$. If A is an Hermitian operator which commutes with the Hamiltonian, so that $A\phi_0 = a_0\phi_0$, then the eigenvalue a_0 may be obtained by averaging $A_L \equiv \Psi_T^{-1} A \Psi_T$ over $\Psi_T \phi_0$. That is,

$$A_m \equiv \int \Psi_T \phi_0 A_L \, d\underline{R} / \int \Psi_T \phi_0 d\underline{R}$$
$$= \int \phi_0 A \Psi_T \, d\underline{R} / \int \phi_0 \Psi_T \, d\underline{R} = a_0 \, . \tag{4.1}$$

Thus, in the terminology used from hereon, the mixed expectation value A_m, is identical to the exact, or "pure", expectation value, $A_p \equiv \langle\phi_0|A|\phi_0\rangle$. (Here, and throughout this chapter, all wave functions are assumed to be normalized.) Note, however, that when $[A, H] \neq 0$, e.g., A is a coordinate operator,

$$\int \phi_0 A \Psi_T \, d\underline{R} / \int \phi_0 \Psi_T \, d\underline{R} \neq \int \phi_0 A \, \phi_0 d\underline{R} \, , \tag{4.2}$$

i.e., $A_m \neq A_p$. Therefore, the QMC approach employed thus far does not give exact values (to the extent that ϕ_0 is exact) of important properties such as dipole and quadrupole moments. As seen from Eq. (4.2), pure expectation values of coordinate operators must be obtained by sampling $|\phi_0|^2$ rather than $\Psi_T \phi_0$.

It should be pointed out, however, that when $A_m \neq A_p$, an improvement over the mixed expectation value may often be obtained.[127] This is seen as follows. The mixed average is accurate to first order in quantities which depend on the difference function, $\delta \equiv \phi_0 - \Psi_T$, i.e., the quantities $\langle\phi_0|\delta\rangle$ and $\langle\phi_0|A|\delta\rangle$. Also accurate to first order in δ is the "trial" expectation value, $A_T \equiv \langle\Psi_T|A|\Psi_T\rangle$. Trial expectation values are computed by sampling from $|\Psi_T|^2$, in a procedure often referred to as variational Monte Carlo (VMC). VMC may be cast in a form algorithmically identical to

diffusion QMC, except that the branching, cf. Eq. (2.23), is suppressed. The usefulness of A_T in this context, is that it can be combined with A_m to obtain an estimate of A_p accurate to *second* order in δ.[127,128] That is,

$$A_s \equiv 2A_m - A_T = A_p + O(\delta^2). \tag{4.3}$$

Though this approximation is generally better than either the trial or mixed values, it may still be poorer than desired for reliable predictions. Given this, it is useful to compute pure expectation values exactly. Algorithms to do this efficiently are explored in this chapter. An alternative approach has also recently been explored.[129]

Approaches for modifying the Green's function Monte Carlo (GFMC) to sample $|\phi_0|^2$ have been proposed and explored by Kalos.[32,33] These approaches, however, do not incorporate importance sampling. Subsequently, Liu *et al.*,[130] in studying quantum hard spheres, presented a GFMC technique with importance sampling which employs the idea of Ref. 33 for obtaining $|\phi_0|^2$. Here, we follow the approach of Liu *et al.* in exploring modifications to the diffusion QMC approach to order to obtain $|\phi_0|^2$, and thereby compute pure expectation values of coordinate operators, for atomic and molecular systems. (Methods for computing pure expectation values of other kinds of operators, e.g., differential operators, are not known.)

The remainder of this chapter is organized in four sections. In Section 4.2 the theory for sampling the ratio ϕ_0/Ψ_T is presented. Also in this section, algorithms are presented for computing pure expectation values using a single QMC walk and trial, mixed and pure expectation values simultaneously employing a VMC walk with QMC "side walks". In studying the relative merits of the algorithms presented in Sec. 4.2, Section 4.3 gives results for the moments of H and H_2 which serve as test cases.

Section 4.4 gives results for the dipole and quadrupole moments are LiH and BH employing the algorithms described below. Section 4.5 concludes this paper with ideas for reducing statistical error in averages of odd functions.

4.2 Algorithms for the Computation of Pure Expectation Values

4.2.1 Pure Expectation Values by a Single QMC Walk

Since the QMC approach described in Chap. 2 yields only mixed expectation values, the ratio ϕ_0/Ψ_T must be sampled in order to obtain pure expectation values. That is,

$$\frac{\langle A\,(\phi_0/\Psi_T)\rangle_f}{\langle \phi_0/\Psi_T\rangle_f} = \frac{\langle \Psi_T\,|A\,(\phi_0/\Psi_T)|\phi_0\rangle}{\langle \Psi_T\,|\phi_0\rangle} \Big/ \frac{\langle \Psi_T\,|\phi_0/\Psi_T\,|\phi_0\rangle}{\langle \Psi_T\,|\phi_0\rangle} \tag{4.4}$$

$$= \langle \phi_0|A\,|\phi_0\rangle = A_p\ .$$

As shown by Liu *et al.*,[130] the ratio ϕ_0/Ψ_T may be obtained from the asymptotic number of descendants resulting from a QMC walk which starts at $\underline{R}$. For completeness, we modify their proof for diffusion QMC.

An initial distribution given by a single point at $\underline{R}$, is

$$f_{\underline{R}}(\underline{R}', 0) = \Psi_T(\underline{R}')\Phi(\underline{R}', 0) = \delta(\underline{R}' - \underline{R})\ . \tag{4.5}$$

Expanding the initial state function, Φ, in the complete set of (normalized) eigenfunctions of H yields

$$\delta(\underline{R}' - \underline{R}) = \Psi_T(\underline{R}')\sum_i c_i\,\phi_i(\underline{R}')\ . \tag{4.6}$$

The expansion coefficients may be obtained by multiplying Eq. (4.6) by $\phi_i(\underline{R}')/\Psi_T(\underline{R}')$

and integrating over $\underline{R}'$, giving

$$c_i = \phi_i(\underline{R})/\Psi_T(\underline{R}) . \tag{4.7}$$

From the asymptotic form of f, Eq. (2.24), we note

$$f_{\underline{R}}(\underline{R}', t) = c_0 \exp[-t(E_0 - E_R)]\Psi_T(\underline{R}')\phi_0(\underline{R}') . \tag{4.8}$$

Integrating Eq. (4.8) over all space "counts" the current number of descendants, or for large t the asymptotic population $P(\underline{R})$ of a walker starting at $\underline{R}$. Thus,

$$P(\underline{R}) \equiv \int f_{\underline{R}}(\underline{R}')d\underline{R}' = \phi_0(\underline{R})/\Psi_T(\underline{R})\exp[-t(E_0 - E_R)]\langle\Psi_T|\phi_0\rangle . \tag{4.9}$$

Returning to Eq. (4.4) for the evaluation of A_p, ϕ_0/Ψ_T may be replaced by P since the overlap integral and time dependence present in P cancel.

In order to count efficiently the descendants of a walker *during* the QMC walk, we have developed the following algorithm. With it, descendants of an arbitrary walker at time t later (or, equivalently, N steps later where $N\tau = t$) may then be counted to obtain $P \propto \phi_0/\Psi_T$. Although this process may be repeated for different points sampled from $\Psi_T\phi_0$, with this algorithm this repetition is unnecessary because the full distribution is generated from a single walker. Moreover, an additional cost would be incurred to propagate each initial point a distance of N time steps to reach the asymptotic domain. Instead, each step of a single (potentially branching) walk is used, that is, by propagating an additional time step to a time $t + \tau$, the "N-distant" offspring of the first generation (at time τ) may be computed. Likewise, at time $t + 2\tau$, N-distant offspring populations may be determined for walkers which may be thought of as beginning at $t = 2\tau$. Therefore, after an initial investment of N steps, additional N-distant estimates of ϕ_0/Ψ_T may be sampled for points at $t = \tau$, 2τ, and so on. Furthermore, since N itself is arbitrary, convergence of the asymptotic population as a func-

tion of N may also be determined.

The branching of the QMC walk, however, requires tagging walkers. When the branching yields several walkers at a time $m\tau$, one must know which of those at $(m+N)\tau$ are descended from which of those at $m\tau$. To store this information a "family tree" is created as the walk progresses, and each walker is labelled so that its location in the tree is specified. This labelling is accomplished with two tags. The first, θ $(0 \leq \theta < \Theta)$, gives (together with t) a walker's location in the tree, while the second, δ, specifies that area (range) within the tree in which descendants of this walker will be placed. The tagging begins by setting $\theta(k=1, t=0) = 0$, where the index k labels the walker, and $\delta(k=1) = \Theta$. If a walker "dies", i.e., has no offspring, then no subsequent values of θ and δ are derived from it. When the kth walker has $n_k(>0)$ immediate offspring, the values of θ and δ assigned to these daughters are obtained from $\theta(k, t)$ and $\delta(k)$ according to

$$\theta(M+l, t+\tau) = \theta(k, t) + \frac{l-1}{n_k}\delta(k) \tag{4.10a}$$

and

$$\delta(M+l) = \delta(k)/n_k, \tag{4.10b}$$

where l ranges from 1 to n_k. To properly sequence the tags of the daughters, $M \equiv \sum_{i=1}^{k-1} n_i$ is the partial sum [up to walker (k-1)] of the number of walkers in the new generation. As the example in Fig. 4.1 shows, this method yields tags that are isomorphic to the tree. Note, in counting the descendants of a walker with $\theta = \theta(k, t)$ and $\delta = \delta(k)$, these descendants lie in the range, S, given by

$$S = [\theta(k, t), \theta(k, t) + \delta(k)) \ . \tag{4.11a}$$

From Eqs. (4.10), $\theta(k, t) + \delta(k) \leq \theta(k+1, t)$; the equality would always hold if no walkers died. Since the descendants of the next walker $(k+1)$ possess values of $\theta \geq \theta(k+1, t)$, the θ values of walkers descended from walker k also lie within

$$S' = [\theta(k, t), \theta(k+1, t)) \ . \tag{4.11b}$$

This choice is more convenient, since only θ values need to be stored to compute the range in which the descendants of a given walker are located.

To verify that the limits in Eq. (4.11a) follow from the procedures described by Eqs. (4.10), we consider the values of θ possessed by the descendants at time $t + N\tau$ of a walker with $\theta = \theta(k, t)$ and $\delta = \delta(k)$. Repeated use of Eqs. (4.10) readily shows that a particular descendant, m, will be located at

$$\theta(m, t+N\tau) = \theta(k, t) + \frac{l_1-1}{n_1}\delta(k) + \frac{l_2-1}{n_2}\cdot\frac{\delta(k)}{n_1} + \cdots$$
$$+ \frac{l_L-1}{n_L}\cdot\frac{\delta(k)}{n_1 n_2 \cdots n_{L-1}} \ . \tag{4.12}$$

The labels l_i, $\{1 \leq l_i \leq n_i\}$, correspond to the ancestral lineage of m, while n_i (> 0) give the number of daughters in each family leading to m. For example, there are n_1 daughters of walker k; daughter l_1, who has n_2 daughters, is the direct ancestor of m in this generation, and so on. In Fig. 4.1, the direct ancestors of point "15" are the points labelled 1, 3, 7 and 12. The minimum value of $\theta(m, t+N\tau)$, corresponding to the initial walker never dying, is obtained by setting $\{l_i\} = 1$, which yields $\theta_{\min}(m, t+N\tau) = \theta(k, t)$. The maximum θ, corresponding to the walker with the largest value of θ at each time step never dying, results from setting each l_i to its maximum value (n_i) yielding

$$\theta_{\max}(m, t+N\tau) = \theta(k, t) + \frac{n_1-1}{n_1}\delta(k) + \frac{n_2-1}{n_2}\cdot\frac{\delta(k)}{n_1} + \cdots + \frac{n_{L-1}}{n_L}\cdot\frac{\delta(k)}{n_1 n_2 \cdots n_{L-1}}$$

$$= \theta(k, t) + \delta(k)\left(1-\frac{1}{n_1 n_2 \cdots n_L}\right) . \tag{4.13}$$

Therefore, the range of walker k is indeed given by Eq. (4.11a). Since this range is contained within $[\theta(k, t), \theta(k+1, t))$, the "descendant spaces" of walkers never overlap.

With the assignment of labels to walkers as described above, asymptotic populations are readily sampled during a QMC walk. For a set of $N(T)$ points, T is an arbitrary time, the number of descendants of the kth point at a later time $T + t$ is given by

$$P(\underline{R}_k, t) = \sum_{i=1}^{N(T+t)} I_i , \quad I_i = 1, \theta(k, t) \le \theta(i, T+t) < \theta(k+1, t) , \tag{4.14}$$

$$I_i = 0, \text{ otherwise,}$$

and $\theta(N+1, t) \equiv \Theta$.

The discussion thus far treats only one tree. To sample asymptotic populations of several independent points simultaneously, and to take advantage of the vector capabilities of current machines, a group of family trees is created. This is easily accomplished for M_p initial points by setting

$$\theta_i(1, 0) = (i-1)\Theta , \qquad 1 \le i \le M_p . \tag{4.15}$$

For each initial point $\delta = \Theta$, new values of θ and δ are computed according to Eqs. (4.10). Again, the descendant spaces are given by either (4.11a) or (4.11b) with integer multiples of Θ added to each bound.

Finally, it should be noted that with this algorithm there is the need for some additional memory overhead in the storage of $\theta(k, t)$. This can become excessive when

the time step τ must be very small to minimize bias. However, because successive values of the function to be sampled, $A(\underline{R})$, are highly correlated, A and P need not be sampled at every point along the walk. By sampling only every n steps, these memory requirements are reduced by a factor of n.

4.2.2 Branching Algorithms with Weighting

Here we explore a variant of the branching generally used in QMC walks to reduce the statistical error in sampled values of ϕ_0/Ψ_T. Upon completing a move, $\underline{R} \rightarrow \underline{R}'$, the most common implementation of branching is to obtain an integer, I_b, which specifies the number of walkers at $\underline{R}'$. The number of copies, I_b, is $\mathrm{int}[b(\underline{R}, \underline{R}') + \xi]$, where $b(\underline{R}, \underline{R}')$ is the weight of $\underline{R}'$ relative to $\underline{R}$ and is given by the branching factor of the short-time Green's function, cf. Eq. (2.27), and ξ is a uniform random variate between 0 and 1. While this rounding is correct on average, i.e., $\overline{I_b} = b$, a "microscopically" exact procedure is to weight each walker by the product of its current weight and the branching factor b. The drawback to this weighting procedure is that, since the product of these weights tends to either 0 or ∞, efficiency is lost with computations on walkers that contribute very little information due to their low (absolute or relative) weights.

A combination of branching and weighting, however, is useful. In this case we omit integer rounding until a weight becomes exceedingly small or large. When the weight w becomes large, an integer I_w is determined from it, as described above; however the daughters are assigned weights of I_w/w, rather than unity, so that no loss of information occurs. When the weight becomes smaller than a threshold value, integer

rounding *is* applied. The benefit of this modification is that the variance of the asymptotic populations, P, now given as the sum of the weights of descendants, is noticeably reduced. This leads to improved precision in pure expectation values. See the Appendix D for an analysis of the variance of rounded versus unrounded numbers.

A further point of interest is the omission of renormalization in this algorithm. Recalling the discussion of the QMC algorithm in Sec. 2.5, the QMC simulation is divided into blocks. At the end of every block the population of walkers, N_p, is reset to the number of walkers at the beginning of the calculation. This step is useful because fluctuations in the ensemble size increase as the simulation proceeds.[40] However, renormalization is not desirable when sampling $|\phi_0|^2$ because the copying or removal of walkers would give an error in the asymptotic populations. That is, for certain points, the asymptotic populations would be arbitrarily changed by creating or removing descendants. Therefore, though the single-walk approach is divided into blocks, the ensemble is not renormalized at the end of each block.

Since renormalization is not employed, the length of the calculation must be kept sufficiently short so that the number of walkers does not become zero or larger than the allocated memory. On the other hand, long runs are desired so that the overhead introduced by converging to asymptotic populations at the beginning of the simulation is relatively small. For all systems studied, the fluctuations in the ensemble size were small enough to allow runs in which the overhead due to convergence is about 5% of the total computational cost.

The final point we discuss concerns the amount of bias in block averages. This bias arises from averaging ratios as opposed to summing numerators and denominators

and then dividing. This bias may be large when the number of points sampled in each block varies widely, due to the absence of renormalization, as we discuss below.

We note that A_p is obtained exactly, by sampling an infinite number of points from $\Psi_T \phi_0$, when computed as

$$A_p = \lim_{N \to \infty} \sum_{i=1}^{N} A(\underline{R}_i) P(\underline{R}_i) / \sum_{i=1}^{N} P(\underline{R}_i) , \tag{4.16}$$

where P are the asymptotic populations. In practice, only a finite number of points may be sampled and statistics are obtained from block averages. Therefore, pure (and other) expectation values are computed from block averages as

$$A_p(\text{block}) = N_b^{-1} \sum_{i=1}^{N_p} A_i , \tag{4.17}$$

where

$$A_i = \sum_{k=1}^{N_i} A(\underline{R}_{ki}) P(\underline{R}_{ki}) / \sum_{k=1}^{N_i} P(\underline{R}_{ki}) , \tag{4.18}$$

and N_i is the number of points sampled in the ith block. Since the average of a set of ratios [Eq. (4.17)] is not generally equal to the ratio of the sum of the numerators over the sum of the denominators [effectively, Eq. (4.16)], then a bias is present in A_p(block). This bias is not significant when the denominators in Eq. (4.18) are roughly the same, which is the case when renormalization is employed, or when the number of points sampled in each block is large enough to yield $A_p(\text{block}) \simeq A_p$. Therefore, a useful way to check for this bias in block averages is to increase the number of points sampled in each block and see if changes in computed averages occur. This is most conveniently accomplished by defining a "block" as the entire run, i.e., compute

$$A_p(\text{run}) = \sum_{i=1}^{N_b}\sum_{k=1}^{N_i} A(\underline{R}_{ki})P(\underline{R}_{ki}) \Big/ \sum_{i=1}^{N_b}\sum_{k=1}^{N_i} P(\underline{R}_{ki}) \ , \tag{4.19}$$

and then compare values of A_p(block) with values of A_p(run). Since the number of points used in computing A_p(run) is an order of magnitude larger than that in each block, a noticeable difference between block and run values will be observed *if* block averages are biased as discussed here. Employing this check, we have found no bias in computations of properties from block averages.

4.2.3 Pure Expectation Values by VMC with QMC "Side Walks"

This approach employs a VMC walk to sample points from the distribution $|\Psi_T|^2$. These walks are very efficient, can employ large time steps, and have no bias. The points obtained are then initial points for QMC "side walks". These side walks are performed to obtain P, and thereby ϕ_0/Ψ_T. Before implementing the QMC walk, the starting coordinates of the initial points are stored so that the VMC walk may be continued after values of ϕ_0/Ψ_T are sampled. Since values of P are computed only for points at the beginning of the QMC walks, the labelling process of the previous section is greatly simplified. For example, for M walkers drawn from $|\Psi_T|^2$, $\theta(i) = i$ and $\delta(i) = 0$. Under these conditions the descendants of the ith walker are simply those at the end of the QMC walk with $\theta = i$, cf. Fig. 4.2.

Since initial points are selected from $|\Psi_T|^2$, the ratio $|\phi_0/\Psi_T|^2$ is necessary to obtain pure expectation values. To do so, at least two *independent* samplings of ϕ_0/Ψ_T must be performed.[129b] Although the asymptotic population P is, on average, equal to ϕ_0/Ψ_T, its square is not an unbiased estimate of $|\phi_0/\Psi_T|^2$ because P possesses statistical error due to the randomness of the QMC walk. To show this we write

$$P(\xi) = \phi_0/\Psi_T + \eta(\xi), \tag{4.20}$$

where ξ corresponds to a specific QMC walk, and for convenience the time dependence is not displayed. Denoting the probability distribution of walks by $h(\xi)$, there results

$$\int h(\xi)P(\xi)d\xi = \phi_0/\Psi_T + \int h(\xi)\eta(\xi)d\xi = \phi_0/\Psi_T \ . \tag{4.21}$$

The term, $\int h(\xi)\eta(\xi)d\xi$ vanishes because on average $\eta = 0$ (i.e., $P = \phi_0/\Psi_T$). Note, however, that squaring the population yields

$$\int h(\xi)P^2(\xi)d\xi = |\phi_0/\Psi_T|^2 + \int h(\xi)\eta^2(\xi)d\xi \ , \tag{4.22}$$

and the second term on the right-hand side of Eq. (4.22) does *not* vanish. On the other hand, the product of two asymptotic populations obtained independently *is* equal to $|\phi_0/\Psi_T|^2$ on average, that is, from Eq. (4.21),

$$\int h(\xi_1)P(\xi_1)d\xi_1 \cdot \int h(\xi_2)P(\xi_2)d\xi_2 = |\phi_0/\Psi_T|^2 \ . \tag{4.23}$$

By sampling A and two values of P at points selected from $|\Psi_T|^2$, trial, mixed and pure expectation values may be computed. These averages are, respectively

$$A_T = \langle\Psi_T|A|\Psi_T\rangle \ ,$$

$$A_m = \langle\Psi_T|[P(\xi_1) + P(\xi_2)]A|\Psi_T\rangle/\langle\Psi_T|[P(\xi_1) + P(\xi_2)]|\Psi_T\rangle \ ,$$

and

$$A_p = \langle\Psi_T|P(\xi_1)P(\xi_2)A|\Psi_T\rangle/\langle\Psi_T|P(\xi_1)P(\xi_2)|\Psi_T\rangle \ . \tag{4.24}$$

Although two walks and two samplings of P are required for A_p, since A and P are sampled from a VMC walk which possesses no time-step bias, large time steps may be taken to sample configuration space efficiently. Another benefit is that in this approach (which we hereafter denote V+QMC) the statistical error in $A_p - A_T$ is generally much smaller than that of A_p alone, since A_p and A_T are correlated. This is

useful because trial expectation values can be quickly computed in a separate VMC calculation to high precision. Therefore, adding $A_p - A_T$ from V+QMC with a value of A_T from VMC alone (referred to hereafter as "correlated V+QMC") can yield a significant reduction in the statistical error of A_p.

4.3 Results for H and H_2

For purposes of evaluating the different approaches we have studied H and H_2. The H atom trial function is chosen as the 1*s* Slater orbital with an exponent detuned to 0.95. For H_2 the trial function is constructed as follows (cf. Table 4.1). A 1*s* Slater orbital is used on each atom and at the midpoint of the internuclear axis. The linear coefficients are obtained from an SCF calculation using the HONDO program.[68] In addition, a simple Jastrow function of electron-electron and electron-nuclear coordinates is also used, namely

$$J(r_{12}, r_{i\alpha}) = \exp\left(\frac{ar_{12}}{1+br_{12}} - \sum_{i\alpha}\frac{\lambda r_{i\alpha}}{1+\nu r_{i\alpha}}\right). \tag{4.25}$$

In Eq. (4.25), Roman indices denote electrons, while Greek indices denote nuclei and the bond function. The final form of the trial function is,

$$\Psi_T(1, 2) = \psi(1)\psi(2)J(1, 2), \tag{4.26}$$

where ψ is the molecular orbital.

Trial, mixed and pure expectation values obtained for H and H_2 are presented in Table 4.2. When analytic values of A_T or A_m are not available, VMC or QMC results are computed, respectively. The data in Table 4.2 indicate the amount of improvement required by each technique.

Pure expectation values of $<r>$, $<z^2>$ and $<r^2>$ for H and of $<z^2>$ and $<r^2>$ for H_2, computed by the methods discussed in Sec. 4.2, are presented in Tables 4.3 and 4.4, respectively. Since efficiency comparisons between the techniques is our primary concern, detailed studies of time-step bias and effects of finite convergence time, t, are not undertaken since accuracies are generally better than 0.5%.

