Title
MONTE CARLO FOR THE ELECTRONIC STRUCTURE OF MOLECULES

Author
Hammond, B.L.

Publication Date
1988-11-01

LBL-26536

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

Materials & Chemical Sciences Division

Monte Carlo for the Electronic Structure of Molecules

B.L. Hammond
(Ph.D. Thesis)

November 1988

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

DISCLAIMER

LBL-26536

Monte Carlo for the Electronic Structure of Molecules

Ph.D Thesis

Brian Lee Hammond

Department of Chemistry
University of California
and
Materials and Chemical Sciences Division
Lawrence Berkeley Laboratory
1 Cyclotron Road
Berkeley, California 94720

November 1988

Monte Carlo for the Electronic Structure of Molecules

Brian Lee Hammond

Abstract

New directions in the rapidly developing quantum Monte Carlo (QMC) electronic structure method are explored for the treatment of chemically interesting systems. Electronic correlation, a difficult problem in standard *ab initio* methods, is accounted for in a simple manner in QMC. Methods developed here are aimed at expanding the capabilities and applicability of QMC. Two general problems are addressed: extending QMC calculations to heavy-atom systems, and obtaining nuclear derivative and potential energy surface information. For heavy-atom systems, two methods have been developed. The first method uses *ab initio* effective core potentials to eliminate the core electrons. The second method retains all electrons explicitly, but achieves a separation of time scales between core and valence to speed computation. These methods are applied to atomic and molecular systems containing Li, Na, Mg, C, Si, Ge, and Fe, and results are compared to other work available. Derivative and PES information is treated in three different ways: by the Hellmann-Feynman theorem, by analytic gradient techniques, and by correlated sampling of finite differences. Each of these approaches is implemented and applied to small diatomic systems.

Acknowledgements

I would like to thank all those people who have touched my life during completion of this work. Formost of these is my advisor and friend, Prof. William A. Lester, Jr. who has provided support and set an example for scientific excellence. I would also like to thank Dr. Peter Reynolds for guiding me though the intracies of Quantum Monte Carlo, and Dr. Michel Dupuis for many discussions of quantum chemistry as a whole. Finally, I wish to thank the other graduate students I have worked closely with, Dr. Randall Grimes, Dr. Christopher Dateo, Dr. Sheng-Yu Huang, and (soon to be Dr.s) Kent Owen, and Robert Barnett.

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

Dedication

To my parents, James and Jean

Table of Contents

Chapter 1

Introduction

1.1 Quantum mechanical background

Quantitative computation of the electronic structure of atoms and molecules remains a great challenge even though the underlying wave mechanics was proposed by Schrödinger [1] in 1926. For present purposes the form of the wave equation used is the time-independent Born-Oppenheimer (fixed-nuclei) non-relativistic Schrödinger equation, which can be written,

$$\mathbf{H}\Phi_I = \mathrm{E}_I\Phi_I \tag{1.1}$$

where Φ_I and E_I are the eigenfunctions and eigenvalues of $\mathbf{H}$, the electronic Hamiltonian operator,

$$\mathbf{H} = \left[-\tfrac{1}{2}\sum_{i=1}^{N_{el}}\nabla_i^2 - \sum_A \frac{Z_A}{r_{iA}} + \sum_A\sum_{B<A}\frac{Z_A Z_B}{R_{AB}}\right] + \left[\sum_{j<i}\frac{1}{r_{ij}}\right] \tag{1.2}$$

$$\equiv \mathbf{H}_1 + \mathbf{H}_2 .$$

The indices i and j refer to electrons while A and B refer to nuclei. Z is the nuclear charge. The functions Φ_I and the values E_I are the stationary states of the electrons, and the corresponding energy levels, Φ_I^2 having the interpretation of a probability density for the electronic distribution.

The major obstacle is the inability to write a closed analytic solution to Eq. 1.1 for more than one electron. The task of an approximate solution, then, is to retain the important interactions while rendering the problem soluble. The most widely used approximate solutions to the electronic Schrödinger equation are derived from either the variational

method[2], perturbation theory[3], density functional theory[4], propagator methods[5], or most recently quantum Monte Carlo methods[6].

By far the most popular of these has been the variational method. Here a distinction is made between methods whose energies are called variational (*i.e.*, the energy is an upper bound to the exact) and the *variational method.* In the variational method a trial function, Ψ_T with suitable boundary conditions and properties is chosen. The energy expectation value of Ψ_T is the functional,

$$E[\Psi_T] = \int d\mathbf{R}\Psi_T \mathbf{H}\Psi_T / \int d\mathbf{R}\Psi_T^2 . \tag{1.3}$$

The variational principle states that for any function Ψ_T, $E[\Psi_T]$ will be the absolute minimum when Ψ_T is the exact ground state solution of the Schrödinger equation. Therefore, one chooses an appropriate Ψ_T with free parameters and then minimizes $E[\Psi_T]$ with respect to these parameters to obtain the best Ψ_T of the chosen form. This yields a powerful method since any set of trial functions can be compared based on their energies, the most accurate having the lowest energy (in contrast with perturbation theory). The difficulty lies in the evaluation of Eq. 1.3. If this integral is to be analytically computable then the form of Ψ_T is greatly restricted.

The most widely used implementation of the variational method has been the Hartree-Fock method[7], where Ψ_T is a product of antisymmetrized one-electron orbitals. This form, however, neglects the cooperative motions of the electrons in avoiding each other. The energy difference between the Hartree-Fock solution, and the exact non-relativistic, Born-Oppenheimer solution is defined as the electronic *correlation energy.* Because of its conceptual simplicity Hartree-Fock has become a standard by which all methods which seek to provide a description of electron correlation are ranked, hence a brief description of the Hartree-Fock method and its descendents is given in Sec. 1.2.

Quantum Monte Carlo (QMC) is a method which has been shown to recover 90-100% of the correlation energy[8-34] in most cases. QMC is a direct numeric method for

solution of the Schrödinger equation which accounts for electronic correlation in a direct physical manner, and, in its most exact form, is completely basis set independent. In this work, the fixed-node approximation[8] is used, creating a basis set dependence; however, this dependence is very weak in comparison to HF and post-HF methods. A brief introduction to QMC methods is given in Sec. 1.3, with a detailed description being the subject of chapter 2. The remaining chapters are devoted to applications and extensions of the QMC method. Chapter 3 describes several methods for dramatically reducing the computation required for heavy atom systems. Both valence electron *pseudopotential* methods and a new all-electron damped-core QMC method are discussed. Chapter 4 is concerned with finding geometrical energy derivative information and energy minimization. Methods considered are the Hellmann-Feynman theorem, finite differences, and analytic derivatives of the QMC energy.

1.2 Hartree-Fock and Post Hartree-Fock methods

In 1927 Hartree[35] proposed an atomic model in which each electron is considered to interact with only the average potential of the other electrons. He was thus able to write a one-electron equation for each electron moving in the field of a non-coulomb central potential obtained from the other electrons. Since the potential of each electron requires the solution of the other electrons, the central potential must be solved iteratively, and thus was called the *self consistent field* (SCF).

Hartree's derivation, however, lacked firm theoretical backing until Fock[36] and Slater[37] independently rewrote Hartree's method in terms of the Schrödinger equation. This was done by writing a properly antisymmetric all-electron trial function based on Hartree's one-electron solutions, as a Slater determinant,

$$\Psi_T = \det\begin{bmatrix} \psi_1(1) & \psi_1(2) & \cdots & \psi_1(n) \\ \psi_2(1) & \psi_2(2) & \cdots & \psi_2(n) \\ \cdots & \cdots & \cdots & \cdots \\ \psi_n(1) & \psi_n(2) & \cdots & \psi_n(n) \end{bmatrix} \tag{1.4}$$

where $\psi_j(i)$ is the j'th one-electron orbital acting on the position of the i'th electron. Eq. 1.4 is commonly abbreviated by writing only the diagonal elements

$$\Psi_T = \det|\psi_1(1)\psi_2(2)\cdots\psi_n(n)|. \tag{1.5}$$

If the Hamiltonian could be written as a sum of one-electron operators, then Eq. 1.4 would be the correct form of the wavefunction. However, the two-electron repulsion, $\mathbf{H}_2$ in Eq. 1.2, prevents any such separation. The great success of the Hartree-Fock (HF) method has been that for most cases an accurate qualitative description is achieved, although quantitative energies are not necessarily obtained. Substituting this Ψ_T into Eq. 1.1 one obtains the Hartree-Fock (HF) equation for the orbitals,

$$\left[2\int \psi_i \mathbf{H}_1 \psi_i + \sum_j^{N_{el}} \left(2\int \psi_i(1)\psi_j(2)\frac{1}{r_{12}}\psi_i(1)\psi_j(2)d\mathbf{r}_1 d\mathbf{r}_2 \right.\right.$$

$$\left.\left. + \int \psi_i(1)\psi_j(2)\frac{1}{r_{12}}\psi_j(1)\psi_i(2)d\mathbf{r}_1 d\mathbf{r}_2 \right) \right] = \varepsilon_i \psi_i \tag{1.6}$$

where the quantity in brackets is the Fock operator, **F**. For atoms, Eq. 1.6 can be solved on a grid, however, for general molecular systems a simpler method is needed.

For molecules Roothaan [38] and Hall [39] proposed writing each molecular orbital (MO) as a linear combination of atomic orbitals (AO's), χ,

$$\psi_i = \sum_{\mu}^{N_{bas}} c_{i\mu}\chi_\mu . \tag{1.7}$$

In actual fact these AO's need not describe the orbitals of the constituent atoms, but are taken to be some set of atom centered functions which loosely resemble the AO's. Typical choices for AO basis functions are the Slater-type functions (STF's)[40] and the Gaussian-type functions (GTF's)[41]. Minimization of the ψ_i's with respect to the

energy leads to the familiar Hartree-Fock-Roothaan equations[7], written in matrix form,

$$\mathbf{FC} = \mathbf{SCE} \tag{1.8}$$

where **C** is the matrix of the linear coefficients, **E** is a diagonal matrix containing the MO energies, **S** is the overlap matrix of the AO's, and **F** is the Fock matrix. As before, the Fock matrix, contains integrals involving the sought-after MO's, hence the solution must be found iteratively: an initial guess is made for **C**, then Eq. 1.7 is solved to produce a new **C** matrix. This process continues until self-consistency is achieved between the input and output **C**'s. This procedure is often referred to as either HF or SCF, however HF formally refers to the exact solution to Eq. 1.6 or 1.8, whereas SCF refers to any finite basis set solution to Eq. 1.8.

Although SCF has had great success in treating simple closed-shell systems, there are some notable failures. The SCF method does not predict electron affinities for many systems for which electron affinities have been measured[42] (as was noted by Hartree himself). SCF also has trouble describing van der Waals systems[43]. For both electron affinities and van der Waals molecules binding occurs though the cooperative motions of the electrons, thus the reason for failure is the neglect of the electronic correlation. So, even though the HF approximation may be valid for many systems, a large number of interesting systems are beyond its reach.

Electronic correlation can be treated without abandoning the HF approach entirely. In the Configuration Interaction (CI) method the exact wavefunction is written as an expansion of Slater determinants, each describing a different electronic state,

$$\Phi_I = \sum_{J=1}^{\infty} C_{IJ} \Psi_J \tag{1.9}$$

where the Ψ_J are the solutions to Eq. 1.8. Thus the trial function is now a double basis set expansion. The 1-particle AO basis set of Eq. 1.6 and the N-particle basis set of Eq. 1.9. The most rigorous use of the variational method involves minimizing the energy with respect to both the $c_{i\mu}$ and the C_{IJ}. This is known as the Multi-configuration SCF

method (MCSCF). Unfortunately, for even moderate numbers of electrons this becomes prohibitive. In practice the largest CI expansions for large systems involve only determinants representing single and double (SDCI) excitation from the HF ground state, although it is usually better to use a determinant constructed from MO's which have been obtained from a limited MCSCF procedure rather than from HF. With modern computers and the use of GTF's, as many as 10^5 to 10^6 determinants may be included in the CI trial function.

The CI procedure and its descendants, however, converge slowly in both the 1-particle and N-particle basis sets. This difficulty has prompted research into the other methods listed above. Rather than exploring all the available methods, it is useful to investigate the slow convergence of CI. This convergence problem was recognised early in the development of CI techniques [44] and is ultimately due to the original one-particle description. A more rapidly converging trial function is that used by Hylleraas [45] for He atom. In the Hylleraas function all the inter-particle distances were explicitly included: in particular the inter-electron distance, r_{ij}. Thus the correct behavior of the wavefunction when two electrons approach each other can be expressed directly, rather than indirectly though the superposition of independent-particle states. The Hylleraas method has been explored for 2-electron systems by Pekeris[46] and others[47] with great success. The draw back of such methods is that the required integration becomes intractable for molecules, and more computation is required for a compact trial function than for a very large CI expansion.

1.3 Monte Carlo methods

One way to avoid analytic integration of trial functions involving r_{ij} is to devise an efficient numeric scheme. For large numbers of electrons grid methods are impossible. Instead Monte Carlo integration can be used. Use of the Metropolis algorithm[48] for this purpose is discussed in detail in Chapter 2. Application of the variational method is still

problematic since numeric integration is slow, compared to analytic integration, making energy optimization expensive when a large numbers of parameters are involved.

The quantum Monte Carlo method simulates the Schrödinger equation directly by random walk in imaginary time, thus avoiding basis sets or grids, and treats the electronic correlation in a simple, physical manner. QMC has had its most wide spread application in many-body physics[49], and only recently has begun to make inroads into quantum chemistry. Several excellent reviews of QMC methods are currently available[49-54]. The growth of usage of QMC for molecular systems has been hampered by the large computational cost involved in the procedure, but the great success of QMC in treating many-body effects in physics is a strong motivation to its development. An additional bonus of the QMC procedure is that not only does it provide highly accurate numerical solutions to the Schrödinger equation, it also provides an independent check of HF-CI wave functions and explicitly correlated functions, and thus insight into the nature of the true wave function can be obtained.

Monte Carlo methods use random numbers to do integrals, simulate random processes, and solve equations[55]. QMC utilizes a stochastic approach to solve the time-dependent Schrödinger equation by writing it in imaginary time and simulating the resulting equation as diffusion with branching[8]. By this means the exact ground state energy can be sampled. However, even though the energy being sampled is exact, it is only known to within a certain statistical precision. In this sense QMC and CI are similar — both provide the precise value of the true energy for infinite computation. An important difference is that in QMC one obtains an accurate estimate of the precision, which is generally not available in CI. Furthermore, since the Schrödinger equation is being solved numerically, QMC avoids the large basis set expansions, both 1-particle and N-particle, required by CI. The strength of CI-based techniques, however, is that in many cases smaller expansions can be used, at low cost, with good cancellation of errors between the systems of interest.

Several different QMC methods are currently in use. Originally a random walk with branching algorithm was used[16], but this suffered from the large statistical variance due to the coulomb singularities. This difficulty can be overcome by introducing an appropriate importance sampling procedure[8]. In importance sampling, an analytic *trial wave function* is introduced to guide the random walk. If the trial function is properly chosen then the coulomb singularities can be "smoothed out"[8].

A second, more difficult problem arises in the treatment of Fermi statistics. The many-body ground state wave function changes sign when two particles are exchanged. However, the wave function plays the role of the *density* of walkers in QMC and thus must be positive. With importance sampling, if one restricts the solution so that the positive and negative regions coincide with those of the trial wave function, the product can be made positive everywhere. This constitutes the fixed-node approximation [30]. This approximation can be relaxed by tracking the walkers in positive and negative regions separately and "releasing" the nodal constraint [31]. Unfortunately, this methods suffers from exponentially increasing statistical noise and becomes entirely unstable for high Z systems (although some interesting recent efforts has been made to address this problem [27]).

An additional computational problem is the nesessity of determining the Green's function for the walk. A simple short-time approximation can be made[8] which must then be extrapolated to zero time step. This time step bias can be eliminated with some additional computational effort[23,31], and coupled with the released node procedure gives exact results for small systems [31]. QMC methods have been further extended into quantum chemistry by algorithms for computing properties[11], analytic geometrical derivatives[11], excited states[13], as well as methods to eliminate the core electrons [15,32-34].

It is the purpose of this investigation to further the development of QMC for systems of chemical interest. Two main lines of research will be discussed. The first is the

extension of QMC to the entire periodic table by some approximate treatment of the core electrons. The second line is to facilitate the study of potential energy surfaces and provide energy derivative information.

CHAPTER 2

Fixed-Node Diffusion Quantum Monte Carlo

2.1 Introduction

In this chapter the quantum Monte Carlo method is described with emphasis on the fixed-node diffusion quantum Monte Carlo (FNDQMC) scheme. The reason for this emphasis is that to date the FNDQMC method has had the greatest impact in the field of quantum chemistry. Yet since much of the discussion is general to all implementations of QMC, where possible the main concepts from the other QMC methods are discussed.

The natural starting point in discussion of Monte Carlo methods is the evaluation of integrals. In particular, one can use Monte Carlo to evaluate the variational energy (Eq. 1.3) of a trial function [56]. This method, called variational Monte Carlo (VMC) here, is the topic of Sec. 2.2. Clearly, in cases where the trial function is analytically integrable, VMC would be a waste of time. The value of the VMC method is in the treatment of explicitly correlated trial functions, *i.e.*, those containing the inter-electronic distance. Such functions were the first to be used in addressing the electronic-correlation problem [45,46], but use dwindled with the development of the relatively faster Configuration-Interaction (CI) techniques [57]. With the advent of supercomputers, however, small CPU-bound algorithms that are vectorizable and can be run in parallel, such as VMC and QMC [58], gain an advantage over I/O-bound CI methods. While VMC is valuable for computing the energy of a given trial function, one can also solve the Schrödinger equation directly using QMC [49]. QMC (Sec. 2.3), although closely related to VMC, has the critical difference of being able to go *beyond* the trial function

energy and obtain the exact ground-state energy.

Three important problems must be addressed in implementing QMC. First, is the problem of evaluation of the Green's function [8,23,31]. The Green's function gives the probability of moving an electron from one position to another, and is not known analytically for the general case. In Sec. 2.4, the short-time approximation to the Green's function is introduced, which is exact in the limit of zero step size [8]. Algorithms using this or similar Green's functions are called diffusion quantum Monte Carlo (DQMC), because of the Gaussian distribution obtained. Other more exact, but more complex, algorithms are also discussed.

The second problem is that of variance reduction. Since QMC is a statistical method, the energy obtained is known only within a given statistical precision, determined by the standard deviation of the mean. As computation time is increased, the statistical precision decreases, but only as the square root of the time[59]. Thus, to reduce the statistical error by an order of magnitude require two orders of magnitude more computation. Rather than computing longer, one can reduce the fluctuations in the simulation method itself. This is done in Sec. 2.5 using the Monte Carlo technique of *importance sampling*.

The final problem is in the treatment of excited states. The simple QMC algorithm will produce only the ground state of the Hamiltonian. The ground state, however, is totally symmetric, both spacially and with respect to exchange of particles (the Bose state). But since electrons are antisymmetric upon exchange, the Fermi state, which lies higher in energy, is required. In Sec. 4.6 methods for obtaining the Fermi state are discussed with emphasis on the *fixed-node* method [8], wherein the nodes of a given trial function are imposed upon the QMC solution to project out Bose solutions.

At this point the ground work of the FNDQMC method has been laid and an algorithm implementing the FNDQMC method is given in Sec. 2.7. Following this several special topics in QMC are discussed: choice of trial function (2.8), the time correlation formalism (2.9), computation of excited state energies (2.10), evaluation of operators

which do not commute with the Hamiltonian (2.11), and the correlated-sampling method for obtaining energy differences (2.12). The final section, 2.13, gives conclusions and identifies future directions for QMC.

2.2 Variational Monte Carlo

The variational method [2], introduced in Sec. 1.1, is a simple and useful way to obtain energies and properties for atoms and molecules from the Schrödinger equation. But for Eq. 1.3 to be analytically integrable Ψ_T must be of a fairly simple form, usually involving only powers and exponentials of Cartesian coordinates and electron-nuclear distances [41]. Of particular difficulty is the integration of trial functions involving the inter-electron distances, r_{ij}, which were shown by Hylleraas [45] to be of prime importance in the Helium wave function. Omission of the r_{ij} term in standard CI trial functions is responsible for their relatively slow convergence to the exact solution [44].

In order to use general forms of Ψ_T some numeric integration scheme must be used. Consider an integral of the form

$$I = \int_a^b g(\mathbf{x}) d\mathbf{x}, \tag{2.1}$$

where $\mathbf{x}$ is the coordinate vector, and g is an arbitrary function of $\mathbf{x}$. By the mean value theorem

$$I = \lim_{N \to \infty} I'_N \tag{2.2}$$

$$I'_N = \frac{b-a}{N} \sum_{i=1}^{N} g(\mathbf{x}_i) . \tag{2.3}$$

For an infinite sum the sample points $\{\mathbf{x}_i\}$ can be chosen to cover all of space, but for a finite sum the value of I'_N will strongly depend on the choice of the $\{\mathbf{x}_i\}$.

A simple choice for $\{\mathbf{x}_i\}$ is an evenly spaced grid in each direction of $\mathbf{x}$ which spans the non-zero realm of $g(\mathbf{x})$. However, as the dimensionality of $\mathbf{x}$ increases, all such grid

methods become costly [60]. Instead, a Monte Carlo method can be used. Specifically, Eq. 2.3 can be used with $\{\mathbf{x}_i\}$ being a set of uniformly distributed random numbers on the interval $[a,b]$. This method will be useless for infinite intervals, or when $g(\mathbf{x})$ is sharply peaked since the probability of choosing an $\mathbf{x}$ in the important (non-zero) range of $g(\mathbf{x})$ will be nearly zero. However, one need not sample $\mathbf{x}$ uniformly, instead *importance sampling* [54,55] can be used whereby one samples $\mathbf{x}$ from a normalized probability density function $P(\mathbf{x})$ chosen to emphasize the "important" regions of space for $g(\mathbf{x})$. This is accomplished by multiplying and dividing by $P(\mathbf{x})$ in the integral

$$I = \int \frac{g(\mathbf{x})}{P(\mathbf{x})} P(\mathbf{x}) d\mathbf{x} \equiv \int f(\mathbf{x}) P(\mathbf{x}) d\mathbf{x} \tag{2.4}$$

where the range of $P(\mathbf{x})$ is assumed to be $[a,b]$. Change variables such that $P(\mathbf{x})d\mathbf{x} = d\mathbf{y}$, *i.e.*

$$\mathbf{y}(\mathbf{x}) = \int_0^{\mathbf{x}} d\mathbf{z}\, P(\mathbf{z}) \tag{2.5}$$

leaving the integral,

$$I = \int f(\mathbf{x}(\mathbf{y})) d\mathbf{y} \sim \frac{1}{N} \sum_i^N f(\mathbf{x}_i) \tag{2.6}$$

where the $\{\mathbf{x}_i\}$ are drawn from the probability distribution function P.

One key feature of the Monte Carlo method is that along with the estimate of the integral, one obtains an estimate of the statistical precision, ε, given by,

$$\varepsilon^2 = (\langle f^2 \rangle_P - \langle f \rangle_P^2)/(N-1) \tag{2.7}$$

where ε is the standard deviation of the mean. The notation $\langle f \rangle_P$ indicates the average value of f with respect to the distribution P. The most efficient sampling occurs when P mimics g precisely, *i.e.*, when $f = 1$ and the $\{\mathbf{x}_i\}$ are distributed such that the number of points per unit volume is proportional to $|g|$. For simple choices of g this can be accomplished by inverting Eq. 2.5. To generate random numbers distributed for a general P, the algorithm of Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller [48]

$(M(RT)^2)$ can be used.

The Metropolis algorithm is based on a *random walk*, rather than a set of randomly chosen points. The goal is to generate a sequence of vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ which are distributed as $P(\mathbf{x})$. The first step is to randomly choose an initial vector, $\mathbf{x}_0$. Next, each element of $\mathbf{x}_0$ is randomly displaced (either one at a time or all at once) to create a trial vector $\mathbf{x}'$. If the value of $P(\mathbf{x}')$ is greater than or equal to $P(\mathbf{x}_0)$ then the move is accepted and $\mathbf{x}_1$ is set to $\mathbf{x}'$. Otherwise a random number between 0 and 1 is generated, and if $P(\mathbf{x}')/P(\mathbf{x}_0)$ is greater than this number then the move is accepted and $\mathbf{x}_1$ is set to $\mathbf{x}'$ else $\mathbf{x}_1$ is set to $\mathbf{x}_0$. The next generation, $\mathbf{x}_2$, is then obtained from $\mathbf{x}_1$ in the same manner, and so forth. An additional feature of the $M(RT)^2$ algorithm is that for non-normalized probability distribution functions the random walk produces points sampled from $P(\mathbf{x})/\int d\mathbf{x} P(\mathbf{x})$, not $P(\mathbf{x})$. For evaluation of expectation values of the form in Eq. 1.3, this is ideal since the normalization of the trial function need not be explicitly evaluated. Unfortunately, if the value of the integral alone (without normalization) is desired, then $\int d\mathbf{x} P(\mathbf{x})$ must be known.

While the random walk is taking place running averages are kept of $f(\mathbf{x}_i) \equiv g/P$, and any other quantities of interest. The optimum size of the random displacement used can be determined empirically from the resulting *acceptance ratio*. The acceptance ratio is the number of moves accepted by the above procedure divided by the total moves attempted. The most efficient sampling occurs when the acceptance ratio is in the 50%-70% range. Initially one must wait for the random walk to equilibrate, *i.e.*, the point when transient behavior caused by choosing a random initial vector has died out (Figure 2.1). Once this point has been reached then averages are kept and the statistical precision, ε, is estimated. The random walk continues until ε is as small as desired. To take advantage of supercomputer vector architecture an ensemble of $\mathbf{x}$'s is maintained at each step of the walk. Thus, like operations can be performed on a large number of indepen-

dent numbers at one time. The effect of increasing the vector length (ensemble size) on a CRAY XMP is shown in Figure 2.2.

A problem arises is this algorithm when estimating the statistical precision [61]. Because each generation, $\mathbf{x}_i$, differs little from the previous generation, $\mathbf{x}_{i-1}$, the correlation between successive values of $g(\mathbf{x})$ will be high and the value of ε calculated by Eq. 2.7 will be too small. One way to overcome this is to average together large "blocks" of data, then calculate ε from the resulting block averages. If the blocks are large enough then the block averages will be de-correlated and ε will be a true measure of the precision of the integral. Another method for calculating ε is to keep the averages for each member of the ensemble separately and use Eq. 2.7 across the ensemble. Since the members of the ensemble are independent of each other an accurate ε will result. Regardless of which method is used to calculate ε, at least 20 to 50 independent averages (block or ensemble) are needed to obtain a reliable value. In Table 2.1 the statistical precision calculated using Eq. 2.7 over block averages is given as a function of block size. The calculated variance has a maximum at a sample size per block of 4000 configurations, with 50 blocks. This maximum is caused by two competing errors in computing the statistical precision: serial correlation is large in small blocks, and for the larger block sizes, the number of blocks is so small that Eq. 2.7 does not have enough data to accurately calculate ε. For a large calculation the second process can be effectively eliminated so that the variance will reach a maximum asymptotically.

Returning to the original variational integral, Eq. 1.3 is rewritten as,

$$E[\Psi_T] = \int d\mathbf{R}\Psi_T^2\left[\frac{\mathbf{H}\Psi_T}{\Psi_T}\right] \tag{2.8}$$

so that Ψ_T^2 is identified as the probability distribution, and the value to be averaged is the *local energy* of Ψ_T, $E_L = \mathbf{H}\Psi_T/\Psi_T$. Note the important property of E_L that if Ψ_T is the exact eigenfunction of state I then $E_L(\mathbf{R}) = E_I$, which is a constant. Thus in the Monte Carlo evaluation of the variational integral the quality of Ψ_T can be judged both by the

energy $E[\Psi_T]$ and by the statistical precision $\varepsilon[\Psi_T]$.

The variational Monte Carlo [56] evaluation of the energy of Ψ_T proceeds as follows. An ensemble of electronic configurations is generated randomly, the positions of the electrons given by $R_k^{(0)}$. Here the subscript denotes the ensemble member and the superscript the generation. Each electron is displaced a distance ΔR in a random direction to form $\mathbf{R}_k'$. The values $\Psi_T^2(\mathbf{R}_k')$ and $\Psi_T^2(\mathbf{R}_k^{(0)})$ are compared as described above and the move is accepted or rejected. An initial equilibration sequence is generated; averages are kept only for the purpose of determining when equilibrium has been reached. Then a longer walk is generated and the average of the energy is evaluated from,

$$E^{(N)}[\Psi_T] = \frac{1}{N}\sum_{i}^{N}\sum_{k} E_L(\mathbf{R}_k^{(i)}) . \tag{2.9}$$

The statistical precision is evaluated by either method (blocking or ensemble) and the walk is either continued or terminated as desired. In addition to the energy, other local operators, such as the dipole and quadrupole moments, can be evaluated and averaged. The one particle density, $\rho(\mathbf{R})$, can be measured by counting the number of walkers which occupy a given volume in space. Thus, for an atom, the radial density can be obtained by putting a grid in the radial direction, then counting the number of walkers in a given grid element over the course of the walk. In this case, however, one obtains $r^2R^2(r)$, $R(r)$ being the radial part of the wavefunction and r the radial coordinate, due to the volume element of the spherical polar coordinates.

In the above VMC algorithm the function Ψ_T was used in two ways: a the probability distribution Ψ_T^2 of the random walk, and as the energy estimator, $\mathbf{H}\Psi_T/\Psi_T$. In fact the same function need not be used for both purposes. Suppose instead one wished to use a function, Ψ_G, to guide the walk, while the energy was sampled from Ψ_T. This is accomplished by multiplying and dividing Eq. 2.8 by Ψ_G^2, and including the proper normalization,

$$E[\Psi_T] = \int d\mathbf{R}\Psi_G^2 \left[\frac{\mathbf{H}\Psi_T}{\Psi_G}\right]\left[\frac{\Psi_T}{\Psi_G}\right] / \int d\mathbf{R}\Psi_G^2 \left[\frac{\Psi_T}{\Psi_G}\right]^2 \tag{2.10}$$

By keeping track of the weight Ψ_T/Ψ_G, the value of the energy of Ψ_T can be determined while sampling from the distribution Ψ_G^2. This is useful in two instances. First, when Ψ_T is very time consuming to compute, this procedure can improve the efficiency of the VMC calculation [16]. Second, when one desires not just the energy of a single trial function, but of a set of functions Ψ_{T_n} [49]. The weight for each trial function is,

$$W_n(\mathbf{R}) = \frac{\Psi_{T_n}(\mathbf{R})}{\Psi_G}. \tag{2.11}$$

In this case Ψ_G can be any one of these functions, or any other appropriate function, and the energies are given by,

$$E[\Psi_{T_n}] = \frac{\frac{1}{N}\sum_i^N \sum_k \left[\frac{\mathbf{H}\Psi_{T_n}}{\Psi_G}\right] W_n}{\frac{1}{N}\sum_i^N \sum_k W_n^2}. \tag{2.12}$$

This not only allows the calculation of many energies at the same time, but more importantly, since the energies are calculated using the same random numbers the fluctuations in the local energies of the Ψ_{T_n} can cancel to a great extent. Hence, the energies are *correlated*, and the energy differences will have an improved statistical precision over what would be obtained with separate calculations. This method of obtaining energy differences is known as differential Monte Carlo, or correlated sampling. Note, however, that the trial functions and the guiding function must all be quite similar, otherwise the ratio of the trial to guiding functions can diverge [31].

2.3 Quantum Monte Carlo

In the above discussion Monte Carlo methods were used to evaluate integrals. However, Monte Carlo can also be used to solve equations directly [55]. In the Born-Oppenheimer approximation the electronic time-dependent Schrödinger equation in

atomic units can be written

$$-\frac{1}{i}\frac{\partial \Psi}{\partial t} = \mathbf{T}\Psi + (\mathbf{V}-E_T)\Psi \tag{2.13}$$

where **T** is the kinetic energy operator,

$$\mathbf{T} = -\tfrac{1}{2}\sum_{i}^{N_{el}} \nabla_i^2 , \tag{2.14}$$

V is the potential energy,

$$\mathbf{V} = \sum_{A}^{N_{nuc}}\sum_{B}^{N_{nuc}} \frac{Z_A Z_B}{R_{AB}} - \sum_{A}^{N_{nuc}}\sum_{i}^{N_{el}} \frac{Z_A}{r_{iA}} + \sum_{i}^{N_{el}}\sum_{j<i}^{N_{el}} \frac{1}{r_{ij}} , \tag{2.15}$$

and E_T is an arbitrary energy offset (Ψ is independent of E_T). Here r and R are inter-particle distances, i and j represent electrons, A and B represent nuclei, and Z is the atomic number. The time dependence can be solved explicitly, yielding

$$\Psi(\mathbf{R},t) = c_0\Phi_0 e^{-i(E_0-E_T)t} + \sum_{I=1}^{\infty} c_I \Phi_I(\mathbf{R})\, e^{-i(E_I-E_T)t} \tag{2.16}$$

where Φ_I and E_I are the eigenstates and eigenvalues of the time-*independent* Schrödinger equation (Eq. 1.1) and Φ_0 designates the ground state.

For Monte Carlo solution of this equation it needs to be related to a process which can be simulated using random numbers. This is done by exploiting the similarity of the imaginary-time Schrödinger equation to the classical diffusion equation, first discussed by Metropolis and Ulam [62] who attributed the observation to Fermi. Writing the time-dependent Schrödinger equation in imaginary time, $\tau \equiv it$, yields,

$$-\frac{\partial \Psi}{\partial \tau} = \mathbf{T}\Psi + (\mathbf{V}-E_T)\Psi \tag{2.17}$$

and

$$\Psi(\mathbf{R},\tau) = c_0\Phi_0 e^{-(E_0-E_T)\tau} + \sum_{I=1}^{\infty} c_I \Phi_I(\mathbf{R})\, e^{-(E_I-E_T)\tau} . \tag{2.18}$$

Eq. 2.17 can be viewed as a sum of two equations: diffusion, and first order rate [16].

If only the kinetic energy is considered, one obtains the familiar diffusion equation,

$\frac{\partial\Psi}{\partial\tau}=-T\Psi$. This equation represents Brownian motion and can be simulated by a random "drunkards" walk of particles through configuration space (the position space of the electrons). On the other hand, if only the potential is considered, then a first order rate equation results, $\frac{\partial\Psi}{\partial\tau}=-(V-E_T)\Psi$, representing branching processes such as radioactive decay, which can be simulated by particle births and deaths. Therefore, the entire equation can be simulated by an ensemble of particles undergoing random motion (based on the kinetic energy) and births and deaths (based on the potential energy). The QMC solution $\Psi(\mathbf{R},\tau)$ is interpreted as the density of walkers (not to be confused with the electronic density, which is still $\Psi^2(\mathbf{R},\tau)$).

The benefit of transforming to imaginary time is that not only is the above analogy possible, but also that Eq. 2.18 exhibits exponential decay in terms of τ, as opposed to the oscillatory behavior of Eq. 2.16, so that the solution decays at large τ to the ground state. This is most easily seen by setting E_T to E_0 in Eq. 2.18 so that the time dependence of the ground state term vanishes, while contributions from all higher energy states decay exponentially with a half life of $(\ln 2/(E_1-E_0))$. Although, if one explicitly requires the solution $\Psi(\mathbf{R},\tau)$ to be orthogonal to Φ_0 (by boundary conditions), then $c_0=0$ in Eq. 2.18, and the first excited state, Φ_1, will result.

The precise manner in which the walkers diffuse and branch is given by the Green's function of the system. The Green's function is the transition probability of moving from $\mathbf{R}$ to $\mathbf{R}'$ in a time τ. Unfortunately, except for certain model systems, the exact Green's function will not be known analytically. Methods for approximating the Green's function and for sampling the exact Green's function are discussed in Sec. 2.4.

2.4 The Green's Function

The solution $\Psi(\mathbf{R},\tau)$ is propagated to large τ using the Green's function, $G(\mathbf{R}\rightarrow\mathbf{R}';\tau)$, of the system,

$$G(\mathbf{R}\rightarrow\mathbf{R}';\tau) = \langle \mathbf{R}' | e^{-\tau \mathbf{H}} | \mathbf{R} \rangle . \tag{2.19}$$

The Green's functions is the probability of finding a particle at position $\mathbf{R}'$ at time τ, given that a particle exists initially at $\mathbf{R}$. Thus the wavefunction at point $\mathbf{R}'$ is related to the wavefunction globally by,

$$\Psi(\mathbf{R}',\tau) = \int d\mathbf{R} G(\mathbf{R}\rightarrow\mathbf{R}';\tau)\Psi(\mathbf{R},0) \tag{2.20}$$

The Green's function is itself a solution to the Schrödinger equation with initial condition $G(\mathbf{R}\rightarrow\mathbf{R}';0) = \delta(\mathbf{R}-\mathbf{R}')$. This means that knowledge of the exact Green's function is equivalent to knowledge of the exact wavefunction.

