LBL-30937
UC-401

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

CHEMICAL SCIENCES DIVISION

Quantum Monte Carlo Methods and Lithium Cluster Properties

R.K. Owen
(Ph.D. Thesis)

December 1990

REFERENCE COPY
Does Not
Circulate
LBL-30937
Copy 1
Bldg. 50 Library.

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

DISCLAIMER

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

LBL-30937
UC-401

Quantum Monte Carlo Methods and Lithium Cluster Properties

by

Richard Kent Owen

DISSERTATION

DOCTOR OF PHILOSOPHY

in

PHYSICS

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

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

Quantum Monte Carlo Methods and Lithium Clusters Properties

by

Richard Kent Owen

Abstract

The properties of small lithium clusters with sizes ranging from $n = 1$ to 5 atoms were investigated using quantum Monte Carlo (QMC) methods. Cluster geometries were found from complete active space self consistent field (CASSCF) calculations. A detailed development of the QMC method leading to the variational QMC (V-QMC) and diffusion QMC (D-QMC) methods is shown. The many-body aspect of electron correlation is introduced into the QMC importance sampling electron-electron correlation functions by using density dependent parameters, and are shown to increase the amount of correlation energy obtained in V-QMC calculations. A detailed analysis of D-QMC time-step bias is made and is found to be at least linear with respect to the time-step.

The D-QMC calculations determined the lithium cluster ionization potentials to be 0.1982(14) [0.1981], 0.1895(9) [0.1874(4)], 0.1530(34) [0.1599(73)], 0.1664(37) [0.1724(110)], 0.1613(43) [0.1675(110)] Hartrees for lithium clusters $n = 1$ through 5, respectively; in good agreement with experimental results shown in the brackets. Also, the binding energies per atom was computed to be 0.0177(8) [0.0203(12)], 0.0188(10) [0.0220(21)], 0.0247(8) [0.0310(12)], 0.0253(8) [0.0351(8)] Hartrees for lithium clusters $n = 2$ through 5, respectively. The lithium cluster one-electron density is shown to have charge concentrations corresponding to nonnuclear attractors. The overall shape of the electronic charge density also bears a remarkable similarity with the anisotropic harmonic oscillator model shape for the given number of valence electrons.

Dedication

I dedicate this work to Jesus Christ,
  who has comforted me;
To my supportive father,
  who has guided me;
To my loving wife Marsha,
  who has stood by me;
To my beautiful children,
  who have been the joy of my life.

Contents

Dedication i

Table of Contents iii

List of Figures v

List of Tables vi

Acknowledgements vii

1 Introduction 1

I Quantum Monte Carlo Theory 3

2 Introduction to Monte Carlo Techniques 5
- 2.1 Monte Carlo Evaluations . . . 6
- 2.2 Time Propagator and the Density Matrix . . . 7

3 Imaginary Time Schrödinger Equation 10
- 3.1 Random Flights . . . 10
- 3.2 Quantum Monte Carlo Method . . . 11

4 Importance Sampling 14
- 4.1 *Ab Initio* Wavefunctions . . . 15
- 4.2 Correlation Functions . . . 18
  - 4.2.1 Electron-Nuclear Cusp Condition . . . 19
  - 4.2.2 Types of Correlation Functions . . . 23
  - 4.2.3 Density Dependent Correlation . . . 24

5 Variational QMC 30
- 5.1 Metropolis Algorithm . . . 31
- 5.2 Variational QMC Algorithm . . . 34
- 5.3 Hydrogen Atom Test of V-QMC . . . 36

6 Diffusion QMC 39
- 6.1 Fermi Statistics . . . 39
  - 6.1.1 Fixed-Node Boundary Conditions . . . 40
  - 6.1.2 Released-Node Boundary Conditions . . . 41

6.2 Diffusion QMC Algorithm 42
6.3 Time-Step Bias 44
6.3.1 Finite Time-Step Perturbation Potential 46
6.3.2 Evaluate Time-Step Errors 48
6.3.3 Time-Step Bias of the Hydrogen Atom 50
6.3.4 Short-Time Approximation and Nodal Boundary Conditions 51

II Alkali Metal Clusters **55**

7 Cluster "Magic" Numbers **57**
7.1 Geometrical Packing 57
7.2 Shell Model 58

8 D-QMC Lithium Cluster Results **64**
8.1 Geometry Optimization 65
8.2 QMC Results 66
8.2.1 Best Estimated Non-Relativistic Total Energies 66
8.2.2 Ionization Potential 70
8.2.3 Binding Energy 71
8.3 Lithium One-Electron Density 76
8.3.1 One-Electron Density and the Anisotropic Harmonic Oscillator 76
8.3.2 One-Electron Density and Nonnuclear Attractors 77
8.4 QMC Computational Scaling 88
8.5 Lithium Cluster Time-Step Bias 89
8.6 Conclusion and Future Work 91

Bibliography **94**

Appendices

A Diffusion/Drift/Branching **104**
A.1 Analytic Derivation of QMC Green's Function 104
A.2 Green's Function Stochastic Simulation 106
A.3 Renormalization 107

B Expectation Values **108**
B.1 Energy Expectation Values 108
B.2 Expectation Values of Other Operators 108
B.2.1 Approximate Expectation Values 109
B.2.2 Exact Expectation Values by Future Walking 109

C Program Documentation **111**

List of Figures

2.1 Hit-or-miss Monte Carlo estimation of π 6

4.1 Non-interacting electron gas states in momentum space 24
4.2 Padé Jastrow range 28
4.3 Double exponential range 29

5.1 Variational QMC walk 36
5.2 Hydrogen atom V-QMC results 38

6.1 Diffusion QMC walk 44
6.2 Hydrogen atom D-QMC time-step bias 52
6.3 Hydrogen atom D-QMC time-step bias with acceptance/rejection 53
6.4 Node cross-recross bias 54

7.1 Icosahedral packing of spheres 61
7.2 Li_8 D_{2d} geometry 62
7.3 Li_8 T_d geometry 63

8.1 Lithium cluster geometries 68
8.2 Lithium cluster ionization potentials 73
8.3 Ionization potential odd-even effect 74
8.4 Lithium cluster binding energies 75
8.5 Lithium dimer electron density 79
8.6 'Obtuse' lithium trimer electron density 80
8.7 'Acute' lithium trimer electron density 81
8.8 Planar lithium quadramer electron density 82
8.9 Planar lithium quadramer QMC scatter plot 83
8.10 Planar lithium quadramer electron density 84
8.11 Triplet lithium quadramer electron density 85
8.12 Planar lithium pentamer electron density 86
8.13 Planar lithium pentamer QMC scatter plot 87
8.14 Normalized CPU times 90

A.1 Generalized Quantum Monte Carlo walk 107

List of Tables

4.1 Slater type orbitals 17
4.2 Electron-nuclear cusp condition 22
4.3 Electron-electron correlation functions 23
4.4 Density dependent correlation parameters 26
4.5 V-QMC correlation function comparison 27

5.1 Alternate Metropolis acceptance functions 33

7.1 Anisotropic harmonic oscillator 59

8.1 Lithium trimer geometries 65
8.2 Lithium cluster data 69
8.3 Lithium cluster ionization potentials 71
8.4 Lithium cluster atomization energies 72
8.5 Lithium cluster densities volumes 77
8.6 QMC CPU times 89
8.7 D-QMC time-step bias 91

A.1 QMC Green's function constants 104

Acknowledgements

I would like to thank those who have assisted me in my research; primarily, Prof. William A. Lester, Jr. for providing support and access to the highest quality computational resources, and for valuable assistance in understanding *ab initio* methods, the various current and past members of the Prof. Lester's research group, Dr. Peter Reynolds, Dr. Brian Hammond, Dr. Zhiwei Sun, Dr. Rob Barnett, Dr. Sheng-yu Huang, Dr. Chris Dateo, Dr. Randy Grimes, and María Soto. I would like to also thank Dr. Rama Krishna, Dr. Walter de Heer, Dr. José L. Martins, and Dr. Berni Alder for insightful conversations.

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

Chapter 1

Introduction

The whole of science is nothing more than a refinement of everyday thinking
Albert Einstein "Physics and Reality"

Clusters are an assembly of atoms or molecules which are aggregated together. The inter-particle forces within clusters are relatively weak compared to covalent bonding. For example, rare gas atoms in clusters are attracted to each other through the weak *London* or *dispersion* forces, and metal cluster atoms are held together by the delocalized nature of the valence electrons in a metallic bond. Clusters have been studied extensively both experimentally and theoretically in recent years. Some areas of interest are the origin of the existence of "magic" numbers, transition to bulk systems, dynamics and physical properties such as polarizability, ionization potentials, and binding energies. There are some fine reviews and journal volumes that extensively cover these aspects[55, 21, 78, 97, 113, 10].

Small alkali metal clusters are interesting systems because, on the one hand, their qualitative properties can be understood from simple models[36, 34, 112], but on the other hand, the fine structure can only be understood from detailed analysis. The smaller lithium clusters, the atom and dimer, have been the focus of many *ab initio* studies to much success. For example, spectroscopic constants[77] have been determined and agree with experiment. However, larger clusters have been calculated with mixed success[14, 100]. Either they obtain good ionization potentials or fairly good binding energies, but not both. The goal of this study is to use a method which has achieved consistently good results for smaller systems, such as atoms and dimers[103], and apply it to lithium clusters.

The first part describes the quantum Monte Carlo theory (QMC). The non-relativistic Schrödinger equation is transformed into a diffusion equation, which describes the average behavour of diffusing particles. In the QMC method, random walkers move about a large dimensional space and the distribution of random walkers corresponds to the ground state solution

of the Schrödinger equation. Importance sampling is introduced to help reduce the statistical uncertainty by guiding the random walk with simple computable functions referred to as trial wavefunctions. Forms of trial wavefunctions are discussed, with the introduction of density dependent electron-electron correlation to account for the many-body electron correlation effects. Two QMC methods, variational QMC (V-QMC) and diffusion QMC (D-QMC), are elaborated. The Metropolis algorithm is discussed in connection with V-QMC. To satisfy Fermi statistics for the electrons, boundary conditions are introduced into D-QMC. The time-step bias of the D-QMC method is examined.

The second part concerns itself with alkali metal clusters. first some simple models are discussed which yield "magic" numbers. One model is geometrical packing, the other is the anisotropic harmonic oscillator (AHO). The method for finding optimum lithium cluster geometries is explained for $n = 1, \ldots, 5$, where n is the number of lithium atoms in the cluster. The lithium cluster D-QMC total energies for the atom, its cation, and the dimer are compared with best estimates of non-relativistic total energies. The D-QMC lithium cluster ionization potentials and binding energies are compared to other *ab initio* results[14, 100] and experiment. The one-electron densities of Li_n, for $n = 2, \ldots 5$, are compared to the AHO model. The one-electron densities also exhibit local concentrations of electronic charge, not centered on the nuclei, for all clusters larger than the atom. The D-QMC computational time to obtain a given precision in the total energies for each of the clusters and cations is found to scale with a simple power law with respect to the number of electrons in the system.

The first appendix shows an analytic derivation of the "short-time" Green's function, and describes the Monte Carlo realization of this Green's function. The second appendix discusses expectation values. The last appendix is the documentation of my QMC program.

Part I

Quantum Monte Carlo Theory

Chapter 2

Introduction to Monte Carlo Techniques

And they said every one to his fellow, Come, and let us cast lots, that we may know for whose cause this evil is upon us. So they cast lots, and the lot fell upon Jonah.
Jonah 1:7

Monte Carlo is a term which is often used in physics to describe some numerical procedure or simulation. To the *un-initiated* it conjures up visions of high-rolling gamblers at some foreign casino winning or losing fortunes. For our purposes here it describes something more mundane. Simply ...skillful aiming. Monte Carlo, or stochastic, methods use a certain amount of 'randomness' to evaluate integral equations. By repeating the 'experiment' a number of times you can arrive at a solution with an associated error. The root of the word, stochastic, comes from the Greek – *stockastikos* – skillful at aiming, guessing[51]. Consequently the meaning: random, as in random processes, random variables. As in marksmanship there is a certain degree of randomness due to external influences on a speeding projectile as it hurls towards the target. However, with skillful aiming the marksman with a steady hand can more reliably get the "bull's eye".

Our purposes here are to develop accurate and precise stochastic methods to 'solve' the non-relativistic Schrödinger equation ...the governing equation of molecular systems. Collectively, these methods will be called quantum Monte Carlo (QMC) methods. The following section will motivate why stochastic methods of evaluations may be preferable over the usual numerical procedures.

2.1 Monte Carlo Evaluations

Stochastic methods can be used to evaluate integral equations. Therefore, it is necessary to pose each problem in terms of an integral or integral equation. The reason is that differential equations require some sort of finite difference scheme to approximate the function derivatives numerically, a local property, to find a solution to the differential equation. Integral equations, on the other hand, require finding a global property, an integral. As we shall see, the integral can be sufficiently and accurately evaluated by a finite set of points, a measure 0 subset, of the volume of integration. A very useful property if the volume of integration is infinite or of high dimensionality.

We illustrate this point with a simple *hit-or-miss* Monte Carlo example: the area of a circle or the stochastic evaluation of π (see Fig. 2.1). The integral is of the form:

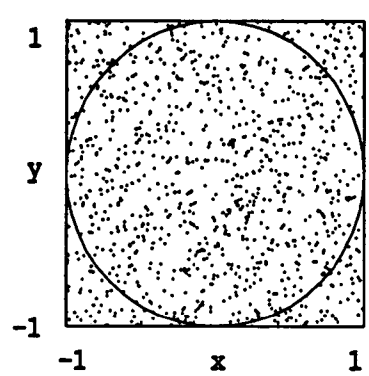
$$\int_{-1}^{1} dx \int_{-1}^{1} dy\, f(x,y) \tag{2.1}$$

where $f(x,y)$ is an indicator function such that

$$f(x,y) = \begin{cases} 1 & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad . \tag{2.2}$$

The usual numerical approach is to partition the area of integration into a set of discrete points

Figure 2.1: Hit-or-Miss estimation of π yields an approximate value of 3.128 for a run of 1000 points.



and evaluate the function at these lattice points. The ratio of the points that fall within the unit circle to all the points within the square will be $\pi/4$, which is the ratio of the area of the circle to the area of the encompassing square. Alternately, the Monte Carlo approach is to uniformly and randomly distribute points within the area of integration and likewise the ratio of the points that fall within the unit circle to all the points within the square also will be $\pi/4$. The first

approach is more accurate for some given n, the number of points, with an error, or uncertainty, decreasing by $1/n$. The latter will have a statistical uncertainty which decreases by $1/\sqrt{n}$.[98]

The advantages of a Monte Carlo approach over the lattice point method are that for high dimensional spaces the number of lattice points increases geometrically with the number of dimensions and to evaluate the integral all points must be sampled before accurate representation of the integral can be found and there is no *a priori* way to determine that a given partitioning will yield the integral with the desired precision. With a Monte Carlo approach you can scatter points randomly through the space, evaluate the mean and standard deviation, and stop when a desired precision has been reached. Techniques of importance sampling, with some *a priori* knowledge of the integrand, can reduce the overall number of points necessary by preferentially sampling the integral in regions of more 'importance'.

A general numerical method to solve a partial differential equation is to break the space into discrete points and evaluate all partial derivatives via finite–difference. Then the solution is determined such that it meets the boundary conditions and satisfies the partial differential equation. The operational words here are "*discrete points*" imply the necessity of partitioning the space into a lattice of points. Since the largest molecular system addressed here is the lithium pentamer with a total of 15 electrons the coordinate space of the electronic wavefunction is 45 dimensional. The wavefunction must be "well-behaved" and normalizable, implying for a bound state solution that the wavefunction has the boundary condition that the electronic wavefunction goes to zero at infinite distance from the Li nuclei.

In general, differential equations can be transformed into integral equations via the Green's function formalism, hence we can speak of 'solving' a partial differential equation when, equivalently, we are evaluating an integral equation. Whether a differential equation or the equivalent integral equation is solved depends on the algorithm or numerical procedure used.

2.2 Time Propagator and the Density Matrix

We impose the Born-Oppenheimer approximation (fixed nuclear positions) and consider the non-relativistic time–dependent electronic Schrödinger equation in position representation,

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V(\vec{R})\psi \equiv H\psi \tag{2.3}$$

where, for N electrons,

$$\begin{aligned}
\vec{R} &= (\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) \\
\nabla &= (\nabla_1, \nabla_2, \ldots, \nabla_N) \\
\nabla^2 &= \sum_{i=1}^{N} \nabla_i^2
\end{aligned}$$

m is the electron mass, $\vec{r}_i$ is the i-th electron position, ∇_i is the usual gradient operator acting on the coordinates of the i-th particle. The total potential V is a sum of Coulomb potentials for all the electrons and nuclei in the system

$$V(\vec{R}) = \frac{1}{2}\sum_{i\neq j}\frac{e^2}{r_{ij}} - \sum_{i\alpha}\frac{Z_\alpha e^2}{r_{i\alpha}} + \frac{1}{2}\sum_{\alpha\neq\beta}\frac{Z_\alpha Z_\beta e^2}{r_{\alpha\beta}} . \tag{2.4}$$

I will use Roman letter indices $i, j, \ldots$ to refer to the electrons and Greek letter indices $\alpha, \beta, \ldots$ to represent the nuclei. By definition, H is the electronic Hamiltonian that represents the energy operator of the system.

The formal solution to Eq. 2.3 exists and is given by[110]

$$\psi(\vec{R}, t) = e^{iHt/\hbar}\psi(\vec{R}, 0) , \tag{2.5}$$

where the exponential factor is the time propagator. Another form of the time propagator can be found from the eigenfunctions and eigenvalues φ_n and ε_n, of the Hamiltonian H to be

$$e^{iHt/\hbar} = \sum_n e^{i\varepsilon_n t/\hbar}|\varphi_n\rangle\langle\varphi_n| . \tag{2.6}$$

In the QMC literature the time propagator is commonly referred to as the Green's function. This causes some confusion with other literature where the Green's function is generally identified as the response function to the time-independent Schrödinger equation[63, 12, 32], $H\psi = E\psi$. The operator form of the Green's function satisfies the following operator equation

$$HG = E\mathbf{1} , \tag{2.7}$$

where $\mathbf{1}$ is the identity operator. The formal solution for the Green's function is

$$G = \frac{E\mathbf{1}}{H} \tag{2.8}$$

and in terms of eigenfunctions and eigenvalues of the Hamiltonian, the Green's function is

$$G = E\sum_n \frac{1}{\varepsilon_n}|\varphi_n\rangle\langle\varphi_n| , \tag{2.9}$$

assuming that the set of eigenfunctions is complete such that

$$\mathbf{1} = \sum_n |\varphi_n\rangle\langle\varphi_n| . \tag{2.10}$$

Mathematically the use of the term "Green's function" for the time propagator is still correct, where solutions to the time-independent Schrödinger equation can be found from the Green's function.

The statistical density matrix operator, ρ, describes the thermal occupation of states for a given inverse temperature $\beta = \frac{1}{kT}$ and is given by

$$\rho = \sum_n w_n|\varphi_n\rangle\langle\varphi_n| \tag{2.11}$$

where $|\varphi_n\rangle\langle\varphi_n|$ is the state projection operator and $w_n = e^{-\beta\varepsilon_n}/Z$ is the Boltzmann distribution function and the partition function $Z = \sum_n e^{-\beta\varepsilon_n}$. The statistical density matrix can be rewritten in an operator form

$$\begin{aligned} \rho &= \frac{1}{Z}\sum_n e^{-\beta\varepsilon_n}|\varphi_n\rangle\langle\varphi_n| \\ &= \frac{1}{Z}\sum_n e^{-\beta H}|\varphi_n\rangle\langle\varphi_n| \\ &= \frac{1}{Z}e^{-\beta H}\sum_n |\varphi_n\rangle\langle\varphi_n| \\ &= \frac{1}{Z}e^{-\beta H}\,, \end{aligned} \tag{2.12}$$

where the eigenfunctions form a complete set. Following the same reasoning, $\mathrm{Tr}\rho = \frac{1}{Z}e^{-\beta H} = 1$, and the partition function $Z = \mathrm{Tr}e^{-\beta H}$. Finally, the operator form of the Boltzmann distribution function is

$$\rho = \frac{e^{-\beta H}}{\mathrm{Tr}e^{-\beta H}}\,. \tag{2.13}$$

The similarity between the time propagator operator and the statistical density matrix is apparent in that both are exponentials of the Hamiltonian operator and the link, $\beta \simeq it/\hbar$, between imaginary time and inverse temperature can be exploited[86]. The QMC method, by going into imaginary time, analytically continues the Schrödinger equation, where the imaginary time propagator becomes a real exponential analogous to the density matrix.

Chapter 3

Imaginary Time Schrödinger Equation

It's as large as life and twice as natural
Lewis Carroll - "Alice in Wonderland"

The fundamental principles of QMC have been around since the 1940's when it has been attributed to Fermi who originally transformed the time independent Schrödinger equation into a diffusion equation by applying the inverse Laplace transform to relate energy E and 'imaginary' time[90]. He also pointed out the property of 'relaxation' to the ground state as it will be discussed below. However, it wasn't until the advent of sophisticated electronic computational machines such as the Los Alamos MANIAC [1] that Monte Carlo solutions could profitably be used for finding the equation of state of a simple two dimensional system[89]. Now with improved hardware and algorithms it is possible to evaluate total energies and properties for simple molecular systems.

3.1 Random Flights

The diffusion equation describes the macroscopic effect of particle diffusion,

$$\frac{\partial \psi}{\partial t} = D\nabla^2\psi \,, \tag{3.1}$$

where D is the 'diffusion' constant, which for particle transport is related to the viscosity. Starting with the random flight problem as described by Chandrasekhar[31], first consider the "drunkard's" walk in one dimension, where the particle either moves to the right or left one step. The description of the probability of displacement from the starting point after N steps is a binomial

[1] Mathematical Analyzer, Numerical Integrator And Computer, see Ref. [40] & [1]

distribution. In the limit of $N \to \infty$ the binomial distribution approaches a Gaussian distribution. The general case of a particle undergoing random steps of arbitrary length at a rate of n steps per unit time can be found by using the Markov method of random flights. The probability that a particle suffers a net displacement $\Delta\vec{r}$ in time Δt is given by

$$\rho(\Delta\vec{r}, \Delta t) = \frac{1}{4\pi D \Delta t^{3/2}} e^{\frac{-|\Delta\vec{r}|^2}{4D\Delta t}} . \tag{3.2}$$

Considering now the probability distribution, $\psi(\vec{r}, t + \Delta t)$, of the particle at some later time $t + \Delta t$ is located at position $\vec{r}$ in terms of the distribution at t is given by

$$\psi(\vec{r}, t + \Delta t) = \int \psi(\vec{r} - \Delta\vec{r}, t)\rho(\Delta\vec{r}, \Delta t)\, d(\Delta\vec{r}) . \tag{3.3}$$

After expanding in a Taylor series in terms of Δt and $\Delta\vec{r}$, in the infinitesimal limit, the diffusion equation is recovered. Since the Brownian motion distribution function satisfies a differential equation that is analogous to the diffusion equation, a relation between the average step length and the diffusion constant can be found, which gives the Einstein relation[46, 47]

$$D = \frac{(\Delta x)^2}{2\Delta t} , \tag{3.4}$$

where Δx is the average step length in one dimension and Δt is the time step size.

3.2 Quantum Monte Carlo Method

To exploit the similarity between the Schrödinger equation and the diffusion equation the time variable in the former must be analytically continued to imaginary time, $\tau = \frac{it}{\hbar}$. This yields the imaginary time Schrödinger equation[2, 4] from Eq. 2.3

$$\frac{\partial\psi}{\partial\tau} = D\nabla^2\psi - \left(V(\vec{R}) - E_T\right)\psi \equiv -(H - E_T)\,\psi , \tag{3.5}$$

where $D = \frac{\hbar^2}{2m}$ is the Einstein diffusion constant and E_T is an arbitrary energy offset. The energy offset, or trial energy E_T as it is called, is introduced for numerical reasons that will be apparent below. Atomic units will be used for the rest of work, where distances are measured in units of Bohrs ($a_0 = \frac{\hbar^2}{m_e e^2} = 0.529$Å), energy will be in units of Hartrees ($= \frac{e^2}{a_0} = 27.2116\, eV$), and imaginary time has dimensions of inverse energy. Consequently, the fundamental constants in the Schrödinger equation have values $\hbar = e = m_e = 1$. The diffusion constant in atomic units has the value of $D = \frac{1}{2}$. The formal solution to Eq. 3.5 is

$$\psi(\vec{R}, \tau) = e^{-(H-E_T)\tau}\psi(\vec{R}, 0) , \tag{3.6}$$

which spectrally resolved in terms of eigenfunctions and eigenvalues is

$$e^{-(H-E_T)\tau} = \sum_n e^{-(\epsilon_n - E_T)\tau}|\varphi_n\rangle\langle\varphi_n| . \tag{3.7}$$

However, it must be stressed that this imaginary time propagator advances a wavefunction forward in imaginary time and hence, does not represent dynamic behavior of the system through phase space. The primary reason for making this transformation becomes evident from the imaginary time propagation of an arbitrary state

$$|\Psi(\tau)\rangle = e^{-(H-E_T)\tau}|\Psi(0)\rangle = \sum_n e^{-(\varepsilon_n - E_T)\tau}|\varphi_n\rangle\langle\varphi_n|\Psi(0)\rangle \,, \tag{3.8}$$

with the long time behavior

$$\lim_{\tau\to\infty}|\Psi(\tau)\rangle = e^{-(\varepsilon_0 - E_T)\tau}|\varphi_0\rangle\langle\varphi_0|\Psi(0)\rangle \,, \tag{3.9}$$

assuming that $n = 0$ represents the lowest energy eigenfunction that is not orthogonal to Ψ. Later it will be shown how this 'decay' to the ground state can be used to determine the ground state energy from the equilibrium distribution. The excited state energies can also be found from this transient behavior[35, 11].

The imaginary time propagator for Eq. 3.5 can be found by using the derivation of Appendix A, and setting $\vec{F} = 0$ and $\alpha = \frac{1}{2}(V(\vec{R}) + V(\vec{R}_0)) - E_T$ as a measure of the potential energy on the path of integration stepping from initial point $\vec{R}_0$ to final point $\vec{R}$. This requires the assumption that for short time steps, equation A.1 is locally correct; even though globally it's not. Hence the imaginary time propagator is

$$\rho(\vec{R}\tau|\vec{R}_0\tau_0) = e^{\alpha(\tau-\tau_0)}\rho_0(\vec{R}\tau|\vec{R}_0\tau_0) \tag{3.10}$$

where we use the familiar free-space imaginary time propagator defined from equation 3.5 with $V(\vec{R}) - E_T = 0$

$$\rho_0(\vec{R}\tau|\vec{R}_0\tau_0) = \left(\frac{1}{4\pi D(\tau-\tau_0)}\right)^{\frac{3N}{2}} \Theta(\tau-\tau_0)e^{\frac{-(\vec{R}-\vec{R}_0)^2}{4D(\tau-\tau_0)}} \tag{3.11}$$

The Monte Carlo realization of equation 3.5 is enumerated as follows:

1. Choose a time step $(\tau - \tau_0)$ sufficiently small.
2. Start with some initial distribution of points in $\mathcal{R}^{3N}$ space.
3. Obtain a new set of points $\vec{R}$ from the initial points $\vec{R}_0$ such that $\vec{R} = \vec{R}_0 + \chi$, where χ is a Gaussian distributed random vector with mean of zero and variance equal to $6ND(\tau - \tau_0)$.
4. Branch or weight each point $\vec{R}$ according to $e^{-\alpha(\tau-\tau_0)}$, where
$$\alpha = \frac{1}{2}\left(V(\vec{R}) + V(\vec{R}_0)\right) - E_T \,. \tag{3.12}$$
5. Adjust E_T such that the branching ratio = 1. Go back to step 3.

The value of the growth energy E_T is a measure of the ground state energy since it is adjusted such that the distribution population growth is constant. This can only occur if $E_T \equiv \varepsilon_0$. However, since this is a stochastic process, there necessarily is some associated uncertainty to the value E_T which can be reduced by sampling E_T after each iteration or block of iterations and averaging. Finally this prescription is only correct for the given time step $\tau - \tau_0$ and a series of evaluations of E_T must be performed at differing time steps and extrapolated to $(\tau - \tau_0) \rightarrow 0$ to obtain the exact energy.

Chapter 4

Importance Sampling

A little knowledge that acts is worth infinitely more than much knowledge that is idle
Kahul Gibran

Rarely is it necessary to approach the solution of the Schrödinger equation without some prior knowledge of the system in question. We already know from the Lebesque square integrability of the wavefunction that the wavefunction should satisfy the boundary condition

$$\Psi \xrightarrow{\vec{R}\to\infty} 0 \tag{4.1}$$

for bound state problems.

The QMC method can benefit from the results of previous *ab initio* calculations. This prior knowledge can be used advantageously to reduce the statistical uncertainty; hence, the computational time necessary to attain a given precision. The question is how to introduce this *a priori* knowledge. This can be accomplished through a method called "importance sampling". As its name implies ... just sample the regions in space where it is more important to do so, in other words, where the wavefunction is large as opposed to regions of small probability (*i.e.* $\Psi = 0$ are the nodes of the wavefunction or regions of zero probability). To this end we define the "trial wavefunction" ψ_T, which is some analytic function that has properties approximating the exact wavefunction. Generally the trial wavefunction will be some modified *ab initio* wavefunction, and should be 'easily' computable with respect to its value, gradient, and Laplacian. Instead of solving for Ψ, we solve for f, which is called the "mixed distribution", where

$$f(\vec{R},\tau) = \psi_T(\vec{R})\Psi(\vec{R},\tau)\,. \tag{4.2}$$

It is a fairly easy exercise to transform equation 3.5 into the imaginary time Schrödinger equation with importance sampling[28, 5, 25, 88]

$$\frac{\partial}{\partial\tau}f \;-\; \underbrace{D\nabla^2 f}_{\text{diffusion}} \;+\; \underbrace{D\nabla\cdot\left(f\vec{F}_Q\right)}_{\text{drift}} \;+\; \underbrace{(E_L - E_T)f}_{\text{branching}} \;=\; 0\,, \tag{4.3}$$

where we define the "quantum force"

$$\vec{F}_Q(\vec{R}) \equiv \frac{2\nabla\psi_T(\vec{R})}{\psi_T(\vec{R})} = \nabla \log\left(\psi_T(\vec{R})^2\right) \tag{4.4}$$

and the "local energy"

$$E_L(\vec{R}) \equiv \frac{H\psi_T(\vec{R})}{\psi_T(\vec{R})} = -\frac{D}{2}\left(\nabla\cdot\vec{F}_Q + \frac{1}{2}\vec{F}_Q^2\right) + V\,. \tag{4.5}$$

The local energy is a scalar energy field that is defined locally at each point in space, and likewise the quantum force is a curlless vector field which can be thought of as the negative gradient of a 'potential' field (*i.e.* $-\log\left(\psi_T(\vec{R})^2\right)$). This last property implies that $\vec{F}_Q$ is a conservative field and that only the end-points are important when evaluating path integrals from some initial point $\vec{R}_0$ to some final point $\vec{R}$.

4.1 *Ab Initio* Wavefunctions

First, we'll consider single determinant wavefunctions such as those obtained from the Hartree-Fock self consistent field (HFSCF) method. The HFSCF wavefunction is an N-particle function comprised of a Slater determinant of 1-particle functions.

$$\Psi_0(\vec{R}) = \det\begin{pmatrix} \phi_1(\vec{r}_1) & \dots & \phi_\kappa(\vec{r}_1) & \dots & \phi_N(\vec{r}_1) \\ \vdots & & \vdots & & \vdots \\ \phi_1(\vec{r}_i) & \dots & \phi_\kappa(\vec{r}_i) & \dots & \phi_N(\vec{r}_i) \\ \vdots & & \vdots & & \vdots \\ \phi_1(\vec{r}_N) & \dots & \phi_\kappa(\vec{r}_N) & \dots & \phi_N(\vec{r}_N) \end{pmatrix} \tag{4.6}$$

The Roothaan approximation[105, 106] is that each 1-particle function, or molecular spin orbital (MSO), is a linear combination of atomic orbitals (LCAO or AO) multiplied by a spin function. The Roothaan procedure changes the Hartree-Fock equation into the Hartree-Fock-Roothaan matrix equation, which is easier to solve for general molecular systems. The set of AOs is called a basis set and incompletely spans the MSO Hilbert space. Some common terminology concerning basis sets are single, double, triple "zeta" which refers to the number AOs with differing ζ-s for each orbital type. Split valence indicates that core orbitals are represented by one , say, single zeta set of orbitals and the valence orbitals are represented by a different number of orbitals, perhaps double zeta. Even though all the AOs contribute to each MSO, in lithium the core orbitals are described primarily by the $1s$ basis functions. By not having double zeta $1s$ AOs one sacrifices some flexibility in obtaining good total energies by not adequately describing the core orbitals, but the practical experience is that the valence orbitals are more

important when considering molecular properties such as dipole moments, ionization potentials, dissociation energies.

Generally I use the restricted Hartree-Fock (RHF) wavefunction, where for each MSO there is another that has the same spatial molecular orbital (MO) function except has opposite spin. If there is an even number of electrons with half of them spin up then it is a closed shell RHF, else it is an open shell RHF. Since the Hamiltonian has no magnetic terms and, therefore, is independent of spin we can relax the spin eigenfunction requirement of the wavefunction, such that expectation values of non-spin dependent observables will be the same. Therefore, a wavefunction which is a product of two Slater determinants, one for the spin up electrons and another for the spin down electrons, will serve adequately to obtain energy expectation values. Defining the MOs in terms of AOs

$$\phi_\kappa(\vec{r}_i) \equiv \sum_\alpha \sum_{p_\alpha} C_{\alpha p_\alpha}\, \vartheta_{p_\alpha}\left(\vec{r}_i - \vec{r}_\alpha\right) \ , \tag{4.7}$$

where each atomic orbital (AO) ϑ_{p_α} is centered on the α-th nuclei and the index p_α labels the AO type (*i.e.* $1s, 2s, 2p_x, \ldots$).

The hierarchy can be continued by making each AO a contraction of Gaussian type orbitals (GTO), of which there are several different GTO basis sets such as Huzinaga's GTO[65], Dunning's contractions of Huzinaga's[43], Pople's 6-31G, etc.[61], and an almost endless list of variants. However, in our case we don't need the integration computational advantages of the GTO, so instead we use the elementary Slater type orbitals (STO)

$$\vartheta_{p_\alpha}\left(\vec{r}_i - \vec{r}_\alpha\right) = R_{nl}^{STO}\left(|\vec{r}_i - \vec{r}_\alpha|\right) Y_{lm}(\Omega) \tag{4.8}$$

where Y_{lm} is the spherical harmonics and R_{nl}^{STO} is defined as follows

$$R_{nl}^{STO}(r) = \sqrt{\frac{8\zeta^2}{(2n)!}}\,(2\zeta r)^{n-1}\, e^{-\zeta r} \tag{4.9}$$

where the ζ parameter may be different for each AO. The STOs are simpler than the hydrogenic orbitals since they don't contain radial nodes for $n > 1$, which correspond to higher excited state, and the form is the same for a given n regardless of azimuthal quantum number l. The orbitals that are actually used for *ab initio* calculations and for the QMC method are linear combinations of the orbitals in Eq. 4.8 for given n, l over differing magnetic quantum number m, such that the orbital is entirely real. See Table 4.1 for a list of the first few.