We point out here, however, that pure expectation values are accurate to only first order in errors introduced by finite t ("convergence bias") and non-zero τ, in contrast to the energy which is accurate to second order. That is, for $\phi_0' = \phi_0 + a(t, \tau)$, where $a(t=\infty, \tau=0) = 0$ and ϕ_0' and ϕ_0 are normalized, there results

$$<\phi_0'|A|\phi_0'> = <\phi_0|A|\phi_0> + <a|A|\phi_0> + <\phi_0|A|a> + <a|A|a> . \qquad (4.27)$$

However, the increased influence of bias on pure expectation values is mitigated by the fact that acceptable accuracies in A_p are much less than those in the energy. For example, if bias in the energy has been reduced to an acceptable level of $\simeq 0.001\%$, then even a two order of magnitude increase in this effect for A_p will not be significant since an error of 0.1% in a dipole or quadrupole moment is quite small. For H and H_2, we see below that convergence and time-step bias are small.

In our data, the worst combination of t and τ, which for the H atom was a time step of 0.05 hartree^{-1} (h^{-1}) and a convergence time of 5 h^{-1}, yields second moments accurate to 1% or better. These small errors at $t = 5\ h^{-1}$ (relative to the 5% error in the mixed averages, $t = 0$) imply that values of P are close to their asymptotic limits, ϕ/Ψ_T. Doubling the convergence time to 10 h^{-1} further increases the accuracy of the second moments to within 0.5%. Though this improvement indicates that further convergence is possible, the smallness of the gain, relative to the total change from the

mixed average, verifies that values of P are well converged. That little time step bias is present may be seen when reducing τ from 0.050 to 0.025 h^{-1}: only a small improvement is obtained (generally less than the statistical error) in the already quite accurate pure expectation values. In the V+QMC approaches, at $t = 5\ h^{-1}$, the errors are noticeably larger than those resulting from the single QMC walk algorithm. This is most likely due to the requirement of sampling (ϕ_0/Ψ_T) twice, compounding the error of incompletely converged values of P. This effect is no longer significant at $t = 10\ h^{-1}$, where the accuracies of all methods are statistically equivalent.

All H_2 calculations were performed with $\tau = 0.01\ h^{-1}$. This time step introduces very little error, as demonstrated by the accuracy of the results of Table 4.4. Since the statistical errors are generally larger than the differences between the means and the exact results, trends are difficult to discern. Generally, it is found that errors in the means are 1% or less at $t = 2\ h^{-1}$, and less than 0.5% at $t = 4$ and $6\ h^{-1}$.

The results of Tables 4.3 and 4.4 show that each algorithm readily produces accurate expectation values, i.e., convergence and time-step bias are small. For a comparison of the efficiency of the various algorithms, all statistical errors in Tables 4.3 and 4.4 correspond to the same amount of computation (10 minutes on a single processor of a Cray/XMP) and therefore provide a direct measure of relative efficiency.

It is immediately apparent that small but consistent improvement in precision results for both H and H_2 when weights are carried in the QMC walk. The average increase in efficiency, (the squared ratios of statistical errors averaged across τ and t), is roughly 60% for H and 130% for and H_2. This improvement does not appear to be strongly dependent on the choices for the upper and lower bounds of the weights. The

weights for all the single-walk calculations were not allowed to exceed a value of 2. The lower bound was chosen as 0.1 or 0.4, and no noticeable change in the efficiency was found for these choices.

We now compare the efficiencies of the V+QMC and correlated V+QMC approaches. As discussed in Sec. 3.5, the difference between a weighted quantity, now A_p, and a VMC value, now A_T, decreases as the convergence time, and hence the statistical error in the weights increase, [cf. Eq. (3.19)]. For H atom (Table 4.3), the increase in efficiency using correlated V+QMC versus V+QMC is about a factor of 10 at $t = 5\ h^{-1}$ and a factor of 6 at $t = 10\ h^{-1}$. For H_2 (Table 4.4), an increase in efficiency of a factor of 6 is obtained at $t = 2\ h^{-1}$ which decreases to 3 at $t = 6\ h^{-1}$. Thus all increases in efficiency obtained by exploiting the correlation between A_p and A_T are substantial.

While the most efficient techniques within the single and double walk algorithms are discernible, the question of which class is better, is not immediately answerable. For H, the smallest statistical errors are obtained by the correlated V+QMC approach. For H_2, on the other hand, the superiority of the correlated V+QMC technique is lost by $t = 6\ h^{-1}$. Since the computation times of A_T and $A_p - A_T$ are in good agreement with the optimum values (not shown), cf. Eq. (3.24), the most probable cause for the loss in performance of correlated V+QMC is the decrease in correlation between the trial and pure expectation values as the convergence time increases. Note that this trend is accelerated for H_2 in comparison to H. For example, at $t = 10\ h^{-1}$ for H, exploiting correlation in the VMC approach increases efficiency by a factor of 6, which is the increase found for H_2 at only $t = 2\ h^{-1}$. Nevertheless, since convergence

to accurate results is obtained for H_2 by $t = 4\ h^{-1}$, the correlated V+QMC technique remains the best approach for H_2 as well as H.

To summarize, employing VMC and QMC walks and exploiting the correlation between trial and pure expectation values is the best approach for the systems studied here. However, the efficiency of the correlated V+QMC method versus a single QMC walk with weighting is strongly dependent on the length of the QMC walk required for convergence of the populations to ϕ_0/Ψ_T, and on how quickly the correlation between A_T and A_p decreases with convergence time. Therefore, the single-walk method may be more competitive in computations of pure expectation values of other systems.

4.4. Properties of LiH and BH

The H and H_2 results of the previous section demonstrate that accurate moments of the electronic charge distribution may be obtained for simple systems. We now consider the computation of pure expectation values for the ground states of LiH and BH at their experimental internuclear separations. These systems represent a greater challenge due to the presence of exchange nodes. As discussed in Chap. 3, the QMC energy resulting from the single-determinant trial functions employed possess a fixed-node error for LiH ($\approx$ 1% of the CE) and BH ($\approx$ 10% of the CE). These trial functions are now used to the compute dipole and quadrupole moments of LiH and BH to determine the effect of fixed-node error on these properties. For LiH, properties are computed with both the single- and double-walk algorithms. In addition to fixed-node error, the effects of the time step and convergence time are also studied.

As described in this chapter, moments are obtained from expectation values of

coordinate operators. (A QMC approach has also been developed which computes a dipole moment from the change induced in the QMC energy by a static electric field.[131]) For an n-electron diatomic molecule with the internuclear axis along x, the dipole and quadrupole operators are given by (in atomic units),

$$\hat{\mu} = \sum_{\alpha} Z_{\alpha} X_{\alpha} - \sum_{i=1}^{n} x_i \ , \tag{4.28}$$

and

$$\hat{\theta} = \frac{1}{2}\sum_{\alpha} Z_{\alpha}(3_{\alpha}^{2} - R_{\alpha}^{2}) - \frac{1}{2}\sum_{i=1}^{n} 3x_i^2 - r_i^2 \ . \tag{4.29}$$

In Eqs. (4.28) and (4.29), Greek indices denote nuclei, Roman indices denote electrons, and the nuclear charges are given by Z_{α}. In the Born-Oppenheimer approximation, i.e., fixed nuclei, the dipole and quadrupole moments, when expressed as expectation values, are

$$\mu \equiv \langle\hat{\mu}\rangle = \sum_{\alpha} Z_{\alpha} X_{\alpha} - n\langle x\rangle \ , \tag{4.30}$$

and

$$\theta \equiv \langle\hat{\theta}\rangle = \frac{1}{2}\sum_{\alpha} Z_{\alpha}(3_{\alpha}^{2} - R_{\alpha}^{2}) - \frac{1}{2}n\langle 3x^2 - r^2\rangle \ . \tag{4.31}$$

Therefore, in the approaches we employ, the dipole and quadrupole moments are obtained from the pure expectation values of x and $3x^2 - r^2$, respectively.

For a charge-neutral species, the computed dipole moment is independent of the origin. However, the quadrupole moment does depend on the coordinate origin when the dipole moment is non-zero. For LiH, the origin was chosen as the midpoint of the internuclear axis. However, it is most common to report quadrupole (and other) moments as computed with respect to the center of mass. For a diatomic molecule,

only the x coordinate differs between the two origins above. Therefore, the center-of-mass (CM) and midpoint (MP) quadrupole moments are related by

$$\theta(\mathrm{CM}) = \theta(\mathrm{MP}) - 2X_{\mathrm{CM}}\mu \,, \tag{4.32}$$

where

$$X_{\mathrm{CM}} = \sum_{\alpha} M_{\alpha}X_{\alpha}/\sum_{\alpha} M_{\alpha} \,. \tag{4.33}$$

For LiH, we employ the experimental internuclear separation of 3.015 *Bohr* and the nuclear masses of the periodic table, $M_{\mathrm{H}} = 1.0079$ and $M_{\mathrm{Li}} = 6.941$ atomic mass units, yielding $X_{\mathrm{CM}} = 1.1252$ *Bohr*.

The LiH trial function is taken from a single Slater determinant, the basis set is given in Ref. 43, and the correlation functions are described by Eqs. (3.4), (3.5), and (3.9). The correlation function parameters ($a_2 = b_2 = 0$) are presented in Table 4.5. The parameters λ_{H} and λ_{Li} are chosen to satisfy their respective cusp conditions, as described in Sec. 3.2, to minimize time-step bias. (A programming error yielded $\lambda_{\mathrm{H}}(\mathrm{cusp}) = 0.1025$ while the correct value, cf. Fig. 3.8, is 0.61. However, as seen in Fig. 3.11, time-step bias is small.)

Chronologically, the first algorithm developed by the author was the one most analogous to diffusion QMC computations of the energy, i.e., the single-walk approach with the branching simulated by integer rounding. As seen in Sec. 4.3, accurate results were obtained for H_2 which motivated further investigation of the capabilities of the single-walk approach with calculations on LiH. To investigate the effects of time-step bias (non-zero τ) and convergence bias (finite t), single-walk computations of mixed and pure properties were performed. Results for these, trial, and second-order expectation values are reported in Table 4.6.

The first significant point is that the mixed dipole moments are found to be worse or equivalent to the trial values. Therefore, the second-order approximation is entirely useless in obtaining improvements over mixed dipole moments, cf. Table 4.6. The second-order quadrupole moments, on the other hand, are seen to give accurate estimates of the pure values.

The limited number of computations in Table 4.6 and the size of the statistical errors does not allow observations to be conclusive. Nonetheless, important trends are indicated by the data. In examining the dipole moments for example, we see that very little time-step bias appears to be present at $\tau = 0.010$ and $0.005\ h^{-1}$. We also observe that at $\tau = 0.020\ h^{-1}$ the mixed dipole moment, μ_m, is somewhat higher than μ_m at the smaller time steps suggesting that some time-step bias is present at $0.020\ h^{-1}$. Since μ_p tends to converge downward from μ_m towards the exact (note that $t = 0$ yields mixed results), a time-step bias which increases μ_m will also increase μ_p. Yet we find that μ_p is most accurate at $\tau = 0.020\ h^{-1}$. This result, therefore, is due to the increase in convergence time which, apparently, more than compensates for the increase in time-step bias. In ascertaining the degree of convergence to asymptotic populations, we now consider values of $|\mu_p - \mu_m|$. Note that $|\mu_p - \mu_m|$ is only 0.023 and 0.020 Debyes (D) at $t = 4\ h^{-1}$ ($\tau = 0.005$ and $0.010\ h^{-1}$, respectively). In contrast, at $t = 8\ h^{-1}$, convergence from the mixed is over three times greater, 0.71 D. This comparison indicates that convergence is not complete at $t = 4\ h^{-1}$ and is therefore the major source of error in these two calculations. A word of caution is in order, however, since the statistical errors are of the order of the differences just discussed. Most important, our most highly converged dipole moment, $t = 8$ and $\tau = 0.02\ h^{-1}$, is

in excellent agreement with experiment, within 0.6(8)%. As a final point, we note that increases in efficiency should be readily obtained by carrying weights, cf. Sec. 4.2.2. Uncompleted calculations indicate that the single-walk algorithm with weighting will yields improvements in efficiency of a factor of 2-4. The calculations presented here each took from 1-3 hours on a Cray/XMP.

In considering the quadrupole moment, the computation of the pure value actually appears to be easier for this quantity. For example, at $t = 4\ h^{-1}$,$|\theta_p - \theta_m| = 0.18\ B\,(B = 10^{-26} esu\text{-}cm^2)$ while this difference increases to only 0.20 B at $t = 8\ h^{-1}$. Thus convergence now appears to be attained at 4 h^{-1} in contrast to the dipole moments which converge at $t \geq 8\ h^{-1}$. Unfortunately, an exact value of θ is unknown, however, we do compare our QMC result with others in Table 4.8.

In addition to single-walk calculations of properties, the V+QMC approach has also been applied to LiH. Two time steps, 0.020 and 0.010 h^{-1}, and several convergence times, $0 - 8\ h^{-1}$, have been employed.

As described in Sec. 4.3, points are sampled from $|\Psi_T|^2$ and then two estimates of the asymptotic population (for several convergence times if desired) are obtained for each point. Subsequently, the VMC walk is continued and a new set of points is sampled from $|\Psi_T|^2$ at a later time and so on. It is highly desirable that points sampled from $|\Psi_T|^2$ are uncorrelated. If correlation does exist between points sampled from $|\Psi_T|^2$, then sampled values of the coordinate operator (A) and the asymptotic population (P) will also be correlated. To the extent that values of A and P are correlated, computational effort is wasted in the relatively time consuming QMC side walks which

sample P. This concern becomes more important as the convergence time increases.

For H and H_2, correlation in points sampled from $|\Psi_T|^2$ is not problematic. Pure expectation values were computed efficiently with guesses, simply based on experience in previous QMC energy computations, of the number of VMC steps between sampling values of A and P from $|\Psi_T|^2$. This is not the case for LiH. In an initial set of calculations, points were sampled from from $|\Psi_T|^2$ after every ten steps with τ (VMC) $= 0.25\ h^{-1}$. The acceptance ratio was 0.6 yielding an actual diffusion time of 1.5 h^{-1} between samplings. While this amount of time between samplings seemed sufficient, apparently it is not as the efficiency of the V+QMC calculations was quite poor. Thus, the question of sampling configuration space in the VMC walk requires greater consideration for LiH.

One method of addressing the concern above is to simply increase the time between samplings from $|\Psi_T|^2$. This approach, however, is a rather time consuming method for ascertaining the best way to sample configuration space in the VMC walk (since QMC side walks are also performed). A, perhaps, better alternative is to simply select a very large ensemble of uncorrelated points from $|\Psi_T|^2$ which can be accomplished quickly. One may then compute the trial values of the energy, dipole and quadrupole moments (as well as other expectation values if this is deemed useful) from this ensemble of points. Comparison of these averages and their statistical errors with those from VMC calculations then allows a judgement to be made concerning the adequacy of the ensemble size. Caution must still be applied, however, since the ensemble averages may be fortuitously close to precisely obtained trial values, and, likewise, statistical error in these averages may be underestimated due to the limited size of the

ensemble.

Following the approach outlined in the preceding paragraph, 30,000 points were randomly chosen and then converged to $|\Psi_T|^2$ in a VMC walk of 25 h^{-1} (the diffusion time was 19 h^{-1}). The ensemble averages of the dipole and quadrupole moments for this ensemble are given in Table 4.7 as the $t = 0$ properties. Comparison of the trial (VMC) moments in Table 4.6 with the ensemble averages in Table 4.7 indicates that the ensemble is sufficiently large. This ensemble of 30,000 points is then divided into 30 subensembles of 1,000 points, and then propagated 8 h^{-1} at $\tau = 0.02$ and 0.01 h^{-1} in sampling $|\phi_0/\Psi_T|^2$. The computational cost for these time steps is 30 and 60 minutes of Cray/XMP computer time. Results are presented in Table 4.7.

The first point of interest, which implies why convergence to an accurate dipole moment can be difficult for LiH, is that the accuracy of this moment actually decreases in the initial stages of converging to asymptotic populations, compare the $t = 2$ and $t = 0$ h^{-1} dipole moments for both time steps. The quadrupole moments, on the other hand, converge monotonically.

The most puzzling aspect of the data is that the $\tau = 0.02$ h^{-1}, $t = 8$ h^{-1} moments disagree with those of the single-walk approach. In considering this, we point out that time-step bias should be equivalent between the single- and double-walk methods. In both approaches the time-step bias enters in twice, in $\Psi_T\phi_0$ and ϕ_0/Ψ_T in the single walk and in $|\phi_0/\Psi_T|^2$ in the V+QMC simulation. However, convergence bias is more significant in the V+QMC algorithm. As discussed in Sec. 4.3, convergence bias is compounded with V+QMC, since a product of asymptotic populations is employed, but not with the single walk. This effect manifested itself in the larger errors obtained

with V+QMC, in comparison to the single walk, at the smaller convergence times for H and H_2. In this light, note that for both time steps in Table 4.7 values of μ_p and θ_p are still changing at the larger convergence times. While these changes appear to be well beneath the statistical noise, in fact they probably are not. This is because the statistical error in, for example, $\mu_p(t{=}6) - \mu_p(t{=}8)$ is roughly an order of magnitude smaller than that given by the statistical error of the two values of μ_p due to the high degree of correlation between values at similar convergence times. Therefore, larger values of t should be employed in future calculations in order to ascertain convergence to $|\phi_0|^2$.

While the $\tau = 0.01\ h^{-1}$ moments are very good. The question of convergence still remains. In addition, larger convergence times would seem to give a dipole moment which is too low contradicting the implications of the single-walk calculations. For this reason, it is important to verify the adequacy of the ensemble size employed here by generating more ensembles of the same size and comparing the accuracy of averages over these ensembles with those of the original.

Finally, we have found that while the statistical error in $\Delta \equiv \mu_p - \mu_T$ is smaller than that of μ_p at the smaller convergence times, these statistical errors were equivalent by $t = 8\ h^{-1}$. Therefore, no benefit appears to be derived from the correlated V+QMC approach in the context of computing moments for LiH. Comparisons of efficiency between the single- and double-walk algorithms will be made when weights are carried in the single-walk approach and when the double-walk calculations are more complete.

Table 4.8 compares our single-walk ($t = 8\ h^{-1}$, $\tau = 0.02\ h^{-1}$) moments with the

exact (when available) and those obtained in other *ab initio* computations. As has already been stated the QMC dipole moment is quite accurate; similar accuracies are also obtained by other theoretical approaches. The differential dipole moment, computed from the difference of statistically correlated QMC energies with and without a small static electric field, is also of reasonable quality. However, this moment was obtained employing the trial function of Ref. 43 which yielded a rather poor QMC energy (since the linear coefficients in the MOs are not optimized) of -8.059(4) h in comparison to ours, -8.0691(4) h (the exact is -8.0702 h). Therefore, it is expected that the differential dipole moment will be improved using a more accurate trail function. In considering the effect of fixed-node error of the dipole moment, we note that this error in the QMC energy is 0.01% of the total energy, and 0.6% in the dipole moment. Thus, although the fixed-node error rises dramatically for the dipole moment, the overall accuracy is still quite good.

The more interesting property studied here is the quadrupole moment for which no exact value exists. A noticeable discrepancy of 0.16 B exists between recent theoretical calculations of this moment, cf. Table 4.8. If the QMC value of θ possesses an accuracy similar to that of μ, then our value of θ should be quite accurate as well (and in favor of the CASSCF result). It will be of great interest to compare our value of the quadrupole moment with an exact estimate, if it is ever obtained, as this would be a significant test of the predicative capability of the QMC approach employed here.

A very limited set of properties calculations has been performed for BH. The trial function is Ψ_1 given in Table 3.15 and the single-walk approach (without carrying weights) was employed with $t = 3\ h^{-1}$ in all calculations. The effect of deleting a

walker upon crossing a node versus rejecting the move is examined. For the deletion approach, calculations were performed at $\tau = 0.005$ and $0.010\ h^{-1}$ which used 2 and 1 hour of Cray/XMP computer time, respectively. No time-step bias is discernible for these two computations. The last calculation used a time step of $0.010\ h^{-1}$ and rejection and took three hours.

Table 4.9 compares QMC results with other *ab initio* values and experiment (apparently only a few computations of properties for BH exist). We see that for the computation times above, the statistical errors in the QMC moments are reasonable ($\approx$ 3-4%). However, the convergence time of only $3\ h^{-1}$ (compare with those of LiH) suggests that convergence is not attained. The most significant result is that the "rejection" value of the dipole moment is noticeably different from the "deletion" moment and in much better agreement with experiment, i.e., now within the (large) experimental error bar. However, since no time-step bias was observed in the deletion properties at the time steps employed this discrepancy, albeit large, must remain a matter of conjecture.

The BH calculations show that a reasonable dipole moment may be obtained (if the rejection value is correct) with a single-determinant trial function. However, the accuracy of the quadrupole moment, which does show a substantial disagreement with the MCSCF value of Bishop *et al.*[136], can not be ascertained given the paucity of theoretical calculations and the absence of an estimated exact value. As noted in Chap. 3, the QMC energy for the trial function employed here gives only 89% of the CE, a much larger error than that of LiH. While it is of interest to determine the effect of this fixed-node error on QMC-computed moments, further pursuit of accurate

moments should include consideration of a multi-determinant trial function which will, hopefully, yield an improved QMC solution.

4.5. Moments with Odd Powers

In this section we explore more efficient approaches for computing expectation values of coordinate operators with odd powers. For such expectation values, cancellations will occur from different regions of space. Therefore, sampling techniques which exploit this cancellation are preferred.

The one-dimensional computation of $<x> = \int x\,\rho(x)dx$, (ρ is assumed to be normalized) serves as an example. As the symmetry of $\rho(x)$ about $x = 0$ increases,[138] the degree of cancellation of the integrand also increases. Therefore, lower variance estimates may be obtained by sampling a new distribution, ρ' (e.g., by sampling ρ'/ρ from the original distribution ρ) which will exploit this cancellation.

Probably the most straightforward approach is to choose ρ' as the antisymmetric component of ρ, namely $\rho'(x) = \rho_A(x) = \frac{1}{2}[\rho(x) - \rho(-x)]$. Since x is an odd function, the symmetric component of ρ vanishes upon integration, yielding

$$<x> = \int x\,\rho_A(x)dx \ . \tag{4.34}$$

In sampling from ρ, $<x>$ may now be computed by averaging $x\,\rho_A(x)/\rho(x)$. Note that if ρ is an even function, all sampled values vanish identically so that the correct result is obtained with no statistical error. More generally, of course, ρ is not fully symmetric, and only a finite statistical error reduction ensues. For example, consider a one-dimensional distribution that, like for a heteronuclear molecule, possesses unequal exponentially decaying tails at large $\pm x$. The antisymmetric part of ρ_A then also

possesses such tails, though now of differing sign. Thus little error reduction is expected in this case. In addition, sampled values of $x\rho_A(x)/\rho(x)$ will be unbounded in regions where $\rho(x)$ is small but $\rho(-x)$ is not, thus increasing statistical error. Therefore, depending on $\rho(x)$, the reduction in statistical error may not be significant.

The problems with the naive choice of $\rho' = \rho_A$ (that the spread in ρ' is not optimally reduced and that ρ'/ρ may be poorly behaved) are addressed by a new choice of ρ', is given by

$$\tilde{\rho}(x) \equiv \rho(x) - \rho(-x) \qquad \text{for } \rho(x) \geq \rho(-x)$$
$$0\,, \qquad \text{for } \rho(x) < \rho(-x)\,. \tag{4.35}$$

It is easily shown that $<x> = \int x\tilde{\rho}(x)dx$. However, the spread in $\tilde{\rho}$ should be less than that of ρ or ρ_A. If tails are present in the original distribution, one of them is immediately eliminated. Also, $\tilde{\rho}$ possesses the zero-variance property (that ρ_A does) when ρ is an even function. Finally, by construction the quantity to be averaged in sampling from ρ, namely $x\tilde{\rho}/\rho$, will remain bounded for small ρ since $0 \leq \tilde{\rho}/\rho \leq 1$.