One method which has had success achieving accurate results without explicit knowledge of the exact Green's function is the short-time approximation (STA) [8]. The short time Green's function is most easily derived from Eq. 2.19. Expanding $e^{-\tau \mathbf{H}}$ in terms of $\mathbf{T}$ and $\mathbf{V}$, Eq. 2.19 becomes

$$\begin{aligned} \langle \mathbf{R}' | e^{-\tau \mathbf{H}} | \mathbf{R} \rangle &= \langle \mathbf{R}' | e^{-\tau(\mathbf{T}+\mathbf{V})} | \mathbf{R} \rangle \\ &= \langle \mathbf{R}' | e^{-\tau \mathbf{T}} e^{-\tau \mathbf{V}} | \mathbf{R} \rangle + O(\tau^2) \end{aligned} \tag{2.21}$$

By neglecting terms order τ^2 the STA Green's function is then

$$G(\mathbf{R}\rightarrow\mathbf{R}';\tau)_{STA} = (4\pi\tau)^{-3N/2} e^{-(\mathbf{R}'-\mathbf{R})^2/4\tau} e^{-(\mathbf{V}(\mathbf{R})-E_T)\tau} \tag{2.22}$$

the first term is derived from the kinetic energy ($e^{-\tau \mathbf{T}}$) describing free particle diffusion, and the second term describes branching ($e^{-\tau(\mathbf{V}-E_T)}$). E_T is adjusted so that the average of the branching term is 1 during the walk. By comparing Eq. 2.22 and Eq. 2.18 it is seen that this is achieved when E_T is equal to E_0. This suggests that the energy can be estimated based on the growth of the ensemble. One way to do this is to note that if the STA green's function were exact then the intial population, N_i, would be related to the population after the time τ, N_f, by

$$N_f = N_i e^{-(E_0-E_T)\tau} . \tag{2.23}$$

Since the Green's function is not exact, the quantity $E_g(\tau)$ is used in place of E_0, which can be computed by inverting Eq. 2.23,

$$E_g(\tau) = \ln(N_i/N_f)/\tau + E_T \tag{2.24}$$

which is known as the *growth estimator* of the energy. Because of the time-step bias of the Green's function, the average of the branching will be 1 when E_T is set to the value of the growth estimator for the given τ, rather than E_0.

Since this Green's function is only exact at $\tau = 0$, propagation to large τ is divided into a large number of small time steps $\delta\tau$ to form a random walk. There will be a range of values near zero in which the linear term dominates all higher order contributions. Hence, computations at several small $\delta\tau$ are carried out then extrapolated to $\delta\tau = 0$ using a linear (or possibly quadratic) fit. A good rule of thumb is that the acceptance ratio (see Sec. 2.2) should not be less than 90% and may need to be greater than 99% to be within the linear regime (Figure 2.3).

A Monte Carlo simulation using the STA Green's function, then, consists of an ensemble of electronic configurations undergoing a random displacement by a spherical gaussian random number with a mean of zero and a standard deviation of $\delta\tau$ (diffusion), with the population increasing or decreasing by $e^{-(\mathbf{V(R)}-E_T)\tau}$ (branching). Several runs at different time steps are needed to estimate the value of the energy (and any other quantities) at $\delta\tau$=0.

Although the STA Green's function is the one used in the FNDQMC algorithm which is being discussed here, there are alternative methods which do not make the short-time approximation: the domain Green's function methods of Kalos, *et al.* [49], and Moskowitz, *et al.* [23], and the exact Green's function method of Ceperley, *et al.* [31]. In these methods, the exact Green's function is sampled by an auxiliary walk in $\mathbf{R}''$ which integrates for example,

$$G(\mathbf{R},\mathbf{R}') = G_T(\mathbf{R},\mathbf{R}') + \int_S d\mathbf{R}''\, G(\mathbf{R},\mathbf{R}')[-\hat{n}\cdot\nabla G(\mathbf{R}',\mathbf{R}'') \tag{2.25}$$

$$+ \int_v d\mathbf{R}''\, G(\mathbf{R},\mathbf{R}'')[U - V(\mathbf{R}'')]G_T(\mathbf{R}'',\mathbf{R}'),$$

as in Ref. 23. U is a constant energy offset, V is the potential, S is the surface of a given "domain" in which the Green's function is being solved, and v is the volume of the domain. Another possibility used by Ceperley [31] is

$$G(\mathbf{R},\mathbf{R}';\tau) = G_T(\mathbf{R},\mathbf{R}';\tau) + \int_0^\tau dt \int d\mathbf{R}'' \, G(\mathbf{R},\mathbf{R}'';\tau-t)[-\mathbf{H}(\mathbf{R}'') - \frac{\partial}{\partial t}]G_T(\mathbf{R}'',\mathbf{R}';t) \, . \quad (2.26)$$

In both cases G_T is a simple guess at the Green's function. The motivation behind these two equations is very similar, the differences are due to the fact that in Eq. 2.25 the Green's function is obtained from $1/(H-E)$ while Eq. 2.26 uses $e^{-\tau\mathbf{H}}$. In this respect Eq. 2.26 is more closely related to DQMC than is Eq. 2.25. One important difference between the two equations is that Eq. 2.25 is contained within a small domain. In systems containing nodes, a domain is defined to be wholly within a nodal volume. This means that use of Eq. 2.25 leads to use of very small time steps, whereas Eq. 2.26 allows larger time steps to be used. While this gives Eq. 2.26 an advantage in sampling efficiency, it also means that the a walker can cross and recross a node (see Sec. 2.7 for a detailed description of the cross-recross error) leading to a time-step bias in the fixed-node walk. This is unfortunate since the purpose of using Eq. 2.26 was to eliminate the time step bias. This bias is eliminated only when the nodes are released, as is done in Ref. 31.

2.5 Importance Sampling

The importance sampling technique used in Sec. 2.2 can also be used in QMC to reduce the statistical fluctuations [8]. A trial function, Ψ_T, is introduced as a systematic bias to the random walk. The chosen Ψ_T is meant to represent the QMC solution $\Psi(\mathbf{R},\tau)$ at long time, and is usually a variational solution to the time independent Schrödinger equation. However, unlike VMC, in QMC Ψ_T will not determine the final energy (extrapolated to $\delta\tau=0$ for DQMC), only its statistical precision, since the Schrödinger equation is being solved here.

The random walk samples space with a new probability distribution $f(\mathbf{R},\tau)$ given by

$$f(\mathbf{R},\tau) = \Psi(\mathbf{R},\tau)\Psi_T(\mathbf{R}) \tag{2.27}$$

which will then decay to $\Phi_0\Psi_T$ at long imaginary time. The Schrödinger equation written in imaginary time with importance sampling[8] is

$$\partial f/\partial\tau = -\tfrac{1}{2}\nabla^2 f + \tfrac{1}{2}\nabla\cdot(\mathbf{F}_Q f) + (E_L - E_T)f\,. \tag{2.28}$$

The two new quantities, E_L, the local energy, and F_Q, the quantum force, are given by,

$$E_L(\mathbf{R}) = \mathbf{H}\Psi_T(\mathbf{R}) / \Psi_T(\mathbf{R}), \tag{2.29}$$

and

$$F_Q(\mathbf{R}) = 2\nabla\Psi_T(\mathbf{R}) / \Psi_T(\mathbf{R}). \tag{2.30}$$

The average value of the local energy is now the energy estimator, and the quantum force is a vector field or drift term pushing the electrons towards regions where Ψ_T is greatest. A very important property of the local energy is that as Ψ_T approaches an eigenstate of the Hamiltonian, $E_L(\mathbf{R})$ will become a constant and the variance tends to zero. In addition, if the electron-electron and electron-nuclear cusp conditions are met (Sec. 2.8) E_L will not diverge at the coulomb singularities.

To show that the introduction of Ψ_T does not change the average value of the final energy, write the average value of E_L as an integral over the probability distribution at long time, f_∞,

$$\langle E_L\rangle_{f_\infty} = \lim_{\tau\to\infty}\int d\mathbf{R}\, f(\mathbf{R})\left[\frac{\mathbf{H}\Psi_T}{\Psi_T}\right] / \int d\mathbf{R}\, f(\mathbf{R}) \tag{2.31}$$

$$= \int d\mathbf{R}\, \Phi_0\mathbf{H}\Psi_T \Big/ \int d\mathbf{R}\Phi_0\Psi_T\,.$$

The Hamiltonian can act to the left upon Φ_0 yielding $E_0\Phi_0$,

$$\langle E_L\rangle_f = \int d\mathbf{R}\, E_0\,\Phi_0\Psi_T \Big/ \int d\mathbf{R}\Phi_0\Psi_T \tag{2.32}$$

returning the ground state energy. Hence, importance sampling improves the efficiency of the QMC simulation by sampling "important" regions of space more frequently without biasing the result.

The short-time approximation Green's function for the function f(**R**)[8] is,

$$G(\mathbf{R}\rightarrow\mathbf{R}',\tau)_{STA} \equiv G_D\ G_B \tag{2.33}$$

$$G_D = (4\pi D\,\tau)^{-3N/2}\exp\{-[\mathbf{R}'-\mathbf{R}-D\,\tau F_Q(\mathbf{R})]^2/4D\,\tau\}$$

$$G_B = \exp[-\tau\{[E_L(\mathbf{R})+E_L(\mathbf{R}')]/2 - E_T\}]$$

As before the results are then extrapolated to zero time step [8] to provide an unbiased estimate of the energy. The full algorithm for conducting the fixed-node diffusion QMC walk is given in Sec. 2.7.

Note that if one does not perform the branching, *i.e.*, set G_B to 1, then the variational energy of Ψ_T will result. This is exactly the VMC algorithm described in Sec. 2.2 with the added drift term, and is in fact a somewhat more efficient algorithm. A critical difference between VMC and QMC, however, is that unlike VMC, in QMC expectation values of operators which do not commute with **H** (such as the dipole and quadrupole moments) are not given in a simple way [11]. This is because the walk samples from the mixed distribution $f\equiv\Phi_0\Psi_T$, rather than the required Φ_0^2 distribution. Sampling of Φ_0^2 can be accomplished with the more complicated *future walking* [11] algorithm (Sec. 2.11). These properties can also be approximated [11] by neglecting terms second order in the difference $\Phi_0-\Psi_T$ with,

$$<A> = 2<A>_{\Phi_0\Psi_T} - <A>_{\Psi_T^2} \tag{2.34}$$

where $<A>_{\Phi_0\Psi_T}$ and $<A>_{\Psi_T^2}$ are the average values of the operator A over the distributions $\Phi_0\Psi_T$ and Ψ_T^2 respectively.

2.6 Fermi statistics and the fixed-node Approximation

As stated in the introduction (Sec. 2.1), care must be taken to obtain a properly antisymmetric solution in QMC. In fact, there are two requirements on $f(\mathbf{R},\tau)$. First, since $f(\mathbf{R},\tau)$ represents the density of walkers in the Monte Carlo simulation, it must be of the same sign everywhere (the value of the sign being arbitrary). On the other hand, because electrons are Fermions the QMC solution, Ψ, must change sign with respect to exchange of electrons. Since the lowest energy eigenstate of the Hamiltonian is symmetric, the

above-described procedure will produce a Bose state, not the desired Fermi state. One way to satisfy both requirements is the fixed-node approximation [8], wherein the Monte Carlo solution is forced to be antisymmetric by imposing the nodes of the importance function Ψ_T onto $\Psi(\mathbf{R},\tau)$ by either deleting [8] or rejecting [63] any walker which attempts to cross a node of Ψ_T. The energy obtained by this procedure will be exact only if the nodes of Ψ_T correspond to the exact nodes. Otherwise, the Schrödinger equation is solved exactly within the imposed nodal boundaries to obtain the fixed-node energy, $E_{FN}[\Psi_T]$, and wavefunction, $\Phi_{FN}[\Psi_T]$. The Monte Carlo distribution is now $f \equiv \Psi_T \Phi_{FN}$ simultaneously insuring that the QMC solution will be both antisymmetric, and of the same sign everywhere. The only modification to the diffusion QMC method above is the additional step to check if Ψ_T changes sign over the course of a move.

When Ψ_T is chosen to describe the nodes of the Fermi ground state, the fixed-node energy can be shown to be a variational upper bound to the true ground-state energy [8]. Excited states can be obtained through the fixed-node approximation by choosing Ψ_T with the proper nodes [13], but the energy is variational with respect to the true excited state energy only in cases where orthogonality with the lower states can be obtained exactly, typically by symmetry considerations.

Cases can occur where an electron crosses two nodes, or the same node twice during a single time step. This has the effect of lowering the energy since the nodes are not being properly enforced. This is the so-called cross-recross error [64]. If one used a sufficiently small time step this move would be rejected. Therefore the cross-recross error becomes part of the overall time step bias and will not affect the extrapolated energy.

The extrapolated fixed-node DQMC energy, then, only depends on the positions of the nodes of Ψ_T and not on the value of the trial function elsewhere, although the overall quality of Ψ_T greatly affects the statistical precision of the energy estimate. Since the electrons have a low probability of being in the region of a node, one can expect the

energy to be weakly dependent on the nodes. At least for first-row atoms, this expectation has been borne out, with fixed-node energies typically returning 90-100% of the correlation energy. Tables 2.2 and 2.3 summarize a number of FNDQMC calculations.

It should also be noted that one can also "release" the nodes [27,31] through the method of transient estimators. This is done by by maintaining two separate ensembles designated "positive" and "negative". The energy of the Fermi state is given by the difference between the energies of the two distributions. The reason this method is called transient is that the variance diverges at long time [31] instead of decreasing as it does in FNDQMC. The amount of sampling time available after equilibrium has been reached and before the variance diverges is directly related to the energy gap between the Bose and Fermi ground states. So while this method works well for light atom systems, convergence is difficult even for systems containing atoms as small as oxygen [31,65].

2.7 Fixed-Node Diffusion QMC Algorithm

Equation 2.28 is simulated by a random walk of an ensemble of $3N_{el}$ dimensional walkers representing the configuration of the N_{el} electrons. Each walker is allowed to drift, diffuse, and branch in such a way as to simulate Eq. 2.28, over a time interval $\delta\tau$. After diffusing for a sufficiently long time the ensemble will be distributed according to an equilibrium distribution given by a product of Ψ_T and the ground-state fixed-node eigenfunction, Φ_{FN}, the latter having the symmetry and nodes of Ψ_T. The local energy of Ψ_T is computed for a large number of configurations taken from this distribution, and averaged to give a Monte Carlo estimate of the ground-state eigenvalue, along with an estimate of the statistical precision.

The algorithm used here is similar to that described in Ref. 8. The steps are as follows:

(0) Choose Ψ_T

(1) Initialize an ensemble of configurations (indexed by m) distributed with probability density f(R,0). This is either chosen at random, taken from a VMC run with Ψ_T, or from a previous QMC run.

(2) Let an electron diffuse for one time step in each configuration. If this move causes the configuration to cross a node, i.e., if the value of Ψ_T changes sign, the move is rejected.

(3) After an electron is moved, accept the move with probability,

$$A(\mathbf{R}\rightarrow\mathbf{R}',\delta\tau) \equiv \min\left[1, \frac{|\Psi_T(\mathbf{R}')|^2 G(\mathbf{R}'\rightarrow\mathbf{R},\delta\tau)}{|\Psi_T(\mathbf{R})|^2 G(\mathbf{R}\rightarrow\mathbf{R}',\delta\tau)}\right] \tag{2.35}$$

where G is the short-time approximation to the Green's function; and R and R′ specify the initial and final configurations respectively. This step is performed so that detailed balance is maintained. Calculate the local energy and other quantities of interest for this electron in each configuration. Select the next electron and return to step (2).

(4) After all electrons in the ensemble have been moved once, increment the time associated with the ensemble by $\delta\tau$.

(5) Determine the branching probability for each configuration, m,

$$M_m = int[A_m + \zeta]$$

where

$$A_m = \exp\left\{-\delta\tau_a\left([(E_L(\mathbf{R}) + E_L(\mathbf{R}')]/2 - E_T\right)\right\} \tag{2.36}$$

and ζ is a random number between 0 and 1. Here $\delta\tau_a$ is the effective time-step, $\delta\tau_a = \langle r^2_{accepted}\rangle / \langle r^2_{total}\rangle$ where $\langle r^2_{accepted}\rangle$ and $\langle r^2_{total}\rangle$ are the mean-squared distance the electrons moved that was accepted and proposed. $E_L = H\Psi_T/\Psi_T$ is the

local energy, and E_T is the trial energy. Note that on average $<M_m>=A_m$.

(6) For each configuration weight E_L and other quantities of interest by A_m. If $M_m = 0$, the m^{th} configuration is eliminated, otherwise, if M_m is 2 or greater, place M_m-1 copies of this configuration in the ensemble.

(7) Repeat (2) thru (6) until the ensemble has reached a target time.

(8) Calculate the weighted mean

$$<E_L> = \sum_m E_L(m) A_m \; / \; \sum_m A_m \tag{2.37}$$

as an estimator of the ground state energy. This average is the expectation value of E_L sampled from the distribution $f_\infty \equiv \Psi_T \hat{\Phi}_0$.

(9) Use the cumulative average of the growth estimator, $<E_g>$, to update the trial energy E_T.

(10) Reset the time counter for the ensemble to 0. This completes one "block".

(11) Repeat steps (2) thru (10) until there is no systematic trend reflected in the single block energies. At this point equilibrium has been reached.

(12) Start a new run using the distribution from (11) and repeat (2) through (10) until the statistical precision in $<E_L>$ has reached the desired level.

The main difference between this algorithm and the one described in Ref. 1 is that each step is completed for the entire ensemble instead of completing (2) thru (6) for each configuration. This change makes it easier to take advantage of the vector capabilities of many supercomputers. The above algorithm has been implemented in the FORTRAN computer program QuantuMagiC. A description of the input to this program is given in Appendix A.

2.8 The Trial Function

The trial function, Ψ_T, plays a central role in both VMC and QMC. The difference between the two methods is that in VMC the energy depends solely on Ψ_T, whereas in FNDQMC the energy only depends on the nodes of Ψ_T, but the statistical precision (and thus efficiency) depends on the global shape of Ψ_T. In addition, when the STA Green's function is used, the trial function will determine the magnitude of the time-step bias. Unfortunately, so far no method for systematically optimizing the nodes alone has been discovered. Therefore, overall improvement of Ψ_T is of concern to both VMC and QMC approaches.

The logical starting point in the search for a suitable trial function is to use an SCF function. The problem with using an SCF function is that the electron interacts only with the average distribution of the other electrons. In the CI form of the trial function, it is observed that the exact wavefunction can be described using an infinite sum of such SCF functions, each describing a different electronic state. While this is a convenient form for the evaluation of analytic integrals, and is capable of treating electronic correlation to arbitrary accuracy, as was stated in Sec. 2.2, the expansion is slowly convergent because the inter-electronic coordinate, r_{ij}, is not included explicitly.

In 1928, Hylleraas [45] showed that a trial function which includes r_{ij} converges rapidly. The reason why CI functions are more popular than Hylleraas functions lies in the relative difficulty in construction and integration of the Hylleraas-type trial functions. The VMC procedure, however, is ideally suited for evaluation of the energy of such functions, although the construction of these functions remains a problem. An alternative procedure which has met with sucess in QMC is a combination of elements from both SCF and explicitly correlated functions, referred to as the "product" trial function here [66].

The product function, as the name suggests, is a product of a SCF or CI function with a electron-electron correlation function (EECF) which explicitly contains factors of r_{ij}. Instead of using the Hylleraas form, a pseudopotential form

$$S = e^{U} \tag{2.38}$$

is employed. This is done so that S introduces no new nodes into Ψ_T, and will have no effect on the fixed-node energy (although such functions can affect the time-step bias). This product form is taken with the expectation that the HF or CI function already describes most of the electronic behavior correctly (in particular placement of the nodes), and that only the regions where the electrons come together need modification. This can be illustrated by looking at one property of the true wavefunction which CI trial functions can only treat with an infinite expansion — the electron-electron cusp condition.

The electron-electron cusp condition is the result of canceling the $1/r_{ij}$ singularity in the potential energy with corresponding terms in the local kinetic energy. Two cases must be considered separately: when the two electrons are of opposite spin, and $\Psi_T(r_{ij}=0)=!0$, and when the two electrons are of the same spin, and $\Psi_T(r_{ij}=0)=0$. For the case of opposite spin electrons, the local kinetic energy of electron i in polar coordinates centered on electron j is,

$$-\tfrac{1}{2}\Psi_T^{-1}\nabla_i^2\Psi_T = -\tfrac{1}{2}\Psi_T^{-1}\left(\frac{2}{r_{ij}}\frac{\partial}{\partial r_{ij}} + \frac{\partial^2}{\partial r_{ij}^2} + \textit{angular terms}\right)\Psi_T \tag{2.39}$$

only one term explicitly depends on $1/r_{ij}$. Adding these terms for ∇_i^2 and ∇_j^2 and setting equal to the potential, one obtains,

$$\tfrac{1}{2}\left(\frac{4}{r_{ij}}\frac{\partial}{\partial r_{ij}}\Psi_T\right) = \frac{\Psi_T}{r_{ij}}\,. \tag{2.40}$$

Evaluating at $r_{ij}=0$ yields the commonly quoted [8] expression for the electron-electron cusp condition for opposite spins,

$$\Psi_T^{-1}\frac{\partial}{\partial r_{ij}}\Psi_T \Big|_{r_{ij}=0} = \tfrac{1}{2}\,. \tag{2.41}$$

Thus, if Ψ_T satisfies Eq. 2.41, then the total local energy will be non-singular when two electrons of opposite spin meet.

For two electrons of like spin, care must be taken when dividing by $\Psi_T(r_{ij}=0)$. First define $r_>$ and $r_<$ to be the greater and lesser of the two electronic vectors. Expanding $r_>$ in a Taylor series around $r_<$ yields $r_> = r_< + r_{ij} + \cdots$. For simplicity consider a 2×2 deter-

minant $\det|\phi_i(i)\phi_j(j)|$. Each molecular orbital, ϕ, depending on $r_>$ can also be expanded around $r_<$,

$$\phi_k(r_>) = \phi_k(r_<) + r_{ij}\frac{\partial\phi_k}{\partial r_<} + \cdots \tag{2.42}$$

the expression for the determinant is then

$$\det|\phi_i(i)\phi_j(j)| = \phi_i(r_<)\left[\phi_j(r_<)+r_{ij}\frac{\partial\phi_j}{\partial r_<}\right] - \phi_j(r_<)\left[\phi_i(r_<)+r_{ij}\frac{\partial\phi_i}{\partial r_<}\right] \tag{2.43}$$

$$= r_{ij}\left[\phi_i(r_<)\frac{\partial\phi_j}{\partial r_<} - \phi_j(r_<)\frac{\partial\phi_i}{\partial r_<}\right]$$

Based on the 2×2 case in general define $\Psi_T = r_{ij}F$ for small r_{ij}. Then the radial terms from the local kinetic energy are

$$-\tfrac{1}{2}\Psi_T^{-1}\left(\frac{2}{r_{ij}}\frac{\partial}{\partial r_{ij}} + \frac{\partial^2}{\partial r_{ij}^2}\right)\Psi_T = -\tfrac{1}{2}F^{-1}\left(\frac{4}{r_{ij}}\frac{\partial F}{\partial r_{ij}} + \frac{2F}{r_{ij}^2} + r_{ij}\frac{\partial^2}{\partial r_{ij}^2}\right) \tag{2.44}$$

Collecting terms depending on r_{ij}^{-1} for both ∇_i^2 and ∇_j^2, the equation analogous to Eq. 2.40 is

$$\tfrac{1}{2}\left(\frac{8}{r_{ij}}\frac{\partial F}{\partial r_{ij}}\right) = \frac{F}{r_{ij}}\,. \tag{2.45}$$

Note that Eq. 2.45 depends on F, not Ψ_T. Breaking F into a determinantal part D and and correlation function S, the derivative of D with respect to r_{ij} is zero by construction, and F in Eq. 2.45 can be replaced by S. In Eq. 2.41 Ψ_T can also be replaced by S since the derivative of the determinant is likewise zero. Therefore a unified equation for both cases can be written as,

$$S^{-1}\frac{\partial}{\partial r_{ij}}S\,\Big|_{r_{ij}=0} = a(s_i,s_j)\,, \tag{4.46}$$

where s_i and s_j are the spins of the electrons and $a(\alpha,\alpha) = a(\beta,\beta) = \frac{1}{4}$, $a(\alpha,\beta) = a(\beta,\alpha) = \frac{1}{2}$. In practice, however, it is common to use the value of ½ for both cases[8]; no appreciable increase in variance is observed.

A common form for U in Eq. 2.38, is the Pade-Jastrow function

$$U = \sum_{i>j} \frac{a_1 r_{ij} + a_2 r_{ij}^2 \cdots}{1 + b_1 r_{ij} + b_2 r_{ij}^2 \cdots} \tag{2.47}$$

where the value of a_1 is fixed by the electron-electron cusp condition. This function has been used in a number of studies [8-12,18,19], usually only retaining the a_1 and b_1 terms. However, one recent study used a related form which involves a polynomial in r_i, r_j, and r_{ij}, with as many as 35 terms [67]. The main difficulty is the optimization of the a_n and b_n. Although this can be done (see Sect. 2.12) it is a non-linear problem, and so the fewer terms needed in the expansion the better. One form which cuts the number of terms in half while still retaining terms to all orders has recently been suggested by Sun, *et al.*,

$$U = -a_0 e^{(-a_1 r_{ij} + a_2 r_{ij}^2 \cdots)} \tag{2.48}$$

this function eliminates the terms in the denominator of Eq. 2.47, which are expected to play a minor role, and the terms a_0 and a_1 are related by the electron-electron cusp condition. This leaves only 2 adjustable parameters for quadratic and 3 for a cubic form of Eq. 2.48, as opposed to 3 and 5 parameters for the corresponding forms of Eq. 2.47.

One expects further improvements in Ψ_T for multi-electron systems if 3-body terms such as $|\mathbf{r}_{ij} \cdot \mathbf{r}_{ik}|$ are included. In the work of Sun, *et al.*, it was suggested that higher-order terms could be represented by making the a_n and b_n depend on the one-electron density $\rho(\mathbf{r})$. Owen, *et.al.* [68], have implemented this method using the linear form of Eq. 2.47 with some success.

An electron-nuclear cusp condition can be derived in an analogous manner to Eq. 2.41,

$$\Psi_T^{-1} \frac{\partial}{\partial r_{iA}} \Psi_T \Big|_{r_{iA}=0} = -Z_A \tag{2.49}$$

where Z_A is the nuclear charge of atom A. In this case, however, the determinant has a non-trivial derivative with respect to r_{iA} if 1S and 2S STF's are used in the atomic basis set. Additionally, the derivative of the determinant will depend on the positions of all the electrons, not just electron i of r_{iA}. This means that as long as $\partial D / \partial r_{iA}$ is non-zero then it will be very difficult to satisfy the electron-nuclear cusp condition exactly [69]. One

way around this is to not use 1S and 2S STF's, instead using GTF's. Unfortunately GTF's have very poor overall behavior which reduces the efficiency of the method. A simple approximate scheme is to satisfy the cusp on average, by first multiplying Eq. 2.49 by Ψ_T^2 and integrating over the positions of all the electrons but one,

$$\int d\mathbf{R}'\Psi_T \frac{\partial}{\partial r_{iA}}\Psi_T \Big|_{r_{iA}=0} = \int d\mathbf{R}'(-Z_A)\Psi_T^2 \tag{2.50}$$

where $d\mathbf{R}'$ indicates that one electron is retained outside the integration. The quantity on the right side of Eq. 2.50 is the one-electron density function, $\rho(\mathbf{r})$, multiplied by $-Z_A$. By the chain rule the left hand side of Eq. 2.45 becomes

$$\tfrac{1}{2}\frac{\partial}{\partial r_{iA}}\left[\int d\mathbf{R}'\Psi_T^2\right] \Big|_{r_{iA}=0} \tag{2.51}$$

which is simply $\frac{1}{2}\partial\rho/\partial r_{iA}$. Thus the average electron-nuclear cusp condition constrains the one-electron density to obey the relationship,

$$\tfrac{1}{2}\rho(\mathbf{r})^{-1}\frac{\partial\rho(\mathbf{r})}{\partial r_A}\Big|_{r_A=0} = -Z_A \tag{2.52}$$

where the electron index i has been dropped, since ρ is a one-electron function.

For the the electron-nuclear cusp, an electron-nuclear correlation function (ENCF) can be used, having either a Pade-Jastrow or the double exponential form (Eq. 2.48), to help satisfy the electron-nuclear cusp condition (substitute r_{iA} for r_{ij} and make a_1 for the Pade-Jastrow or a_0 for the double exponential, negative). The ENCF also has another important use. In most cases the HF density is a good representation of the true density, but the addition of an EECF causes a global expansion of the electron density. The ENCF then can be used to rescale the density back towards that of the HF function. The adverse effect of the EECF on the electronic density illustrates the need to optimize the product trial function with respect to all the parameters.

The efficiency of a computation depends in part on the amount of time needed to evaluate quantities necessary for the random walk, *e.g.*, the local energy and the quantum force. Practical considerations concerning the computation of Ψ_T, E_L, and $\mathbf{F}_Q$ are discussed separately in Appendix B. However, since computational cost increases linearly

with basis-set size and with the number of determinants, the efficiency of such expansions will depend on the trade-off between the increased cost and the variance reduction. In most previous computations, performed with single determinant and small basis set Ψ_T's, high accuracy fixed-node energies resulted, i.e. 90-100% of the correlation energy was obtained. Expansion of the basis set past DZ has, thus far, had a negligible effect on the fixed-node energy; however, some recent results on the F + H_2 reaction [19] suggest that larger basis sets could be useful in heavy atom systems.

Multi-determinant trial functions are necessary both in cases where the true wavefunction is not adequately described by an SCF function [70], and to describe excited states [13]. However, it should be emphasized here, that in comparison to more standard techniques, experience with QMC is limited. Increases in the number of determinants and basis set size may be of greater importance in future calculations.

2.9 Time correlation form of the DQMC equations

Instead of starting from the imaginary-time Schrödinger equation, one can express the above method in a somewhat more general form in terms of expectation values. Define,

$$N(\tau) = \langle \Psi_T | e^{-\tau \mathbf{H}} | \Psi_T \rangle \tag{2.53}$$

$$H(\tau) = \langle \Psi_T | \mathbf{H} e^{-\tau \mathbf{H}} | \Psi_T \rangle \tag{2.54}$$

Then the exact ground state energy is given by

$$E_0 = \lim_{\tau \to \infty} H(\tau) / N(\tau). \tag{2.55}$$

This can be related to the standard QMC method by noting that

$$\lim_{\tau \to \infty} e^{-\tau \mathbf{H}} \Psi_T = c_0 | \Phi_0 \rangle . \tag{2.56}$$

Thus,

$$\lim_{\tau \to \infty} H(\tau) = c_0 \langle \Phi_0 | \mathbf{H} | \Psi_T \rangle \tag{2.57}$$

and,

$$\lim_{\tau\to\infty} N(\tau) = c_0 \langle \Phi_0 | \Psi_T \rangle, \tag{2.58}$$

and the ratio $H(\tau)/N(\tau)$ is precisely the QMC mixed estimator of the ground state energy. However, the form of Eq. 2.53 and Eq. 2.54 suggest that rather than using the diffusion analogy of the imaginary-time Schrödinger equation (Eq. 2.28), an equivalent procedure is to sample the operators $e^{-\tau \mathbf{H}}$ and $\mathbf{H}e^{-\tau \mathbf{H}}$ during a VMC walk [71].

To accomplish this, first consider integration of $H(\tau)$ to $H(\tau+\delta\tau)$. Introduce two complete sets of position states into Eq. 2.54,

$$H(\tau+\delta\tau) = \int d\mathbf{R} \int d\mathbf{R}' \langle \Psi_T(0) | \mathbf{R}' \rangle \langle \mathbf{R}' | e^{-\delta\tau \mathbf{H}} | \mathbf{R} \rangle \langle \mathbf{R} | \mathbf{H} \Psi_T(\tau) \rangle \tag{2.59}$$

$$= \int d\mathbf{R} d\mathbf{R}' \Psi_T(\mathbf{R}',0) G(\mathbf{R} \to \mathbf{R}';\delta\tau) \mathbf{H} \Psi_T(\mathbf{R},\tau) .$$

If the integration over $\mathbf{R}$ was performed, the QMC solution would result (Eq. 2.20), which at long time decays to Φ_0. On the other hand, multiplying and dividing the integrand by a guiding function Ψ_G in both $\mathbf{R}$ and $\mathbf{R}'$ yields,

$$H(\tau+\delta\tau) = \int d\mathbf{R} d\mathbf{R}' \frac{\Psi_T(\mathbf{R}',0)}{\Psi_G(\mathbf{R}')} \left[\Psi_G(\mathbf{R}') G(\mathbf{R} \to \mathbf{R}';\delta\tau) \Psi_G(\mathbf{R})^{-1} \right] \frac{\mathbf{H}\Psi_T(\mathbf{R},\tau)}{\Psi_G(\mathbf{R})} \Psi_G(\mathbf{R})^2 \tag{2.60}$$

$$= \int d\mathbf{R} d\mathbf{R}' \Psi_G(\mathbf{R})^2 \frac{\Psi_T(\mathbf{R}',0)}{\Psi_G(\mathbf{R}')} G'(\mathbf{R} \to \mathbf{R}';\delta\tau) \frac{\mathbf{H}\Psi_T(\mathbf{R},\tau)}{\Psi_G(\mathbf{R})}$$

where $G'(\mathbf{R} \to \mathbf{R}';\delta\tau)$, the quantity in brackets, is the Green's function guided by Ψ_G. This is analogous to the introduction of the guiding function in Sec. 2.2 (Eq.'s 2.10, 2.11). The formula for $N(\tau+\delta\tau)$ follows in a similar manner from Eq. 2.53.

Using Eq. 2.60, evaluation of $H(\tau)$ and $N(\tau)$ for large τ takes place during the course of a VMC walk guided by Ψ_G. The quantities $\Psi_T(\tau)/\Psi_G$ ($\equiv F(\tau)$) and $\mathbf{H}\Psi_T(\tau)/\Psi_G$ ($\equiv E(\tau)$), are needed as well as the branching multiplicity of Ψ_G,

$$M_{0\to\tau} = \exp\left\{ \int_0^\tau (E_L - E_T) d\tau' \right\}. \tag{2.61}$$

Using the trapezoid rule for integration of Eq. 2.61,

$$M_{0\to\tau_i} = \exp\left\{ \frac{\delta\tau}{n_{\delta\tau}} \left[\tfrac{1}{2} \sum_{i=0}^{n_{\delta\tau-1}} (E_L(\tau=(i+1)\delta\tau) - E_L(\tau=i\,\delta\tau)) \right] - E_T \right\} \tag{2.62}$$

where E_L is the local energy of Ψ_G, and $n_{\delta\tau}$ is the number steps, $\tau/\delta\tau$. The quantities $H(\tau)$ and $N(\tau)$ are then constructed by,

$$H(\tau) = \tfrac{1}{2}\left[F(\tau) M_{0\to\tau} E(0) + F(0) M_{0\to\tau} E(\tau) \right] \tag{2.63}$$

$$N(\tau) = F(\tau) M_{0\to\tau} F(0) \tag{2.64}$$

where the fact that $H(\tau)$ is symmetric in τ (*i.e.*, forward and reverse times are equivalent after equilibrium has been reached) has been used to improve sampling. An example of the convergence of this method with respect to time is given in Table 2.4 for the H_2^+ molecule.

The problem with the above procedure is that it will only provide a transient estimate of the QMC energy; the branching multiplicity $M_{0\to\tau}$ either goes to 0 or ∞ at long time, and that space is sampled less efficiently than in the branching algorithm since configurations in low probability regions are kept, rather than being discarded. These problems are exacerbated if the nodes of Ψ_T and Ψ_G do not coincide. If the ground state energy is all that is desired, then the FNDQMC method described earlier is much more efficient. The value of this method is that since the branching multiplicity is kept as a weight, rather than changing the ensemble size, a method for using this weight to obtain other quantities is easily derived, and since Ψ_G is used, the energies of a set of trial functions can be evaluated, as in Eq. 2.12. In Sec.'s 2.10, 2.11, and 2.12, the time correlation formalism is applied to the problems of computing excited states, expectation values of operators which do not commute with the Hamiltonian, and direct determination of energy differences.

2.10 Calculation of Excited states

The introduction of a guiding function in Eq. 2.60 allows the quantities $H(\tau)$ and $N(\tau)$ to be generalized, by using a different function on either side of the operator,

$$N_{IJ}(\tau) = \langle \Psi_I | e^{-\tau \mathbf{H}} | \Psi_J \rangle \tag{2.65}$$

$$H_{IJ}(\tau) = \langle \Psi_I | \mathbf{H} e^{-\tau \mathbf{H}} | \Psi_J \rangle \tag{2.66}$$

where,

$$|\Psi_I\rangle = \sum_k C_{Ik} \Psi_{T_k} \tag{2.67}$$

the Ψ_{T_k} being a set of trial functions. Eq. 2.67 is used here, as opposed to using the trial functions directly, for added flexibility later on. Ceperley and Bernu [72], showed that the solutions of

$$\det | H_{IJ}(\tau) - \lambda N_{IJ}(\tau) | = 0 \tag{2.68}$$

will converge to the exact eigenstates, Φ_I, of $\mathbf{H}$ for $\tau \to \infty$, as long as each $|\Psi_I\rangle$ has non-zero overlap with its corresponding eigenstate.