Since the HFSCF wavefunction is a product of 1-particle functions that were determined in the 'average' field of the other particles, it is necessarily deficient in describing electron-electron correlation, where the dynamics of one electron depends on the state of the other electrons. The best that HFSCF can do is called the Hartree-Fock limit. The difference between the Hartree-Fock limit and the exact non-relativistic energy is defined as the correlation energy, the energy

Table 4.1: Real Slater type orbitals through $n = 3$, with $x_i, x_j = x, y, z$ and $i \neq j$.

orbital	STO
$1s$	$\frac{1}{\sqrt{\pi}}\zeta^{3/2}e^{-\zeta r}$
$2s$	$\frac{1}{\sqrt{3\pi}}\zeta^{5/2}re^{-\zeta r}$
$2p_{x_i}$	$\frac{1}{\sqrt{\pi}}\zeta^{5/2}x_ie^{-\zeta r}$

orbital	STO
$3s$	$\sqrt{\frac{2}{45\pi}}\zeta^{7/2}r^2e^{-\zeta r}$
$3p_{x_i}$	$\sqrt{\frac{2}{15\pi}}\zeta^{7/2}x_ire^{-\zeta r}$
$3d_{x_ix_j}$	$\sqrt{\frac{2}{3\pi}}\zeta^{7/2}x_ix_je^{-\zeta r}$
$3d_{x_i^2}$	$\sqrt{\frac{2}{9\pi}}\zeta^{7/2}x_i^2e^{-\zeta r}$

(4.10)

due to electron-electron correlation. The goal of *ab initio* methods then is to recover as much of the correlation energy as possible. To this end the configuration interaction (CI) expansion is employed, which uses a sum of N-particle wavefunctions.

The CI expansion[118] is a linear combination of determinantal wavefunctions from the HF MOs, obtained from solving the Roothaan equation. The number of MOs generated is equal to the number of AOs, hence the maximum number of N-particle Slater determinants generated is just the number of combinations of $2N_{MO}$ objects taken N at a time, $\begin{pmatrix} 2N_{MO} \\ N \end{pmatrix}$. Then the CI wavefunction is of the form

$$|\Phi\rangle = c_0 |\Psi_0\rangle + c_S |S\rangle + c_D |D\rangle + c_T |T\rangle + c_Q |Q\rangle + \cdots , \tag{4.11}$$

where $S, D, T, Q, \ldots$ represent the set of single, double, triple, quadruple, *etc.* excitations out of the Hartree-Fock determinant. If all possible excitations are taken then that wavefunction represents the 'exact' solution to the many-electron problem ... for that given basis set, and is called a full CI. However, even for relatively small basis sets usually some truncation of the expansion must be made to accommodate computational limitations. These truncations then introduce problems such as the size consistency bias. For example if we do a doubly excited CI (DCI) on a dimer $A + B$, where A and B are infinitely separated, this is clearly different from doing a DCI on monomer A and then on monomer B and adding the total energies. The latter case allows both A and B to be doubly excited which is excluded when $A + B$ are considered together. Hence, the $A + B$ total energy does not equal the total energy of A summed with the total energy of B.

Another problem related to basis set expansions is the basis set superposition error. Since each AO contributes to each MO to some degree then, for example, do an HFSCF on dimer $A + B$ and on monomers A and B. The larger basis set of $A + B$ influences the quality

of the total energy. To obtain the dissociation energy the naïve approach is to subtract the total energy of $A+B$ from the sum of total energies of A and B. However, in the latter case the total HFSCF energy of A was determined with a smaller basis set without the contribution of AOs centered on B, and likewise for B with respect to A. The proper way to overcome the basis set superposition bias is to compute the total energy of A in the presence of 'ghost' orbitals centered where B would be, and conversely for B with respect to A. The QMC method is unaffected by such questions since it will be shown later in chapter 6 that, except for the trial wavefunction nodes, the D-QMC energy is independent of basis set and wavefunction type.

An alternate method based on CI is the multi-reference DCI (MRD-CI), where double excitations are allowed out of, not one, but several reference configurations to partially account for higher excitation terms. Another approach to limiting the CI expansion is to take a given subset of the Slater determinants, say, the single and double excitations out of the HF determinant, and then the N-particle coefficients and the MOs contained within the determinants are both optimized in a self consistent manner. This is called the multi-configuration self-consistent-field (MCSCF) method. If all excitations are allowed for a given subspace of the molecular orbitals and electrons then it is called a complete active space self-consistent-field (CASSCF).

4.2 Correlation Functions

Since the QMC trial wavefunction can be any arbitrary analytic function[103] such that the gradient and Laplacian can be easily computed, we can take advantage of this and introduce correlation functions that can partially account for the particle-particle correlation that is lost from SCF-type wavefunctions. A handy form is a product of the antisymmetric SCF wavefunction Ψ_0 and a symmetric form[23]

$$\psi_T = e^U \Psi_0 \,, \tag{4.12}$$

where $U = \sum'_{ij} f_{ij}(r_{ij}) + \sum_{i\alpha} f_{i\alpha}(r_{i\alpha})$ and the prime on the sum is a short hand way to indicate that the sum excludes $i = j$ terms. The Fermi correlation, due to the Fermi statistics, is present from the antisymmetric SCF wavefunction. The functions $f_{ij}(r_{ij})$ and $f_{i\alpha}(r_{i\alpha})$ are, respectively, electron-electron and electron-nuclear correlation functions which will be defined later. We will also assume that the trial wavefunction and parts are real.

The electron-electron correlation functions are generally 'repulsive' due to the mutual electron-electron Coulomb repulsion. The electrons tend to correlate their behavior such that they avoid each other, of which the SCF wavefunction fails to adequately describe since the 'motion' or one-particle function of one electron was only considered in the average field of the other electrons.

The electron-nuclear correlation functions are included to give some added flexibility to the trial wavefunction such that the electron-nuclear cusp can be satisfied. We ascribe the following properties to the electron-nuclear correlation factor

$$\lim_{r_{i\alpha}\to 0} \left.\frac{\partial f_{i\alpha}(r_{i\alpha})}{\partial r_{i\alpha}}\right|_{r_{i\alpha}} = -Z_\alpha a_{i\alpha} \,. \tag{4.13}$$

In other words the linear term of the electron-nuclear correlation function between the i-th electron and the α-th nucleus is $-Z_\alpha a_{i\alpha}$ as the distance decreases to zero.

4.2.1 Electron-Nuclear Cusp Condition

The electron-nuclear cusp condition can be easily understood from the Schrödinger equation (2.3) by considering that the i-th electron is approaching 'too' close to the α-th nucleus (*i.e.* $r_{i\alpha}$). Then the electronic Hamiltonian can be naïvely approximated by

$$H \approx -D\nabla_i^2 - \frac{Z_\alpha}{r_{i\alpha}} \equiv H_i \tag{4.14}$$

The Laplacian then will be $\nabla_i^2 \approx \frac{\partial^2}{\partial r_{i\alpha}{}^2} + \frac{2}{r_{i\alpha}}\frac{\partial}{\partial r_{i\alpha}} + \cdots$. The $\frac{1}{r_{i\alpha}}$ term in the Laplacian and the potential will become infinitely large unless the two parts cancel each other out. Eliminating the $\frac{1}{r_{i\alpha}}$ cusp as $r_{i\alpha} \to 0$ leads to the following condition

$$-2D\frac{\partial \psi_T}{\partial r_{i\alpha}} - Z_\alpha \psi_T = 0 \qquad \Longrightarrow \qquad \frac{1}{\psi_T}\frac{\partial \psi_T}{\partial r_{i\alpha}} = \frac{-Z_\alpha}{2D} \,, \tag{4.15}$$

and if the wavefunction can be approximated to be constant except for the corresponding correlation factor,

$$\psi_T \approx C e^{f_{i\alpha}(r_{i\alpha})} \,. \tag{4.16}$$

Having the approximate electronic Hamiltonian acting on the approximate trial wavefunction requires with the provisions of Eq. 4.13 that $a_{i\alpha} = \frac{1}{2D}$ to eliminate singularities in $r_{i\alpha}$. However, the assumption that the determinantal part of the wavefunction is constant as $r_{i\alpha} \to 0$ is generally false and requires more careful examination if one wishes to satisfy the cusp condition.

Satisfying the Cusp Condition on Average.

The cusp condition can be satisfied on average by considering the average effect of the other electrons[58]. The one-particle density is defined as

$$\rho(\vec{r}_1) = \int \psi_T^2 \, d\vec{r}_2 \cdots d\vec{r}_N \,. \tag{4.17}$$

We shall assume that the other correlation functions can be neglected when considering the density near the α-th nuclei.

$$\rho(r_{i\alpha}) \xrightarrow{r_{i\alpha}\to 0} Ce^{2f_{i\alpha}(r_{i\alpha})}\hat{\rho}(r_{i\alpha}) \tag{4.18}$$

where C is a constant representative of the effect of the other correlation functions, $e^{2f_{i\alpha}(r_{i\alpha})}$ is due to the correlation function between the i-th electron and the α-th nucleus, and $\hat{\rho}$ is the the SCF one-particle density

$$\hat{\rho}(r_{i\alpha}) = \sum_{\kappa} n_{\kappa}^{occ}\phi_{\kappa}^{2}(r_{i\alpha}) \ , \tag{4.19}$$

n^{occ} is the occupation number for each MO. (for SCF wavefunctions $n^{occ} = 0, 1, 2$, and may be non-integer for CI wavefunctions.)

Take the cusp equation 4.15 and multiply by ψ_T and integrate over all the other electrons except for the i-th electron

$$\int \psi_T \frac{\partial \psi_T}{\partial r_{i\alpha}} dR_{(i)}^{N-1} = \frac{Z_\alpha}{2D}\int \psi_T^2 \, dR_{(i)}^{N-1} \tag{4.20}$$

where $dR_{(i)}^{N-1} = d\vec{r}_1 \cdots d\vec{r}_{i-1} d\vec{r}_{i+1} \cdots d\vec{r}_N$, then

$$\frac{1}{2}\frac{\partial}{\partial r_{i\alpha}}\rho(r_{i\alpha}) = \frac{Z_\alpha}{2D}\rho(r_{i\alpha}) \tag{4.21}$$

with the approximate one-particle density of equation 4.18 as $r_{i\alpha} \to 0$ and determining the linear term of $f_{i\alpha}(r_{i\alpha})$

$$Z_\alpha a_{i\alpha} \quad = \quad \lim_{r_{i\alpha}\to 0}\left(\frac{1}{2\hat{\rho}(r_{i\alpha})}\frac{\partial}{\partial r_{i\alpha}}\hat{\rho}(r_{i\alpha}) + \frac{Z_\alpha}{2D}\right) \tag{4.22}$$

In general the antisymmetric part of the trial wavefunction will partially satisfy the electron-nuclear cusp condition. The value of $a_{i\alpha}$ will tend to be smaller with larger basis sets on each nuclei.

Exact Electron-Nuclear Cusp Condition.

The electron-nuclear cusp condition can be determined exactly for the trial wavefunction of equation 4.12 by using the properties of determinants: $\det AB = \det A \det B$, and $\det(e^B) = e^{\mathrm{Tr}(B)}$. We can combine the electron-nuclear part of the correlation functions with the single-determinant part of the wavefunction. The electron-electron part doesn't contribute to the derivation and will be lumped together in coefficient C. Define the matrix elements $B_{\kappa i} = e^{\sum_\beta f_{i\beta}}\delta_{\kappa i}$ then

$$e^{\sum_{i\beta} f_{i\beta}} = e^{\mathrm{Tr}(B)} = \det(e^B) = \det\begin{pmatrix} e^{\sum_\beta f_{1\beta}} & & & 0 \\ & \ddots & & \\ & & e^{\sum_\beta f_{i\beta}} & \\ & & & \ddots \\ 0 & & & e^{\sum_\beta f_{N\beta}} \end{pmatrix} . \tag{4.23}$$

The wavefunction can be rewritten as

$$\psi_T(\vec{R}) = C \det \begin{pmatrix} \phi_1(\vec{r}_1)e^{\sum_\beta f_{1\beta}} & \dots & \phi_\kappa(\vec{r}_1)e^{\sum_\beta f_{1\beta}} & \dots & \phi_N(\vec{r}_1)e^{\sum_\beta f_{1\beta}} \\ \vdots & & \vdots & & \vdots \\ \phi_1(\vec{r}_i)e^{\sum_\beta f_{i\beta}} & \dots & \phi_\kappa(\vec{r}_i)e^{\sum_\beta f_{i\beta}} & \dots & \phi_N(\vec{r}_i)e^{\sum_\beta f_{i\beta}} \\ \vdots & & \vdots & & \vdots \\ \phi_1(\vec{r}_N)e^{\sum_\beta f_{N\beta}} & \dots & \phi_\kappa(\vec{r}_N)e^{\sum_\beta f_{N\beta}} & \dots & \phi_N(\vec{r}_N)e^{\sum_\beta f_{N\beta}} \end{pmatrix} . \tag{4.24}$$

From this we can see immediately that any operation concerning the i-th electron only affects the i-th row of the determinant.

The cusp condition as given by equation 4.15 can be specified in another way which explicitly depends on the Hamiltonian

$$\lim_{r_{i\alpha}\to 0} r_{i\alpha} H_i \psi_T = 0 . \tag{4.25}$$

Using the LCAO description of the MOs of equation 4.7, The effect of this second cusp condition on an element of the i-th row of equation 4.24 can be found from straight-forward algebra

$$\lim_{r_{i\alpha}\to 0} r_{i\alpha} H_i \phi_\kappa(\vec{r}_i) e^{\sum_\beta f_{i\beta}} = \tag{4.26}$$
$$\left[-2D \sum_{p_\alpha} C_{\alpha,p_\alpha} \vartheta'_{p_\alpha}(r_{i\alpha}=0) + Z_\alpha \left(2Da_{i\alpha} - 1\right) \phi_\kappa(\vec{r}_i = \vec{r}_\alpha)\right] e^{\sum_\beta f_{\alpha\beta}} ,$$

where the above equation follows from equation 4.13 and defining $f_{\alpha\beta} \equiv f_{i\beta}(\vec{r}_i = \vec{r}_\alpha)$. The prime refers to the first derivative to the argument of the AO, and if the AOs are STOs then the only contributions comes from $\vartheta'_{1s}(r_{i\alpha}=0) = \frac{-\zeta^{5/2}}{\sqrt{\pi}}$ and $\vartheta'_{2s}(r_{i\alpha}=0) = \frac{\zeta^{5/2}}{\sqrt{3\pi}}$.

A general solution to equation 4.25 is difficult to find except for special cases. A particular solution can be found by setting each element of the i-th row equal to zero, or in other words find conditions such that equation 4.26 can be made to vanish. This can be accomplished by using AOs of the type that $\vartheta'_{p_\alpha} = 0$, and setting $a_{i\alpha} = \frac{1}{2D}$.

Cuspless Orbitals

One way to satisfy the vanishing-derivative condition is to use AOs that are a sum of GTOs, since each GTO has a quadratic argument in the exponential the first derivative necessarily vanishes.

$$\vartheta_{p_\beta}(\vec{r}_{i\beta}) = \sum_a c_a e^{-\gamma_a(\zeta r_{i\beta})^2} \tag{4.27}$$

To get a fairly good description of a STO usually requires a sum of at least 6 Gaussians noted as STO-6G. For QMC this becomes too computationally intensive and generally only produces

adequate results for V-QMC, but gives excessive branching near the nuclei for D-QMC which in turn underestimates the total energy[1].

A second method is to 'add' a $1s$ STO such that the derivative vanishes.

$$\begin{aligned}\vartheta_{1s}^{\mathrm{C}}(\zeta,r) &= \vartheta_{1s}^{\mathrm{STO}}(\zeta,r)+a_{1s}\vartheta_{1s}^{\mathrm{STO}}(\alpha_{1s},r)\\ \vartheta_{2s}^{\mathrm{C}}(\zeta,r) &= \vartheta_{2s}^{\mathrm{STO}}(\zeta,r)+a_{2s}\vartheta_{1s}^{\mathrm{STO}}(\alpha_{2s},r)\end{aligned}$$

where the $1s$ and $2s$ STO are given from Table 4.1. Since only one condition is given, namely the derivative at $r=0$, then some flexibility exists in the choice for the other condition. Either the intercept ϑ_0 or the range r_0 of the cusplessness may be chosen. (See Table 4.2.) In this study

Table 4.2: Electron-nuclear cusp condition satisfied by addition of 'extra' $1s$ STO to each $1s$ and $2s$ AOs found in the STO basis set.

condition	α_{1s}	a_{1s}	α_{2s}	a_{2s}
range $= r_0$	$\frac{1}{r_0}$	$-\left(\frac{\zeta}{\alpha_{1s}}\right)^{5/2}$	$\frac{1}{r_0}$	$\frac{1}{\sqrt{3}}\left(\frac{\zeta}{\alpha_{2s}}\right)^{5/2}$
intercept $= \vartheta_0$	$\frac{\zeta}{1-\sqrt{\pi}\vartheta_0\zeta^{-3/2}}$	$-\left(\frac{\zeta}{\alpha_{1s}}\right)^{5/2}$	$\frac{\zeta^{5/2}}{\sqrt{3\pi}\vartheta_0}$	$\frac{1}{\sqrt{3}}\left(\frac{\zeta}{\alpha_{2s}}\right)^{5/2}$

the range was specified for the cusplessness and the electron-nuclear correlation function range was set equivalently and the linear term was set to satisfy the electron-nuclear cusp condition.

Another possibility is to construct special functions for the $1s$ and $2s$ such that they have vanishing derivative at the origin and behave similarly to the STO away from the origin. For this purpose I have defined $1c$ and $2c$ AOs

$$\begin{aligned}\vartheta_{1c}(\zeta,r) &= \frac{1}{\sqrt{\pi}}\zeta^{3/2}e^{-\zeta\left(r+\frac{c^2}{r+c}\right)}\\ \vartheta_{2c}(\zeta,r) &= \frac{1}{\sqrt{3\pi}}\zeta^{5/2}\left(r+\frac{c^2}{r+c}\right)e^{-\zeta\left(r+\frac{c^2}{r+c}\right)}\,.\end{aligned}$$

These have the advantage that computationally they are only moderately more complicated than the equivalent STO, whereas the technique of adding compensating $1s$ STOs doubles the

[1]Excessive branching can be understood from observing the difference between STOs, $\vartheta_{1s}^{\vartheta^{\mathrm{STO}}} \propto e^{-\zeta r}$, and GTOs, $\vartheta_{1s}^{\mathrm{GTO}} \propto e^{-\xi r^2}$. The Laplacian for each is

$$\begin{aligned}\nabla^2\vartheta_{1s}^{\vartheta^{\mathrm{STO}}} &= \left[\zeta^2-\frac{2\zeta}{r}\right]\vartheta_{1s}^{\vartheta^{\mathrm{STO}}}\\ &\textit{and}\\ \nabla^2\vartheta_{1s}^{\mathrm{GTO}} &= \left[4\xi^2r^2-6\xi\right]\vartheta_{1s}^{\mathrm{GTO}}\,.\end{aligned}$$

For the STO the Laplacian gets negatively large as $r\to\infty$, and hence the kinetic energy gets positively large thereby offsetting the increase of the potential energy, however, this is not the case for the GTO where the kinetic energy remains bounded. (see chapter 6 for further details on "branching".)

amount of computation done for each $1s$ and $2s$ which comprise the major orbital types for small molecules.

4.2.2 Types of Correlation Functions

Correlation functions can be of any arbitrary form, but it is expedient to choose simple forms for ease of evaluation; however, they should have certain properties[116].

1. They satisfy the cusp condition to some degree, on average or exactly if possible.

2. As the interparticle separation decreases the correlation functions should describe the correct behavior appropriate to the given pair of particles. (*i.e.* electron-electron correlation should be repulsive and maximally so in the limit of vanishing inter-electron distance.)

3. The pseudo-force between the two given particles should vanish as the particles become infinitely separated. This corresponds to the correlation function approaching a constant as interparticle distance increases to infinity.

The correlation functions most often used in QMC are listed in Table 4.3. A wide range of correlation functions have been used for the electron-electron correlation. Currently, only the Padé-Jastrow is used for the electron-nuclear correlation functions.

Table 4.3: Common electron-electron correlation functions used in QMC with SCF wavefunctions. Linear cusp condition is satisfied by a (a_1 for Q P-J).

description	notation	form
Padé-Jastrow	P-J	$\exp\left(\frac{ar}{1+br}\right)$
quadratic Padé-Jastrow	Q P-J	$\exp\left(\frac{a_1r+a_2r^2}{1+b_1r+b_2r^2}\right)$
Sun's double exponential[116]	exp exp	$\exp\left[-b\exp\left(-ar/b\right)\right]$
quadratic double exponential	Q exp exp	$\exp\left[-b\exp\left(-ar/b-cr^2\right)\right]$
Sun's 1 - exponential[116]	(1-exp)	$1-b\exp\left(-a(1-b)r/b\right)$
quadratic 1 - exponential	Q (1-exp)	$1-b\exp\left(-a(1-b)r/b-cr^2\right)$

4.2.3 Density Dependent Correlation

A simple model is proposed to relate the electron density and the electron-electron correlation range. This relation manifests itself by modifying some of the electron-electron correlation functions listed in Table 4.3 to use parameters which are not constant but are dependent on the one-electron density.

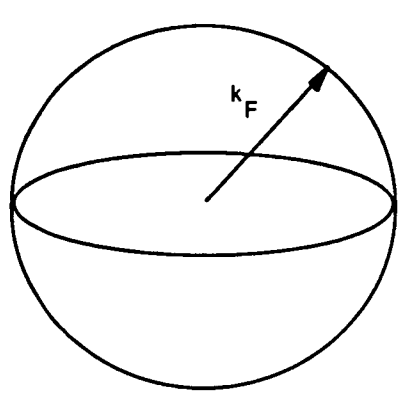
In an earlier paper[116] a simple position-dependent form was proposed for the lithium dimer, such that the b parameter for the (1-exp) electron-electron correlation function was varied depending on the midpoint of the two electrons involved. Roughly speaking if the midpoint is near the bond center then the effective correlation range is shorter than if the two electrons are at some distance away from the molecule where they experience a stronger repulsive correlation.

Density dependent correlation can be easily understood from the simple picture of the free electron gas of N electrons confined to a cubic box of length L. In momentum space (or wavenumber space), at zero temperature, the allowed electronic momentum states form a dense cubic lattice of points that fit within a Fermi sphere of radius k_F, or the Fermi wavenumber. The lattice point separation is $2\pi/L$. The Fermi wavenumber, k_F, is easily found from $N = 2\frac{\frac{4\pi}{3}k_F^3}{(2\pi/L)^3}$ to be

$$k_F = \sqrt[3]{3\pi^2 n} \tag{4.28}$$

where $n = N/L^3$ is the electron number density.

Figure 4.1: Fermi sphere of allowed non-interacting electron gas states in momentum space at T=0.



Due to the electronic many-body effect the electronic Coulomb interaction is screened by other electrons. A simple description of the electron gas response to electric fields is given by the Thomas-Fermi dielectric function in k space

$$\epsilon(k) = 1 + \frac{k_S^2}{k^2} \tag{4.29}$$

where $1/k_S$ is the screening length which is given by

$$k_S^2 = \frac{6\pi n e^2}{\frac{\hbar^2 k_F^2}{2m}} = 4\sqrt[3]{\frac{3}{\pi}}\frac{n^{1/3}}{a_0} \tag{4.30}$$

where a_0 as the Bohr radius. Taking the Fourier transform of the Coulomb potential in momentum space with the Thomas-Fermi dielectric function, leads to the screened Coulomb potential

$$V(r) = \frac{e}{r}\exp(-k_S r)\ . \tag{4.31}$$

The screening term $\exp(-k_S r)$ suggests that the electrons don't fully experience their mutual repulsion unless they are within the screening distance $1/k_S$, hence the range of correlation r_0 depends on the electronic density

$$r_0 \propto \frac{1}{k_S} \propto n^{-1/6}\ . \tag{4.32}$$

However, in a molecule the electron density is non-uniform and can be given by the one-electron density as given by equation 4.19. Furthermore, since the one-electron density may be, and probably is, different at the locations of each electron, a measure of the screened electronic density can be given by the geometric mean of the one-electron density for each pair of electrons that the correlation function is being evaluated for.

The ranges for the given correlation functions can be approximated from Figs. 4.2 and 4.3. The range of the Padé-Jastrow is approximately given by $r_0 \propto 1/b$, and for the double exponential form is $r_0 \propto b$. The final forms used for b in the electron-electron correlation functions are given in Table 4.4, where parameter γ is thrown in to guard against numerical overflows if, say, one of the electrons is a great distance away from the molecule, hence its one-electron density is vanishingly small. However, during parameter optimizations, γ usually vanished and succeeding trials ran with no ill effect.

An example of the effectiveness of density dependent correlation can be seen from a comparison of five trial wavefunctions for the lithium dimer and pentamer, each with identical antisymmetric parts, varying only the electron-electron correlation. 'V-QMC' refers to variational QMC which is a direct measure of how well the wavefunction approaches the exact ground state by how much the correlation energy is recovered. This will be explained later in Chapter 5. We seek a trial wavefunction that 'best' reproduces the exact ground state and provides the least statistical uncertainty. A trade-off must be made between the more computationally intensive forms and less precise simple forms[27]. Some of the efficiency of the density dependent form lies

Table 4.4: Functional forms for the correlation function b parameter corresponding to Table 4.3, such that the correlation function range becomes dependent on the one-electron density $n(r_i)$ for the i-th electron. In the correlation function notation n refers to density dependent correlation and the remainder corresponds to the notation of Tables 4.3.

Notation	$b(n(r_i), n(r_j))$
n-P-J	$\beta(n(r_i)+\gamma)^{1/12}(n(r_j)+\gamma)^{1/12}$
n-exp exp	$\frac{\beta}{(n(r_i)+\gamma)^{1/12}(n(r_j)+\gamma)^{1/12}}$

in the fact that all the necessary information to construct the one-electron density already exists since the antisymmetric part of the wavefunction is a Slater determinant of spin molecular orbitals such that the one-electron density can be simply given by equation 4.19. The results shown in Table 4.5 indicate that correlations functions do increase the amount of correlation energy obtained and is consistent with previous results of Sun *et.al.* [116]. The n-exp exp correlation function returns 2% - 3% less correlation energy than the more sophisticated correlation functions of Sun *et.al.*; however, this form contains only one free parameter as opposed to six, making parameter optimization particularly easy.

In addition, we note that the density dependent form of the Padé-Jastrow consistently performs worse than the globally constant form. This can be attributed to the assumption that the correlation range is inversely proportional to parameter b. For lithium the density dependent value of b is typically of order unity, which corresponds to the region of Fig. 4.2 where $r_0 \propto r^{-\nu}$ such that ν may not be sufficiently close to unity as supposed. To make effective use of the density dependent Padé Jastrow form may also require the optimization of ν, thus eliminating some of the advantage of using a density dependent form in the first place.

Table 4.5: Electron-electron correlation functions are compared for a couple of systems, Li_2 and Li_5. The anti–symmetric part of the trial wavefunction is the same for each with a double 'zeta' HFSCF wavefunction. The lithium dimer internuclear separation is $5.05a_0$, and energies are in atomic units. The lithium pentamer geometry is the same as pictured in Fig. 8.1. The V-QMC energies are in Hartrees and $\%E_{corr}$ is the percentage of correlation energy gained from inclusion of the electron-electron correlation function. In the list of correlation functions the n refers to density dependent correlation and the remainder corresponds to the notation of Tables 4.3 and 4.4.

Correlation	Li_2		Li_5	
Function	V-QMC	$\%E_{corr}$	V-QMC	$\%E_{corr}$
none	-14.8716(16)	-0	-37.2037(23)	0
P-J	[a] -14.9321(8)	49	[b] -37.3415(15)	44
n-P-J	[c] -14.9126(14)	33	[d] -37.2689(21)	21
exp exp	[e] -14.9317(8)	49	[f] -37.3598(14)	50
n-exp exp	[g] -14.9390(8)	55	[h] -37.3718(18)	54

[a] $b = 1.3$
[b] $b = 2.0$
[c] $\beta = 3.0, \gamma = 0.$
[d] $\beta = 4.2, \gamma = 0.$
[e] $b = .250$
[f] $b = .250$
[g] $\beta = .250, \gamma = 0.$
[h] $\beta = .220, \gamma = 0.$

Figure 4.2: The electron-electron Padé Jastrow correlation function parameter b as a function of range r_0. $a = \frac{1}{4}$, or $\frac{1}{2}$ whether like or unlike spins respectively. All values are in atomic units.

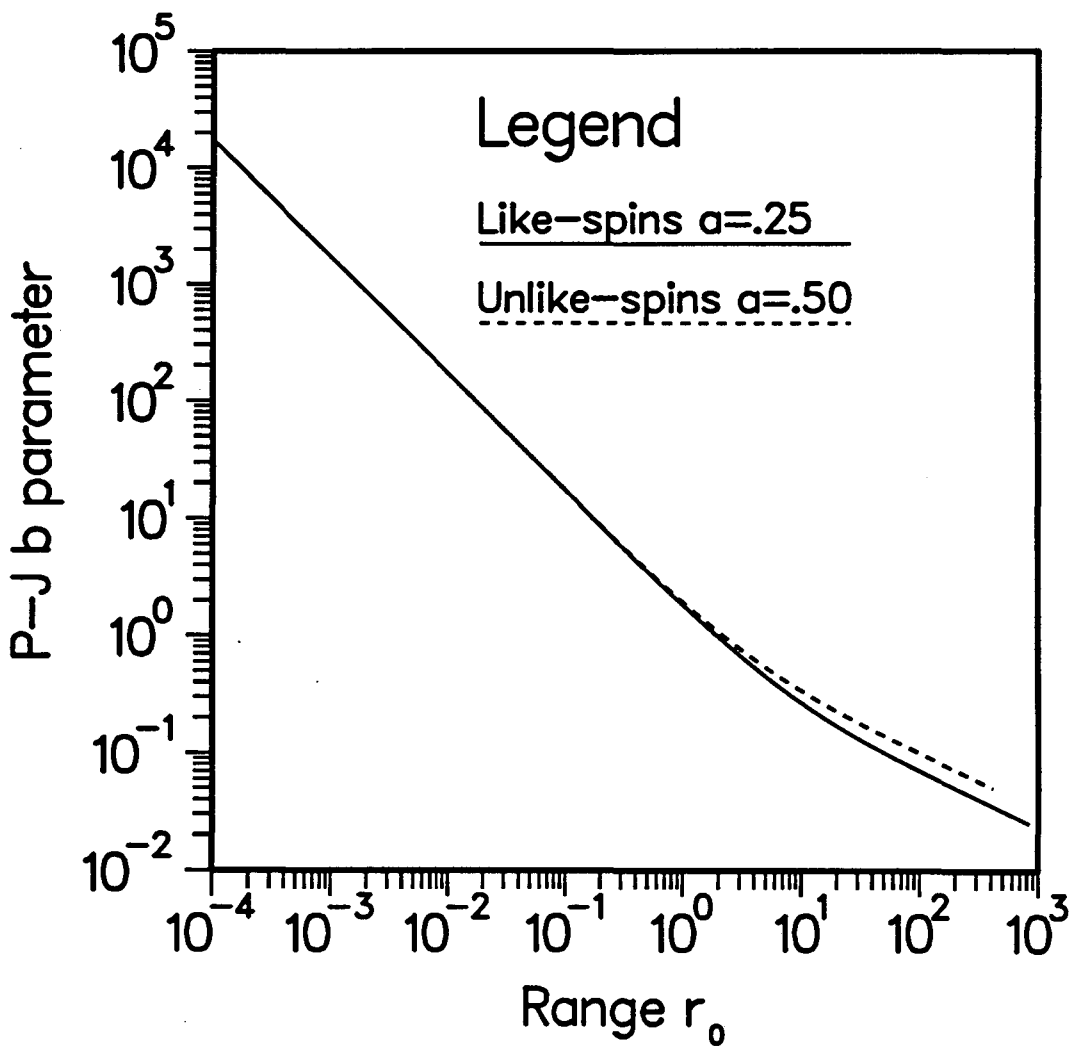
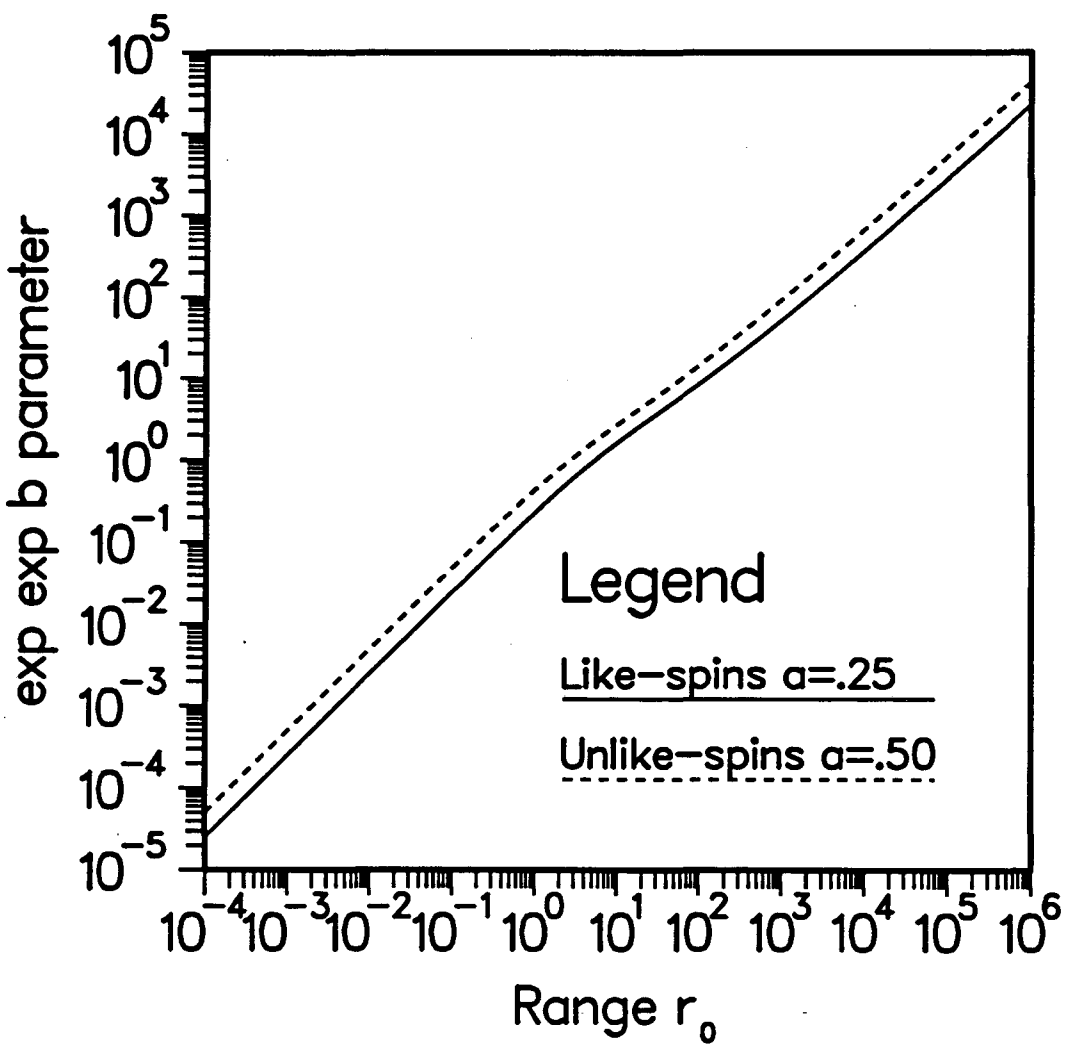


Figure 4.3: The electron-electron double exponential correlation function parameter b as a function of range r_0. $a = \frac{1}{4}$, or $\frac{1}{2}$ whether like or unlike spins respectively. All values are in atomic units.