To take a simple example, consider $\rho(x) = \frac{1}{2}(1 + x)$, $|x| \leq 1$. From the definitions above, $\rho_A(x) = x/2$ on $|x| \leq 1$ and $\tilde{\rho}(x) = x$ on $0 \leq x \leq 1$. The variance obtained from $\tilde{\rho}$ is given by

$$V_{\tilde{\rho}} \equiv \int_{-1}^{1} \rho(x)[x\tilde{\rho}(x)/\rho(x)]^2 dx - <x>^2 = 0.11\,. \tag{4.36}$$

Efficiency is thus doubled over sampling from ρ directly, for which $V_\rho \approx 0.22$. In contrast, V_{ρ_A} is infinite because of the singularity in ρ_A/ρ at $x = -1$.

A more realistic example is offered in Fig. 4.3, where ρ is the square of a trial function for LiH at a fixed distance from the internuclear axis (x). Comparing the

solid line, ρ, to the dotted, $\tilde{\rho}$, shows the narrowing of the distribution achieved, even though ρ is far from an even function. A numerical computation for the variance yields $V_{\tilde{\rho}}(x) \approx \frac{1}{2}V_{\rho}(x)$. Applying these ideas is most readily accomplished within the V+QMC approach. Specifically, values of ϕ_0/Ψ_T for each point $\underline{R}$ sampled from $|\Psi_T|^2$ are required at $\underline{R}$ and at a symmetry point (or points) related to $\underline{R}$ by reflection or inversion. If the symmetry is sufficiently high, the extra computation resulting from the necessary sampling of added values of $|\phi_0|^2$ will be compensated by reductions in the statistical error in the averages. The most pronounced reductions will be found for charge distributions which are nearly symmetric. For example, for molecules such as CO, substantial reductions in statistical error, versus a straightforward sampling of $|\phi_0|^2$, should be observable employing the difference technique described here.

Table 4.1. Parameters for the H_2 trial function.

c (H)	0.48610	a	0.50
c (BF[a])	0.11089	b	0.50
Exponent	1.19000	λ	0.15
		ν	1.00

[a] The bond function is located at the midpoint of the internuclear axis.

Table 4.2. Comparison of trial, mixed and pure expectation values for H and H_2.[a]

	H		
	$\langle r \rangle$	$\langle z^2 \rangle$	$\langle r^2 \rangle$
Trial	1.5789	1.1080	3.3241
Mixed	1.5385	1.0519	3.1558
Second-order[b]	1.4981	0.9958	2.9875
Pure	1.5000	1.0000	3.0000

	H_2	
	$\langle z^2 \rangle$	$\langle r^2 \rangle$
Trial (VMC)	1.0787(6)	2.6228(11)
Mixed (QMC)[c]	1.0491(8)	2.5809(14)
Second-order[b]	1.0195(10)	2.5390(18)
Pure (exact)[d]	1.0230	2.5464

[a] Results for the H atom are exact, analytic values.

[b] See Eq. (4.5).

[c] Time step for QMC results is 0.01 h^{-1}.

[d] Ref. 10.

Table 4.3. Pure expectation values for H atom by various Monte Carlo methods.[a]

Method	(t, τ^b)	$\langle r \rangle$	$\langle z^2 \rangle$	$\langle r^2 \rangle$
Single QMC walk: Integer rounding	(5, 0.050)	1.5058(14)	1.0075(32)	3.025(6)
	(10, 0.050)	1.5038(15)	1.0066(31)	3.015(7)
	(10, 0.025)	1.5025(21)	1.0029(41)	3.010(9)
Single QMC walk: Integer rounding *and* weighting	(5, 0.050)	1.5052(11)	1.0084(25)	3.024(5)
	(10, 0.050)	1.5040(12)	1.0024(25)	3.018(6)
	(10, 0.025)	1.5023(15)	1.0040(33)	3.009(7)
V+QMC	(5, 0.050)	1.5074(15)	1.0150(40)	3.033(6)
	(10, 0.050)	1.5032(22)	1.0018(51)	3.012(9)
	(10, 0.025)	1.5026(23)	1.0008(53)	3.012(11)
correlated V+QMC	(5, 0.050)	1.5094(5)	1.0144(11)	3.045(2)
	(10, 0.050)	1.5024(7)	1.0029(21)	3.010(4)
	(10, 0.025)	1.5020(9)	1.0025(24)	3.009(5)
Exact[c]		1.5000	1.0000	3.000

[a] Units for length and time are bohr and h^{-1}, respectively. Statistical errors, in parentheses, represent one standard deviation in the mean of a Gaussian distribution, and are normalized to correspond to 10 minutes of computation time on a Cray/XMP.

[b] t denotes convergence time; τ is the time step.

[c] t See Table 4.2.

Table 4.4. Pure expectation values for H_2.[a]

Method	t	$\langle z^2 \rangle$	$\langle r^2 \rangle$
Single QMC walk: Integer rounding	2	1.028(9)	2.554(14)
	4	1.024(8)	2.549(14)
	6	1.021(9)	2.549(15)
Single QMC walk: Integer rounding *and* weighting	2	1.034(6)	2.554(10)
	4	1.020(5)	2.535(9)
	6	1.028(6)	2.545(12)
V+QMC	2	1.034(5)	2.563(9)
	4	1.022(8)	2.548(12)
	6	1.027(8)	2.550(21)
correlated V+QMC	2	1.033(2)	2.562(4)
	4	1.026(3)	2.549(6)
	6	1.026(6)	2.550(10)
Exact[b]		1.023	2.546

[a] Units are given in Table 4.3. The time step is 0.01 h^{-1} in all calculations shown; the statistical errors correspond to 10 minutes of computation on a Cray/XMP.

[b] Ref. 10.

Table 4.5. LiH correlation function parameters.

a_1	0.50000
b_1	1.50000
λ_H	0.10250
ν_H	2.00000
λ_{Li}	0.03075
ν_{Li}	0.02500

Table 4.6. Single-walk properties for LiH.

(t, τ^{a})	Dipole Moments[b]			Quadrupole Moments[c]		
	Mixed	Pure	Second-order[d]	Mixed	Pure	Second-order[d]
(8, 0.020)	5.932(18)	5.861(46)	5.951(36)	-4.25(5)	-4.05(17)	-4.12(10)
(4, 0.010)	5.909(32)	5.886(46)	5.905(64)	-4.17(9)	-3.98(13)	-3.96(18)
(4, 0.005)	5.914(18)	5.894(28)	5.915(36)	-4.23(7)	-4.07(13)	-4.08(16)
Trial	5.913(6)			-4.38(1)		
Exact[e]	5.828			---		

[a] t is the convergence time in obtaining asymptotic populations and τ is the time step, both are in h^{-1}.
[b] Units are 10^{-18} *esu–cm* (*Debyes*).
[c] Units are 10^{-26} $esu\text{–}cm^2$ and are reported with respect to the center of mass. [d] See Eq. (4.3).
[e] A vibrationless value derived from experiment in Ref. 132 by extrapolating to −1/2 vibrational quantum number.

Table 4.7. V+QMC properties for LiH.[a]

	$\tau = 0.02$	
t	μ_p	θ_p
0[b]	5.910(24)	-4.40(5)
2	5.944(27)	-4.37(7)
4	5.932(32)	-4.31(8)
6	5.918(38)	-4.29(9)
8	5.910(42)	-4.24(10)
Exact[c]	5.828	---

	$\tau = 0.01$	
t	μ_p	θ_p
0[b]	5.910(24)	-4.40(5)
2	5.923(26)	-4.33(7)
4	5.902(31)	-4.28(9)
6	5.865(33)	-4.22(10)
8	5.835(32)	-4.16(10)
Exact[c]	5.828	---

[a] Units for the convergence time (t), time step (τ), dipole moment (μ_p), and quadrupole moment (θ_p) are given in Table 4.6.

[b] $t = 0$ corresponds to the trial values.

[c] Ref. 132.

Table 4.8. Properties of LiH by various approaches.[a]

Method	μ	θ
HF[b]	6.00	-4.51
CI[c]	5.85	-4.16
Largest CI[d]	5.86	---
CASSCF[e]	5.83	-4.13
Coupled-Cluster[f]	5.87	-4.29
Differential QMC[g]	5.77(8)	---
QMC[h]	5.86(5)	-4.05(17)
Exact[i]	5.83	---

[a] Units are given in Table 4.6.

[b] Ref. 117.

[c] Ref. 133.

[d] Ref. 112.

[e] Ref. 134.

[f] Ref. 21.

[g] Ref. 131.

[h] This work. Single walk with $t = 8\ h^{-1}$ and $\tau = 0.02\ h^{-1}$.

[i] Ref. 132.

Table 4.9. Properties for BH.[a]

Method	μ	θ
HF[b]	1.74	-3.59
MCSCF[c]	1.32	-3.09
Trial[d]	1.75(4)	-4.39(2)
QMC(deletion)	1.68(5)	-3.83(9)
QMC(rejection)	1.44(6)	-4.02(12)
Experiment[e]	1.27(21)	---

[a] Units are given in Table 4.6 and quadrupole moments are with respect to the center of mass.

[b] Ref. 135.

[c] Ref. 136.

[d] Trial function Ψ_1, see Table 3.15.

[e] Ref. 137.

Figure 4.1. Tagging algorithm for single QMC walk method. A "family tree" for a single walker starting at the origin is shown. Each ring outward corresponds to one generation (or an increase in time by τ). Location in the diagram identifies a walker with its $\theta(k, t)$ label. All descendants of a walker will be in the range of angles from θ to $\theta + \delta$, making identification of progeny possible for all future generations. This enables one to determine convergence to asymptotic populations in a single calculation. An ensemble of trees may be readily treated as described in the text.

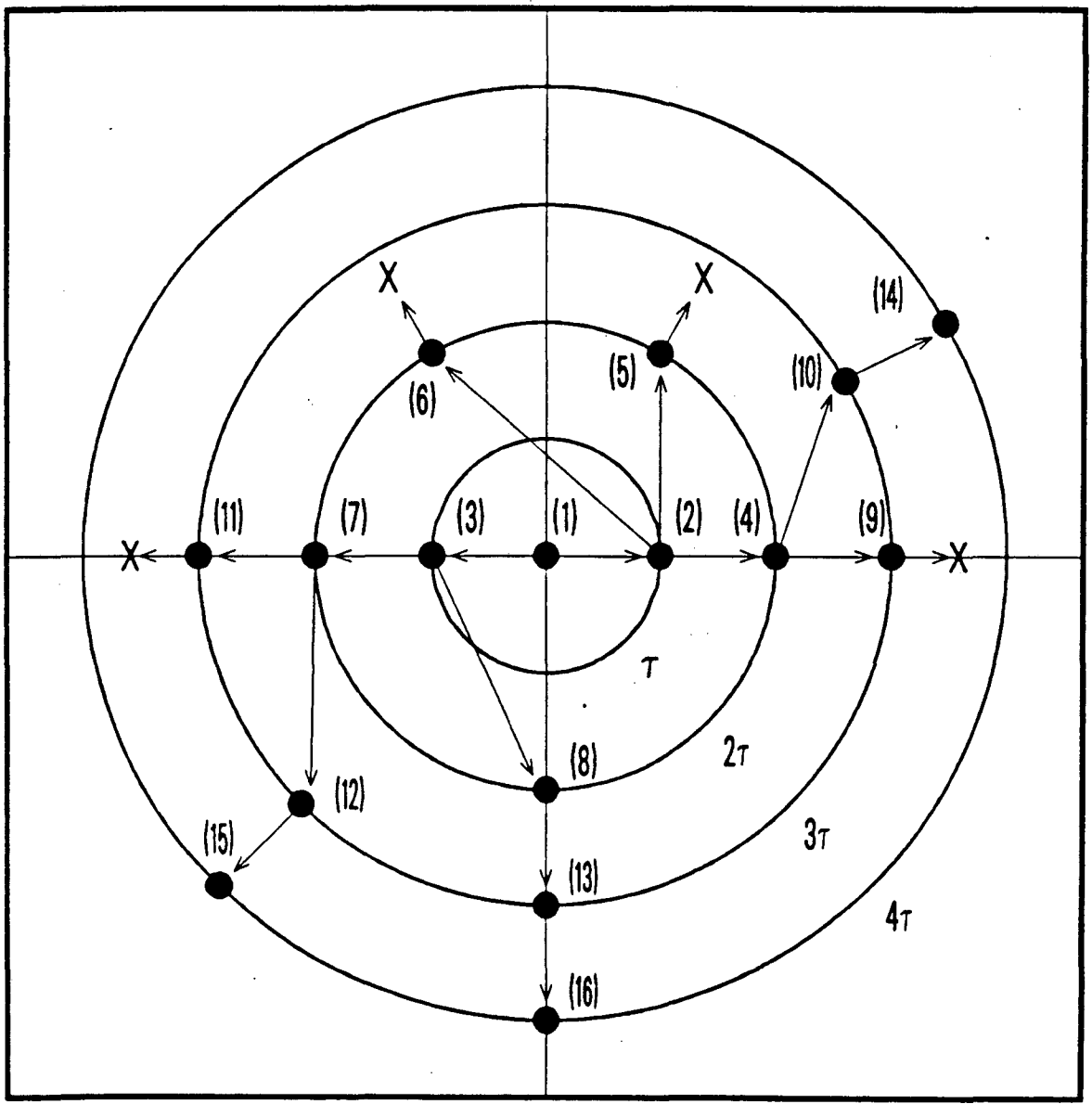


Figure 4.2. Tagging algorithm for VMC with QMC side walks method. The family trees generated are shown for of five points sampled from $|\Psi_T|^2$. Since only the descendants of the points beginning each QMC walk are tracked, the tagging algorithm is very simple, as indicated here and discussed in the text. Weights are carried with the branching walkers so that the asymptotic population of a point is the sum of the weights of its descendants at sufficiently large t. For each initial point, *two* QMC walks are employed in order to obtain two statistically independent samplings of ϕ_0/Ψ_T and, therefore, an unbiased estimate of $|\phi_0|^2/|\Psi_T|^2$. As in the single QMC walk approach, results may readily be computed for several convergence times (t).

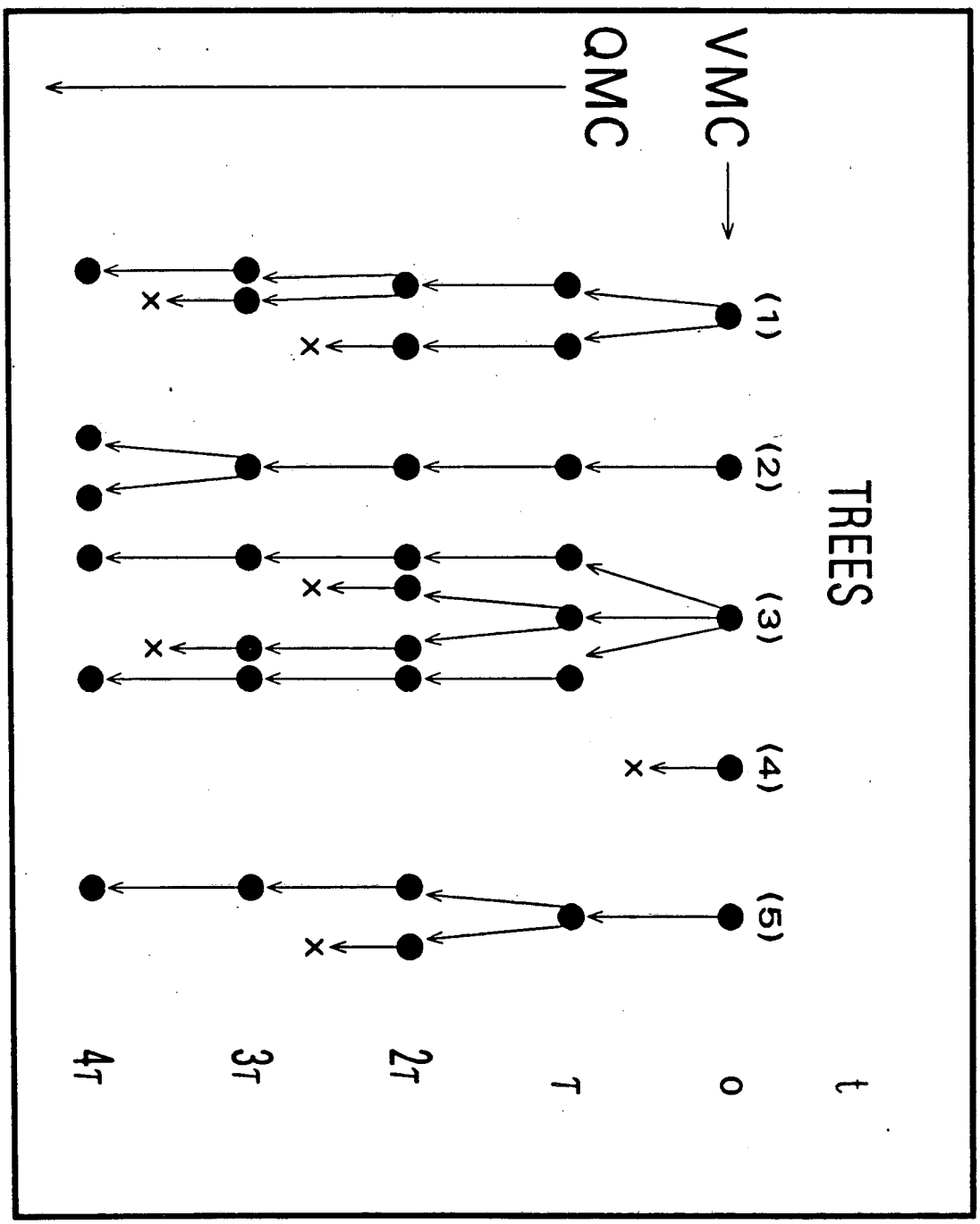
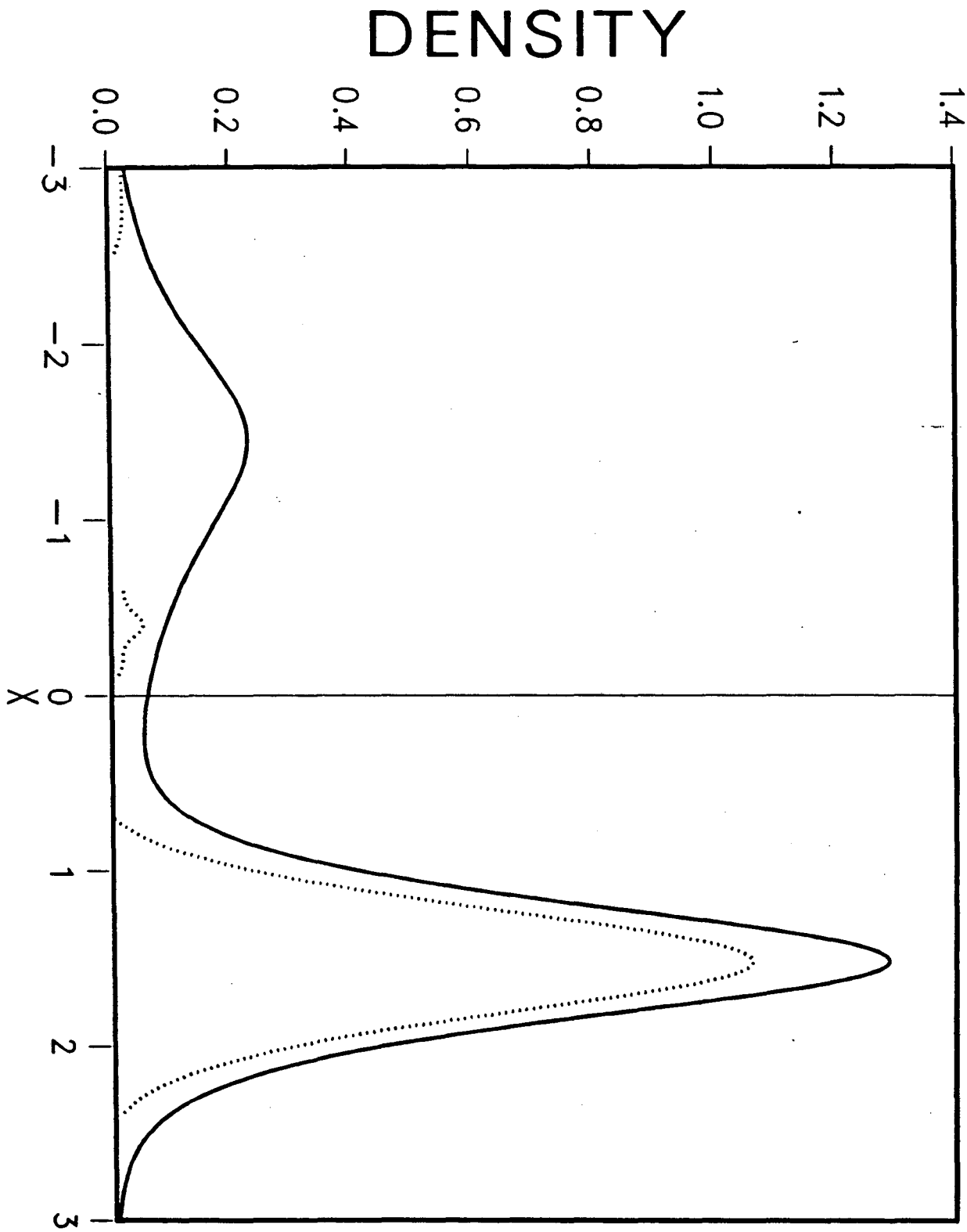


Figure 4.3. Equivalent distributions for the computation of $\langle x \rangle$. The solid line represents the original distribution, ρ, while the dotted one represents the difference distribution $\tilde{\rho}$, cf. Eq. (4.35). The average value of x is identical over both distributions, but the variances are not.



Chapter 5

The Computation of Transition Dipole Moments

5.1. Introduction

The work discussed thus far began with the implementation of the well-developed[39,43], though not widely applied, fixed-node diffusion QMC approach for computing electronic energies, cf. Chap. 3. Subsequently, novel algorithms, designed to sample $|\phi_0|^2$, have been developed and explored in calculations on small atoms and molecules, cf. Chap. 4. The approaches and applications detailed in Chaps. 3 and 4 have considerably broadened the scope of single-state computations by diffusion QMC. In this chapter, we explore an entirely new area – the sampling of several states simultaneously so that multi-state properties may be computed. The techniques described below allow the computation of transition matrix elements. Since transition dipole moments are by far the most significant class of transition matrix elements, we focus on, but are not limited to, the computation of these moments.

Transition dipole moments are of great interest in chemistry and physics since they appear in the expressions for transition rates and probabilities for photon- or electric field-induced atomic and molecular state changes. From first-order time-dependent perturbation theory in the dipole approximation, the probabilities per unit time for absorption, induced emission and spontaneous emission are proportional to the square

of the transition dipole moment between the two states of interest.[139] This moment can be written in several forms.[140] For a transition from state m to state n, the transition dipole moment in the frequently used length form is expressed as

$$|\vec{r}_{mn}|^2 = |\langle\phi_m|\sum_i \vec{r}_i|\phi_n\rangle|^2, \tag{5.1}$$

where the sum is over all electrons.

Also related to the dipole transition moments are experimentally observable quantities such as oscillator strengths, f, and mean lifetimes, τ. The oscillator strength is defined as the ratio of the absorption coefficient integrated over frequency to the value this quantity would have if each molecule (atom for elemental substances) of the absorbing material were replaced by an harmonically bound electron. For a transition from state m to state n, the oscillator strength may be written as,[141]

$$f_{mn} = \frac{2}{3}(E_n - E_m)|\vec{r}_{mn}|^2 \tag{5.2}$$

where $E_n > E_m$, and the transition dipole moment is averaged over the initial and summed over the final state degeneracies. The mean lifetime of a state n obtained by excitation from state m, is given by[141]

$$\tau_n = \frac{3}{2} g_n/(g_m f_{mn} \bar{\nu}^2). \tag{5.3}$$

Here τ_n is in seconds, $\bar{\nu}$ is the energy of the absorbed photon in cm^{-1}, and g_i is the degeneracy of state i.