This can be seen by dividing the operator $e^{-\tau \mathbf{H}}$ into two operators of $\frac{1}{2}\tau$, and moving one to each side of the Hamiltonian (since $e^{-\tau \mathbf{H}}$ and $\mathbf{H}$ commute),

$$H_{IJ} = \langle \Psi_I e^{-\frac{1}{2}\tau \mathbf{H}} | \mathbf{H} | e^{-\frac{1}{2}\tau \mathbf{H}} \Psi_J \rangle . \tag{2.69}$$

At long time each of the states will decay to the exact state of lowest energy which is not orthogonal to $|\Psi_I\rangle$,

$$\lim_{\tau \to \infty} e^{-\frac{1}{2}\tau \mathbf{H}} | \Psi_I \rangle = a_I | \Phi_I \rangle \tag{2.70}$$

so that,

$$\lim_{\tau \to \infty} H_{IJ}(\tau) = a_I a_J \langle \Phi_I | \mathbf{H} | \Phi_J \rangle \tag{2.71}$$

and,

$$\lim_{\tau \to \infty} N_{IJ}(\tau) = a_I a_J \langle \Phi_I | \Phi_J \rangle \tag{2.72}$$

which will yield the exact eigenstates.

Of primary importance is to choose a Ψ_G which has maximal overlap with each Ψ_I. This can be done either with a simple function chosen to overlap all states, such as a single diffuse Gaussian, or one can use,

$$\Psi_G = \sum_I \left[b_I \Psi_I^n \right]^{1/n} \tag{2.73}$$

where the b_I and n are chosen to adjust convergence to the various states [72]. Both H_{IJ}

and N_{IJ} are needed only for $\tau \to \infty$, but will be calculated at a series of times, τ_i, so that convergence can be observed. The values $\Psi_I/\Psi_G \equiv F_I$, $\mathbf{H}\Psi_I/\Psi_G \equiv E_I$ and $M_{0\to\tau_i}$ are calculated as above, and the matrices are constructed,

$$H_{IJ}(\tau_i) = 0.25\Big[F_I(\tau_i)M_{0\to\tau_i}E_J(0) + F_J(\tau_i)M_{0\to\tau_i}E_I(0) + F_I(0)M_{0\to\tau_i}E_J(\tau_i) + F_J(0)M_{0\to\tau_i}E_I(\tau_i)\Big] \tag{2.74}$$

$$N_{IJ}(\tau_i) = 0.5\Big[F_I(\tau_i)M_{0\to\tau_i}F_J(0) + F_J(\tau_i)M_{0\to\tau_i}F_I(0)\Big] \tag{2.75}$$

where now H_{IJ} and N_{IJ} are symmetric in both I, J as well as τ.

Because of the transient nature of the energy estimator, care must be taken to assure convergence to the proper states [72]. One critical step to insure proper convergence is the construction of a new set of trial functions for each time interval by diagonalizing $H(\tau_i) - \lambda N(\tau_i)$ so that maximum overlap with the true states is maintained. Another consequence of the transient estimate is that the higher excited states will be more poorly converged than the lower states. Thus this method is best suited for Bose systems (such as the vibrational problem [73]) where the energy gap between ground and excited states is relatively small. It is less clear how well this method will work for Fermi states, since the convergence will get worse as the gap between the Bose and Fermi ground states increases. In this respect, the time correlation method is similar to the released-node method [31], and can be expected to be applicable to the same range of systems.

2.11 Calculation of operators which do not commute with the Hamiltonian

In standard QMC the mixed estimator, $\Phi_0\Psi_T$, is obtained. However, for operators, such as the dipole moment, which do not commute with the Hamiltonian, the Φ_0^2 distribution is required [11]. Using the time correlation formalism the matrix elements of a property, Q_{IJ}, can be computed by

$$Q_{IJ}(\tau) = \langle \Psi_I | e^{-\frac{1}{2}\tau H} Q e^{-\frac{1}{2}\tau H} | \Psi_J \rangle / \sqrt{N_{II}N_{JJ}} \tag{2.76}$$

However, unlike in the case of the energies, assuming **Q** and **H** do not commute, the two time propagators must be evaluated in the order indicated. This means that the state I

will be computed at a time τ, the operator at time $\frac{1}{2}\tau$, and state J at time 0. Explicitly,

$$Q_{IJ} = \lim_{t\to\infty} < \tfrac{1}{2}(F_I(\tau)\,\mathrm{Q}(\tau/2)\,F_J(0) + F_J(\tau)\,\mathrm{Q}(\tau/2)\,F_I(0))> \qquad (2.77)$$

where again the symmetry of I and J has been used. This is equivalent to the future walking method of Reynolds, *et al.* [11], the only difference being that in the future walking method one typically performs a QMC walk, allowing the population to fluctuate with the branching, whereas here the quantities are evaluated during a VMC walk. An approach similar to this has been explored by Barnett, *et al.*, for computing transition dipole moments [74].

2.12 Correlated sampling of energy differences

In Sec. 2.2 a method for obtaining energy differences by correlated sampling was described (Eq.'s 2.11, 2.12). Note that the quantities in Eq. 2.12, $\mathrm{H}\Psi_{T_n}/\Psi_G$ and W_n, are precisely E_n and F_n above, and that the energy $H_{nn}(0)/N_{nn}(0)$ is the same as that obtained from Eq. 2.12. Therefore, one method for sampling QMC correlated differences is to use a set of trial functions, as was done in Sec. 2.10 for excited states, but only computing the diagonal elements to obtain the energies. The energy differences themselves will be the same whether the average of the differences or the difference of the averages (*i.e.* $<E_1-E_2>=<E_1>-<E_2>$) is kept. However, the variance of the energy difference must be computed from the energy differences at each step (or at each block), in order to compute the correlated difference variance.

One use of such a scheme is in optimization of the trial function [49]. Using Eq. 2.12, the variational energy of a trial function, Ψ_{T_0}, can be optimized with respect to a set of parameters by changing each parameter slightly to form Ψ_{T_1}, Ψ_{T_2}, an so forth. Then a VMC walk with Ψ_{T_0} as the guiding function is performed to find the energy differences. As stated earlier, if the Ψ_{T_n} are closely related, then the correlated differences will have a lower variance than differences between the averages. The results of the VMC calculation are analysed and a new set of Ψ_{T_n} are chosen, until some convergence criteria is reached. It has often been noted that the best optimization procedure should take into

account not only the variational energy but also the variance of the energy [67,75,76]. Since the correlated difference of any function of Ψ_T can be evaluated in the above procedure, this can easily be done. One can also optimize Ψ_T with respect to some molecular property. This has the drawback that the property must be known. Use of the Hellmann-Feynman *force sum* (the sum of the forces on the nuclei), which is zero for the exact wavefunction, has been suggested by Kern and Karplus[77] for this purpose. Optimization methods have been implemented by Umrigar, *et al.* [67], using the variance only, and Huang and Lester [78] using only the energy. Both of these two extremes show good behavior for small molecules (systems up to 4 electrons are considered).

The time-correlation energy differences (or variance differences) can also be used to optimize the trial function. The advantage of the time-correlated energy differences over the VMC scheme is that one can optimize the fixed-node energy. Although the VMC and FNDQMC energies are related, it is possible that for a given trial function, the nodes of the VMC optimised function may not correspond to the best nodes obtainable in the FNDQMC calculation. To date this procedure has not been used because of the difficulty in dealing with the nodes.

Another kind of energy difference desired involves not just changing the trial function, but also the Hamiltonian. Methods for doing this have been published by Wells [79], and recently by Traynor and Anderson [80]. When using a VMC walk, the same correlated energy difference scheme as above can be used in this case, but now the energies correspond to different Hamiltonians. This can be done because the VMC distribution only depends on the guiding function and its Hamiltonian. All other trial functions or Hamiltonians will be averaged over this distribution. A much different situation arises in the case of the QMC walk. In the QMC walk, different Hamiltonians can have different branching multiplicities during the walk. Thus, for two Hamiltonians, $\mathbf{H}_a$ and $\mathbf{H}_b$, which could represent different molecular geometries [80], or the presence of an external field [79], the quantity $M_{0\to\tau}$ must be evaluated for both. If a branching QMC algorithm is used, then the differential energies must be weighted by the ratio of the two $M_{0\to\tau}$'s. The

VMC and time-correlated energy difference schemes are applied to the molecular geometry optimization problem in chapter 4.

2.13 Conclusions and future directions

QMC in all its forms is just beginning to have significant impact upon quantum chemistry. The strong points of QMC are its simplicity, accuracy, relative basis set independence, and its ability to take advantage of supercomputer architecture. However, there are a number of aspects which need to be improved for QMC to continue its entry into quantum chemistry. First, a robust method for eliminating the time step bias in FNDQMC is needed. This could consist of either a more generally applicable exact-Green's function algorithm than currently available, or a gaining a deeper understanding of the time-step dependence of the energy in FNDQMC. A second area of useful research is into more accurate forms of trial function. To date the ability of Monte Carlo methods to handle any differentiable function has not been widely exploited. Third, all the above QMC schemes are at the present time computationally intractable for atoms beyond the first row of the periodic table (Li-Ne). Methods for dealing with this by focusing on the valence electrons is currently a very active area of research, and new methods for addressing heavy-atom systems are the subject of chapter 3. Finally, QMC methods for molecular geometry optimization have not yet been developed. In chapter 4 directions for doing so are explored, including the calculation of analytic gradients.

Table 2.1 Dependence of the estimated statistical precision (ε) on block size in the H_2^+ molecule (for equal computation).

Number of Blocks	Block size[a]	ε/Hartree
1000	200	0.00069
500	400	0.00084
200	1000	0.00104
100	2000	0.00113
50	4000	0.00124
20	10000	0.00114
10	20000	0.00113
5	40000	0.00108
2	100000	0.00084

(a) Block size is the ensemble size × the number of time steps. The ensemble used consisted of 200 walkers in all cases. The time step was 0.1 hartree^{-1} in all cases.

Table 2.2 Molecular Properties Calculated by FNDQMC

Property	Best Variational	QMC	Experiment or Exact
$H + H_2$ barrier height (kcal/mol)	<9.86[a]	<9.70(13)[g]	9.65(8)[k]
H_2 quadrupole moment (a.u.)	0.493[b]	0.45(2)[h]	0.457[l]
N_2 quadrupole moment (a.u.)	-0.907[c]	-1.06(14)[h]	-1.04(7)[c]
F electron affinity (eV)	3.368[d]	3.45(11)[i]	3.399[m]
CH_2 singlet-triplet splitting (kcal/mol)	9.9[e]	9.5(2.3)[j]	9.55[n]
N_2 binding energy (kcal/mol)	212.9[f]	217.1(2.4)[h]	228.4[o]

(a) B. Liu, J. Chem. Phys. **80**, 581 (1984)

(b) A. D. McLean, M. Yoshimine, J. Chem. Phys. **45**, 3676 (1966)

(c) F. P. Billingsley II, M. Krauss, J. Chem. Phys. **60**, 2767 (1974)

(d) B. H. Botch, T. H. Dunning, Jr., J. Chem. Phys. **76**, 6046 (1982)

(e) H.-J. Werner, E.-A. Reinsch, J. Chem. Phys. **76**, 3144 (1982)

(f) P. E. M. Siegbahn, Int. J. Quant. Chem. **23**, 1869 (1983)

(g) R. N. Barnett, P. J. Reynolds, W. A. Lester, Jr., J. Chem. Phys. **82**, 2700 (1985)

(h) P. J. Reynolds, R. N. Barnett, B. L. Hammond, W. A. Lester, Jr., J. Stat. Phys. **43**, 1017 (1986)

(i) R. N. Barnett, P. J. Reynolds, W. A. Lester, Jr., J. Chem. Phys. **84**, 4992 (1986)

(j) P. J. Reynolds, M. Dupuis, W. A. Lester, Jr., J. Chem. Phys. **82**, 1983 (1985)

(k) D. M. Ceperley, B. J. Alder, J. Chem. Phys. **81**, 5833 (1984)

(l) W. Kolos, L. Wolniewicz, J. Chem. Phys. **43**, 2429 (1965)

(m) H. Hotop, W. C. Lineberger, J. Phys. Chem. Ref. Data **4**, 539 (1975)

(n) 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)

(o) A. Lofthus, P. H. Krupenie, J. Phys. Chem. Ref. Data **6**, 113 (1977)

Table 2.3 Total Electronic Energies (in hartrees) Calculated by FNDQMC

System	Best Variational	QMC	Experiment /Exact
H_2	-1.1737[a]	-1.1745(8)[j]	-1.17447[q]
He (1s2s)	-2.14307[b]	-2.14493(7)[k]	-2.14599[r]
He (1s3s)	-2.06036[b]	-2.06119(7)[k]	-2.06128[r]
Be	-[c]	-14.6657(7)[l]	-14.66733[s]
LiH	-8.06904[d]	-8.0700(4)[m]	-8.07050[t]
CH_4	-40.4584[e]	-40.5063(22)[n]	-40.514(2)[u]
N	-54.5133[f]	-54.5765(12)[k]	-54.5895[v]
N_2	-109.3658[f]	-109.4835(37)[k]	-109.535[w]
H_2O	-76.3683[g]	-76.377(7)[j]	-76.4376[x]
F	-99.7166[h]	-99.7005(21)[o]	-99.7313[y]
F^-	-99.8312[h]	-99.8273(34)[o]	-99.857(3)[z]
FH_2	-100.7892[i]	-100.8861(17)[p]	-100.906[h]

(a) B. Liu, J. Chem. Phys. **58**, 1925 (1973)

(b) Z. Ritter, R. Pauncz, J. Chem. Phys. **32**, 1820 (1960)

(c) Be variational

(d) N. C. Handy, R. J. Harrision, P. J. Knowles, H. F. Schaefer III, J. Phys. Chem. **88**, 4852 (1984)

(e) W. Meyer, J. Chem. Phys. **58**, 1017 (1973)

(f) P. E. M. Siegbahn, Int. J. Quantum Chem. **23**, 1869 (1983)

(g) W. Meyer, Int. J. Quant. Chem. Symp. No. 5, 341 (1971)

(h) F. Sasaki, M. Yoshimine, Phys. Rev. **9A**, 17,26 (1974)

(i) M. J. Frisch, J. S. Binkley, H. F. Schaefer III, J. Chem. Phys. **81**, 1882 (1984)

(j) P. J. Reynolds, D. M. Ceperley, B. J. Alder, W. A. Lester, Jr., J. Chem. Phys. **77**, 5593 (1982)

(k) P. J. Reynolds, R. N. Barnett, B. L. Hammond, W. A. Lester, Jr., J. Stat. Phys. **43**, 1017 (1986)

(l) R. J. Harrison, N. C. Handy, Chem. Phys. Lett. **113**, 257 (1985)

(m) R. N. Barnett, P. J. Reynolds, W. A. Lester, Jr., J. Phys. Chem. **91**, 2004 (1987)

(n) D. R. Garmer, J. B. Anderson, J. Chem. Phys. **86**, 4025 (1987).

(o) R. N. Barnett, P. J. Reynolds, W. A. Lester, Jr., J. Chem. Phys. **84**, 4992 (1986)

(p) D. R. Garmer, J. B. Anderson, J. Chem. Phys. **86**, 7237 (1987).

(q) W. Kolos, L. Wolniewicz, J. Chem. Phys. **43**, 2429 (1965)

(r) C. E. Moore, Natl. Bur. Standards Circ. **467**, I (1949)

(s) C. F. Bunge, Phys. Rev. **14A**, 1965 (1976)

(t) S. Larsson, Phys. Rev. **169**, 49 (1968)

(u) J. A. Pople, J. S. Binkley, Molec. Phys. **29**, 599 (1975)

(v) F. Sasaki, M. Yoshimine, Phys. Rev. **133A**, 419 (1964)

(w) F. Grimaldi, J. Chem. Phys. **43**, S59 (1965)

(x) B. J. Rosenberg, I. Shavitt, J. Chem. Phys. **63**, 2162 (1975)

(y) R. C. Kelley, *Atomic and Ionic Spectrum Lines Below 2000 Angstroms*, Oakridge National Laboratory-5922 (1982)

(z) H. Hotop, W. C. Lineberger, J. Phys. Chem. Ref. Data **4**, 539 (1975)

Table 2.4 Energy vs. time for time-correlation VMC walk on H_2^+. All quantities in a.u.

Time	Energy[a]
0.	-0.44079(36)
5.	-0.45324(34)
10.	-0.45307(35)
15.	-0.45306(36)
20.	-0.45321(37)
25.	-0.45360(38)
30.	-0.45382(39)
35.	-0.45461(40)
40.	-0.45514(41)
45.	-0.45587(42)
50.	-0.45597(42)

(a) Evaluated with a time step of 0.1 hartree^{-1}

Figure 2.1. Equilibration of VMC and QMC walks starting from a random initial configuration for the CCH^- molecule. All quantities are in a.u.

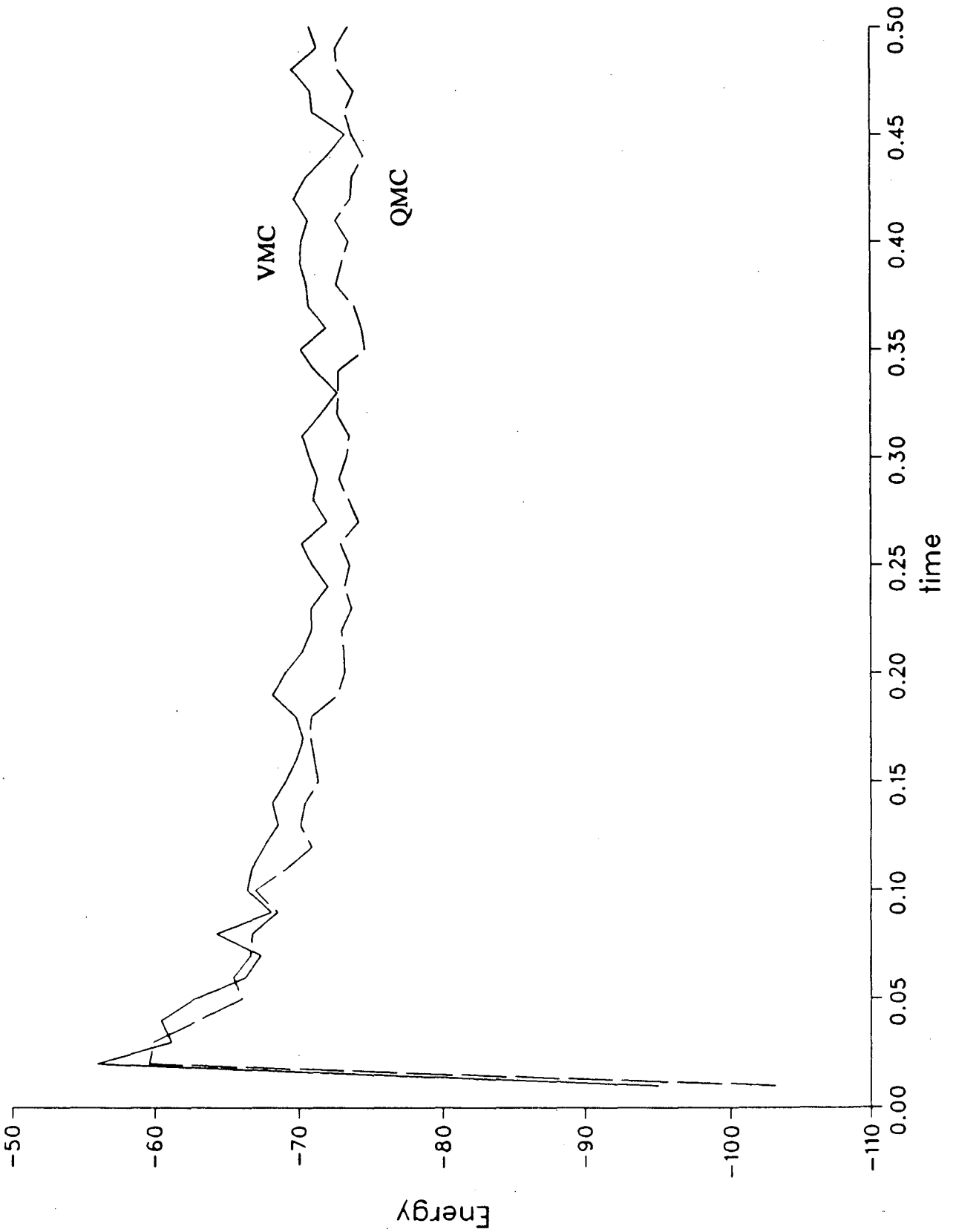


Figure 2.2. CPU time on a CRAY XMP as a function of vector length (ensemble size) for a VMC calculation on the H_2^+ molecule at an internuclear separation of 1.0 bohr. The trial function consisted of a single 1S-STF on each atom with an exponent of 1.53.

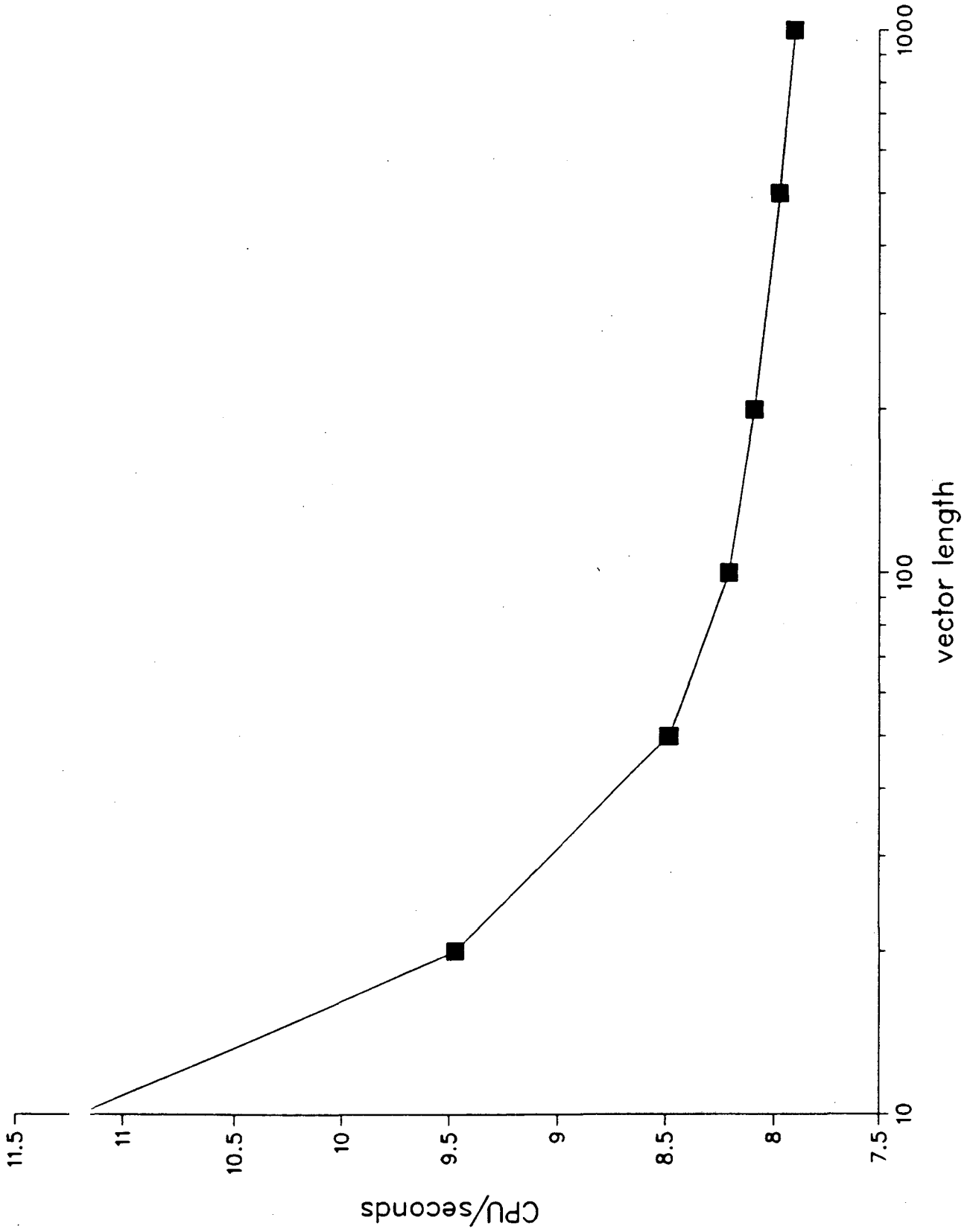
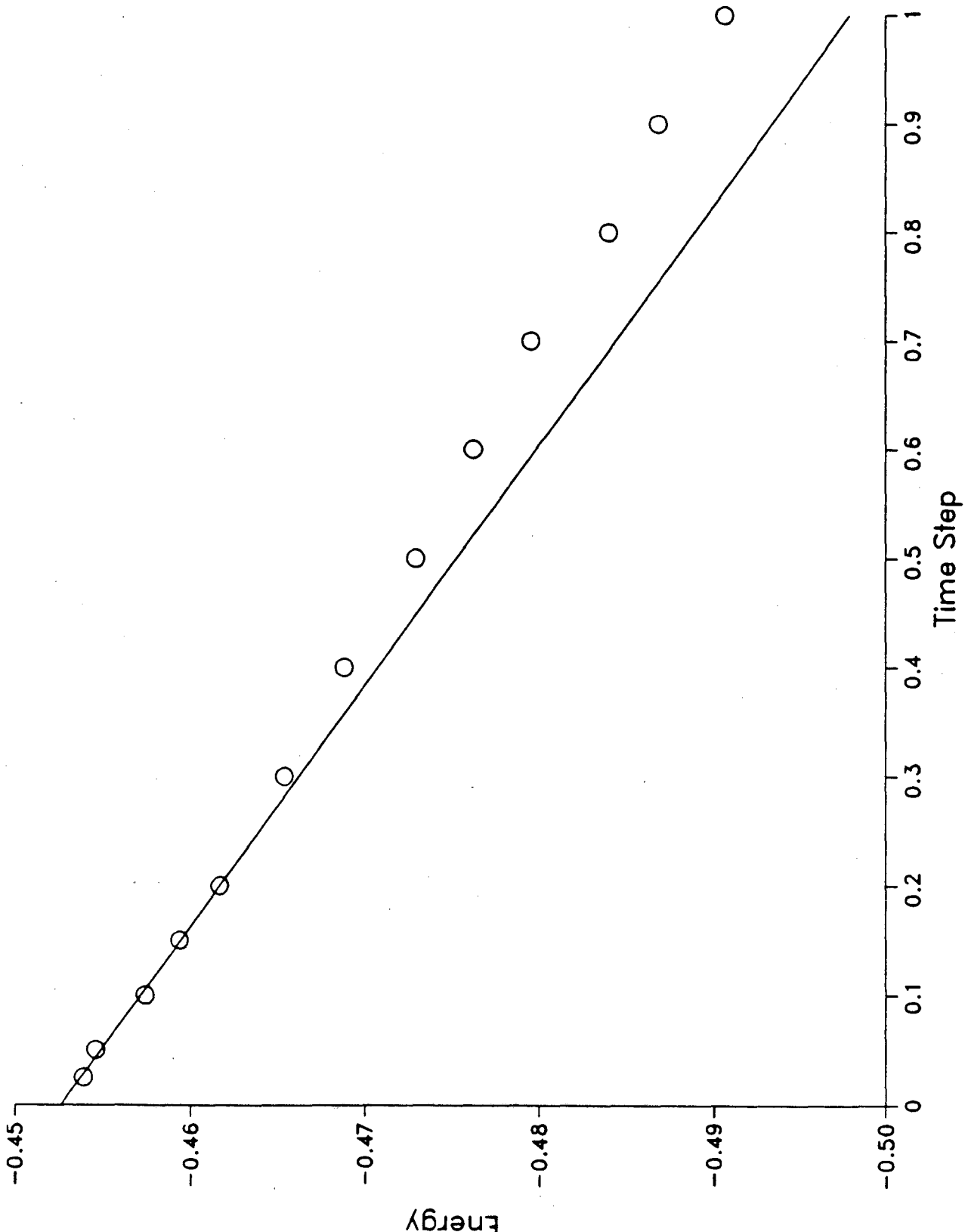


Figure 2.3. Time step bias curve for the H_2^+ molecule at an internuclear separation of 1.0 bohr. The trial function is the same as that used in Table 2.2. The circles include the error bars on the energies. The line represents a linear fit to the data from time steps 0.025, 0.050, 0.100, 0.150, and 0.200. All quantities are in a.u.



CHAPTER 3

Valence Quantum Monte Carlo

3.1 Background: The Heavy-Atom Problem

To date, no all-electron QMC calculations have been published that included atoms heavier than F. This is because of the sharp increase in the computation time required to compute heavy-atom systems [15,81,82]. A rough estimate of the dependence of CPU time on Z (atomic number) for a typical calculation can be obtained as follows [15]. Asymptotically, the CPU time required to compute the energy is determined by the evaluation of the Slater determinant and the two-body potential. To update N electrons requires times of order N^3 and N^2 respectively. The evaluation of the two-body potential dominates at small N, while for large N the evaluation of the Slater matrix will dominate. Thus, for a neutral atom of nuclear charge $Z = N$, if the energy is sampled M times during a simulation, the CPU time, T, scales as

$$T \sim M Z^3 , \tag{3.1}$$

for large Z. The value of M is determined by the statistical uncertainty, ε, required to resolve the energy of interest, E. Specifically, $\varepsilon^2=\sigma^2/M$, where the variance, σ^2, of the underlying distribution may be approximated as a sum of a fluctuation term V_1 [81] and a time-step dominated serial correlation term V_2 [82]. Doll [81] has argued that V_1 is proportional to $E\Delta E$, where ΔE may be regarded as the difference between the variational and exact energies. Ceperley [81] argues that the dominant contribution to ε^2 (at small τ) is from V_2 which is proportional to ΔE. Considering both terms, we obtain

$$M \sim \frac{E\Delta E+\tau^{-1}\Delta E}{\varepsilon^2} , \tag{3.2}$$

where the τ^{-1} dependence results from the reduction in phase space sampled with shorter time steps, in the limit of complete serial correlation.

For energies of chemical interest (*e.g.* binding energies, ionization potentials, *etc.*), ε is essentially constant, while the total energy scales roughly as $E \sim Z^2$. The quantity ΔE can be taken to be the correlation energy, which scales as $Z^{1.5}$ for $Z<20$ [82]. Finally, to determine the dependence of τ on Z note that the radius of the inner electrons decreases as Z^{-1}. (The hydrogenic wavefunction, e^{-Zr}, contracts by this factor.) To avoid crossing nodes, one must decrease τ to keep step sizes sufficiently small. Since the distance traveled in one time step is $\Delta R \sim \tau^{1/2}$, then $\tau \sim \Delta R^2 \sim Z^{-2}$. Thus E and τ^{-1} scale in the same way with Z. Combining Eqs. (3.1) and (3.2), with these substitutions, leads to

$$T \sim Z^{6.5}. \tag{3.3}$$

A related argument has been given by Ceperley, and yields a lower bound of $Z^{5.5}$ [82].

It is the innermost (core) electrons which require these small time steps and are responsible for the steep rise in the energy. Yet it is the valence electrons which largely determine chemical properties such as bond strengths, polarizabilities, electron affinities, and ionization potentials, as well as molecular geometries. Following the above analysis, but treating only the valence electrons, a very different picture emerges. The CPU time no longer depends on Z but on the screened nuclear charge, $Z^{eff}=Z-N_{core}$, where N_{core} is the number of core electrons. Moving across a row of the periodic table, Z^{eff} increases (just as does Z). So, by the argument leading to (3.3), E increases and τ decreases as before. However, the dependence of these quantities on Z^{eff} is considerably weaker than on Z. A fit to the SCF valence energies of the second row elements from Ref. 83 yields $E \sim (Z^{eff})^{0.7}$, compared to the all-electron Z^2 dependence. Since the valence electrons determine the covalent radius, and this radius scales roughly as $(Z^{eff})^{-0.35}$ (again based on a fit of second row atoms), $\tau \sim (Z^{eff})^{-0.7}$. This is consistent with the earlier finding that τ^{-1} scales as E. The quantity ΔE, as indicated above, should have a somewhat weaker

dependence on Z^{eff} than E. Thus, as an upper bound assume $\Delta E \sim (Z^{eff})^{0.7}$. Finally, in the range of values taken by Z^{eff}, calculation of the two-body potential dominates the Slater matrix inversion [8], and $T \sim M \cdot (Z^{eff})^2$. Combining these results leads to a CPU time-dependence that should rise no faster than

$$T \sim (Z^{eff})^{3.4}. \tag{3.4}$$

This is a significantly lower power than the previous $Z^{6.5}$. More importantly however, unlike Z, Z^{eff} remains a small number for *all* atoms. For example, moving down a column of the periodic table, Z^{eff} may generally be held constant. In this situation the valence electrons become more diffuse, E actually *decreases*, and τ may be increased, thus decreasing the required CPU time. Hence, for the valence-only problem, scaling relations such as Eqs. 3.3 and 3.4 do not remain important considerations in determining CPU time.

Another difficulty all-electron calculations face at large Z stems from the fixed-node approximation. Although the fixed-node error is typically a very small percentage of the total energy (*e.g.* 0.02% for CH_2 [9]), when the total energy is large, the fixed-node error can be a significant fraction of the dissociation energy. This was found for N_2 [84]. Thus the placement of the inner nodes can dominate the accuracy of the calculation. Treating only the valence electrons leads to a smaller calculated energy and eliminates the core nodes.

QMC is not alone in its difficulties with heavy-atoms. Other areas of physics have dealt with the problem via the use of pseudopotentials. Such methods are long established in solid state physics and quantum chemistry for large scale computations involving heavy atoms. In standard analytic *ab initio* techniques Hellmann [85] and Gombas [86] were the first to replace the core electrons by an effective potential which contains the core-valence repulsion and orthogonality condition. Since then, two general schemes using pseudopotentials in molecular structure calculations have been most extensively

developed. Using the formalism of Phillips and Kleinman [87], workers such as Goddard and Melius [88], Kahn, Baybutt and Truhlar [89], and Christiansen, Lee and Pitzer [90], have developed an Effective Core Potential (ECP) method which has been extensively tested and used over the past several years in conventional, electronic structure studies [91]. On the other hand, Huzinaga, *et al.* [92], have developed a Model Potentials (MP) approach which is closely related to the frozen-core SCF method.

The goal of the present investigation is to explore and implement methods which exploit the relatively weak Z-dependence of the valence energy in QMC. This aim has been accomplished in two different ways. The first is to eliminate the core electrons entirely by the development of a QMC method implementing *ab initio* ECP's (ECP-QMC)[15]. The ECP-QMC method has also been independently implemented by Hurley and Christiansen [33,93]. In addition, two other methods for eliminating the core electrons have been published: that of Yoshida and Iguchi[34,94] using the Huzinaga MP's, and that of Bachelet, Ceperley, and Chiocchetti[32], using a local pseudo-Hamiltonian (PH) method based on modification of the momentum operator.

While each of these approaches has been successfully used to bring QMC to bear beyond the first row elements, one must always keep in mind the underlying approximation that the core electrons are inert. For many cases this assumption is not entirely true. To treat cases where the core and valence electrons interact strongly, a second direction has been investigated in which the core electrons are explicitly retained, but treated in a separate, less accurate walk. This is the all-electron damped-core (DC) method[95], which shares many of the benefits of the pseudopotential methods (*i.e.* low Z dependence of the CPU time), while mostly avoiding the frozen-core approximation. In the following sections the formalism of ECP's (Sec. 3.2.1), and MP's (Sec. 3.2.2), are applied to QMC. The PH method is outlined in Sec. 3.3, and the DC method is presented in Sec. 3.4. In section 3.5 results obtained from these methods will be compared. Finally in Sec. 3.6 conclusions and future work will be outlined.

3.2 Pseudopotential methods

The basic ECP and MP formalisms can be found in Ref.'s 89 and 92. Although these methods are very similar, there are some important differences regarding implementation in QMC. In both cases the first step is to introduce the core-valence separation,

$$\Psi = \Psi_{core}\,\Psi_{val}\,, \tag{3.5}$$

and then write a valence Schrödinger equation,

$$\mathbf{H}_{val}\Psi_{val} = E_{val}\Psi_{val}\,, \tag{3.6}$$

with

$$\mathbf{H}_{val} = \sum_{i=1}^{N_{val}}\left[-\frac{1}{2}\nabla_i^2 - \sum_A \frac{Z_A^{eff}}{r_{iA}} + \sum_{j<i}^{N_{val}} \frac{1}{r_{ij}}\right] + \sum_A \sum_{B<A} \frac{Z_A^{eff} Z_B^{eff}}{R_{AB}} + \mathbf{U}\,. \tag{3.7}$$

The indices i and j refer to valence electrons while A and B refer to nuclei. Z^{eff} is the screened nuclear charge seen by the valence electrons and other nuclei, and is taken to be $Z - N_{core}$. The core-valence repulsion and the core-valence orthogonality condition are replaced by a *non-local* pseudopotential, **U**. In Sec's 3.2.1 and 3.2.2 the different forms of the pseudopotentials are discussed. Non-local potentials present a problem to QMC [32]; thus a local potential is constructed by integration over the trial function. Because of this, the core-valence antisymmetry will only be represented to the accuracy of Ψ_T, allowing the possibility of the valence QMC solution collapsing into the core. An empirical method for dealing with this is presented in Sec. 3.2.3, using a damped-core Green's function. The computational difficulty in obtaining the local potential is large in the case of ECP's, and details of the analytic integration are left to Appendix C. The issue of using non-analytically-integrable Ψ_T's is discussed in Sec. 3.2.4.