Chapter 5

Variational QMC

Our wisdom and deliberation for the most part follow the lead of chance
Michel Eyquem de Montaigne

The topic of trial wavefunctions has been quite expansive thus far in the previous chapter. However, the topic would be pointless unless there existed some way to objectively evaluate the relative merits of each class of wavefunctions. An appropriate measure can be found via the Rayleigh-Ritz variational principle

$$\varepsilon_0 \leq \frac{\langle \psi_T | H | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle} . \tag{5.1}$$

Therefore, the wavefunction ψ_T which yields the lowest value for the expectation value on the right then 'best' approximates the exact ground state and the equality only holds if the wavefunction ψ_T is the exact ground state for the Hamiltonian[1]. The stochastic realization of the right-hand integral is most easily accomplished using sample-mean Monte Carlo methods[108] with importance sampling techniques,

$$\frac{\langle \psi_T | H | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle} = \frac{\int \frac{H\psi_T}{\psi_T} \psi_T^2 \, d\vec{R}}{\int \psi_T^2 \, d\vec{R}} \approx \frac{1}{\mathrm{N}} \sum_{i=1}^{\mathrm{N}} E_L(\vec{R}_i) \tag{5.2}$$

where the points $\vec{R}_i$ are distributed according to the probability distribution function (pdf) ψ_T^2.

The generation of sample points according to ψ_T^2 can be accomplished once a differential equation is found such that ψ_T^2 represents the solution. The most likely choice is to look at the imaginary time Schrödinger equation with importance sampling and see what modifications need to be made such that ψ_T^2 will be a solution. Equation 4.3, the partial differential equation for f (now ψ_T^2) returns $(E_L - E_T)f$. Now by subtracting this term from the partial differential equation one then obtains the Fokker-Plank equation

$$\frac{\partial}{\partial \tau} f - D\nabla^2 f + D\nabla \cdot \left(f \vec{F}_Q \right) = 0 , \tag{5.3}$$

[1] Alternately, the statistical variance can be minimized such that the 'best' trial wavefunction would have the lowest statistical uncertainty, where the exact ground state would give a constant value for the local energy and, hence, zero variance (see [42] and [117])

where instead of the mixed distribution as given by Eq. 4.2 the homogenous solution is given by

$$f(\vec{R},\tau) = \psi_T(\vec{R})^2 \,. \tag{5.4}$$

Making the finite time approximation that the quantum force $\vec{F}_Q$ is locally constant (see Appendix A), what follows is the discrete-time Langevin equation

$$\vec{R} = \vec{R}_0 + D\Delta\tau\vec{F}_Q(\vec{R}_0) + \chi \,, \tag{5.5}$$

which simulates the transition probability $T(\vec{R}\tau|\vec{R}_0\tau_0)$ as given in Appendix A, where $\Delta\tau = \tau - \tau_0$. However, due to the short-time approximation there still is a time-step bias, which prevents us from sampling from a distribution of points $\vec{R}_i$ which are distributed according to pdf ψ_T^2. The next section addresses this problem and shows how the short-time approximation of the imaginary time propagator can be circumvented.

5.1 Metropolis Algorithm

In an influential 1953 paper Nicholas Metropolis, *et.al.*[89] proposed a novel technique for generating random variates from a conditional probability distribution function (cpdf), where the probability distribution function (pdf) is known. Providing the known pdf represents the limiting density of the cpdf, *i.e.*, satisfies a homogenous Fredholm equation of the second kind,

$$f(\mathcal{X}) = \int G(\mathcal{X}|\mathcal{X}_0) f(\mathcal{X}_0)\, d\mathcal{X}_0 \,, \tag{5.6}$$

where $\mathcal{X}, \mathcal{X}_0 \in \mathcal{R}^N$, the kernel $G(\mathcal{X}|\mathcal{X}_0)$ is the cpdf to sample from with respect to pdf $f(\mathcal{X})$.

Simply stated, Metropolis sampling is to:

1. propose a move,
2. either accept it or reject it.

Necessary for the Metropolis algorithm is a transition probability $T(\mathcal{X}|\mathcal{X}_0)$ and an acceptance probability $A(\mathcal{X}|\mathcal{X}_0)$. The transistion probability has the property that moves from $\mathcal{X}_0$ will end somewhere in the space,

$$\int T(\tilde{\mathcal{X}}|\mathcal{X}_0)\, d\tilde{\mathcal{X}} = 1 \,. \tag{5.7}$$

The acceptance probability lies between 0 and 1 (and, of course, the rejection probability is given by $1 - A(\mathcal{X}|\mathcal{X}_0)$). The choice of $A(\mathcal{X}|\mathcal{X}_0)$ determines the cpdf. Even though any $T(\mathcal{X}|\mathcal{X}_0)$ can be used, the prudent choice for the transition probability is one that can best approximate the cpdf.

Suppose one has a distribution of points which will be labeled the n-th generation pdf f_n, then the $n+1$-th generation pdf f_{n+1} can be found with the Metropolis algorithm[66],

$$\begin{aligned} f_{n+1}(\mathcal{X}) &= \int A(\mathcal{X}|\mathcal{X}_0)T(\mathcal{X}|\mathcal{X}_0)f_n(\mathcal{X}_0)\,d\mathcal{X}_0 \\ &+ f_n(\mathcal{X})\int \left[1 - A(\tilde{\mathcal{X}}|\mathcal{X})\right] T(\tilde{\mathcal{X}}|\mathcal{X})\,d\tilde{\mathcal{X}} \,. \end{aligned} \tag{5.8}$$

The first term on the right-hand side of the above equation is the accepted move, where some of the points of the n-th generation distribution are moved with respect to the transition probability T and form part of the $n+1$-th generation. The second term on the right-hand side above are the points of the n-th generation which didn't move and now form the rest of the $n+1$-th generation distribution. In the limit as $n \to \infty$, the limiting density is given by $f = \lim_{n\to\infty} f_n$. The exact cpdf $G(\mathcal{X}|\mathcal{X}_0)$ embedded in the Metropolis algorithm of Eq. 5.8 can be expressed by

$$G(\mathcal{X}|\mathcal{X}_0) = A(\mathcal{X}|\mathcal{X}_0)T(\mathcal{X}|\mathcal{X}_0) + \delta(\mathcal{X} - \mathcal{X}_0)\int \left[1 - A(\tilde{\mathcal{X}}|\mathcal{X}_0)\right] T(\tilde{\mathcal{X}}|\mathcal{X}_0)\,d\tilde{\mathcal{X}} \,. \tag{5.9}$$

Metropolis showed that the choice of

$$A(\mathcal{X}|\mathcal{X}_0) = \min\left[1, q(\mathcal{X}|\mathcal{X}_0)\right] \,, \tag{5.10}$$

where

$$q(\mathcal{X}|\mathcal{X}_0) = \frac{T(\mathcal{X}_0|\mathcal{X})f(\mathcal{X})}{T(\mathcal{X}|\mathcal{X}_0)f(\mathcal{X}_0)} \,, \tag{5.11}$$

will generate a cpdf according to $G(\mathcal{X}|\mathcal{X}_0)$ and satisfies detailed balance,

$$G(\mathcal{X}|\mathcal{X}_0)f(\mathcal{X}_0) = G(\mathcal{X}_0|\mathcal{X})f(\mathcal{X}) \,, \tag{5.12}$$

or that transitions from $\mathcal{X}_0 \to \mathcal{X}$ in the equilibrium distribution are as likely as from $\mathcal{X} \to \mathcal{X}_0$. There are other choices that can be made for $A(\mathcal{X}|\mathcal{X}_0)$, but this one is efficient and physically intuitive. Assume for the moment that $T(\mathcal{X}|\mathcal{X}_0)f(\mathcal{X}_0)$ corresponds to a Boltzmann distribution at some temperature T with a uniform density of states noted as $\mathcal{X}$, then Metropolis sampling would correspond to a transition from one state $\mathcal{X}_0$ to another $\mathcal{X}$, where if the new state has lower energy than the first state then it is accepted ($A(\mathcal{X}|\mathcal{X}_0) = 1$), or else it is accepted with a probability ($A(\mathcal{X}|\mathcal{X}_0) < 1$). When viewed in this fashion, and letting $T \to 0$, the process is called simulated annealing[72, 125]. Some of the possible forms of the acceptance probability are shown in Table 5.1, which might be more efficient for some types of transition probabilities T than others. All produce results near the exact expectation value and commensurate uncertainties; however, this has not been tested for more extensive systems. If the second expression for the acceptance function of Table 5.1 is taken to the limit of large ν

$$\lim_{\nu\to\infty} \frac{1}{(1 + q(\mathcal{X}|\mathcal{X}_0)^\nu)^{1/\nu}} = \min[q(\mathcal{X}|\mathcal{X}_0), 1] \,, \tag{5.13}$$

Table 5.1: Alternate Metropolis acceptance functions, $A(q)$, that satisfy detailed balance $A(q) = A(1/q)q$, where the ratio q is given by equation 5.11. It can be shown the for any $s(q)$, such that $\lim_{q\to\infty} q\, s(q) = c_{-1}$ and $\lim_{q\to 0} s(q) = c_0$, where $0 \leq s(q) \leq 1$ and $c_{-1} + c_0 = 1$; then an acceptable choice for $A(q)$ is given by $A(q) = q\, s(q) + s(1/q)$. V-QMC results are tabulated for the ground state of the hydrogen atom with a trial wavefunction $\psi_T(r) = \frac{\zeta^{3/2}}{\sqrt{\pi}} e^{-\zeta r}$, where $\zeta = 1.1 a_0^{-1}$ and energies are in Hartrees.

$s(q)$	$A(q)$	Energy
$\frac{1}{2(1+q)}$	$\frac{q}{1+q}$	$-.494733$ (233)
$\frac{1}{2(1+q^\nu)^{1/\nu}}$	$(1+q^{-\nu})^{-1/\nu}$	$-.494942$ (242)[a]
$\Theta(1-q)$	$\min[q,1]$	$-.494977$ (217)
exact		$-.495$

[a] $\nu = 5$

then one can then retrieve the Metropolis' form of the acceptance function. This may be useful in theoretical considerations when the discontinuous Metropolis acceptance function of equation 5.10 can lead to intractable integrals.

From a careful examination of the Metropolis algorithm it becomes clear that the key requirement is that $f(\mathcal{X})$ must be known analytically. The Metropolis algorithm can only be performed in a variational calculation where $f(\mathcal{X})$, by definition, is given by equation 5.4. Therefore, the Metropolis algorithm allows the evaluation of the Rayleigh-Ritz variational integral without resorting to extrapolation to zero time step. Hence any arbitrary time step can be used, with the requirement that the acceptance ratio should have some reasonable value. As a typical example, the V-QMC results shown in Table 8.2 had a time step of $\tau = 0.100$ and an acceptance ratio approximately 85%, because it was desirable to have the rejection ratio (15% in this case) to be less than the inverse of the number of electrons for the atom with the highest nuclear charge. For lithium clusters where the highest nuclear charge is $Z = 3$, this corresponded to a cut-off rejection of 33%. This contrasts with the general wisdom of adjusting the time step to obtain an acceptance ratio of 50%[8]. This generally allows all the electrons in the molecule to move reasonably often, and minimizes the chance that certain core electrons are 'frozen' and their moves are constantly rejected. The electrons nearest to the nuclei generally have large quantum

forces, which are trying to move them towards the nuclei and, if the time step is large, through the nuclear center to a distance beyond. If the distance is large then the quantum force at the final point may be small, and the ratio $q(\vec{R}\tau|\vec{R}_0\tau_0)$ or the acceptance $A(\vec{R}\tau|\vec{R}_0\tau_0)$ will also be small.

5.2 Variational QMC Algorithm

Collecting all the pieces sections together, simulating a diffusion/drift equation (Appendix A) and Metropolis sampling, leads to the variational QMC (V-QMC) algorithm, where $f \equiv \psi_T^2$. The notation $\langle \mathcal{O} \rangle_f$ is understood to mean the expectation value of some quantity $\mathcal{O}$ with respect to distribution f. The QMC algorithm is readily realized as follows:

1. Choose a time step $(\tau - \tau_0)$ sufficiently small.

2. Start with some initial distribution of points in $\mathcal{R}^{3N}$ space.

3. Obtain a new set of points $\vec{R}$ from the initial points $\vec{R}_0$ such that $\vec{R} = \vec{R}_0 + D(\tau - \tau_0)\vec{F}_Q(\vec{R}_0) + \chi$, where χ is a Gaussian distributed random vector with mean of zero and variance equal to $6ND(\tau - \tau_0)$.

4. Accept or reject the move with probability

$$A(\vec{R}\tau|\vec{R}_0\tau_0) = \min\left[1, q(\vec{R}\tau|\vec{R}_0\tau_0)\right] , \tag{5.14}$$

where

$$q(\vec{R}\tau|\vec{R}_0\tau_0) = \frac{T(\vec{R}_0\tau_0|\vec{R}\tau)\psi_T(\vec{R})^2}{T(\vec{R}\tau|\vec{R}_0\tau_0)\psi_T(\vec{R}_0)^2} , \tag{5.15}$$

and

$$T(\vec{R}\tau|\vec{R}_0\tau_0) = \left(\frac{1}{4\pi D(\tau - \tau_0)}\right)^{3N/2} \Theta(\tau - \tau_0)\, e^{\frac{-(\vec{R} - \vec{R}_0 - D(\tau - \tau_0)\vec{F}(\vec{R}_0))^2}{4D(\tau - \tau_0)}} . \tag{5.16}$$

5. Increase the total for N, evaluate and sum $E_L(\vec{R})$ to the total $\sum_{i=1}^{N} w_i E_L(\vec{R}_i)$, where the weighting factor $w_i = 1$ for a V-QMC evaluation.

6. Go back to step(3), unless a sufficient number of points have been sampled.

7. Return $E = \frac{1}{N}\sum_{i=1}^{N} E_L(\vec{R}_i)$, as an estimate for $\frac{\langle H \rangle_f}{\langle 1 \rangle_f}$

To obtain an estimate for the variance of the mean then call the above procedure a 'block' with $E_j \leftarrow E$ and maintain new sums $\sum_{j=1}^{N_{block}} E_j$ and $\sum_{j=1}^{N_{block}} E_j^2$. After a sufficient number of blocks

have been computed then return the expectation value

$$E = \frac{1}{\mathrm{N}_{block}} \sum_{j=1}^{\mathrm{N}_{block}} E_j \,, \tag{5.17}$$

and the standard deviation of the mean

$$\sigma_E = \sqrt{\frac{\left(\frac{1}{\mathrm{N}_{block}} \sum_{j=1}^{\mathrm{N}_{block}} E_j^2\right) - E^2}{\mathrm{N}_{block} - 1}} \,. \tag{5.18}$$

A diagrammatical representation of a random walk is given by Fig. 5.1, where a single configuration or random walker is alternately propagated through $3N$ dimensional space by drifts (depicted with solid vectors) and diffusive steps (shown as dashed lines).

In addition to generating a set of points $\vec{R}_i$ that are distributed according to pdf ψ_T^2, the algorithm also explicitly states how expectation values of the energy with respect to wavefunction ψ_T (see Eq. 5.2) and the variance of the energy can be determined. The energy and variance can be part of a feed-back loop that either changes the trial wavefunction or signals the end of the sampling.

Other expectation values can be similarly evaluated to as precise an uncertainty as desired. There are two expectation values which are are evaluated in different ways, but are equivalent if the set of points $\vec{R}_i$ are distributed according to ψ_T^2. The expectation values are

$$\begin{aligned} \langle \psi_T \,|\vec{p}^{\,2}|\, \psi_T \rangle &= \int \psi_T^\dagger \vec{p}^{\,2} \psi_T \, d\vec{R} && (5.19) \\ &= \hbar^2 \int \frac{\nabla^2 \psi_T}{\psi_T} \psi_T^2 \, d\vec{R} && (5.20) \\ &\equiv \hbar^2 \left\langle \frac{\nabla^2 \psi_T}{\psi_T} \right\rangle_{\psi_T^2} && (5.21) \end{aligned}$$

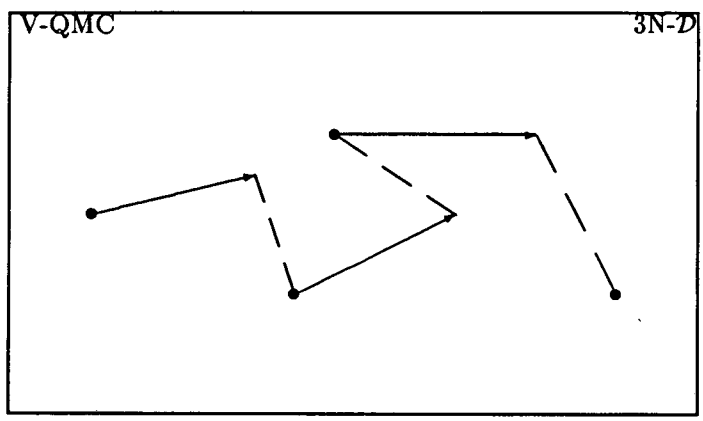
and

$$\begin{aligned} \langle \psi_T \,|\vec{p}^{\,2}|\, \psi_T \rangle &= \int \psi_T^\dagger \vec{p}^{\,\dagger} \cdot \vec{p} \psi_T \, d\vec{R} && (5.22) \\ &= \hbar^2 \int \left(\frac{\nabla \psi_T}{\psi_T} \right)^2 \psi_T^2 \, d\vec{R} && (5.23) \\ &\equiv -\hbar^2 \left\langle \left(\frac{\nabla \psi_T}{\psi_T} \right)^2 \right\rangle_{\psi_T^2} \,, && (5.24) \end{aligned}$$

where ψ_T is assumed to be real and normalized, $\langle \psi_T | \psi_T \rangle = 1$, Since the momentum operator $\vec{p}$ is Hermitian then the expectation value of the local properties $-\left(\frac{\nabla \psi_T}{\psi_T}\right)^2$ and $\frac{\nabla^2 \psi_T}{\psi_T}$ are identical. The expectation values for either of the quantities should be identical within statistical uncertainty if and only if the distribution has reached equilibrium with $f \equiv \psi_T^2$. However, as personal experience has shown for the lithium clusters calculations (see section 8.2), one in five randomly

distributed ensembles will not reach equilibrium under any circumstances. Typically when a recalcitrant ensemble of this type occurs the expectation value $\langle(\nabla\psi_T)^2\rangle_{\psi_T^2}$ will be several orders of magnitude greater than $-\langle\psi_T\nabla^2\psi_T\rangle_{\psi_T^2}$. This can be understood by assuming that one of the electrons is near a node of ψ_T or near a nuclear center. Then the initial quantum force $\vec{F}_Q(\vec{R}_0)$ is large, but the final quantum force $\vec{F}_Q(\vec{R})$ may not be. This will cause the acceptance ratio to be quite low, since transitions from $\vec{R}$ back to $\vec{R}_0$ will be unlikely. To remedy such behavior either the time step can be reduced significantly or just try another ensemble until equilibrium can be established.

Figure 5.1: A schematic representation of a variational QMC walk, where a random walker drifts (solid arrows) according to trial wavefunction ψ_T, then diffuses (dashed lines) with some root-mean-square distance, and at each point Metropolis accept or reject the drift-diffuse move. The diagram shows 3 successful steps.



5.3 Hydrogen Atom Test of V-QMC

The hydrogen atom, the constant friend of the physicist, provides a simple, well studied, and analytically solvable atomic system that lends itself to verification of numerical procedures. The Hamiltonian for the hydrogen atom in atomic units is

$$H = D\nabla^2 - \frac{1}{r}\,, \tag{5.25}$$

where r is the radial distance from the nucleus to the electron.

The exact ground state wavefunction is $\varphi_0(\vec{r}) = \frac{1}{\sqrt{\pi}}e^{-r}$ with energy $\varepsilon_0 = \frac{-1}{2}$ Hartrees. We assume that the trial wavefunction has a similar form such as

$$\psi_T(\vec{r}) = \frac{\zeta^{3/2}}{\sqrt{\pi}}e^{-\zeta r}\,, \tag{5.26}$$

where the local energy field is found from the product of the reciprocal trial wavefunction with the operation of the Hamiltonian on the trial wavefunction (see Eq. 4.5) to be

$$E_L(\vec{r}) = -D\zeta^2 + \frac{2D\zeta - 1}{r}\,, \tag{5.27}$$

and the quantum force is given by twice the gradient of the logarithm of the trial wavefunction (see Eq. 4.4) to be

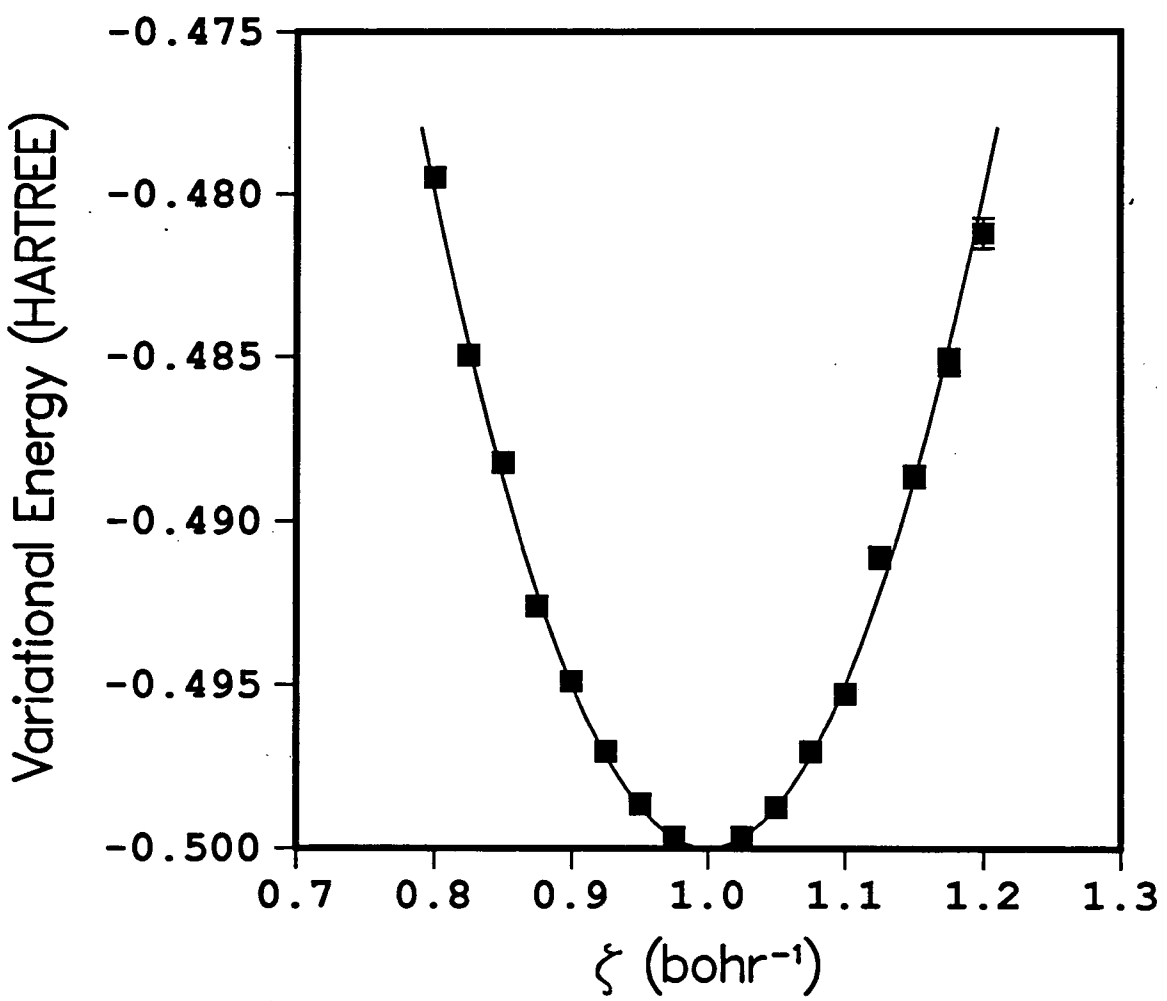
$$\vec{F}_Q(\vec{r}) = -2\zeta\hat{r}\,, \tag{5.28}$$

the unit vector $\hat{r} = \frac{\vec{r}}{r}$ is in the radial direction away from the nucleus. The energy expectation value with respect to the ψ_T^2 distribution is easily found to be

$$\langle H \rangle_{\psi_T^2} = \zeta\,(D\zeta - 1)\ . \tag{5.29}$$

The V-QMC test results are given in Fig. 5.2 and shows excellent agreement between numerical evaluations and the exact analytic result.

Figure 5.2: The hydrogen atom V-QMC test case with differing parameterized trial wavefunctions of the form $\psi_T(\vec{r}) = \frac{\zeta^{3/2}}{\sqrt{\pi}}\exp(-\zeta r)$. The solid line is the exact analytic variational energy in atomic units $[\zeta((\zeta/2-1)\,\text{Hartree}]$. The blocks are the V-QMC results, all computed with a time step $(\tau - \tau_0) = 0.1\,\text{Hartree}^{-1}$.



Chapter 6

Diffusion QMC

Probabilities direct the conduct of the wise man
Cicero - "De Natura Deorum"

6.1 Fermi Statistics

Up to this point little has been said concerning spin statistics of the electrons for the non-relativistic electronic Hamiltonian. It is well known that for a fermion system the total wavefunction needs to be totally antisymmetric on the interchange of any two identical particles (*i.e.* the electrons). Previously, it has been shown in section 3.2 that solutions to the imaginary time Schrödinger equation 3.5 will asymptotically approach the lowest eigenstate of the same symmetry as the initial wavefunction. However, the ground state for an arbitrary Hamiltonian is the bosonic ground state since the fermion statistics are not inherent to the Hamiltonian but are an artifact of the imposition of auxiliary boundary conditions through the Pauli exclusion principle.

The goal is to stochastically simulate solutions to the Schrödinger equation by a finite set of points such that expectation values may be performed. Suppose the initial distribution of points, $\vec{R}_i$, in the $3N$ dimensional space is closely, or exactly, distributed according to the square of the fermion ground state wavefunction. Clearly for *any* finite distribution of discrete points the distribution is given by

$$f(\vec{R}) = \sum_{i=1}^{N} \delta(\vec{R} - \vec{R}_i) \,, \tag{6.1}$$

and the spectral resolution of any finite sum of Dirac delta functions must necessarily yield an infinite expansion of all eigenfunctions. Hence, with any finite distribution of points there will always be a component of the bosonic ground state present and, with long enough imaginary time

propagation, it will come to dominate. One way this dilemma can be eliminated is by imposing boundary conditions such that the nodes of the trial wavefunction represent the nodes of the solution, then by imposing the condition that the the trial wavefunction satisfies antisymmetry, the ground state solution will also satisfy antisymmetry and will be a linear combination of all the antisymmetric eigenfunctions

$$\varphi_0{}' = \sum_{\{a\}} C_a \varphi_a \ , \tag{6.2}$$

where $\{a\}$ represents the set of all antisymmetric eigenfunctions of the Hamiltonian. It is then desirable to find ψ_T such that its nodes are in some measure 'close' to that of the desired eigenstate such that its coefficient C_0 is nearly unity.

6.1.1 Fixed-Node Boundary Conditions

With this imposition of fermion boundary conditions (called the fixed–node approximation) it's clear to see that that only the energy expectation value of the antisymmetric ground state will be a strict upper bound. The excited state energy expectation values will underestimate the exact excited state energies. This can be understood from considering each of the different nodal regions, or each volume of $3N$ dimensional space encompassed by a $3N-1$ dimensional surface. Suppose each region is equally populated with random walkers, then let each random walker propagate forward in imaginary time with branching. As shown in Appendix A, branching favors those regions with lower average local energy E_L. After a long enough time, due to repeated renormalization of the population, the only random walkers left are in the region of lowest average energy. Since the average of energy from all the regions is assumed to be relatively close to the exact energy, the energy expectation value computed will underestimate the excited state energy. This does not apply to the antisymmetric ground state since the only nodes of the trial wavefunction, and hence the approximate ground state solution, are due to the electron exchange and not to excited states. It is argued that each nodal region must be identical in volume since each can be mapped onto another by exchange. Furthermore, the energy expectation value is identical for each region since the local energy E_L and the trial wavefunction ψ_T can also be mapped identically by exchange in each region. Therefore, the average energy expectation values of each nodal region will be the same hence the aggregate energy expectation value will be unbiased and will represent an upper bound to the antisymmetric ground state.

The fixed-node D-QMC energy obtained will quadratically vary from the exact energy for 'small' nodal displacement[26]. This can most easily be seen from a simple example of two non-interacting spin $\frac{1}{2}$ fermions both with spin up in a box of length $2a$. We set the left and right boundaries at $-a+\delta a$ and $a+\delta a$ respectively thus offsetting the box by δa. All the one particle eigenfunctions will have nodes at the left and right boundaries. We take for the trial wavefunction

a Slater determinant of the following orbitals the first is exactly the $n = 1$ eigenstate with energy E_1 and the second is a 'close' approximation to the $n = 2$ eigenstate with energy E_2 where the middle node is displaced by δa,

$$\phi_1(x) = \frac{1}{\sqrt{a}} \cos\left(\frac{\pi(x - \delta a)}{2a}\right) , \tag{6.3}$$

and

$$\phi_2(x) = \begin{cases} \frac{1}{\sqrt{a}} \sin\left(\frac{\pi x}{a - \delta a}\right) & -a + \delta a < x < 0 \\ \frac{1}{\sqrt{a}} \sin\left(\frac{\pi x}{a + \delta a}\right) & 0 < x < a + \delta a \end{cases} . \tag{6.4}$$

In each nodal region our choice of trial wavefunction also corresponds to an eigenfunction, hence our trial wavefunction is just the ground state wavefunction that a fixed-node D-QMC would return. After fairly lengthy evaluations the expectation value of the particles-in-a-box Hamiltonian $H = \frac{p_x^2}{2m}$ can be found to be

$$E_{\text{tot}} = E_1 + \frac{E_2}{1 - \delta^2} , \tag{6.5}$$

where $\delta = \delta a / a$. A Taylor expansion of $(1 - \delta^2)^{-1} = 1 + \delta^2 + \cdots$ explicitly shows that the fixed-node D-QMC energy varies quadratically for small nodal displacement from the exact energy. The *caveat* is that for high dimensional spaces there's no guarantee that *ab initio* generated nodal surfaces will have the same topology as the exact nodal surface[74, 6, 19] and the question of small nodal displacement becomes vague.

Nodal boundary conditions are imposed on the D-QMC walk by requiring random walkers that cross over a node to vanish[31]. This has the effect of depleting the number of random walkers in the neighborhood of the node, thus making the probability distribution become vanishingly small at the node.

6.1.2 Released-Node Boundary Conditions

An alternate to additional boundary conditions is to allow some sort of transient estimator that lets random walkers cross the nodes and negatively contribute to the expectation values. This method is called "released-node", where an initial configuration of random walkers is distributed close to the fermion ground state distribution, and as the distribution propagates forward in imaginary time, the symmetric part is projected out.

The initial distribution will correspond to a wavefunction that is a sum of all eigenfunctions of the Hamiltonian

$$\varphi' = \sum_{\{s\}} C_s \varphi_s + \sum_{\{a\}} C_a \varphi_a , \tag{6.6}$$

where s is the set of symmetric eigenfunctions and a is the antisymmetric, and the fermion ground state coefficient C_{a0} is assumed large. If this wavefunction is propagated in imaginary

time according to the operator form of the imaginary time propagator $\exp(-(H-\varepsilon_{a0})\tau)$, where the trial energy is set to the antisymmetric ground state energy, ε_{a0}; then φ' will approach

$$\varphi' \rightarrow C_{s0}e^{-(\varepsilon_{s0}-\varepsilon_{a0})\tau}\varphi_{s0} + C_{a0}\varphi_{a0} \,. \tag{6.7}$$

It has also been assumed that the symmetric and antisymmetric states will decay to the boson and fermion ground states, respectively. For long enough imaginary times the bosonic ground state will come to dominate and all traces of the fermionic nature will disappear. Therefore, the transient estimation can only allow the propagation of the initial wavefunction some fixed time or given number of steps[26, 29, 69]. If an antisymmetric trial wavefunction, ψ_T, is projected onto the wavefunction then the symmetric part will not contribute, by symmetry, to either the overlap, $\langle\psi_T|\varphi'\rangle = C_{a0}$, or the expectation value of the Hamiltonian, $\langle\psi_T|H|\varphi'\rangle = C_{a0}\varepsilon_{a0}$. Thus the antisymmetric ground state energy can be determined from

$$\frac{\langle\psi_T|H|\varphi'\rangle}{\langle\psi_T|\varphi'\rangle} = \varepsilon_{a0} \approx \frac{\sum_i \sigma_i E_L(\vec{R}_i)}{\sum_i \sigma_i} \,, \tag{6.8}$$

where $\sigma_i = \pm 1$ is negative if the random walker crossed an odd number of nodes and positive if the converse, and the points $\vec{R}_i$ are distributed according to $|\psi_T\varphi'|$. Since the distribution of points represents a probability distribution and by definition can not have negative values, the sign of $\psi_T\varphi'$ is transferred to E_L via σ_i. The problem with transient estimators is that the variance grows with time and may become too large before the expectation value approaches the antisymmetric ground state energy.

Alternate schemes to release the nodal boundary conditions rely on propagating a dual set of 'positive' and 'negative' random walkers. The random walkers may, somehow, interact by either cancelling each other out when sufficiently close[39] or by 'repelling' each other through a non-linear interaction in the Hamiltonian[20].