Despite the importance of transition dipole moments, computing them remains problematic for all but the simplest systems. This is because $\vec{r}_{mn}$ is *very* sensitive to the quality of the wavefunctions used,[22] so much so that wavefunctions which yield

very accurate energies can still give inaccurate transition moments. Also, the convergence properties of the transition dipole moment with basis set and configuration expansion are generally not known. These difficulties encountered by standard quantum mechanical methods is a strong motivation for investigating the possibility that accurate transition dipole moments can be computed by QMC. Towards this end, we investigate approaches for sampling from more than one state. Though our approach uses diffusion QMC with the fixed-node approximation for treating Fermi statistics, two of the methodologies developed here may be applied directly to other QMC methods.

In the next section, we demonstrate how approximations to the transition dipole moment, accurate to first order in the error in the trial function, may be combined to obtain second-order estimates. In Sections 5.3-5.5, three different methods for computing exact transition dipole moments are discussed. A presentation and discussion of results for the $1s \rightarrow 2p_x$ transition dipole moment of the hydrogen atom concludes each of these sections. In Section 5.6, the three methods are compared and their relative merits and deficiencies are discussed. The final section describes oscillator strength calculations for the $2^2S \rightarrow 2^2P$ transition of Li.

5.2. Approximate Transition Dipole Moments

Let us define ϕ_1 and ϕ_2 as the two states between which we wish to compute a transition dipole moment, and Ψ_{T_1} and Ψ_{T_2} as the trial functions for these states. Several approximations to the transition dipole moment may be computed using QMC. These approximations are

$$A_{tt} = \langle \Psi_{T_1} | A | \Psi_{T_2} \rangle \ , \tag{5.4a}$$

$$A_{1t} = \langle \phi_1 | A | \Psi_{T_2} \rangle \ , \tag{5.4b}$$

and

$$A_{2t} = \langle \Psi_{T_1} | A | \phi_2 \rangle \ , \tag{5.4c}$$

where, for notational convenience the equations above are in terms of normalized wavefunctions. Since asymptotic populations (to be employed in the next two sections) are written in terms of normalized eigenfunctions, ϕ_i, and in other expressions ϕ_i appears to equal powers in both the numerator and denominator, QMC eigenfunctions are taken to be normalized without loss of generality. The normalization of trial functions will be discussed when required.

The "trial" matrix element, A_{tt}, is readily computed using VMC by averaging $\Psi_{T_1} A \Psi_{T_2} / |\Psi_T|^2$ over $|\Psi_T|^2$. The "mixed" matrix elements, A_{1t} and A_{2t}, are obtained from from a QMC walk. For example, to compute A_{1t}, $\Psi_{T_1}^{-1} A \Psi_{T_2}$ is averaged over $\Psi_{T_1} \phi_1$. This average is then multiplied by $\langle \Psi_{T_1} | \phi_1 \rangle$ to give A_{1t}. (Methods for computing overlap integrals are described later.)

Writing $\Psi_{T_i} = \phi_i + \delta_i$, the above three approximations are all accurate to first order in the error in the trial functions. However, they may be combined to obtain a "second-order" estimate of $\langle \phi_1 | A | \phi_2 \rangle$. In particular,

$$A_s \equiv A_{1t} + A_{2t} - A_{tt} = \langle \phi_1 | A | \phi_2 \rangle - \langle \delta_1 | A | \delta_2 \rangle \ . \tag{5.5}$$

To demonstrate the quality obtainable from these second order estimates, we compute A_s analytically for the transition dipole moment ($A = x$) of the $1s \rightarrow 2p_x$ transition of the hydrogen atom. Results are presented in Table 5.1. The trial functions, Ψ_{T_1} and Ψ_{T_2} are $1s$ and $2p_x$ Slater orbitals with exponents of 0.90 and 0.45, respectively.

The exponents are chosen such that the variational energy, $<\Psi_{T_i}|H|\Psi_{T_i}>$, differs by 1% from the exact value. This choice was made to simulate the accuracy of trial functions generally used in QMC when studying more complex systems. As seen in Table 5.1, the error in the second-order approximation is an order of magnitude less than the errors in any of the first-order approximations.

5.3. Exact Transition Dipole Moments: QMC Walks within QMC Walks (Method 1)

In this section, we discuss a method which involves propagating a QMC walk for each state under study. Though this approach (as well as the following two) can be applied to any number of states simultaneously, for simplicity, we limit the discussion to the study of only two states.

We wish to compute the matrix elements $<\phi_1|A|\phi_2>$ where specifically we focus on $A = x, y$ or z, and ϕ_1 and ϕ_2 are two different energy eigenfunctions. This method begins with a straightforward implementation of QMC using Ψ_{T_1} as the trial function to generate the distribution $f_1 \propto \Psi_{T_1}\phi_1$ (cf. Chap. 2). Here, Ψ_{T_1} is a trial function chosen to describe state "1", and ϕ_1 is the exact (fixed-node QMC) solution for this state. We now seek an operator, O, sampled from f_1 such that

$$<O>_{f_1} \equiv \frac{<\Psi_{T_1}|O|\phi_1>}{<\Psi_{T_1}|\phi_1>} = <\phi_1|A|\phi_2> . \tag{5.6}$$

From Eq. (5.6), the transition matrix element is obtained when

$$O \equiv s_1 A \phi_2/\Psi_{T_1} , \tag{5.7}$$

where overlap integrals are written as $s_i = <\Psi_{T_i}|\phi_i>$ for normalized trial functions.

The first step in obtaining values of O is to sample a second eigenfunction employing an auxiliary walk guided by a second trial function, Ψ_{T_2}. This walk commences with points selected from $\Psi_{T_1}\phi_1$. The structure of the two walks is shown schematically in Fig. 5.1. As shown in Chap. 4, for a QMC walk beginning at $\underline{R}$, the number of points at a later time is asymptotically proportional to the ratio of the QMC solution to the trial function at $\underline{R}$, in this case $\phi_2(\underline{R})/\Psi_{T_2}(\underline{R})$. Thus values of ϕ_1 are sampled from a standard QMC walk, which yields $\Psi_{T_1}\phi_1$, and values of ϕ_2 from an auxiliary QMC walk, guided by Ψ_{T_2}, which yields ϕ_2/Ψ_{T_2}.

Recasting Eq. (4.9) in terms of ϕ_2 and Ψ_{T_2}, the asymptotic populations resulting from the auxiliary walk are given by

$$P_2(\underline{R}, t) \equiv \int f_{\underline{R}}(\underline{R}', t)d\underline{R}' = \frac{\phi_2(\underline{R})}{\Psi_{T_2}(\underline{R})}\exp[-t(E_2-E_{R_2})]s_2 \ . \tag{5.8}$$

Thus, Eqs. (5.7) and (5.8) imply

$$O = \exp[t(E_2-E_{R_2})]s_1 s_2^{-1}\Psi_{T_1}^{-1}\Psi_{T_2}P_2 A = s_1 A\,\phi_2/\Psi_{T_1} \tag{5.9}$$

yields the transition dipole moment in Eq. (5.6).

We point out that for Eqs. (5.6)–(5.9) to be consistent, Ψ_{T_1} and Ψ_{T_2} must be normalized to unity when computing the ratio Ψ_{T_2}/Ψ_{T_1}. If correlation functions are included in the trial functions, the normalization of the trial function can not be obtained analytically. Since the correlation functions we employ are desirable, a method of normalizing trial functions is of interest.

When the trial function is given by $\mathrm{Det}^{\alpha}\mathrm{Det}^{\beta}F$, where the MOs in Det^{α} and Det^{β} are taken from SCF calculations and F is the correlation function, the trial function can be normalized as follows. Defining $\Psi_{\mathrm{SCF}} \equiv \mathrm{Det}^{\alpha}\mathrm{Det}^{\beta}$, we compute

$$K \equiv \int |\Psi_{SCF}|^2 |\Psi_T/\Psi_{SCF}|^2 d\underline{R} / \int |\Psi_{SCF}|^2 d\underline{R}$$

$$= \langle\Psi_T|\Psi_T\rangle / \langle\Psi_{SCF}|\Psi_{SCF}\rangle \ , \tag{5.10}$$

by employing a VMC walk guided by Ψ_{SCF}. Note that the ratio of Ψ_T to Ψ_{SCF} is simply F. Since the SCF MO's are orthonormal,

$$\langle\Psi_{SCF}|\Psi_{SCF}\rangle = n_\alpha! n_\beta! \ , \tag{5.11}$$

where n_α and n_β are the number of spin-up and spin-down electrons, respectively. (The normalization of Ψ_{SCF} is not $(n_\alpha + n_\beta)!$ because the Slater determinant is reduced to $\mathrm{Det}^\alpha\ \mathrm{Det}^\beta$.) Therefore, if Ψ_T is not normalized, a normalized trial function $\tilde{\Psi}_T$ may be obtained from

$$\tilde{\Psi}_T = [K n_\alpha! n_\beta!]^{\frac{1}{2}} \Psi_T \ . \tag{5.12}$$

In addition to values of the trial functions and their normalizations, the time dependence and overlap integrals in Eq. (5.9) must be computed. The time dependence is readily obtained by monitoring the change in the asymptotic populations at large t. That is,

$$P(t) \equiv \frac{\int \Psi_{T_1}(\underline{R})\phi_1(\underline{R})P_2(\underline{R}, t)d\underline{R}}{\int \Psi_{T_1}(\underline{R})\phi_1(\underline{R})d\underline{R}} = \exp[-t(E_2 - E_{R_2})]s_2 I / s_1 \ , \tag{5.13}$$

where

$$I \equiv \int \Psi_{T_1}(\underline{R})\phi_1(\underline{R})(\phi_2(\underline{R})/\Psi_{T_2}(\underline{R}))d\underline{R} \ . \tag{5.14}$$

Since the nodes of ϕ_i are those of Ψ_{T_i}, the integrand in Eq. (5.14) is non-negative so that I is non-zero and ϕ_2/Ψ_{T_2} is bounded at the nodes of Ψ_{T_2}. Unless Ψ_{T_2} is chosen so poorly that $\Psi_{T_2} < \Psi_{T_1}\phi_1\phi_2$ at large $|\underline{R}|$, I is bounded from above as well. Thus, the time dependence may be computed from

$$\exp[-t(E_2-E_{R_2})] = [P(t+\Delta t)/P(t)]^{t/\Delta t} \ . \tag{5.15}$$

The overlap integrals of normalized trial functions differ from unity by only second order in the error of the trial function, δ. For $\Psi_{T_i} = \phi_i + \delta_i$, $s_i = 1 + <\delta_i|\phi_i>$, and from

$$<\Psi_{T_i}|\Psi_{T_i}> = 1 = 1 + <\delta_i|\phi_i> + <\phi_i|\delta_i> + <\delta_i|\delta_i> \ , \tag{5.16}$$

there results (when Ψ_{T_i}, and therefore ϕ_i, is real)

$$|<\delta_i|\phi_i>| \leq \frac{1}{2}<\delta_i|\delta_i> \ . \tag{5.17}$$

Thus, the overlap integrals and the ratio s_1/s_2 are unity to order $<\delta_i|\delta_i>$. If necessary, these integrals may be obtained in separate calculations by sampling the asymptotic population for points initially distributed as $|\Psi_{T_i}|^2$. This may be seen by noting that [cf. Eq. (5.8)]

$$\begin{aligned} <P_i>_{|\Psi_{T_i}|^2} &\equiv <\Psi_{T_i}|P_i|\Psi_{T_i}> \\ &= \exp[-t(E_i-E_{R_i})]s_i^{\,2} , \end{aligned} \tag{5.18}$$

and so

$$s_i = \left[<P_i>_{|\Psi_{T_i}|^2}\exp[t(E_i-E_{R_i})]\right]^{\frac{1}{2}} . \tag{5.19}$$

Once again the time dependent factor may be computed by Eq. (5.15).

To summarize, $\Psi_{T_1}\phi_1$ and P_2 are sampled using two independent QMC walks, while the time dependence and overlap integrals are obtained by Eqs. (5.15) and (5.19), respectively. Since the overlap integrals and time dependence factor out of integrals involving O, their computed values are multiplied into averages of $\Psi_{T_1}^{-1}\Psi_{T_2}P_2A$ at the end of the calculation.

The methodology described here, in the context of diffusion QMC, is also applicable with exact QMC approaches, i.e., Green's function Monte Carlo (GFMC). (For a description of GFMC, see Refs. 35-36 and 41-42.)

To evaluate the efficiency of the present method, the $1s \rightarrow 2p_x$ transition dipole moment of the hydrogen atom is computed using the trial functions described in Section 5.2. State "1" in these equations corresponds to the ground state while state "2" is the excited state. Since the ground state is without nodes, and the $2p_x$ state has a node which may be specified exactly as $x = 0$, no fixed-node error results for the trial functions employed. The time steps used are 0.050 h^{-1} for the ground state walk and 0.025 h^{-1} for the secondary, excited-state, walk. Biases in the energy and mixed expectation values of r and r^2, resulting from using time steps of this magnitude in standard QMC calculations, are listed in Table 5.2. Note that the bias in $<r>$ and $<r^2>$ is an order of magnitude greater than that in the energies. However, since all biases are very small (0.3% or less), the time steps above are deemed suitable for the transition dipole moment calculations.

Results for the transition dipole moment and excited state energy for several different convergence times are presented in Table 5.3. Each energy is computed by averaging the local energy over the points in the secondary QMC walk at the time in question. In this approach the transition dipole moment appears to converge at about two standard deviations above the exact value – an error of 5%. The inaccuracy of the transition dipole moment may result from not having fully attained the $2p_x$ distribution. Poor convergence is indicated by the QMC energies for the $2p_x$ state which have stabilized significantly below the correct value.

The convergence behavior of the excited-state energies requires explanation. The last five energies in Table 5.3 are all very well described by an exponential which agrees nicely with QMC theory. The asymptotic value of ~ $-0.1272h$ obtained from this fit, however, is in significant disagreement with the QMC energy in Table 5.2 computed with the same time step and trial function. The explanation lies in the fact that $|\Psi_{T_2}|^2$ describes the initial distribution for the calculation of the excited-state energy of Table 5.2 whereas $\Psi_{T_1}\phi_1$ (i.e., a ground-state distribution) is the initial distribution leading to the results in Table 5.3. Though, the initial distribution does not affect the fixed-node energy, it can certainly modify the convergence to the final distribution. Specifically, the Metropolis acceptance/rejection step, in our algorithm, which maintains detailed balance, causes the probability of moving away from a point sufficiently near a node to be exceedingly small for time steps of the sizes generally used. Since the distribution $\Psi_{T_1}\phi_1$ need not be small at a node of Ψ_{T_2}, sampling points from $\Psi_{T_1}\phi_1$ as the initial distribution in a walk to obtain $\Psi_{T_2}\phi_2$ can result in several walkers being ''trapped'' near a node of Ψ_{T_2} for exponentially long periods of time.

The significance of this effect was verified by performing separate QMC calculations of the excited-state energy using Ψ_{T_2} and with an initial ensemble of points distributed according to $\Psi_{T_1}\phi_1$. With the time step employed above, 0.025 h^{-1}, the energy converged to and remained at -0.12827(73) h – even for times up to t = 75,000 h^{-1}. Since the probability of moving away from a node increases rapidly as the time step decreases, a run using $\tau = 0.01\ h^{-1}$ quickly yielded a much improved result of -0.125302(50) h, yet this result is still biased by trapped points. After a time of

$t = 18{,}000\ h^{-1}$, the remaining trapped points escaped, as indicated by a sudden increase in the energy to the statistically correct value of -0.125023(23)h.

The trapping effect described above may be avoided by omitting the acceptance/rejection step for points which do not move. However, this step, to the extent that the trial function is exact, eliminates time-step bias that would otherwise be present, cf. Fig. 3.1, and is therefore useful. Given the relatively small number of points for which the acceptance/rejection step would not be employed, the omission of this step for trapped points should, hopefully, not yield a noticeable increase in time-step bias.

5.4. Exact Transition Dipole Moments: VMC Walk with QMC Side Walks (Method 2)

As seen in the previous section, convergence from a distribution describing one state to that describing another can be problematic due to points trapped near a node. Furthermore, if Ψ_{T_1} also has nodes, the computed quantities may depend on the number of walkers in each nodal volume of this trial function. For example, let Ψ_{T_1} and Ψ_{T_2} have the nodes of the $2s$ and $2p_x$ states of the H atom, respectively. Since Ψ_{T_1} has a spherical node at $r = 2$, points sampled from f_1 will be confined in either of two nodal volumes, $r < 2$ or $r > 2$, during the first QMC walk. The average of $O = s_1 x \phi_2/\Psi_{T_2}$ is different in the inner nodal volume than in the outer, and will depend on the number of points in each volume. In practice the relative number of points in each nodal volume which is correct may be difficult to determine. To avoid this ambiguity and obtain better convergence to each state, another method is explored.

(Note, however, that when Ψ_{T_1} only possesses exchange nodes, the transition dipole moment is independent of the nodal volume sampled since $PO = O$.)

In Method 2, shown schematically in Fig. 5.2, quantities are sampled from a VMC walk in which walkers are not confined within nodal boundaries. For a VMC walk guided by Ψ_g, we now seek an operator O such that

$$<O>_{|\Psi_g|^2} \equiv <\Psi_g|O|\Psi_g> = <\phi_1|A|\phi_2> , \tag{5.20}$$

where again we are primarily interested in $A = x, y$, or z. Eq. (5.20) implies that the desired form of O is

$$0 = \Psi_g^{-1}\phi_1 A\,\phi_2\Psi_g^{-1} . \tag{5.21}$$

In manner analogous to Method 1, QMC solutions are obtained from asymptotic populations, P. Here, however, initial points are sampled from $|\Psi_g|^2$ and *two* QMC "side walks", one guided by Ψ_{T_1} and the other by Ψ_{T_2}, to sample P_1 and P_2, and thereby ϕ_1 and ϕ_2. These walks, and therefore this method, are generalizable to other QMC approaches.

The form of P_2 is given by Eq. (5.8) and that of P_1 is analogous. Therefore, the desired operator O for this method becomes

$$O \equiv (s_1 s_2)^{-1} T_1 T_2 (\Psi_{T_1}\Psi_{T_2}/|\Psi_g|^2) P_1 P_2 A , \tag{5.22}$$

where the time dependence, $\exp[t_i(E_i - E_{R_i})]$, is abbreviated as T_i. Also, t_1 and t_2 are the times required to converge to the asymptotic populations for the two states in question.

In deriving the operator O which satisfies Eq. (5.20), both trial functions and Ψ_g are assumed to be normalized. If this is not the case, then averages of O over $|\Psi_g|^2$ must be divided by $(\eta_1\eta_2)^{\frac{1}{2}}$, where $\eta_i \equiv <\Psi_{T_i}|\Psi_{T_i}>/<\Psi_g|\Psi_g>$. In contrast to Method 1,

normalizations do not need to be computed separately since η_i may be obtained from averages over $|\Psi_g|^2$ as

$$\eta_i = \int |\Psi_g|^2 |\Psi_{T_i}/\Psi_g|^2 d\underline{R} / \int |\Psi_g|^2 d\underline{R} \ . \tag{5.23}$$

The remaining quantities necessary for computing the transition dipole moment are the time dependences and the overlap integrals. The time dependences may be obtained as described in Sec. 5.3, cf. Eq. (5.15). The overlap integrals may be computed from the asymptotic populations once the time dependence is known. That is, for

$$\gamma_i \equiv \int |\Psi_g|^2 P_i |\Psi_{T_i}/\Psi_g|^2 d\underline{R} / \int |\Psi_g|^2 d\underline{R} \ , \tag{5.24}$$

there results, (the trial functions are not assumed to be normalized)

$$\gamma_i = \langle \Psi_{T_i} | \Psi_{T_i} \rangle \langle \Psi_g | \Psi_g \rangle^{-1} s_i^2 T_i^{-1} \ , \tag{5.25}$$

or

$$s_i = (T_i \gamma_i / \eta_i)^{\frac{1}{2}} \ . \tag{5.26}$$

As seen from Eqs. (5.23)–(5.26), all quantities necessary for the computation of the transition dipole moment may be obtained from the VMC and QMC walks. Since the overlap integrals, time dependences, and normalizations factor out of the integral involving O, computed values of these quantities are employed as weighting factors, according to Eq. (5.22), at the end of the calculation.

An additional benefit of Method 2 is that the asymptotic populations may be used to compute trial, mixed, and pure expectation values of any coordinate operator for each state. In particular,

$$\frac{<|\Psi_{T_i}/\Psi_g|^2 Q P_i^n>_{|\Psi_g|^2}}{<|\Psi_{T_i}/\Psi_g|^2 P_i^n>_{|\Psi_g|^2}} = <\Psi_{T_i}|Q|\Psi_{T_i}>/<\Psi_{T_i}|\Psi_{T_i}> \quad n = 0 \text{ (trial)} , \tag{5.27a}$$

$$= <\Psi_{T_i}|Q|\phi_i>/<\Psi_{T_i}|\phi_i> \quad n = 1 \text{ (mixed)} , \tag{5.27b}$$

$$= <\phi_i|Q|\phi_i> \quad n = 2 \text{ (pure)} , \tag{5.27c}$$

For example, to compute a quadrupole moment, $Q = \frac{1}{2}\sum_i (3z_i^2 - r_i^2)$.

As discussed in Chap. 4, two independent samplings of the asymptotic population must be performed to obtain $|\phi_i/\Psi_{T_i}|^2$. Therefore, to compute pure expectation values, each of the QMC side walks must be performed twice. Since pure expectation values possess a greater dependence on the degree of convergence than do mixed results, the behavior of pure values is most useful in ascertaining convergence to asymptotic populations, which may be critical for accurately computing transition dipole moments.

In addition to yielding pure expectation values, the statistical error of P_i is reduced by $\sqrt{2}$ when it is sampled twice. Also, the time dependences do not need to be computed separately. Note that, employing two independent samplings of each asymptotic population and computing

$$\kappa_i \equiv \int |\Psi_g|^2 P_i P_i^{\dagger} |\Psi_{T_i}/\Psi_g|^2 d\underline{R} / \int |\Psi_g|^2 d\underline{R} , \tag{5.28}$$

there results

$$\kappa_i = T_i^{-2} s_i^2 <\Psi_{T_i}|\Psi_{T_i}>/<\Psi_g|\Psi_g> . \tag{5.29}$$

Thus, the time dependences, overlap integrals, and normalizations are now given directly by

$$(s_1 s_2)^{-1} T_1 T_2 \langle \Psi_g | \Psi_g \rangle [\langle \Psi_{T_1} | \Psi_{T_1} \rangle \langle \Psi_{T_2} | \Psi_{T_2} \rangle]^{-\frac{1}{2}} = (\kappa_1 \kappa_2)^{-\frac{1}{2}} . \tag{5.30}$$

When separate computations of the time dependences are not required, continuation of the QMC walks after convergence to asymptotic populations is no longer necessary. Therefore, for the reasons discussed above, the additional computation incurred by two samplings of P_i may be justified.

Finally, the first and second-order approximations discussed in Sec. 5.2 may also be obtained, namely

$$\langle \Psi_{T_1} | A | \Psi_{T_2} \rangle = \langle \Psi_{T_1} \Psi_g^{-1} A \Psi_g^{-1} \Psi_{T_2} \rangle_{|\Psi_g|^2}, \qquad A_{tt} \tag{5.31a}$$

$$\langle \phi_1 | A | \Psi_{T_2} \rangle = s_1^{-1} T_1 \langle \Psi_{T_1} \Psi_g^{-1} P_1 A \Psi_g^{-1} \Psi_{T_2} \rangle_{|\Psi_g|^2}, \qquad A_{1t} \tag{5.31b}$$

$$\langle \Psi_{T_1} | A | \phi_2 \rangle = s_2^{-1} T_2 \langle \Psi_{T_1} \Psi_g^{-1} P_2 A \Psi_g^{-1} \Psi_{T_2} \rangle_{|\Psi_g|^2}, \qquad A_{2t} \tag{5.31c}$$

and

$$A_s = A_{1t} + A_{2t} - A_{tt} . \tag{5.31d}$$

[If functions are not normalized, the approximate moments may be obtained in terms of normalized functions by dividing by $(\eta_1 \eta_2)^{\frac{1}{2}}$.]