3.2.1 Effective Core Potentials

The theory of Effective core potentials (ECP's) is now well established in conventional *ab initio* approaches [89], and has been reviewed by Krauss and Stevens [96].

Here the approach of Kahn, *et. al.* [89] is followed, writing the non-local pseudopotential as,

$$\mathrm{U}^{ECP} = \sum_{A}\sum_{i=1}^{N_{val}}\left[U^{A}_{l_{max}+1}(r_{iA}) + \sum_{l=0}^{l_{max}}\sum_{m=-l}^{l} |Y_{lm}(\Omega_{iA})> U^{A}_{l}(r_{iA}) <Y_{lm}(\Omega_{iA})|\right], \tag{3.8}$$

where the sum on A is over only those centers having a pseudopotential. For center A, Ω_{iA} is the solid angle of electron i from A, $l_{\max}$ is the largest orbital angular momentum among the core electrons, and U^{A}_{l} is a radial pseudopotential for atom A which depends only on the electron-nuclear distance r_{iA} and the angular momentum l. The spherical harmonics, Y_{lm}, act as projection operators, insuring the correct orthogonality between the missing core and the valence wavefunction [97]. The functions U^{A}_{l} are generally obtained in numerical form from atomic Hartree-Fock calculations. Typically they are then fit to functions of the form,

$$U_l(r) = \frac{1}{r^2}\sum_{k} d_{kl}\, r^{n_{kl}} e^{-b_{kl} r^2}, \tag{3.9}$$

with n_{kl} taking on values of 0, 1, and 2.

For QMC the non-local ECP operator must be made into a local form. One way to accomplish this is to allow U^{ECP} to act on Ψ_{val} in analogy to the local energy [15,33], $\mathrm{H}\Psi_{val}/\Psi_{val}$, leading to,

$$U^{ECP}_{L}[\Psi_{val}] = \tag{3.10}$$
$$\sum_{A}\sum_{i=1}^{N_{val}}\left[U^{A}_{l_{max}+1}(r_{iA}) + \sum_{l=0}^{l_{max}}\sum_{m=-l}^{l} Y_{lm}(\Omega_{iA})\, U^{A}_{l}(r_{iA}) <Y_{lm}(\Omega_{iA})|\Psi_{val}>/\Psi_{val}\right],$$

where Ψ_{val} is the QMC valence importance function. Although at first this procedure seems exact, there is in fact a subtle approximation being made. In standard QMC, because the Hamiltonian is hermitian, the exact eigenvalue can be obtained from the

mixed estimator, $<\Phi|\mathbf{H}|\Psi_{val}>$ where Φ is the eigenfunction sampled in the fixed-node QMC random walk [8]. This would seem to indicate that the exact value of $<\Phi|\mathbf{U}^{ECP}|\Psi_{val}>$ can be calculated in the usual manner. However, this will rarely be the case because one cannot in general sample the exact wave function Φ, without knowing $<\Phi|\mathbf{U}^{ECP}|\Psi_{val}>$. Thus the usual QMC walk will generate a distribution self-consistent with $U_L^{ECP}[\Psi_{val}]$ rather than the the exact distribution one would obtain with the non-local $\mathbf{U}^{ECP}$. However, this approximation will be small if Ψ_{val} is a good approximation to Φ. Note that for variational QMC there is no approximation, *i.e.*, a random walk simulation using $U_L^{ECP}[\Psi_{val}]$ will self-consistently produce the distribution Ψ_{val}^2.

The form of the proper local potential can be found by consideration of the ECP matrix element in QMC,

$$<\Phi|\mathbf{U}^{ECP}|\Psi_{val}> = \int d\mathbf{R}\ \Phi\mathbf{U}^{ECP}\Psi_{val} \tag{3.11}$$

Let $\tilde{\mathbf{R}}$ be the coordinates of all the electrons *except* electron i, and let atom A be the center of the coordinate system (so that $r_{iA} \equiv r_i$). Thus the ECP operator centered on atom A which acts on electron i has the matrix element

$$<\Phi|\mathbf{U}^{ECP}(iA)|\Psi_{val}> = \sum_{lm} \int d\tilde{\mathbf{R}} \int d\Omega_i \int dr_i\ r_i^2\ \Phi(\tilde{\mathbf{R}}, r_i, \Omega_i) Y_{lm}(\Omega_i) U_l(r_i) \tag{3.12}$$

$$\times \int d\Omega'_i\ Y_{lm}(\Omega'_i) \Psi_{val}(\tilde{\mathbf{R}}, r_i, \Omega'_i)$$

$$= \sum_{lm} \int d\tilde{\mathbf{R}} \int dr_i\ r_i^2\ <\Phi|Y_{lm}> U_l(r_i) <Y_{lm}|\Psi_{val}>$$

The $<\Phi|Y_{lm}>$ and $<Y_{lm}|\Psi_{val}>$ integrals are functions only of r_i and $\tilde{\mathbf{R}}$, so Ω_i can be reintroduced by multiplying by $1 = \int Y_{lm}^2$,

$$<\Phi|\mathbf{U}^{ECP}(iA)|\Psi_{val}> = \sum_{lm} \int d\tilde{\mathbf{R}} \int dr_i\ r_i^2\ <\Phi|Y_{lm}> U_l(r_i) <Y_{lm}|\Psi_{val}> \int d\Omega_i\ Y_{lm}^2 \tag{3.13}$$

$$= \int d\mathbf{R}\,\Phi\left[\sum_{lm}\frac{<\Phi|Y_{lm}>Y_{lm}^2(\Omega_i)U_l(r_i)<Y_{lm}|\Psi_{val}>}{\Phi\Psi_{val}}\right]\Psi_{val}\ .$$

The sum of the quantity in brackets over all electrons and ECP centers is defined to be the local pseudopotential for the mixed distribution, $U_L^{ECP}[\Phi\Psi_{val}]$. One could similarly argue that the pure estimator,

$$U_L^{ECP}[\Phi^2] = \sum_{iA}\sum_{lm} Y_{lm}^2 U_l <\Psi_{val}|Y_{lm}>^2/\Phi_{val}^2\ , \tag{3.14}$$

should be used. In either case the expressions for $U_L^{ECP}[\Phi\Psi_{val}]$ and $U_L^{ECP}[\Phi^2]$ are consistent with the claim made above that one must know Φ in order to represent the non-local potential exactly. Although it may be possible to sample this local potential by some iterative future-walking algorithm [11], the computational effort will most likely be just as bad, if not worse, than the original all-electron problem. However, Eqs. 3.13 and 3.14 suggest a second approximate local potential based on the VMC estimator, namely,

$$U_L^{ECP}[\Psi_{val}^2] = Y_{lm}^2 U_l <\Psi_{val}|Y_{lm}>^2/\Psi_{val}^2 \tag{3.15}$$

which will have different properties from the original $U_L^{ECP}[\Psi_{val}]$ (Eq. 3.10), and thus may prove useful.

Evaluation of the $<\Psi_{val}|Y_{lm}>$ is itself a rather demanding task, and is described in detail in Appendix C. Because of the difficulty of obtaining these integrals, Appendix C also outlines a method by which these integrals are fit to a grid for later use during the walk.

3.2.2 Model Potentials

An alternative approach to ECP's is the model potential method developed principally by Huzinaga and co-workers [92], and used in QMC by Yoshida and Iguchi [34,94]. The MP method is unique in pseudopotential methods in that it is based upon the Hartree-Fock-Roothaan method with the frozen-core approximation, rather than on the Phillips-Kleinman method. The important feature here is that the MP method is capa-

ble of reproducing the *ab initio* valence nodes, rather than relying on smoothed out pseudo-valence orbitals. This approach could possibly have great benefit in QMC if this results in a smoother E_L surface.

The method uses the level shift operator,

$$\sum_c |\phi_c> B_c <\phi_c| , \tag{3.16}$$

which shifts the energy of a core orbital, ϕ_c, above the valence orbitals so that the resulting lowest eigenstate will describe the valence electrons. The level shift operator is functionally similar to the angular momentum projection operators used in ECP's. Added to this is a radial potential to represent the coulomb repulsion of the core,

$$V_{MP} = -\frac{Z^{eff}}{r}(\sum_p A_p r^{n_p} e^{-\alpha_p r^2}) . \tag{3.17}$$

Thus, one representation of the local potential would be

$$U_L^{MP}[\Psi_T] = V_{MP} + \sum_c \phi_c B_c <\phi_c | \Psi_T> / \Psi_T . \tag{3.18}$$

As in the case of ECP's the ϕ_c's will not be able to enforce exact orthogonality unless they describe the exact core, but there are possible advantages to MP's over ECP's. The first of which, already mentioned, is that the *ab initio* valence nodal structure is maintained. In addition, the integrals $<\phi_c | \Psi_T>$ are notably simpler to perform than the angular integrals needed in ECP's. In fact one may construct Ψ_T such that $<\phi_c | \Psi_T> = 0$. This means that the QMC walk would then rely entirely on the V_{MP} and the nodes of Ψ_T (which have been chosen orthogonal to $\{\phi_c\}$) to maintain orthogonality. Using the nodes of Ψ_T to maintain orthogonality in the fixed-node method has been previously used successfully in treating electronic excited states [13].

3.2.3 Damped-core Green's Function

As was observed above, the local pseudopotential does not keep the QMC solution orthogonal to the "exact" core, thus allowing collapse of the valence solution into the

core region. Since the local potential reproduces the "exact" non-local operator as $\Psi_{val} \rightarrow \Phi_0$, one can increase both the 1-particle and n-particle basis sets in hope of rectifying this behavior. There is no way of knowing how large Ψ_{val} will have to be for this solution to work. Very large basis sets (either 1- or N-particle) would reduce the attractiveness of the QMC method.

Rather than going to large trial functions one can instead modify the QMC method. Since the branching is at fault, the multiplicity can be damped out in the core region using a "damped-core" Green's function (DCGF). The DCGF is simply the Green's function of the system with the branching term modulated by a cutoff function $\gamma(\mathbf{R})$. For the FNDQMC algorithm the DCGF is

$$G(\mathbf{R}\rightarrow\mathbf{R}',\tau) = G_D\ [G_B]^{\gamma(\mathbf{R})} \tag{3.19}$$

$$G_D = (4\pi D\,\tau)^{-3N/2}\exp\{-[\mathbf{R}'-\mathbf{R}-D\,\tau F_Q(\mathbf{R})]^2/4D\,\tau\}$$

$$G_B = \exp[-\tau\{[E_L(\mathbf{R})+E_L(\mathbf{R}')]/2 - E_T\}]\ ,$$

G_D and G_B being the standard diffusion and branching terms[8]. By damping the inner portion of the valence distribution where overlap with the core is greatest, the QMC solution remains orthogonal to the approximate core represented by the pseudopotential.

The simplest choice of γ would be to make it zero for any configuration which has one or more electrons within a given cutoff radius of the ECP nucleus, and unity elsewhere. This gives a conceptual definition of what a core-like configuration is, but will also cause a discontinuity in the QMC solution. Instead we take γ defined as

$$\gamma(\mathbf{R}) \equiv \prod_{A}\prod_{i} [1+\exp(-(r_{iA}-\rho_A)/\alpha_A)]^{-1} \tag{3.20}$$

where A is the damped-core atom, i loops over electrons, and ρ and α are the cutoff radius and width. When a valence electron approaches an atomic core, $\gamma(\mathbf{R})$ will approach zero and the branching will go to unity. With a branching of unity, we perform precisely the VMC walk, yielding the variational solution Ψ^2_{val} in this region, which is orthogonal to the core by construction. This approximation smoothly connects the pure

fixed-node QMC solution, which corresponds to *i.e.* $\rho=0$, to the variational solution, corresponding to $\rho=\infty$. Unfortunately, this also means that the result will depend on the choice of ρ and α.

3.2.4 Slater-type functions and correlation functions

Rendering the non-local potentials in local form requires integration over Ψ_{val}, either explicitly in the case of ECP's, outlined in Appendix C, or implicitly in the case of MP's, using the orthogonality of ϕ_c and Ψ_{val}. In Appendix C, GTF's have been used to evaluate the ECP integrals because the results are relatively simple, just as GTF's are standard in conventional *ab initio* molecular-orbital approaches because electronic integrals take on a particularly simple analytic form [98].

One of the great strengths of standard all-electron QMC, however, is that no analytic integration is necessary. Instead, what is most important is a smooth local energy surface to reduce the variance of the random walk, and thus the computation time needed to achieve the desired statistical uncertainty. In this respect GTF's are inferior to STF's. Thus it has been common in QMC to use determinants constructed from STF's, and to multiply the determinant by electron-electron and electron-nuclear correlation functions (EECF's and ENCF's). For analytic integration of the local ECP, only GTF's without correlation functions have been used in Appendix C. In principle STF's and correlation functions can also be included analytically or numerically, although at added computational cost. Note that STF's are only a problem in treating molecules by ECP-QMC [15] — STF's are not a problem either for MP-QMC or in the case of isolated atoms in ECP-QMC.

For the general case, when STF's are included in the basis set, there are several options for evaluating the angular integrals. One option is to analytically expand all STF's [99] about the ECP center, as is done for the GTF's. This expansion has the drawback that it is much more cumbersome than for GTF's. Another approach for including

STF's involves numerical integration. Possibly the best option is to expand the STF's in terms of GTF's and to proceed straightforwardly. This could be done either by a least squares fit, as in the STO-nG basis sets [100], or by a Gaussian integral transform [101]. Of these alternatives, the least squares method is the simplest, and has been implemented for all the cases involving STF's described here.

Correlation functions can be treated in a manner similar to the STF's. In particular, a suitable expansion or numerical technique can be employed. It can be argued, however, that the correlation functions need not be included in the integrals at all. This stems from the fact that both the ENCF and EECF are essentially constant away from the region where the pair of particles meet, which enables the correlation functions to be factored out of the angular integral (*c.f.* Eq. 3.10). Consider an ECP atom A. An ENCF centered on A is effectively constant unless an electron is inside the core of A. However, a valence electron should not be found frequently inside an ECP core due to the strong repulsive form of the ECP. In this situation, it is a good approximation to factor the ENCF out of Eq. 3.10. On the other hand, an ENCF centered on a *different* core cannot be factored out of the angular integral — but the radial potential, U_l^A, is exponentially small away from the core of A. Thus there is effectively no contribution to U_L^{ECP} in either case. A similar result is found for the EECF, which is constant except when two electrons are close together. Both electrons must be in the core region of a nucleus with an ECP for a contribution to U_L^{ECP}. As with the ENCF's, it is a good approximation to factor EECF's out of the angular integral. Factoring correlation functions can be viewed either as an approximation to the integral or as a modification of the pseudopotential. In either case, the effect on the energy must be examined to insure that the results are not biased by this treatment.

3.2.5 Treatment of Core-valence interactions

A number of cases arise, particularly in the alkali metals, for which the core and

valence electrons interact strongly. In such cases, one anticipates errors due to the frozen-core approximation. The only manner in which these effect can be treated in ECP or MP is through the use of an additional pseudopotential designed to take into account the mutual polarization of the atomic cores, and the core-valence correlation. Müller, *et al.* [102], use a core polarization potential (V_{CPP}) based on the electric field operator, $\mathbf{f}_c$,

$$V_{CPP} = -\tfrac{1}{2}\sum_A \alpha_A \mathbf{f}_A \cdot \mathbf{f}_A \,. \tag{3.21}$$

The electric field acting on atomic core A due to the electrons i and other cores B is given to be,

$$\mathbf{f}_A = \sum_i \frac{\mathbf{x}_{iA}}{r_{iA}^3} C(r_{iA}, \rho_A) - \sum_{B \neq A} \frac{\mathbf{X}_{AB}}{R_{AB}^3} \,, \tag{3.22}$$

where $\mathbf{x}$ and $\mathbf{X}$ are the Cartesian coordinated of the electron and nucleus, respectively, and $C(r_{iA}, \rho_A)$ is a cutoff function. Eq. 3.22 is a function of the coordinates only, and is thus easily evaluated in QMC.

There is, however, a possible problem in the QMC implementation of Eq. 3.21. The quantity $\mathbf{f}_c \cdot \mathbf{f}_c$ goes as $1/r^4$ near the origin, and its variance goes as $1/r^8$. Unless the combination of the cutoff function and the trial function contribute a factor of r^6 or greater, the variance of Eq. 3.21, and possibly the quantity itself will be divergent. In fact, for the ECP formalism, the valence trial function should go to zero as r^{l+2} where l is the angular momentum of the pseudopotential U_l. Thus for first row atoms, for which only the U_0 term is present, the trial function should go as r^2 near the nucleus, which will contribute an overall factor of r^4 to the variational integral. Note, however, that in order to insure r^2 behavior in the trial function, only 3S, 3P, 3D, and higher STF's can be used in the basis set. Assuming the trial function has the proper behavior, the cutoff function must go to zero as at least r^2 for both V_{CPP} and its variance to be convergent. In ref. 102 a cutoff of the form

$$C(r,\rho_c) = (1 - e^{(r/\rho_A)^2})^2 \tag{3.23}$$

is suggested. Expanding about r=0, one can see that this cutoff function goes to zero as r^4, which satisfies the above criteria.

Implementation of this method, then, entails calculation of the potential V_{CPP}, as part of the local energy. The question still remains as to how big the variance of V_{CPP} will be, even though here it is shown to be finite.

3.3 Pseudo-Hamiltonian method

Because of the ambiguity introduced by the local approximation it is useful to find a local Hamiltonian which replaces the core electrons. One method for doing this has been recently introduced by Bachelet, Ceperley, and Chiocchetti [32]. They first define three properties of all-electron QMC which they wish to maintain in a valence QMC method. First, the fixed-node Green's function should be positive or zero everywhere. This allows simulation of many-Fermion systems. Second, the fixed-node energy should be a variational upper bound to the true Fermion energy. And third, only the positions of the nodes should determine the energy, to maintain relative basis set independence. The second and third criteria exclude the use of non-local operators, since the energy can be either above or below the exact, and the pseudopotential depends globally on Ψ_{val}, not just the nodes.

To satisfy these criteria, they modify the kinetic energy, instead of the potential energy. This has the effect of changing the electron mass in the core region. The pseudo-Hamiltonian used is,

$$\mathbf{H} \equiv \sum_{i<j} \frac{1}{r_{ij}} - \tfrac{1}{2}\sum_i \nabla_i^2 + \sum_{i,A} h_A(r_{iA}), \tag{3.24}$$

where,

$$h_A(r) \equiv -\tfrac{1}{2}\nabla a(r)\nabla + \frac{b(r)L^2}{2r^2} + v(r). \tag{3.25}$$

The functions $a(r)$ and $b(r)$ are required to go to zero outside the core region and $v(r)$

goes to the standard valence electron-nuclear potential. By writing the one-electron radial Schrödinger equation, $a(r)$, $b(r)$, and $v(r)$, are related, such that only one of these functions need be specified. At this point $a(r)$ is approximated to be $a_0 e^{-(r/r_c)^k}$, where a_0, r_c and k are chosen to match atomic valence orbitals (s,p,d). Then $b(r)$ and $v(r)$ are solved for in terms of $a(r)$. Note, that here the problem is to define the functions $a(r)$, $b(r)$, and $v(r)$, rather than the pseudopotential and trial function in ECP- and MP-QMC.

This has the advantage over both ECP and MP formalisms in that it is local, and no further approximations need be made. But on the other hand, the authors report that they may not be able to effectively treat first row atoms or transition metals because of the inherent non-locality of the core-valence interaction. Also note, that this is still a frozen-core method, so in cases where the core and valence electrons interact strongly, none of the above methods (ECP, MP, or PH) will be accurate without some sort of correction term.

3.4 All-electron Damped-Core QMC

As noted above, the use of pseudopotentials is known to lead to inaccuracies in cases where the core electrons significantly influence the valence electrons through either electronic correlation or polarization effects. In addition, each different partitioning of core and valence electrons requires computation of a new pseudopotential, a task requiring additional effort and great care. One method to partially correct for these effects was discussed in Sec. 3.2.5.

Here a novel QMC approach is presented, based on separating the energies and time scales of the core and valence electrons, which treats correlation and polarization effects in a consistent fashion. The main goal is to reduce the Z-dependence of the QMC method which is largely attributable to the core electrons [15]. The two main effects are that the core energy rises as Z^2, requiring longer computation to achieve the same absolute statistical uncertainty, and that the time scale of the core electrons decreases as Z^{-2},

necessitating small time steps which decrease the efficiency of the method. Clearly, one way around this is to replace the core electrons by a pseudopotential, which does result in a much smaller Z-dependence[15,33,34].

Instead, one can stay within the framework of QMC to deal with the Z-dependence by separating the core- and valence-electron walks such that each can move with its own appropriate time scale. First, the wavefunction is approximately partitioned into core and valence parts (Eq. 3.5) allowing distinction between core and valence electrons. Since most of the chemical information is contained in the valence electrons the core wavefunction is approximated to be a known core trial function, Ψ_{core}. This is done using a VMC walk for the core electrons guided by Ψ_{core}. The valence electrons are then treated by fixed-node diffusion QMC in the potential of the core electrons, to obtain the valence energy of the system. This addresses both of the main large Z effects mentioned above since the valence energy was shown to have has a much lower-order dependence on Z in Sec. 3.1, and the valence electrons are released from the time-scale of the core allowing appropriately larger time steps to be used in the fixed-node walk.

The core and valence trial functions for the QMC procedure are obtained by performing an all-electron SCF calculation (as is commonly done in standard QMC) and using the corresponding molecular orbitals. For a system containing N_{core} core electrons and N_{val} valence electrons, Ψ_{core} and Ψ_{val} are

$$\Psi_{core} = \det|\psi_1(1)\cdots\psi_{N_{core}}(N_{core})| \tag{3.26}$$

$$\Psi_{val} = \det|\psi_{N_{core}+1}(N_{core}+1)\cdots\psi_{N_{val}+N_{core}}(N_{val}+N_{core})|. \tag{3.27}$$

The respective Hamiltonians are,

$$\mathbf{H}_{core} = \sum_{s=1}^{N_{core}}\left[-\frac{1}{2}\nabla_s^2 - \sum_A \frac{Z_A}{r_{sA}} + \sum_{t<s}\frac{1}{r_{st}}\right] + V_{NN} - V_{NN}^{val} \tag{3.28}$$

and,

$$\mathbf{H}_{val}=\sum_{i=1}^{N_{val}}\left[-\frac{1}{2}\nabla_i^2-\sum_A\frac{Z_A}{r_{iA}}+\sum_{j<i}\frac{1}{r_{ij}}+\sum_s^{N_{core}}\frac{1}{r_{is}}\right]+V_{NN}^{val} \tag{3.29}$$

The indices s and t refer to core electrons, i and j to valence electrons, and A and B to nuclei. V_{NN} is the standard repulsive nuclear potential energy and V_{NN}^{val} is taken to be,

$$\sum_A\sum_{B<A}\frac{Z_A^{eff}Z_B^{eff}}{R_{AB}}, \tag{3.30}$$

Z^{eff} being $Z-N_{core}$. Note that the valence Hamiltonian is the same as in the pseudopotential case (Eq. 3.7), with the pseudopotential, U identified as,

$$\mathbf{U}=\sum_{i=1}^{N_{val}}\left[\sum_A\frac{Z_A^{eff}}{r_{iA}}-\sum_A\frac{Z_A}{r_{iA}}+\sum_s^{N_{core}}\frac{1}{r_{is}}\right] \tag{3.31}$$

Because of the factorization in Eq. 3.5, strict orthogonality between the QMC valence solution and VMC core is not maintained, *i.e.* $\langle\Phi_{val}\,|\,\Psi_{core}\rangle\neq 0$ (as in the case of ECP- and MP-QMC). At long time the valence QMC solution will be able to collapse into the core region. So the DCGF can be used in the valence walk to maintain approximate orthogonality between core and valence solutions. A subtle point in using the DCGF is that since the DCGF sets the branching to be unity in the core region, the trial energy must be adjusted so that the average branching in the valence region is also greater than or, preferably, equal to unity. Otherwise, rather than keeping the valence walkers out of the core region, $\gamma(\mathbf{R})$ will instead insure that valence walkers in the core region will survive, while those in the valence region die out. Thus use of the DCGF results in a parameter dependence on E_T, unlike in the case of all-electron QMC.

It should be noted that the all-electron SCF function could be used as Ψ_{val}, avoiding the above separation. (Clearly this function cannot be used for Ψ_{core}, else no separation of core and valence would be possible.) However, such a Ψ_{val} leads to the following problem. Suppose we take a single core and valence walker. The nodes of Ψ_{core} depend only on the position of the core walker, but the nodes of Ψ_{val} now depend on both the valence and the core walkers. Thus when the core walker moves it can cause a valence

node to approach the valence walker *without itself feeling the influence of the node.* In the all-electron walk the core electron would be repelled from this node. Hence, the above factorization has been used here. Explicit core-valence EECF's can be used in both Ψ_{val} and Ψ_{core}, however, because these functions have no nodes of their own and will not affect the nodes of either the valence or core walks.

Implementation involves two separate, but linked, random walks: a core VMC walk, and a valence fixed-node diffusion QMC walk. Here, the core is chosen to be represented by a single walker. Using a single core walker is advantageous for systems containing a large number of core electrons, so that the evaluation of the core walk remains a small portion of the overall calculation. The core random walker is unaffected by the valence electrons since the core Hamiltonian and trial function are chosen to contain no valence terms (assuming core-valence EECF's are not used). The valence electrons, in turn, all use this walker to evaluate the core-valence two electron potential. Because the core electrons are represented explicitly, core-valence electronic correlation effects are taken into account. Core-core electronic correlations may be treated using EECF's or multi-determinant core trial functions. The atomic cores can also polarize in the molecular environment to the accuracy of the method used to calculate the trial wavefunction. Note that if one wished to reproduce the frozen-core, an ensemble of core walkers could be used, allowing the valence electrons to see only the average two-electron potential. In the limit of an infinite number of core walkers the DC method greatly resembles the MP-QMC method, with the pseudopotential given by the average of the coulomb interactions with the core-ensemble.

For large systems in which there are either too many electrons to be treated in an all-electron SCF, or too many valence electrons to be treated in the above described procedure, a hybrid between pseudopotentials and damped-core QMC is possible. For example, for first-row transition metals it is common practice to include the 3s and 3p shells in the valence space because of their strong overlap with the 3d orbitals. In this

case one can represent the 1s, 2s and 2p electrons using either ECP's or MP's, and use DC to represent the 3s and 3p electrons retaining only the 4s and 3d electrons in the fixed-node walk.

Another strategy which can be employed to reduce the variance in the core-valence potential is to use the same set of configurations to represent the core electrons in the molecular environment as was used in the atom (or between neutral and ion). This is an application of the correlated sampling method, and should result in a reduced variance in the energy difference between the systems calculated in this manner.

Also in large systems where relativistic effects cause a large contraction of the core, a relativistic Hartree-Fock Ψ_{core} used in DC will reflect these effects. This is, in principle, identical to using a relativistic effective core potential in QMC. Other relativistic effects can be estimated on the VMC level [103], however no procedure has been published for doing so in the QMC walk.

3.5 Results and Discussion

Although the first valence QMC paper was published only a year ago, there are now numerous, and sometimes conflicting, results available for the various techniques outlined above. The ECP-QMC method has been pursued by Hammond, *et al.*[15], and Christiansen *et al.*[33,93], the MP-QMC method by Yoshida and Iguchi [34,94], the PH method by Bachelet, *et. al.*[32], and the DC method by Hammond, *et al.*[95]. Here an in depth discussion of the results of these methods is given.

The discussion is broken into two categories of atoms and molecules: those containing one- or two- valence electrons, and those with three or more. This is done in part because of simplifications of the theories in the one- and two-valence electron cases, and because the greatest body of data exists for such systems. For the many-valence electron case fewer results have been published both because of the added computational

requirements, and, as is pointed out later, difficulties with use of non-local pseudopotentials.

3.5.1 One- and Two-valence electron systems

One- and two-valence electron systems treated to date include Li[15,33], Na[15,32], K[33], Mg[33,34], Ca[34], Sr[34], and their respective anions or cations, Be[93], and the diatomics NaH[15] and Na_2[15,32]. Table 3.1 compares ionization potentials (IP's) and electron affinities(EA's) of Li and Na, and the IP of Mg, using HF, CI, MP-QMC, ECP-QMC, PH and DC with the experimental values. The IP results for Na and Li are taken to be the respective valence energies. While this results in good IP's in the DC method, the valence energies in the ECP-QMC and PH methods do not agree well with the IP. This is most likely due to errors in the pseudopotentials (or pseudo-Hamiltonian) used. These errors seem to cancel out when two valence energies are subtracted, as in the case of the EA's of Li and Na, and the IP of Mg, all of which are close to the experimental values. The approximation of replacing the core electrons by a pseudopotential is found only to be accurate for relative energies, making direct comparison of valence energies between methods impossible. This is reflected in the work of Christiansen [93] on Mg and Be where even the relative energies were found to be very sensitive to the trial function used in the local approximation.

The great time savings predicted in Sec. 3.1 for valence only techniques can be readily seen in the cases of the NaH and Na_2 molecules. Tables 3.2 and 3.3, and Fig. 3.1 give the binding energies of Na_2, NaH, and and the potential curve of Na_2. In order to compute the binding energy of Na_2 by all-electron QMC, with a total energy of nearly -320 hartree [104] and an experimental binding energy of only 0.0275 hartree [105], a statistical uncertainty of less than 5-10 parts per million would be required in a QMC calculation of this 22 electron system. Furthermore, the potential energy curve has a very broad minimum, and resolution of energies at bond lengths differing by 0.8 bohr would

further require a statistical uncertainty of less than 2 parts per million. However, for ECP-QMC the valence energy is only 0.3925 hartree, meaning that the present statistical uncertainty of 0.2% is sufficient to resolve the binding energy and points on the potential energy curve in Fig. 3.1.

For the Na_2 molecule, both ECP-QMC[15] and PH[32] formalisms have been used. As seen in table 3.2, however, the PH are not in agreement with the ECP-QMC and experimental values. This is alarming since the PH value for the EA of Na was within its error bars of the experimental value. This does not mean the ECP-QMC method is without its problems. It was noted in Ref. 15 that in the case of NaH use of a GTF basis set for the H atom led to large variances with respect to what is expected from a STO basis set. The need for both an STO-6G basis and an ENCF led to the ECP-QMC procedure described above, where STO basis sets are used throughout, with the necessary angular integrals being computed with the corresponding STO-6G expansion. Christiansen [93] has also found strong dependence on the trial function. Thus, in the ECP-QMC approach great care must be taken in selecting a trial function, which determines the local pseudopotential, whereas in the PH case, the difficulty rests mainly with determination of the PH itself.

For all the pseudopotential approaches, agreement with experiment will be difficult for cases with a large core-valence interaction, as in the heavier alkali and alkaline earth metals. This trend is indicated in the results of Yoshida and Iguchi [34], whose IP for Mg was within 0.075(32) of the experimental, yet Ca and Sr were off by 0.235(32) and 0.122(27) respectively. These difficulties should be alleviated in the DC approach, even though the computational requirement is greater.

The DC results given in Table 3.1 are obtained from a core ensemble of a single walker while the valence ensemble was 100 to 400. No core-valence terms were included in the trial wavefunction. The trial functions were constructed from SCF functions using atomic basis sets taken from Clementi [106] multiplied by linear electron-

electron and electron-nuclear Pade-Jastrow functions (Sec. 2.8). Separate values for b (inversely proportional to the EECF range) were allowed for Ψ_{val} and Ψ_{core}, representing the core-core and valence-valence electronic correlations independently, while only a single value for v (the ENCF parameter corresponding to b is called v here, which is inversely proportional to the ENCF range) was used for each atom since this parameters depends most strongly on the core electrons. The values used in the atomic basis sets and Pade-Jastrow functions are given in Table 3.6 and 3.7 respectively.

The fact that both DC-computed IP's and EA's, show good agreement with experiment is an indication that the core-valence interaction is being correctly described in an absolute sense, unlike in the pseudopotential cases. The quoted error bars should not be considered the limit achievable by this method, since investigation of trial functions for these systems was not exhaustive and computation was restricted.

The DCGF was found to play a role only for propagation to very large imaginary times. Thus it was possible to set ρ to a nominal value (0 to 0.1 with α=0.025) and propagate the ensemble only far enough to achieve equilibrium. The ensemble is then expanded by taking configurations generated at time intervals, and propagated a relatively small number of steps. This process is facilitated by the large valence time steps used in the DC technique.

3.5.2 Many-electron systems

Multi-electron systems treated to date include the atoms C[95], Si[32,95], Cl[32,94], Ge[95], the diatomics of Si[32] and Cl[32], and the reaction $SiH_4 \rightarrow SiH_2 + H_2$. However, none of these studies are ECP-QMC studies. Based on the success of the ECP-QMC method for two-valence electron systems, one would expect to see many such studies, however, the ECP-QMC method as described earlier gives poor results on multi-electron systems. The problem is that in the local approximation to the non-local ECP operator, use of standard small SCF functions results in large variance in the generated local

pseudopotential. The large variance comes from the fact that the nodes of the trial wavefunction induce singularities into the local ECP. These singularities can be both positive and negative. For the exact solution of the valence Hamiltonian these singularities would be cancelled by the kinetic energy. However, for the simple Ψ_{val} used, this cancellation is very poor, causing the local ECP potential to oscillate rapidly. In fact, this behavior can be observed in a VMC-ECP run, for which the local ECP is exactly equivalent to the non-local U^{ECP}. Use of more accurate trial functions (*e.g.* small CAS-SCF) suggested by Christiansen [93] improves the behavior of the variational walk, but only small gains were found for the fixed-node. Since such singularities in the kinetic energy also exist in the all-electron calculation one may wonder why they present such problems here. In the all-electron case, however, the core electrons would "bump" out such an electron either by their potential, or by moving the nodes of the trial wavefunction. In fact, for singularities in the valence region this does occur, but when the offending electron is in the core, and is spacially separated from the rest, it becomes more difficult to "bump".

Normally, a large variance would simply require more computation time, however, in this case it leads to collapse of the fixed-node solution into the core. The cause of this collapse can be seen in the difference between $U_L^{ECP}[\Phi\Psi_{val}]$ and $U_L^{ECP}[\Psi_{val}^2]$. In the variational walk (as well as in standard *ab initio* techniques) $U_L^{ECP}[\Psi_{val}^2]$ is interchangeable with the non-local ECP, since Ψ_{val} is known analytically. But in the fixed-node walk, the branching causes the distribution to change from Ψ_{val}^2. If $U_L^{ECP}[\Phi\Psi_{val}]$ were used, then the local potential would change to maintain core-valence orthogonality. Neither $U_L^{ECP}[\Psi_{val}^2]$ nor $U_L^{ECP}[\Psi_{val}]$ (since it does not contain *global* information about Φ) have this feedback effect. This problem can be overcome by the use of the DCGF to prevent the QMC solution from changing from the VMC solution in the core region, but the ECP-QMC procedure remains problematic for more than 2 electrons.

In principle the MP-QMC will also suffer from the same problems as ECP-QMC, but

since the projection operators are zero in the local approximation, the calculation relies on the smooth V^{MP} potential and the nodes of the trial function (which are the same as in the all-electron case) to prevent core collapse. This is analogous to the excited state work of Grimes, *et al.*[13], where the fixed-node approximation prevented any appreciable collapse of the excited state solution to the ground state. The most convincing proof, however, is the Cl atom EA calculation of Yoshida and Iguchi[94] in which they obtained 3.617(198) eV which reproduces the experimental value of 3.615 eV [107] to within their error bars. This is in contrast to the PH result for the EA of Cl (3.76(11) eV [32]) which is in error by 0.14(11) eV, although the error bars of the two calculations overlap to a great extent.

Table 3.4 presents calculations of the IP and EA of C, Si, and Ge. Good agreement with experiment is observed for both DC and for PH (for Si). In the DC method note that adding valence electrons has a greater effect on the variance of the energy than adding core electrons, as evidenced by comparing the error bars of the ionization potential of Li, Na, and C.

The DC results were obtained in exactly the same manner as those for Li and Na, with basis sets and Jastrow parameters given in Tables 3.6 and 3.7. Note that while the value of b (which is inversely related to the correlation range) for the valence electrons is about the same for all the systems, for the core electrons' b increases with the nuclear charge. This bears out the conclusions of Sun, *et al.*[108], that b should depend on the electronic density.

Table 3.5 presents data from representative runs of C, Si, and Ge, for roughly equal amounts of computation, by which the performance of the DC method can be analyzed. The most important finding, of course, is that for each atom the statistical uncertainty of the valence energy is reduced significantly with respect to the core energy. Thus in the case of C atom, only modest computation time savings are achieved in comparison to all-electron QMC. In the cases of Si and Ge, however, the statistical uncertainty has been

reduced by a factor of 50 to 100, achieving variances in the valence energies which would be impossible by all-electron QMC. Although this difference is in part due to the relative sizes of the core and valence ensembles, it is offset by the greater efficiency gained by using core time steps 10 to 100 times larger than in all-electron fixed-node QMC. In fact, for the heavy atom cases, use of large time steps becomes the dominant factor in computational speed. The smallest core time step used here was 0.01 hartree^{-1}. This can be compared to the recent work of Barnett, *et al.*,[12] on F and F^- in which the smallest time step was 0.0005 hartree^{-1}. Using this time step for F, and the Z^{-2} dependence of the time scale mentioned above, one can expect that a corresponding time step no larger than 0.00004 hartree^{-1} could be used for Ge in an all-electron computation. This represents a relative speed increase of 5000 over an all-electron QMC simulation.