6.2 Diffusion QMC Algorithm

The diffusion QMC procedure is almost identical to the V-QMC algorithm of chapter 5, except for the inclusion of branching as given by the local energy exponential. In the imaginary time Schrödinger equation with importance sampling (Eq. 4.3) there is the "branching" term that, neglecting all other spatial derivative terms, yields a rate equation with the growth $\approx e^{-(E_L-E_T)\tau}$. For (E_L-E_T) large and negative, the growth becomes large, while conversely, for (E_L-E_T) large and positive, there is rapid decay. In other words, the sign of the quantity $E_L - E_T$ represents a source or sink of diffusing particles respectively.

As with the V-QMC method of chapter 5, the short time approximation is made for the imaginary time propagator, where the path of propagation is from $\vec{R}_0$ to $\vec{R}$ with a measure

of the branching $E_L - E_T$ along the path of propagation given by setting α of Appendix A to

$$\alpha = \frac{E_L(\vec{R}) + E_L(\vec{R}_0)}{2} - E_T \ . \tag{6.9}$$

Some other definition could be employed, but the expediency of this choice will be outlined later when the time-step bias error is discussed. (see section 6.3).

The D-QMC algorithm is identical to the V-QMC algorithm of page 34, except for the inclusion of an extra step between steps 4 and 5 which we shall label 4.5:

4.5 Branch configuration by factor $w_i = e^{-\alpha(\tau-\tau_0)}$; either by making $m = [w_i + u]$ copies, where u is a uniform random variate between 0 and 1 and $[z]$ is the largest integer less than or equal to z (then $w_i = 1$ in the following step); or carrying along a weight factor w_i associated with the configuration.

In this study the weighting factor w_i is simulated by creating m copies of the current random walker[103, 22] and can be shown that on average

$$w_i = \int_0^1 [w_i + u]\, du \ . \tag{6.10}$$

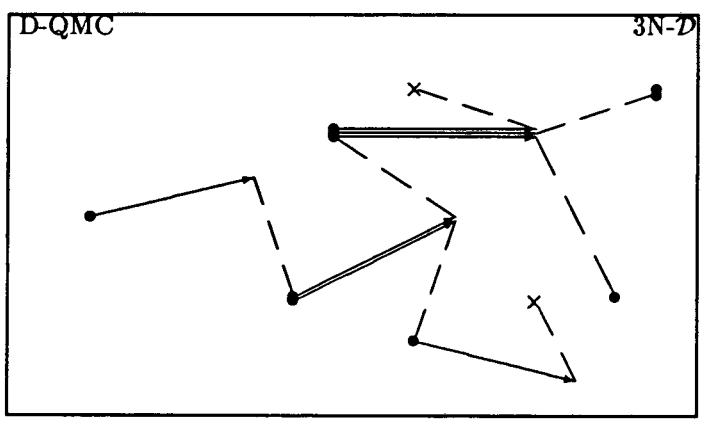
A non-integer weight factor without branching can be used, which ties into the generalized Feynman-Kac formalism developed by Michel Caffarel, *et.al.*[16, 17, 18, 19]. Even though both methods of biasing the random walk are equivalent there are advantages and disadvantages to each. The advantages of branching is that there are more walkers in the more important regions, those in the least important regions eventually die off. It is the branching that biases the distribution from ψ_T^2 to the mixed distribution $\psi_T \varphi_0{}'$. If branching is rare then longer runs are required to adequately sample from the mixed distribution. Carrying weights along overcomes the last problem, but may result in many random walkers with very small weights. These random walkers will use up computer resources, but not contribute much to the expectation value.

Note that it is possible for $w_i < 1$ such that m can be zero. The random walker is said to be "killed" and its random walk is terminated. This is no problem since an entire ensemble of random walkers is being tracked. On the other hand, if $m > 1$ then more than one copy of the random walker is made and this is referred to as the "birth" of additional walkers. If the trial energy E_T is set to approximate the correct ground state energy, then on average the number of random walkers in the ensemble remains constant, which means that the number of deaths is approximately equivalent to the number of births.

The Metropolis acceptance/rejection step 4 is not strictly necessary but generally improves the time-step bias (see section 6.3) and guarantees that the D-QMC will return the exact energy when the exact wavefunction is used. The effect of including step 4, or not, is shown in the hydrogen atom time-step bias plots (Figs. 6.3 and 6.2).

A diagrammatic representation of a random walker in $3N$ dimensional space undergoing a D-QMC process is shown in Fig. 6.1; where, as before in the V-QMC case, the solid vectors represent the drift guided by the trial wavefunction, and the dashed lines depicts diffusive steps. However, after each Metropolis accepted step the branching term is computed and either one or more copies of the random walker is kept or the random walker is killed.

Figure 6.1: A schematic representation of a diffusion QMC walk, where random walkers drift (solid arrows) according to trial wavefunction ψ_T, then diffuse (dashed lines) with some root-mean-squared distance, and at each point Metropolis accept or reject the drift-diffuse move, then either branch or weight each point according to $e^{-\alpha(\tau-\tau_0)}$ (branching is shown). The symbol × denotes a configuration that has 'died', and multiple copies are shown as multiple • points. The diagram shows 3 generations or successful steps.



6.3 Time-Step Bias

A time-step bias exists in the D-QMC algorithm[104, 107] due to the use of the short-time approximation to the imaginary time propagator. The "time-step bias" is defined as the slope of energy expectation value as a function of time step. A motivation for the time-step bias can be found from the Trotter approximation[121, 52]. The imaginary time propagator for the imaginary time Schrödinger equation is given by

$$\rho = e^{-H\tau} , \tag{6.11}$$

where the Hamiltonian H is the sum of two non-commuting operators T and V, the kinetic and potential energy operators respectively. The imaginary time propagator can be split into a product of two exponentials via the Campbell-Baker-Hausdorff formula, where the other terms

will be at least order τ^2

$$e^{-H\tau} = e^{-T\tau}e^{-V\tau} + O(\tau^2) \,. \tag{6.12}$$

Slightly different forms can result in accuracies greater than order τ^2[56, 57]. The realization of the above operator results in a process where diffusive steps are alternated with branching. Therefore, it is expected that for the short-time approximation that treats the diffusive step separately from the branching that some some sort of systematic bias is introduced. The magnitude of the time-step bias will be elaborated below.

The equilibrium distribution for the exact imaginary time propagator satisfies the imaginary time Schrödinger equation (Eq. 4.3) which we shall rewrite as

$$Lf(\vec{R},\tau) = 0 \,, \tag{6.13}$$

where L is a parabolic partial differential operator, first order in imaginary time and second order in space,

$$L = \frac{\partial}{\partial\tau} - D\nabla^2 + D\nabla\cdot\vec{F}_Q + D\vec{F}_Q\cdot\nabla + (E_L - E_T) \,. \tag{6.14}$$

The equilibrium distribution $f(\vec{R},\tau)$ satisfies the integral equation,

$$f(\vec{R},\tau) = \int \rho(\vec{R}\tau|\vec{R}_0\tau_0)f(\vec{R}_0,\tau_0)\,d\vec{R}_0 \,. \tag{6.15}$$

However, since a short-time imaginary time propagator, which we label $\rho_D(\vec{R}\tau|\vec{R}_0\tau_0)$, is used a slightly different problem has also been solved. This new equilibrium distribution f^0, which we shall call the short-time equilibrium distribution is generated from

$$f^0(\vec{R},\tau) = \int \rho_D(\vec{R}\tau|\vec{R}_0\tau_0)f^0(\vec{R}_0,\tau_0)\,d\vec{R}_0 \,, \tag{6.16}$$

which will satisfy a different differential operator L^0, where

$$L^0 f^0(\vec{R},\tau) = 0 \,. \tag{6.17}$$

This differential operator L^0 can be found by having operator L act on the short-time distribution and subtract the extra terms.

$$\begin{aligned} Lf^0(\vec{R},\tau) &= \int L\rho_D(\vec{R}\tau|\vec{R}_0\tau_0)f^0(\vec{R}_0,\tau_0)\,d\vec{R}_0 && (6.18)\\ &= \int\left[\frac{\partial\rho_D}{\partial\tau} - D\nabla^2\rho_D + D\nabla\cdot\left(\rho_D\vec{F}_Q\right) - (E_T - E_L)\rho_D\right] f^0 d\vec{R}_0 && (6.19)\\ &= V'(\vec{R})f^0(\vec{R},\tau) \,. && (6.20) \end{aligned}$$

Finally the operator L^0 is found

$$L^0 = L - V'(\vec{R}) \,, \tag{6.21}$$

such that f^0 represents a homogeneous solution. The 'potential' V' can be estimated by actually performing this set of manipulations and will be shown in the following sections to produce a time-step bias which is at least first order with respect to time step.

6.3.1 Finite Time-Step Perturbation Potential

By following the steps outlined in section 6.3 we will arrive at the form for the perturbation potential $V'(\vec{R})$, where this perturbation is obtained from the exact electronic Hamiltonian of Eq. 2.3. Therefore, this perturbed Hamiltonian represents what is actually 'solved' in a finite time-step D-QMC process, where moves aren't accepted or rejected via the Metropolis algorithm (Step 4 in the D-QMC algorithm of pages 34 and 43). The derivation is somewhat tedious and can be skipped over to the last of section 6.3.2, Eq. 6.39, which gives the final form of the perturbation potential in terms of the trial wavefunction.

The derivation relies on Taylor expansions of the terms resulting in the integral of Eq. 6.19 around $\vec{R}$ and then regrouping terms with respect to $\Delta_i = (\vec{R} - \vec{R}_0 - D\tau\vec{F})_i$ (set $\tau_0 = 0$ for convenience). The necessity of this convoluted procedure is dictated by the need to analytically evaluate the resulting Gaussian integrals. The first step is to expand in a Taylor series and regroup the short-time distribution function

$$f^0(\vec{R}_0) = f^0(\vec{R}) - \sum_i \Delta_i \frac{\partial f^0}{\partial \vec{R}_i}\bigg|_{\vec{R}} + \left[\frac{1}{2}\sum_{ij} \Delta_i\Delta_j \frac{\partial^2 f^0}{\partial \vec{R}_i \partial \vec{R}_j}\bigg|_{\vec{R}} - D\tau \sum_i \vec{F}_i \frac{\partial f^0}{\partial \vec{R}_i}\bigg|_{\vec{R}}\right] + \cdots . \quad (6.22)$$

However, the drift $\vec{F}$ itself is also a function of the end points of the path (*i.e.* $\vec{F} = \vec{F}(\vec{R}, \vec{R}_0)$), and a decision has to be made on how to proceed. We therefore expand and regroup the second argument of $\vec{F}$, the initial point of the path, around the final end point

$$\begin{aligned} \vec{F}(\vec{R}, \vec{R}_0) &= \vec{F}(\vec{R}, \vec{R}) - \sum_i \Delta_i^{(1)} \frac{\partial \vec{F}}{\partial \vec{R}_{0i}}\bigg|_{\vec{R}_0 = \vec{R}} \\ &+ \left[\frac{1}{2}\sum_{ij} \Delta_i^{(1)}\Delta_j^{(1)} \frac{\partial^2 \vec{F}}{\partial \vec{R}_{0i}\partial \vec{R}_{0j}}\bigg|_{\vec{R}_0 = \vec{R}} - D\tau \sum_i \vec{F}_i \frac{\partial \vec{F}}{\partial \vec{R}_{0i}}\bigg|_{\vec{R}_0 = \vec{R}}\right] + \cdots \end{aligned} \quad (6.23)$$

where we define $\Delta_i^{(1)} = (\vec{R} - \vec{R}_0 - D\tau\vec{F}(\vec{R}, \vec{R}))_i$. We have expanded out all dependence on the starting point $\vec{R}_0$ and have simplified the integral in Eq.6.19. The difference between $\Delta^{(1)}$ and Δ does not occur until $O(\tau^{3/2})$ terms, which is shown by

$$\Delta_i = \Delta_i^{(1)} + D\tau \sum_j \Delta_j^{(1)} \frac{\partial \vec{F}_i}{\partial \vec{R}_{0j}}\bigg|_{\vec{R}_0 = \vec{R}} + O(\tau^2) \quad (6.24)$$

$$\Delta^2 = \Delta^{(1)^2} + 2D\tau \sum_{ij} \Delta_i^{(1)}\Delta_j^{(1)} \frac{\partial \vec{F}_i}{\partial \vec{R}_{0j}}\bigg|_{\vec{R}_0 = \vec{R}} + O(\tau^3) , \quad (6.25)$$

where each *Root-Mean-Square* power of $\Delta^{(1)}$ in the Gaussian integral contributes an order $\tau^{1/2}$.

Continuing with Eq. 6.22, gradients and Laplacians for the finite time-step mixed distribution $f^0 = \psi_T \varphi_0'$ can be approximated in terms of the quantum force, where φ_0' is the ground

state eigenfunction of the perturbed Hamiltonian,

$$\frac{\partial_i f^0}{f^0} = \frac{1}{2}\vec{F}_i + \frac{\partial_i \varphi_0{}'}{\varphi_0{}'} \approx \vec{F}_i \tag{6.26}$$

$$\frac{\partial_i \partial_j f^0}{f^0} \approx \vec{F}_i \vec{F}_j + \partial_i \vec{F}_j \ , \tag{6.27}$$

where we have just assumed that ψ_T is approximately equivalent to the ground state eigenfunction $\varphi_0{}'$, and define $\partial_i = \frac{\partial}{\partial \vec{R}_{0i}}$.

By expanding and regrouping, the mixed distribution f^0 becomes

$$f^0(\vec{R}_0) \approx f^0(\vec{R}) \left\{ 1 - \sum_i \Delta_i \vec{F}_i \Big|_{\vec{R}} \right. \tag{6.28}$$

$$\left. + \left[\frac{1}{2} \sum_{ij} \Delta_i \Delta_j (\vec{F}_i \vec{F}_j + \partial_i \vec{F}_j) \Big|_{\vec{R}} - D\tau \sum_i \vec{F}_i \vec{F}_i \Big|_{\vec{R}} \right] + \cdots \right\} ,$$

where $\vec{R}$ after the vertical bar means that $\vec{R}_0$ is set to $\vec{R}$ after the partial derivative (if any) with respect to $\vec{R}_0$ is performed. Respective terms of the above equation are labeled $\hat{f}^0$, $\hat{f}^{1/2}$, and $\hat{f}^1$ for corresponding orders of τ, where $\hat{f}^0 = f^0(\vec{R}_0)$, *etc.* This notation will be used in section 6.3.2 to help delineate the various orders of τ.

At this time, to avoid being completely buried by algebra, it is convenient to fix α and $\vec{F}$ to a particular choice which will be shown later in section 6.3.2 to have no bias of order τ^0.

$$\alpha = \frac{1}{2}\left[E_L(\vec{R}) + E_L(\vec{R}_0)\right] - E_T \qquad \text{and} \qquad \vec{F}(\vec{R}, \vec{R}_0) = \vec{F}_Q(\vec{R}_0) \ . \tag{6.29}$$

This choice of local energy measure α has pointwise behavior that cancels out to $O(\tau)$ in equation 6.19. Other choices for the quantum force will on average leave terms $\propto D\nabla \cdot \vec{F}_Q$[93, 94, 104]. The vectors $\vec{F}_Q$ and $\vec{F}$ are the quantum forces evaluated at the points $\vec{R}$ and $\vec{R}_0$ respectively.

Evaluating the terms within the Eq. 6.19 brackets, yields the exact expression

$$-\alpha + E_L(\vec{R}) - E_T + \left[D\nabla - \frac{1}{2\tau}\Delta\right] \cdot (\vec{F}_Q - \vec{F}) - \Delta \cdot \nabla\alpha + D\tau\nabla^2\alpha - D\tau\vec{F}_Q \cdot \nabla\alpha - D\tau^2(\nabla\alpha)^2 \ . \tag{6.30}$$

After expanding and regrouping the above expression gives

$$\left[D\nabla \cdot \vec{F}_Q - \frac{1}{2\tau} \sum_{ij} \Delta_i^{(1)} \Delta_j^{(1)} \partial_j \vec{F}_i \Big|_{\vec{R}} \right]^{(\hat{h}^0)} \tag{6.31}$$

$$+ \left[\frac{1}{4\tau} \sum_{ijk} \Delta_i^{(1)} \Delta_j^{(1)} \Delta_k^{(1)} \partial_j \partial_k \vec{F}_i \Big|_{\vec{R}} - \frac{D}{2} \sum_{ij} \vec{F}_i \Delta_j^{(1)} \partial_i \vec{F}_j \Big|_{\vec{R}} \right]^{(\hat{h}^{1/2})}$$

$$+\left[\frac{D\tau}{2}\sum_i \vec{F}_i\partial_i E_L\Big|_{\vec{R}} - \frac{1}{4}\sum_{ij}\Delta_i^{(1)}\Delta_j^{(1)}\partial_i\partial_j E_L\Big|_{\vec{R}}\right.$$

$$-\frac{1}{12\tau}\sum_{ijkl}\Delta_i^{(1)}\Delta_j^{(1)}\Delta_k^{(1)}\Delta_l^{(1)}\partial_j\partial_k\partial_l\vec{F}_i\Big|_{\vec{R}}$$

$$\left.+\frac{D}{2}\sum_{ijk}\Delta_i^{(1)}\Delta_j^{(1)}\left(\vec{F}_k\partial_j\partial_k\vec{F}_i\Big|_{\vec{R}} - \partial_i\vec{F}_k\partial_j\vec{F}_k\Big|_{\vec{R}}\right) + \frac{D\tau}{2}\nabla^2 E_L\right]^{(\hat{h}^1)} + \cdots \quad ,$$

where the superscripts on the square brackets label the respective orders of τ with $\hat{h}^0$, $\hat{h}^{1/2}$, or $\hat{h}^1$ for the terms within the brackets.

Lastly, the D-QMC imaginary time propagator (Eq. A.10) must be expanded and grouped around $\vec{R}$. With the help of Eq. 6.23 it is found to be

$$N\exp\left(\frac{-\Delta^2}{4D\tau}\right) = N\exp\left(\frac{-\Delta^{(1)^2}}{4D\tau}\right)\left[1 - \frac{1}{2}\sum_{ij}\Delta_i^{(1)}\Delta_j^{(1)}\partial_j\vec{F}_i\Big|_{\vec{R}} + \cdots\right] , \tag{6.32}$$

where N is the normalizing constant which includes the $e^{\alpha\tau}$ factor, which adds only $O(\tau^{3/2})$ terms when Taylor expanding from $\vec{R}_0$ to $\vec{R}$. The terms of the Green's function also respectively denoted $\hat{g}^0$, and $\hat{g}^1$ respectively referring to orders of τ. Green's function or imaginary time propagator has no $\tau^{1/2}$ terms in this expansion.

6.3.2 Evaluate Time-Step Errors

Now that the ground work has been established in the previous section the respective orders of τ can be examined from suitable combinations of $\hat{f}$, $\hat{g}$, $\hat{h}$ and evaluating the Gaussian integrals. It is assumed in the following equations that $\vec{R}_0$ has been set to $\vec{R}$ after the ∂_i operator has been performed. Gaussian integrals will be denoted by $\langle\cdots\rangle$, where for some arbitrary function $\mathcal{O}(\Delta^{(1)})$

$$\left\langle\mathcal{O}(\Delta^{(1)})\right\rangle \equiv \int \mathcal{O}(\Delta^{(1)})N\exp\left(\frac{-\Delta^{(1)^2}}{4D\tau}\right)d\vec{R}_0 \tag{6.33}$$

over all $3N$ dimensional space.

There is only one zero order term given by the combination $\hat{f}^0\hat{g}^0\hat{h}^0$. The evaluation of this term with respect to Eq. 6.19 is

$$\left\langle f^0(\vec{R})\left[D\nabla\cdot\vec{F}_Q - \frac{1}{2\tau}\sum_{ij}\Delta_i^{(1)}\Delta_j^{(1)}\partial_j\vec{F}_i\right]\right\rangle = 0 \, . \tag{6.34}$$

This shows that the $O(1)$ term vanishes and there is no systematic bias introduced by the D-QMC finite time-step approximation. The integrals can easily be done since all the terms have been appropriately expanded and grouped in terms of $\Delta^{(1)}$, the integrals are now just

various moments of a Gaussian, where $\langle 1 \rangle = 1$, $\langle \Delta_i^{(1)} \Delta_j^{(1)} \rangle = 2D\tau\delta_{ij}$, $\langle \Delta_i^{(1)} \Delta_j^{(1)} \Delta_k^{(1)} \Delta_l^{(1)} \rangle = 4D^2\tau^2 [\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}]$, and all odd powers of $\Delta^{(1)}$ moments are equal to zero. This shows that, on average, all half-integer powers of τ terms cancel out! Therefore, extrapolations to zero time step need not consider polynomials in $\tau^{1/2}$[103].

The next possible contribution to the perturbation potential are the $O(\tau)$ terms $\hat{f}^1\hat{g}^0\hat{h}^0 + \hat{f}^0\hat{g}^1\hat{h}^0 + \hat{f}^0\hat{g}^0\hat{h}^1 + \hat{f}^{1/2}\hat{g}^0\hat{h}^{1/2}$. We will look at each term in the sum individually. The first term in the sum, $\hat{f}^1\hat{g}^0\hat{h}^0$, gives for the integral equation of Eq. 6.19

$$\left\langle \left[\frac{1}{2} \sum_{ij} \Delta_i^{(1)} \Delta_j^{(1)} \partial_i \partial_j f^0 - D\tau \sum_i \vec{F}_i \partial_i f^0 \right] \left[D\nabla \cdot \vec{F} - \frac{1}{2\tau} \sum_{ij} \Delta_i^{(1)} \Delta_j^{(1)} \partial_i \vec{F}_i \right] \right\rangle \tag{6.35}$$

$$= \quad -2D^2\tau \sum_{ij} \partial_i \partial_j f^0 \partial_i \vec{F}_j \quad ,$$

for the next term $\hat{f}^0\hat{g}^1\hat{h}^0$ gives

$$\left\langle f^0 \left[-\frac{1}{2} \sum_{ij} \Delta_i^{(1)} \Delta_j^{(1)} \partial_j \vec{F}_i \right] \left[D\nabla \cdot \vec{F} - \frac{1}{2\tau} \sum_{ij} \Delta_i^{(1)} \Delta_j^{(1)} \partial_j \vec{F}_i \right] \right\rangle \tag{6.36}$$

$$= \quad 2D^2\tau f^0 \sum_{ij} \left(\partial_i \vec{F}_j \right)^2 \quad ,$$

and for the third term $\hat{f}^0\hat{g}^0\hat{h}^1$ gives

$$\left\langle f^0 \left[\frac{D\tau}{2} \sum_i \vec{F}_i \partial_i E_L - \frac{1}{4} \sum_{ij} \Delta_i^{(1)} \Delta_j^{(1)} \partial_i \partial_j Eloc \right.\right.$$

$$- \frac{1}{12\tau} \sum_{ijkl} \Delta_i^{(1)} \Delta_j^{(1)} \Delta_k^{(1)} \Delta_l^{(1)} \partial_j \partial_k \partial_l \vec{F}_i$$

$$\left.\left. + \frac{D}{2} \sum_{ijk} \Delta_i^{(1)} \Delta_j^{(1)} \left(\vec{F}_k \partial_j \partial_k \vec{F}_i - \partial_i \vec{F}_k \partial_j \vec{F}_k \right) \right] \right\rangle \tag{6.37}$$

$$= D^2\tau f^0 \left\{ \frac{1}{2D} \sum_i \left(\vec{F}_i \partial_i E_L - \partial_i^2 E_L \right) - \sum_{ij} \left[\partial_i^2 \partial_j \vec{F}_j + \left(\partial_i \vec{F}_j \right)^2 - \vec{F}_i \partial_i \partial_j \vec{F}_j \right] \right\} \quad ,$$

finally for the last term $\hat{f}^{1/2}\hat{g}^0\hat{h}^{1/2}$ gives

$$\left\langle \left[-\sum_i \Delta_i^{(1)} \partial_i f^0 \right] \left[\frac{1}{4\tau} \sum_{ijk} \Delta_i^{(1)} \Delta_j^{(1)} \Delta_k^{(1)} \partial_j \partial_k \vec{F}_i - \frac{D}{2} \sum_{ij} \vec{F}_i \Delta_j^{(1)} \partial_i \vec{F}_j \right] \right\rangle \tag{6.38}$$

$$= \quad D^2\tau \sum_{ij} \left[\frac{1}{2} \partial_i f^0 \partial_i \left(\vec{F}_j^2 \right) - 3\partial_i f^0 \partial_i \partial_j \vec{F}_j \right] \quad .$$

The sum of all the above terms gives the perturbation potential

$$V'(\vec{R}) = D^2\tau \left[\frac{1}{2D}\vec{F}_Q \cdot \nabla E_L - \nabla^2 E_L - \nabla^2\nabla \cdot \vec{F}_Q \right. \tag{6.39}$$

$$\left. -\vec{F}_Q \cdot \nabla\nabla \cdot \vec{F}_Q - \frac{1}{2}\vec{F}_Q \cdot \nabla \vec{F}_Q^2 - \frac{1}{2}\nabla^2 \vec{F}_Q^2 \right] ,$$

where the approximations to ∇f^0 and $\nabla^2 f^0$ of Eqs. 6.26 and 6.27 have been used.

This perturbation potential has been shown to be linear with respect to time step, and there is no systematic error introduced for the given choice of local energy and quantum force. It will, therefore, vanish in the limit as time-step approaches zero. However, it will be shown in the next section that, for certain choices of trial wavefunctions, the perturbation potential has terms which can not be easily corrected for.

6.3.3 Time-Step Bias of the Hydrogen Atom

The hydrogen atom is the only system simple enough to compare the computational and analytic time-step bias. Using a trial wavefunction of the form

$$\psi_T(\vec{r}) = \frac{\zeta^{3/2}}{\sqrt{\pi}} e^{-\zeta r} \tag{6.40}$$

from section 5.3, it can be shown by algebraic manipulation that the time-step bias perturbation potential from Eq. 6.39 for the hydrogen atom is

$$D^2\tau \left[(-12\zeta - 4)\pi\delta(r) + 2\zeta(5\zeta - 1)\frac{1}{r^2}\right] . \tag{6.41}$$

This form is archetypical of all STO type trial wavefunctions, in that it contains a Dirac delta function at the nuclear center. Thus attempts to correct the effect of the perturbation potential by including the negative of it in the Hamiltonian will turn a K step random walk to a K^2 step random walk. Because of the delta function, the random walk at each step will require the evaluation of the wavefunction at that position. This requires an auxiliary random walk to be performed to stochastically compute a value for the wavefunction[20].

Perturbation theory predicts the first-order perturbation energy for the ground state energy level to be $-\tau\,\delta(6 + 5\delta)$, in atomic units, where $\delta = \zeta - 1$ (See Fig. 6.2). We see qualitatively good agreement with the numerical evaluation from a series of D-QMC runs, where the Metropolis acceptance/rejection step was disabled and the trial wavefunction ζ was varied. Each run consists of a series of energy expectation values at time steps ranging from .01 to .2 inverse Hartrees. For each run the best linear fit is made through each set of points and the slope determines the time-step bias.

Another set of runs was performed with the Metropolis acceptance/rejection step enabled with the same set of time steps and ζ's and the results are shown in Fig. 6.3. The time-step

bias is at least an order of magnitude less than for the previous case in Fig. 6.2. The time-step bias has also a parabolic behavior with respect to ζ and peaks with a value of zero at $\zeta \approx 1$. This corresponds to the exact ground state eigenfunction, where $\zeta = 1$.

6.3.4 Short-Time Approximation and Nodal Boundary Conditions

With the short time approximation a random walker starts at some initial point $\vec{R}_0$ and then drifts and diffuses according to the short-time approximation of the imaginary time propagator to another point $\vec{R}$. This represents one step. However, it is entirely possible that the path taken from point $\vec{R}_0$ to $\vec{R}$ may have crossed a nodal boundary (see Fig. 6.4). As required by the imposition of the fixed-node constraint, all such random walkers must be terminated. The effect of not taking this into account is to effectively enlarge the nodal volume, such that the curvature of the wavefunction is diminished. In other words, this has the effect of decreasing the average kinetic energy and, hence, the total energy.

Empirically, every STO basis set that I have tried has had a negative time-step bias, such that the energy expectation value approaches the zero time-step limit from 'below'. Any study that just uses a single time step, no matter how small, will under-estimate the extrapolated zero time-step limit and will apparently 'recover' more of the correlation energy than it should.

There are *ad hoc* ways to account for the cross-recross bias by computing the probability of node crossing and terminating random walks accordingly[3]. This requires a detailed knowledge of nodal positions such that distances from nodes can be computed. Another way is to sample from the exact imaginary time propagator by either the Domain's Green's function Monte Carlo method (DGFMC)[67, 68, 70] or the improved Green's function Monte Carlo method (GFMC)[29, 24].

Figure 6.2: The hydrogen atom D-QMC test case with differing parameterized trial wavefunctions of the form $\psi_T(\vec{r}) = \frac{1}{\sqrt{\pi}}\exp(-\zeta r)$. The solid line is the best quadratic fit with respect to the time-step bias vs. ζ data and is $-.582(3)\delta - .846(29)\delta^2$ in atomic units, where $\delta = \zeta - 1$.

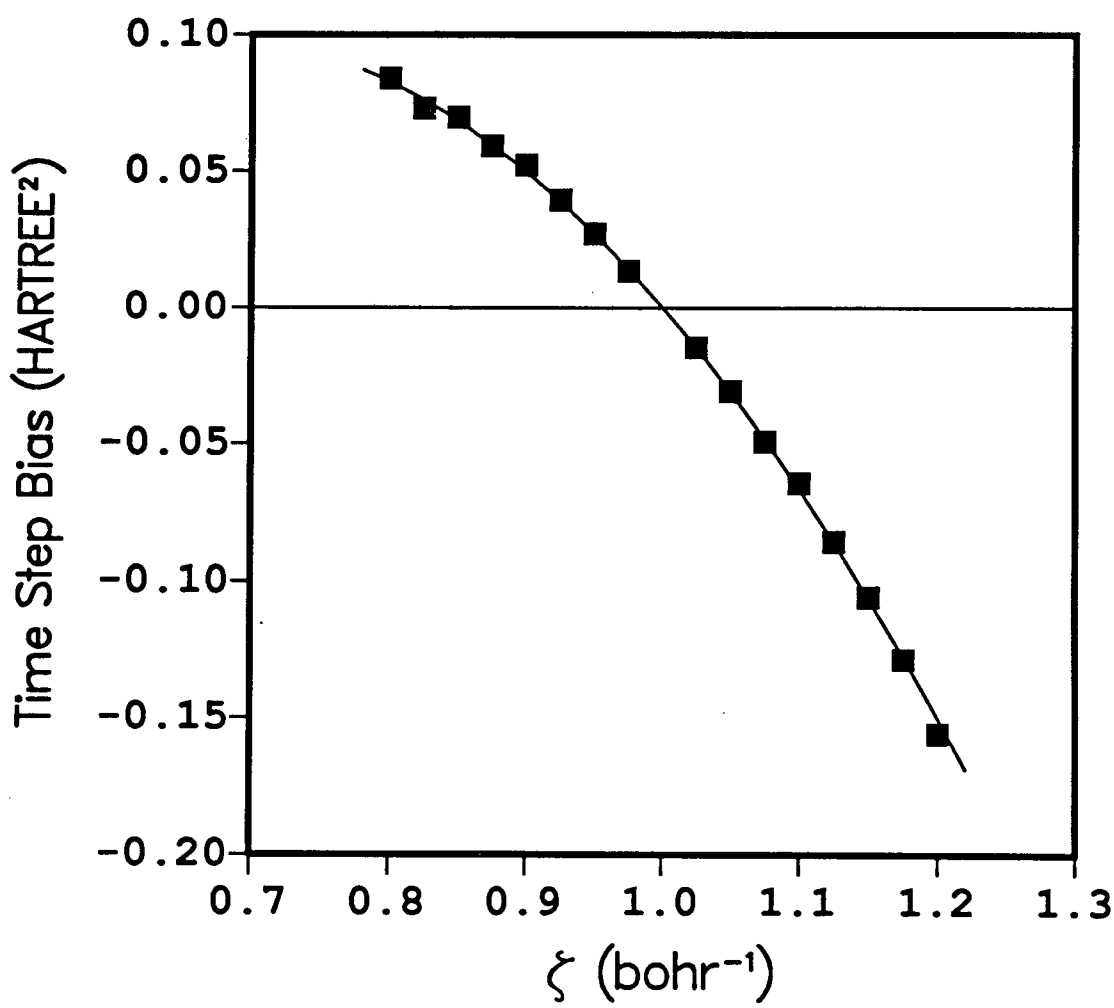


Figure 6.3: Another hydrogen atom D-QMC test case with Metropolis acceptance/rejection and differing parameterized trial wavefunctions of the form $\psi_T(\vec{r}) = \frac{1}{\sqrt{\pi}}\exp(-\zeta r)$. The solid line is the best quadratic fit with respect to the time-step bias vs. ζ data and is $-.0004(2) - .0132(36)\delta - .1810(323)\delta^2$ in atomic units, where $\delta = \zeta - 1$.