Having discussed the means by which the transition dipole moment is computed, we now consider the choice of Ψ_g. The goal in choosing Ψ_g is to be able to converge to the asymptotic populations of both states without t_1 and t_2 being excessively large. The rate of convergence for a given state depends, in part, on how well it is described by the guiding function Ψ_g. Therefore, each state should be represented to some extent in Ψ_g. Since the weighting factor $\Psi_{T_1} \Psi_{T_2} / |\Psi_g|^2$ is incorporated into averages, for purposes of reducing the statistical error, it is desirable that this quantity be well behaved where Ψ_g is small, i.e., when (or if) Ψ_g goes to zero, both Ψ_{T_1} and Ψ_{T_2} should

as well. A convenient form of Ψ_g which satisfies both these conditions is given by

$$\Psi_g \equiv \left[\sum_{i=1}^{2} |c_i \Psi_{T_i}|^2 \right]^{\frac{1}{2}} . \tag{5.32}$$

The ratio, $(c_1/c_2)^2$, affects the rate of convergence for each state. As the coefficient for the trial function describing a given state is increased, the time needed to obtain its asymptotic population is correspondingly decreased. It is thus useful to increase the coefficient for a state showing a slower convergence rate.

Results for the QMC energy, transition dipole moment between the $1s$ and $2p_x$ states, and other expectation values for the hydrogen atom, computed by Method 2, are presented in Tables 5.4 and 5.5. Calculations were undertaken for several time steps and convergence times to investigate convergence and time step bias. Below, removal of bias is measured by the accuracy in the computed energy and mixed expectation values, $<r>_m$.

Examining the results for the energy and $<r>_m$ in Table 5.4 immediately reveals faster convergence for the ground state than for the excited state. Decreasing c_1 thus still allows accurate ground-state energies to be obtained for convergence times of only about 30 h^{-1}. Also note that time step bias is not apparent: varying the time step for similar convergence times and the same c_1 has little effect on the ground-state results. However, the dependence of the excited-state quantities on the convergence time and time step is markedly different. The first four energies in Table 5.4 reveal that, both at $t = 20$ and 60 h^{-1}, the energies at $\tau = 0.025$ and 0.050 h^{-1} are noticeably different. More importantly, convergence to the exact energy is extremely slow – a noticeable error persists at $t = 60\ h^{-1}$. The trapping of walkers near the node of the excited state

appears to once again be affecting the energy. However, the effect is not as large in comparison to Method 1; at $\tau = 0.025\ h^{-1}$ and at $t = 60\ h^{-1}$, for example, the $2p_x$ energies from Methods 1 and 2 are -0.12745(7) and -0.12645(15) h, respectively. On the other hand, the trapping of points near $x = 0$ is not significant for $<r>_m$ as the accuracy of this quantity depends only on the convergence time.

Obviously, convergence to the $2p_x$ state distribution is difficult when Ψ_g has equal mixtures of the ground and excited state trial functions ($c_1 = c_2$). The last three energies for each state demonstrate the advantage of increasing the excited state representation relative to the ground state. Accurate ground-state energies are still obtained, but now the the $2p_x$ QMC energies are much better than the best $c_1 = c_2$ energy (0.4% error versus 1.2%). Even at $t = 15\ h^{-1}$, the $2p_x$ energy (for $c_1 = 0.16$, $c_2 = 0.84$) has roughly half the $c_1 = c_2$ error at a convergence time of 60 h^{-1} and the same time step. Therefore, judicious choices of the coefficients yield efficient convergence for both states.

Values for the transition dipole moments are shown in the last column of Table 5.4. As expected for a quantity which depends on two different eigenstates, there is overall less accuracy in the transition dipole moment when good convergence is not obtained for both states. The first four computations in Table 5.4 yield the least accurate energies for each state and transition dipole moments of inferior quality in comparison to the final two computations. Increasing the representation of the excited state in Ψ_g and the convergence time for the ground state improves the energies of both states. However, for first of these results, the fifth entry, the relatively short convergence time of 15 h^{-1} for the excited state yields the poorest value of $<r>_m$. As indicated by the

last two calculations in Table 5.4, further increasing the excited-state convergence time to 30 h^{-1} produced little change in the (accurate) energy, but significantly improved the value of $<r>_m$. Correspondingly, the most accurate transition dipole moments are obtained – 0.5% error.

To summarize, the highest quality transition dipole moment is obtained when each state is described most accurately, as measured by the energy *and* the mixed expectation value of r. (Other expectation values should be useful as well.) This conclusion is not surprising since the energy is only one, and not necessarily the best, way to judge the accuracy of the sampling of a given distribution. Therefore, it is best to study the behavior, as function of the convergence time and time step employed in each QMC walk, of expectation values in addition to that of the energy.

In Table 5.5 we report pure expectation values obtained for r and r^2 for both states using the present method. Since only a single value of P_1 and P_2 is sampled, the pure expectation values are computed employing the square of the asymptotic population rather than the product of two independent estimates of the quantity. Therefore, an unknown error is introduced in sampling $|\phi_i/\Psi_{T_i}|^2$. The accuracy of the results in Table 5.5 implies that this error is quite small.

For the ground state, all the values of $<r>$ are statistically equivalent as are all those for $<r^2>$. This indicates that these expectation values are fairly insensitive to the time step, and that sufficient convergence times have been achieved. Note also that the accuracies are quite good, in each case showing only a 0.3% error. The results for $<r>$ and $<r^2>$ of the $2p_x$ state also appear to be quite insensitive to the time step, but, in contrast to the $1s$ state values, show a marked convergence-time dependence. Note

that as t is increased from 15 to 60 h^{-1}, the errors in $<r>$ and $<r^2>$ are reduced from 3.2% to 0.6% and 7.1% to 0.7%, respectively. Comparing the $c_1 = c_2$ excited-state expectation values with the ($c_1 = 0.16$, $c_2 = 0.84$) results reveals that, similar to the case of $<r>_m$ but unlike that of the energy, increasing the representation of the excited state in Ψ_g is not necessary to obtain accuracy in pure expectation values of r and r^2 for the $2p_x$ state. Finally, we point out that the error in $<\Psi_{T_1}|x|\Psi_{T_2}>$ is 11% and that the errors in the trial expectation values of r and r^2 are, for each state, 11% and 23%, respectively. Therefore, reducing these errors to under 1% using QMC represents a significant improvement.

5.5. Exact Transition Dipole Moments: Green's Function Approach (Method 3)

In the previous two methods, values of a QMC solution, ϕ_i, were sampled using an auxiliary walk guided by a trial function, Ψ_{T_i}. These walks were propagated in imaginary time until the number of descendants of the point starting the secondary walk converged to a value proportional to ϕ_i/Ψ_{T_i}. When convergence times are large, the computational cost is prohibitive. Therefore, a method which does not depend on asymptotic populations to compute transition dipole moments might be significantly more efficient. Such a method, discussed in this section, is based on an idea due to Kalos[32] for sampling a single eigenstate by averaging an analytic expression for the Green's function over a suitable distribution. Here this approach is implemented so that the eigenstate required for computing the transition dipole moment is sampled by averaging a short-time approximation to Green's function over the appropriate QMC

distribution. If low variance averages of the Green's function are readily obtained, this method will present a viable alternative to the other two.

We begin the theory with the time development of a distribution, f_i,

$$f_i(\underline{R}', t+\tau) = \int G(\underline{R} \rightarrow \underline{R}', \tau) f_i(\underline{R}, t) d\underline{R} , \tag{5.33}$$

which implies for the Green's function average

$$\begin{aligned} \overline{G}^{(i)}(\underline{R}', \tau, t) &\equiv \int G(\underline{R} \rightarrow \underline{R}', \tau) f_i(\underline{R}, t) d\underline{R} / \int f_i(\underline{R}, t) d\underline{R} \\ &= f_i(\underline{R}', t+\tau) / \int f_i(\underline{R}, t) d\underline{R}. \end{aligned} \tag{5.34}$$

At large imaginary time, i.e., $f_i \propto \Psi_{T_i}\phi_i$, the dependence on t cancels in the numerator and denominator of Eq. (5.34) leaving

$$\overline{G}^{(i)}(\underline{R}', \tau) = \exp[-\tau_i(E_i - E_{R_i})] \Psi_{T_i}(\underline{R}')\phi_i(\underline{R}')/s_i . \tag{5.35}$$

Thus the Green's function average, like the distribution itself, is proportional to $\Psi_{T_i}\phi_i$.

Here the short-time Green's function, G_a, cf. Eq. (2.26), is employed. Note that, G_a yields time-step bias in both the distribution it generates in the QMC walk, f_i, and the numerical values of $\Psi_{T_i}\phi_i$ computed by Eq. (5.34). Therefore, as in Methods 1 and 2, time-step bias must be removed.

In the present Green's function approach, transition dipole moments are computed by sampling from two QMC distributions, in analogy to Method 1. (A Green's function approach analogous to Method 2 was studied and found to be less satisfactory.) Here, the first state is obtained from Green's function averages, and the second from $\Psi_{T_2}\phi_2$, as depicted in Fig. 5.3. We now seek an operator O involving $\overline{G}_a^{(1)}$ such that

$$\langle\Psi_{T_2}|O|\phi_2\rangle / \langle\Psi_{T_2}|\phi_2\rangle = \langle\phi_1|A|\phi_2\rangle . \tag{5.36}$$

By inspection of Eqs. (5.35) and (5.36) O is found to be

$$O = s_1 s_2 \exp[\tau_1(E_1 - E_{R_1})] \Psi_{T_1}^{-1} A \Psi_{T_2}^{-1} \overline{G}_a^{(1)} , \tag{5.37}$$

where Ψ_{T_1} and Ψ_{T_2} are assumed to be normalized. Likewise a quantity analogous to O in Eq. (5.37) with $\overline{G}_a^{(2)}$ may be used to obtain an independent estimate of the transition moment.

Since $\tau_1 |E_1 - E_{R_1}|$, a product of two small quantities, is generally less than 0.001, the exponential factor can be neglected when one is interested in only 0.1% accuracy. The overlap integrals may be computed from the same distributions used to evaluate $<O>$. In particular,

$$s_i = [<|\Psi_{T_i}|^{-2} \overline{G}_a^{(i)}>_{\Psi_{T_i}\phi_i}]^{-\frac{1}{2}} , \tag{5.38}$$

where we again neglect the factor of $\exp[\tau_i(E_i - E_{R_i})] \approx 1$. In addition to the transition moment, the pure expectation values can be computed also since

$$Q_p = <|\Psi_{T_i}|^{-2} \overline{G}_a^{(i)} A>_{f_i} / <|\Psi_{T_i}|^{-2} \overline{G}_a^{(i)}>_{f_i} . \tag{5.39}$$

For this method to be viable, Green's function averages should be precise for distributions of reasonable size. In addition, the time step bias in $\overline{G}_a$ must not be excessively large. Because G_a approaches a delta function as τ goes to zero, the variance in $\overline{G}_a$ diverges thus precluding the use of very small time steps. However, since the average of G_a evolves the distribution $\Psi_T \phi$ for only time τ, if $\Psi_T \phi$ possesses little time step bias then one expects the bias in $\overline{G}_a$ to also be small.

In our computation of the $1s \rightarrow 2p_x$ transition dipole moment for the H atom using the present method, we simply obtained 50,000 points distributed as $\Psi_T \phi$ for each state by diffusion QMC. The trial functions used to guide these walks are the same as those described in Sec. 5.2. The ground- and excited-state time steps were 0.050 and

0.025 h^{-1}, respectively. To ascertain the quality of each distribution, we compare averages of various quantities obtained from them with exact analytic averages over $\Psi_T\phi$. Table 5.6 shows this comparison and demonstrates that the distributions appear to be quite good, i.e., errors are less than 0.2% for the ground state and 1.0% for the excited state.

To test the quality of the averages, $\overline{G}_a$, overlap integrals and pure expectation values [cf. Eqs. (5.38) and (5.39)] are computed for each state. This is accomplished by evaluating Ψ_T and $\overline{G}_a$ for each point in the QMC distribution, $\Psi_T\phi$. (While single-state properties may be computed from Green's function averages and by Method 2, neither of these approaches appear to be as efficient in this regard to the algorithms described in Chap. 4.) When computing $\overline{G}_a(\underline{R}_i)$ for single-state properties, $\underline{R}_i$ is omitted from the average, i.e.

$$\overline{G}_a(\underline{R}_i) = (N-1)^{-1}\sum_{k\neq i}^{N} G_a(\underline{R}_k \rightarrow \underline{R}_i, \tau) \ . \tag{5.40}$$

This omission is employed because the two samplings of ϕ, one from $\Psi_T\phi$ and the other from $\overline{G}_a$, must be independent as shown in Chap. 4. In a sense, including $\underline{R}_i$ in the computation of $\overline{G}_a(\underline{R}_i)$ corresponds to an artificial "clustering" of the distribution about $\underline{R}_i$. This effect was verified in preliminary calculations of pure expectation values in which including the point $\underline{R}_i$ in $\overline{G}_a(\underline{R}_i)$ yielded poor results.

As the number of points, N, in Eq. (5.40) is increased, values of $\overline{G}_a$ become more precise. Therefore, to obtain precise results yet a rough estimate of statistical error, the 50,000 points were blocked into two groups of 25,000 in computing the overlap integral and pure expectation values of each state. (Given that only two values are

averaged, the statistical errors in the pure expectation values are somewhat uncertain, probably by a factor of two at most.) The results of these computations are presented in Table 5.7. The accuracy of computed quantities depends strongly on the accuracy of $\bar{G}_a$ which in turn depends on the magnitude of time-step bias. Therefore, this bias was studied for each state by computing at several values of τ.

For the ground state, biases are small but noticeable at $\tau = 0.20$ and $0.10\ h^{-1}$. By $\tau = 0.05\ h^{-1}$ it appears that the values of $\bar{G}_a$ are accurate. This is implied by the better than 0.5% agreement of the expectation values obtained from Green's function averages with those obtained by substituting analytic values of $\Psi_T\phi/s$ for $\bar{G}_a$. Also, these results are within 0.5% of the exact pure expectation values. Thus, for 50,000 points sampled from $\Psi_T\phi$, where Ψ_T is only of modest accuracy, the overlap integral and pure expectation values are of high quality for the ground state of H.

We now consider the behavior of the Green's function averages for the $2p_x$ state at $x = 0$, $\bar{G}_a(0)$ where "0" signifies $x = 0$. Since $\bar{G}_a(0)$ should be proportional to the $2p_x$ state distribution, $\Psi_{T_2}(0)\phi_2(0)$, $\bar{G}_a(0)$ should vanish. However, note that

$$\bar{G}_a(0) = \int G_a(\underline{R}_k \to 0, \tau)\Psi_{T_2}(\underline{R}_k)\phi_2(\underline{R}_k)d\underline{R}_k / \int \Psi_{T_2}(\underline{R}_k)\phi_2(\underline{R}_k)d\underline{R}_k \tag{5.41}$$

will not vanish for non-zero τ because G_a and $\Psi_{T_2}\phi_2$ are everywhere positive and only zero at $x = 0$. The node only appears for $\tau = 0$ when G_a is a delta function. Therefore, we have investigated a modification of $\bar{G}_a$ such that it vanishes at $x = 0$ for non-zero τ as follows. Since $G_a(\underline{R}_k \to 0, \tau)$ is symmetric about $x_k = 0$ [employ $\Psi_T(\underline{R}_k) = x_k \exp(-\gamma r_k)$ in $G_a(\underline{R}_k \to 0, \tau)$ in Eq. (2.26)] then averages of $G_a(\underline{R}_k \to \underline{R}_i, \tau)$ over $\Psi_{T_2}(\underline{R}_k)\phi_2(\underline{R}_k)$, which is also symmetric about $x_k = 0$, will become zero as $x_i \to 0$ when weighted by $\mathrm{sign}(x_i x_k)$. Finally, since G_a becomes a

delta function as $\tau \rightarrow 0$, weighted and unweighted averages are identical because only $\underline{R}_k = \underline{R}_i$ [where $\mathrm{sign}(x_k x_i) = 1$] contributes to these averages at $\tau = 0$.

Computations (not shown) were performed with and without this weighting on the $2p_x$ state with $0.05 \leq \tau \leq 0.50\ h^{-1}$. The effect of this modification on values of the quantities reported in Table 5.7, always less than 0.2%, decreases with decreasing time step.

The behavior of the expectation values in Table 5.7 shows that time-step bias in these quantities is larger for the ground state. Pure expectation values for the excited state are in statistical agreement with each other, and very accurate (generally less than 1% error), for $0.10 \leq \tau\ 1.00\ h^{-1}$. In contrast time-step bias is obviously present in all the computed pure expectation values of the ground state at $\tau = 0.20\ h^{-1}$. Time steps which are large, in terms of the resulting bias, for the $1s$ state are small, by the same criterion, for the $2p_x$ state. This concept of different time scales is also reflected in the behavior of the statistical errors. As stated previously the Green's function becomes a delta function as $\tau \rightarrow 0$ so that the averages of G_a possess large variances (and poor accuracy) at small time steps. This effect manifests itself in low accuracies and large statistical errors in computed expectation values as the time step becomes "small". A time step of $0.05\ h^{-1}$ yields high accuracy and small statistical errors in results for the ground state. On the other hand, this time step gives the poorest agreement with the exact results, and diverging statistical errors, for the excited state.

Turning now to the transition dipole moment, we discuss the factors which affect the computed result. The primary consideration is that values of $\overline{G}_a$ give good agreement with

$$\overline{G}_a(\underline{R}) = \Psi_T(\underline{R})\phi(\underline{R})/s \ . \tag{5.42}$$

That is, both the explicit τ–bias, $\exp[\tau(E-E_R)]$, and the implicit τ–bias, arising from the short-time approximation, should be small. Furthermore, the number of points used to compute the Green's function average must be large enough to yield good statistical accuracy. This aspect is investigated by performing computations with $\Psi_T\phi/s$, the analytic form of $\overline{G}_a$ when it is exact, in place of $\overline{G}_a$. In other words, we compare averages of $\Psi_{T_1}^{-1}x\Psi_{T_2}^{-1}\overline{G}_a^{(i)}$ with averages of $\Psi_{T_1}^{-1}x\Psi_{T_2}^{-1}(\Psi_{T_i}\phi_i/s_i)$, the latter are referred to as "Analytic $\overline{G}$" in Table 5.8, in ascertaining the accuracy of Green's function averages. (The errors in the Analytic $\overline{G}$ results arise from statistical error due to the finite number of points sampled from $\Psi_{T_i}\phi_i$, and the time step used in the QMC walk yielding this distribution.)

Results for the transition dipole moment are presented in Table 5.8. Two separate sets of calculations have been carried out at several values of τ. The top set samples values of $\Psi_{T_1}\phi_1$ from $\overline{G}_a^{(1)}$, while the bottom set samples $\Psi_{T_2}\phi_2$ from $\overline{G}_a^{(2)}$. In both sets of calculations, $\overline{G}_a^{(i)}$ is obtained from all 50,000 points distributed as $\Psi_{T_i}\phi_i$. Averages of $\Psi_{T_1}^{-1}x\Psi_{T_2}^{-1}\overline{G}_a^{(i)}$ are grouped into ten blocks of 5,000 points each from which statistics are obtained. The first column of results ignores both the explicit τ–dependence and the overlap integrals. Therefore, these and the Analytic $\overline{G}$ (this column) results allow a direct analysis of the errors in $\overline{G}_a$ from time-step bias and limited ensemble size. In the second column of results, the explicit τ–dependence is removed by multiplying first-column results by $\exp[\tau_i(E_i-E_{R_i})]$ (this factor is unity for the Analytic $\overline{G}$ values). This dependence is taken from previous QMC computations on the $1s$ and $2p_x$ states of H, cf. 3.2, with the trial functions employed here. Finally, the transition dipole

moments are obtained by multiplying the second-column results by the overlap integrals s_1 and s_2. The values of s_1 and s_2 are those computed at $\tau = 0.05$ and $0.20\ h^{-1}$, respectively, the smallest values of τ before large increases in the observed statistical error of these quantities occur, cf. Table 5.7.

The results in Table 5.8 demonstrate that values computed with $\overline{G}_a^{(2)}$ possess less time-step bias than do those of $\overline{G}_a^{(1)}$, as also found in the computation of pure expectation values. However, the statistical error in the $\overline{G}_a^{(2)}$ results is seen to be roughly twice as large, making calculations with $\overline{G}_a^{(2)}$ less preferable since, as seen in Table 5.8, time-step bias is readily removed. The difference in statistical errors is caused by the greater inherent statistical uncertainty in sampling $\Psi_{T_2}\phi_2/s_2$, an analytic comparison of the statistical errors, σ, yields

$$\sigma[<\Psi_{T_1}^{-1}x\,\Psi_{T_2}^{-1}(\Psi_{T_2}\phi_2/s_2)>_{\Psi_{T_1}\phi_1}] > 3\sigma[<\Psi_{T_1}^{-1}x\,\Psi_{T_2}^{-1}(\Psi_{T_1}\phi_1/s_1)>_{\Psi_{T_2}\phi_2}]\ . \qquad (5.42)$$

Computing the transition dipole moment by sampling values of the $1s$ distribution (from $\overline{G}_a^{(1)}$) and averaging over the $2p_x$ distribution is most efficient in Method 3. Results for expectation values computed in this way, top part of Table 5.8, are now analyzed in detail. Comparing the second and third columns in Table 5.8, reveals that the explicit τ–dependence is insignificant when other non-zero–τ effects are also negligible, $\tau \leq 0.20\ h^{-1}$. We also see, from the first column, that Green's function averages are reliable for $\tau \leq 0.20\ h^{-1}$ as the discrepancy between $\overline{G}_a^{(1)}$ and Analytic $\overline{G}$ results is less than 0.9%. The last column in Table 8 gives the transition dipole. Though the time-step bias in the computed transition dipole moment is large at $\tau = 1\ h^{-1}$, 27% error, by $\tau = 0.20\ h^{-1}$ convergence to statistically constant and accurate values, for example only 0.63(62)% error at $\tau = 0.10\ h^{-1}$, is obtained. Thus, distributions of

reasonable size (50,000 points from $\Psi_{T_1}\phi_1$ and $\Psi_{T_2}\phi_2$) used in the "Green's function" approach yield accurate and precise single-state expectation values and transition dipole moments for the $1s$ to $2p_x$ transition in the hydrogen atom.

5.6. Evaluation of Each Method

The results presented here show that Methods 2 and 3 are clearly superior to Method 1. The 4.4% error in the transition dipole moment obtained from Method 1 is an order of magnitude worse than the ~ 0.4% error obtained by the other two. This disparity might be caused by the lack of convergence to the QMC distribution of the $2p_x$ state, as indicated by the energies in Table 5.3. In addition, the statistical error with Method 1 is relatively large. For example, to obtain a transition dipole moment of 0.778(17) using the Method 1 took four times the computation time as Method 2 took to yield a value of 0.7425(35). This poor efficiency results from the ratio Ψ_{T_2}/Ψ_{T_1} becoming unbounded as r becomes large [cf. Eq. (5.9)]. In general, such ratios of trial functions describing different states will possess singularities and cause large statistical errors in transition dipole moments computed by Method 1. For this reason, when using several walks, a VMC walk with QMC side walks is more efficient for the computation of transition dipole moments.

As we have seen, the latter two methods yield highly accurate single-state expectation values *and* transition dipole moments. In considering the relative efficiency of these two methods, the time-step biases and difficulty in correctly selecting other parameters (e.g., convergence time or distribution size) must be examined.