It is also found that as the valence radius increases, larger valence time steps can be used, as was anticipated in Ref. 15. There are two important caveats, however, in choosing time steps. First, in order for the valence walk to reach equilibrium, the core time step must be smaller than the smallest valence time step to be used. Empirically a value of one fifth to one half the smallest valence time step was found to be adequate. Second, since the valence trial function has a relatively large number of nodes, excessively large time steps can lead to instabilities and inaccuracies due to nodal cross-recross and trapping effects [64].

Preliminary results on the $SiH_4 \rightarrow SiH_2 + H_2$ reaction are given in Table 3.8. Excellent agreement is found between the DC binding energy of SiH_4 and other theoretical results, and the error bar on the DC energy is very low (2.0 kcal/mole). Results on the SiH_2 molecule are not as satisfactory. The DC-computed binding energy is about 20 kcal /mole to small. At this point it is not clear why this discrepancy occurs, although it could be due to errors in time step extrapolation, or perhaps the SiH_2 molecule is not adequately described by the trial function used.

As was mentioned in Sec. 3.4, correlated sampling can be used in the DC method by

using the same set of core configurations. Results of using correlated sampling in DC are shown in Table 3.9 to compute the ionization potential of C and Si. The core configurations were generated by a VMC simulation of the quadruple ions C^{4+} and Si^{4+}, and stored. Then these configurations were used for the core walk in the DC simulation for both the neutral atom and the cation. All calculations were done at a single time step of 0.1 hartree^{-1}. For C atom, the correlated sampling procedure did not result in significant reduction of the error bar on the ionization potential. On the other hand, for Si atom, correlated sampling lowered the error bar of the ionization potential by roughly a factor of 2. These results indicate that the cancellation of fluctuations is more effective with a greater number of core electrons. Thus, one can expect a still larger reduction in the error bar for Ge atom.

3.5.3 Composite DC-pseudopotential QMC

In Table 3.10 a comparison of preliminary results on the Fe atom for various QMC treatments is presented. Five cases are considered, Fe as an 8 valence-electron atom with both ECP-QMC and DC, Fe as a 16 valence-electron atom with ECP-QMC and DC, and for the 16 valence-electron case, a hybrid ECP-DC method was implemented, damping the 3s and 3p electrons as suggested in Sec. 3.4. The ECP-QMC procedure used here is the same as that described in Ref. 15, with two modifications: the local potential was described by $H_L^{ECP}[\Psi_{val}^2]$, and the DCGF was used to prevent collapse of the valence distribution into the core.

The ECP's used were taken from Stevens, *et al.*[83] for the 16 valence-electron case, and Hay and Wadt [109] for the 8 valence-electron case. It can be seen for both ECP and DC treatments that large variances result in the 8 valence-electron case. The reason for this is that the "valence" 3d electrons interact strongly with the "core" 3s and 3p electrons. This is indirect confirmation of the currently held view in standard ECP methods

that the 3s, 3p, and 3d electrons overlap too much to allow separation between these shells. The best method appears to be DC with 16 valence electrons, which has both the lowest variance, and is in agreement with the SCF valence energy for 16 electrons. The ECP-QMC method with 16 valence electrons also has a good variance, but the overly low energy compared to the SCF indicates collapse into the core, and the DCGF parameters would have to be carefully adjusted to give meaningful results. The method with the largest variance is in fact the composite DC/ECP method, which instead of gaining the best of both methods, seems in this particular case to have gotten the worst of both methods.

3.6 Conclusions and future directions

Valence QMC methods have already allowed the study of atoms and molecules once out the reach of all-electron QMC. While the above results are promising, much work remains to be done. In the case of the ECP-QMC method, a great deal of work is needed, for as of yet ECP-QMC has not been demonstrated to be effective beyond two-valence electron systems. On the other hand, the MP-QMC approach has been successfully applied to the Cl and Cl^- atoms, and looks to be the most fruitful pseudopotential method of the two presented here. The PH method has also shown its effectiveness on many-electron atoms. However, in this case much work remains in the parametrization of the pseudo-Hamiltonian.

All these methods, ECP, MP, and PH, are frozen core. Thus the DC method offers the added capability to treat core-valence interactions in a systematic manner. In addition, no parametrization is needed of either pseudopotentials of pseudo-Hamiltonians. Such parametrizations took on the order of a decade in the case of the ECP method. The only disadvantage of the DC method with respect to the ECP-QMC, MP-QMC, and PH methods is that the added accuracy requires more computation.

In all of the methods described here, the remaining challenge is that the core-valence separation is not as well understood as will be needed for routine use of these methods. Much more experience and research is required to bring valence QMC methods to the status of standard valence methods have achieved over the past 15 to 20 years. However, the heavy atom problem is clearly on the way to being solved.

Table 3.1. Electron affinities and ionization potentials of one- and two-valence-electron atoms. Energies in eV. Uncertainties in parentheses.

Atom	Quantity	Analytic		QMC				Exp.
		HF[a]	CI[b]	MP[c]	ECP	PH[g]	DC[h]	
Li	I.P.	5.342		-	5.412(8)	-	5.386(8)	5.392[i]
	E.A.	-0.122	0.615	-	0.611(20)[d]	-	0.600(27)	0.620(7)[j]
					0.557(16)[e]			
Na	I.P.	4.947		-	4.967(4)	5.309(3)	5.120(22)	5.139[i]
	E.A.	-0.116	0.539	-	0.555(21)[d]	0.555(19)	0.569(54)	0.546(5)[j]
Mg	I.P	6.613	7.524	7.571(32)	7.637(26)[d]	7.753(8)	-	7.646[i]
					7.546(8)[f]			

(a) E. Clementi, A. D. McLean, Phys. Rev. **133**, A419 and A1274 (1964).

(b) A. W. Weiss, Phys. Rev. **166**, 70, (1962).

(c) T. Yoshida, K. Iguchi, J. Chem. Phys. **88**, 1032 (1988).

(d) B. L. Hammond, P. J. Reynolds, W. A. Lester, Jr., J. Chem. Phys. **87**, 1130. (1987)

(e) M. M. Hurley, P. A. Christiansen, J. Chem. Phys. **86**, 1069 (1987).

(f) P. A. Christiansen, L. A. LaJohn, Chem. Phys. Lett. **146**, 162 (1988).

(g) G. B. Bachelet, D. M. Ceperley, M. G. B. Chiocchetti, preprint.

(h) present work.

(i) H. Hotop, W. C. Lineberger, J. Phys. Ref. Data. **4**, 539 (1975).

(j) C. E. Moore, *National Standard Reference Data Series*, Vol. 34, (U.S. Government Printing Office, Washington D.C., 1970).

Table 3.2. Binding energy for Na_2

	Method	D_e (eV)
All electron	HF[a]	-0.022
	MCSCF[b]	0.732
Valence electron	CI[c]	0.73
	CVCI[d]	0.747
	PH-QMC[e]	0.844(16)
	ECP-QMC[f]	0.746(20)
Experiment	56 vib. spect.[g]	0.7469(7)

(a) Near Hartree-Fock limit, R. L. Martin, E. R. Davidson, Mol. Phys. **35**, 1713 (1978).

(b) Multi-configuration SCF, W. J. Stevens, M. M. Hessel, P. J. Bertoncini, A. C. Wahl, J. Chem. Phys. **66**, 1477 (1977).

(c) Configuration interaction, A. Valance, Q. N. Tuan, J. Phys. B **15**, 17 (1982).

(d) of the core-valance interaction, G. Jeung, J. Phys. B **16**, 4289 (1983).

(e) Pseudo-Hamiltonian, G. B. Bachelet, D. M. Ceperley, M. G. B. Chiocchetti, preprint.

(f) ECP-QMC described in Sec. 3.3.1, B. L. Hammond, P. J. Reynolds, W. A. Lester, Jr., J. Chem. Phys. **87**, 1130 (1987).

(g) RKR analysis of vibrational spectrum using levels through v''=56, K. K. Verma, J. T. Bahns, A. R. Rajaei-Rizi, W. C. Stwalley, W. T. Zemke, J. Chem. Phys. **78**, 3599 (1983).

Table 3.3. Binding energy for NaH

	Method	D_e(eV)
All electron	HF[a]	0.932
	CI[b]	1.922
Valence electron	MCSCF[c]	1.885
	ECP-QMC[d]	1.954(73)
Experiment	11 vib. spect.[e]	1.9714

(a) Near Hartree-Fock limit, P. E. Cade, W. M. Huo, J. Chem. Phys. **47**, 649 (1967).

(b) Configuration interaction, R. E. Olson, B. Liu, J. Chem. Phys. **73**, 2817 (1980).

(c) Pseudopotential multi-configuration SCF, W. J. Stevens, A. M. Karo, J. R. Hiskes, J. Chem. Phys. **74**, 3989 (1981).

(d) ECP-QMC described in Sec. 3.3.1, B. L. Hammond, P. J. Reynolds, W. A. Lester, Jr., J. Chem. Phys. **87**, 1130 (1987).

(e) RKR analysis of vibrational spectrum using levels through $v''=11$, W. T. Zemke, R. E. Olson, K. K. Verma, W. C. Stwalley, B. Liu, J. Chem. Phys. **80**, 356 (1984).

Table 3.4. Electron affinities and ionization potentials of C, Si, and Ge. Energies in eV.

Atom	Quantity	Analytic		QMC		Exp.
		HF[a]	CI	PH[f]	DC[g]	
C	I.P.	10.786		-	11.2(2)	11.26[h]
	E.A.	0.549	1.14[b]	-	1.3(2)	1.268(5)[i]
Si	I.P.	7.658	8.083[c]	8.19(3)	8.3(2)	8.151[h]
	E.A.	0.956	1.32[d]	1.39(3)	1.3(2)	1.385(5)[i]
Ge	I.P	7.434	7.68[e]	-	8.0(3)	7.899[h]
	E.A.	0.950	1.12[e]	-	1.3(3)	1.233(3)[i]

(a) E. Clementi, A. D. McLean, Phys. Rev. 133, A419 and A1274 (1964).

(b) J. Kalcher, R. Janoschek, Chem. Phys. 104, 251 (1986).

(c) L. A. Curtiss, J. A. Pople, Chem. Phys. Lett. 144, 38 (1989).

(d) M. Lewerenz, P. J. Bruna, S. D. Peyerimhoff, R. J. Buenker, J. Phys. B. 16, 4511 (1983).

(e) L. G. M. Pettersson, P. E. M. Siegbahn, Chem. Phys. 105, 355 (1986).

(f) G. B. Bachelet, D. M. Ceperley, M. G. B. Chiocchetti, preprint.

(g) B. L. Hammond, P. J. Reynolds, W. A. Lester, Jr., "Damped core", submitted.

(h) H. Hotop, W. C. Lineberger, J. Phys. Ref. Data. 4, 539 (1975).

(i) C. E. Moore, *National Standard Reference Data Series*, Vol. 34, (U.S. Government Printing Office, Washington D.C., 1970).

(j) T. M. Miller, A. E. S. Miller, W. C. Lineberger, Phys. Rev. A 33, 3558 (1986).

Table 3.5. Damped-core QMC energies and parameters for C, Si, and Ge, for equal computation. Quantities are in atomic units. N is the number of electrons, E the energy, and r the rms radius obtained in the simulations.

	Carbon	Silicon	Germanium
N_{val}	4	4	4
N_{core}	2	10	28
E_{val}	-5.397(2)	-3.815(7)	-3.798(27)
E_{core}	-32.418(10)	-285.2(5)	-2070.(8)
r_{val}	1.95(1)	2.77(3)	2.88(8)
r_{core}	0.312(8)	0.547(14)	0.598(26)

Table 3.6. STF basis set exponents used in Damped Core QMC for Li, Na, C, Si, and Ge. The principle quantum number is n, and l is the angular momentum quantum number.

Atom	l	n			
		1	2	3	4
Li	S	4.6170 2.4620	1.9630 0.6700		
Na	S	13.1474 9.7154	3.9098 2.6039	1.2594 0.75485	
	P		5.4964 2.5713		
C	S	7.9690 5.2310	1.8200 1.1680		
	P		2.7260 1.2560		
Si	S	14.5372 9.6094	6.0540 4.1338	2.1000 1.2000	
	P		7.8130 4.0852	1.8626 1.0934	
Ge	S	36.3802 30.7540	26.6899	12.7300 7.8067 5.2103	2.8423 1.7747
	P		13.2000	8.0305 5.2052	3.7000 2.1400 1.2600
	D			6.0456 3.7850	

Table 3.7. Jastrow parameters[a] used for Li, Na, C, Si, and Ge.

Atom	b_{val}	b_{core}	λ	ν
Li	2.0	5.0	0.0125	0.0800
Na	2.0	11.0	0.0250	0.125
C	1.0	5.0	0.0250	0.150
Si	1.0	10.0	0.0125	0.0625
Ge	1.0	20.0	0.0125	0.0750

(a) Linear form of Eq. 2.42, a_1 is chosen as 0.5 in all cases to satisfy the electron-electron cusp condition, $b_{c,\nu}$ is b_1, λ and ν correspond to a_1 and b_1 for the electron-nuclear correlation function.

Table 3.8. Relative energies of the $SiH_4 \rightarrow SiH_2+H_2$ reaction. Energies in kcal.

Quantity	Previous Theory	DC
SiH_4 Binding Energy	304.4[a] 303.4[b] 306.8[c]	306.5(2.0)
SiH_2 Binding Energy	140.1[a] 147.2[b] 142.9[c]	120(8)
$SiH_4 \rightarrow SiH_2+H_2$ Reaction Energy	59.9[a] 59.3[d]	77(8)
TST Energy[e]	56.9[d]	-

(a) P. Ho, M. E. Coltrin, J. S. Binkley, C. F. Melius, J. Phys. Chem. **89**, 4647 (1985).

(b) D. S. Horowitz, W. A. Goddard III, Theochem **163**, 207 (1988).

(c) L. A. Curtiss, J. A. Pople, Chem. Phys. Lett. **144**, 38 (1989).

(d) M. S. Gordon, D. R. Gano, J. S. Binkley, M. J. Frisch, J. Am. Chem. Soc. **108**, 2191 (1986).

(e) Energy of transition state geometry given in (d) relative to SiH_4.

Table 3.9. Ionization potentials (IP) computed with and without correlated sampling (CS) for C and Si atoms[a].

Atom	IP without CS	IP with CS
C	11.07(20)	11.07(15)
Si	8.09(22)	8.09(10)

(a) Values are not time-step extrapolated. Valence time step used was 0.1 hartree^{-1}, and the core time step was 0.02 hartree^{-1} in all cases. The values with and without CS were generated in the same run, thus the values of the IP are trivially the same.

Table 3.10. QMC results for Fe atom.

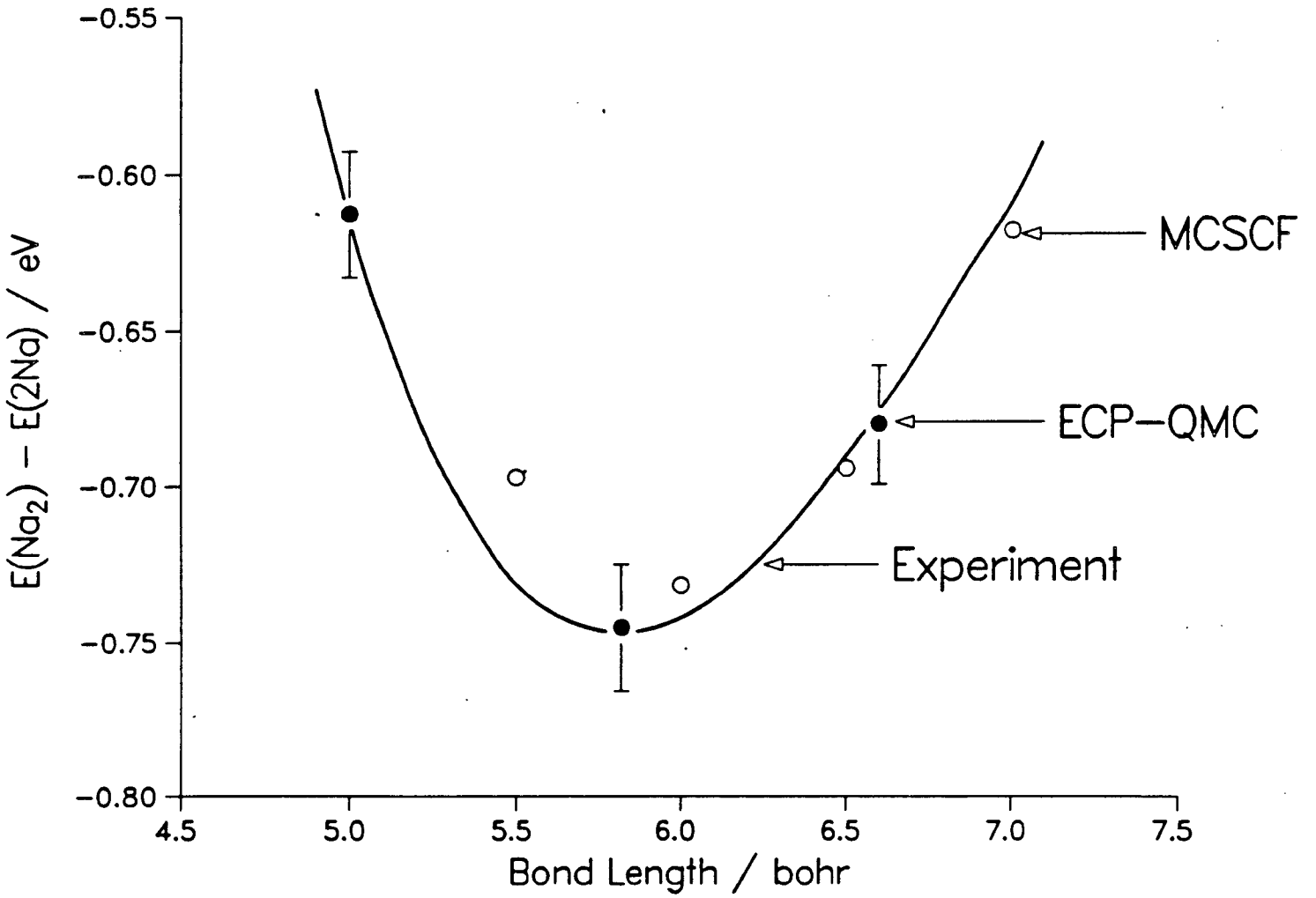
Core	No. Valence Electrons	Energy[a]	CPU[b]	SCF energy	Scaled error[c]
ECP	8	-22.64(20)	987	-21.5493	0.200
ECP	16	-131.646(75)	1779	-122.2233	0.100
DC	8	-21.56(38)	1276	-	0.430
DC	16	-122.781(57)	2363	-	0.088
DC/ECP	8/16	24.22(40)	1543	-	0.500

(a) FNDQMC energy using a valence time step of 0.01 hartree^{-1}. For the DC results, a core time step of 0.002 hartree^{-1} was used.

(b) CPU time in CRAY XMP seconds.

(c) Statistical error of the energy quoted in column 3, scaled to represent 1000 seconds of CPU time.

Figure 3.1. Interaction energy of Na_2 molecule as a function of internuclear distance. Solid circles, ECP-QMC [a]; open circles, all-electron MCSCF[b]; solid line, spline fit to experimental points [c].

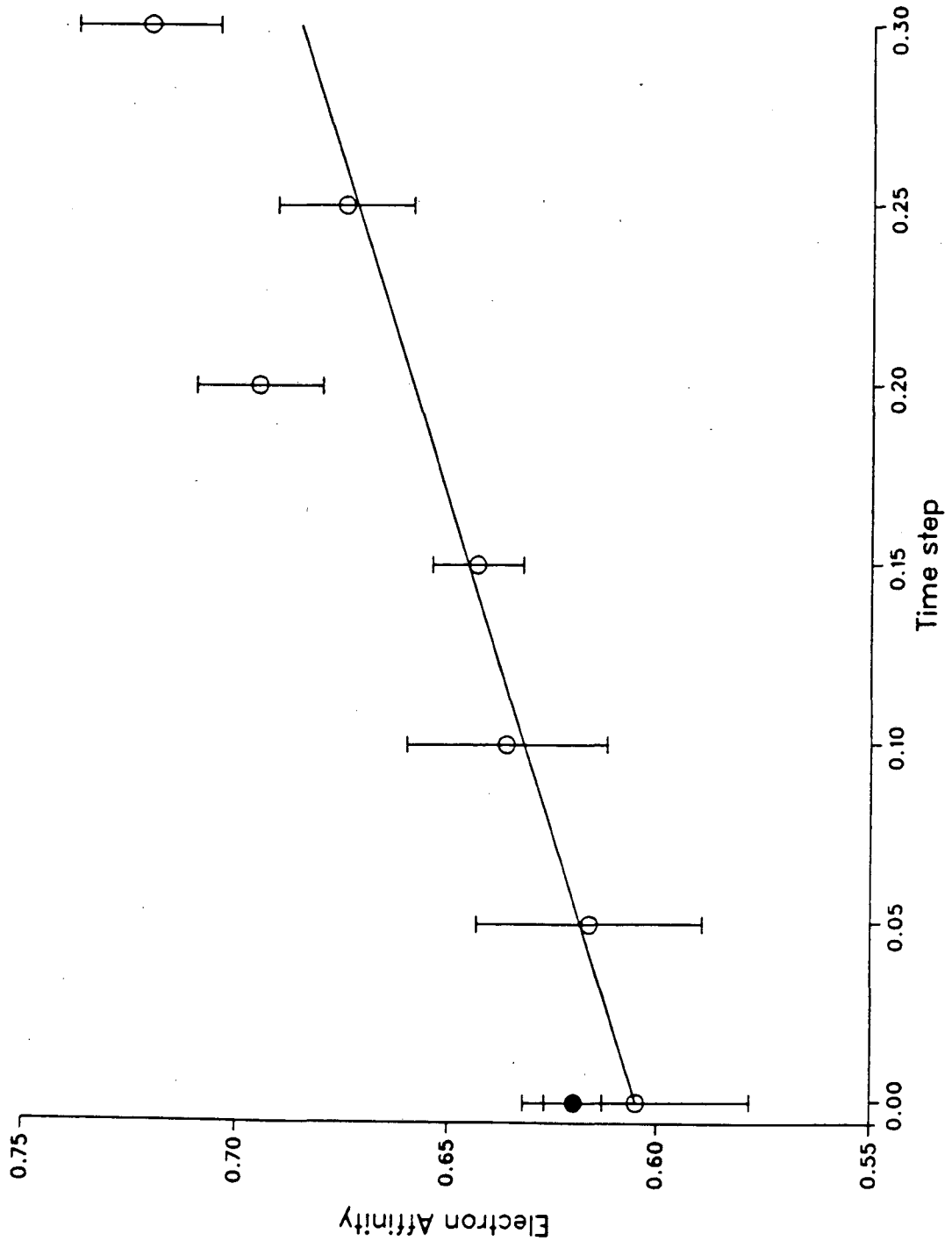


(a) B. L. Hammond, P. J. Reynolds, W. A. Lester, Jr., J. Chem. Phys. **87**, 1130. (1987).

(b) W. J. Stevens, M. M. Hessel, P. J. Bertoncini, A. C. Wahl, J. Chem. Phys. **66**, 1477 (1977).

(c) K. K. Verma, J. T. Bahns, A. R. Rajaei-Rizi, W. C. Stwalley, W. T. Zemke, J. Chem. Phys. **78**, 3599 (1983).

Figure 3.2. Time step bias curve of the electron affinity of Li atom using the DC method. The line represents a linear fit to the data from time steps 0.050, 0.100, and 0.150. The filled circle represents the experimental value. The electron affinity is given in units of eV, and the time step is given in hartree^{-1}. The y-axis has been displaced for clarity.



Chapter 4

Derivatives of the quantum Monte Carlo Energy

4.1 Background

Most often in quantum chemistry one requires more information than a single energy at a given molecular geometry. Of particular interest are the calculation of equilibrium and transition-state geometries, and vibrational spectra (both infrared and Raman). Such information can be obtained by mapping out the potential-energy surface (PES) of a molecule in the region of interest; however, following the work of Pulay [110] a great deal of interest has been focused on the use of gradient methods. In particular, first and second derivatives of the energy with respect to the nuclear positions are powerful tools for characterizing PES's and determining force constants and harmonic vibrational frequencies. Pulay stressed the inherent accuracy and efficiency of *analytic* gradients, that is, derivative expressions calculated directly from the energy expression for the method being used, as opposed to finite difference methods, or use of the Hellmann-Feynman theorem. Since that time analytic energy derivatives have become a mainstay of quantum chemists, and have been implemented in nearly all commonly used *ab initio* computer programs. For an excellent review of gradient methods see the recent chapter by Pulay [111].

In QMC, however, energy derivatives have not yet been exploited. This is because the three methods for calculating derivatives — finite differences, the Hellmann-Feynman theorem (HFT), and analytic derivatives — all have major difficulties to be overcome in QMC. In the case of finite differences the difficulty is one of the QMC statistical uncertainty. Using a two-point finite difference formula for the first derivative, the statistical uncertainty of the derivative, ε_{deriv}, is related to the statistical uncertainty

of the energy, ε_{energy}, by,

$$\varepsilon_{deriv} \sim \sqrt{2}\frac{\varepsilon_{energy}}{\delta R} \tag{4.1}$$

where δR is the step size in the difference scheme. Typical values for δR are in the range 0.01 to 0.05, thus the error of the derivative will be 20 to 100 times that of the energy, making this process useless.

In the case of the HFT, it is known to give poor results in standard *ab initio* calculations [111]. For QMC, it is shown in Sec. 4.2 that the variance of the HFT estimator is unbounded, and consequently the HFT estimator cannot be obtained directly. In addition the true HFT estimator must be sampled from the Φ_0^2 distribution, rather than the usual $\Phi_0\Psi_T$ distribution.

Finally, in the case of analytic derivatives, which have been so successful in standard methods, QMC is presented with two problems. The first is the formal necessity of differentiating Φ_0, which is not known analytically. The second is that the resulting expressions are cumbersome, and the calculation of the analytic derivatives can significantly increase the required computation time.

In this chapter the various methods available to QMC for calculation of derivatives and energy minimization are explored. The Hellmann-Feynman theorem is discussed in Sec. 4.2., analytic gradients in Sec. 4.3., and finite differences in Sec. 4.4. In each case the above issues are addressed and example results are given. Finally, in Sec. 4.5 future directions, such a second derivatives, are discussed.

4.2 The Hellmann-Feynman theorem

The HFT estimator of the energy gradient is given by the average of the gradient of the potential;

$$\partial E/\partial X_A = \int d\mathbf{R}\Psi_T \frac{\partial V}{\partial X_A}\Psi_T . \tag{4.2}$$

(Throughout this chapter upper case variables refer to nuclei and lower case to electrons; X and x refer to any Cartesian coordinate). As stated above the HFT is not commonly used in standard methods, although some recent progress has been made [112].

The reason for the failure of the HFT can be seen readily from the derivation of Eq. 4.2. The first step is to write the derivative of the variational energy expectation value (Eq. 1.3) with respect to X_A,

$$\frac{\partial <E>}{\partial X_A} = \frac{\partial}{\partial X_A}\int d\mathbf{R}\ \Psi_T \mathbf{H}\Psi_T \tag{4.3}$$

$$= \int d\mathbf{R}\Psi_T\left[\frac{\partial \mathbf{H}}{\partial X_A}\right]\Psi_T + \int d\mathbf{R}\left[\frac{\partial \Psi_T}{\partial X_A}\right]\mathbf{H}\Psi_T + \int d\mathbf{R}\Psi_T \mathbf{H}\left[\frac{\partial \Psi_T}{\partial X_A}\right].$$

Since **H** is a Hermitian operator it can act either to the left or right, so that the last two terms of Eq. 4.3 combine to become

$$2\int d\mathbf{R}\left[\frac{\partial \Psi_T}{\partial X_A}\right]\mathbf{H}\Psi_T . \tag{4.4}$$

If Ψ_T is an eigenfunction of **H** then $\mathbf{H}\Psi_T = E\Psi_T$, making Eq. 4.4

$$2E\int d\mathbf{R}\left[\frac{\partial \Psi_T}{\partial X_A}\right]\Psi_T = E\frac{\partial}{\partial X_A}\int d\mathbf{R}\Psi_T^2. \tag{4.5}$$

The integral of Ψ_T^2 will be a constant for any eigenfunction so that Eq. 4.5 is equal to zero. Eq. 4.3 then reduces to Eq. 4.2 by noting that the Hamiltonian is a sum of kinetic and potential energies, but since the kinetic energy operator does not depend on the nuclear positions its gradient is zero. Note that the critical assumption in the derivation is that Ψ_T is an eigenfunction of the Hamiltonian. However, this need not be the exact Hamiltonian. Thus the exact Hartree-Fock solution (*i.e.* the Hartree-Fock limit) will satisfy the above assumption [114]. The problem with the HFT gradient is illustrated in Table 4.1, which presents the SCF energy, HFT derivative, and analytic derivative for H_2^+ for STF basis sets, as well as the exact value[113]. Clearly the HFT derivative is inferior

to the analytic derivative even for very large basis sets. The best agreement for small basis sets between HFT and analytic derivatives is in in the final two basis sets, designated (1/1)* and (1/1/1)*, which consisted of a 1S/2P or a 1S/2P/3D basis with a single exponent shared by all the STF's. These basis sets are similar to those recommended in Ref. 112, where it was proved that the Hellmann-Feynman theorem is satisfied for basis sets which include all the derivatives of the basis functions, while in practice they used only the first derivative. For GTF's this is a simple procedure since the derivative of a 1S function is a 2P, and so on. While this is not strictly true for STF's, inclusion of the 2P and 3D shared exponent functions has better agreement between HFT and analytic derivatives than does the largest standard basis set.

In spite of the above difficulties the HFT is of interest in QMC because of its simplicity. In addition to energy gradients, the HFT can also be used to judge the quality of the trial function, since the sum of the HFT forces should go to zero for an accurate trial function [114] (even if the forces themselves are not accurate). Unfortunately, although the quantity $\langle\Phi_0|d\mathbf{H}/dX_A|\Phi_0\rangle$ may be a good estimator of the energy derivative, in standard QMC it is the mixed estimator, $\langle\Phi_0|\partial\mathbf{H}/\partial X_A|\Psi_T\rangle$, which is actually evaluated. Unless Ψ_T is also very close to the exact wavefunction, Eq. 4.4 will not hold. Thus the first difficulty in using the HFT derivatives is that the pure estimator (using the future walking method) must be used, although the second order approximation (Eq. 2.26) has given good results for some properties [11], and may prove to be useful here.

A second difficulty arises when one evaluates the variance of the HFT estimator. Expanding Eq. 4.2 in terms of the potential energy, one obtains

$$\frac{\partial \mathbf{V}}{\partial X_i} = \sum_A \left[\sum_B \frac{Z_A Z_B}{R_{AB}^3}(X_B - X_A) - \sum_i \frac{Z_A}{r_{iA}^3}(x_{iA}) \right] . \tag{4.6}$$

The first part is the derivative due to the nuclei, and is trivial to evaluate. The second quantity, the derivative of the electron-nuclear potential, is straightforward to evaluate by QMC, but its variance is another matter. Consider the H atom with $\Psi_T = e^{-ar}$. Even

though the HFT derivative is trivially zero, the *variance* of the HFT derivative in the x direction will be

$$\varepsilon^2 = \left\langle \frac{\hat{x}^2}{r^4} \right\rangle - \left\langle \frac{\hat{x}}{r^2} \right\rangle^2 \tag{4.7}$$

where,

$$\left\langle \frac{\hat{x}^2}{r^4} \right\rangle = \int d\mathbf{R}\, e^{-2ar}\, \frac{\hat{x}^2}{r^4}, \tag{4.8}$$

and $\hat{x}$ is the angular part of the cartesian coordinate x (*i.e.* x/r). Eq. 4.8 can be evaluated in polar coordinates,

$$\left\langle \frac{\hat{x}^2}{r^4} \right\rangle = \int d\Omega \hat{x}^2 \int_0^\infty dr\, r^2\, e^{-2ar}\, \left(\frac{1}{r^4}\right) \tag{4.9}$$

the angular integral is equal to $4\pi/3$ (See Appendix C, Eq. C.14), however the radial part will be infinite (except for cases when Ψ_T^2 goes to zero quadradically at the origin). During the QMC simulation this infinite variance will effect the HFT estimator in the form of infrequent, but large, "spikes" in the derivative, such that no reliable estimate can be obtained.

A scheme for suppressing these spikes and approximately evaluating the Hellmann-Feynman derivative can be derived by noting that sufficiently near the nucleus the charge distribution will be spherical. By partitioning the expectation value of the HFT derivative into a region within a sphere of radius ρ, and a region ω, which is the rest of space, one obtains

$$\int d\mathbf{R}\Psi_T^2 \frac{\partial V}{\partial X_i} = \int_\rho d\mathbf{R}\Psi_T^2 \frac{\partial V}{\partial X_i} + \int_\omega d\mathbf{R}\Psi_T^2 \frac{\partial V}{\partial X_i}. \tag{4.10}$$

The integral over ω can be evaluated by QMC, and the integral over the region contained within ρ is approximate to be zero. The result can either be extrapolated to ρ=0, or, as is done here, ρ is made sufficiently small that the error is below the statistical error of the QMC evaluation.

This method has been tested for H_2^+ and LiH. First the HFT derivative was evaluated

analytically for an SCF function, and then by VMC for the same trial function with different values of ρ. The results are presented in Table 4.2 for H_2 and in Table 4.3 for LiH. A sufficiently small value of ρ for H is found to be 0.05. For Li atom the value 0.15 is found to be sufficiently small. The next question is whether or not the exact derivative can be obtained from the HFT using QMC. With the ρ values found above, the HFT derivative was calculated by FNDQMC with the short time approximation. Table 4.4 present results for H_2^+ at various time steps, as well as the exact result of H. Wind [113]. The extrapolated value of the HFT (0.526(4)) is in very good agreement with the exact value (0.5207). Results for LiH are presented in Table 4.5 at a bond length of 2.5 bohr. Here the results are not as conclusive as in H_2^+, however the HFT estimator systematically improves from the SCF level to the approximate Φ^2 result. The biggest problem with the HFT estimator is its variance. Even for H_2^+ significant amounts of computation were required, and for LiH, the variance is so large that the result is useless.

4.3 Analytic energy derivatives

Analytic energy derivative methods have become ubiquitous in *ab initio* computations. The reason for this is two-fold. First, as seen in Table 4.1, the HFT derivative is not reliable for SCF wavefunctions. Second, analytic derivatives can be evaluated at little additional cost to the energy calculation, and are accurate for second and third derivatives, where finite difference schemes would be very costly, and of dubious reliability. A key difference between standard methods and QMC, however, is that in standard methods most of the work is expended in diagonalization of the Hamiltonian. Once this is done, properties involving the wavefunction are relatively inexpensive to compute. For QMC, all properties must be sampled during the walk, and therefore complicated expressions slow the calculation significantly. One way to speed computation is to sample only the energy at every step of the walk, and to sample other properties at less frequent

intervals[16]. Even though this results in a smaller sample size, the samples taken will be less correlated, and thus more efficiently placed.

Starting with the QMC energy estimator,

$$E_{QMC} = \langle E_L \rangle_{\Phi_0\Psi_T} = \int d\mathbf{R}\, \Phi_0 E_L\, \Psi_T. \tag{4.11}$$

The derivative of this expression is

$$\frac{\partial \langle E_L \rangle}{\partial X_A} = \langle \frac{\partial E_L}{\partial X_A} \rangle + \langle E_L \Psi_T^{-1} \frac{\partial \Psi_T}{\partial X_A} \rangle - \langle E_L \rangle \langle \Psi_T^{-1} \frac{\partial \Psi_T}{\partial X_A} \rangle$$

$$+ \langle E_L \Phi^{-1} \frac{\partial \Phi}{\partial X_A} \rangle - \langle E_L \rangle \langle \Phi^{-1} \frac{\partial \Phi}{\partial X_A} \rangle . \tag{4.12}$$

(The subscript $\Phi_0\Psi_T$ will be assumed for all averages unless otherwise stated.) The derivatives of Φ_0 can be evaluated using the relationship

$$|\Phi_0\rangle = \lim_{\tau\to\infty} e^{-\tau \mathbf{H}} |\Psi_T\rangle . \tag{4.13}$$

Substituting Eq. 4.13 into Eq. 4.12 the analytic QMC energy derivative is obtained,

$$\frac{\partial \langle E_L \rangle}{\partial X_A} = \langle \frac{\partial E_L}{\partial X_A} \rangle + 2\langle E_L \Psi_T^{-1} \frac{\partial \Psi_T}{\partial X_A} \rangle - 2\langle E_L \rangle \langle \Psi_T^{-1} \frac{\partial \Psi_T}{\partial X_A} \rangle$$

$$+ \langle E_L M^{-1} \frac{\partial M}{\partial X_A} \rangle - \langle E_L \rangle \langle M^{-1} \frac{\partial M}{\partial X_A} \rangle , \tag{4.14}$$

where M is given by

$$M(\tau) = e^{-\int_0^\tau d\tau' (E_L - E_T)} , \tag{4.15}$$

which is the value of the branching multiplicity integrated over the walk. The difficulty with this expression, is that the future walking method must be used to evaluate the terms involving $M(\tau)$. An alternate approach is to note that for a Ψ_T sufficiently close to Φ_0 the branching is close to 1 at all times, and $M^{-1}\frac{\partial M}{\partial X_A}$=0, leaving the approximate relationship,

$$\frac{\partial \langle E_L \rangle}{\partial X_A} = \langle \frac{\partial E_L}{\partial X_A} \rangle + 2\langle E_L \Psi_T^{-1} \frac{\partial \Psi_T}{\partial X_A} \rangle \tag{4.16}$$

The derivatives of Ψ_T are obtained readily from the analytic form of Ψ_T chosen. The remaining term $\frac{\partial E_L}{\partial X_A}$, however, is more difficult.