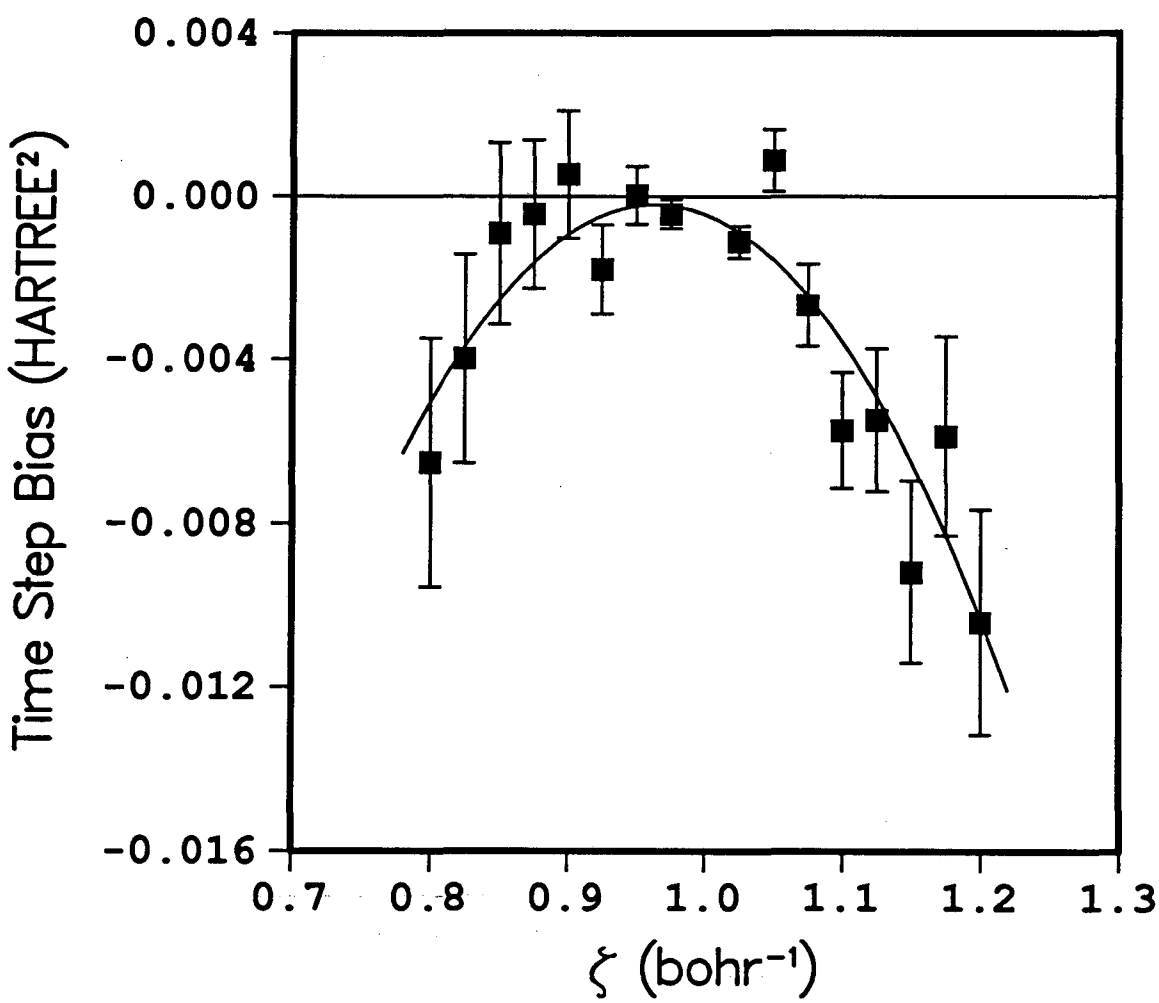
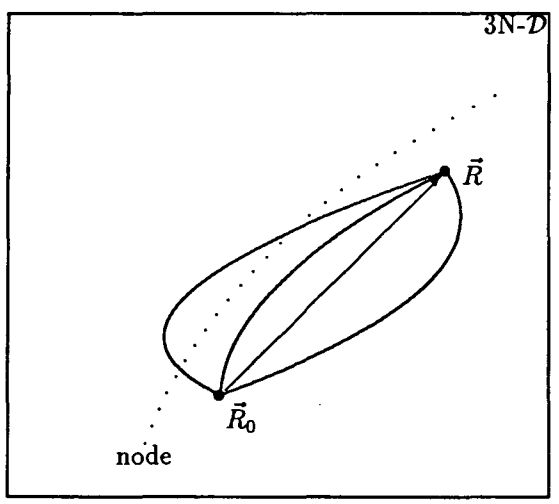


Figure 6.4: A schematic depiction of a random walker that may have, in a step, crossed and recrossed a node in violation the fixed-node boundary conditions. A step is shown in $3N$ dimensional space with a vector from $\vec{R}_0$ to $\vec{R}$. Only three of the many possible paths from $\vec{R}_0$ to $\vec{R}$ are shown, where one of the paths crosses and recrosses the nodal boundary.



Part II

Alkali Metal Clusters

Chapter 7

Cluster "Magic" Numbers

Research is the process of going up blind alleys to see if they are blind
Marston Bates

In recent years there has been a profusion of theoretical and experimental work regarding atomic clusters. Many articles[21, 78, 97] and even whole journal volumes[113, 10] have been devoted to this topic. One of the most interesting properties that has been found is the appearance of "magic" numbers in the experimental mass spectra. This is to say that clusters with certain numbers of atoms or molecules have enhanced stability and will preferentially nucleate in supersonic molecular beam experiments[114]. This enhanced stability can be attributed to some degree to closed shell structures. For rare gas clusters the magic numbers correspond to geometrical packing of hard spheres into icosahedral units. For ionic-bonded clusters the more stable configurations represent some compact subset of the bulk crystalline structure[109]. For alkali metals, their magic numbers closely follow the shell filling of valence electrons in a bag model potential reflecting the delocalized nature of the valence electrons.

7.1 Geometrical Packing

Rare gas atoms are characterized as having filled electronic shells and as such are chemically inert and except for rare instances do not bond chemically. Typically the interaction between two rare gas atoms can be adequately described by a Van der Waals interaction that results from the electron-electron correlation effect. A simple and often used potential is the Lennard-Jones 6-12 potential which models the long range inter-atomic attraction, due to the *London* or *dispersion force*, by providing the appropriate attractive interaction that for large distances scales as $-1/R^6$, where R is the inter-atomic separation. There is also a strongly repulsive core, $1/R^{12}$, to prevent the two rare gas atoms from interpenetrating each others' core. The rare gas atoms, as

a result, act like solid spheres and try to maximize the number of nearest neighbors to increase the binding energy. Packing the rare gas atoms into icosahedral units provides an optimal hard sphere packing which yields "magic" numbers that closely match with experiment. These could be the Mackay icosahedra, which have a five-fold axis of rotation at each vertex. The number of atoms for the first few Mackay icosahedra are $n = 1, 13, 55, 147, \ldots$. Substructure in the rare gas magic numbers occurs when 'caps' of atoms[49] are attached to stable structures. An example of this is given by Fig. 7.1. The entire sequence of magic numbers for the first few rare gas clusters are $n = 7, 13, 19, 23, 26, \ldots$, which matches with numerical thermodynamic calculations[50] and with experiment[45, 49].

7.2 Shell Model

Some of the earliest experimental evidence of alkali metal shell structure comes from the mass spectra of supersonic cluster beams[75, 76]. Major features in the mass spectra exist at numbers $N = 2, 8, 20, 40, \ldots$ corresponding to predicted shell closings in spherical jellium models[9, 33, 34]. The axially symmetric ellipsoidal Nilsson-Cleminger model provides some of the fine structure in the mass spectra by using a total-angular-momentum modified axially symmetric harmonic oscillator potential for the valence electrons[36, 37]. The assumption is that the alkali cluster properties are dictated primarily by the valence electrons and the effects of the cores can be smeared uniformly over a cluster-wide potential.

A simpler model that can be used for small alkali metal clusters is the anisotropic harmonic oscillator potential for the valence electrons[112]. The one-electron Hamiltonian that each valence electron experiences in the alkali metal clusters is of the form:

$$H_i = \frac{\vec{p}_i^{\,2}}{2m} + \frac{1}{2m}\left[\omega_x^2 x_i^2 + \omega_y^2 y_i^2 + \omega_z^2 z_i^2\right] , \tag{7.1}$$

where the index i refers to the i-th valence electron. The total Hamiltonian for the system is given by

$$H = \sum_i H_i , \tag{7.2}$$

where the valence electrons are weakly interacting and the orbitals are given by one-particle eigenfunctions of the separable Hamiltonian H_i. Using the harmonic oscillator number basis the orbital energies are given by

$$E(\vec{n}, \vec{\omega}) = \hbar\omega_x \left(n_x + \frac{1}{2}\right) + \hbar\omega_y \left(n_y + \frac{1}{2}\right) + \hbar\omega_z \left(n_z + \frac{1}{2}\right) , \tag{7.3}$$

where $\vec{n} = (n_x, n_y, n_z)$ are the orbital quantum numbers.

For a given positive energy the shape of the potential is an ellipsoid, which is parameterized by $\vec{\omega} = (\omega_x, \omega_y, \omega_z)$. Assuming that the cluster has constant volume with respect to

distortion, then the product $\omega_x\omega_y\omega_z = \omega_0^3$ is a constant. If we define the deformation parameter δ_x such that $\omega_x = \omega_0(1+\delta_x)$, and similarly for δ_y and δ_z, then the constant volume condition requires

$$(1+\delta_x)(1+\delta_y)(1+\delta_z) = 1\,. \tag{7.4}$$

Assume that there are N weakly interacting fermions of spin $\frac{1}{2}$, then the occupation for any orbital is $c_{\vec{N}} = 0, 1$, or 2. The sum of the occupied orbital energies gives the total energy of the system

$$E_{tot} = \sum_{\vec{n}} c_{\vec{n}}\, E(\vec{n}, \vec{\delta})\,, \tag{7.5}$$

with the constraint

$$N = \sum_{\vec{n}} c_{\vec{n}}\,. \tag{7.6}$$

The total energy can be found by minimizing equation 7.5 subject to the constraints of equations 7.6 and 7.4 by the method of Lagrange multipliers The results are reported in Table 7.1. An immediate prediction of this model is that $n = 8$, should be a magic number due to the 'p'

Table 7.1: Anisotropic harmonic oscillator for N weakly interacting spin $\frac{1}{2}$ particles, where the orbital occupations are given by $c_{\vec{n}}(n_x n_y n_z)$ and the total energy is given in units of $\hbar\omega_0$. The deformations $\delta_x, \ldots$ are related to the anisotropic frequencies by $\omega_x = \omega_0(1+\delta_x), \ldots$. Complete shells are denoted by [n], where n is the number of weakly interacting particles that fill the shell, and only used for $N > n$ to reduce table space.

N	Occupation	Deformation			Energy
		δ_x	δ_y	δ_z	($\hbar\omega_0$)
1	1(000)	0.0000	0.0000	0.0000	1.5000
2	2(000)	0.0000	0.0000	0.0000	3.0000
3	[2] 1(100)	-0.2886	0.1856	0.1856	5.3353
4	[2] 2(100)	-0.3700	0.2599	0.2599	7.5595
5	[2] 2(100) 1(010)	-0.2440	-0.0280	0.3608[a]	10.2061
6	[2] 2(100) 2(010)	-0.1566	-0.1566	0.4057	12.6515
7	[2] 2(100) 2(010) 1(001)	-0.0647	-0.0647	0.1431	15.4324
8	[2] 2(100) 2(010) 2(001)	0.0000	0.0000	0.0000	18.0000
9	[8] 1(200)	-0.1638	0.0935	0.0935	21.3241
10	[8] 2(200)	-0.2602	0.1626	0.1626	24.4147

[a]For $N = 5$ has the only significantly different deformation from the Nilsson-Clemenger model for small N, which predicts $\delta_x = \delta_y = -.140$ and $\delta_z = .352$.

shell closure, or filling of the $n_x, n_y, n_z = 1$ orbitals, which have similar nodal characteristics as

the atomic p orbitals. The characteristic of a complete shell is that the ellipsoid has spherical symmetry as observed for the $n = 2, 8$ cases, where all the deformations $\delta_x = \delta_y = \delta_z = 0$. Therefore, we would expect the $n = 8$ alkali metal cluster to to have a high degree of symmetry. J.L. Martins, in his local spin density functional theory calculations for sodium[83, 82, 84], obtained an optimized D_{2d} geometry for Na_8 similar to that shown in Fig. 7.2. Using this same geometry (suitably scaled for lithium) I performed a minimum basis set SCF geometry optimization that ended in a T_d symmetry $^1A_1$ state. Such optimized geometries for $n = 8$ has been published by Koutecký and workers[78, 14, 13]. Referring to Fig. 7.3, one has a core of four inner lithiums forming a tetrahedron, the four outer lithiums sit interstitially with respect to the four faces of the tetrahedron and out from the faces. A higher degree of symmetry could be obtained if the lithium atoms form into a perfect cube with O_h symmetry; however, if a model pair potential[95] is postulated then the T_D symmetry geometry is preferable over that of the O_h symmetry. This can be understood from simple nearest neighbor bond counting. The O_h cube has 12 edges, whereas, the final geometry shown in Fig. 7.3 has 18 edges; therefore, the T_d symmetry geometry will produce a larger binding energy, hence lower total energy if the pair potential is attractive.

Figure 7.1: Close icosahedral packing of spheres. The bottom group correspond to the $n = 13$ Mackay icosahedron. The top group represents a 'cap' of 6 atoms that can be added to the bottom group to obtain a dense $n = 19$ packing.

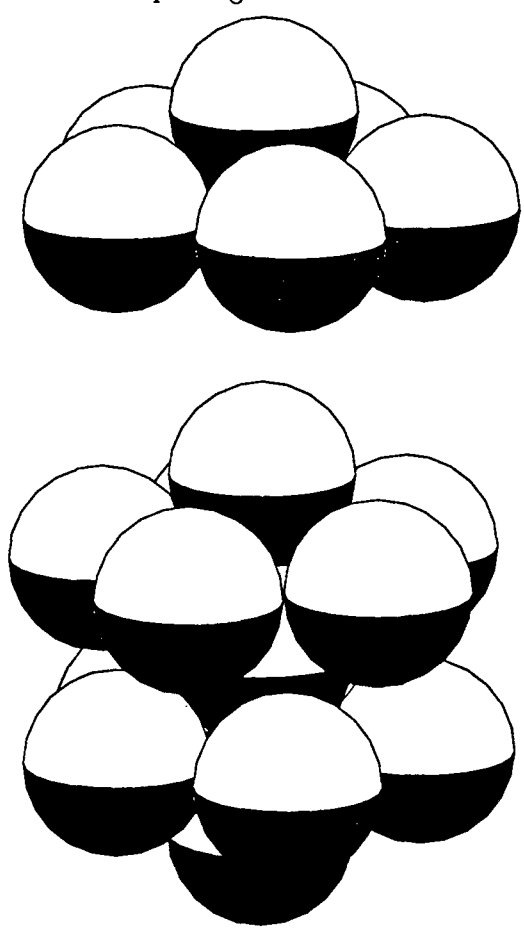


Figure 7.2: The initial geometry for Li_8 is shown for a minimum basis set SCF geometry optimization. The figure is a scaled version of the LSDFT optimized geometry of J.L. Martins, *et.al.* with D_{2d} symmetry.

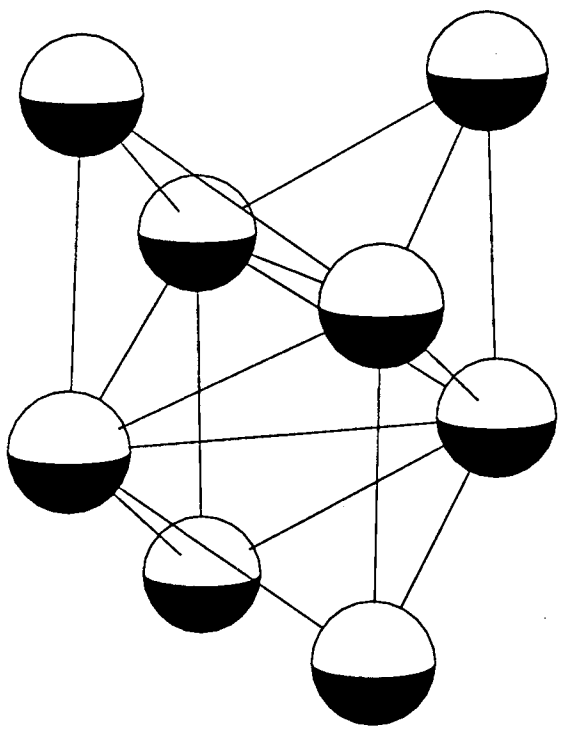
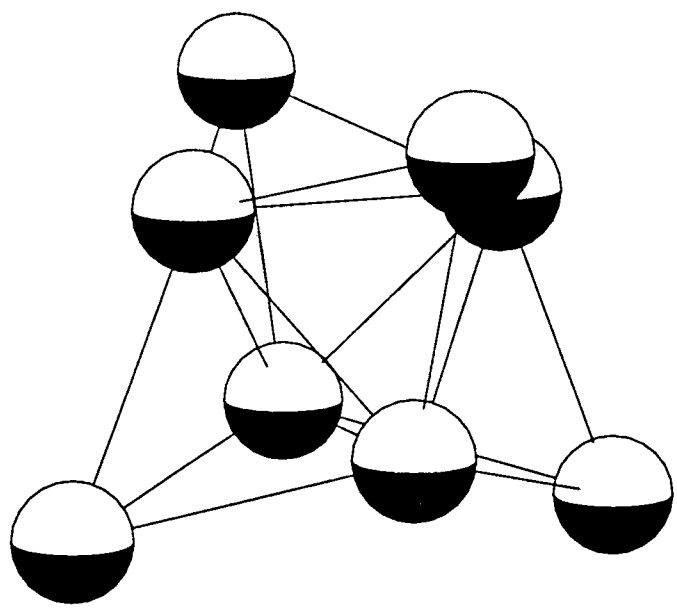


Figure 7.3: The final geometries for Li_8 are shown for a minimum basis set SCF geometry optimization, which has T_d symmetry.



Chapter 8

D-QMC Lithium Cluster Results

The wise man turns chance into good fortune
Thomas Fuller, M.D.

This chapter details the necessary steps to obtain V-QMC and D-QMC determinations of the lithium cluster total energy. First, optimum geometries must be found for each cluster. Secondly, simple and effective trial wavefunctions are found to guide the QMC walks. For the D-QMC method, the zero time-step total energy expectation value is extrapolated from several total energy expectation values performed with differing time steps. For some of the smaller clusters the total energies are compared to the best estimated non-relativistic total energies. The difference between the neutral and cation D-QMC total energies are compared to experimental ionization potentials. The difference in energy of the clusters and an equivalent number of isolated atoms gives the binding energy, which is also compared to experiment.

A section is devoted to the lithium cluster one-electron density. The general overall shape of the one-electron density bears remarkable resemblance to the anisotropic harmonic oscillator shape. Several one-electron density plots are shown. These plots show the existence of charge concentrations which are not located on a nuclear center.

Some computational aspects of QMC are discussed. First, the complexity of the QMC method is discussed. Total computer time for each lithium cluster is given and the empirical computational time scaling is shown. The scaling is compared to other *ab initio* methods. secondly, the time-step bias for the lithium clusters are given. Finally, the results are summarized and some future work is proposed.

8.1 Geometry Optimization

The lithium cluster geometries were found by using a CASSCF geometry optimization procedure, using a double zeta STO basis set from Clementi & Roetti[38] approximated by a contraction of 6 GTOs[62]. The exponential values are $\zeta = 4.61679, 2.46167$ for the $1s$ and $\zeta = 1.96299, .067198$ for the $2s, 2p$ shell. The level of CI excitation was equivalent to the number of electrons to allow core-valence orbital interactions which are important in the case of lithium. The final geometries are shown in Fig. 8.1.

Others have used various levels of theory to optimize geometries. The two that I will key on are from B.K. Rao, *et.al.*[99] where Rao uses a STO-6G minimum basis set at a DCI level of theory, and Boustani, *et.al.*[14] who use an augmented minimum basis set for a SCF geometry optimization followed by a MRD-CI to uniformly scale the geometry and obtain a MRD-CI energy minimum.

Table 8.1 shows a comparison of differing Li_3 'obtuse' geometries in the literature, where 'obtuse', in this case, is understood to mean that the apex angle of the isosceles triangular geometry is greater than 60°. We observe pretty good agreement between the differing methods, except for Boustani's which is largely influenced by the minimum basis set SCF geometry optimization. The MRD-CI scaling of Boustani only decreases the inter-nuclear distances by 2%, although it, strangely enough, produces remarkably good ionization potentials.

Table 8.1: Comparison of lithium trimer geometries between differing methods for the 'obtuse' geometry (apex angle > 60°). Angles are in degrees and lengths are in units of a_0. The length b refers to the base of the isosceles triangle and a is the length of the side.

Method	angle	a	b
Boustani, *et.al.* SCF/MRD-CI[14]	87.0	5.46	7.50
Rao, *et.al.* D-CI[99]	75.0	5.07	6.17
Rao, *et.al.* UHF[100]	75.8	5.41	6.65
Rao, *et.al.* SD-CI[101]	75.2	5.07	6.18
A.K. Ray, *et.al.* SD-CI[102]	72.2	5.33	6.28
J.L. Martins, *et.al.* PP-LSD[85, 119]	71.6	5.26	6.15
M.H. McAdon, *et.al.* GVB[87]	68.1	5.42	6.07
Gerber, *et.al.* CEPA[53, 119]	70.1	5.25	6.03
R.K. Owen, CASSCF	71.2	5.31	6.18

8.2 QMC Results

After geometry optimization, a double-zeta restricted Hartree-Fock (RHF) wavefunction was used for the antisymmetric part of the QMC trial wavefunction. The electron-nuclear correlation cusp condition was satisfied on average with the correlation range set to $0.04a_0$, where the Padé-Jastrow functional form is used for the electron-nuclear correlation function. The density dependent "*exp exp*" form of Tables 4.3 and 4.4 is used for electron-electron correlation. All correlation function parameters were found from V-QMC optimization.

The V-QMC calculations used a time step $\tau = 0.1$ Hartree^{-1} where each block was 4.0 Hartree^{-1} long. Each result was a statistical sum of 6 runs, where each run was 32 blocks long and was initialized with a pre-randomizing run of 8 blocks to reduce statistical correlation between blocks. Pre-randomizing runs are exactly the same as regular runs, but expectation values are not kept. The expectation values $\langle -(\psi_T^{-1}\nabla\psi_T)^2\rangle_{\psi_T^2}$ and $\langle \psi_T^{-1}\nabla^2\psi_T\rangle_{\psi_T^2}$ agreed within statistical uncertainty; therefore, the random walker distribution was sufficiently equilibrated for accurate expectation values with respect to the ψ_T^2 distribution. The statistical addition of each run's variance-weighted energy results in the mean energy

$$\bar{E} = \frac{\sum_i \frac{E_i}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2}}, \tag{8.1}$$

where E_i is the energy for the i-th run and σ_i^2 is its variance. The uncertainty of the mean energy σ_E is given by

$$\sigma_E^2 = \frac{1}{\sum_i \frac{1}{\sigma_i^2}}. \tag{8.2}$$

The D-QMC calculations used time steps $\tau = 0.05, 0.025, 0.010$ with block lengths of 2.0 Hartree^{-1}, except for the lithium dimer which used additional time steps of $\tau = 0.005, 0.015, 0.020$. The initial configuration of random walkers for the D-QMC calculation used the final equilibrium configurations from the V-QMC calculation. At each time step a set of three runs with lengths of 20 blocks were done and then statistically added together as described above for the V-QMC calculation.

Prior to each set of three runs, a pre-randomizing run of 4 blocks was performed to let the distribution reach equilibrium at the given time step. The D-QMC zero time-step results were extrapolated from the least χ^2 linear fit to the data points. The QMC results, in addition to the RHF and CASSCF total energies, are shown in Table 8.2.

8.2.1 Best Estimated Non-Relativistic Total Energies

The best estimated non-relativistic total energies shown in the footnotes of Table 8.2 have been determined from the following experimental and theoretical numbers. The neutral

and cation atomic energies were obtained from the lithium first, second, and third ionization potentials[92, 48], where the relativistic correction, Lamb shift, and center of mass have been accounted for[60, 122]. The estimated atomic total energy is -7.4781 Hartrees and the cation energy is -7.2800 Hartrees. The atomic total energy, obtained in this way, agrees almost exactly with other extensive theoretical calculations[71]. We note the excellent agreement of the D-QMC results, which match within statistical uncertainty. Similarly, the lithium dimer non-relativistic total energy, neglecting zero-point vibrational energy[64], is estimated to be -14.9941 Hartrees, which agrees with Ceperley's released-node GFMC total energies[29]. The D-QMC energy falls shy of the exact energy by 0.003 Hartree, but is sufficiently good to capture 98% of the correlation energy.

Figure 8.1: Lithium cluster geometries are obtained from a CASSCF geometry optimization with a "double zeta" basis set. The geometries are ordered left to right, top to bottom and correspond to the entries in Table 8.2, except the atomic case which is not shown. The 3 digit numbers are the internuclear distances in angstroms, and the 2 digit numbers are some angles of interest in degrees. All are planar except for the second Li_4.

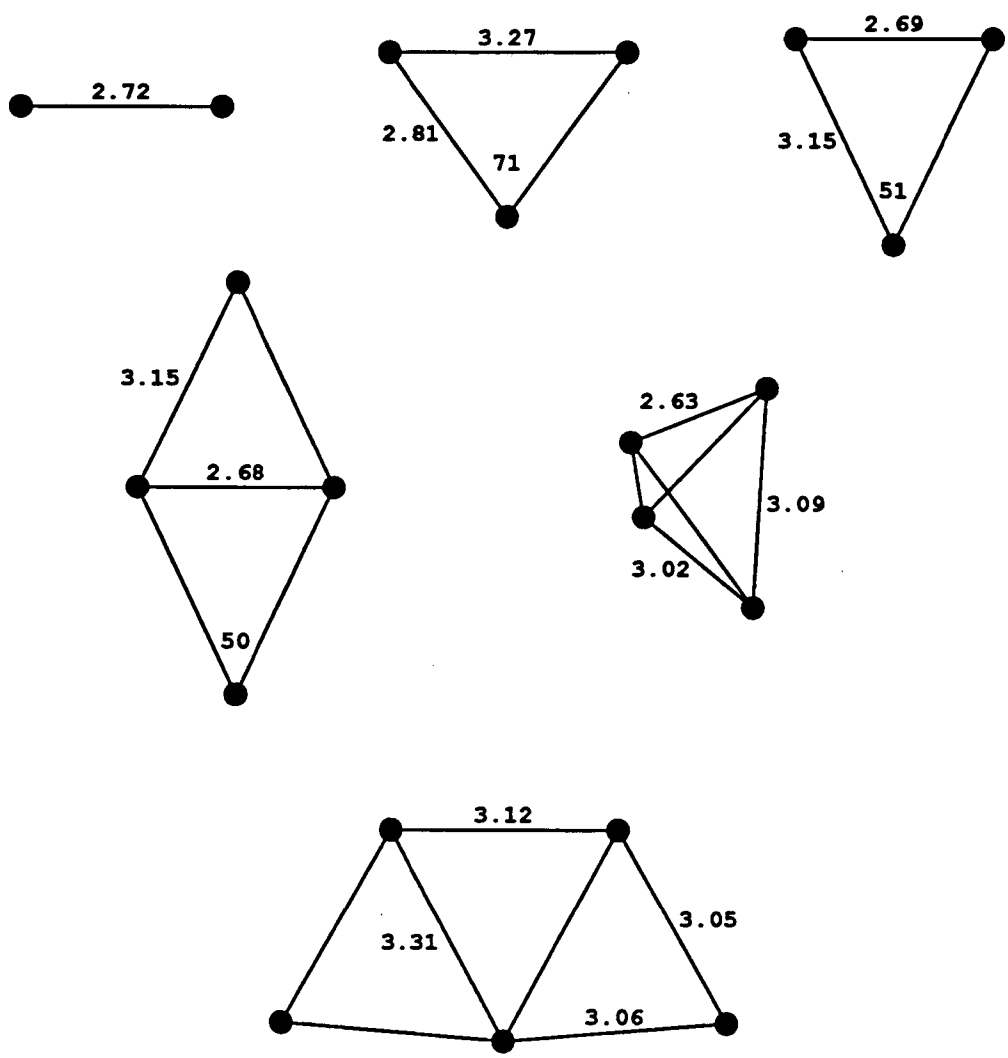


Table 8.2: Neutral and cationic total energies in Hartrees are tabulated for different *ab initio* methods. All geometries used were optimized at the CASSCF level. (see Fig. 8.1).

Neutral Lithium Clusters						
N	Symmetry	State	RHF	CASSCF	V-QMC[a]	D-QMC
1		^2S	-7.4307	-7.4439	[b] -7.4655(8)	-7.4779(10)
2	$D_{\infty h}$	$^1\Sigma^+{}_g$	-14.8675	-14.8919	[c] -14.9409(11)	-14.9912(6)
3	'obtuse' C_{2v}	$^2B_2$	-22.2997	-22.3246	[d] -22.4043(15)	-22.4900(23)
3	'acute' C_{2v}	$^2A_1$	-22.3016	-22.3269	[e] -22.4075(14)	-22.4898(23)
4	D_{2h}	^1A_g	-29.7524	-29.7676	[f] -29.8900(16)	-30.0101(27)
4	C_{2v}	$^3B_2$	-29.7438	-29.7640	[g] -29.8794(14)	-29.9968(26)
5	C_{2v}	$^2A_1$	-37.1988	-37.2324	[h] -37.3718(18)	-37.5160(31)

Cationic Lithium Clusters						
N	Symmetry	State	RHF	CASSCF	V-QMC	D-QMC
1		^1S	-7.2344	-7.2486	[i] -7.2666(5)	-7.2797(10)
2	$D_{\infty h}$	$^2\Sigma^+{}_g$	-14.6996	-14.7364	[j] -14.7627(12)	-14.8017(6)
3	'obtuse' C_{2v}	$^1A_1$	-22.1613	-22.1829	[k] -22.2633(11)	-22.3370(25)
3	'acute' C_{2v}	$^1A_1$	-22.1616	-22.1831	[l] -22.2673(13)	-22.3391(19)
4	D_{2h}	$^2B_{2u}$	-29.6064	-29.6248	[m] -29.7393(16)	-29.8437(25)
4	C_{2v}	$^2A_1$	-29.5975	-29.6134	[n] -29.7249(19)	-29.8319(27)
5	C_{2v}	$^1A_1$	-37.0451	-37.0779	[o] -37.2151(16)	-37.3547(30)

[a] All QMC runs satisfied the electron-nuclear cusp condition on average with a Padé-Jastrow correlation function that had a range of $0.04a_0$ (see Sect. 4.2.1). All QMC runs used the density dependent '*exp exp*' electron-electron correlation function with $\gamma = 0$ (see Tables 4.3 and 4.4).

[b] $\beta = .260$, best estimated total energy = -7.4781 (see section 8.2.1)
[c] $\beta = .240$, best estimated total energy = -14.9941 (see section 8.2.1)
[d] $\beta = .230$
[e] $\beta = .230$
[f] $\beta = .220$
[g] $\beta = .200$
[h] $\beta = .220$
[i] $\beta = .200$, best estimated total energy = -7.2800 (see section 8.2.1)
[j] $\beta = .260$
[k] $\beta = .210$
[l] $\beta = .230$
[m] $\beta = .230$
[n] $\beta = .180$
[o] $\beta = .220$

8.2.2 Ionization Potential

The vertical ionization potential (IP) is defined as the difference in energy between the system's neutral ground state at the optimal geometry and the cation at the same geometry, and is defined as

$$IP = E_n^+ - E_n \ , \tag{8.3}$$

where E_n and E_n^+ are respectively the total energy of the neutral and cationic cluster with n lithium atoms. An adiabatic ionization potential takes the difference of total energies with respect to a geometry optimized cationic ground state and a geometry optimized neutral ground state. In both definitions of the ionization potential the zero point vibrational energies are assumed to be comparable[96]. In this study we compute the vertical ionization potential and compare it to experimental values in the literature. The D-QMC results are given in Table 8.3 and Fig. 8.2 shows the comparison of D-QMC with experiment, Rao's D-CI[99], and Boustani's SCF/MRD-CI[14, 15] (denoted as MRD-CI+ on the plot), and the work function of the classical conducting drop[126, 91, 21]. The classical conducting drop work function is the amount of energy required to extract a single charge from a conducting sphere, and is given by the expression

$$W_n = W_\infty + (W_1 - W_\infty)n^{-1/3} \ , \tag{8.4}$$

where $W_1 = 5.39eV$, $W_\infty = 2.32eV$, and n is the number of atoms in the lithium cluster. Note that there is perfect agreement for the lithium atom by definition. The W_∞ term is the work function value when the cluster has infinite radius, or in other words, when the charge is escaping from a plane surface. The classical conducting drop model describes the size effects on ionization due to the finite size of clusters. The results of Boustani and the D-QMC results agree very closely with experiment and the Rao results consistently underestimate the experimental ionization potential.

The ionization potentials also exhibit an odd-even effect. Comparing to the neighboring clusters and ignoring finite size effects, the even-numbered clusters for lithium require more energy to ionize. The odd-numbered clusters, which are in a doublet spin state, will have a single unpaired electron in an orbital and will be relatively easy to ionize. This can be seen by considering just the $2s$ valence orbital. If two lithiums are brought together the two $2s$ atomic orbitals split their degeneracy and form a bonding and anti-bonding pair of molecular orbitals. However, for three they split to form three molecular orbitals with the middle one comparable in energy to the $2s$ atomic orbital energy. It should, therefore, take nearly the same energy to ionize the odd-numbered clusters as it is to ionize the atom. The even-numbered clusters will have a net lower energy and, therefore, require greater energy to ionize. However, as the cluster size increases the difference should decrease as the orbital spacing decreases (See Fig. 8.3).

Table 8.3: Lithium cluster Ionization Potentials comparing experimental results and D-QMC vertical Ionization Potentials tabulated from Table 8.2. Ionization Potentials are in Hartrees.

Ionization Potentials		
n	Exp	D-QMC
1	[a] 0.1981(0)	0.1982(14)
2	[b] 0.1874(4)	0.1895(9)
3	[c] 0.1599(73)	0.1530(34)
4	[d] 0.1724(110)	0.1664(37)
5	[e] 0.1675(110)	0.1613(43)

[a]See Ref. [71]
[b]See Ref. [111]
[c]See Ref. [127]
[d]See Ref. [128]
[e]See Ref. [129]

8.2.3 Binding Energy

The binding energy (BE), also called the atomization energy, is the amount of energy required to infinitely separate the atoms and is defined as

$$BE = nE_1 - E_n \; ; \tag{8.5}$$

where E_n, as before, is the total energy of the n-th cluster. Generally the binding energy is a measure of the cluster stability and in the case of lithium is an increasing function. It is useful to define the binding energy per atom BE/n, which is a measure of the 'relative' stability. It might be energetically favorable for a cluster to break up into smaller clusters if the binding energy per atom decreases for larger clusters. The theoretical and experimental results (see Fig. 8.4) for small lithium clusters $n = 2, 3, 4, 5$ shows that BE/n is increasing, indicating that Li_5 is more stable with respect to the smaller clusters.

The binding energy can also be considered a measure of when the cluster is large enough such that it exhibits properties of the bulk. For solids the atomization energy is equivalently called the "cohesive energy", the energy required per atom to form separated neutral atoms in their electronic ground state from the solid at 0 K and 1 atm. For metallic lithium, the cohesive energy is 0.060 Hartrees[73]. This is approximately twice the binding energy of the lithium pentamer. Therefore, Li_5 may be more stable than the smaller clusters, but the larger clusters, and particularly the bulk form, represent more stable systems.

The D-QMC binding energy per atom results are approximately 19% below the experimental values. Better trial wavefunctions could reduce this discrepancy by a few percent. Even

if the best estimated non-relativistic total energies are used, the lithium dimer binding energy per atom would be .0189 Hartree, or 93% of experiment, in contrast to the D-QMC results which is 91% of the experimental results.

Table 8.4: Lithium cluster atomization energies or binding energies, BE, and binding energy per atom, BE/n, tabulated from Table 8.2. Binding Energies are in Hartrees.

Binding Energies				
	Experimental		D-QMC	
n	BE	BE / n	BE	BE / n
2	[a] 0.0406(24)	0.0203(12)	0.0355(15)	0.0177(8)
3	[b] 0.0661(64)	0.0220(21)	0.0564(29)	0.0188(10)
4	[c] 0.1240(48)	0.0310(12)	0.0987(33)	0.0247(8)
5	[d] 0.1753(40)	0.0351(8)	0.1267(38)	0.0253(8)

[a]See Ref. [127]
[b]See Ref. [127]
[c]See Ref. [128]
[d]See Ref. [129]

Figure 8.2: Lithium cluster vertical ionization potentials from D-QMC are compared to D-CI[99], SCF/MRD-CI[14] (noted as MRD-CI+ on plot), the classical conducting drop[126, 91, 21] and experiment (see Table 8.3 for references).