The time-step bias of the two methods is largely equivalent. While small time

steps were not required in computing Green's function averages, the time steps employed in obtaining $\Psi_{T_1}\phi_1$ and $\Psi_{T_2}\phi_2$ should be the same as those used in the QMC side walks of Method 2. With Method 2, an ensemble of points sampled from $|\Psi_g|^2$ is converged to the distributions $\Psi_{T_1}\phi_1$ and $\Psi_{T_2}\phi_2$ in sampling asymptotic populations. In Method 3, each QMC distribution is obtained by converging from $|\Psi_T|^2$. Since $|\Psi_{T_i}|^2$ yields a more accurate description of $\Psi_{T_i}\phi_i$ than does $|\Psi_g|^2$, convergence in Method 3 is about an order of magnitude faster than in Method 2. However, employing a Green's function average instead of asymptotic populations gives an additional summation in the Green's function approach. Therefore the relative efficiency between the two methods is strongly dependent on the number of points required to obtain high-quality estimates of $\Psi_T\phi$ from $\overline{G_a}$. Table 5.8 shows that an accurate transition dipole moment is obtained employing moderately sized distributions of 50,000 points for each state. In addition, the selection of c_1, c_2, and convergence times in Method 2 is not trivial. This may be appreciated by noting the number of computations and the varying quality of the transition dipole moment in Table 5.4. For these reasons the Green's function averaging method is the most efficient of the three studied here for the computation of the $1s \rightarrow 2p_x$ transition dipole moment of the hydrogen atom.

It is possible that larger systems may be more difficult to treat using Method 3 than Method 2. As the dimensionality of a problem increases, larger distributions are required for statistically accurate averages of an increasingly complex G_a. Such distributions will reduce efficiency versus Method 2 due to the increased computation time in evaluating $\overline{G_a}$. It is also notable that the time step bias can be eliminated more easily using Method 2, although smaller time steps yields longer QMC side walks (to

converge to asymptotic weights) making calculations costly at small τ, the limitation on the smallness of the time step appears to be more severe for Green's function averaging. This is due, as seen in calculations on the $2p_x$ state, to statistical errors rapidly increasing as the time step becomes small and the Green's function approaches a delta function.

In conclusion, both Methods 2 and 3 hold promise for computing transition dipole moments of larger atoms.

5.7 Calculation of the Li $2^2S \rightarrow 2^2P$ Oscillator Strength

As seen in the previous two sections, Methods 2 and 3 yield an accurate transition dipole moment for H. Here, we investigate the capability of Method 2 to compute the $2^2S \rightarrow 2^2P$ transition dipole moment from which is obtained the experimentally observable oscillator strength and excited-state lifetime, cf. Eqs. (5.2) and (5.3). (Computations on Li employing Method 3 have not yet been undertaken.) The lithium atom serves as an excellent test case because it is the simplest atom for which fixed-node error must be considered for both the ground and first-excited state. Furthermore, a near-exact theoretical computation of the oscillator strength has not been performed.

For the S and P states, the trial function consists of a Slater determinant multiplied by an E-E correlation function. As discussed in Chap. 3, the basis sets are of near-HF quality and are given in Ref. 77. The E-E correlation function parameters, VMC, QMC, and estimated exact energies are presented in Table 5.9. As seen from comparing the QMC and exact energies, fixed-node error in the QMC energy is not observable. Therefore, these trial functions make an excellent starting point for computing

the oscillator strength, f.

We now turn to the initial calculations of the oscillator strength. Since no time-step bias was visible in the P–state energy at $\tau = 0.01\ h^{-1}$, cf. Sec. 3.4, this time step was used in the QMC side walks of both states. Also, for all computations discussed in this section, moves were rejected upon crossing a node. Since carrying the weights when sampling asymptotic populations was advantageous in the single-walk calculations of pure expectation values, cf. Chap. 4, weights are carried in the QMC side walks employed here. The minimum for the weights is set at 0.1 and the maximum at 2.0. In addition, asymptotic populations are only sampled once. Thus, an error is present in the pure expectation values (which was quite small for H). Nonetheless, these quantities are expected to show the most marked dependence on convergence to asymptotic populations and are therefore considered useful.

Table 5.10 presents results for the single-state properties and the oscillator strength. The energies are computed by averaging the local energy over points sampled in the QMC walk at times greater than $t/2$. The coefficients, c_1^2 and c_2^2, are defined by Eq. (5.32) and "1" denotes the ground state and "2" denotes the excited state. Several convergence times are employed and the trial functions are weighted differently at $t = 20\ h^{-1}$. The non-QMC results for r_p and r_p^2 (ground state) are taken from Ref. 11 in which a 352-term Hylleraas expansion is employed to yield an almost exact energy of -7.47806 h (the exact is -7.47807 h). Therefore, the expectation values of r and r^2 obtained in Ref. 11 should be quite accurate.

While the pure expectation values of the ground state appear to have converged, and are in good agreement with the results of Ref. 11, the energies are too low, by

more than statistical error, for $t \geq 10\ h^{-1}$. As seen in Secs. 5.3 and 5.4, trapping effects can play a role in preventing the convergence to asymptotic populations. In addition, these effects may be visible in one quantity but not in another. (The overall effect of trapped points in biasing an average depends on the extremity of the values near a node of the quantity averaged.)

The excited-state energies, on the other hand, appear to converge by $t = 10\ h^{-1}$ and give statistical agreement with the exact (as did the QMC energy computed straightforwardly, cf. Table 5.9). However, the statistical errors are so large that the discrepancies found in the ground state energies are within the statistical errors of the excited-state energies. The pure expectation values have also reached a reasonable degree of convergence. However, no accurate estimates are available for comparison.

We now turn to the oscillator strengths, which are computed from the transition dipole moment and ΔE $[\equiv E(P) - E(S)]$ according to Eq. (5.2). Note that the three-fold degeneracy of the P state yields a factor of three in Eq. (5.2), and for all QMC (and VMC) oscillator strengths, $\Delta E = 0.06778(30)\ h$ is obtained from the QMC $S-$ and $P-$state energies (the exact difference is 0.06791 h). For all calculations, the second-order approximation is more accurate than the pure. In addition, all but the $t = 20\ h^{-1}$ calculation yielded poor values of the oscillator strength. The relatively high quality of the $t = 20\ h^{-1}$ oscillator strength may be due to increasing the excited-state representation in Ψ_g since this was found to be beneficial for the H calculations. However, the reasonable quality of the $t = 20\ h^{-1}$ oscillator strength versus the $t = 30\ h^{-1}$ value is somewhat puzzling given the agreement of the single-state expectation values. Of course, such an occurrence is possible given the uniqueness of

the oscillator strength. In any event, our most accurate pure oscillator strength, f_p, is not of the quality desired as it only reproduces the previously most accurate theoretical value of $f = 0.748$[12], obtained from a Hylleraas expansion. The second-order f–value is somewhat better, within 0.4% of experiment.

While an accurate value of f is computed above, we wish to further improve the quality of the calculations. Specifically, convergence to asymptotic populations does not appear to have reached the point such that excellent agreement with experiment is observed. While this lack of agreement may be due to fixed-node error, it is of interest to facilitate convergence to asymptotic populations to insure that converged results are obtained. Furthermore, time-step bias has not been studied. While it is expected that time-step bias is not significant at $\tau = 0.01\ h^{-1}$, it is useful to know if bias is small at larger time steps since such values of τ would allow more efficient QMC side walks.

Lack of convergence may be due to the trapping of points selected from $|\Psi_g|^2$ which are close the node of Ψ_{T_1} or Ψ_{T_2} at the start of a QMC walk. For Li, explicit measurements of the number of accepted moves for each walker has shown that a small, but non-negligible, percentage of the points sampled from $|\Psi_g|^2$ do not move during the entire course of the QMC walk. The results in Secs. 5.3 and 5.4 also strongly indicate that adverse effects are caused by the trapping of points near a node. Here, we pursue an idea to greatly reduce trapping effects without removing the Metropolis acceptance/rejection step which is generally very desirable.

To reduce trapping effects, two time steps are employed during the course of the walk. The initial time step is very small and is used for only a brief period to allow

points to move away from nodes. The final time step is of the size typically used in QMC calculations for the system at hand, and the majority of the QMC walk is implemented with the larger value of τ. The results of such calculations are presented in Table 5.11.

For all the calculations presented in Table 5.11, the initial time step is 0.002 h^{-1}, i.e., about an order of magnitude smaller than those normally used for Li. The initial propagation time with the first time step is denoted by t_i, and t is the total convergence time. Also, $c_1 = c_2 = 0.5$.

In addition to the quantities reported in Table 5.10, the "growth" energy, E_g, is given as well. The growth energy estimates the QMC energy from the change in the ensemble size as the QMC walk proceeds. Recalling Eq. (2.24), the time dependence of the QMC distribution is given by

$$f(\underline{R}, t) = c_0 \exp[-t(E - E_R)]\Psi_T(\underline{R})\phi(\underline{R}) , \tag{5.43}$$

from which the population is obtained as

$$P(t) = \int f(\underline{R}, t) d\underline{R} = c_0 \exp[-t(E - E_R)] \langle \Psi_T | \phi \rangle . \tag{5.44}$$

Thus, E_g is computed from populations at t and $t + T$ as

$$E_g = E_R - T^{-1} \ln[P(t + T)/P(t)] . \tag{5.45}$$

While E_g is a valid estimator of the QMC energy, it most often possesses a significantly larger amount of time-step bias than the average of the local energy and is, therefore, generally not of interest. However, in the current context the sensitivity of E_g is a useful indicator of the remaining effects of time-step bias. Since Eqs. (5.44) and (5.45) require that convergence to $\Psi_T \phi$ is obtained, the growth energy is computed by measuring changes in the population 1–2 h^{-1} before the end of each

QMC walk.

The most significant results in Table 5.11, are those of the oscillator strength. We now see that employing two time steps, 0.002 h^{-1} and 0.010 h^{-1}, yields an oscillator strength in excellent agreement with experiment, 0.741(7) (second-order) and 0.742(7) (pure) versus 0.742(1). That convergence is enhanced by using two time steps is best seen by comparing the results above with those using the same final time step and a longer convergence time (12 h^{-1}), but without a small initial time step. This comparison shows that the oscillator strength and all single-state properties for $\tau = 0.01\ h^{-1}$ and $t = 12\ h^{-1}$ are of poorer quality than those for which a smaller time step is used in conjunction with $\tau = 0.01\ h^{-1}$ (except for the P–state QMC energies which are in close statistical agreement). It is also seen, by comparing the $\tau = 0.025$, $t = 10\ h^{-1}$ results with those at $\tau = 0.010$, $t = 10\ h^{-1}$, that the larger time step generally yields less accuracy, indicating the presence of time-step bias. These differences are most noticeable for the for the oscillator strength, growth energies, and the excited-state values of r_p and r_p^2.

In concluding this chapter, Table 5.12 compares QMC results with other theoretical values. In addition to oscillator strengths, excited-state lifetimes are also presented. The QMC value of the excited-state lifetime is obtained from the $\tau = 0.01$, $t = 10\ h^{-1}$ value of the oscillator strength which is combined with the QMC-computed energy according to Eq. (5.3). Since the other theoretical oscillator strenghts are computed using the exact energy difference, the corresponding excited-state lifetimes are as well. As seen in Table 5.12, the QMC result is the most accurate and is in excellent agreement with recent and precise experimental measurements. However, the accuracy

above was not easily obtained as care, i.e., a small initial time step, had to be taken to allow the walkers to propagate to the asymptotic populations. Furthermore, it is desirable to increase the efficiency of the approach since the 0.9% precision in our value required seven hours on a Cray/XMP. Improvements in efficiency will require more accurate trial functions which will yield asymptotic populations more quickly and reduce time-step bias. In addition, alternative forms of $|\Psi_g|^2$ should be explored.

Table 5.1. Approximations to the transition dipole moment. Four approximations to the transition dipole moment between the $1s$ ("1") and $2p_x$ ("2") states of the hydrogen atom are compared. The trial functions, Ψ_{T_i}, are given in the text and are approximations to the exact hydrogenic wave functions, ϕ_i, accurate to 1% for the energy expectation value. All functions are normalized. The approximations reported here correspond to A_{tt}, A_{1t}, A_{2t} and A_s $(A = x)$, respectively. The first three expressions are accurate to first order in the error in the trial functions, while the superscripted quantity is accurate to second order.

Quantity	Analytic Expectation Value	| % Error |
$<\Psi_{T_1}\|x\|\Psi_{T_2}>$	0.8277	11.1
$<\Psi_{T_1}\|x\|\phi_2>$	0.8981	20.6
$<\phi_1\|x\|\Psi_{T_2}>$	0.6782	9.0
$<\phi_1\|x\|\phi_2>^{(2)}$	0.7485	0.5
$<\phi_1\|x\|\phi_2>_{Exact}$	0.7449	0

Table 5.2. Time-step biases for the $1s$ and $2p_x$ states of the hydrogen atom. Atomic units are used throughout: energy is in hartrees, distance in bohrs, and time in inverse hartrees. Statistical errors, representing one standard deviation of a Gaussian distribution, are shown in parentheses. The time step for the ground state calculation is 0.050 h^{-1} and that for the excited state is 0.025 h^{-1}. All expectation values are computed with respect to the mixed distribution $\Psi_T\phi$ for the state in question.

	$1s$ state		
	$-E$	$\langle r \rangle$	$\langle r^2 \rangle$
QMC	0.49984(4)	1.5819(8)	3.3345(38)
Analytic	0.50000	1.5789	3.3241
% Bias	-0.032(9)	0.19(5)	0.31(11)
	$2p_x$ state		
QMC	0.124999(19)	5.2637(33)	33.251(41)
Analytic	0.125000	5.2632	33.241
% Bias	-0.001(15)	0.01(6)	0.03(12)

Table 5.3. Transition dipole moment of the hydrogen atom by Method 1. The trial functions describing the $1s$ and $2p_x$ states are given in the text. The values for zero convergence time are computed analytically.

Convergence time	$-E\ (2p_x)$	$<1s\ \lvert x\ \rvert 2p_x>$
0.0	0.19625	0.678
7.5	0.13601(7)	0.720(15)
15.0	0.13143(7)	0.731(15)
22.5	0.12959(7)	0.742(15)
30.0	0.12866(7)	0.756(16)
37.5	0.12812(7)	0.768(16)
45.0	0.12780(7)	0.776(17)
52.5	0.12759(7)	0.777(17)
60.0	0.12745(7)	0.778(17)
Exact	0.12500	0.745

Table 5.4. Energies and transition dipole moment for the $1s$ and $2p_x$ states of the hydrogen atom by Method 2. These quantities are computed for several values of the time step (τ), convergence time t, and trial function coefficient [c, see Eq. (5.32)].

$1s$ State			$2p_x$ State			$\langle 1s \lvert x \rvert 2p_x \rangle$
(τ, t, c^2)	$-E$	$\langle r \rangle_m$	(τ, t, c^2)	$-E$	$\langle r \rangle_m$	
(0.050,20,0.50)	0.49889(3)	1.5787(15)	(0.050,20,0.50)	0.12803(21)	5.325(26)	0.7267(20)
			(0.050,60,0.50)	0.12708(14)	5.285(12)	0.7498(20)
			(0.025,20,0.50)	0.12741(9)	5.325(12)	0.7366(20)
(0.025,20,0.50)	0.49888(6)	1.5876(39)	(0.025,60,0.50)	0.12645(15)	5.264(10)	0.7564(30)
(0.040,32,0.16)	0.49979(5)	1.5772(36)	(0.025,15,0.84)	0.12565(4)	5.356(8)	0.7292(23)
(0.040,40,0.16)	0.49989(4)	1.5822(51)	(0.025,30,0.84)	0.12551(8)	5.291(11)	0.7425(35)
(0.025,35,0.16)	0.49987(3)	1.5778(42)	(0.015,30,0.84)	0.12538(2)	5.294(7)	0.7410(18)
Exact	0.50000	1.5789	---	0.12500	5.263	0.7449

Table 5.5. Pure expectation values for the $1s$ and $2p_x$ states of the hydrogen atom by Method 2. The parameters τ, t and c are described in the text and take the same values as those in Table 5.4.

$1s$ State			$2p_x$ State		
(τ,t,c^2)	$\langle\phi\|r\|\phi\rangle$	$\langle\phi\|r^2\|\phi\rangle$	(τ,t,c^2)	$\langle\phi\|r\|\phi\rangle$	$\langle\phi\|r^2\|\phi\rangle$
(0.050,20,0.50)	1.5018(17)	3.008(7)	(0.050,20,0.50)	5.102(22)	31.42(32)
			(0.050,60,0.50)	5.028(11)	30.33(12)
			(0.025,20,0.50)	5.105(10)	31.40(15)
(0.025,20,0.50)	1.5005(23)	2.999(12)	(0.025,60,0.50)	4.969(19)	29.93(13)
(0.040,32,0.16)	1.5070(39)	3.010(13)	(0.025,15,0.84)	5.160(7)	32.13(11)
(0.040,40,0.16)	1.5064(65)	3.023(20)	(0.025,30,0.84)	5.045(11)	30.61(12)
(0.025,35,0.16)	1.5042(33)	3.008(11)	(0.015,30,0.84)	5.056(6)	30.70(7)
Exact	1.5000	3.000	---	5.000	30.00

Table 5.6. Expectation values over distributions used in Method 3. For each state, the distribution consists of 50,000 points sampled from $\Psi_T \phi$. Statistical errors are obtained by averaging ten values of $\langle\Psi_T|Q|\phi\rangle/\langle\Psi_T|\phi\rangle$, each of which is an average over 5000 points.

Q	$1s$ state		$2p_x$ state	
	From distribution	Exact	From distribution	Exact
x^2	1.1098(74)	1.1080	19.80(10)	19.94
r	1.5816(33)	1.5789	5.240(10)	5.2663
r^2	3.328(12)	3.324	32.91(11)	33.24
E_L	-0.49942(50)	-0.50000	-0.125036(67)	-0.125000

Table 5.7. Expectation values from Green's function averages. For these single-state properties, averages are obtained from two calculations of 25,000 points each. Here, and in Table 5.8, points in the ground- and excited-state distributions were obtained employing time steps of 0.050 and 0.025 h^{-1}, respectively.

		1s state		
τ	$\langle\Psi_T\|\phi\rangle$	$\langle x^2\rangle$	$\langle r\rangle$	$\langle r^2\rangle$
0.20	0.99699(19)	1.0172(53)	1.5249(30)	3.0528(11)
0.10	0.99625(33)	1.0034(48)	1.5113(33)	3.0247(34)
0.05	0.99599(22)	0.9975(90)	1.5066(34)	3.0093(29)
Analytic $\bar{G}$	0.99613(18)	1.0038(47)	1.5030(22)	3.0062(10)
Exact	0.99585	1.0000	1.5000	3.0000
		$2p_x$ state		
1.00	0.99492(35)	18.06(13)	4.997(10)	29.77(23)
0.75	0.99408(31)	17.98(14)	4.984(10)	29.63(24)
0.50	0.99322(29)	17.90(17)	4.973(11)	29.54(25)
0.20	0.99150(19)	17.80(20)	4.973(6)	29.60(17)
0.10	0.9893(24)	17.70(21)	4.993(17)	29.92(21)
0.05	0.9848(79)	17.50(35)	5.048(79)	30.79(120)
Analytic $\bar{G}$	0.99207(31)	17.93(7)	4.982(4)	29.74(11)
Exact	0.99309	18.00	5.000	30.00

Table 5.8. Transition dipole moment of H by method 3. The Green's function is averaged over 50,000 points, and the expectation value, $<\Psi_{T_1}^{-1} x \Psi_{T_2}^{-1} \overline{G}_a^{(i)}>_{\Psi_{T_j}\phi_j}$, is computed by averaging over ten distributions of 5000 points drawn from $\Psi_{T_j}\phi_j$. The labels "1" and "2" correspond to the $1s$ and $2p_x$ states, respectively.

τ	$<\Psi_{T_1}^{-1} x \Psi_{T_2}^{-2} \overline{G}_a^{(1)}>_{\Psi_{T_2}\phi_2}$	$T(\tau)^a <\Psi_{T_1}^{-1} x \Psi_{T_2}^{-2} \overline{G}_a^{(1)}>_{\Psi_{T_2}\phi_2}$	$<1s\|x\|2p_x>$
1.00	0.9370(52)	0.9566(53)	0.9446(52)
0.75	0.8441(45)	0.8535(45)	0.8428(44)
0.50	0.7858(41)	0.7900(41)	0.7801(40)
0.20	0.7527(41)	0.7539(41)	0.7445(40)
0.10	0.7494(47)	0.7496(47)	0.7402(46)
0.05	0.7496(62)	0.7497(62)	0.7403(61)
Analytic $\overline{G}$	0.7561(39)	0.7561(39)	0.7469(38)
Exact	0.7532	0.7532	0.7449

τ	$<\Psi_{T_1}^{-1} x \Psi_{T_2}^{-2} \overline{G}_a^{(2)}>_{\Psi_{T_1}\phi_1}$	$T(\tau)<\Psi_{T_1}^{-1} x \Psi_{T_2}^{-2} \overline{G}_a^{(2)}>_{\Psi_{T_1}\phi_1}$	$<1s\|x\|2p_x>$
1.00	0.7364(88)	0.7371(88)	0.7279(87)
0.75	0.7399(89)	0.7403(89)	0.7310(88)
0.50	0.7430(90)	0.7432(90)	0.7339(89)
0.20	0.7463(93)	0.7463(93)	0.7370(92)
0.10	0.748(10)	0.748(10)	0.739(10)
0.05	0.750(12)	0.750(12)	0.741(12)
Analytic $\overline{G}$	0.7426(86)	0.7426(86)	0.7336(85)
Exact	0.7532	0.7532	0.7449

[a] The factor $\exp[\tau(E - E_R)]$, cf. Eq. (5.37), where E is the growth energy at time step τ.

Table 5.9. Correlation function parameters and energies of Li.[a]

	S State	P State
a_1	0.5	0.5
b_1	3.5	3.0
E (VMC)	-7.4506(10)	-7.3865(10)
E (QMC)	-7.47809(24)	-7.41031(22)
E_{exact}	-7.47807[b]	-7.41016[c]

[a] All energies are in h.
[b] Ref. 116.
[c] Ref. 12.

Table 5.10. Single-state properties and oscillator strengths for Li, first set of calculations.[a]

	Single-State Properties (S State)		
t	$-E$	r_p	r_p^2
0 (VMC)	7.4506(10)	1.6811(18)	6.247(14)
5	7.47810(28)	1.6684(23)	6.167(20)
10	7.47830(21)	1.6663(53)	6.169(52)
20	7.47841(22)	1.6639(44)	6.124(33)
30	7.47839(13)	1.6610(66)	6.104(63)
Exact or estimated	7.47807[b]	1.6623[c]	6.118[c]

	Single-State Properties (P State)		
t	$-E$	r_p	r_p^2
0 (VMC)	7.3685(10)	1.998(4)	9.703(5)
5	7.4186(18)	1.957(18)	9.63(7)
10	7.4107(6)	1.963(9)	9.36(12)
20	7.4098(12)	1.955(9)	9.26(10)
30	7.4095(20)	1.951(10)	9.21(12)
Exact[d]	7.4102	---	---

	Oscillator Strengths		
t	c_1^2, c_2^2	Second-order[e]	Pure
0 (VMC)	---	0.780(4)	0.780(4)
5	0.50, 0.50	0.757(7)	0.762(7)
10	0.50, 0.50	0.771(9)	0.778(10)
20	0.33, 0.67	0.745(7)	0.748(8)
30	0.50, 0.50	0.767(11)	0.775(11)
Experiment[f]	---	0.742(1)	

[a] The unit of time is h^{-1}, the unit of energy is h, and distances are in *bohr*.
[b] Ref. 116.
[c] Ref. 11.
[d] Ref. 12.
[e] See Eq. (5.5).
[f] Ref. 142.