Full details of the evaluation of the terms in the analytic derivative are given in Appendix D. However note that the derivative of E_L can be broken into kinetic and potential energy terms, the potential energy term being exactly the same as Eq. 4.6 for the HFT. However, the derivative of the kinetic energy is a two-electron operator,

$$\sum_i \frac{\partial}{\partial x_{iA}} \sum_j \nabla_j^2 \Psi_T \,. \tag{4.17}$$

Such operators are cumbersome to evaluate for determinantal wavefunctions. For one electron operators, $A \equiv \sum_i A_i$, one can use the expression (see Appendix B),

$$\langle A_i \rangle = \sum_k a_{ik} D_{ki}^{-1} \tag{4.18}$$

A_i being the operator acting on electron i, a_{ij} is A_i acting on the molecular orbital k, and D^{-1} is the inverse of the Slater matrix. However, in the present case, with the exception of the diagonal elements of Eq. 4.17, the D_{ki}^{-1} will depend on the coordinates of electron j. Eq. 4.17, then, must be evaluated by first allowing either the operator for electron i or j to act upon the determinant, then calculating the inverse matrix of the result for use in Eq. 4.18. This requires either one matrix inversion per electron or three per atom (see Appendix D). This will be a costly expression to evaluate during the walk. An alternative to this is to evaluate the derivative of the kinetic energy by finite difference. This would require six additional kinetic energy evaluations per atom for the general case, but would not require any Slater matrix inversions. This would be of greatest advantage in systems with many electrons but only a few atoms.

Another important point is to insure that $\partial E_L / \partial X_A$ has a finite variance. Since the derivative of the potential is the same as in the HFT, it has an infinite variance. In this

case, however, the derivative of the kinetic energy can cancel the derivative of the potential energy at the origin if

$$\frac{x_{iA}}{r_{iA}^3}\frac{\partial}{\partial r_i}\Psi_T\Big|_{r_{iA}=0} = -Z_A\frac{x_{iA}}{r_{iA}^3}\Psi_T\,, \tag{4.19}$$

which reduces to the electron-nuclear cusp condition of Sec. 2.8. In addition, the the method of removing the derivative within radius ρ of each nucleus, as was done for the Hellmann-Feynman theorem, can be done here as well. Even if the singularities are not cancelled exactly between the kinetic and potential energy derivatives, one expects that partial cancelation will lead to a reduced variance in comparison to the HFT derivative where only the potential term is present.

In order to use the exact expression, Eq. 4.14, the derivatives of the branching must be evaluated. This can be done by the future walking method [11], or by the time correlation method discussed in Sec. 2.9. Using the time-correlation method one conducts a non-branching (VMC) walk and integrates the weight $M(\tau)$ for several values of τ during the walk. For the derivatives, one must also compute

$$M^{-1}\frac{\partial M}{\partial X_A} = -\int_0^\tau d\tau'\frac{\partial E_L}{\partial X_A}\,. \tag{4.20}$$

The derivative of E_L is already available, so the only additional work required is to keep track of the integral for a set of τ values.

The approximate expression, Eq. 4.16, has been implemented for H_2[11,84], and LiH. Note that in the case of H_2, the derivatives of the kinetic energy are simple. For H_2 the results are given in Table 4.6. One aspect of having derivatives of the energy available is that this information can be used in addition to the energy values when constructing an analytic potential energy curve. Fig. 4.1 shows a simple polynomial fit to the energies in table 4.6, exhibiting spurious oscillatory behavior. On the other hand, when the first derivatives from Table 4.6 are included in the fit, the line in Fig. 4.2 results. This line is visually indistinguishable from the exact curve of Kolos and Wolniewicz [115].

Another important point is that in each case the variance of the analytic derivative was found to be a factor of 2 to 10 smaller than the variance of the HFT derivative. Finally, the effect of the approximation can be seen at R=1.9 bohr, where the HF trial function is relatively poor, and thus the derivative calculated by Eq. 4.16 is off.

Analytic energy derivatives for LiH are given in Table 4.7. Significant improvement of the derivative over the HFT is observed. In addition, the analytic derivative was found to have a variance 2 to 3 times smaller than that of the HFT derivative computed during the same run.

4.4 Finite difference energy derivatives and Correlated Sampling

Finite difference methods approximate the derivative of a function, $f(x_0)$ by,

$$\partial^m f(x_0)/\partial x^m = \sum_\lambda d_\lambda^m f(x_\lambda) \tag{4.21}$$

where the x_λ are typically a set of grid points surrounding x_0. The coefficients d_λ^m can be obtained by solving for the derivatives of f in the Taylor series expansion of f about x_0. As stated in Sec. 4.1, obtaining finite difference energy derivatives in QMC from independent runs is costly because the error of the derivative will be 20 to 100 (or more) times that of the energy.

A more profitable approach is to use the correlated sampling QMC method [79] in which the energy difference is determined directly from a set of correlated walks. For a VMC calculation this can be accomplished by the re-weighting scheme discussed in Sec. 2.9. A VMC walk is conducted in the usual manner guided by a trial function describing the geometry of interest $\Psi_T(X_0)$. The energies of several closely related nuclear geometries are obtained from

$$\langle E(X_\lambda)\rangle = \langle E_\lambda W_\lambda\rangle / \langle W_\lambda^2\rangle \tag{4.22}$$

where,

$$E_\lambda \equiv \frac{H\Psi_T(X_\lambda)}{\Psi_T(X_0)} \tag{4.23}$$

and

$$W_\lambda \equiv \frac{\Psi_T(X_\lambda)}{\Psi_T(X_0)}, \tag{4.24}$$

where the electronic positions are drawn from a VMC walk guided by $\Psi_T(X_0)$. The variance of the energy differences obtained in this manner will be smaller than the variance obtained from independent runs as long as the statistical fluctuations of the E_λ are correlated. Note that while the energy differences may be computed from the average values of $W_\lambda E_\lambda$ and W_λ^2 at the end of the run, the variance must be obtained by computing the differences after each block, or by keeping the ensemble averages separately. The main difficulty here is that W_λ will diverge where the nodes of $\Psi_T(X_\lambda)$ have moved away from the nodes of $\Psi_T(X_0)$. If δR is small enough this may not be a great problem, since the electrons avoid the nodes.

Evaluating the correlated energies in the QMC walk is more complicated due to the branching. In Sec. 2.12 it was stated that correlated energy differences can be obtained in a manner similar to the excited states, namely,

$$\langle E(X_\lambda)\rangle = \lim_{\tau\to\infty} \langle H_\lambda(\tau)\rangle / \langle N_\lambda(\tau)\rangle \tag{4.25}$$

$$= \lim_{\tau\to\infty} \tfrac{1}{2}\langle W_\lambda(\tau) M_{0\to\tau}^{(\lambda)} E_\lambda(0) + W_\lambda(0) M_{0\to\tau}^{(\lambda)} E_\lambda(\tau)\rangle / \langle W_\lambda(\tau) M_{0\to\tau}^{(\lambda)} W_\lambda(0)\rangle$$

Unlike the excited states case, in Eq. 4.25 the branching multiplicity depends upon the value of X_λ. Consequently, separate values of $M_{0\to\tau}^{(\lambda)}$ are computed by Eq. 2.57 for each λ for use in Eq. 4.25. If one desires to use the branching algorithm rather than the VMC to conduct the central walk then rather than using the full value of $M_{0\to\tau}^{(\lambda)}$, the differential systems are weighted by $(M_{0\to\tau}^{(\lambda)} - M_{0\to\tau}^{(0)})$, the latter quantity being the branching multiplicity of the QMC walk. The branching algorithm has been implemented by Wells [79] to compute the dipole moment of LiH in the finite field approximation. A correlated

energy difference of -0.000227(3) was reported, while the energy itself was -8.0595(30) [79]. Hence an energy difference smaller than the statistical error of the total energy was obtained in this fashion.

Traynor and Anderson [80] have recently implemented a related method treating geometrical energy differences in H_3^+. Results similar to those in Ref. 79 were obtained in that energy differences smaller than the statistical error of the total energy were measured to good accuracy. Using Traynor and Anderson's data, the derivative of the energy for the equilateral configuration with the bond distance of 1.65 is either 0.0011(8) using a δR of 0.001, or 0.0016(8) for a δR of 0.05 using a two-point finite difference formula. It is interesting to note that the statistical precision is the same in both cases even though the energy difference in the first case is 500 times smaller than in the second case. This is because the statistical error of the energy differences sampled increases as the difference from the reference system increases. In both the cases treated, dipole moment and H_3^+, there were no problems due to the nodes of the trial functions shifting from those of the reference function. It remains to be seen whether cases with non-trivial nodal behavior can be treated.

Often the gradient itself is not important, rather one wishes to find equilibrium or transition state geometries for which the gradient is zero. For this case, rather than computing the derivative, then moving the coordinates, the information on the surrounding finite difference can be used directly to find the lowest energy point. Such a method has been implemented for the optimization of the trial function parameters in a VMC walk by Huang, *et al.* [78], with good results. For a geometry optimization, use of the above method is relatively straight forward for VMC. However, when optimizing the geometry, one would prefer to use the QMC energy. Here again the problem of changing nodes could make such an optimization very difficult. Using the VMC method, the above procedure has been used to optimize the geometry of H_2^+. The convergence of a run is shown in Table 4.8. In this case, optimization took place on a grid with spacing δR =0.05 bohr.

The energy differences were accumulated over the optimization so that successively better guesses at the minimum were obtained. Again, for this system without nodes, good behavior is obtained. Good agreement with the SCF geometry and derivative are obtained, however, the second derivative is exceptionally noisy, and is of little use.

4.5 Conclusions and future directions

Each of the above methods, HFT, analytic gradients, and finite differences hold great promise in QMC for first derivatives. And, each of the above methods needs a great deal of development. The HFT is attractive because of its simplicity, but provides a relatively noisy estimator. This behavior can be corrected by using analytic derivatives, but the required expressions become cumbersome and costly to evaluate. Correlated sampling allows the use of finite difference methods, but it is not known whether systems with non-trivial nodes can be treated.

A particularly interesting question is whether or not second and higher derivatives can be measured. For both the HFT and analytic derivative methods, the answer is definitely not. This is because for second derivatives one would have to sample the quantity x^2/R^5. Not only are the average values of the quantity and its variance unbounded, but the variance cannot be reduced by either cutting out the origin or by cancelation with the kinetic energy. Thus one must use a finite difference method, and either compute the second (and higher) derivatives directly, or using the analytic of HFT first derivative.

Table 4.1. SCF comparison of Hellmann-Feynman Theorem Derivatives for Hydrogen Molecule Ion. The internuclear separation (R) is 1.0 bohr. Basis sets are designated as (s/p/d), basis sets with a "*" are shared exponent.

Basis set	Energy	dV/dR	dE/dR
(1/0)	-0.44087	0.6158	0.5081
(3/1)	-0.44549	0.5797	0.5150
(5/2)	-0.44954	0.5561	0.5229
(10/2)	-0.45059	0.5441	0.5213
(12/4/1)	-0.45167	0.5214	0.5210
(1/1)*	-0.44902	0.5350	0.5161
(1/1/1)*	-0.45038	0.5135	0.5142
Exact[a]	-0.45179	-0.5207	0.5207

(a) Derivative obtained from an 8'th degree polynomial fit to the data of H. Wind, J. Chem. Phys. 42, 2371 (1965).

Table 4.2. VMC Hellmann-Feynman Theorem Derivatives for Hydrogen Molecule Ion. Internuclear separation (R) is 1.0 bohr. The analytic HFT derivative is -0.6158.

ρ	dV/dR
0.25	-0.628(1)
0.10	-0.620(2)
0.05	-0.618(3)
0.025	-0.619(4)

Table 4.3. VMC Hellmann-Feynman Theorem Derivatives for LiH Internuclear separation (R) is 2.0 bohr, the analytic HFT derivative is 0.0141.

ρ	dV/dR
0.25	-0.028(7)
0.15	0.021(34)

Table 4.4. QMC Hellmann-Feynman Theorem Derivatives for Hydrogen Molecule Ion. Internuclear separation (R) is 1.0 bohr, ρ_H=0.05.

τ	dV/dR	dV/dR ("Φ^2")[a]
SCF	0.6158	
0.20	0.544(3)	0.472(3)
0.15	0.554(3)	0.491(3)
0.10	0.557(3)	0.499(3)
0.05	0.565(3)	0.513(3)
0.00	0.571(4)	0.526(4)
Exact[b]		0.5207

(a) Approximate Φ^2 result using Eq. 2.34.

(b) Derivative obtained from an 8'th degree polynomial fit to the data of H. Wind, J. Chem. Phys. 42, 2371 (1965).

Table 4.5. QMC Hellmann-Feynman Theorem Derivatives for LiH Molecule. Internuclear separation (R) is 2.5 bohr, ρ_{Li}=0.15, ρ_H=0.05.

τ	dV/dR		dV/dR ("Φ^2")[a]	
	Li	H	Li	H
SCF	0.09417	-0.05929		
0.10	0.007(20)	-0.055(5)	-0.079(20)	-0.050(5)
0.05	0.036(34)	-0.052(6)	-0.021(34)	-0.045(6)
0.025	0.044(67)	-0.056(9)	-0.006(67)	-0.053(6)
0.00	0.059(70)	-0.055(10)	0.024(70)	-0.051(10)
HF[b]			0.06068	-0.06068
CI[c]			0.05866	-0.05866
"Exact"[d]			0.05763	-0.05763

(a) Approximate Φ^2 result using Eq. 2.34.

(b) Derivative obtained from an 5'th degree polynomial fit to the data of P. E. Cade, W. M. Huo, J. Chem. Phys. 47, 614 (1967).

(c) 7'th degree polynomial fit, D. M. Bishop, L. M. Cheung, J. Chem. Phys. 78, 1396 (1983).

(d) 7'th degree polynomial fit to estimated "exact" data of G. C. Lie, E. Clementi, J. Chem. Phys. 60, 1275 (1974).

Table 4.6. Analytic energy derivatives of H_2 as a function of internuclear distance (R) using standard techniques (SCF and CI) and QMC. All energies are in hartrees, derivatives are in hartrees/bohr.

Method	R=0.4 bohr		R=0.9 bohr		R=1.4011 bohr		R=1.9 bohr	
	E	$\frac{dE}{dR}$	E	$\frac{dE}{dR}$	E	$\frac{dE}{dR}$	E	$\frac{dE}{dR}$
SCF[a]	-0.0787	-5.030	-1.0433	-0.5145	-1.1336	0.0049	-1.1017	0.0970
CI[b]	---	---	-1.0826	-0.5101	-1.1737	0.00073	-1.1461	0.0851
QMC[c]	-0.1192(13)	-5.297(12)	-1.0831(12)	-0.5053(45)	-1.1745(12)	0.0009(24)	-1.1467(16)	0.1028(55)
Exact[d]	-0.1202	-5.3066	-1.0836	-0.50074	-1.17447	0.0000	-1.14685	0.0852

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

(b) B. Liu, J. Chem. Phys. 58, 1925 (1973).

(c) The QMC trial function consisted of a double zeta SCF wavefunction and an electron-electron correlation function of the form $(1-a\exp(br_{12}+cr_{12}^d)$.

(d) W. Kolos and L. Wolniewicz, J. Chem. Phys. 43, 2429 (1965).

Table 4.7. QMC Analytic Derivatives for Lithium Hydride. Internuclear separation (R) is 2.5 bohr, ρ_{Li}=0.15, ρ_H=0.05.

τ	E	dE/dR
0.100	-8.1136(16)	-0.109(18)
0.050	-8.0909(29)	-0.081(21)
0.025	-8.0769(31)	-0.069(16)
0.000	-8.0656(36)	-0.055(21)
CI[a]	-8.0524	-0.05866
"Exact"[b]	-8.0577	-0.05763

(a) 7'th degree polynomial fit, D. M. Bishop, L. M. Cheung, J. Chem. Phys. 78, 1396 (1983).

(b) 7'th degree polynomial fit to estimated "exact" data of G. C. Lie, E. Clementi, J. Chem. Phys. 60, 1275 (1974).

Table 4.8. Convergence of correlated-sampling VMC optimization. The SCF minimum is at 1.9900 bohr. The final entry is the average value of the last 7 points.

R	$E(R)$	dE/dR	d^2E/dR^2
2.050	-0.5178(108)	-0.118(43)	-0.45(3.35)
2.040	-0.5215(117)	-0.139(45)	-3.09(3.15)
2.030	-0.5272(122)	-0.123(23)	0.32(0.87)
2.020	-0.5315(121)	-0.135(25)	-0.57(0.87)
2.010	-0.5331(120)	-0.134(35)	1.87(0.77)
2.000	-0.5314(117)	-0.214(78)	1.55(2.05)
1.990	-0.5284(120)	-0.097(38)	2.02(1.27)
2.010	-0.5311(118)	-0.120(71)	2.90(2.07)
2.010	-0.5307(91)	-0.230(65)	-90.98(112.01)
2.000	-0.5317(122)	-0.152(32)	0.89(0.60)
2.000	-0.5309(118)	-0.205(54)	4.15(4.50)
2.000	-0.5353(70)	-0.151(16)	0.56(0.90)
1.995	-0.5344(70)	-0.063(32)	-1.19(1.92)
1.990	-0.5397(72)	-0.371(15)	1.16(0.42)
1.985	-0.5322(69)	0.308(49)	0.98(1.54)
1.990	-0.5370(70)	0.022(19)	1.57(0.78)
1.990	-0.5361(70)	0.075(77)	9.26(11.04)
1.990	-0.5360(71)	0.048(28)	1.64(2.32)
1.990	-0.5366(72)	0.052(12)	-0.26(0.57)
1.990	-0.5362(69)	0.049(20)	-0.68(1.12)
1.990	-0.5361(71)	0.026(36)	1.82(2.11)
1.990	-0.5359(70)	0.035(34)	-1.41(1.66)
< 1.990 >	-0.5367(4)	0.044(6)	0.88(0.44)

Figure 4.1. Polynomial fit to the four QMC energies of H_2 given in Table 4.6

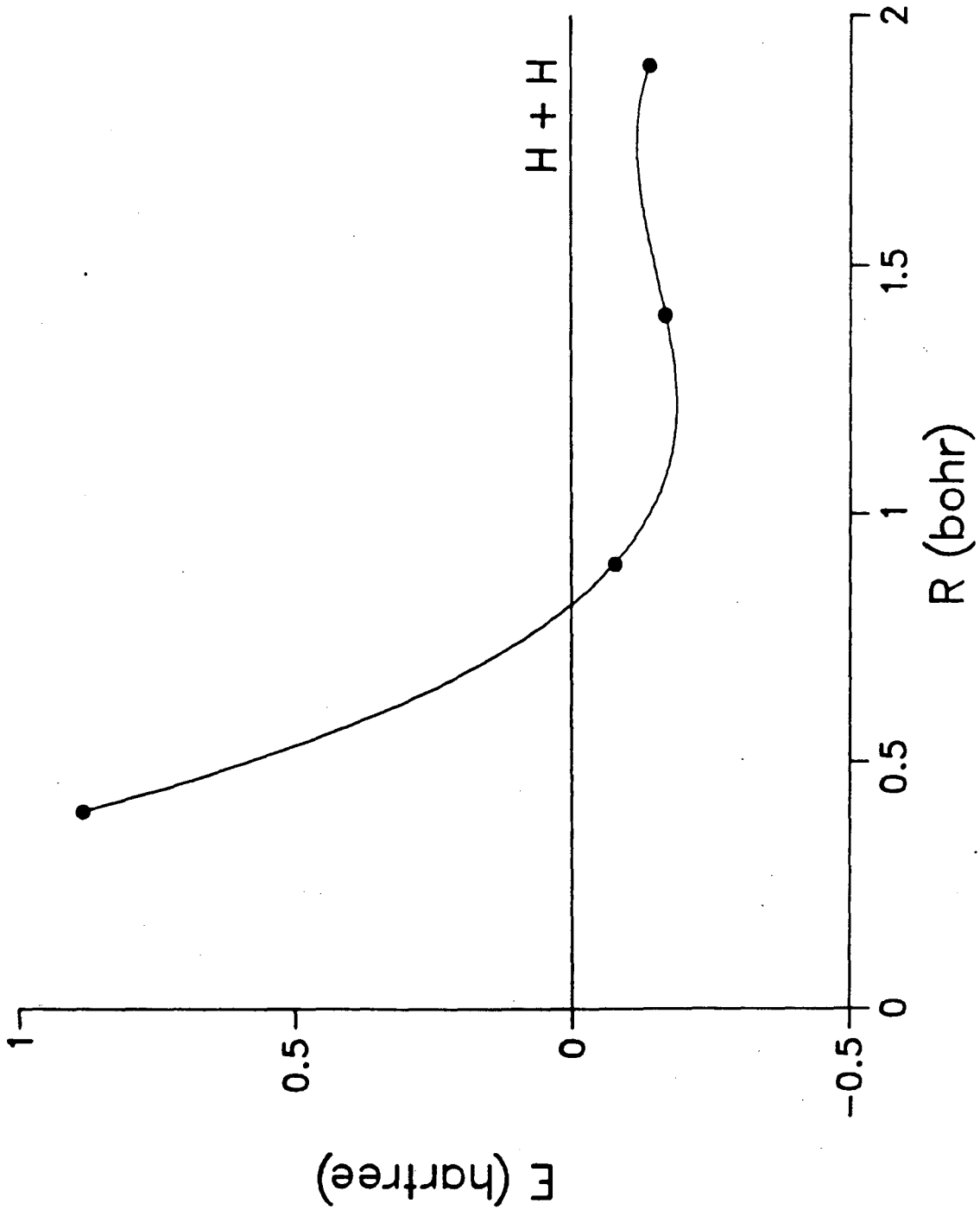
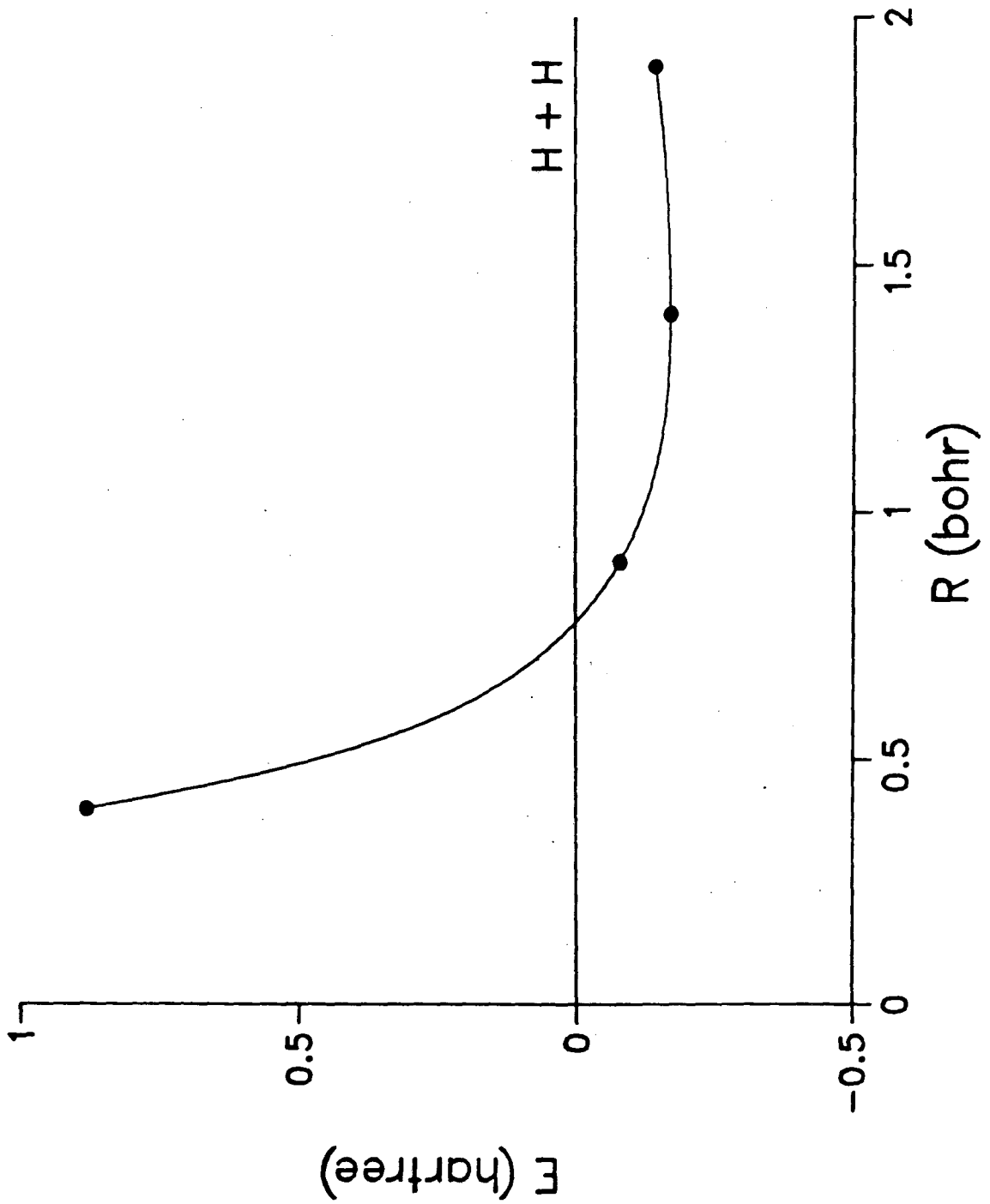


Figure 4.2. Polynomial fit to the four QMC energies and analytic derivatives of H_2 given in Table 4.6. The given curve lies visually on top of the exact result.



APPENDIX A

Input and specifications for QMagiC version 7.2

The input consists of two files. Sections A.1 to A.5 deal with the first of these files. This file specifies the wavefunction to be used (A.1), the Monte Carlo parameters (A.2), generation and I/O of the initial and final ensemble (A.3), and general control parameters (A.4) of the program, and input to the special modules ECP, FZC, HFT, and DRV(A.5). The second file is only needed when the present run is to startup where a previous run left off (section A.6). This is called the XX file, and it contains the electronic configurations and random numbers of a previous run. Sec. A.7 is a brief outline of how to start using the program, and Sec. A.8 is a sample input and output for the methane molecule.

Input is organized into namelists and data groups with descriptive titles. Any information outside a namelist or data group is considered to be a comment, and is ignored. Also, a namelist or data group can be deactivated by adding an extra $ to the title ($$NAMELIST). This makes it possible to keep old values of the parameters in the same file for easy switching and reference. In data groups, all integers can have at most 10 digits, and real variables 18 digits (not counting decimal point and sign) 8 of which can be to the right of the decimal point. All values must be given in atomic units (hartree, bohr, etc.).

All parameters default values are set in BLOCK DATA DEFAULT. At present the maximum number of atoms expected by the input is 20, the maximum number of primitive functions per a.o. shell is 30, and the maximum number of a.o. shells is 100.

A.1 The trial wavefunction.

NAMELIST $PSIT

Function:

Reads in the parameters for the electron-electron correlation function.

Parameters:

EECFLG	Flag specifying form to use 0...use Pade-Jastrow function (default) 1...use double-exponential function
EECF(4)	Parameters of correlation function

Example:

```
$PSIT  EECFLG=0, EECF=0.50,1.50,0.0,0.0, $END
```

DATA $BASIS

Function:

Input geometry, basis set of atomic orbitals for determinant, and e-n jastrow parameters. Each atomic center is specified separately. First the nuclear information, and a electron-nuclear correlation function, followed by the atomic orbitals associated with this center. The basis set consists of Cartesian Slater-type or Gaussian-type functions. Typically, the STO basis set is generated using a standard GTO-SCF program with an STO-NG basis set (N≥5).

Line	Parameters	Format and comments
-1-	TITLE	format(A80)
-2-	SYMB,CHARGE(i),X(i),Y(i),Z(i), LAMBDA(i),NU(i)	free format
	SYMB	atomic symbol (up to 3 characters)
	CHARGE	atomic charge (real number)
	X,Y,Z	cartesian coordinates of atom
	LAMBDA,NU	parameters on e-n Jastrow
-3-	TYPE(l),ZETA(l)	free format
	TYPE	type of atomic function ex. 1S, 2S, 2P, 3S, 3P, 3D for Slaters G1S, G2S, ... for Gaussians. if CONT is specified then a contracted basis is read.
	ZETA	exponent of atomic orbital not used if TYPE=CONT

-4- if TYPE=CONT specify ZETA and COEF of contracted basis, end with blank line.

-5- REPEAT -3- and -4- for each atomic orbital on this center.

-6- leave one BLANK line at end of each atomic center, including the last.

-7- REPEAT -2- to -6- for each atom

-7- $END

Example:

```
$BASIS
   Methane -- double-zeta basis from Clementi
C       6.     0.    0.     0.      0.125 0.625
    1S        7.969
    1S        5.231
    2S        1.820
    2S        1.168
    2P        1.256
    2P        2.726

H1      1.     1.6738  1.1836   0.
    1S        1.640
    1S        1.120

H2      1.    -1.6738  1.1836   0.
    1S        1.640
    1S        1.120

H3      1.     0.     -1.1836   1.6738
    1S        1.640
    1S        1.120

H4      1.     0.     -1.1836  -1.6738
    1S        1.640
    1S        1.120

$END
```

DATA $WFN

Function:

Defines orbital occupancy of each determinant. Each line specifies one determinant: coefficent, then occupation of each m.o. Multiple determinants are only available in the QMC72MD program.

-1-	DETCO(i),(SPIN(i,j),j=1,NMO)	free format
	DETCO	coefficient of this Slater determinant
	SPIN	orbital occupancy of each m.o. in Slater det. SPIN take on values of, UOC ... unoccupied oribital ALP ... alpha occupied orbital BET ... beta occupied orbital DOC ... doubly occupied orbital
-2-	REPEAT -1- for each Slater det.	
-3-	$END	

Example:

```
-EXAMPLE C atom triplet -
 $WFN
1.0  DOC DOC ALP ALP
 $END

-EXAMPLE CH2 open shell singlet -
 $WFN
0.9512  DOC DOC DOC ALP BET UOC
0.3086  DOC DOC DOC ALP UOC BET
 $END
```

DATA $VEC

Function:

Reads in atomic orbital coefficient matrix.

Format depends on VECFLG in $CNTRL (see A.4)

VECFLG = 0 (default)

Standard HONDO format. Matrix is printed 5 a.o.'s per line. Each line starts with the m.o. number and the number of this line in the m.o.

-1-	IORB,ICARD,COEF(5*(ICARD-1)+1,IORB), COEF(5*(ICARD-1)+2,IORB),...	format(I2,I3,5E15.8)
	IORB	m.o. number
	ICARD	line number in this m.o.
	COEF	coefficient of the 5*(icard)+j atomic orbital
-2-	REPEAT -1-	
-3-	$END	

Example:

See sample input, Sec. A.8

A.2 Monte Carlo parameters.

NAMELIST $WALK

Function:

Controls the random walk. Multiple walks can be achieved by putting the appropriate values into the arrays RUNFLG, NUMBLK, BLKTIM and TSTEP. Up to 10 separate walks are accepted. The run is terminated by the first 0 value found in either NUMBLK, or BLKTIM. The ensemble is returned to the value KONORM when the number of configuations reaches either KONMIN or KONMAX.

Parameters:

RUNFLG(10)	type of walk to be performed 0 ... variational walk 1 ... fixed-node walk
NUMBLK(10)	number of blocks in walk
BLKTIM(10)	time span of a block
TSTEP(10)	size of time step
KONORM	number of configurations in normalized ensemble (Default=100)
KONMIN	minimum number of configurations in the ensemble (Default=KONORM/4)
KONMAX	maximum number of configurations in the ensemble Default=KONORM for VMC Default=2*KONORM for FNDQMC or if IGUESS=-1
ETRIAL	inital trial energy if no value (or 0.) is supplied, the value in the XX file is used, otherwise a crude guess is made.
ETW	weight used to update the trial energy from the growth estimator. ETW=0. means trial energy is never updated
RAN	inital random number (in range 0. to 1.) RAN = 0. ... created from time and date, or read in from XX file 0<RAN<1.0 ... RAN is used

Example:

```
$WALK
    RUNFLG=  0,   1,
    NUMBLK=  5,  50,
    BLKTIM= 10.0, 100.0,
    TSTEP= 0.20, 0.20,
    KONORM= 20,
    KONMAX= 50,
    ETRIAL= -40.1960,
    ETW= 0.5,
    RAN= 0.,
$END
```

A.3 Generation and I/O of the electronic coordinates.

The I/O unit numbers are contained in IR (input), IW (output), IXX1 (input XX file) and IXX2 (output XX file). At present these are, IR = 5, IW = 6, IXX1 = 1, IXX2 = 2. These can be changed if desired. Flag PUNFLG in $CNTRL controls whether the final XX file is output to IXX2. See also Sec. A.6 which details the content of the XX file.

NAMELIST $GUESS

Function:

Specifies how the initial configurations are generated. Note that if IGUESS=-1 then ICHECK and IFILL will be set to 1.

Parameters:

IGUESS	-1 ... configs created randomly
	1 ... configs read in on unit IXX1 (see A.5)
ICHECK	1 ... checks the ensemble using the tolerances in $CNTRL
	0 ... do not check ensemble (default)
IFILL	1 ... fills out the ensemble from however many configurations were found in the XX FILE to KONORM (default)
	0 ... do not fill out ensemble. If not enough configurations were supplied an error message is printed and termination of the program occurs

A.4 General control parameters for the program.

NAMELIST $CNTRL

Function:

Reads in parameters controling the run.

Parameters:

VECFLG
0 ... a.o.'s are in HONDO format (Default)
1 ... a.o.'s are in user supplied format (see $VEC)

IPRINT
0 ... normal printing (Default)
1 ... Prints info after each block
2 ... debug printing

PUNFLG
-1 ... don't print XX FILE
0 ... normal printing of XX FILE (Default)
1 ... print the configurations at the end of each block
this will result in NBLOCK*KONORM configurations in
the XX FILE. Useful for enlarging the ensemble.

CSPFLG
0 ... do not evaluate the average electron-nuclear cusp of the
Slater determinant
1 ... evaluate the average electron-nuclear cusp of the
Slater determinant (Default)

STATOL
the smallest a.o. coefficient to be included in m.o. (Default=1.E-06)

DETOL
the value of the determinant which is to be considered
as zero when checking for nodes (Default=1.E-25)
NOTE: set this to the precision of your machine.

REETOL
the smallest e-e distance allowed in random
generation of the ensemble or if ICHEK=1 in $GUESS (Default=1.E-01)

RENTOL
the smallest e-n distance allowed in random
generation of the ensemble or if ICHEK=1 in $GUESS (Default=1.E-01)

PSITOL
the smallest absolute value of Ψ_T allowed in random
generation of the ensemble or if ICHEK=1 in $GUESS (Default=0.)

FQTOL
the largest absolute value of F_Q allowed during the
random walk (Default=200.)

NAMELIST $AP

Function:

Reads in information on parallel runs. Can be used in sequential mode to do multiple independent runs (an XX file must be supplied for each run).

Parameters:

NPROC	the number of independant runs to be performed. Does not need to be the same as the number of actual processors, so in sequential mode this produces nproc independant runs. Note: there must be a complete XX file for each run.
NAP	the number of actual processors to be used
IPROC	identity of processors

A.5 Input to optional modules.

The following describes input to the ECP (effective core potentials), FZC (damped-core), HFT (Hellmann-Feynman derivatives), and DRV (analytic derivatives) modules.

NAMELIST $ECPCTL

Function:

Reads in parameters controling the ECP run.

Parameters:

NSTEP	the number of points calculated on the ECP grid
NGRID	the number of points printed to IW (default=0)
IPRNT	0 ... normal printing 1 ... Prints the ECP grid
RESTRT	0 ... calculate the ECP grid 1 ... read in ECP grid from unit FOR003
ECPTOL	L,M component of ECP is considered to be zero if all points on grid are less that ECPTOL

DATA $ECP

Function:

Specifies the ECP parameters. For each angular momentum value (L) the potential is given by,

$$U_l(r) = \frac{1}{r^2}\sum_k d_{kl} r^{n_{kl}} e^{-b_{kl}r^2},$$

Parameters:

-1- Atomic symbol of ECP atom as in $BASIS and true atomic number.
Each ECP atom must have a unique symbol, even if identical to another.

The true atomic number is used only in $GUESS if IGUESS=-1.

-2- L value of this potential (S, P, D,...). Note combinations can be used, S-D, and the lowest L value is used.

-3- Gaussian parameters D(I),N(I),B(I)

For each center all the L values are specified without blank lines. Each center is separated from the others by a blank line.