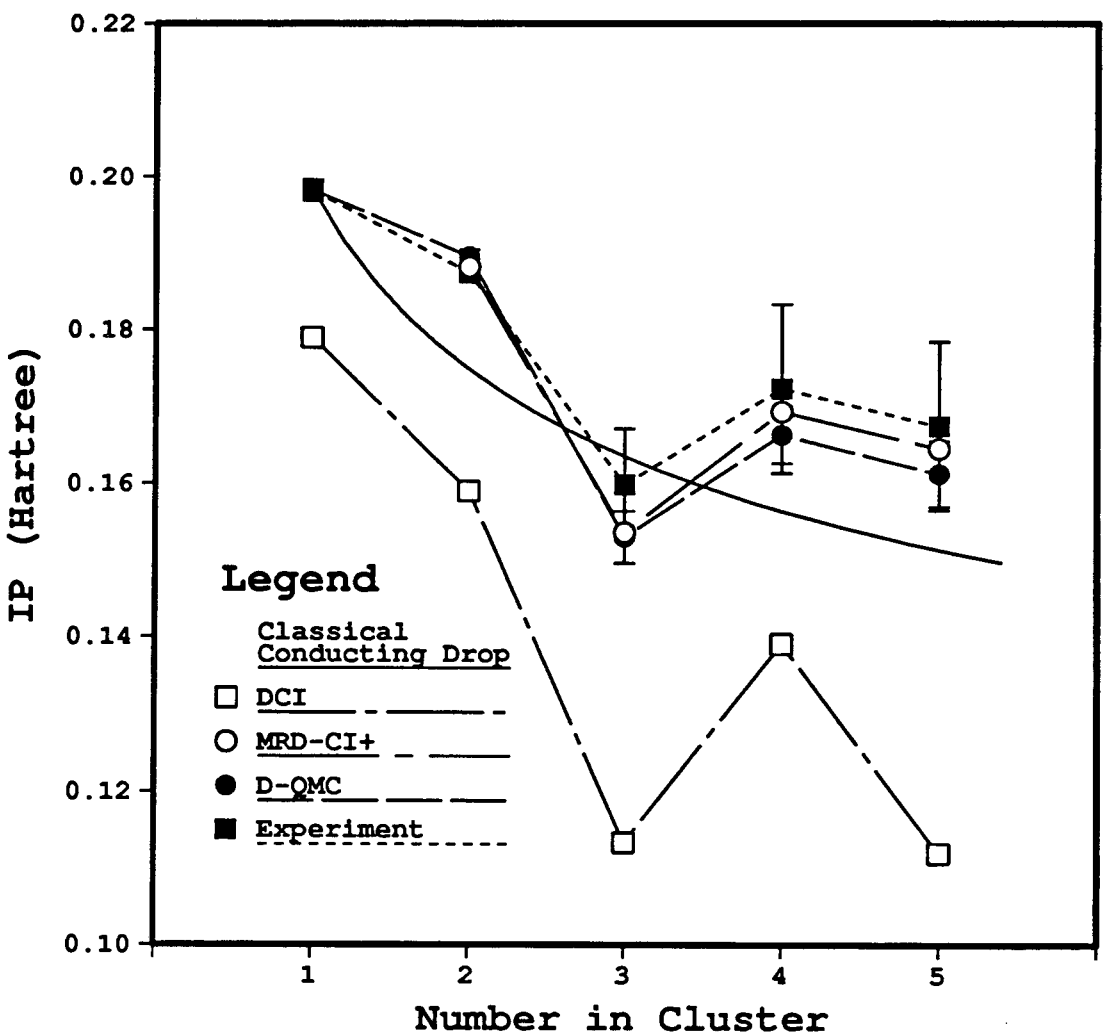


Figure 8.3: An odd-even effect is shown with respect to bonding of a half filled $2s$ orbital, where N is the number of atoms and the arrows represent electrons with up and down spins. In the case of 2 atoms the degeneracy of the $2s$ orbitals split to form the bonding and the anti-bonding orbitals, where the system will find it energetically favorable to fill the bonding orbital. The same happens for the larger clusters, where the width of the 'band' is governed by the overlap of the $2s$ atomic orbitals. In the infinite limit the translational periodicity of the system determines the exact shape of the band and for such systems, as shown, would become metallic with a half filled conduction band where the zero temperature energy surface is given by the Fermi energy ϵ_F.

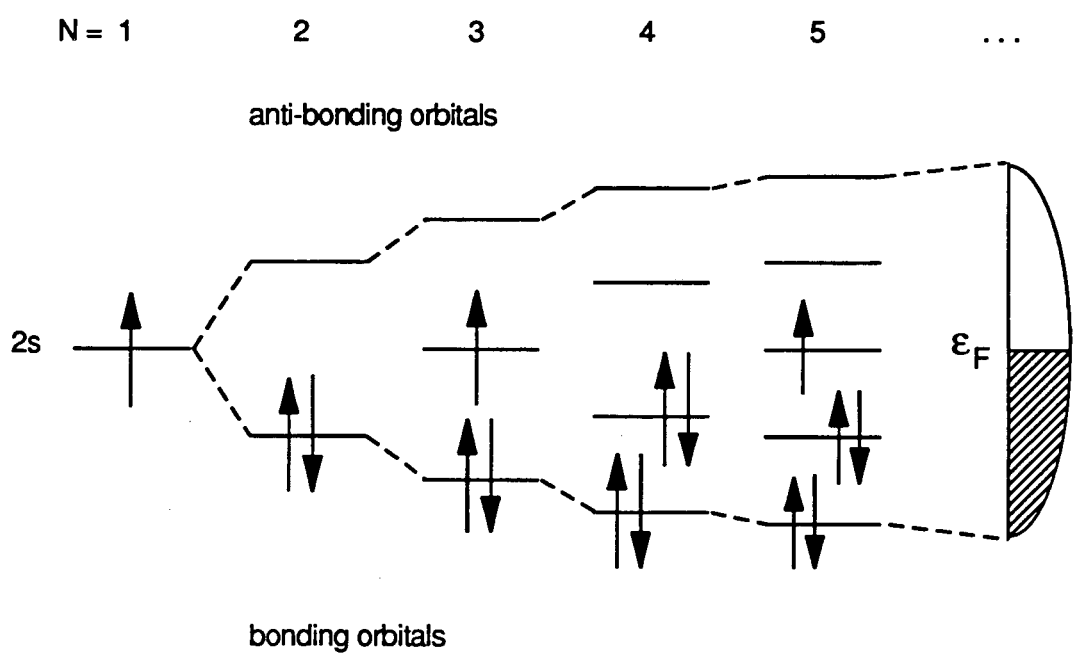
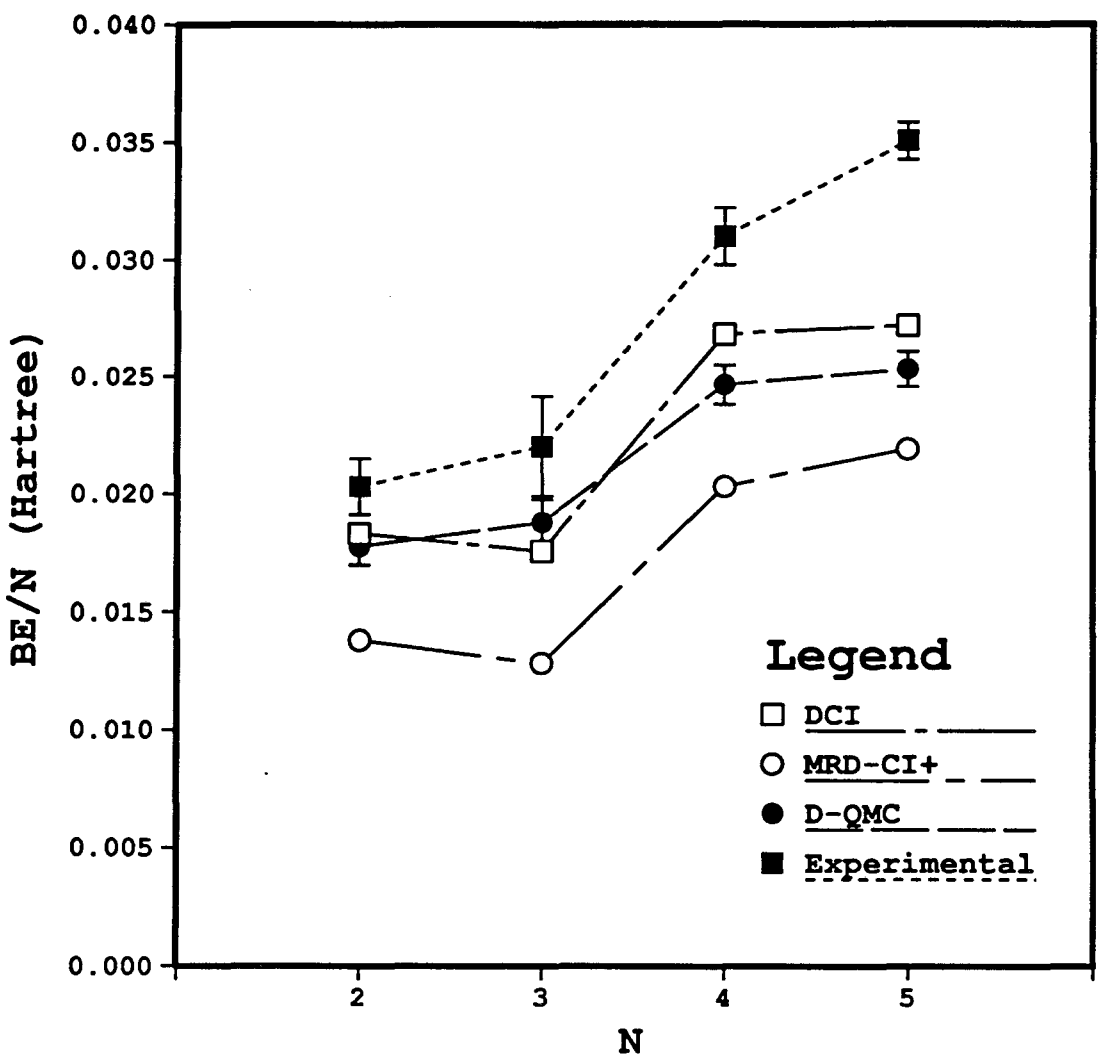


Figure 8.4: Lithium cluster binding energies per atom from D-QMC are compared to D-CI[99], SCF/MRD-CI[14] (noted as MRD-CI+ on plot), and experiment (see Table 8.4 for references).



8.3 Lithium One-Electron Density

The one-electron density is related to the probability of finding an electron in an infinitesimal volume $d\vec{r}$ at some given position $\vec{r}$. Since the location of the other electrons is not essential, they can be averaged over. This is accomplished by integrating the electronic wavefunction over all the coordinates except one. It doesn't matter which electronic coordinate is chosen since the probability density $\left|\psi(\vec{R})\right|^2$ is invariant under particle exchange. Hence, the one-electron density can be found from the expression

$$\rho(\vec{r}) = N \int \left|\psi(\vec{R})\right|^2 d\vec{r}_1\, d\vec{r}_2 \cdots d\vec{r}_{N-1} \tag{8.6}$$

$$= \sum_{\kappa} n_\kappa \left|\phi_\kappa(\vec{r})\right|^2 , \tag{8.7}$$

where N is the number of electrons, κ ranges over the molecular orbitals ϕ_κ, and $n_\kappa = 0, 1, 2$ is the occupation number of the κ-th MO in a HFSCF one-particle basis set.

8.3.1 One-Electron Density and the Anisotropic Harmonic Oscillator

The one-electron density is important for density-dependent electron-electron correlation (see section 4.17); however, notice the smallest value contour line for the lithium dimer, Fig. 8.5, the contour surface describes a region that is nearly spherical, which corresponds to the predicted deformation of the anisotropic harmonic oscillator (AHO) shell model of section 7.2. Plots 8.5 through 8.12 are the one-electron density plots for the various optimized geometries. The contour lines are in units of number of electrons per angstrom cubed. Integrating the one-electron density over all space gives the number of electrons N. The density contours near the nuclei are too dense and are not plotted. Overall, the one-electron densities range five orders of magnitudes with respect to the density confined within the 0.001 contour surface, which represents approximately 99.95% of the electronic charge. Measure the principal axes of the volume confined within the 0.001 contour surface, and if the ratio of the longest axis is taken to the shorter axes then a striking similarity can be observed between the one-electron density and the AHO deformation parameters (see Table 8.5).

The lithium dimer one-electron density volume is approximately spherical as predicted by the AHO model. The other geometry optimized clusters exhibit this same close association with the AHO model, where the trimer and quadramer are approximately axially symmetric. The Nilsson-Clemenger model on the other hand, which is constrained to be axially symmetric, predicts different deformation parameters than the AHO model for $N = 5$; however, the lithium pentamer density volume deformation which shows definite asymmetry is in good agreement with respect to the AHO model.

The range of the density volume is due to the valence electrons of the lithium cluster

and not to the core electrons, which are confined to regions near the nuclei. Since the bonding within the cluster is mediated by the valence electrons, the simple AHO shell model, which only considers the valence electrons as weakly interacting in a one-particle potential, should be effective in predicting some of the valence electron properties. The deformation of the one-electron density volume qualitatively agrees with this simple model.

Table 8.5: The lithium cluster density volume (the volume contained within the 0.001Å^{-3} contour surface) is compared to the anisotropic harmonic oscillator (AHO) model. The table lists the ratios of the longest principal axis length to the two shorter axes lengths. The lithium cluster ratios are obtained from the density plots as shown in Figs. 8.5 through 8.12. The AHO model volume ratios are found from the Table 7.1 deformation parameters $\delta_x, \delta_y, \delta_z$, where the ratio is given by $\frac{1+\delta_i}{1+\delta_m}$, δ_m is the minimum deformation parameter and δ_i is one of the other two. The ratio pairs are arranged in increasing order.

Deformations		
N	AHO	Cluster
2	1.00,1.00	1.1, 1.1
3	1.67,1.67	1.5, 1.6
4	2.00,2.00	1.9, 1.9
5	1.29,1.80	1.3, 2.0

8.3.2 One-Electron Density and Nonnuclear Attractors

Plot 8.5 shows an interesting phenomenon; midway between the two nuclei, there is a local maximum of the one-electron density. This corresponds to the build-up of electronic charge not associated with a nuclear center. This phenomenon is called a "nonnuclear attractor", or also referred to as a "pseudoatom". The studies of R. Glaser, *et.al.*[54], which compared various basis sets and methods, found the existence of nonnuclear attractors for the lithium dimer regardless of method or basis set. He also formulated some general guidelines as to when nonnuclear attractors would most likely occur.

Nonnuclear attractors are found for the larger lithium clusters. For the 'obtuse' lithium trimer (see Fig. 8.6) the nonnuclear attractor spans from the approximate bond centers of the two equal sides. However, for the 'acute' lithium trimer (see Fig. 8.7) there are two attractors, one is bond centered and the other resides more to the interior of triangle. The lithium quadramer (see Fig. 8.8) has two nonnuclear attractors which are within the two triangular faces. At the

center of the geometry, between the two closest lithium nuclei, is also a region of decreased one-electron density. This can also be observed in the D-QMC scatter plot (see Fig. 8.9), where the final random walker distribution for the given time step are projected onto the two dimensional plane that contains the lithium nuclei. The Li_4 D-QMC scatter plot shows, other than the nuclear centers, two regions of greater electron concentrations and in between a diminished concentration. The lithium pentamer (see Fig. 8.12) has a total of three nonnuclear attractors, where two are similar to those of the quadramer and one is bond-centered like the dimer. As with the quadramer, the pentamer D-QMC scatter plot Fig. 8.13 shows a definite region in the center with few electrons.

Figure 8.5: Lithium dimer one-electron density in units of number of electrons per angstrom cubed. Linear distances are in bohr.

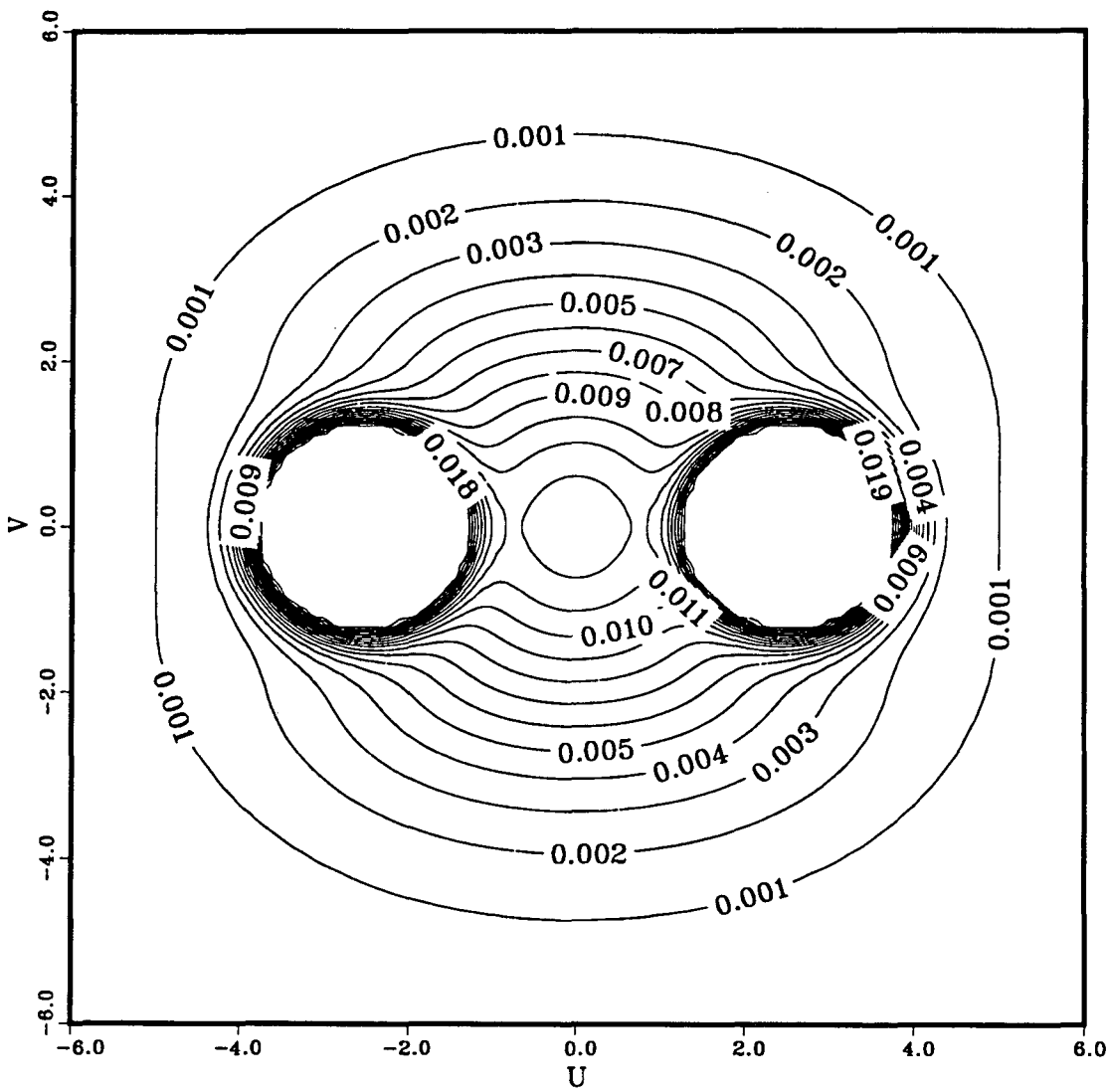


Figure 8.6: 'Obtuse' lithium trimer one-electron density in units of number of electrons per angstrom cubed. Linear distances are in bohr.

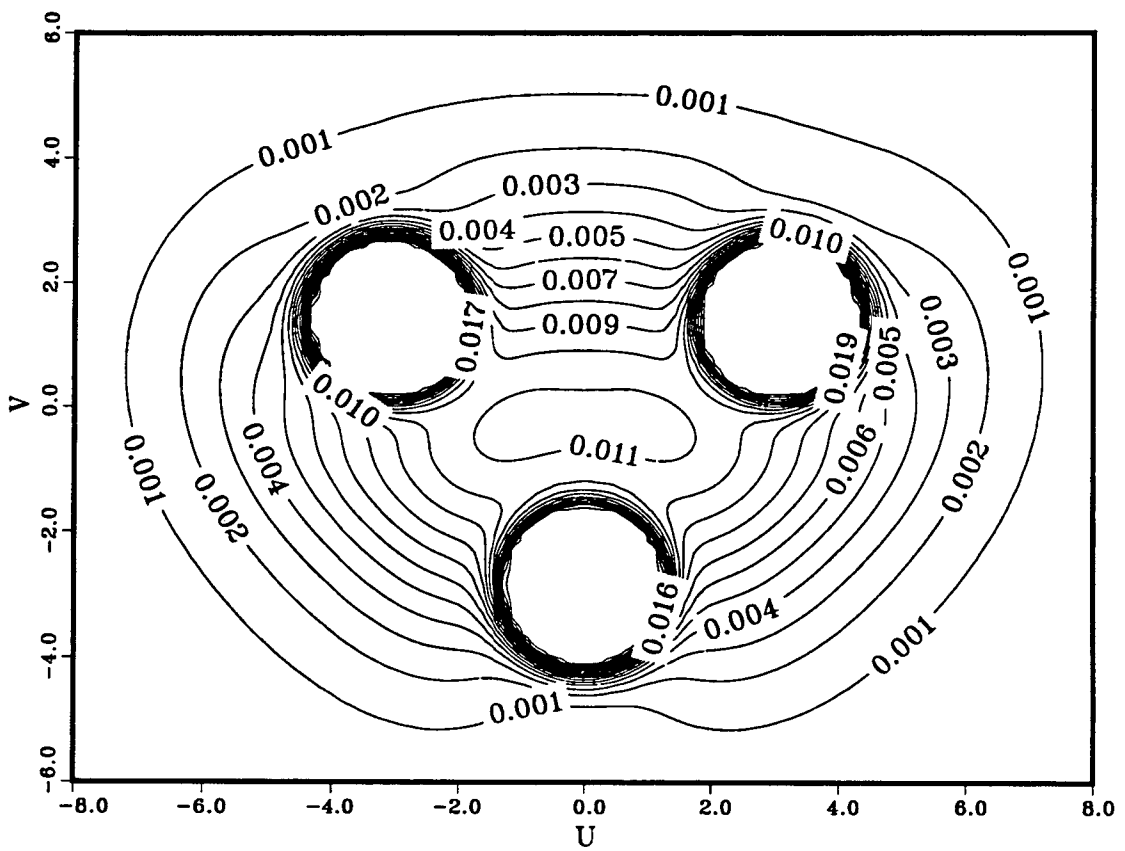


Figure 8.7: 'Acute' lithium trimer one-electron density in units of number of electrons per angstrom cubed. Linear distances are in bohr.

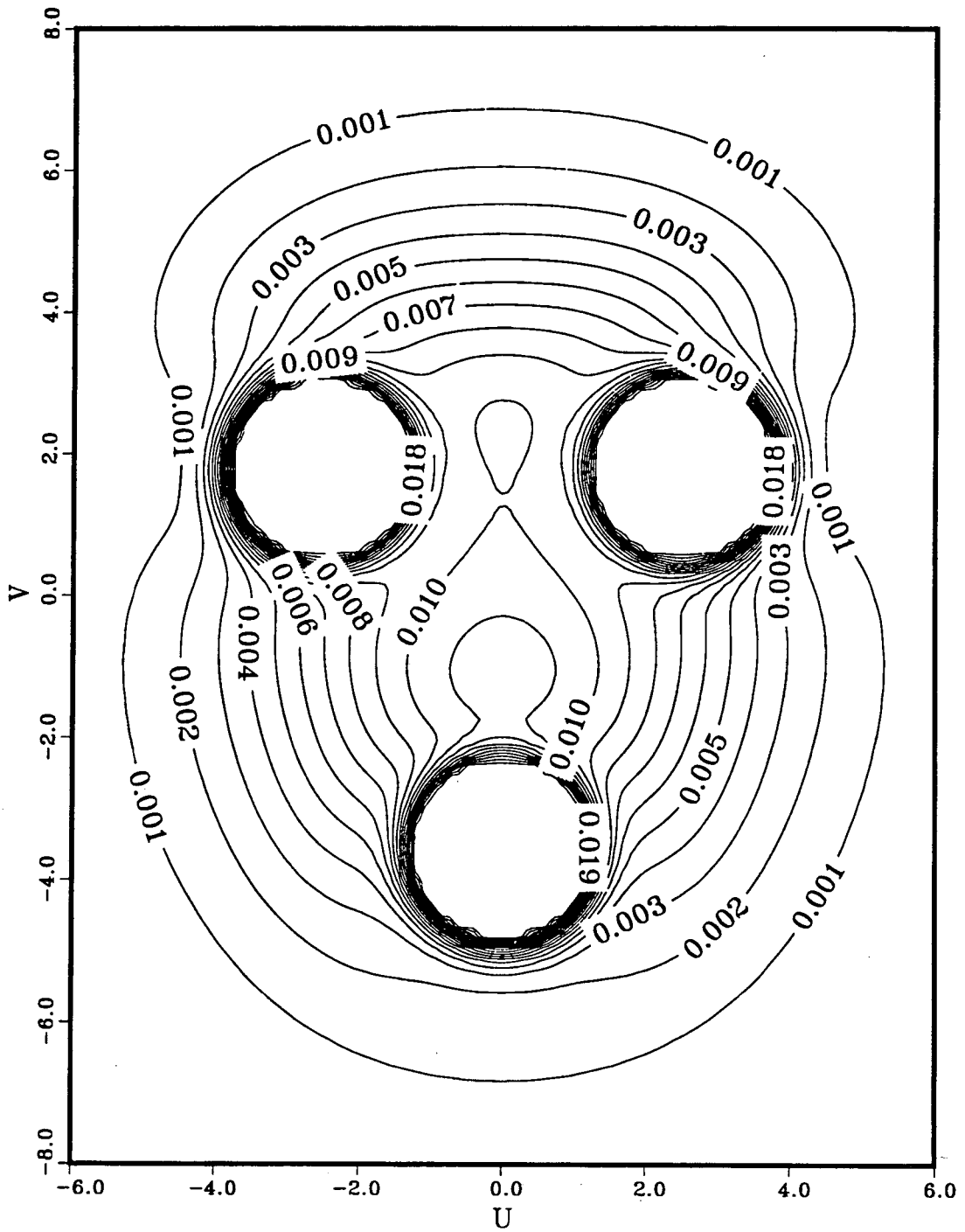


Figure 8.8: Planar lithium quadramer one-electron density in units of number of electrons per angstrom cubed. Linear distances are in bohr.

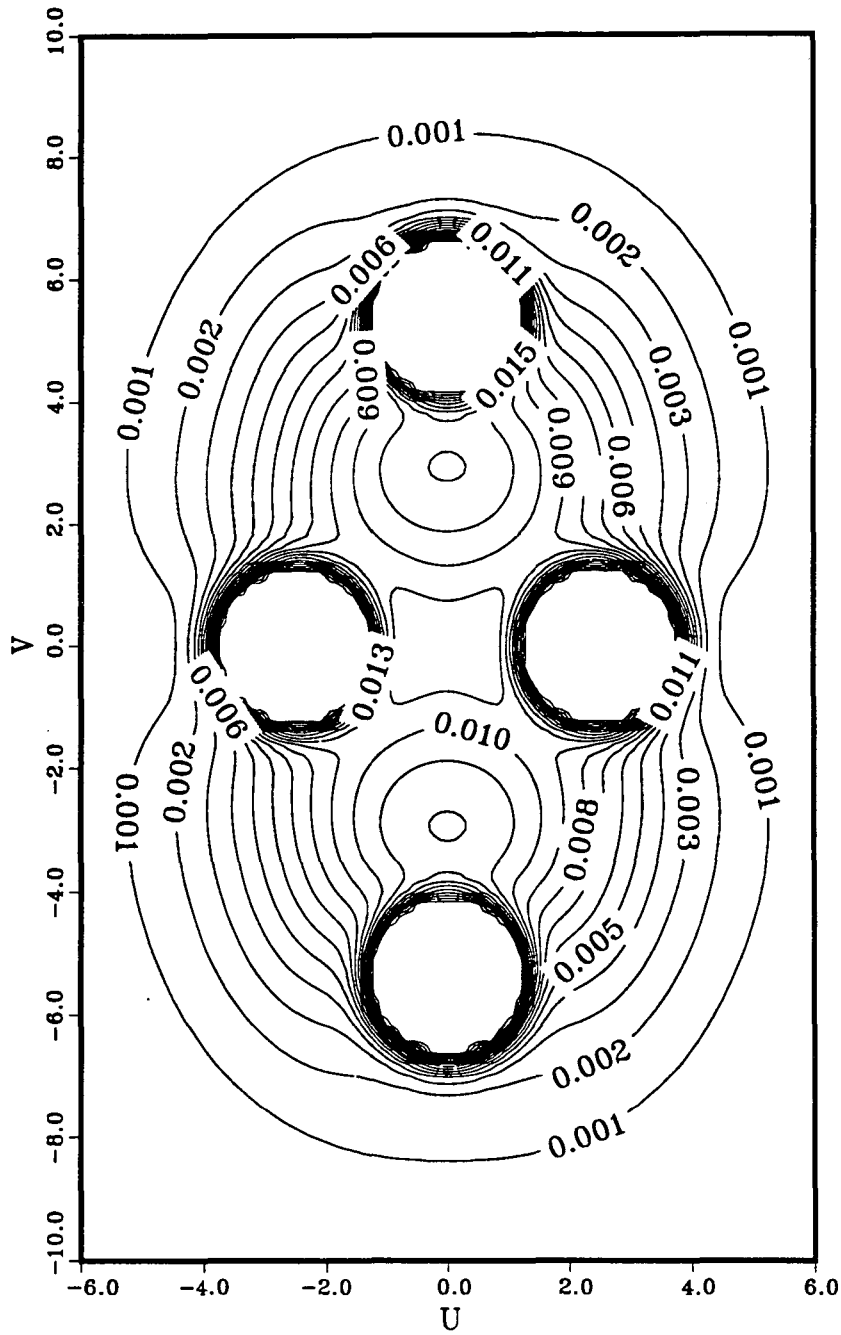


Figure 8.9: Planar lithium quadramer scatter plot of the $\tau = 0.01$ equilibrium distribution of D-QMC random walker positions, which have been projected onto the two dimensional plane that contains the lithium nuclei. This corresponds to the one-electron density integrated with respect to the axis normal to the plane. The small white circles are the nuclear positions.

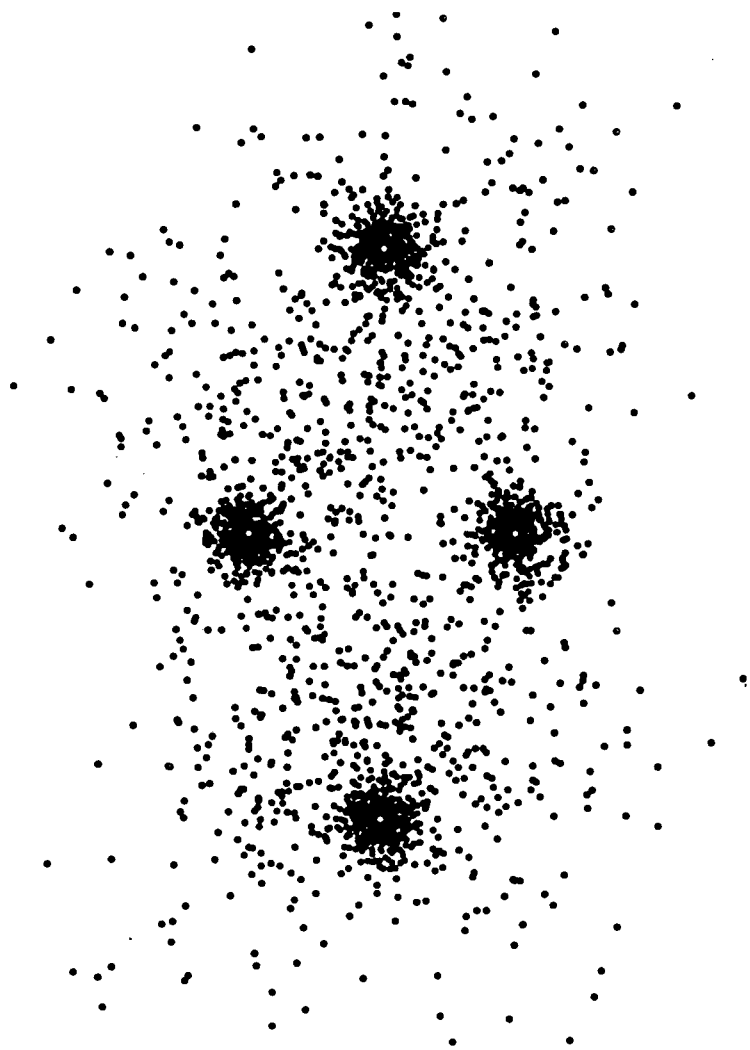


Figure 8.10: Planar lithium quadramer one-electron density in units of number of electrons per angstrom cubed. Linear distances are in bohr. Two nuclei are on the x axis, and the other two are on the y axis. The contours lines are on a plane that cuts vertically and diagonally across the quadramer. The diameter of the spheres correspond to the lithium core diameter. The contour line increment is 0.002.