Table 5.11. Single-state properties and oscillator strengths for Li, second set of calculations.[a]

	Single-State Properties (S State)			
(τ,t_i,t^b)	$-E$	$-E_g$	r_p	r_p^2
(0.025,0.2,4)	7.4800(17)	7.47290(70)	1.6753(63)	6.238(65)
(0.025,0.5,10)	7.47896(63)	7.47284(42)	1.6644(52)	6.138(45)
(0.010,0.5,10)	7.47820(29)	7.47571(31)	1.6659(33)	6.147(34)
(0.010,0.0,12)	7.47750(51)	7.47497(54)	1.6714(38)	6.211(42)
Exact or estimated	7.47807[c]	7.47807[c]	1.6623[d]	6.118[d]

	Single-State Properties (P State)			
(τ,t_i,t)	$-E$	$-E_g$	r_p	r_p^2
(0.025,0.2,4)	7.4061(17)	7.4039(8)	1.989(9)	9.67(11)
(0.025,0.5,10)	7.4102(8)	7.4037(6)	1.968(7)	9.48(10)
(0.010,0.5,10)	7.4140(18)	7.4087(3)	1.958(7)	9.33(8)
(0.010,0.0,12)	7.4127(18)	7.4080(4)	1.970(9)	9.74(14)
Exact[e]	7.4102	7.4102	---	---

	Oscillator Strengths	
(τ,t_i,t)	Second-order[f]	Pure
(0.025,0.2,4)	0.765(8)	0.768(8)
(0.025,0.5,10)	0.763(10)	0.770(11)
(0.010,0.5,10)	0.741(7)	0.742(7)
(0.010,0.0,12)	0.758(9)	0.762(10)
Experiment[g]	---	0.742(1)

[a] The units are given in Table 5.10.
[b] All initial time steps are 0.002 h^{-1}, τ is the final time step, t_i is the propagation time with $\tau = 0.002\ h^{-1}$, and t is the total convergence time.
[c] Ref. 116.
[d] Ref. 11.
[e] Ref. 12.
[f] See Eq. (5.5).
[g] Ref. 142.

Table 5.12. Oscillator strengths and excited-state lifietimes for the $2^2S \rightarrow 2^2P$ transition of Li.[a]

Method	Oscillator Strength	2^2P Lifetime
Hartree-Fock[b]	0.768	26.36
CI[b]	0.753	26.89
Hylleraas expansion[c]	0.748	27.07
QMC[d]	0.7419(69)	27.41(35)
Experiment[e]	0.7416(12)	27.29(4)

[a] Excited-state lifetimes are in nanoseconds.
[b] Ref. 111.
[c] Ref. 12.
[d] The QMC results are derived using the QMC-computed energy difference of 0.06778(30) h while the remaining theoretical results are obtained using the exact energy difference of 0.06791 h.
[e] Ref. 142.

Figure 5.1. QMC walk within a QMC walk. The primary (vertical) QMC walk is guided by Ψ_{T_1} and generates the distribution $\Psi_{T_1}\phi_1$. During the primary walk, secondary (horizontal) QMC walks, guided by Ψ_{T_2}, are performed to sample asymptotic populations, proportional to ϕ_2/Ψ_{T_2}. Combining this weight with an operator (e.g. x) allows the evaluation of the transition matrix element.

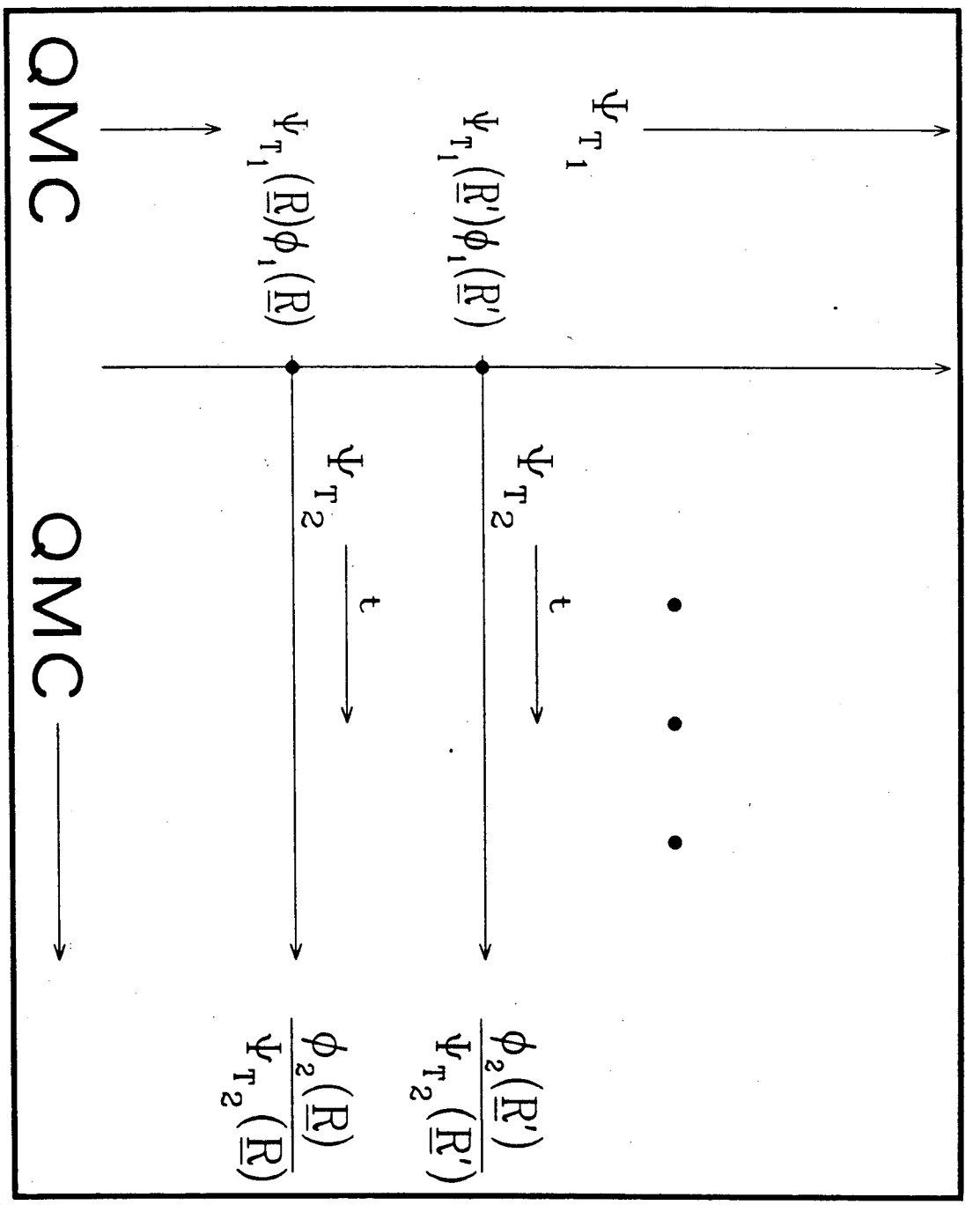


Figure 5.2. VMC walk with two QMC "side walks". Configuration space is sampled using a guided Metropolis (VMC) walk (vertical) guided by Ψ_g. Starting with points sampled during the Metropolis walk, two separate QMC walks guided by Ψ_{T_i} (horizontal) are evolved forward in imaginary time. The asymptotic populations of the QMC walks are proportional to ϕ_1/Ψ_{T_1} and ϕ_2/Ψ_{T_2}. These factors may be combined with multiplicative operators to obtain transition matrix elements.

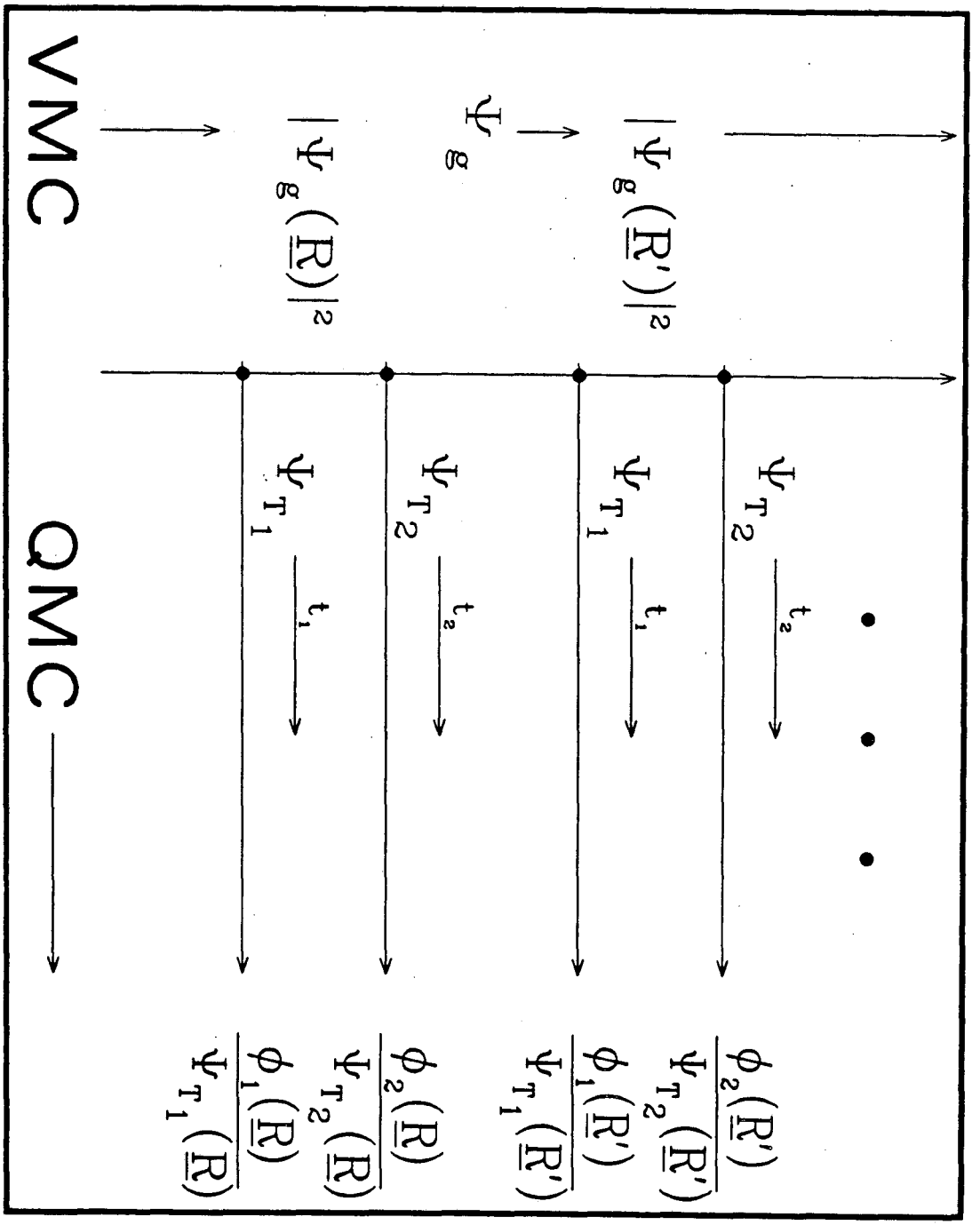
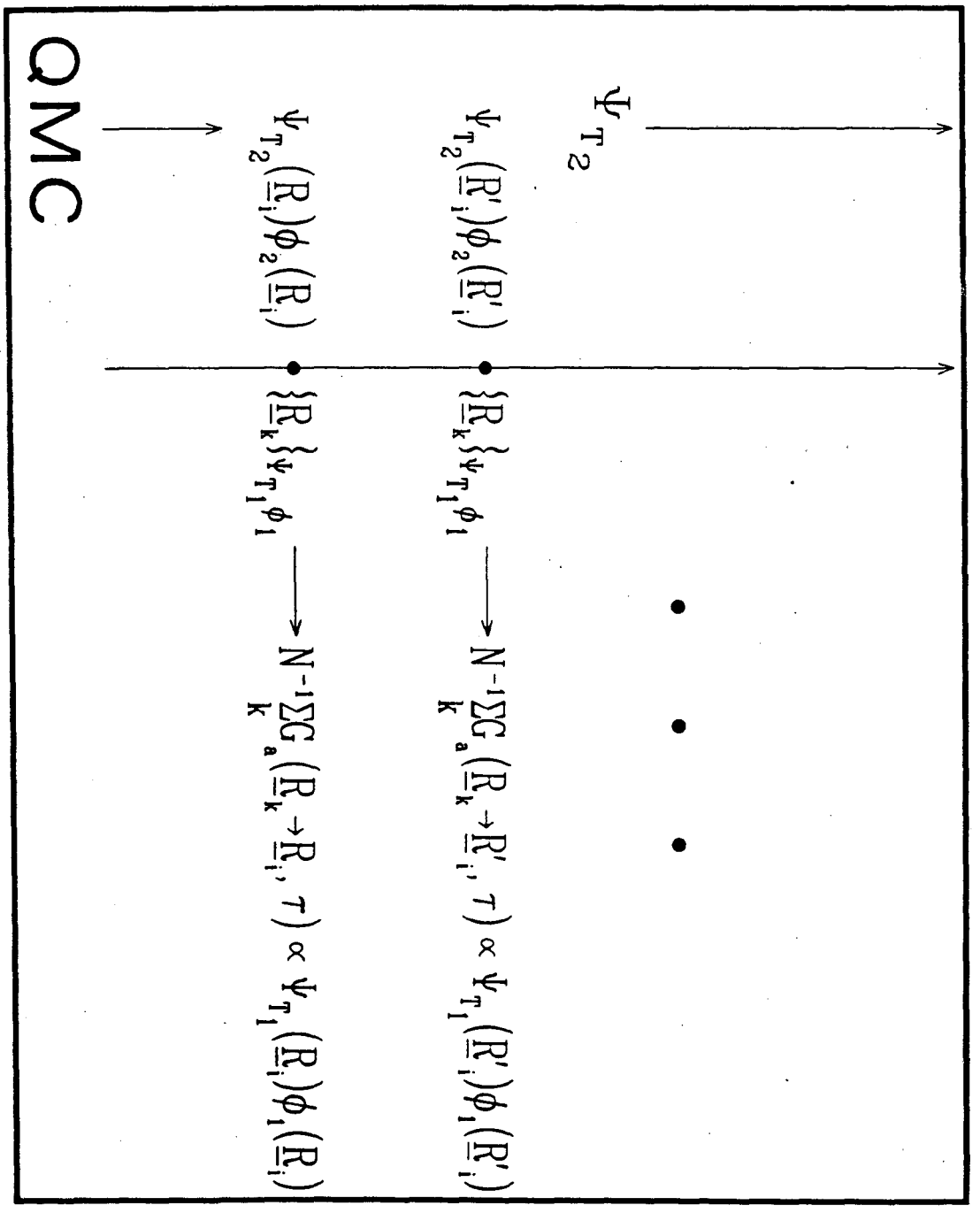


Figure 5.3. Green's function averaging approach. As in Method 1, the distribution $\Psi_{T_2}\phi_2$ is sampled using a QMC walk (vertical) guided by Ψ_{T_2}. For a point $\underline{R}_i$ sampled from $\Psi_{T_2}\phi_2$, a short-time approximation to the Green's function is averaged over points $\underline{R}_k$ drawn from the QMC distribution $\Psi_{T_1}\phi_1$. This average is proportional to $\Psi_{T_1}(\underline{R}_i)\phi_1(\underline{R}_i)$. Combining these with the operator of interest results in an estimator of the transition matrix element.



Chapter 6

Summary and Conclusions

This thesis has described three areas of application of the fixed-node diffusion QMC approach: energies (Chap. 3), single-state properties other than the energy (Chap. 4), and multi-state properties (Chap. 5).

Computations of the energy have been performed for several first-row atoms and molecules. It has been found that simple trial functions, generally derived from a single Slater determinant, have yielded accurate energies. QMC computations on small systems, i.e., H_3, Li, LiH, and Li_2, have yielded much success – obtaining 98-100% of the correlation energy (CE). For the remaining higher-Z systems that have been studied in detail, BH, N, N_2, H_2O, F and F^-, single-determinant trial functions yield approximately 90% of the CE. Thus if near-exact energies are desired on a general basis, trial functions with improved nodal structures will be required.

As has already been observed for Be[56] and Li_2, improvement upon a single-determinant trial function, so that nearly exact energies are obtained, is readily accomplished by employing a small multi-determinantal trial function which describes effects known to be important. This is encouraging as it implies that similar improvements, over single-determinant energies, may be obtained for other systems. However, it is important to realize that that increasing the complexity of the trial function along standard *ab initio* guidelines does not necessarily yield an improved nodal structure. A

simple example of this was found for H_3 where a near-HF basis set introduced an additional node.

Therefore, it is also useful to consider less standard trial function forms. An important advantage possessed by QMC methods is that only trial function derivatives are required, rather than integrals, allowing a much greater flexibility in the choice of the trial function than standard theoretical approaches. The extent to which this flexibility can be exploited has been demonstrated by Umrigar *et al.*[24] in obtaining a trial functions for Li and Be whose VMC energies yielded more than 99% of the CE. If trial functions approaching this accuracy, e.g., 90% of the CE obtained by VMC, can be found in a more general context, then near-exact QMC energies should result.

In addition to fixed-node error, the time-step bias resulting from the short-time approximation to the Green's function must be considered. That is, estimates of the unbiased energy must be obtained from calculations of biased energies. When time-step bias is large, unbiased estimates are difficult to obtain and may require computations at small time steps which is quite costly. (See, for example, the discussion of the F–F^- calculations in Chap. 3.) We have found that how well the electron-nuclear cusp condition is satisfied has a large effect on time-step bias. Choosing trial function parameters to satisfy this condition (exactly or approximately depending on the number of electrons) has resulted in QMC energies with a greatly reduced time-step bias allowing reliable estimates of the $\tau = 0$ value. For example, for H_2O we were able to compute at values of τ for which bias was not discernible, without incurring excessive computational cost.

In addition to the energy, the computation of other single-state properties has been

explored. Both single- and double-walk algorithms have been presented, and modifications which improve efficiency have been implemented, i.e., carrying weights or exploiting the correlation between pure and trial expectation values. Thus far, small systems such as H_2 and LiH can be treated with high accuracy. Computations on larger systems have not yet been implemented, and, as discussed above, trial functions more accurate than the single-determinant variety are desired.

The final avenue of exploration involved the computation of transition dipole moments. Three algorithms were presented and implemented on the H atom. These calculations introduced new considerations, such as trial function normalizations, overlap integrals, and other quantities. The approaches described in Secs. 5.3 and 5.4 were found to be the most viable.

Subsequent to the H atom calculations, Method 2 (VMC and two QMC side walks) was employed in computing the transition dipole moment, and thereby the oscillator strength and excited-state lifetime, for the $2^2S \rightarrow 2^2P$ transition of Li. The resulting oscillator strength and excited-state lifetime were found to be of excellent quality. The high computational cost of the calculations, however, emphasizes, as with the other applications, that more accurate trial functions are desired to increase the applicability of the approaches developed here.

Appendix A

A Condition for the Equivalence of Nodal Volumes

Here, we show that when there exists a one-to-one transformation between two nodal volumes, which does not change the Hamiltonian, then the nodal volume energies will be identical.

Consider two nodal volumes v_α and v_β and the corresponding nodal volume energies ε_α and ε_β, respectively. These energies are given by

$$\varepsilon_\alpha = \int_{\underline{R}\in v_\alpha} d\underline{R}\, \phi_\alpha^*(\underline{R}) H\, \phi_\alpha(\underline{R}) / \int_{\underline{R}\in v_\alpha} d\underline{R}\, \phi_\alpha^*(\underline{R})\phi_\alpha(\underline{R}) \tag{A.1}$$

and

$$\varepsilon_\beta = \int_{\underline{R}'\in v_\beta} d\underline{R}\, \phi_\beta^*(\underline{R}') H\, \phi_\beta(\underline{R}') / \int_{\underline{R}\in v_\beta} d\underline{R}'\, \phi_\beta^*(\underline{R}')\phi_\beta(\underline{R}') \,. \tag{A.2}$$

Let there exist a one-to-one transformation, T, acting on electron coordinates such that for $\underline{R} \in v_\alpha$ and $\underline{R}' \in v_\beta$, $\underline{R} = T\underline{R}'$ and $TH = H$. Since T is one-to-one, there exists an inverse, T', such that $\underline{R}' = T'\underline{R}$ and $T'H = H$.

Employing the transformation T in Eq. (A.1) yields

$$\varepsilon_\alpha = \int_{T\underline{R}'\in v_\alpha} d(T\underline{R}')\phi_\alpha^*(T\underline{R}')TH\, \phi_\alpha(T\underline{R}') / \int_{T\underline{R}'\in v_\alpha} d(T\underline{R}')\phi_\alpha^*(T\underline{R}')\phi_\alpha(T\underline{R}') \,. \tag{A.3}$$

Since T is one-to-one, the integrals above may be written as an integration over $\underline{R}'$. Also, for every $\underline{R}' \in v_\beta$, we may define a new function in v_β as $\phi'_\beta(\underline{R}') = \phi_\alpha(\underline{R})$ for $\underline{R} = T\underline{R}'$. Thus, there results

$$\varepsilon_\alpha = \int_{\underline{R}'\in v_\beta} d\underline{R}'\phi'^*_\beta(\underline{R}')H\phi'_\beta(\underline{R}')/\int_{\underline{R}'\in v_\beta} d\underline{R}'\phi'^*_\beta(\underline{R}')\phi'_\beta(\underline{R}') \ . \tag{A.4}$$

Since ε_β is the minimum energy obtainable in v_β, $\varepsilon_\alpha \geq \varepsilon_\beta$. However, substituting $T'\underline{R}$ into Eq. (A.2) yields by the same reasoning above, $\varepsilon_\beta \geq \varepsilon_\alpha$ and, therefore, $\varepsilon_\alpha = \varepsilon_\beta$. Finally, if ϕ_α and ϕ_β are unique, i.e., no degeneracy exists, then $\phi'_\beta(\underline{R}') = \phi_\beta(\underline{R}') = \phi_\alpha(\underline{R})$. [Note that we may also obtain $\phi_\beta(\underline{R}') = -\phi_\alpha(\underline{R})$.]

Appendix B

The Computation of the Trial Function and its Derivatives

As discussed in Sec. 2.5, a move $\underline{R} \rightarrow \underline{R}'$ is performed in a step-wise fashion by obtaining new electron coordinates for each electron one at a time. Therefore, we now consider the move above only to consist of $\underline{r}_1 \rightarrow \underline{r}'_1$; computations concerning the moves of the other electrons are completely analogous. The first n_α electrons are assumed to be spin-up and the remaining $N - n_\alpha$ are spin-down. For each electron move, we desire an efficient algorithm to compute the quantities necessary for the QMC simulation.

For notational convenience in the equations below, we define,

$$\underline{R}_\alpha \equiv (\underline{r}_1, \underline{r}_2, \cdots, \underline{r}_{n_\alpha}) ,$$

$$\underline{R}'_\alpha \equiv (\underline{r}'_1, \underline{r}_2, \cdots, \underline{r}_{n_\alpha}) , \quad \text{and}$$

$$\underline{R}_\beta \equiv (\underline{r}_{n_\alpha+1}, \underline{r}_{n_\alpha+2}, \cdots, \underline{r}_N) . \tag{B.1}$$

From Eq. (2.38), $\Psi_T(\underline{R})$ may be written as

$$\Psi_T(\underline{R}) = \mathrm{Det}[\mathbf{A}^\alpha(\underline{R}_\alpha)]\mathrm{Det}[\mathbf{A}^\beta(\underline{R}_\beta)]F(\underline{R}) , \tag{B.2}$$

and analogously at $\underline{R}'$. F is the product of the E-E and E-N correlation functions, and $\mathbf{A}$ is the matrix formed from the MOs. The quantities of interest are $\Psi_T(\underline{R}')/\Psi_T(\underline{R})$, and $\Psi_T^{-1}\nabla_1\Psi_T$ and $\Psi_T^{-1}\nabla_1^2\Psi_T$ at $\underline{R}$ and $\underline{R}'$. From Eq. (B.1), the spin-down (β) determinant cancels from the quantities under consideration. Also, the correlation functions and their derivatives are computed straightforwardly, leaving the computation of the

spin-up determinant and its derivatives as the present concern.