Example:

```
------- EXAMPLE FE ATOM 16 VALENCE ELECTRONS --------
$ECPCTL NSTEP=100, IPRNT=0, RESTRT=0, NGRID=0, $END
$ECP
FE1    26.
  D
 -3.89423    1.  17.61910
  S-D
  3.81706    0.   2.33150
 172.05349    2.   6.08365
-144.70056    2.   5.26659
  P-D
  4.13737    0.  49.40456
  82.73696    2.  11.40183

FE2    26.
  D
 -3.89423    1.  17.61910
  S-D
  3.81706    0.   2.33150
 172.05349    2.   6.08365
-144.70056    2.   5.26659
  P-D
  4.13737    0.  49.40456
  82.73696    2.  11.40183

$END
```

NAMELIST $FZC

Function:

Reads in parameters controling the Damped core run. NOTE: at present these parameters are contained in namelists $CNTRL, $PSIT, and $WALK and data list $WFN.

Parameters:

$CNTRL RENTOL	the cutoff radius of the DCGF (Sec. 3.2.3)
$CNTRL RENA	the cutoff width of the DCGF
$PSIT EECF(4)	these four values specify the "a" and "b" parameters of the linear Pade-Jastrow function, first for the core, then valence.
$WALK KONORC	the number of core configurations
$WALK TSTEPC	the core time step
$WFN COR	indicates a doubly occupied core m.o.

```
------- EXAMPLE SODIUM DIMER --------
$WALK
    RUNFLG=  1,  1,
    NUMBLK=  50,  0,
    BLKTIM= 20.0, 20.,
    TSTEP=  0.05,0.10,
    KONORM= 50, KONMAX=100, ETRIAL=-.200, ETW=0.5,
    KONORC=1, TSTEPC=0.025,
$END
$CNTRL IPRINT=1, PUNFLG=0, DETOL=1.E-1000,
    PSITOL=1.E-10, RENTOL=0.0, RENA=0.02, $END
$PSIT  EECFLG=0, EECF=0.5,11.0,0.5,2.0, $END
$WFN
1. COR COR COR COR COR  DOC
$END
```

NAMELIST $HFTD

Function:

Reads in parameters controling the calculation of HFT derivatives.

Parameters:

EPSLN(10) the cutoff radius of the HFT (Sec. 4.2) for each atom.

------- EXAMPLE LITHIUM HYDRIDE --------
$HFTD EPSLN=0.25,0.05, $END

NAMELIST $DRV

Function:

Reads in parameters controling the calculation of analytic derivatives.

Parameters:

RHO(10) the cutoff radius (Sec. 4.3) for each atom.

NSAMP the deviative will only be calculated every NSAMP time steps

------- EXAMPLE LITHIUM HYDRIDE --------
$DRV RHO=0.25,0.05, NSAMP=10, $END

A.6 Description of the XX file.

Function:

Reads the electronic configurations and random number table generated by a previous run.

*** NOTE:

This file is SEPARATE from the input file. It is read in on unit IXX1 and written out to IXX2 (see section A.3). YOU DO NOT NEED TO CREATE OR CHANGE THIS FILE. THE PROGRAM WILL GENERATE THIS FILE AUTOMATICALLY AFTER A SUCCESSFUL RUN (see NOXX in $CNTRL if you do not want this file created).

*** WARNING:

THE PROGRAM DOES NOT CHECK WHETHER THIS FILE IS COMPATIBLE WITH THE PRESENT RUN (E.G. THE SAME MOLECULE, OR WAVEFUNCTION WAS USED IN BOTH RUNS). THIS PROVIDES FLEXIBILITY, BUT IT IS UP TO YOU TO KEEP THINGS STRAIGHT. A MIXUP CAN RESULT IN GARBAGE.

DATA $TITLE

Descriptive title of xx-file. This is the title from namelist $BASIS when output. Ends with $END.

DATA $ETRIAL

Trial energy from last run. Will be used only if ETRIAL in $WALK is 0. Ends with $END.

DATA $RAND

The random number or random number table, depending on which random number generator is being used. Ends with $END.

DATA $XX

The electronic coordinates listed one electron at a time in 3F25.20 format. Ends with $END.

A.6 Conversion Hints

This code has several features intended to aid in conversion from one machine to another. The only machine dependent subroutines are SYSIO, SYSTIM, SYSCPU, SYSCLOSE, and SETCRN. These can be replaced by versions which work on your machine, and no further modifications should be nessary. In addition, there are a few other subroutines you may want to replace to optimize the code. These are RDVEC, and SETRAN (which has entry points SETRAN, GETRAN and PUTRAN). RDVEC reads in the atomic orbital coefficients in standard HONDO format (see section A.3). If you have a SCF code which produces another format, it is problably easier to rewrite RDVEC than to change the format of the coefficients. SETRAN is specific to the random number generator provided (RAND). This is problably slower than a native generator on your machine. To convert to a new generator, change all calls to RAND, and rewrite SETRAN.

A.7 Getting started

To start running from scratch the following steps are needed:

(1) Generate a trial function. Typically a DZ to DZP SCF function using STF's. The QuantuMagiC series is designed to be compatible with HONDO. Thus in HONDO specify "STO 6" or "STO 5" and set the scale parameter to the desired ZETA. Then in QMC72, just specify the ZETA, and include the .PUN (FOR007) from HONDO. Note that the VMC energy should be slightly lower than the SCF energy since QMC72 uses real STF's not STO-NG's.

(2) Generate an initial set of configurations. For small systems the automatic procedure invoked by $GUESS IGUESS=-1, almost always works. For large numbers of electrons the value of $CNTRL PSITOL may need to be set to a very small value in order for the program to pass the checking stage. If difficulties occur in generation, obtain a single configuration by any means (either by

IGUESS=-1 or create an XX file), and do a short VMC walk. Then do another VMC walk with KONORM=10 and IFILL=1, to obtain 10 initial configurations.

(3) Test the CPU time requirements. With a small ensemble, very short values of $WALK BLKTIM, NUMBLK, and a large TSTEP (say 0.1), run a VMC walk to see how long it takes. The CPU time scales linearly with BLKTIM, NUMBLK, and KONORM.

(4) Equilibrate this ensemble for a VMC run. Equilibrium is usually achieved in 1-10 hartree^{-1}. For cases with very slow convergence, one can run a short QMC run to help push the ensemble along.

(5) Optimize the parameters. Knowing the CPU requirements, increase BLKTIM, NUMBLK, and KONORM to achieve optimum performance. Also find a TSTEP which produces an acceptance ratio of 50–80. The ensemble can be filled out by setting $CNTRL PUNFLG=1, then NUMBLK*KONORM configurations will result which will be relatively decorrelated. The values of BLKTIM, NUMBLK, and KONORM must be balanced so that each is as large as possible, yet still keep the CPU down. In order for the statistical precision to be accurate NUMBLK should be at least 20 and BLKTIM/TSTEP should be 50-100 steps. Hence, if necessary, KONORM can be kept at 1 for a VMC run, so as to decrease equilibration time and increase the reliability of the computed statistical precision. At this point moderate size VMC runs are performed to optimize the EECF and ENCF parameters. For the linear Pade-Jastrow EECF a=0.5, and a good starting value of b is the atomic number of the heaviest nuclei present. For the linear Pade-Jastrow ENCF a is set to the cusp value produced by $CNTRL CSPFLG=1, and a good starting value of b is 2 to 3 times a. A problem arises when the ENCF cusp is negative. This has the effect of pushing out the electronic distribution, whereas the original intent of the ENCF was to contract the distribution. In such cases, the values which produce the best energy and variance, while remaining close to the cusp value (or the absolute value of the cusp) is recomended.

(6) Time step bias exploration in the FNDQMC. At this point one chooses several time steps which give acceptance ratios of >90% and allow the ensembles to equilibrate. Run long enough to either detect the time step bias or determine that it is negligable. From these determine the target values of the time step for large calculation

(7) Production runs. Calculate the energies of these time steps until the desired statistical precision has been reached. Extrapolate to zero time step, and analyse the results as desired.

A.8 Example: Methane, all electron.

---------- *Input file* ----------

```
$BANNER
      ********* Methane all electron -test - no Jen cusp *********
$END
$GUESS IGUESS=1, IFILL=1, ICHECK=0, $END
$WALK RUNFLG=1, NUMBLK=5, BLKTIM=2.0, TSTEP=0.2,
   KONORM=10,KONMAX=10,ETRIAL=-40.1960, ETW=0.5, $END
$CNTRL IPRINT=1, CSPFLG=1, PSITOL=2.E-20, $END
$PSIT  EECFLG=0, EECF=0.50,5.50,0.0,0.0, $END
$WFN
1.0  DOC  DOC  DOC  DOC  DOC
$END
$BASIS
  Methane -- double-zeta basis from Clementi
C       6.     0.          0.          0.    0.0285  0.0750
    1S         7.069
    1S         4.631
    2S         2.300
    2S         1.280
    2P         2.300
    2P         1.280

H1      1.     1.6738      1.1836      0.
    1S         1.566
    1S         0.888

H2      1.    -1.6738      1.1836      0.
    1S         1.566
    1S         0.888

H3      1.     0.         -1.1836      1.6738
    1S         1.566
    1S         0.888

H4      1.     0.         -1.1836     -1.6738
    1S         1.566
    1S         0.888

$END
----- ORBITALS FROM -RHFCL- -----
$VEC
1  1 0.40981966E+00 0.61742212E+00-0.38604892E-01 0.38174342E-01 0.00000000E+00
1  2 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00
1  3 0.45771361E-02-0.92455994E-02 0.45771361E-02-0.92455994E-02 0.45771361E-02
1  4-0.92455994E-02 0.45771361E-02-0.92455994E-02
2  1 0.15555769E-01 0.22493452E+00-0.32587279E+00-0.58663590E+00 0.00000000E+00
2  2 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00
2  3-0.17254099E+00 0.59328564E-01-0.17254099E+00 0.59328564E-01-0.17254099E+00
2  4 0.59328564E-01-0.17254099E+00 0.59328564E-01
3  1 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00
3  2 0.25624597E+00 0.27592231E-10 0.00000000E+00 0.37114981E+00 0.30461063E-10
3  3 0.13106609E+00 0.16795116E+00 0.13106609E+00 0.16795116E+00-0.13106609E+00
3  4-0.16795116E+00-0.13106609E+00-0.16795116E+00
```

```
4  1 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00-0.60357528E-01
4  2-0.29102469E-10 0.24903558E+00-0.87423206E-01-0.30465920E-10 0.36070875E+00
4  3-0.43658178E-01-0.55950322E-01 0.43658178E-01 0.55950322E-01 0.18013394E+00
4  4 0.23085142E+00-0.18013394E+00-0.23085142E+00
5  1 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00-0.24903558E+00
5  2 0.00000000E+00-0.60357528E-01-0.36070875E+00 0.00000000E+00-0.87423206E-01
5  3-0.18013394E+00-0.23085142E+00 0.18013394E+00 0.23085142E+00-0.43658178E-01
5  4-0.55950322E-01 0.43658178E-01 0.55950322E-01
$END
```

---------- *End of input file* ----------

---------- *Output file* ----------

12-OCT-87 13:56:16

AMOUNT OF INTEGER CORE REQUESTED 163, AMOUNT AVAILABLE 5000
AMOUNT OF REAL CORE REQUESTED 3928, AMOUNT AVAILABLE 300000

********* Methane all electron -test - no Jen cusp *********

1

Methane -- double-zeta basis from Clementi

DEBUG PRINTING LEVEL = 1

ATOMIC PARAMETERS

ATOM	CHARGE	X	Y	Z	LAMBDA	NU	A.O.S
C	6	0.0000	0.0000	0.0000	0.028500	0.075000	10
H1	1	1.6738	1.1836	0.0000	0.000000	0.000000	2
H2	1	-1.6738	1.1836	0.0000	0.000000	0.000000	2
H3	1	0.0000	-1.1836	1.6738	0.000000	0.000000	2
H4	1	0.0000	-1.1836	-1.6738	0.000000	0.000000	2

INTERNUCLEAR DISTANCES

	C	H1	H2	H3	H4
C	0.0000	2.0500	2.0500	2.0500	2.0500
H1	2.0500	0.0000	3.3476	3.3477	3.3477
H2	2.0500	3.3476	0.0000	3.3477	3.3477
H3	2.0500	3.3477	3.3477	0.0000	3.3476
H4	2.0500	3.3477	3.3477	3.3476	0.0000

XCM = 0.0000 YCM = 0.0000 ZCM = 0.0000

ATOMIC BASIS SET

ATOM NUMBER	ORBITAL TYPE	EXPONENT	NORM	COEFFICIENT (UNNORM)
1 C				
1	1S	7.0690	1.000000	(10.603812)
2	1S	4.6310	1.000000	(5.622601)
3	2S	2.3000	1.000000	(2.613268)
4	2S	1.2800	1.000000	(0.603795)
5	2PX	2.3000	1.000000	(4.526314)
6	2PY	2.3000	1.000000	(4.526314)
7	2PZ	2.3000	1.000000	(4.526314)
8	2PX	1.2800	1.000000	(1.045803)
9	2PY	1.2800	1.000000	(1.045803)
10	2PZ	1.2800	1.000000	(1.045803)
2 H1				
11	1S	1.5660	1.000000	(1.105637)

```
      12   1S        0.8880   1.000000 (  0.472111)
3 H2
      13   1S        1.5660   1.000000 (  1.105637)
      14   1S        0.8880   1.000000 (  0.472111)
4 H3
      15   1S        1.5660   1.000000 (  1.105637)
      16   1S        0.8880   1.000000 (  0.472111)
5 H4
      17   1S        1.5660   1.000000 (  1.105637)
      18   1S        0.8880   1.000000 (  0.472111)


NUMBER OF NUCLEI          =  5
NUMBER OF ELECTRONS FOR NEUTRAL =  10
NUMBER OF ATOMIC ORBITAL SHELLS =  14
TOTAL NUMBER OF ATOMIC ORBITALS =  18
NUCLEAR POTENIAL ENERGY      =   13.49960
FLAG FOR E-E CORRELATION FUNTION=  0
THE EECF VALUES ARE   0.50000  5.50000  0.00000  0.00000


ORBITAL OCCUPATION OF DETERMINANTS
 DET.  M.O.  COEF.  OCCUPANCY
  1     1  1.000   DOC
  1     2  1.000   DOC
  1     3  1.000   DOC
  1     4  1.000   DOC
  1     5  1.000   DOC


NUMBER OF SLATER DETERMINANTS  =  1
NUMBER OF MOLECULAR ORBITALS   =  5
ELECTRONIC CHARGE          =  0
STATE MULTIPLICITY         =  1
NUMBER OF ALPHA ORBITALS      =  5
NUMBER OF BETA ORBITALS       =  5


UNNORMALIZED TRIAL VECTORS:

                1       2       3       4       5

C  1 1S      1 4.34565  0.16495  0.00000  0.00000  0.00000
     1S      2 3.47152  1.26472  0.00000  0.00000  0.00000
     2S      3 -0.10088 -0.85159  0.00000  0.00000  0.00000
     2S      4  0.02305 -0.35421  0.00000  0.00000  0.00000
     2PX     5  0.00000  0.00000  0.00000 -0.27320 -1.12721
     2PY     6  0.00000  0.00000  1.15985  0.00000  0.00000
     2PZ     7  0.00000  0.00000  0.00000  1.12721 -0.27320
     2PX     8  0.00000  0.00000  0.00000 -0.09143 -0.37723
     2PY     9  0.00000  0.00000  0.38815  0.00000  0.00000
```

```
    2PZ    10  0.00000  0.00000  0.00000  0.37723 -0.09143
H1  2 1S    11  0.00506 -0.19077  0.14491 -0.04827 -0.19916
    1S     12 -0.00436  0.02801  0.07929 -0.02641 -0.10899
H2  3 1S    13  0.00506 -0.19077  0.14491  0.04827  0.19916
    1S     14 -0.00436  0.02801  0.07929  0.02641  0.10899
H3  4 1S    15  0.00506 -0.19077 -0.14491  0.19916 -0.04827
    1S     16 -0.00436  0.02801 -0.07929  0.10899 -0.02641
H4  5 1S    17  0.00506 -0.19077 -0.14491 -0.19916  0.04827
    1S     18 -0.00436  0.02801 -0.07929 -0.10899  0.02641
```

ATOMIC CUSP CONDITIONS

ATOM#	DENS. DERIV.	CUSP	DIFF
1	-0.75637E+03	-5.988482	0.011518
2	-0.48914E+00	-1.379483	-0.379483
3	-0.48914E+00	-1.379483	-0.379483
4	-0.48914E+00	-1.379483	-0.379483
5	-0.48914E+00	-1.379483	-0.379483

MAXIMUM SIZE OF ENSEMBLE = 20
SIZE ENSEMBLE NORMALIZED TO = 10
INITIAL TRIAL ENERGY = -0.401960E+02
NEW TRIAL ENERGY WEIGHT = 0.500000E+00
HBAR**2/2*MASS = 0.500000E+00
UNIT CHARGE**2 OR E**2 = 0.100000E+01

ELAPSED TIME: 3.16, TOTAL TIME: 3.16

1

ENSEMBLE READ IN

FIXED-NODE DIFFUSION MONTE CARLO

TIME STEP = 0.20000
NUMBER OF STEPS IN BLOCK = 10
TIME IN ONE BLOCK = 2.00
VALUE OF ABLOCK IF NO BRANCHING = 100
NUMBER OF BLOCKS = 5
TOTAL TIME FOR THIS WALK = 10.00

ELAPSED TIME: 0.06, TOTAL TIME: 3.22

```
   TITLE OF XX FILE:
Methane -- double-zeta basis from Clementi
   RANDOM NO. FROM XX FILE      = 0.780164631668875064

 ELAPSED TIME:        0.56, TOTAL TIME:        3.78

N    ETRIAL     ELOCAL     ABLOCK     EGROWTH   UNDER OVER
1   -40.19600  -41.20721    116.1   -40.56722    0   1
2   -40.38161  -41.13084    114.2   -40.39947    0   1
3   -40.38607  -41.20777    122.6   -40.50010    0   1
4   -40.40508  -41.44207    149.7   -40.72518    0   2
5   -40.44509  -40.89898    131.7   -40.65080    0   3
   FINAL RANDOM NO., TO XX FILE  = 0.937170131751068933

 ELAPSED TIME:      179.78, TOTAL TIME:      183.56


 ####### AP #  1, BLOCKS  5, CPU    179.560 #######

      BLOCK     EBLOCK     ABLOCK
     -------   --------   --------
       1     -41.20721     116.1150
       2     -41.13084     114.1863
       3     -41.20777     122.5674
       4     -41.44207     149.6910
       5     -40.89898     131.7377
             --------   --------
      TOTAL  -41.17737     634.2974 CPS =    3.53251
      ERROR    0.08705



 ------------------------------------------------
 |                                              |
 | THE TOTAL ENERGY OF THIS RUN IS   -41.177374 |
 |    WITH AN ESTIMATED ERROR OF      0.087047  |
 |                                              |
 ------------------------------------------------
        QUANTITY          AVERAGES      ERROR
 1  E-E + E-N CORR FUNCTIONS  0.708448E+01 0.211931E-02
 2  E-E POTENTIAL ENERGY      0.242755E+02 0.142960E+00
 3  E-N POTENTIAL ENERGY     -0.103084E+03 0.103256E+01
 4  E-E DISTANCE              0.253498E+01 0.283991E-01
 5  DEL**2 PSI + (DEL PSI)**2 0.307088E+02 0.216045E+01
 6  KINETIC ENERGY            0.241315E+02 0.101340E+01
 7  POTENTIAL ENERGY         -0.653089E+02 0.108562E+01
 8  BLOCK ENERGY ESTIMATE    -0.411774E+02 0.870471E-01
 9  CONTINOUS ENERGY ESTIMATE -0.411440E+02 0.649126E-01
 10  GROWTH ENERGY ESTIMATE   -0.405686E+02 0.568217E-01
 11  X - X(CM)                -0.242485E-01 0.248725E-01
 12  Y - Y(CM)                 0.282972E-01 0.354548E-01
 13  Z - Z(CM)                 0.270116E-01 0.294174E-01
 14  (X - X(CM))**2            0.121234E+01 0.398093E-01
 15  (Y - Y(CM))**2            0.122264E+01 0.298828E-01
 16  (Z - Z(CM))**2            0.127316E+01 0.470064E-01
 17  (X - X(CM))*(Y - Y(CM))  -0.191703E-02 0.308030E-01
```

```
 18  (X - X(CM))*(Z - Z(CM))   -0.506386E-01  0.427385E-01
 19  (Y - Y(CM))*(Z - Z(CM))   -0.145393E+00  0.343368E-01
 20  E-CM DISTANCE             0.163158E+01  0.149786E-01
 21  NODES FOUND/BLOCK         0.140000E+01  0.748331E+00
 22  VIRIAL THEOREM (2T+V)     -0.170458E+02  0.943687E+00
 23  VIRIAL THEOREM (-V/2T)    0.135959E+01  0.381755E-01
 24  ACCEPTANCE RATIO          0.757463E+00  0.394883E-02
 25  INT(MULT+X) / MULT        0.102781E+01  0.106373E-01
TOTAL NUMBER OF TIMES A NODE WAS CROSSED IS      7
   ENSEMBLE UNDERFLOWS:   0, OVERFLOWS:   3

ELAPSED TIME:        0.25, TOTAL TIME:        183.81
```

---------- *End of output file* ----------

Appendix B

Evaluation of the Trial Function, and the Local Energy

B.1 Quantities required by the FNDQMC algorithm

The random walk moves each configuration from **R** to **R′** one electron at a time by diffusion, drift, and then acceptance or rejection of the move. This is followed by branching of the configuration when all the electrons have been moved and then updating then averages being kept. For the diffusion step only a spherical gaussian random number is needed, the drift step requires $\mathbf{F}_Q$, and the branching requires the local energy. The only additional quantity needed by the acceptance and averaging is the ratio $\Psi_T(\mathbf{R}')/\Psi_T(\mathbf{R})$ used in the Green's function. Thus the following quantities depending on Ψ_T must be computed: $\Psi_T(\mathbf{R}')/\Psi_T(\mathbf{R})$, $\nabla\Psi_T(\mathbf{R}')/\Psi_T(\mathbf{R})$, and $\nabla^2\Psi_T(\mathbf{R}')/\Psi_T(\mathbf{R})$. Note that in each case the numerator depends on **R′** and the denominator on **R**. This is because it is easier to compute the ratio of two determinants, that differ only by one electron, than the determinants themselves. For the gradient and Laplacian terms, the value with **R′** in both the numerator and denominator is obtained by dividing by $\Psi_T(\mathbf{R}')/\Psi_T(\mathbf{R})$.

B.2 Form of Ψ_T considered

QMC utilizes a Ψ_T which is a factorization of the full NxN Slater determinant by assigning spins to the electrons. The product of α and β determinants is multiplied by a product of electron-electron and electron-nuclear correlation functions to obtain

$$\Psi_T = D^{\alpha} D^{\beta} S_{ee} S_{en} \tag{B.1}$$

The determinants are constructed of molecular orbitals, obtained from either a SCF or MCSCF calculation. The correlation functions are either of the Pade-Jastrow form [8], or the double exponential form [108]. Both of these are of the general form,

$$S = \prod_{u,v} e^{U_{uv}}, \tag{B.2}$$

where u and v are either two electrons or an electron and a nucleus.

B.3 Expressions in terms of D^{α}, D^{β}, S_{ee}, and S_{en}

First the various expressions needed in Sec. B.1 will be expressed in terms of Eq. B.1, then in Sec.'s B.4 and B.5 the explicit forms will be given. Evaluating the ratio $\Psi_T(\mathbf{R}')/\Psi_T(\mathbf{R})$ is simply a matter of evaluating the product of the ratios of each component in Eq. B.1, namely

$$\frac{\Psi_T(\mathbf{R}')}{\Psi_T(\mathbf{R})} = \frac{D^{\alpha}(\mathbf{R}')}{D^{\alpha}(\mathbf{R})}\frac{D^{\beta}(\mathbf{R}')}{D^{\beta}(\mathbf{R})}\frac{S_{ee}(\mathbf{R}')}{S_{ee}(\mathbf{R})}\frac{S_{en}(\mathbf{R}')}{S_{en}(\mathbf{R})}. \tag{B.3}$$

The gradient of the trial function is similarly simple,

$$\sum_i \frac{\nabla_i \Psi_T(\mathbf{R}')}{\Psi_T(\mathbf{R})} = \sum_i \left[\frac{\nabla_i D^{\alpha}(\mathbf{R}')}{D^{\alpha}(\mathbf{R})} + \frac{\nabla_i D^{\beta}(\mathbf{R}')}{D^{\beta}(\mathbf{R})} + \frac{\nabla_i S_{ee}(\mathbf{R}')}{S_{ee}(\mathbf{R})} + \frac{\nabla_i S_{en}(\mathbf{R}')}{S_{en}(\mathbf{R})} \right], \tag{B.4}$$

the index i being the electron.

The term $\nabla^2\Psi_T(\mathbf{R}')/\Psi_T(\mathbf{R})$ is evaluated by the definition $\nabla^2 \equiv \nabla\cdot\nabla$. Hence, one evaluates $\nabla\cdot(\nabla\Psi_T(\mathbf{R}')/\Psi_T(\mathbf{R}))$, obtaining,

$$\begin{aligned}\sum_i \frac{\nabla_i^2 \Psi_T(\mathbf{R}')}{\Psi_T(\mathbf{R})} = \sum_i \Bigg[& \frac{\nabla_i^2 D^{\alpha}(\mathbf{R}')}{D^{\alpha}(\mathbf{R})} + \frac{\nabla_i^2 D^{\beta}(\mathbf{R}')}{D^{\beta}(\mathbf{R})} + \frac{\nabla_i^2 S_{ee}(\mathbf{R}')}{S_{ee}(\mathbf{R})} + \frac{\nabla_i^2 S_{en}(\mathbf{R}')}{S_{en}(\mathbf{R})} \\ & + \left[\frac{\nabla_i D^{\alpha}(\mathbf{R}')}{D^{\alpha}(\mathbf{R})} + \frac{\nabla_i D^{\beta}(\mathbf{R}')}{D^{\beta}(\mathbf{R})} \right] \cdot \left[\frac{\nabla_i S_{ee}(\mathbf{R}')}{S_{ee}(\mathbf{R})} + \frac{\nabla_i S_{en}(\mathbf{R}')}{S_{en}(\mathbf{R})} \right] \Bigg].\end{aligned} \tag{B.5}$$

Each one of these terms can now be expressed in terms of the explicit form of the various functions.

B.4 Evaluation of the determinant

As stated above, finding the ratio of two determinants which differ only by a single electron (corresponding to a single column) is a relatively simple operation. First define the determinant to be

$$D^{\alpha,\beta} = \det \begin{bmatrix} \psi_1(1) & \psi_1(2) & \cdots & \psi_1(n) \\ \psi_2(1) & \psi_2(2) & \cdots & \psi_2(n) \\ \cdots & \cdots & \cdots & \cdots \\ \psi_n(1) & \psi_n(2) & \cdots & \psi_n(n) \end{bmatrix} \tag{B.6}$$

abbreviated,

$$D^{\alpha,\beta} = \det |\psi_1(1)\psi_2(2)\cdots\psi_n(n)|. \tag{B.7}$$

If electron i is moved, or if the operation ∇_i or ∇_i^2 is performed, then only a single column of the determinant is affected, for example, $\nabla_i D$ (either α or β) is

$$\det |\psi_1(1)\cdots\nabla\psi_i(i)\cdots\psi_n(n)|. \tag{B.8}$$

By expanding this in cofactors of the i'th column, C_{ij}, one obtains

$$\nabla_i D = \sum_j \nabla\psi_j(i) C_{ij} \tag{B.9}$$

Note that the inverse of any matrix A can be expressed as the adjoint of the matrix of cofactors divided by the determinant,

$$A_{ij}^{-1} = C_{ji}/\det|A|. \tag{B.10}$$

Therefore, the *ratio* of $\nabla_i D$ to D is given in a straight forward manner,

$$\nabla_i D/D = \sum_j \nabla\psi_j(i) D_{ji}^{-1}. \tag{B.11}$$

In the case where the operator is evaluated at $\mathbf{R}'$ but the inverse determinant is known for position $\mathbf{R}$ then the ratios $\nabla_i D(\mathbf{R}')/D(\mathbf{R})$, $\nabla_i^2 D(\mathbf{R}')/D(\mathbf{R})$, and $\nabla_i^2 D(\mathbf{R}')/D(\mathbf{R})$, are obtained from Eq. B.11. Using this method has the distinct advantage that the inverse matrix only needs to be calculated at $\mathbf{R}$, yet all the values $\nabla_i D(\mathbf{R})/D(\mathbf{R})$, $\nabla_i^2 D(\mathbf{R})/D(\mathbf{R})$, $\nabla_i D(\mathbf{R}')/D(\mathbf{R})$, $\nabla_i^2 D(\mathbf{R}')/D(\mathbf{R})$, and $D(\mathbf{R}')/D(\mathbf{R})$, needed during the electron move are given by Eq. B.11. Only if the move is accepted is the inverse matrix updated (see Sec. B.6).

The molecular orbitals (MOs), ψ are given as a linear combinations of atomic orbitals (AOs), χ. For Slater-type functions (STF's) the following relationships apply

$$\chi_{STF} = x^a y^b z^c r^n e^{-\zeta r} \tag{B.12}$$

$$\chi_{STF}^{-1}\frac{\partial\chi_S}{\partial x}=\frac{a}{x}-\zeta\frac{x}{r}+\frac{nx}{r^2} \tag{B.13}$$

$$\chi_{STF}^{-1}\nabla^2\chi_S=\frac{a(a-1)}{x^2}+\frac{b(b-1)}{y^2}+\frac{c(c-1)}{z^2}+\zeta^2-2\zeta\frac{l}{r}+\frac{2n(l-n/2-½)}{r^2} \tag{B.14}$$

where $l=a+b+c+n+1$. For Cartesian Gaussian-type functions (GTF's), the analogous equations are

$$\chi_{GTF}=x^a y^b z^c \exp(-\zeta r^2), \tag{B.15}$$

$$\chi_{GTF}^{-1}\frac{\partial}{\partial x}\chi_G=\frac{a}{x}-2\zeta x \tag{B.16}$$

$$\nabla^2\chi_{GTF}=\{a(a-1)x^{a-2}+b(b-1)y^{b-2}+c(c-1)z^{c-2} \tag{B.17}$$

$$-2\zeta((2a+1)x^a+(2b+1)y^b+(2c+1)z^c)+4\zeta^2(x^{a+2}+y^{b+2}+z^{c+2})\}e^{-\zeta r^2}$$

In both Eq. B.12, and B.15, the normalization has been left out. The normalization for an STF is

$$N_{STF}(a,b,c,\zeta)=\mathrm{I}(2a,2b,2c)\frac{(2l)!}{(2\zeta)^{2l+1}} \tag{B.18}$$

the function I (a,b,c) being given by Eq. C.14. For a GTF the normalization is most easily given by expanding the GTF into x, y, and z components, *i.e.* the x-component is $x^a exp(-\zeta x^2)$. The normalization for this component is

$$N_{GTF}(a,\zeta)=(\frac{2\zeta}{\pi})^{¼}(4\zeta)^{½a}[(2a-1)!!]^{-½}. \tag{B.19}$$

The total normalization, then, is given by the product of $N_G(a,\zeta)$, $N_G(b,\zeta)$, and $N_G(c,\zeta)$.

B.5 Evaluation of the correlation functions

Two forms of correlation function will be discussed here: the Pade-Jastrow form (Eq. 2.47), and the double-exponential form (Eq. 2.48). For the Pade-Jastrow form the expression considered here is

$$S_{PJ}=e^{\sum_{i>j}\frac{ar_{ij}}{1+br_{ij}}}, \tag{B.20}$$

where only the linear terms of Eq. 2.47 have been retained. The gradient of this is

$$S_{PJ}^{-1}\nabla S_{PJ} = \sum_{i>j}\frac{\mathbf{x}}{r}\frac{a}{(1+br_{ij})^2}\,, \tag{B.21}$$

and the Laplacian of S_{PJ} is,

$$S_{PJ}^{-1}\nabla^2 S_{PJ} = \sum_{i>j} a\,(2r_{ij}^{-1}(1+br_{ij})+a)(1+br_{ij})^{-4}\,. \tag{B.22}$$

For the double-exponential form the linear expression is,

$$S_{DE} = e^{\sum_{i>j}-ae^{(-br_{ij})}} \tag{B.23}$$

The gradient is

$$S_{DE}^{-1}\nabla S_{DE} = \sum_{i>j}\frac{\mathbf{x}}{r}abe^{(-br_{ij})} \tag{B.21}$$

and the Laplacian of S_{DE} is,

$$S_{DE}^{-1}\nabla^2 S_{DE} = \sum_{i>j} ae^{(-br_{ij})}(b^2(abe^{(-br_{ij})}-1)+2br_{ij}^{-1})\,. \tag{B.22}$$

Clearly, S_{DE} is much simpler to evaluate, and quadratic terms can be included by replacing b by $b+cr$ are using the chain rule.

B.6 Computing the Inverse Slater Matrix

Computation of the inverse Slater matrix is the single most time consuming step in the QMC walk. Assume that an initial inverse matrix has already been computed for electronic positions, **R**. If only a single particle is moved, Ceperley, *et al.* [28], showed that not all of the elements of the inverse matrix need to be re-computed. Let the vector **R′** be equal to **R** except that electron i has been moved. Using Eq. B.10, the ratio of the determinant at **R′** to the determinant at **R** is

$$d = \sum_j \psi_j(\mathbf{R}')D_{ji}^{-1}(\mathbf{R}). \tag{B.23}$$

If in the Monte Carlo algorithm this move is rejected then no further computation is needed. If the move is accepted then the new inverse matrix is related to the old inverse matrix by

$$D_{ji}^{-1}(\mathbf{R}') = D_{ji}^{-1}(\mathbf{R})/d, \tag{B.24}$$

and

$$D_{jk}^{-1}(\mathbf{R}') = D_{jk}^{-1}(\mathbf{R}) - D_{ji}^{-1}/d\sum_{l} D_{lj}^{-1}(\mathbf{R}')\psi_l(\mathbf{R}') \tag{B.25}$$

where $k \neq 1$. To compute the initial inverse matrix one can start with the identity matrix and the "move" each electron to its initial position using Eqs. B.24 and B.25.

Appendix C

Evaluation of the ECP

C.1 The angular integrals

In this section a method for the evaluation of $U_L^{ECP}[\Psi_T]$ is presented; extension to $U_L^{ECP}[\Psi_T^2]$ is straightforward. The expression for the local approximation is (Eq. 3.10),

$$U_L^{ECP}[\Psi_T] = \tag{C.1}$$

$$\sum_A^{N_{val}} \sum_{i=1} \left[U^A_{l_{max}+1}(r_{iA}) + \sum_{l=0}^{l_{max}} \sum_{m=-l}^{l} Y_{lm}(\Omega_{iA})\, U_l^A(r_{iA}) <Y_{lm}(\Omega_{iA}) | \Psi_{val}> / \Psi_{val} \right] .$$

Although the theory can be applied to multi-determinant wave functions, it is sufficient for present purposes to consider a single determinant,

$$\Psi_{val} = \det | \phi_1(1) \cdots \phi_i(i) \cdots \phi_n(n) | \ , \tag{C.2}$$

where the ϕ_i are molecular orbitals. Since the local ECP is a sum over one-electron operators, one need consider only a single column at a time of the Slater matrix. Thus, by expanding the determinant in co-factors of electron i, the angular integral may be written

$$<Y_{lm}(\Omega_{iA}) | \Psi_{val}> / \Psi_{val} = \sum_j D_{ji}^{-1} <Y_{lm}(\Omega_{iA}) | \phi_j(i)> , \tag{C.3}$$

where we have used the property that the elements of the inverse of the Slater matrix, D_{ji}^{-1}, are the transpose of the co-factors divided by the determinant. Next, expanding the molecular orbitals in single-particle basis functions, χ, we obtain

$$<Y_{lm}(\Omega_{iA}) | \phi_j(i)> = \sum_{p=1}^{N_{bas}} c_{jp} \int d\Omega_{iA}\, Y_{lm}(\Omega_{iA}) \chi_p(\mathbf{r}_i) \ . \tag{C.4}$$

The basis functions, χ, are taken to be Cartesian Gaussian-type functions (GTF's), see Appendix B. In Eq. C.4, the sum over basis functions can be usefully broken up into functions centered on atom A (with the ECP) and functions centered on all remaining atoms. Thus the right hand side of Eq. C.4 becomes

$$\sum_{p=1}^{N_{bas}^{A}} c_{jp} \int d\Omega_{iA}\, Y_{lm}(\Omega_{iA})\chi_p(\mathbf{r}_{iA}) + \sum_{B\neq A}\sum_{p=1}^{N_{bas}^{B}} c_{jp} \int d\Omega_{iA}\, Y_{lm}(\Omega_{iA})\chi_p(\mathbf{r}_{iB}) \qquad \text{(C.5)}$$

where N_{bas}^{A} and N_{bas}^{B} are the number of basis functions centered on atoms A and B, respectively.