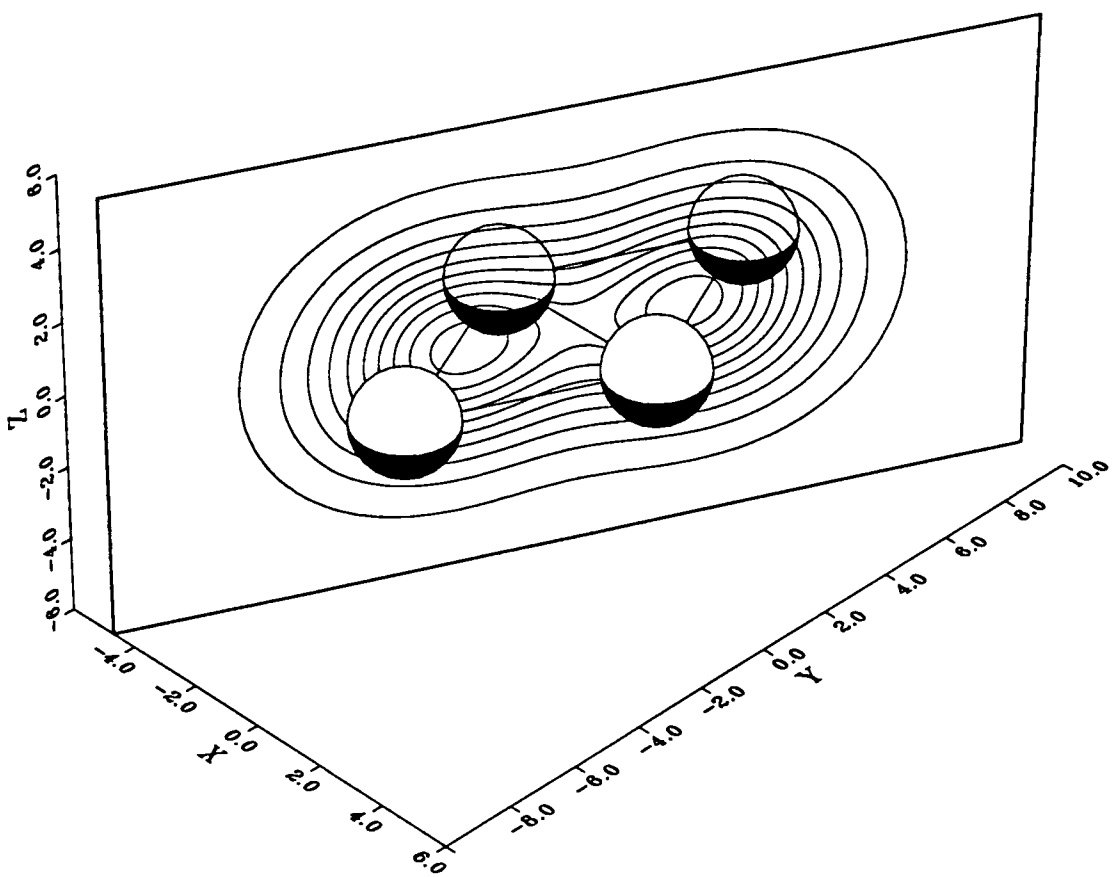


Figure 8.11: Triplet lithium quadramer one-electron density in units of number of electrons per angstrom cubed. Linear distances are in bohr. Two nuclei are located on the xz plane, and the other two are on the yz plane. The contours lines are on a pair of planes, one diagonally cuts across and the other horizontally slices through the quadramer. The diameter of the spheres correspond to the lithium core diameter. The contour line increment is 0.002.

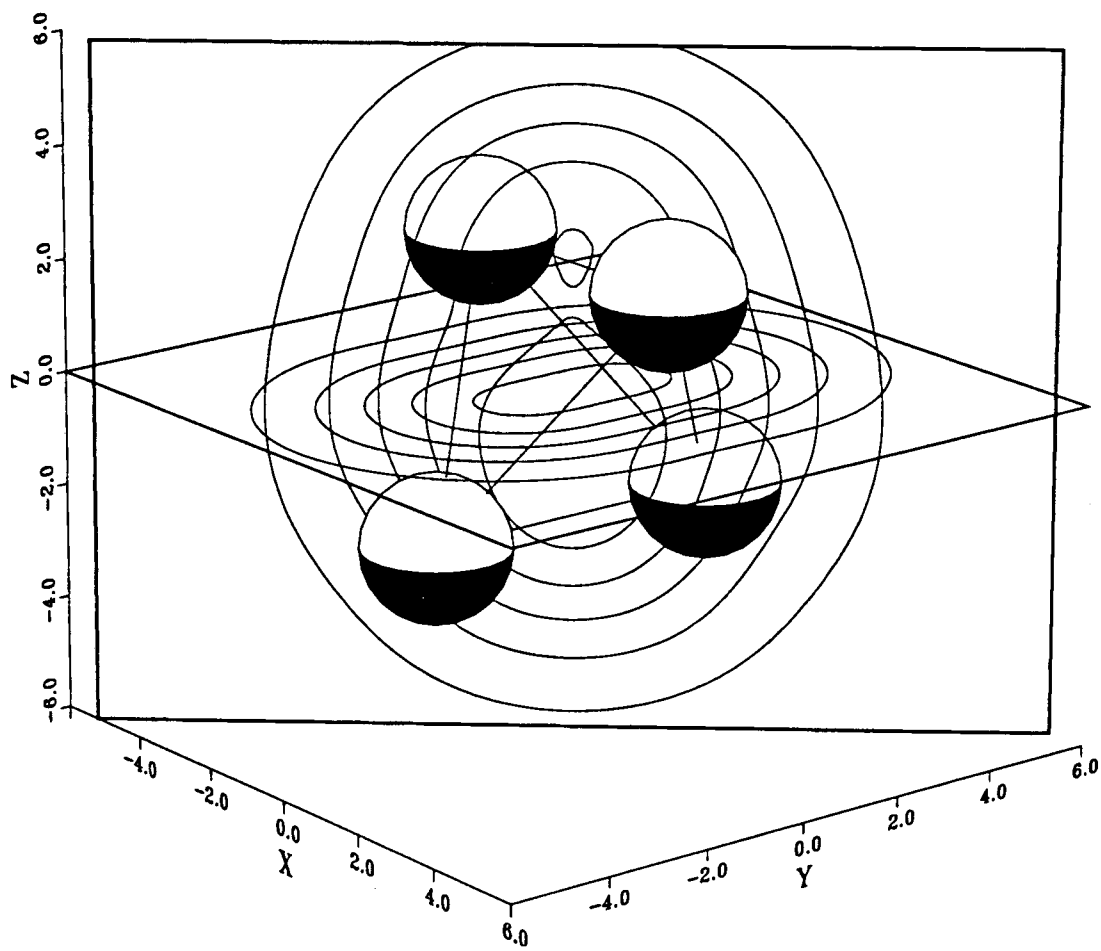


Figure 8.12: Planar lithium pentamer one-electron density in units of number of electrons per angstrom cubed. Linear distances are in bohr.

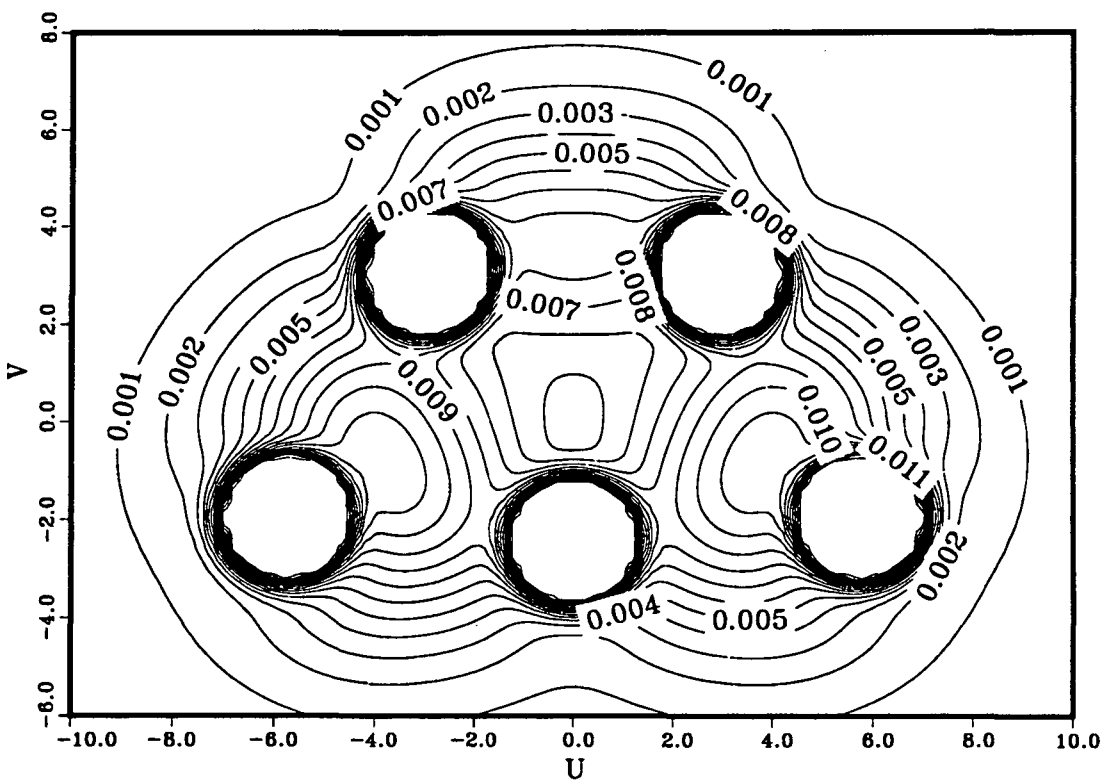
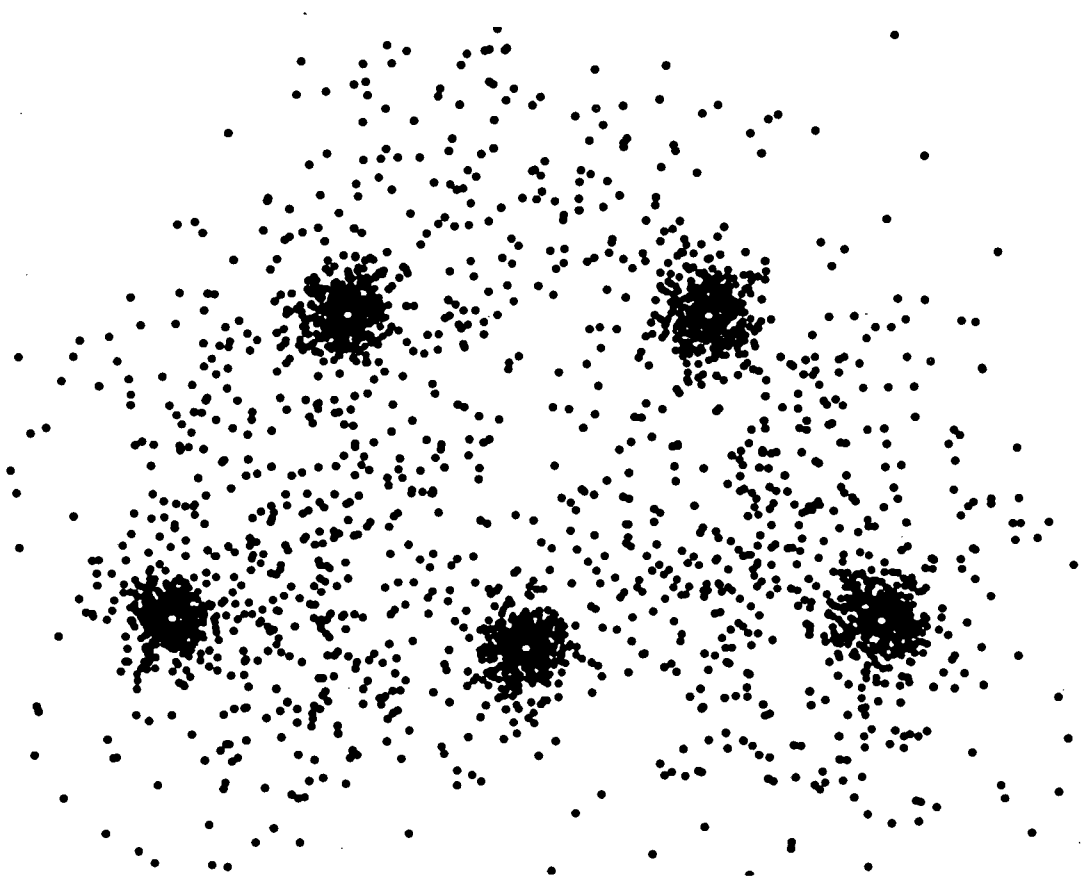


Figure 8.13: Planar lithium pentamer scatter plot of the $\tau = 0.01$ equilibrium distribution of D-QMC random walker positions, which have been projected onto the two dimensional plane that contains the lithium nuclei. This corresponds to the one-electron density integrated with respect to the axis normal to the plane. The small white circles are the nuclear positions.



8.4 QMC Computational Scaling

The efficacy of any method depends on the size of the problems it can address. A computational method can be fantastic for small systems, but become hopelessly impossible for larger systems. One measure of the scalability of an *ab initio* method is the dependence of the CPU time on the number of electrons or some other parameter such as the number of AO basis functions. The limiting operation in Hartree-Fock is computing the various integrals over the basis functions. The number of integrals needed goes as N^4, where N is the number of AOs. The configuration interaction method (CI), in addition to integral evaluations, requires reordering and sorting which scales by at least N^5[41]. Orbital symmetry can be used to cut down the number of integrals to be evaluated, but the scaling still applies. As explained in section 4.1, CI expansions of N-particle wavefunctions are truncated at some level due to computational limits. Reliable estimates of total energies for a system usually requires a set of calculations with increasing numbers of basis functions and higher levels of theory until 'convergence' to a stable value can be obtained[80]. Therefore, the combination produces a time scaling somewhat greater than N^5.

The complexity of the QMC method with importance sampling depends on evaluation of the trial wavefunction. The trial wavefunction requires the evaluation of a determinant, which scales as N^3, where N is the number of MOs. The number of MOs is proportional to the number of electrons. The total energy depends on the nodal surface and does not depend on the number of basis functions except to the degree the basis functions establish the nodal surface. There have been cases where larger basis sets have produced no better results than smaller basis sets[115].

Table 8.6 shows the 'raw' CPU times[1] used to arrive at the present lithium cluster D-QMC results. However, Table 8.6 by itself does not provide an accurate scaling of the CPU time with the size of the problem addressed. The CPU times shown must be normalized with respect to the uncertainty to provide a basis for comparison. Statistically the uncertainty scales as $1/\sqrt{n}$, where n is the number of samples taken. Since the number of samples is directly related to the amount of computer time used, then the CPU time can be adjusted to reflect this. The normalized CPU time estimates how much time is required to obtain a given uncertainty. The uncertainty is considered "fixed" if it needs to be computed to a known uncertainty and "relative" if it's relative to the total energy. Fig. 8.14 shows the times of Table 8.6 normalized with respect to the uncertainties listed in Table 8.2. The times are adjusted such that the two electron case, the lithium atom cation, requires unit time. The lines are the best fits for the $n \geq 4$ data. The fixed uncertainty case shows that the CPU time scales as $N^{3.4}$, where N is the number of electrons. The best that can be done is N^3. An earlier non-optimized version of the

[1] CPU = Computer Processor Unit, and CPU times represents a measure of actual computational time used and approximately measures the complexity of the calculation.

code that was used in a preliminary study had a fixed uncertainty time scaling that went as $N^{5.7}$. The large difference can be attributed to the lack of optimization, the machine architecture and the use of slow computational algorithms. The relative uncertainty time scaling should be two orders of magnitude less than the fixed uncertainty case, since the lithium cluster total energy is approximately linear with respect to the number of atoms and, hence, the number of electrons.

There is also a dependence on the nuclear charge Z [27, 59] that goes approximately as $Z^{5.5}$ to $Z^{6.5}$. This comes from the necessity of using smaller time steps for the core electrons which are more tightly bound to the nucleus. In the case of lithium clusters, this is not an issue since the the nuclear charge remains the same.

Table 8.6: CRAY-2 CPU times are shown for the lowest energy geometry optimized lithium clusters for the neutral and cationic species. The CPU times are measured in minutes. The first two columns refer to the number of atoms N and number of electrons in the cluster n.

CRAY-2 CPU times			
N	n	V-QMC	D-QMC
1	2	1.7	4.5
	3	3.6	8.1
2	5	10.2	203.5
	6	13.4	271.8
3	8	27.3	58.0
	9	37.0	82.6
4	11	60.9	146.3
	12	67.1	148.7
5	14	109.6	268.0
	15	133.1	291.3
total (min)		463.8	1482.8
total (hours)		7.7	24.7

8.5 Lithium Cluster Time-Step Bias

This section shows the empirical lithium cluster time-step bias with respect to the number of electrons, where the time-step bias is defined as the slope of the energy expectation value as a function of time step. The time-step bias is approximately linear with respect to the number of electrons and the numerical values are shown in Table 8.7.

Figure 8.14: Normalized lithium cluster CPU times are plotted. For the "fixed uncertainty" data, the CPU time is normalized such that a computer run of that length of time should return an energy expectation value with the given uncertainty. The "relative uncertainty" is the same except the energy expectation value uncertainty is relative to the energy value. The times for each set of points are adjusted such that the lithium atom cation computational time is unity. The best fits shows that the "fixed uncertainty" CPU time scales as $N^{3.4}$, where N is the number of electrons. The "relative uncertainty" CPU time scales as $N^{1.6}$.

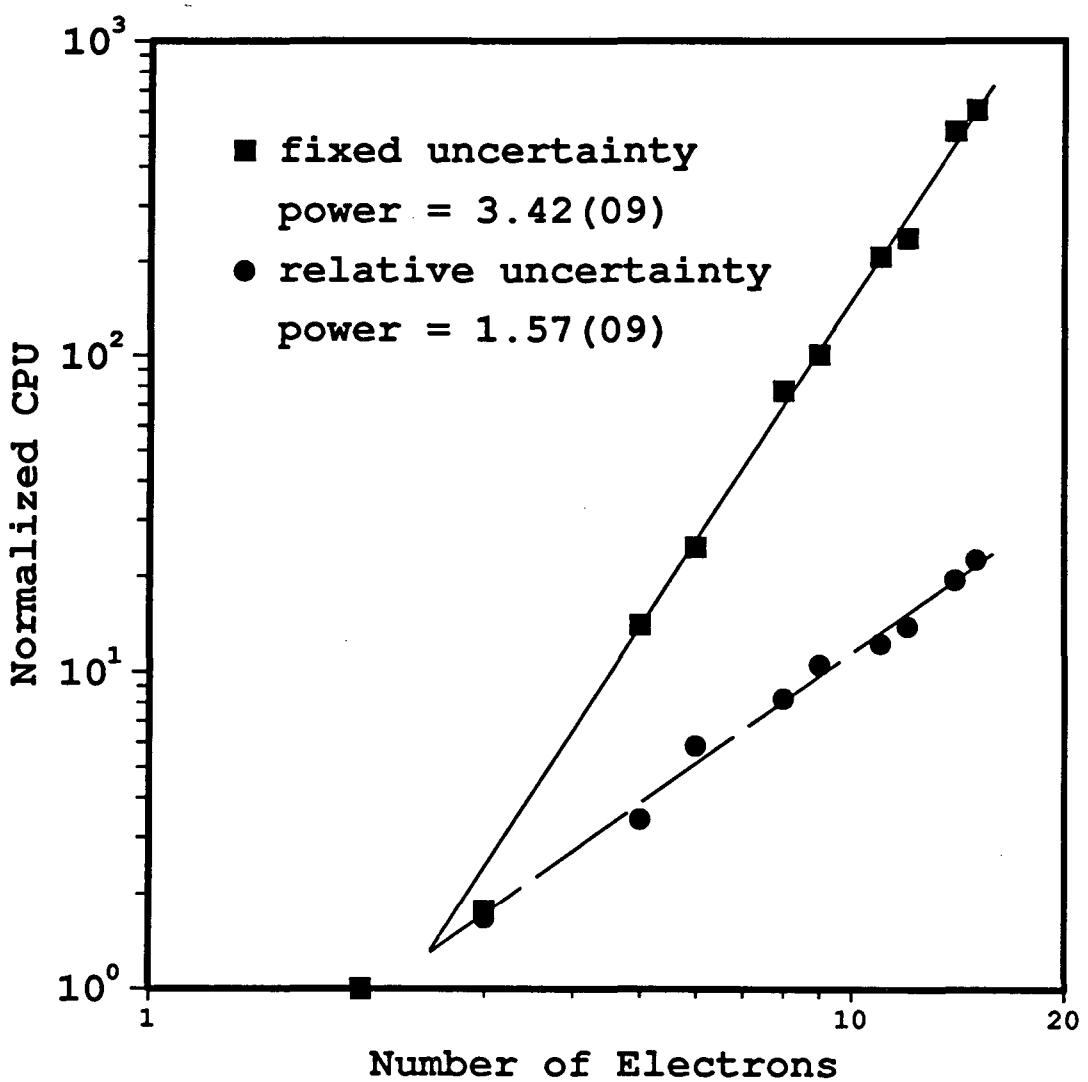


Table 8.7: The D-QMC time-step bias is shown for neutral and cationic lithium clusters, where time-step bias refers to the slope of the energy expectation value as a function of time-step. The units of the time-step bias are Hartree squared. See Table 8.2 the extrapolated total energies.

Time-Step Bias		
N	Neutral	Cationic
1	-0.2152(320)	-0.1114(326)
2	-0.4628(266)	-0.4063(239)
3 'obtuse'	-0.8339(688)	-0.4492(762)
3 'acute'	-0.6453(742)	-0.4940(623)
4 D_{2h}	-1.0696(820)	-0.9746(787)
4 C_{2v}	-0.6731(816)	-0.6618(901)
5	-1.1101(993)	-1.0511(993)

8.6 Conclusion and Future Work

In this work I have introduced density dependent correlation that can partially account for the many-body effects of correlation on electron-electron pairs. A simple form of the electron-electron correlation function, with density dependent parameters and only one free parameter to be optimized, can obtain 56% of the correlation energy in V-QMC calculations for lithium clusters from the dimer to the pentamer. Since the molecular orbitals are already needed for the trial wavefunction, the implementation of density dependence can be efficiently performed with little added computational overhead.

Solutions to the exact electron-nuclear cusp condition are shown to lead to an intractable set of equations for the AO coefficients. A special case can be found if the AOs have vanishing slope as the electron-nuclear distance diminishes, then the electron-nuclear cusp condition can be satisfied by the electron-nuclear correlation function. Alternatively, the electron-nuclear correlation function can satisfy the cusp condition on average, which gives adequate results for V-QMC and D-QMC lithium cluster energies.

The Metropolis algorithm for V-QMC is shown to allow a wide variety of acceptance probabilities (see Table 5.1). Each of the acceptance probabilities were demonstrated to yield equivalent results for the hydrogen atom. The time-step bias, or expectation value of the total energy as a function of time step, for D-QMC is shown to have no constant or half-integer terms with respect to the time step. The first term is shown to be linear with respect to the time step. Exact zero time-step total energies can be obtained by extrapolating a set of D-QMC calculations to zero time-step.

Total energies found by QMC represent a strict upper-bound to the ground state energy.

The ionization potential and the binding energy depend on differences of total energies and, therefore, a strict upper or lower bound does not exist for them. Good agreement with experiment can be due to fortuitous cancellation of errors. However, the QMC method has demonstrated its merits for small molecular systems in yielding total energies that compare well with the atom and lithium dimer best estimated non-relativistic energies. The D-QMC calculation of total energies for small lithium clusters reproduces the available experimental results for the ionization potentials, and to a lesser degree the binding energies. It is clear though that the total energies, thus obtained, may be improved at greater cost by either using trial wavefunctions that more closely approximate the exact nodal surface or by using released-node techniques.

The one-electron density of lithium clusters exhibit remarkable correspondence with the valence electron AHO. It was found that the one-electron contour volume agreed fairly close to the AHO potential shape. Furthermore, the one-electron density also exhibited the phenomenon of nonnuclear attractors for all lithium clusters beyond the atom.

The limiting factor in this lithium cluster study was not the D-QMC CPU time, but the geometry optimization. Geometries were obtained from a CASSCF procedure that limited the CI excitation level to 15 electrons. These calculations required a large amount of disk space, such that Li_5 represented the practical available limit. The geometries found agree with other geometries found in the literature for alkali metal clusters[83, 99, 78]. The trimer geometry is 'obtuse' since the apex angle is greater than 60°. The difference in energy between the 'obtuse' and 'acute' arrangement is found to be less than a milli-Hartree. The quadramer and pentamer are found to be planar in agreement with other work.

Finally, the D-QMC CPU time was shown to scale as $N^{3.4}$ where N is the number of electrons in the system. This compares favorably with other *ab initio* methods that scale N^5 or greater. This lower scaling for D-QMC makes it useful for larger systems.

In the future, further work remains to be done for QMC to provide alternative ways of obtaining potential energy surface (PES) gradients with respect to nuclear positions, for larger clusters and for improved geometries. Independently calculating the total energy for two slightly different geometries can not obtain meaningful finite difference gradients due to the inherent statistical uncertainty in both energies. PES gradients can be evaluated from the Helmann-Feynman theorem or from analytic energy derivatives[79], but the variance in the expectation value can be infinite. Alternatively, a generalized correlated random walk scheme can be used to find PES gradients by using a non-Euclidean metric, g^{ij}, suitably parameterized such that each of the infinitesimally different geometries[120] can be sampled and the random walks can be appropriately weighted. This scheme will require the replacement of gradients and Laplacians with the equivalent covariant forms, for example

$$\nabla^2\psi \rightarrow g^{ij}\frac{\partial^2\psi}{\partial x^i \partial x^j} - g^{ij}\Gamma^k_{ij}\frac{\partial\psi}{\partial x^k}\,, \tag{8.8}$$

where Γ^k_{ij} is the Christoffel symbol. The task will be to find metric tensors g^{ij} that can infinitesimally deform simple geometries, such as trimers, quadramers, etc., along the normal modes.

Bibliography

[1] H.L. Anderson. Scientific uses of the MANIAC. *Jour. Stat. Physics*, 43(5-6):731, June 1986.

[2] James B. Anderson. Random walk simulation of the Schrödinger equation: H_3^+. *Jour. Chem. Physics*, 63(4):1499, August 1975.

[3] James B. Anderson. Quantum chemistry by random walk. H ^2P, H_3^+ D_{3h} $^1A_1$, H_2 $^3\Sigma_u^+$, H_4 $^1\Sigma_g^+$, Be ^1S *Jour. Chem. Physics*, 65(10):4121, 15 November 1976.

[4] James B. Anderson. Quantum Chemistry by Random Walk: H_4 Square. *Int.J.Quant.Chem.*, 15(1):109, January 1979.

[5] James B. Anderson. Quantum chemistry by random walk: Higher accuracy. *Jour. Chem. Physics*, 73(8):3897, 15 October 1980.

[6] James B. Anderson. Symmetry of the nodal hypersurface for $1s2p$ ^3P helium. *Physical Review A*, 35(8):3550, 15 April 1987.

[7] Robert Nicholas Barnett. *Quantum Monte Carlo for Atoms and Molecules.* PhD thesis, University of California, Berkeley, Dept. of Chemistry, November 1989.

[8] Ghassan George Batrouni and Peter J. Reynolds. Accelerated Green's Function Monte Carlo: Avoiding Critical Slowing Down in Simulations Containing Large-Z Atoms. *(Preprint)Physical Review Lett.*, 1990.

[9] D.E. Beck. Self-consistent calculations of the polarizability of small jellium spheres. *Physical Review B*, 30(12):6935, 15 December 1984.

[10] G. Benedek, T.P. Martin, and G. Pacchioni, editors. *Elemental and Molecular Clusters, Proceedings of the 13th International School, Erice, Italy*, volume 6 of *Materials Science.* Springer-Verlag, 1 - 15 July 1987. Whole volume is devoted to various cluster topics.

[11] B. Bernu, D.M. Ceperley, and W.A. Lester, Jr. The Calculation of Excited States with Quantum Monte Carlo. II. Vibrational Excited States. *Jour. Chem. Physics*, 93(1):552, 1 July 1990.

[12] S.M. Blinder. On Green's Functions, Propagators, and Sturmians for the Nonrelativistic Coulomb Problem. *Int.J.Quant.Chem.*, 18:293, 1984.

[13] V. Bonačić-Koutecký, P. Fantucci, and J. Koutecký. Systematic *ab initio* configuration-interaction study of alkali-metal clusters. *Physical Review B*, 36(6):4369, 15 March 1988-II.

[14] I. Boustani, W. Pewstorf, P. Fantucci, V. Bonačić-Koutecký, and J. Koutecký. Systematic *ab initio* configuration-interaction study of alkali-metal clusters: Relation between electronic structure and geometry of small Li clusters. *Physical Review B*, 35(18):9437, 15 June 1987-II.

[15] Ihsan Boustani and Jaroslav Koutecký. Investigation of the electronic and geometric structure of small Li anionic clusters with quantum chemical CI procedure. *Jour. Chem. Physics*, 88(9):5657, 1 May 1988.

[16] Michel Caffarel and Pierre Claverie. Treatment of the Schrodinger Equation Through a Monte Carlo Method Based upon the Generalized Feynman-Kac Formula. *Jour. Stat. Physics*, 43(5-6):797, June 1986.

[17] Michel Caffarel and Pierre Claverie. Development of a pure diffusion quantum Monte Carlo method using a full generalized Feynman-Kac formula. I. Formalism. *Jour. Chem. Physics*, 88(2):1088, 15 January 1988.

[18] Michel Caffarel and Pierre Claverie. Development of a pure diffusion quantum Monte Carlo method using a full generalized Feynman-Kac formula. II. Applications to simple systems. *Jour. Chem. Physics*, 88(2):1100, 15 January 1988.

[19] Michel Caffarel, Pierre Claverie, C. Mijoule, J. Andzelm, and D.R. Salahub. Quantum Monte Carlo method for some model and realistic coupled anharmonic oscillators. *Jour. Chem. Physics*, 90(2):990, 15 January 1989.

[20] J. Carlson and M.H. Kalos. Mirror potentials and the fermion problem. *Physical Review C*, 32(5):1735, November 1985.

[21] A.W. Castleman, Jr. and R.G. Keesee. Clusters: Properties and Formation. *Ann. Rev. Phys. Chem.*, 37:525, 1986.

[22] D. Ceperley. Simulation of Quantum Systems with Random Walks: A New Algorithm for Charged Systems. *Jour. Comp. Physics*, 51(3):404, September 1983.

[23] D. Ceperley, G.V. Chester, and M.H. Kalos. Monte Carlo simulation of a many-fermion study. *Physical Review B*, 16(7):3081, 01 October 1977.

[24] David Ceperley and Berni Alder. Quantum Monte Carlo. *Science*, 231:555, 7 February 1986.

[25] David Ceperley, M.H. Kalos, and Joel L. Lebowitz. Computer Simulation of the Static and Dynamic Properties of a Polymer Chain. *Macromolecule*, 14(5):1472–1479, Sep-Oct 1981.

[26] D.M. Ceperley. Stochastic Solution fo the Many-Body Schrodinger Equation for Fermions. In J. Zabolitzky, M. de Llano, M. Fortes, and J.W. Clark, editors, *Recent progress in Many-Body Theories*, pages 262–269. Springer, Berlin, 1981.

[27] D.M. Ceperley. The Statistical error of Green's function Monte Carlo. *Jour. Stat. Physics*, 43(5-6):815, June 1986.

[28] D.M. Ceperley and B.J. Alder. Ground State of the Electron Gas by a Stochastic Method. *Physical Review Lett.*, 45(7):566, 18 August 1980.

[29] D.M. Ceperley and B.J. Alder. Quantum Monte Carlo for molecules: Green's function and nodal release. *Jour. Chem. Physics*, 81(12):5833, 15 December 1984.

[30] D.M. Ceperley and M.H. Kalos. Quantum Many-Body Problems. In Kurt Binder, editor, *Topics in Current Physics 7, Monte Carlo Methods in Statistical Physics*, page 145. Springer, Berlin, 1979.

[31] S. Chandrasekhar. Stochastic Problems in Physics and Astronomy. *Rev.Mod.Phys.*, 15(1):1, January 1943.

[32] L. Chetouani and T.F. Hammann. Coulomb Green's functions, in an n-dimensional Euclidean space. *J.Math.Phys.*, 27(12):2944, December 1986.

[33] M.Y. Chou, Andrew Cleland, and Marvin L. Cohen. Total energies, abundances, and electronic shell structure of lithium, sodium, and potassium clusters. *Solid State Comm.*, 52(7):645, 1984.

[34] M.Y. Chou and Marvin L. Cohen. Electronic shell structure in simple metal clusters. *Physics Letters*, 113A(8):420, 13 January 1986.

[35] Pierre Claverie and Armelle Denis. The Representation of Functions Through the Combined use of Integral Transforms and Padé Approximates : Padé - Laplace Analysis of Functions as Sums of Exponetials. *Computer Physics Reports*, 9(5):249, July 1989.

[36] Keith Clemenger. Ellipsoidal shell structure in free-electron metal clusters. *Physical Review B*, 32(2):1359, 15 July 1985.

[37] Keith Levern Clemenger. *Spheroidal Shell Structure and Static Electric Polarizabilities of Alkali Metal Clusters.* PhD thesis, University of California, Berkeley, 1986.

[38] Enrico Clementi and Carlo Roetti. Neutral and Ionized Atoms $Z \leq 54$. *Atomic Data and Nuclear Data Tables*, 14(3-4):177, September-October 1974.

[39] D.F. Coker and R.O. Watts. Quantum simulation of systems with nodal surfaces. *Molecular Physics*, 58(6):1113, 1986.

[40] Ellen T. Crowley, editor. *Acronyms, Initialisms, & Abbreviations Dictionary.* Gale Research Co., Detroit, Mich, 8th edition, 1983. MANIAC stood for 'Mathematical Analyzer, Numerical Integrator And Computer'.

[41] Ernest E. Davidson. Configuration Interaction Method. In *Proc. of the Workshop Computational methods for Molecular Structure Determination: Theory and Technique*, number 8, pages 13–1, held at Indiana University, 13-24 August 1979. National Resource for Computation in Chemistry and Quantum Chemisty Program Exchange.

[42] J.D. Doll. Monte Carlo based electronic structure techniques: Analysis and application. *Chem.Physics Letters*, 81(2):335, 15 July 1981.

[43] T.H. Dunning. Gaussian basis functions for use in molecular calculations. I. Contraction of $(9s5p)$ atomic basis sets for the first row atoms. *Jour. Chem. Physics*, 53:2823, 1970.

[44] M. Dupuis, J.D. Watts, H.O. Villar, and G.J.B. Hurst. `HONDO`: version 7.0, 1987. available from: IBM Corporation, Scientific And Engineering Computations, Dept. 48B / MS 428, Neighborhood Road, Kingston, N.Y. 12401, USA.

[45] O. Echt, K. Sattler, and E. Rechnagel. *Physical Review Lett.*, 47:1121, 1981.

[46] A. Einstein. *Ann.Phys.*, 17:549, 1905.

[47] A. Einstein. *Ann.Phys.*, 19:371, 1906.

[48] G.W. Erickson. *Jour. Physical Chem. Ref.Data*, 6:831, 1977.

[49] J. Farges, M.F. de Feruady, B. Raoult, and G. Torchet. Cluster models made of double icosahedron units. *Surf. Sci.*, 156:370, 1985.

[50] David L. Freeman and Jimmie D. Doll. Quantum Monte Carlo study of the thermodynamic properties of argon clusters: The homogeneous nucleation of argon in argon vapor and

"magic number" distributions in argon vapor. *Jour. Chem. Physics*, 82(1):462, 1 January 1985.

[51] Gordon R. Freeman. Emergence of Stochastic Theories: What Are They and Why Are They Special. *Jour.Chem.Educ.*, 61(11):944, November 1984.

[52] R.M. Fye. Study of Trotter-like Approximations. *Jour. Stat. Physics*, 43(5-6):827, June 1986.

[53] Werner H. Gerber. *Theorie des dynamischen Jahn-Teller-Effekts in Li_3 und Untersuchung von Lithium-Molekularstrahlen.* PhD thesis, Universität Bern auf Antrag, 13 November 1980.

[54] Rainer Glaser, Roy F. Waldron, and Kenneth B. Wiberg. Origin and Consequences of the Nonnuclear Attractor in the Ab Initio Electron Density. *Jour. Physical Chem.*, 94(19):7357, 1990.