Specifically, we must evaluate $\mathrm{Det}[\mathbf{A}^{\alpha}(\underline{R}'_{\alpha})]/\mathrm{Det}[\mathbf{A}^{\alpha}(\underline{R}_{\alpha})]$ and $\mathrm{Det}[\mathbf{A}^{\alpha}]^{-1}(\nabla_1, \nabla_1^2)\mathrm{Det}[\mathbf{A}^{\alpha}]$ at $\underline{R}_{\alpha}$ and $\underline{R}'_{\alpha}$. The first step towards this end is to obtain the MOs, their gradient, and Laplacian at $\underline{r}_1$ and $\underline{r}'_1$. As can be seen above, a total of ten determinants are required. If each is computed individually, the computational cost scales as $10n_{\alpha}^2$.

A more efficient alternative is implemented using the inverse of $\mathbf{A}^{\alpha}$. Dropping the superscript, α, $\mathbf{A}$ is given by

$$\mathbf{A}(1, 2, \cdots, n_{\alpha}) = \begin{bmatrix} \psi_1(1) & \psi_2(1) & \cdots & \psi_{n_{\alpha}}(1) \\ \psi_1(2) & \psi_2(2) & \cdots & \psi_{n_{\alpha}}(2) \\ \cdots & \cdots & \cdots & \cdots \\ \psi_1(n_{\alpha}) & \psi_2(n_{\alpha}) & \cdots & \psi_{n_{\alpha}}(n_{\alpha}) \end{bmatrix} . \tag{B.3}$$

The inverse of $\mathbf{A}$, $\mathbf{A}^{-1}$, is computed at the start of the random walk and is updated for each electron move with an algorithm scaling as n_{α}^2. Each column of $\mathbf{A}^{-1}$ is

$$\mathbf{A}_i^{-1} = \mathrm{Det}(\mathbf{A})^{-1}(M_{i1}, M_{i2}, \cdots, M_{in_{\alpha}}) , \tag{B.4}$$

where M_{ij} is the minor of $a_{ij} = \psi_j(i)$.

Returning to the MOs, we define

$$O_1\psi(1) \equiv [O_1\psi_1(1), O_1\psi_2(1), \cdots, O_1\psi_{n_{\alpha}}(1)] , \tag{B.5}$$

where $O_1 \equiv (\nabla_1, \nabla_1^2)$. We now see that the quantities depending on Det(A) may be conveniently computed as scalar products, which scale as n_{α}. Namely,

$$\mathrm{Det}[\mathbf{A}(\underline{R}_{\alpha})]^{-1}\mathrm{Det}[\mathbf{A}(\underline{R}'_{\alpha})] = \psi(\underline{R}'_{\alpha})\cdot\mathbf{A}_1^{-1} ,$$

$$\mathrm{Det}[\mathbf{A}(\underline{R}_{\alpha})]^{-1}O_1\mathrm{Det}[\mathbf{A}(\underline{R}_{\alpha})] = [O_1\psi(\underline{R}_{\alpha})]\cdot\mathbf{A}_1^{-1} ,$$

and

$$\mathrm{Det}[\mathbf{A}(\underline{R}'_\alpha)]^{-1} O_1 \mathrm{Det}[\mathbf{A}(\underline{R}'_\alpha)] = [O_1 \psi(\underline{R}'_\alpha)] \cdot \mathbf{A}_1^{-1} / [\psi(\underline{R}'_\alpha)] \cdot \mathbf{A}_1^{-1} \ . \tag{B.6}$$

From Eq. (B.6), and the fact that $\mathbf{A}^{-1}$ must be updated for each electron moved, the computational cost now scales as $n_\alpha^2 + 9n_\alpha$, which is far more efficient than computing the determinants individually, $10n_\alpha^2$. The computational cost, T, incurred by moving all the electrons is now seen to be

$$T = n_\alpha(n_\alpha^2 + 9n_\alpha) + n_\beta(n_\beta^2 + 9n_\beta) \ , \tag{B.7}$$

or for $n_\alpha \approx n_\beta \approx N/2$,

$$T \sim \frac{1}{4}N^3 + \frac{9}{2}N^2 \ . \tag{B.8}$$

Appendix C

Antisymmetry Nodes of H_3

It is useful to calculate the nodal locations of the trial functions, since these nodes ultimately determine the accuracy of the QMC energy. Calculations of the exchange nodes lead to the diagrams shown in Figs. 3.12-3.14.

Here we demonstrate that a nodal surface, obtained by finding the zeros of the 2×2 determinant of the trial function at a fixed position of one like-spin electron, remains invariant for the previously fixed electron *anywhere* on this surface. In other words, the surface is the same independent of the positions of the like-spin electrons on it. This property greatly simplifies the picturing of the exchange nodes.

In the 2×2 case considered here, the vanishing of the determinant implies

$$\psi_1(\underline{r}_1)\psi_2(\underline{r}_2) = \psi_1(\underline{r}_2)\psi_2(\underline{r}_1) \ , \tag{C.1}$$

where ψ_1 and ψ_2 are the two MO's. If we now move electron 2 to $\underline{r}_2'$, still on the nodal surface for fixed $\underline{r}_1$,

$$\psi_1(\underline{r}_1)\psi_2(\underline{r}_2') = \psi_1(\underline{r}_2')\psi_2(\underline{r}_1) \ . \tag{C.2}$$

Dividing Eq. (C.1) by Eq. (C.2) gives

$$\frac{\psi_2(\underline{r}_2)}{\psi_2(\underline{r}_2')} = \frac{\psi_1(\underline{r}_2)}{\psi_1(\underline{r}_2')} \ , \tag{C.3}$$

which may be written as

$$\psi_1(\underline{r}_2')\psi_2(\underline{r}_2) = \psi_1(\underline{r}_2)\psi_2(\underline{r}_2') \ . \tag{C.4}$$

Thus we find that if $(\underline{r}_1, \underline{r}_2)$ and $(\underline{r}_1, \underline{r}_2')$ are *both* on a nodal surface, the determinant

also vanishes at $(\underline{r}_2, \underline{r}_2')$. So any pair points on the locus of points mapped out at fixed $\underline{r}_1$ is also on the nodal surface.

Appendix D

Variance of Rounded Versus Unrounded Weights

Here, we compare the variance obtained for the weight when it is integer rounded with the correct expectation value versus when its full real value is kept. The difference is significant when it is necessary to sample asymptotic populations.

Consider a number w and its associated probability density function, $f(w)$. The mean and variance of w are given by

$$\overline{w} = \int w f(w) dw$$

and

$$V_w = \int w^2 f(w) dw - \overline{w}^2 . \tag{D.1}$$

Let $I_w(\xi) \equiv \mathrm{int}(w+\xi)$ be the integer rounded weight. Here ξ is a uniform random variate between 0 and 1. The mean value of $I_w(\xi)$ over the uniform distribution of ξ's is given as

$$\overline{I}_w = \int_0^1 I_w(\xi) d\xi = \int_0^1 \mathrm{int}(w+\xi) d\xi = w . \tag{D.2}$$

In comparison to Eq. (D.1), the variance variance of the rounded weight is

$$V_{I_w} = \int [\int_0^1 I_w^2(\xi) d\xi] f(w) dw - [\int \overline{I}_w f(w) dw]^2. \tag{D.3}$$

Defining the remainder $r(w)$ by $w = \mathrm{int}(w) + r(w)$ one obtains for the integral over ξ,

$$\int_0^1 I_w^2(\xi)d\xi = \int_0^{1-r} [\text{int}(w + \xi)]^2 d\xi + \int_{1-r}^1 [\text{int}(w + \xi)]^2 d\xi ,$$

$$= (1-r)[\text{int}(w)]^2 + r[\text{int}(w) + 1]^2$$

$$= w^2 + r - r^2 . \tag{D.4}$$

Eq. (D.3) now becomes

$$V_{\bar{I}_w} = \int [w^2 + r(w) - r^2(w)] f(w) dw - \bar{w}^2 , \tag{D.5}$$

or

$$V_{\bar{I}_w} = V_w + \int [r(w) - r^2(w)] f(w) dw. \tag{D.6}$$

Since $r(w) - r(w)^2 \geq 0$ and $f(w)$ is, by definition positive definite, $V_{\bar{I}_w} \geq V_w$. This quantifies the lower variance in the weights when integer rounding is avoided. Though the additional variance in a single integer rounded weight is small, the reduction in the variance of the asymptotic populations (which are essentially the products of several weights) can be significant.

References

1. E. Schrödinger, Phys. Rev. **28**, 1049 (1926).

2. H. Erying, J. Walter, and G. E. Kimball, *Quantum Chemistry*, (John Wiley and Sons, Inc., New York, 1944).

3. Many descriptions exist; for a review see S. T. Epstein, *The Variational Method in Quantum Chemistry*, (Academic Press, N. Y., 1974).

4. R. J. Bartlett, Ann. Rev. Phys. Chem. **32**, 359 (1981).

5. R. O. Jones, Adv. Chem. Phys. **67**, 413 (1987).

6. L. Pauling and E. B. Wilson, *Introduction to Quantum Mechanics*, (McGraw-Hill, New York, 1955).

7. J. K. L. MacDonald, Phys. Rev. **43**, 830 (1933).

8. E. A. Hylleraas, Z. Physik **48**, 469 (1928).

9. C. L. Pekeris, Phys. Rev. **112**, 1649 (1958).

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

11. F. W. King and V. Shoup, Phys. Rev. A **33**, 2940 (1986).

12. T. Alhenius and S. Larsson, Phys. Rev. A **8**, 1 (1973).

13. F. Sasaki and M. Yoshimine, Phys. Rev. A **9**, 17, 26 (1974).

14. D. R. Hartree, Proc. Cambridge Phil. Soc. **24**, 89 (1928).

15. V. Fock, Z. Physik **61**, 126 (1930).

16. J. C. Slater, Phys. Rev. **35**, 210 (1930).

17. C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).

18. P.-O. Löwdin, in *Advances in Chemical Physics, Vol. II*, I. Prigogine, ed. (Interscience Publishers, Inc., New York, 1959), p. 235.

19. *Ibid.*, pp. 259-262.

20. T. L. Gilbert, Rev. Mod. Phys. **35**, 491 (1963).

21. C. E. Dykstra, S. Liu, and D. J. Malik, to appear in *Advances in Chemical Physics.*

22. R. Crossley, Phys. Scripta **T8**, 117 (1984).

23. H. Conroy, J. Chem. Phys. **41**, 1331 (1964).

24. C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).

25. Z. Sun, S. Huang, R. N. Barnett, and W. A. Lester, Jr., to be published.

26. M. H. Kalos and P. A. Whitlock, in *Monte Carlo Methods, Vol. I: Basics* (Wiley,

27. J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods* (Methuen, London, 1964), pp. 52-59.

28. P. R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences* (McGraw-Hill, 1969).

29. G. Goertzel and M. H. Kalos, *in* "Progress in Nuclear Energy" (D. J. Hughes, Ed.), 368, Series I, Vol. 2, (Pergammon, New York, 1958).

30. N. Metropolis and S. Ulam, J. Amer. Stat. Assoc. **44**, 335 (1949).

31. M. H. Kalos, Phys. Rev. **128**, 1791 (1962).

32. M. H. Kalos, J. Comp. Phys. **1**, 257 (1966).

33. M. H. Kalos, Phys. Rev. A **2**, 250 (1970).

34. M. H. Kalos, D. Levesque, and L. Verlet, Phys. Rev. A **9**, 250 (1974).

35. D. M. Ceperley and M. H. Kalos, in *Monte Carlo Methods in Statistical Physics*, K. Binder, Ed. (Springer-Verlag, N. Y., 1979).

36. K. E. Schmidt and M. H. Kalos, in *Monte Carlo Methods in Statistical Physics II*, K. Binder, Ed. (Springer-Verlag, N. Y., 1985).
N. Y., 1986).

37. D. M. Ceperley and B. J. Alder, Science **231**, 555 (1986).

38. D. M. Arnow, M. H. Kalos, M. A. Lee, and K. E. Schmidt, J. Chem. Phys. **77**, 5562 (1982).

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

40. D. M. Ceperley, J. Comp. Phys. **51**, 404 (1983).

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

42. K. E. Schmidt and J. W. Moskowitz, J. Stat. Phys. **43**, 1027 (1986).

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

44. D. M. Ceperley and B. J. Alder, Phys. Rev. B **36**, 2092 (1987).

45. J. B. Anderson, J. Chem. Phys. **63**, 1499 (1975).

46. J. B. Anderson, J. Chem. Phys. **65**, 4121 (1976).

47. J. B. Anderson, Int. J. Quantum Chem. **15**, 109 (1979).

48. S. M. Rothstein and J. Vrbik, J. Chem. Phys. **87**, 1902 (1987).

49. J. B. Anderson and D. R. Garmer, J. Chem. Phys. **87**, 1903 (1987).

50. P. J. Reynolds, R. K. Owen, and W. A. Lester, Jr., J. Chem. Phys. **87**, 1905 (1987).

51. D. J. Klein and H. M. Pickett, J. Chem. Phys. **64**, 4811 (1976).

52. P. J. Reynolds, M. Dupuis, and W. A. Lester, Jr., J. Chem. Phys. **82**, 1983 (1985).

53. R. N. Barnett, P. J. Reynolds, and W. A. Lester, Jr., J. Chem. Phys. **82**, 2700 (1985).

54. R. N. Barnett, P. J. Reynolds, and W. A. Lester, Jr., J. Chem. Phys. **84**, 4992 (1986).

55. R. N. Barnett, P. J. Reynolds, and W. A. Lester, Jr., J. Phys. Chem. **91**, 2004 (1987).

56. R. J. Harrison and N. C. Handy, Chem. Phys. Lett. **113**, 257 (1985).

57. D. R. Garmer and J. B. Anderson, J. Chem. Phys. **86**, 4025 (1987).

58. *Ibid.*, 7237 (1987).

59. J. B. Anderson, J. Chem. Phys. **82**, 2662 (1985).

60. G. E. P. Box and M. E. Mueller, Ann. Math. Statist. **29**, 610 (1958).

61. H. M. Frey and G. J. Kennedy, J. Chem. Soc. Chem. Commun. **233**, (1975); J. Chem. Soc. Faraday Trans. 1 **73**, 164 (1977).

62. P. F. Zittel, G. B. Ellsion, S. V. O'Neil, E. Herbst, W. C. Lineberger, and W. P. Reinhardt, J. Am. Chem. Soc. **98**, 3731 (1976); P. C. Engelking, R. R. Corderman, J. J. Wendoloski, G. B. Ellsion, S. V. O'Neil, and W. C. Lineberger, J. Chem. Phys. **74**, 5460 (1981).

63. R. K. Lengel and R. N. Zare, J. Am. Chem. Soc. **100**, 7495 (1978); C. C. Hayden, D. M. Neumark, K. Shobatake, R. K. Sparks, and Y. T. Lee, J. Chem. Phys. **76**, 3607 (1982); A. R. W. McKellar, P. R. Bunker, T. J. Sears, K. M. Evenson, R. J. Saykally, and S. R. Langhoff, J. Chem. Phys. **79**, 5251 (1983); D. G. Leopold, K. K. Murray, and W. C. Lineberger, J. Chem. Phys. **81**, 1048 (1984).

64. D. G. Truhlar and R. E. Wyatt, Ann. Rev. Phys. Chem. **27**, 1 (1976).

65. B. Liu, J. Chem. Phys. **58**, 1925 (1973).

66. P. Siegbahn and B. Liu, J. Chem. Phys. **68**, 2457 (1978).

67. B. Liu, J. Chem. Phys. **80**, 581 (1984).

68. NRCC Software Catalog **1**, 60 (1980), LBL-10811, UC-4. [Available from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22161; Price Code: A05.]

69. P. J. Reynolds, R. N. Barnett, and W. A. Lester, Jr., Int. J. of Quant. Chem.: Quant. Chem. Symp. **18**, 709 (1984).

70. E. Clementi, J. Chem. Phys. **38**, 2248 (1963).

71. E. Clementi and A. D. McLean, Phys. Rev. **133**, A419 (1964).

72. C. W. Scherr, J. N. Silverman, and F. A. Matsen, Phys. Rev. **128**, 2675 (1962).

73. C. Moore, "Atomic Energy Levels" (National Bureau of Standards, Washington, D.C., 1958).

74. I. Oksuz and O. Sinanoglu, Phys. Rev. **181**, 54 (1969).

75. H. F. Schaefer III, R. A. Klemm, and F. E. Harris, J. Chem. Phys. **51**, 4643

(1969).

76. A. Veillard and E. Clementi, J. Chem. Phys. **49**, 2415 (1968).

77. A. W. Weiss, Phys. Rev. A **3**, 126 (1971).

78. C. M. Moser and R. K. Nesbet, Phys. Rev. A **4**, 1336 (1971); **6**, 1710 (1972).

79. S. Fraga, J. Karwowski, and K. M. S. Saxena, *Handbook of Atomic Data* (Elsevier, Amsterdam, 1976).

80. K. Jankowski, P. Malinowski, and M. Polasik, Phys. Rev A **22**, 51 (1980).

81. S. A. Kucharski, Y. S. Lee, G. D. Purvis III, and R. J. Bartlett, Phys. Rev. A **29**, 1619 (1984).

82. B. H. Botch and T. H. Dunning, Jr., J. Chem. Phys. **76**, 6046 (1982) and references therein.

83. B. O. Roos, A. J. Sadlej, and P. E. M. Siegbahn, Phys. Rev. A **26**, 1192 (1982).

84. K. Raghavachari, J. Chem. Phys. **82**, 4142 (1985).

85. L. Adamowicz and R. J. Bartlett, J. Chem. Phys. **84**, 6837 (1986).

86. B. Edlén, J. Chem. Phys. **33**, 98 (1960).

87. R. S. Berry and C. W. Reiman, J. Chem. Phys. **38**, 1540 (1963).

88. H. Hotop and W. C. Lineberger, J. Phys. Chem. Ref. Data **4**, 539 (1975).

89. E. Clementi and C. Roetti, Atomic Data and Nuclear Data Tables **14**, 177 (1974).

90. R. C. Kelly, *Atomic and Ionic Spectrum Lines Below 2000* Å, Oak Ridge National Laboratory-5922 (1982). [Available from National Technical Information Service, U. S. Department of Commerce, 5285 Port Royal Road, Springfield, Virginia 22161; price code: A19.]

91. P. E. Cade, K. D. Sales, and A. C. Wahl, J. Chem. Phys. **41**, 1976 (1966).

92. P. E. M. Siegbahn, Int. J. Quant. Chem. **23**, 1869 (1983).

93. W. C. Ermler and C.-W. Huang, Chem. Phys. Lett. **120**, 159 (1985).

94. J. S. Wright and R. J. Buenker, J. Chem. Phys. **83**, 4059 (1985).

95. C. W. Bauschlicher, Jr., and S. R. Langhoff, J. Chem. Phys. **86**, 5595 (1987).

96. H.-J. Werner and P. J. Knowles, J. Chem. Phys. **89**, 5803 (1988).

97. M. J. Frisch, R. Krishnan, and J. A. Pople, Chem. Phys. Lett. **75**, 66 (1980).

98. J. S. Binkley and M. J. Frisch, Int. J. Quant. Chem.: Quant. Chem. Symp. **17**, 331 (1983).

99. M. J. Frisch, J. A. Pople, and J. S. Binkley, J. Chem. Phys. **80**, 3265 (1984).

100. K. A. Peterson and R. C. Woods, J. Chem. Phys. **87**, 4409 (1987).

101. W. D. Laidig, P. Saxe, and R. J. Bartlett, J. Chem. Phys. **86**, 887 (1986).

102. A. D. Becke, J. Chem. Phys. **76**, 6037 (1982); **78**, 4787 (1983).

103. T. T. Rantala, B. Wastberg, and A. Rosen, Chem. Phys. **109**, 261 (1986).

104. K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure. IV. Constants of Diatomic Molecules*, (Van Nostrand Rheinhold, New York, 1979).

105. P. A. Christiansen and E. A. McCullough, Jr., J. Chem. Phys. **67**, 1877 (1977).

106. F. Grimaldi, J. Chem. Phys. **43**, 559 (1965).

107. Z. Ritter and R. Pauncz, J. Chem. Phys. **32**, 1820 (1960).

108. C. E. Moore, Natl. Bur. Standards Circ. No. 467, I (1949).

109. M. Caffarel and P. Claverie, J. Chem. Phys. **88**, 1088 (1988).

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

111. A. Weiss, Astrophys. J. **138**, 1262 (1963).

112. N. C. Handy, R. J. Harrison, P. J. Knowles, and H. F. Schaefer III, J. Phys. Chem. **88**, 4852 (1984).

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

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

115. S.-Y. Huang, Z. Sun, and W. A. Lester, Jr., J. Chem. Phys., submitted.

116. J. Muszynska, D. Papierowska, J. Pipin, and W. Woznicki, Int. J. Quant. Chem. **22**, 1153 (1982).

117. L. Laaksonen, P. Pyykko, and D. Sundholm, Chem. Phys. Lett. **96**, 1 (1983).

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

119. K. K. Verma, M. E. Koch, and W. C. Stwalley, J. Chem. Phys. **78**, 3614 (1983).

120. C. F. Bunge, Phys. Rev. A **14**, 1968 (1976).

121. P. A. Christiansen and E. A. McCullough, Jr., Chem. Phys. Lett. **55**, 439 (1978).

122. W. Meyer and P. Rosmus, J. Chem. Phys. **63**, 2356 (1975).

123. B. J. Rosenberg and I. Shavitt, J. Chem. Phys. **63**, 2162 (1975).

124. S. Aung, R. M. Pitzer, and S. I. Chan, J. Chem. Phys. **49**, 2071 (1968), wave function II in this refernce.

125. B. L. Hammond, Ph. D. thesis, pp. 32-33.

126. J. Stat. Phys. **43**, (1986).

127. P. A. Whitlock, D. M. Ceperley, G. V. Chester, and M. H. Kalos, Phys. Rev. B **19**, 5598 (1979).

128. P. J. Reynolds, R. N. Barnett, B. L. Hammond, and W. A. Lester, Jr., J. Stat. Phys. **43**, 1017 (1986).

129. (a) A. L. L. East, S. M. Rothstein, and J. Vrbik, J. Chem. Phys. **89**, 4880 (1988); P. J. Reynolds, J. Chem. Phys., accepted. This comment points out a deficiency in Ref. 16(a).

130. K. S. Liu, M. H. Kalos, and G. V. Chester, Phys. Rev. A **10**, 303 (1974).

131. B. H. Wells, Chem. Phys. Lett. **115**, 89 (1985).

132. L. Wharton, L. Gold, W. Klemperer, J. Chem. Phys. **37**, 2149 (1962).

133. C. F. Bender and E. R. Davidson, Phys. Rev. **183**, 23 (1969).

134. B. O. Roos and A. J. Sadlej, Chem. Phys. **94**, 43 (1985).

135. G. Maroulis and D. M. Bishop, Chem. Phys. **96**, 409 (1985).

136. D. M. Bishop, J. Pipin, and B. Laum, Chem. Phys. Lett. **127**, 377 (1986).

137. J. W. C. Johns, and F. A. Grimm, and R. F. Porter, J. Mol. Spectrosc. **22**, 435 (1977).

138. Though not required, for simplicity, only symmetry about 0 is considered.

139. See, for example, L. I. Schiff, *Quantum Mechanics*, Third edition (McGraw-Hill, New York, 1968), Chaps. 11 and 14.

140. H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Academic Press, New York, 1957), pp. 251-53.

141. H. Okabe, *Photochemistry of Small Molecules* (Wiley-Interscience, New York, 1978), pp. 25-26.

142. A. Gaupp, P. Kuske, and H. J. Andra, Phys. Rev. **26**, 3351 (1982).

LAWRENCE BERKELEY LABORATORY
TECHNICAL INFORMATION DEPARTMENT
1 CYCLOTRON ROAD
BERKELEY, CALIFORNIA 94720