To evaluate the integral involving only atomic center A, note that

$$\sum_{m=-l}^{l} Y_{lm}(\Omega_{iA}) \int d\Omega_{iA}\, Y_{lm}(\Omega_{iA})\chi_p(\mathbf{r}_{iA}) = \delta_{l,l_p}\, \chi_p(\mathbf{r}_{iA}) \qquad \text{(C.6)}$$

where l_p is the electronic orbital-angular-momentum quantum number of basis function p. Thus, combining Eqs. C.1, C.2, C.3, and the first term in Eq.C.5, performing the sum over m and l, and using C.6, the single-center term for electron i in molecular orbital j is

$$u_{ij}^{AA} \equiv D_{ji}^{-1} \sum_{l=0}^{l_{max}} \sum_{m=-l}^{l} Y_{lm}(\Omega_{iA})\, U_l^A(r_{iA}) \sum_{p=1}^{N_{bas}^{A}} c_{jp} \int d\Omega_{iA}\, Y_{lm}(\Omega_{iA})\chi_p(\mathbf{r}_{iA}) \qquad \text{(C.7)}$$

$$= D_{ji}^{-1} \sum_{p=1}^{N_{bas}^{A}} c_{jp}\, \chi_p(\mathbf{r}_{iA}) U_{l_p}^A(r_{iA}) .$$

For the integrals involving two centers, basis functions on center B must be expanded around center A. To this end Kahn, *et. al.* [89] have shown that

$$x_{iB}^{u} y_{iB}^{v} z_{iB}^{w} e^{-\zeta r_{iB}^2} = e^{-\zeta r_{iA}^2} \Bigg\{ e^{-\zeta R_{AB}^2} \sum_{r=0}^{u}\sum_{s=0}^{v}\sum_{t=0}^{w} \binom{u}{r}\binom{v}{s}\binom{w}{t} X_{AB}^{u-r} Y_{AB}^{v-s} Z_{AB}^{w-t}\, r_{iA}^{r+s+t}$$

$$\times 4\pi \sum_{\lambda=0}^{\infty} M_\lambda(2\zeta R_{AB} r_{iA}) \sum_{\mu=-\lambda}^{\lambda} Y_{\lambda\mu}(\Omega_{AB})\, \hat{x}_{iA}^{r} \hat{y}_{iA}^{s} \hat{z}_{iA}^{t}\, Y_{\lambda\mu}(\Omega_{iA}) \Bigg\} \qquad \text{(C.8)}$$

where $\hat{x} = x/r$ (the angular part of the Cartesian coordinate x), and M_λ is a modified spherical Bessel function of the first kind [115],

$$M_\lambda(z) = \sqrt{\tfrac{1}{2}\pi/z}\, I_{\lambda+\frac{1}{2}}(z). \qquad \text{(C.9)}$$

Substituting Eq. C.8 into the second term of Eq. C.5, and temporarily ignoring all the

radial terms, the angular integral is of the form

$$\int d\Omega\, Y_{lm}(\Omega)\, \hat{x}^r\, \hat{y}^s\, \hat{z}^t\, Y_{\lambda\mu}(\Omega) \equiv \Omega^{rst}_{\lambda\mu lm}. \tag{C.10}$$

The Y_{lm} may be expressed in terms of $\hat{x}$, $\hat{y}$, and $\hat{z}$,

$$Y_{lm} = \sum_{a=0}^{a+b+c=l} \sum_{b=0} \sum_{c=0} g^{lm}_{abc}\, \hat{x}^a\, \hat{y}^b\, \hat{z}^c\ , \tag{C.11}$$

which leads to

$$\Omega^{rst}_{\lambda\mu lm} = \sum_{a,b,c}^{a+b+c=l} \sum_{d,e,f}^{d+e+f=\lambda} g^{lm}_{abc}\, g^{\lambda\mu}_{def}\, \mathrm{I}(r+a+d, s+b+e, t+c+f)\,.$$

The elementary integral $\mathrm{I}(i,j,k)$ [116] is

$$\mathrm{I}(i,j,k) \equiv \int d\Omega\, \hat{x}^i\, \hat{y}^j\, \hat{z}^k = \begin{cases} \dfrac{4\pi(i-1)!!(j-1)!!(k-1)!!}{(i+j+k+1)!!} & \text{i,j,k all even} \\ 0 & \text{otherwise} \end{cases} \tag{C.12}$$

where $n!! \equiv 1\cdot 3\cdot 5 \cdots n$, 0!!=1, and (-1)!!=1. Note that the elements of $\Omega^{rst}_{\lambda\mu lm}$ are constants, and so may be calculated once and tabulated.

Thus, the two-center term for electron i in molecular orbital j is

$$u^{AB}_{ij} = D^{-1}_{ji} \sum_{p=1}^{N^B_{bas}} c_{jp}\, N_{u_p v_p w_p}(\zeta_p)\, e^{-\zeta_p r^2_{iA}}$$

$$\times \sum_{l=0}^{l_{max}} \sum_{m=-l}^{l} \sqrt{4\pi}\, Y_{lm}(\Omega_{iA})\, U^A_l(r_{iA})\, \mathrm{F}^p_{lm}(r_{iA}, \mathbf{R}_{AB}). \tag{C.13}$$

All the complicated behavior of Eq. C.8 has been lumped into the single term, F^p_{lm}, which is,

$$\mathrm{F}^p_{lm}(r,\mathbf{R}) \equiv e^{-\zeta_p R^2} \sum_{u}^{u_p} \sum_{v}^{v_p} \sum_{w}^{w_p} \begin{bmatrix} u_p \\ u \end{bmatrix} \begin{bmatrix} v_p \\ v \end{bmatrix} \begin{bmatrix} w_p \\ w \end{bmatrix} X^{u_p-u}\, Y^{v_p-v}\, Z^{w_p-w}\, r^{u+v+w}$$

$$\times \sum_{\lambda=0}^{u_p+v_p+w_p+l} M_\lambda(2\zeta_p R r) \sum_{\mu=-\lambda}^{\lambda} \sqrt{4\pi}\, Y_{\lambda\mu}(\Omega)\, \Omega^{uvw}_{\lambda\mu lm}\ , \tag{C.14}$$

where (X,Y,Z), are the Cartesian and (R,Ω), are the polar coordinates of the internuclear

vector **R**, and r is the distance from the ECP atom (A) to the electron.

Returning to the original expression for the local ECP, Eq. C.1 can now be written in terms of u_{ij}^{AA} and u_{ij}^{AB},

$$U_{Local}^{ECP} = \sum_{A}\sum_{i=1}^{N_{val}}\left[U_{l_{max}+1}^{A}(r_{iA}) + \sum_{j}\left[u_{ij}^{AA} + \sum_{B\neq A} u_{ij}^{AB}\right]\right]. \qquad \text{(C.15)}$$

For an atom, the u_{ij}^{AB} term does not appear, and the expressions are greatly simplified. In the molecular case, to evaluate u_{ij}^{AB} at each step of the walk takes roughly 15 times longer to compute than the rest of E_{local}. To reduce the time required, the r dependence of u_{ij}^{AB} (Eq. C.15) is fit numerically to exponential splines prior to the QMC run, as described in Sec C.2. This procedure must be carried out for each choice of importance function and geometry.

C.2 Pre-calculating the ECP

The first step is to isolate a set of calculations which may be stored easily, yet eliminate most of the work needed at run time. Examining Eqs. C.1 and C.2, one finds a multi-electron part (D_{ji}^{-1}), a one-electron angular part ($Y_{lm}(\Omega_{iA})$), and a one-electron radial part, namely,

$$R_1(r,l,m,j,A) = U_l^A(r)\langle Y_{lm}(\Omega)\,|\,\phi_j(i)\rangle. \qquad \text{(C.16)}$$

$R_1(r,l,m,j,A)$, meets the above requirements since its evaluation is the bulk of the calculation, yet it can be stored as a one-dimensional grid (or fit to a grid). The formula required at each step of the walk is

$$U_L^{ECP}[\Psi_T] = \sum_{A}\sum_{i}\sum_{j} D_{ji}^{-1}\sum_{l=0}^{l_{max}}\sum_{m=-l}^{l}\sqrt{4\pi}Y_{lm}(\Omega_{iA})R_1(r_{iA},l,m,j,A). \qquad \text{(C.17)}$$

Both the atomic and molecular cases should now require equivalent amounts of work to evaluate the ECP during the walk.

Several different schemes for storing R_1 were investigated. The best overall scheme involves transforming the radial coordinate, r, onto a finite variable, t, by

$$t = \frac{r}{1+r}, \tag{C.18}$$

then fitting the result using a cubic-exponential spline

$$R_1 = \begin{cases} A_k + B_k t + C_k t^2 + D_k t^3 & \text{if } p_k = 0 \\ A_k + B_k t + C_k e^{p_k t} + D_k e^{-p_k t} & \text{if } p_k \neq 0 \end{cases} \tag{C.21}$$

where $t_k \leq t < t_{k+1}$, and p_k is a tension parameter selected to eliminate spurious inflection points [117].

Appendix D
Evaluation of analytic gradients of the QMC energy

The estimator for the analytic derivative of the QMC energy is $\partial \langle E_L \rangle / X_A$, where X_A is a cartesian degree of freedom of a particular nucleus, A. Expanding this requires evaluaton of $\partial E_L / \partial X_A$, and $\partial \Psi_T / \partial X_A$. In the following sections analytic forms for the various terms are derived.

D.1 Evaluation of $\Psi_T^{-1} \frac{\partial \Psi_T}{\partial X_A}$

Let

$$\Psi_T = S \cdot A = S_{ee} S_{en} |\phi_1 \phi_2 \cdots \phi_{n_\alpha}\rangle |\bar{\phi}_1 \bar{\phi}_2 \cdots \bar{\phi}_{n_\beta}\rangle \tag{D.1}$$

where $|\phi_1 \phi_2 \cdots \phi_{n_\alpha}\rangle$ is the Slater determinant of the α electrons. An operator of the form $\mathbf{O} = \sum_i \mathbf{O}_i$ (i=electron) acting on A, the antisymmetric part of Ψ_T, can be represented in matrix form as,

$$\frac{\mathbf{O}_i A}{A} = \sum_{l=1}^{N_a} O_{il} \bar{A}_{li} \tag{D.2}$$

and,

$$\frac{\mathbf{O} A}{A} = \sum_i \frac{\mathbf{O}_i A}{A} = Tr(O\bar{A}) \tag{D.3}$$

where, $O_{il} = \mathbf{O}_i \phi_l(\mathbf{r}_i)$, and $\bar{A}A = \mathbf{1}$. Furthermore, for a particular generation of the ensemble these quantities will also depend on the electronic configuration. Thus in the case at hand, let

$$\alpha_{il} \equiv \frac{\partial}{\partial X_A} \phi_l(\mathbf{r}_i) \tag{D.4}$$

Therefore, the desired expression is

$$\frac{1}{\Psi_T} \frac{\partial \Psi_T}{\partial X_A} = \sum_i \sum_l \alpha_{il} \bar{A}_{li} = Tr(\alpha \bar{A}). \tag{D.5}$$

D.2 Evaluation of $\partial E_L/\partial X_A$

Given the definition of E_L this becomes,

$$\frac{\partial E_L}{\partial X_A} = \frac{\partial}{\partial X_A}\left[\frac{\mathbf{H}\Psi_T}{\Psi_T}\right] = \frac{\partial}{\partial X_A}\mathbf{V} + \frac{\partial}{\partial X_A}\left[\frac{\mathbf{T}\Psi_T}{\Psi_T}\right] \tag{D.6}$$

where,

$$\frac{\mathbf{T}\Psi_T}{\Psi_T} = \sum_j \left[-\frac{1}{2}\frac{\nabla_i^2 A}{A} - \frac{1}{2}\frac{\nabla_i^2 S_{ee}}{S_{ee}} - \frac{1}{2}\frac{\nabla_i S_{en}}{S_{en}} - \left(\frac{\nabla_j S_{ee}}{S_{ee}} + \frac{\nabla_j S_{en}}{S_{en}}\right)\cdot\frac{\nabla_i A}{A}\right] \tag{D.7}$$

and,

$$\mathbf{V} = V_{ee} + V_{en} + V_{nn}. \tag{D.8}$$

The derivative of the potential is

$$\frac{\partial \mathbf{V}}{\partial X_A} = -\sum_B \frac{Z_A Z_B}{R_{AB}^3}(X_A - X_B) - \sum_i \frac{Z_A}{r_{iA}^3}(x_i - X_A) \tag{D.9}$$

The derivative of the kinetic energy operator is considerably more complex,

$$\begin{aligned}
\frac{\partial}{\partial X_A}\left[\frac{\mathbf{T}\Psi_T}{\Psi_T}\right] = &\sum_i \left[-\frac{1}{2}\frac{1}{A}\frac{\partial}{\partial X_A}\nabla_i^2 A - \left(\frac{\nabla_i S_{ee}}{S_{ee}} + \frac{\nabla_i S_{en}}{S_{en}}\right)\cdot\frac{1}{A}\frac{\partial}{\partial X_A}\nabla_i A\right] \\
&- \sum_i \left[-\frac{1}{2}\frac{\nabla_i^2 A}{A} - \left(\frac{\nabla_i S_{ee}}{S_{ee}} + \frac{\nabla_i S_{en}}{S_{en}}\right)\cdot\frac{\nabla_i A}{A}\right]\frac{1}{A}\frac{\partial}{\partial X_A}A \\
&+ \sum_i \left[-\frac{1}{2}\frac{1}{S_{en}}\frac{\partial}{\partial X_A}\nabla_i^2 S_{en} - \left(\frac{\nabla_i S_{ee}}{S_{ee}} + \frac{\nabla_i A}{A}\right)\cdot\frac{1}{S_{en}}\frac{\partial}{\partial X_A}\nabla_i S_{en}\right] \\
&- \sum_i \left[-\frac{1}{2}\frac{\nabla_i S_{en}}{S_{en}} - \left(\frac{\nabla_i S_{ee}}{S_{ee}} + \frac{\nabla_i A}{A}\right)\cdot\frac{\nabla_i S_{en}}{S_{en}}\right]\frac{1}{S_{en}}\frac{\partial}{\partial X_A}S_{en} \\
&\equiv T'_A + T'_S
\end{aligned} \tag{D.10}$$

T'_A can be evaluated by expanding the determinant as in Eq. D.2. The first term, the derivative of $\dfrac{\nabla_i^2 A}{A}$ becomes,

$$-\frac{1}{2}\frac{1}{A}\frac{\partial}{\partial X_A}\nabla_i^2 A = \sum_l \kappa_{il}\bar{A}_{li} + \sum_{j\neq i}\sum_l \alpha_{il}\bar{K}_{lj}^i \tag{D.11}$$

where,

$$\kappa_{il} \equiv -\frac{1}{2}\frac{\partial}{\partial X_A}\nabla^2\phi_l(\mathbf{r}_i) \tag{D.12}$$

and,

$$\bar{K}_{lj}^i = \text{inverse matrix of } -\frac{1}{2}\nabla_i^2 \times t_i \ , \tag{D.13}$$

where t_i is the local kinetic energy of electron i. A similar expression holds for the gradient term with κ_{il} replaced by γ_{il}, and $\bar{K}_{lj}^i$ replaced by $\bar{G}_{lj}^i$, where

$$\gamma_{il} \equiv -\left(\frac{\nabla_i S_{ee}}{S_{ee}} + \frac{\nabla_i S_{en}}{S_{en}}\right)\frac{\partial}{\partial X_A}\nabla\phi_l(\mathbf{r}_i), \tag{D.14}$$

and,

$$\bar{G}_{lj}^i \equiv \text{inverse matrix of} -\left(\frac{\nabla_i S_{ee}}{S_{ee}} + \frac{\nabla_i S_{en}}{S_{en}}\right)\nabla_i \times g_i \ , \tag{D.15}$$

where g_i is the local gradient of electron i.

Substituting back into equation D.10,

$$T_A' = \sum_i \left[\sum_l \kappa_{il}\bar{A}_{li} + \sum_{j\neq i}\sum_l \alpha_{il}\bar{K}_{lj}^i + \sum_l \gamma_{il}\bar{A}_{li}\right. \tag{D.16}$$

$$\left. + \sum_{j\neq i}\sum_l \alpha_{il}\bar{G}_{lj}^i\right] - T_A\left[\sum_i\sum_l \alpha_{il}\bar{A}_{li}\right]$$

where

$$T_A \equiv \sum_i \left[-\frac{1}{2}\frac{\nabla_i^2 A}{A} - \left(\frac{\nabla_i S_{ee}}{S_{ee}} + \frac{\nabla_i S_{en}}{S_{en}}\right)\cdot\frac{\nabla_i A}{A}\right], \tag{D.17}$$

is the kinetic energy associated with A. Regrouping this gives,

$$\frac{\partial}{\partial X_A}\left[\frac{\mathrm{T}\Psi_T}{\Psi_T}\right] = \sum_i\sum_l \left[\kappa_{il} + \gamma_{il} - T_A\alpha_{il}\right]\bar{A}_{li} \tag{D.18}$$

$$+ \sum_j\sum_{j\neq i}\sum_l \left[\bar{K}_{lj}^i + \bar{G}_{lj}^i\right]\alpha_{il}$$

Let $d_{il} \equiv \kappa_{il} + \gamma_{il} - T_A\alpha_{il}$ and, $\bar{D}_{li} \equiv \sum_{j\neq i}(\bar{K}_{lj}^i + \bar{G}_{lj}^i)$. Then, the final result looks like the chain rule,

$$\frac{\partial}{\partial X_A}\left[\frac{T\Psi_T}{\Psi_T}\right]=\sum_i\sum_l\left[d_{il}\bar{A}_{li}+\alpha_{il}\bar{D}_{li}\right] \tag{D.19}$$

$$=Tr(d\bar{A}+\alpha\bar{D})$$

The remaining term, T_S, depends on the form of S_{en} chosen. Note, however, that its form will be much less complicated than that of T_A, since it does not invlove a determinant. Thus these terms are on par with the evaluation of α_{il}, γ_{il}, and κ_{il}.

D.3 Evaluation of the derivatives of molecular orbitals

The derivative of the molecular oribital is expanded in terms of the constituent atomic orbitals,

$$\frac{\partial\phi}{\partial X_A}=\frac{\partial}{\partial X_A}\sum_{\mu}^{N_A}c_{i\mu}\chi_\mu \tag{D.20}$$

where n_A indicates that only those atomic orbitals centered on nucleus A need be considered. Similar expressions hold for the gradient and kinetic energy operators. Obtaining the derivatives of the $c_{i\mu}$ are beyond the scope of the present work. Instead since in QMC Ψ_T serves only as an importance function, it can be taken to be such that $\partial c_{i\mu}/\partial X_A$ is zero. This will not affect the nodes of Ψ_T and therefore will only impact the energy derivative. Such terms can be evaluated if deemed necessary, by finite difference from SCF wavefunctions. The a.o.'s are functions of the electron-nuclear vector, which has the x component $x=x_i-X_A$ so that

$$\frac{dx}{dx_i}=1 \tag{D.21}$$

and,

$$\frac{dx}{dX_A}=-1\,. \tag{D.22}$$

Thus the derivative with respect to X_A is just the negative value of that with respect to x_i, and the a.o. derivatives presented in section B.4 can be applied. In addition the derivatives of the operators ∇_i and ∇_i^2 are required. For STF's these are,

$$\chi=x^a y^b z^c r^n e^{-gr} \tag{D.23}$$

$$\chi^{-1}\frac{\partial^2\chi}{\partial x^2}=\frac{a(a-1)}{x^2}-\frac{g(2a+1)}{r}+\frac{g^2x^2+(2a+1)n}{r^2}-\frac{gx^2(2n-1)}{r^3}+\frac{nx^2(n-2)}{r^4} \tag{D.24}$$

$$\chi^{-1}\frac{\partial^2\chi}{\partial x\partial y}=(\frac{a}{x}-\frac{gx}{r}+\frac{nx}{r^2})(\frac{b}{y}-\frac{gy}{r}+\frac{ny}{r^2})+\frac{gxy}{r^3}-\frac{2nxy}{r^4} \tag{D.25}$$

$$\chi^{-1}\frac{\partial}{\partial x}\nabla^2\chi=-\frac{2a(a-1)}{x^3}+\frac{2glx}{r^3}-\frac{4n(l-n/2-½)x}{r^4}+\chi^{-1}\nabla^2\chi\chi^{-1}\frac{\partial\chi}{\partial x} \tag{D.26}$$

where $l=a+b+c+n+1$. Relationships for GTF's can be simularly derived from Eq. B.15.

D.4 Evaluation of derivatives of the correlation functions.

From Eq. D.10, one can see that the only terms not covered in Sec. B.5 are the derivatives of the operators ∇_i and ∇_i^2 on S_{en}. Because of their complexity derivatives involving the Pade-Jastrow form of S_{en} are not presented here. Instead, the focus will be on the double-exponential form,

$$S_{en}=\exp\left\{\sum_{iA}-\lambda_a e^{-\nu_A r_{iA}}\right\}. \tag{D.27}$$

Here the variables λ and ν are used instead of a_0 and a_1 or a and b to avoid confusion with the nuclear index A. In addition only a single term in the sum over A need be considered so the index A on λ and ν will be assumed. For simplicity define

$$S=e^{-U}, \tag{D.28}$$

$$U=\lambda e^{-V}, \tag{D.29}$$

$$V=\nu r, \tag{D.30}$$

$$V'=dV/dr. \tag{D.31}$$

The final definition is made so that more general forms of V, *i.e.*, including quadratic terms, can be easily subsituted for the form used here. The derivatives given in Sec. B.5 become,

$$\nabla S=\frac{\mathbf{x}}{r}U\,V'\,S \tag{D.32}$$

$$\nabla^2 S=U\,(V''+(U-1)(V')^2+\frac{2}{r}V')\,S \tag{D.33}$$

and the derivatives of these with respect to x are,

$$\frac{\partial^2 S}{\partial x\partial y}=\frac{xy}{r^2}U\,(V''+(U-1)(V')^2-\frac{1}{r}V')\,S \tag{D.34}$$

$$\frac{\partial^2 S}{\partial x^2} = \frac{x^2}{r^2} U\,(V'' + (U-1)(V')^2 - \frac{1}{r}V')\,S + \frac{1}{r}U\,V'\,S \tag{D.35}$$

$$\frac{\partial \nabla^2 S}{\partial x} = \frac{x}{r} U \left[V''' + \frac{2}{r}V'' + 3(U-1)V''V' - \frac{2}{r^2}V' \right. \tag{D.36}$$

$$\left. + \frac{2}{r}(U-1)(V')^2 + (U^2-3U+1)(V')^3 \right]$$

Using Eqs. D.21 and D.22, then, derivatives such as $\partial^2/\partial X_A\,\partial y_i$ can be evaluated.

REFERENCES AND FOOTNOTES

(1) E. Schrödinger, Phys. Rev. **28**, 1049 (1926).

(2) Applications of the variational principle have recently been reviewed by S. T. Epstein, *The Variational Method in Quantum Chemistry*, (Academic Press, N.Y., 1974).

(3) R. J. Bartlett, Ann. Rev. Phys. Chem. **32**, 359 (1981).

(4) R. O. Jones, Adv. Chem. Phys. **67**, 413 (1987).

(5) J. Oddershede, Adv. Chem. Phys. **69**, 201 (1987).

(6) P. J. Reynolds, W. A. Lester, Jr., Physics Today, S-13 (Jan 1986); D. M. Ceperley, B. J. Alder, Science **231**, 555(1986).

(7) An excellent description of the Hartree-Fock method is given by J. A. Pople, D. L. Beveridge, *Approximate Molecular Orbital Theory* (McGraw-Hill, N.Y., 1970).

(8) P. J. Reynolds, D. M. Ceperley, B. J. Alder, W. A. Lester, Jr., J. Chem. Phys. **77**, 5593 (1982).

(9) P. J. Reynolds, M. Dupuis, W. A. Lester, Jr., J. Chem. Phys. **82**, 1983 (1983).

(10) R. N. Barnett, P. J. Reynolds, W. A. Lester, Jr., J. Chem. Phys. **82**, 2700 (1985).

(11) P. J. Reynolds, R. N. Barnett, B. L. Hammond, R. M. Grimes, W. A. Lester, Jr., Int. J. Quantum Chem. **29**, 589 (1986).

(12) R. N. Barnett, P. J. Reynolds, W. A. Lester, Jr., J. Chem. Phys. **84**, 4992 (1986).

(13) R. M. Grimes, B. L. Hammond, P. J. Reynolds, W. A. Lester, Jr., J. Chem. Phys. **85**, 4749 (1986).

(14) P. J. Reynolds, R. K. Owen, W. A. Lester, Jr., J. Chem. Phys. **87**, 1905 (1987).

(15) B. L. Hammond, P. J. Reynolds, W. A. Lester, Jr., J. Chem. Phys. **87**, 1130. (1987)

(16) J. B. Anderson, J. Chem. Phys. **63**, 1499 (1975); J. Chem. Phys. **65**, 4121 (1976); J. Chem. Phys. **82**, 2662 (1985).

(17) F. Mentch, J. B. Anderson, J. Chem. Phys. **74**, 6307 (1981); J. Chem. Phys. **80**, 2675 (1984).

(18) D. R. Garmer, J. B. Anderson, J. Chem. Phys. **86**, 4025 (1987).

(19) D. R. Garmer, J. B. Anderson, J. Chem. Phys. **86**, 7237 (1987).

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

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

(22) D. W. Skinner, J. W. Moskowitz, M. A. Lee, P. A. Whitlock, K. E. Schmidt, J. Chem. Phys. **83**, 4668 (1985).

(23) J. W. Moskowitz, K. E. Schmidt, J. Chem. Phys. **85**, 2868 (1986).

(24) J. Vrbik, S. M. Rothstein, J. Comp. Phys. **63**, 130(1986); J. Comp. Phys. **74**, 127 (1988).

(25) J. Vrbik, M. F. DePasquale, S. M. Rothstein, J. Chem. Phys. **88**, 3784(1988).

(26) D. F. Coker, R. E. Miller, R. O. Watts, J. Chem. Phys. **82**, 3554 (1985).

(27) D. F. Coker, R. O. Watts, Mol. Phys. **58**, 1113 (1986).

(28) D. M. Ceperley, G. V. Chester, M. H. Kalos, Phys. Rev. A **16**, 3081 (1977).

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

(30) B. J. Alder, D. M. Ceperley, P. J. Reynolds, J. Chem. Phys. **86**, 1200(1982.)

(31) D. M. Ceperley, B. J. Alder, J. Chem. Phys. **81**, 5833 (1984).

(32) G. B. Bachelet, D. M. Ceperley, M. G. B. Chiocchetti, preprint

(33) M. M. Hurley, P. A. Christiansen, J. Chem. Phys. **86**, 1069 (1987).

(34) T. Yoshida, K. Iguchi, J. Chem. Phys. **88**, 1032 (1988).

(35) D. R. Hartree, Proc. Cambridge Phil. Soc. **24**, 89 (1928).

(36) V. Fock, Z. Physik **61**, 126 (1930).

(37) J. C. Slater, Phys. Rev. **35**, 210 (1930).

(38) C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).

(39) G. G. Hall, Proc. Roy. Soc. (London) **205A**, 541 (1951).

(40) J. C. Slater, Phys. Rev. **36**, 57 (1930); C. Zener, Phys. Rev. **36**, 51 (1930).

(41) S. F. Boys, Proc. Roy. Soc. (London) **200A**, 5421 (1950); M. Dupuis, J. Rys, H. F. King, J. Chem. Phys. **65**, 111 (1976).

(42) E. Clementi, A. D. McLean, Phys. Rev. **133A**, 419 (1964).

(43) J. H. van Lenthe, J. G. C. M. van Duijneveldt-van de Rijdt, F. B. van Duijneveldt, Adv. Chem. Phys. **69**, 521 (1987).

(44) T. L. Gilbert, Rev. Mod. Phys. **35**, 491 (1963).

(45) E. Hylleraas, Z. Physik **48**, 469 (1928).

(46) C. L. Pekeris, Phys. Rev. **112**, 1649 (1958).

(47) R. McWeeny, B. T. Sutcliffe, Proc. Roy. Soc. (London) **273A**, 103 (1963).

(48) N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. M. Teller, E. Teller, J. Chem. Phys. **21**, 1087 (1953).

(49) D. M. Ceperley, M. H. Kalos, in *Monte Carlo Methods in Statistical Physics,* K. Binder, ed. (Springer-Verlag, N.Y., 1979).

(50) J. W. Moskowitz, K. E. Schmidt, in *Monte Carlo Methods in Quantum Problems,* M. H. Kalos, ed., (D. Reidel, N.Y., 1984).

(51) K. E. Schmidt, M. H. Kalos, in *Monte Carlo Methods in Statistical Physics II,* K. Binder, ed. (Springer-Verlag, N.Y., 1984).

(52) M. H. Kalos, P. A. Whitlock, in *Monte Carlo Methods, Vol. 1: Basics* (Wiley, N.Y., 1986).

(53) B. H. Wells, in *Methods in Computational Chemistry* **1**, S. Wilson, ed. (Plenum, N.Y., 1987).

(54) S. E. Koonin, *Computational Physics,* (Addison-Wesley, N.Y., 1986).

(55) R. Y. Rubenstein, *Simulation and the Monte Carlo Method,* (Wiley, N. Y. 1981).

(56) J. W. Moskowitz, M. H. Kalos, Int. J. Quantum Chem. **20**, 1107 (1981)

(57) For an excellent introduction to configuration-interaction based methods see, A. Szabo, N. S. Ostlund, *Modern Quantum Chemistry,* (MacMillan, N. Y., 1982)

(58) P. J. Reynolds, S. Alexander, D. Logan, W. A. Lester, Jr., Lecture Notes in Chemistry **44**, M. Dupuis, ed., 280 (1986)

(59) P. J. Davis, P. Rabinowitz, *Methods of Numerical Integration,* (Academic Press, N. Y., 1975

(60) The cross over range between grid methods and Monte Carlo methods has been estimated to be between dimension 6 or 7 to about 12. See Ref. 59.

(61) R. Bianchi, P. Cremaschi, G. Morosi, C. puppi, Chem. Phys. Lett. **148**, 86 (1988).

(62) N. Metropolis, S. M. Ulam, J. Am. Stat. Assoc. **44**, 247 (1949).

(63) P. J. Reynolds, D. M. Ceperley, W. A. Lester, Jr., in preparation.

(64) See the second article in Ref. 16.

(65) C. W. Bauschlicher, Jr., S. R. Langhoff, H. Partridge, P. R. Taylor, J. Chem. Phys. **85**, 3407 (1986).

(66) Such functions have a long history in quantum chemistry, see, for instance, C. C. J. Roothaan, A. W. Weiss, Rev. of Modern Phys. **32**, 194 (1960).

(67) C. J. Umrigar, K. G. Wilson, J. W. Wilkins, "Optimised Trial Functions for Quantum Monte Carlo", Phys. Rev. Lett., submitted.

(68) R. K. Owen, W. A. Lester, private communication.

(69) J. Cho, J. Ng, R. N. Barnett, B. L. Hammond, R. K. Owen, P. J. Reynolds, W. A. Lester, Jr., "The role of the electron-nuclear Cusp in QMC", in preparation.

(70) R. J. Harrison, N. C. Handy, Chem. Phys. Lett. **113**, 257 (1985).

(71) See also, T. Yoshida, K. Iguchi, Chem. Phys. Lett. **143**, 329 (1988), for a description of a method to compute "exact" energies from a VMC walk using the connected moments expansion.

(72) D. M. Ceperley, B. Bernu, "The calculation of excited state properties with QMC.", preprint.

(73) B. Bernu, D. M. Ceperley, W. A. Lester, Jr.,"The calculation of excited state properties with QMC. II. Vibrational excited states.", preprint.

(74) R. N. Barnett, P. J. Reynolds, W. A. Lester, Jr., preprint.

(75) H. M. James, A. S. Coolidge, Phys. Rev. **51**, 860 (1937).

(76) K. McDowell, Int. J. Quantum Chem. Symp. No. 15, 177 (1981).

(77) C. W. Kern, M. Karplus, J. Chem. Phys. **40**, 1374 (1964).

(78) S.-Y. Huang, W. A. Lester, Jr., preprint.

(79) B. H. Wells, Chem. Phys. Lett. **115**, 89 (1985).

(80) C. A. Traynor, J. B. Anderson, Chem. Phys. Lett. **147**, 389 (1988).

(81) J. D. Doll, Chem. Phys. Lett. **81**, 335 (1981).

(82) D. M. Ceperley, J. Stat. Phys. **43**, 815 (1986).

(83) W. J. Stevens, H. Basch, M. Krauss, J. Chem. Phys. **81**, 6026 (1984).

(84) P. J. Reynolds, R. N. Barnett, B. L. Hammond, W. A. Lester, Jr., J. Stat. Phys. **43**, 1017 (1986).

(85) H. Hellmann, J. Chem. Phys. **3**, 61 (1935); H. Hellmann, W. Kassatutschkin, J. Chem. Phys. **4**, 325 (1936).

(86) P. Gombas, Z. Physik **94**, 473 (1935).

(87) J. C. Phillips, L. Kleinman, Phys. Rev. **116**, 287 (1959).

(88) W. A. Goddard III, Phys. Rev. **174**, 659 (1968); C. F. Melius, W. A. Goddard III, Phys. Rev. A **10**, 1528 (1974).

(89) L. R. Kahn, P.Baybutt, D. G. Truhlar, J. Chem. Phys. **65**, 3826 (1976).

(90) P. A. Christiansen, Y. S. Lee, K. S. Pitzer, J. Chem. Phys. **71**, 4445 (1979).

(91) See, for example, C. M. Rohlfing, P. J. Hay, J. Chem. Phys. **83**, 4641 (1985); M. Krauss, W. J. Stevens, J. Chem. Phys. **82**, 5584 (1985); S. W. Wang, K. S. Pitzer, J. Chem. Phys. **79**, 3851 (1983).

(92) V. Bonifacic, S. Huzinaga, J. Chem. Phys. **60**, 2779 (1974); S. Huzinaga, L. Seijo, Z. Barandiaran, M. Klobukowski, J. Chem. Phys. **86**, 2132 (1987).

(93) P. A. Christiansen, J. Chem. Phys. **88**, 4867 (1988); P. A. Christiansen, L. A. LaJohn, Chem. Phys. Lett. **146**, 162 (1988).

(94) T. Yoshida, Y. Mizushima, K. Iguchi, "Electron Affinity of Cl: A model potential-quantum Monte Carlo study", preprint.

(95) B. L. Hammond, P. J. Reynolds, W. A. Lester, "Damped Core QMC: an effective treatment of large Z systems", Phys. Rev. Lett., submitted.

(96) M. Krauss, W. J. Stevens, Ann. Rev. Phys. Chem. **35**, 357 (1984).

(97) C. F. Melius, W. A. Goddard III, L. R. Kahn, J. Chem. Phys. **56**, 3342 (1972).

(98) S. F. Boys, Proc. Roy. Soc. Ser. A **200**, 542 (1950); M. Dupuis, J. Rys, H. F. King, J. Chem. Phys. **65**, 111 (1976).

(99) F. J. Corbato, A. C. Switendick,in *Methods in Computational Physics* **2**, B. Alder, S. Fernbach, M. Rotenberg, eds., (Academic Press, N.Y., 1963).

(100) W. J. Hehre, R. F. Stewart, J. A. Pople, J. Chem. Phys. **51**, 2657 (1969).

(101) I. Shavitt, in *Methods in Computational Physics* **2**, B. Alder, S. Fernbach, M. Rotenberg, eds., (Academic Press, N.Y., 1963).

(102) W. Müller, J. Flesch, W. Meyer, J. Chem. Phys. **80**, 3297 (1984).

(103) Jan Vrbik, M. F. DePasquale, S. M. Rothstein, J. Chem. Phys. **88**, 3784 (1988).

(104) D. D. Konowalow, M. E. Rosenkrantz, M. L. Olsen, J. Chem. Phys. **72**, 2612 (1980).

(105) K. K. Verma, J. T. Bahns, A. R. Rajaei-Rizi, W. C. Stwalley, W. T. Zemke, J. Chem. Phys. **78**, 3599 (1983).

(106) E. Clementi, IBM J. Res. Dev. Sup. **9**, 2 (1975).

(107) H. Hotop, W. C. Lineberger, J. Phys. Chem. Ref. Data **4**, 539 (1975).

(108) Z. Sun, P. J. Reynolds, R. K. Owen, W. A. Lester, Jr., "Monte Carlo study of electron correlation functions for small molecules", submitted to Theoretica Chimica Acta.

(109) P. J. Hay, W. R. Wadt, J. Chem. Phys. **82**, 270 (1985).

(110) P. Pulay, Mol. Phys. **17**, 197 (1969); ibid.**17**, 473 (1970).

(111) P. Pulay, Adv. Chem. Phys. **69**, 241 (1987).

(112) H. Nakatsuji, K. Kanda, T. Yonezawa, Chem. Phys. Lett. **75**, 340 (1980); H. Nakatsuji, T. Hayakawa, M. Hada, ibid. **80**, 94 (1981); H. Nakatsuji, K. Kanda, M. Hada, T. Yonezawa, J. Chem. Phys. **77**, 3109 (1982); H. Nakatsuji, K. Kanda, T. Yonezawa, ibid., 1961 (1982).

(113) H. Wind, J. Chem. Phys. **42**, 2371 (1965).

(114) C. W. Kern, M. Karplus, J. Chem. Phys. **40**, 1374 (1964).

(115) M. Abramowitz, I. A. Stegun, *Handbook of Mathmatical functions*, (Dover, New York, 1964).

(116) L. E. McMurchie, E. R. Davidson, J. Comp. Phys., **44**, 289 (1981).

(117) A. K. Cline, Communications of the Assoc. for Computing Machinery **17**, 218 (1974).

LAWRENCE BERKELEY LABORATORY
TECHNICAL INFORMATION DEPARTMENT
1 CYCLOTRON ROAD
BERKELEY, CALIFORNIA 94720