[55] J.L. Gole and E.C. Stwalley, editors. Number 179 in Amer.Chem.Soc.Symposium Series. Amer.Chem.Soc., 1982.

[56] R.C. Grimm and R.G. Storer. A New Method for the Numerical Solution of the Schrodinger Equation. *Jour. Comp. Physics*, 4(2):230, August 1969.

[57] R.C. Grimm and R.G. Storer. Monte–Carlo Solution of Schrödinger's Equation. *Jour. Chem. Physics*, 7(1):134, February 1971.

[58] Brian Hammond, 1988. (private communication).

[59] Brian Lee Hammond. *Monte Carlo for the Electronic Structure of Molecules.* PhD thesis, University of California, Berkeley, Dept. of Chemistry, November 1988.

[60] E. Hartmann and E. Clementi. *Physical Review*, 133:A1295, 1964.

[61] W.J. Hehre, L. Radom, P.v.R. Schleyer, and J.A. Pople. *Ab Initio Molecular Orbital Theory.* Wiley, New York, 1986.

[62] W.J. Hehre, R.F. Stewart, and J.A. Pople. Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals. *Jour. Chem. Physics*, 51(6):2657, 15 September 1969.

[63] Levere Hostler. Runge-Lenz Vector and the Coulomb Green's Function. *J.Math.Phys.*, 8(3):642, March 1967.

[64] K.P. Huber and G. Herzberg. *Molecular Spectra and Molecular Structure*, volume 2. Van Nostrand Reinhold Co., 1979.

[65] Sigeru Huzinaga. Gaussian-Type Functions for Polyatomic Systems. I. *Jour. Chem. Physics*, 42(4):1293, 15 February 1965.

[66] Malvin H. Kalos and Paula A. Whitlock. *Monte Carlo Methods, Volume I: Basics.* John Wiley & Sons, 1986. Page 73, a good comprehensive explanation of the Metropolis algorithm.

[67] M.H. Kalos. Monte Carlo Calculations of the Ground State of Three- and Four-Body Nuclei. *Jour. Chem. Physics*, 128(4):1791, 15 November 1962.

[68] M.H. Kalos. Energy of a Boson Fluid with Lennard-Jones Potentials. *Physical Review A*, 2(1):250, July 1970.

[69] M.H. Kalos. Optimization and the Many-Fermion Problem. In Malvin H. Kalos, editor, *Monte Carlo Methods in Quantum Problems*, page 19. D.Reidel Pub. Co., 1984.

[70] M.H. Kalos, D. Levesque, and L. Verlet. Helium at zero temperature with hard-sphere and other forces. *Physical Review A*, 9(5):2178, May 1974.

[71] Frederick W. King and Victor Shoup. Calculations on the ^2S ground state of the lithium atom. *Physical Review A*, 33(5):2940, May 1986.

[72] S. Kirkpatrick, Jr. C.D. Gelatt, and M.P. Vecchi. Optimization by Simulated Annealing. *Science*, 220:671–680, 1983.

[73] Charles Kittel. *Introduction to Solid State Physics.* John Wiley & Sons, 5-th edition, 1976. page 74, Table 3.1, "Cohesive Energies of the Elements".

[74] D.J. Klein and H.M. Pickett. Nodal hypersurfaces and Anderson's random-walk simulation of the Schrodinger equation. *Jour. Chem. Physics*, 64:4811, 01 June 1976.

[75] W.D. Knight, Keith Cleminger, Walt A. de Heer, Winston A. Saunders, M.Y. Chou, and Marvin L. Cohen. Electronic Shell Structure and Abundances of Sodium Clusters. *Physical Review Letters*, 52(24):2141, June 1984.

[76] W.D. Knight, Walt A. de Heer, and Keith Cleminger. Electronic shell structure in potassium clusters. *Solid State Comm.*, 53(5):445, 1985.

[77] D.D. Konowalow and M.E. Rosenkrantz. Electronic Structure and Spectra of Light Alkali Diatomic Molecules and Their Molecular Cations. In Gole and Stwalley [55], page 3.

[78] Jaroslav Koutecký and Piercarlo Fantucci. Theoretical Aspects of Metal Atom Clusters. *Chem. Rev.*, 86(3):539, June 1986.

[79] William A. Lester, Jr. and Brian L. Hammond. Quantum Monte Carlo for the Electronic Structure of Atoms and Molecules. *Ann.Rev.Phys.Chem*, 41:283, 1990.

[80] Bowen Liu. Configuration Interaction: Results. In *Proc. of the Workshop Computational methods for Molecular Structure Determination: Theory and Technique*, number 8, pages 17–1, held at Indiana University, 13-24 August 1979. National Resource for Computation in Chemistry and Quantum Chemisty Program Exchange.

[81] K.S. Liu, M.H. Kalos, and G.V. Chester. *Physical Review A*, 10:303, 1974.

[82] José Luís Martins. *Calcul des proprietes structurales et electroniques d'agregats de metaux alcalins*. PhD thesis, Ecole Polytechnique Federale De Lausanne, 1984.

[83] José Luís Martins, Jean Buttet, and Roberto Car. Equilibrium Geometries and Electronic Structures of Small Sodium Clusters. *Physical Review Letters*, 53(7):655, 13 August 1984.

[84] José Luís Martins, Jean Buttet, and Roberto Car. Electronic and structual properties of sodium clusters. *Physical Review B*, 31(4):1804, 15 February 1985.

[85] José Luís Martins, Roberto Car, and Jean Buttet. Electronic properties of alkali trimers. *Jour. Chem. Physics*, 78(9):5646, 1 May 1983.

[86] T. Matsubara and H. Matsuda. *Prog.Theor.Phys.(Kyoto)*, 16:416, 1956.

[87] Mark H. McAdon and William A. Goddard, III. Generalized Valence Bond Studies of Metallic Bonding: Naked Clusters and Applications to Bulk Metals. *Jour. Physical Chem.*, 91(10):2607, 1987.

[88] Frank Mentch and James B. Anderson. Quantum chemistry by random walk: Importance sampling for H_3^+. *Jour. Chem. Physics*, 74:6307, 1 June 1981.

[89] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculation by Fast Computing Machines. *Jour. Chem. Physics*, 21(6):1087, June 1953.

[90] Nicholas Metropolis and S. Ulam. Monte Carlo Method (The). *J.Amer.Stat.Assoc.*, 44(247):335, September 1949.

[91] H.B. Michaelson. *J.Appl.Phys.*, 48:4729, 1977.

[92] C.E. Moore. *NSRDS-NBS*, 34, September 1970.

[93] Jules W. Moskowitz and K.E. Schmidt. The domain Green's function method. *Jour. Chem. Physics*, 85(5):2868, 01 September 1986.

[94] Jules W. Moskowitz and K.E. Schmidt. Erratum: The domain Green's function method [JCP 85,2868(1986)]. *Jour. Chem. Physics*, 87(3):1906, 1 August 1987.

[95] J.N. Murrel. The Many-body Expansion of the Potential Energy Function for Elemental Clusters. *Int.J.Quantum Chem.*, 37:95, 1990.

[96] H. Nakatsuji, J. Ushio, and T. Yonezawa. Cluster expansion of the wavefunction. Potential energy curves of the ground, excited, and ionized states of Li_2. *Can.J.Chem.*, 63:1857, 1985.

[97] J.C. Phillipps. Chemical Bonding, Kinectics, and the Approach to Equilibrium Structures of Simple Metallic, Molecular, and Network Microclusters. *Chem. Rev.*, 86(3):539, June 1986.

[98] William H. Press and Saul A. Teukolsky. Quasi- (that is, Sub-) Random Numbers. *Computers in Physics*, 3(6):76, Nov-Dec 1989. The Sobol's sequence results in Quasi- or Sub-random numbers that maximally fill the space at the same time they are "maximally avoiding" each other, This leads to uncertainties that scale approximately $1/n$.

[99] B.K. Rao and P. Jena. Physics of small metal clusters: Topology, Magnetism, and electronic structure. *Physical Review B*, 32(4):2058, 15 August 1985.

[100] B.K. Rao, P. Jena, and M. Manninen. Relationship between topological and magnetic order in small metal clusters. *Physical Review B*, 32(1):477, 1 July 1985.

[101] B.K. Rao, S.N. Khanna, and P. Jena. Evolution of the electronic and structural properties of microclusters. *Physical Review B*, 36(2):953, 15 July 1987-I.

[102] A.K. Ray, J.L. Fry, and Charles W. Myles. A configuration interaction study of small lithium clusters. *J.Phys.B: At.Mol.Phys*, 18:381, 1985.

[103] Peter J. Reynolds, David M. Ceperley, Berni J. Alder, and William A. Lester, Jr. Fixed Node quantum Monte Carlo for molecules. *Jour. Chem. Physics*, 77:5593, 1 December 1982.

[104] Peter J. Reynolds, R.K. Owen, and William A. Lester, Jr. Is there a zeroth order time-step error in diffusion quantum Monte Carlo? *Jour. Chem. Physics*, 87(3):1905, 1 August 1987.

[105] C.C.J. Roothaan. *Reviews of Mod.Phys.*, 23(2):69, April 1951.

[106] C.C.J. Roothaan. *Reviews of Mod.Phys.*, 32(2):179, April 1960.

[107] Stuart M. Rothstein, Narayan Patil, and Jan Vrbik. Time Step Error in Diffusion Monte Carlo Simulations: An Empirical Study. *Jour. Comp. Chem.*, 8(4):412, September 1987.

[108] Reuven Y. Rubenstein. *Simulations and the Monte Carlo Method.* John Wiley & Sons, 1981.

[109] K. Sattler. Experimental research on free metallic, ionic and Van der Waals clusters. *Surf. Sci.*, 156:292, 1985.

[110] Leonard I. Schiff. *Quantum Mechanics.* McGraw-Hill Book Company, third edition, 1955. page 169.

[111] E. Schumacher, W.H. Gerber, H.P. Härri, M. Hofmann, and E. Scholl. Preparation, Electronic Spectra, and Ionization of Metal Clusters. In Gole and Stwalley [55], page 83.

[112] Kathy Selby, Michael Vollmer, Jun Masui, Vitaly Kresin, Walt A. de Heer, and W.D. Knight. Surface plasma resonances in free metal clusters. *Physical Review B*, 40(8):5417, 15 September 1989.

[113] *Surf. Sci.*, 156, 1985. Whole volume is devoted to various cluster topics.

[114] G.D. Stein. Cluster beam sources; predictions and limitations of the nucleation theory. *Surf. Sci.*, 156:44, 1985.

[115] Zhiwei Sun, 1990. (private communication).

[116] Zhiwei Sun, Peter J. Reynolds, R. Kent Owen, and William A. Lester, Jr. Monte Carlo study of electron correlation functions for small molecules. *Theo.Chim.Acta*, 75:353–368, 1989.

[117] Zhiwei Sun, Sheng yu Huang, Robert N. Barnett, and William A. Lester, Jr. Wave function optimization with a fixed sample in quantum Monte Carlo. *Jour. Chem. Physics*, 93(5):3326, 1 September 1990.

[118] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry.* MacMillan Publishing Co., 1982. A good introductory book into *ab initio* methods.

[119] Todd C. Thompson, Grant Izmirlian, Jr., Stephen J. Lemon, and Donald G. Truhlar. Consistent analytic representation of the two lowest potential energy surfaces for Li_3, Na_3, and K_3. *Jour. Chem. Physics*, 82(12):5597, 15 June 1985.

[120] C.A. Traynor and J.B. Anderson. Parallel Monte Carlo Calculations to Determine Energy Differences Among Similar Molecular Structures. *Chem.Physics Letters*, 147(4):389, 10 June 1988.

[121] H.F. Trotter. *Proc. Am. Math. Soc.*, 10:545, 1959.

[122] A. Veillard and E. Clementi. *Jour. Chem. Physics*, 49:2415, 1968.

[123] Jan Vrbik, Daniel A. Legare, and Stuart M. Rothstein. Infinitesmal differential diffusion quantum Monte Carlo: Diatomic molecular properties. *Jour. Chem. Physics*, 92(2):1221, 15 January 1990.

[124] Bryan H. Wells. The Differential Green's function Monte Carlo method. The Dipole moment of LiH. *Chem.Physics Letters*, 115(1):89, March 1985.

[125] L.T. Wille. Searching potential energy surfaces by simulated annealing. *Nature*, 324:46–48, 1986.

[126] D.W. Wood. *Physical Review Letters*, 46:749, 1981.

[127] C.H. Wu. Thermochemical properties of gaseous Li_2 and Li_3. *Jour. Chem. Physics*, 65(8):3181, 15 October 1976.

[128] C.H. Wu. Experimental Investigation of a Stable Lithium Cluster. The Thermochemical Study of the Molecule Li_4. *Jour. Physical Chem.*, 87(9):1534, 1983.

[129] C.H. Wu. Thermochemical properties of the lithium cluster Li_5. *Jour. Chem. Physics*, 91(1):546, 1 July 1989.

Appendix A

Diffusion/Drift/Branching

A.1 Analytic Derivation of QMC Green's Function

This Appendix forms the backbone to the various QMC methods in that the short-time Green's function for each particular method can be found by appropriately setting the given constants. The term "short-time" is understood to mean that for short enough time steps in imaginary time the fields $\vec{F}$ and α are slowly varying and that they can be approximated with constants on the path of propagation. There will regions where this is not so, but overall the short-time approximation provides an adequate beginning to obtain an approximate Green's function. Table A.1 lists the typical definitions of the constants $\vec{F}$ and α for each of methods.

Table A.1: Several definitions for the Green's function constants $\vec{F}$ and α for each of the various QMC methods. The spatial vectors $\vec{R}_0$ and $\vec{R}$ refer to the initial and final points of the integral path respectively.

Method	$\vec{F}$	α	see
QMC - no importance sampling	0	$\frac{1}{2}(V(\vec{R})+V(\vec{R}_0))-E_T$	section 3.2
V-QMC	$\vec{F}_Q(\vec{R}_0)$	0	chapter 5
D-QMC	$\vec{F}_Q(\vec{R}_0)$	$\frac{1}{2}(E_L(\vec{R})+E_L(\vec{R}_0))-E_T$	chapter 6

The Green's function is the general solution to the following parabolic partial differential equation, where n is the dimensionality of the space, $\vec{R}\in\mathcal{R}^n$ is a vector in that space,

$$\frac{\partial}{\partial\tau}f \;-\; \underset{\text{diffusion}}{D\nabla^2 f} \;+\; \underset{\text{drift}}{D\vec{F}\cdot\nabla f} \;+\; \underset{\text{branching}}{\alpha f} \;=\; 0 \;, \tag{A.1}$$

where $\vec{F}$ and α are constants with respect to $\vec{R}$ and τ on the integral path, but may depend on the end points. The constant $D=\frac{1}{2}$ is in atomic units. Parts of Eq. A.1 are labeled with respect to their operational interpretation. The effects of the inclusion of each term are most easily understood from the resulting density matrix (Eq. A.10).

The Green's function is defined from the solution of the following partial differential equation

$$\frac{\partial G}{\partial \tau} - D\nabla^2 G + D\vec{F}\cdot\nabla G + \alpha G = \delta(\vec{R}-\vec{R}_0)\delta(\tau-\tau_0)\,. \tag{A.2}$$

The solution to Eq. A.2 can be obtained by taking the Fourier transform with respect to $\vec{R}-\vec{R}_0$, and noting the following relationships

$$\begin{aligned} \delta(\vec{R}-\vec{R}_0) &= \int \frac{d\vec{K}}{(2\pi)^n} e^{i\vec{K}\cdot(\vec{R}-\vec{R}_0)} \\ G(\vec{R}\tau|\vec{R}_0\tau_0) &\equiv \int \frac{d\vec{K}}{(2\pi)^n} e^{i\vec{K}\cdot(\vec{R}-\vec{R}_0)} g(\vec{K},\tau-\tau_0) \\ \nabla &\rightarrow i\vec{K}\,, \end{aligned} \tag{A.3}$$

where $\vec{K}\in\mathcal{R}^n$ is a vector in the reciprocal space, Then the Green's function equation A.2 is reduced to a differential equation in time

$$\left(\frac{\partial}{\partial\tau}+\gamma^2\right) g = \delta\left(\tau-\tau_0\right) \tag{A.4}$$

where

$$\gamma^2 = D\vec{K}^2 + iD\vec{F}\cdot\vec{K} + \alpha\,. \tag{A.5}$$

The next step is to take the Fourier transform with respect to $\tau-\tau_0$, which transforms the above equation to

$$\left(i\omega+\gamma^2\right)\mathcal{G} = 1 \tag{A.6}$$

where the transformation relates the following quantities

$$\begin{aligned} \delta(\tau-\tau_0) &= \int_{-\infty}^{\infty}\frac{d\omega}{2\pi} e^{i\omega(\tau-\tau_0)} \\ g(\vec{K},\tau-\tau_0) &= \int_{-\infty}^{\infty}\frac{d\omega}{2\pi} e^{i\omega(\tau-\tau_0)}\mathcal{G}(\vec{K},\omega) \\ \frac{\partial}{\partial\tau} &\rightarrow i\omega\,. \end{aligned}$$

In the reciprocal space of $\vec{R}$ and τ the differential equation becomes a purely algebraic one where the solution to $\mathcal{G}$ is simply

$$\mathcal{G} = \frac{-i}{\omega - i\gamma^2}\,. \tag{A.7}$$

The transform back to $g(\vec{K},\tau-\tau_0)$ results in a complex contour integration along the real ω axis that encloses the upper half ω-plane if $\tau-\tau_0$ is positive or the lower half ω-plane if $\tau-\tau_0$ is negative, such that the exponential in either case then becomes vanishingly small. The pole of the integrand is at $\omega = i\gamma^2$ with a residue $= e^{-\gamma^2(\tau-\tau_0)}$. The pole is enclosed only if $\tau-\tau_0>0$. (The behavior if $\tau-\tau_0=0$ is ill defined and the principle value can be taken resulting in half the residue.) Therefore, the transform back to τ space yields

$$g(\vec{K},\tau-\tau_0) = e^{-\gamma^2(\tau-\tau_0)}\Theta\left(\tau-\tau_0\right) \tag{A.8}$$

where Θ is the Heaviside step function

$$\Theta(t) = \begin{cases} 1 & t > 0 \\ \frac{1}{2} & t = 0 \\ 0 & t < 0 \end{cases} . \tag{A.9}$$

The transform back to $\vec{R}$ space via equation A.3 is trivially done by completing the square in the argument with respect to $\vec{K}$, thus introducing an exponential quadratic in $\vec{R} - \vec{R}_0$.

Finally the resulting form of the Green's function, or density matrix, is

$$G(\vec{R}\tau|\vec{R}_0\tau_0) = \left(\frac{1}{4\pi D(\tau - \tau_0)}\right)^{n/2} \Theta(\tau - \tau_0)\, e^{-\alpha(\tau-\tau_0)} e^{\frac{-(\vec{R}-\vec{R}_0-D(\tau-\tau_0)\vec{F})^2}{4D(\tau-\tau_0)}} . \tag{A.10}$$

The diffusion/drift/branching is evident from equation A.10 in that the branching gives the predictable exponential growth related to α, and the drift and diffusion by a Gaussian of $\vec{R} - \vec{R}_0 - D(\tau - \tau_0)\vec{F}$.

A.2 Green's Function Stochastic Simulation

The stochastic realization of the density matrix $G(\vec{R}\tau|\vec{R}_0\tau_0)$ is made easier if the exponential term in α is factored out such that

$$G(\vec{R}\tau|\vec{R}_0\tau_0) = e^{-\alpha(\tau-\tau_0)} T(\vec{R}\tau|\vec{R}_0\tau_0) . \tag{A.11}$$

The second term, $T(\vec{R}\tau|\vec{R}_0\tau_0)$, represents a conditional, or a transition probability distribution function for transitions to $\vec{R}, \tau$ if initially at $\vec{R}_0, \tau_0$ and has the usual properties

$$\begin{aligned} T(\vec{R}\tau|\vec{R}_0\tau_0) &\geq 0 \\ \int d\vec{R}\, T(\vec{R}\tau|\vec{R}_0\tau_0) &= 1 . \end{aligned}$$

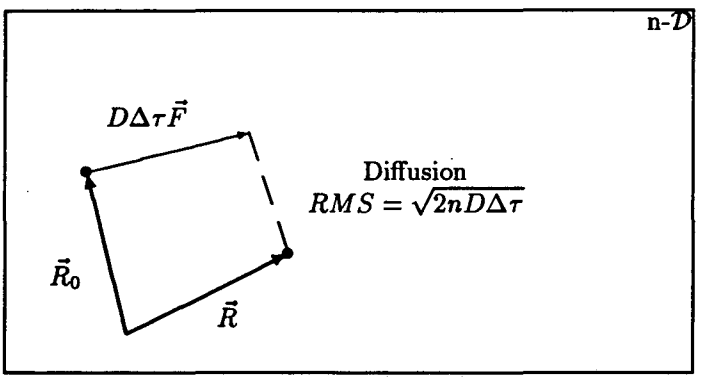
These relations just say that $T(\vec{R}\tau|\vec{R}_0\tau_0)$ is positive semi-definite transition probability and starting at any given point always yields a transition somewhere else in space. The stochastic realization of this conditional probability distribution function is:

1. Set time step $\Delta\tau = \tau - \tau_0$.
2. Start at given point $\vec{R}_0$.
3. Generate a Gaussian distributed vector $\chi \in \mathcal{R}^n$ with mean $< \chi >= 0$ and variance $< \chi^2 >= 2Dn\Delta\tau$.
4. Return the new point $\vec{R}$ distributed according to transition probability $T(\vec{R}\tau|\vec{R}_0\tau_0)$, where $\vec{R} = \vec{R}_0 + D\Delta\tau\vec{F} + \chi$.

See figure A.1 for a diagrammatic description of the above process, where the light solid vector corresponds to a drift of $D\Delta\tau\vec{F}$ and the dashed line represents a diffusive step of χ, in the transitions from $\vec{R}_0$ to $\vec{R}$.

The exponential $m = e^{-\Delta\tau\alpha}$ can either be used to 'branch', or copy, the new point according to the multiplicity m, or as a numerical weight carried along with the point, or a combination of both. Branching is necessarily an integer process and can be realized by creating $[m + u]$ number of copies of the point, where u is a uniformly distributed random number on the interval $[0, 1)$,and the truncation operator $[a]$ returns the largest integer less than or equal to a.

Figure A.1: Quantum Monte Carlo algorithm: heavy vectors are n dimensional coordinate vectors for the initial and final points, the light vector is the importance sampling drift according to trial wavefunction $\psi_T(\vec{R})$, the dashed line represents diffusion or Brownian motion with a given root-mean-squared length.



A.3 Renormalization

Operationally, you have a collection of random walkers located at the set of points $\vec{R}_{0i}$ which are distributed in some fashion throughout the n dimensional space. Each random walker is propagated via the above stochastic algorithm to a new set of points $\vec{R}_i$. Then m_i copies of each point are made and this set of points becomes the initial set of points $\vec{R}_{0i}$ for the next generation. This process can be repeated several times as necessary.

Renormalization of the sample population is necessary at times due to the finite memory available in current computers and their storage devices. Because of the branching term, $e^{-\alpha\Delta\tau}$, in the density matrix (Eq. A.10), it's either possible for the population to grow beyond the available array storage or to diminish to nothing depending whether the exponential argument is respectively positive or negative overall. In this implementation renormalization was done sparingly and only on demand when the population size threatened to overrun the available array storage or was close to vanishing.

Appendix B

Expectation Values

This appendix is devoted to the various methods that can be employed to obtain expectation values from the QMC sampled population. All eigenfunctions shown in this appendix have the proviso that they are eigenfunctions of the Hamiltonian for a given set of boundary conditions (*i.e.* fixed-node).

B.1 Energy Expectation Values

Exact Expectation values of the ground state energy from the D-QMC mixed distribution can be easily found by advantageously using the hermiticity of the Hamiltonian operator, H, and letting it act on the eigenstate directly,

$$\frac{\langle\varphi_0|H|\psi_T\rangle}{\langle\varphi_0|\psi_T\rangle} = \varepsilon_0 \tag{B.1}$$

$$= \frac{\int \varphi_0 H \psi_T \, d\vec{R}}{\int \varphi_0 \psi_T \, d\vec{R}} \tag{B.2}$$

$$= \frac{\int \frac{H\psi_T}{\psi_T} \varphi_0 \psi_T \, d\vec{R}}{\int \varphi_0 \psi_T \, d\vec{R}} \tag{B.3}$$

hence

$$\varepsilon_0 \approx \frac{1}{\mathrm{N}} \sum_{i=1}^{\mathrm{N}} E_L(\vec{R}_i) \tag{B.4}$$

where the points $\vec{R}_i$ are distributed according to the mixed distribution pdf $f = \varphi_0 \psi_T$.

B.2 Expectation Values of Other Operators

Expectation values of operators that don't commute with the Hamiltonian can be especially difficult. Generally, the ground state does not represent an eigenstate of the operator in question. To obtain exact expectation values for some operator $\mathcal{O}$,

$$\langle\mathcal{O}\rangle_{\varphi_0^2} \equiv \frac{\langle\varphi_0|\mathcal{O}|\varphi_0\rangle}{\langle\varphi_0|\varphi_0\rangle}, \tag{B.5}$$

requires finding some way to sample the square of the exact ground state. However, all that is easily available is the mixed distribution and the expectation value associated with it,

$$\langle \mathcal{O} \rangle_f \equiv \frac{\langle \varphi_0 | \mathcal{O} | \psi_T \rangle}{\langle \varphi_0 | \psi_T \rangle} . \tag{B.6}$$

Since the ground state and trial function are not generally eigenfunctions of the operator $\mathcal{O}$, then the mixed expectation value is appreciably different from the exact expectation value.

B.2.1 Approximate Expectation Values

The following approximation to the exact expectation value for a non-commuting operator was due to Ceperley and Kalos[30]. Beginning with the assumption that the trial function ψ_T is fairly 'close' to the exact eigenstate, then the difference

$$|\delta\rangle = |\varphi_0\rangle - |\psi_T\rangle , \tag{B.7}$$

can be considered 'small' and we can neglect terms to some given order of δ. Writing the variational and mixed-distribution expectation values to second order becomes

$$\langle \mathcal{O} \rangle_f = \langle \varphi_0 | \varphi_0 \rangle^{-1} \left[\langle \varphi_0 | \mathcal{O} | \varphi_0 \rangle - \langle \varphi_0 | \mathcal{O} | \delta \rangle - \frac{\langle \varphi_0 | \mathcal{O} | \varphi_0 \rangle \langle \varphi_0 | \delta \rangle}{\langle \varphi_0 | \varphi_0 \rangle} \right] + O(\delta^2) \tag{B.8}$$

and

$$\langle \mathcal{O} \rangle_{\psi_T^2} = \langle \varphi_0 | \varphi_0 \rangle^{-1} \left[\langle \varphi_0 | \mathcal{O} | \varphi_0 \rangle - 2 \langle \varphi_0 | \mathcal{O} | \delta \rangle - 2 \frac{\langle \varphi_0 | \mathcal{O} | \varphi_0 \rangle \langle \varphi_0 | \delta \rangle}{\langle \varphi_0 | \varphi_0 \rangle} \right] + O(\delta^2) . \tag{B.9}$$

By taking the following difference an approximate expectation value will differ from the exact to order δ^2,

$$\langle \mathcal{O} \rangle_{\text{approx}} \equiv 2 \langle \mathcal{O} \rangle_f - \langle \mathcal{O} \rangle_{\psi_T^2} \tag{B.10}$$

$$= \frac{\langle \varphi_0 | \mathcal{O} | \varphi_0 \rangle}{\langle \varphi_0 | \varphi_0 \rangle} + O(\delta^2) . \tag{B.11}$$

An advantage to this approximate expectation value is that both quantities are easily computable via QMC methods. Either two independent runs one a V-QMC and the other a D-QMC can be executed, or a 'weighted' D-QMC and 'un-weighted' V-QMC can be performed. The meaning of the last statement is that instead of using branching to bias the random walk, weights are propagated along with each random walker to produce the mixed-distribution[123]. The expectation values can be computed with and without weights, thus producing a correlated random walk[124] The resulting D-QMC and V-QMC expectation values will have an uncertainty in the difference which is far less than the uncertainty in the expectation values.

B.2.2 Exact Expectation Values by Future Walking

We need not be content with approximate expectation values. There is a method to obtain the exact expectation value at the expense of simplicity. This can be done by the method of future walking[81, 7].

We want to sample from φ_0^2 rather than $f = \psi_T \varphi_0$ so we 'correct' the distribution with some sort of weighting factor $P(\vec{R})$ such that

$$P(\vec{R}) = \frac{C \varphi_0(\vec{R})}{\psi_T(\vec{R})} , \tag{B.12}$$

where constant C has yet to be determined. The operator $\mathcal{O}$ in the expectation value need not be a scalar operator in the representation chosen. As with the local energy a local scalar field can be constructed such that $\mathcal{O}(\vec{R}) = \frac{\mathcal{O}\psi_T(\vec{R})}{\psi_T(\vec{R})}$, which will not change the following derivation. The desired exact expectation value can be arrived at from the mixed distribution with this weight factor such that

$$\langle \mathcal{O} \rangle_{\varphi_0^2} = \int P(\vec{R})\varphi_0(\vec{R})\mathcal{O}(\vec{R})\varphi_0(\vec{R})\, d\vec{R} \tag{B.13}$$

$$= \frac{\int P(\vec{R})\varphi_0(\vec{R})P(\vec{R})\mathcal{O}(\vec{R})\psi_T(\vec{R})\, d\vec{R}}{\int P\varphi_0(\vec{R})P(\vec{R})\psi_T(\vec{R})\, d\vec{R}}, \tag{B.14}$$

where we have assumed that the wavefunctions are real.

For the moment we consider

$$\hat{\rho}(\vec{R}|\vec{R}_0;\tau) = \frac{\psi_T(\vec{R})}{\psi_T(\vec{R}_0)} \sum_i e^{-(\varepsilon_i - E_T)\tau} \varphi_i(\vec{R})\varphi_i(\vec{R}_0) \tag{B.15}$$

which has the property

$$\hat{\rho}(\vec{R}|\vec{R}_0;0) = \delta(\vec{R} - \vec{R}_0) \tag{B.16}$$

and in the long-time limit approaches $\frac{\psi_T(\vec{R})}{\psi_T(\vec{R}_0)} e^{-(\varepsilon_0 - E_T)\tau} \varphi_0(\vec{R})\varphi_0(\vec{R}_0)$. Then it follows that

$$\int d\vec{R}\hat{\rho}(\vec{R}|\vec{R}_0;\tau) \xrightarrow{\tau\to\infty} \left[e^{-(\varepsilon_0 - E_T)\tau} \int d\vec{R}\psi_T(\vec{R})\varphi_0(\vec{R}) \right] \frac{\varphi_0(\vec{R}_0)}{\psi_T(\vec{R}_0)} \equiv \frac{C\varphi_0(\vec{R})}{\psi_T(\vec{R})}. \tag{B.17}$$

The evaluation of the above integral of $\hat{\rho}$ can be realized by starting a single configuration at $\vec{R}_0$ and propagating forward in imaginary time some arbitrarily period of time, say T. The integral with respect to $\vec{R}$ then represents the number, N_T, of random walkers descended from this lone configuration at $\vec{R}_0$. Return N_T as an estimate for $P_T(\vec{R}_0)$ where

$$\lim_{T\to\infty} P_T(\vec{R}) = P(\vec{R}). \tag{B.18}$$

Before contributing to the sum for the expectation value of $\langle \mathcal{O} \rangle_{\varphi_0^2}$ keep track of how many configurations at some later time are descended from that point then weight the evaluation $\mathcal{O}(\vec{R})$ accordingly.

Appendix C

Program Documentation

This appendix lists the documentation for the program **q903cd**. The program itself is not listed since it is approximately 250 pages/12000 lines long and can be obtained directly from the author. The program has been written in **FORTRAN** and is designed to be modular with respect to the trial wavefunction evaluation. The code is **FORTRAN 77** compliant except for the use of **NAMELIST**s, which is fairly standard on most compilers. At the beginning of each subroutine a detailed explanation of the subroutine purpose and argument list is given, which will aid the interested hacker.

The program has been vectorized for the CRAY architecture, but will run under the VAX/VMS operating system and be easily adapted to other environments by linking in a different **SYS90CD** file that contains the system dependent calls.

Some of the input file specifications closely follow the input specifications of the *ab initio* quantum chemistry code called **HONDO** of M. Dupuis, J.D. Watts, H.O. Villar, G.J.B. Hurst[44].

```
              INPUT SPECIFICATIONS FOR PROGRAM QMC90CD

                 VECTORIZED V-QMC, FN-DQMC, RN-GFMC
                        -- VERSION 9.03CD --
                     +----------------------+
                     | By Richard Kent Owen |
                     +----------------------+

      and B.L.Hammond, R.N.Barnett, S.Alexander, P.J.Reynolds,
                        PROF. Wm. A. Lester,Jr.
                  UNIVERSITY OF CALIFORNIA, BERKELEY
                                and
                     LAWRENCE BERKELEY LABORATORY
                      BERKELEY, CALIFORIA  94720

  THIS PROGRAM IS BASED ON THE FIXED-NODE DIFFUSION QUANTUM MONTE
  CARLO ALGORITHM (SEE P.J.REYNOLDS, D.M.CEPERLEY, B.J.ALDER, AND
  W.A.LESTER, J.CHEM.PHYS. 77 (1982) 5593  AND
  THE RELEASED-NODE GREEN'S FUNCTION MONTE CARLO (D.M. CEPERLEY,
  B.J. ALDER, J.CHEM.PHYS. 81 (1984) 5833.

  DISCLAIMER:
      THE AUTHOR MAKES NO GUARANTEES AND/OR WARRANTIES EITHER
  EXPRESSED OR IMPLIED WITH RESPECT TO THIS CODE, DOCUMENTATION,
  OR THE RESULTS OBTAINED.  THE USER OF THIS CODE, ETC. DOES SO
  AT HIS OWN RISK, AND THE AUTHOR IS IN NO WAY RESPONSIBLE FOR
  ANY DAMAGES WHICH MAY RESULT FROM PROPER OR IMPROPER USE.

  PROPRIETARY STATEMENT:
      THE DEVELOPMENT OF THIS PROGRAM AND ITS DOCUMENTATION ARE
  COPY-RIGHTED BUT ARE PUT INTO THE GENERAL PUBLIC USE WITH SOME
  LIMITATIONS. THE USER MAY MAKE CHANGES TO THIS CODE AND
  DOCUMENTATION AS HE SEES FIT.  HOWEVER, THE AUTHOR LIST SHOULD
  REMAIN INTACT AS IS.  IF THE CODE IS CHANGED SIGNIFICANTLY THE
  PROGRAMMER MAY ADD HIS NAME AS A COAUTHOR FOLLOWING THE PRINCIPLE
  AUTHOR.  ANY PAPERS THAT ARE PUBLISHED WITH RESULTS FROM THIS
  PROGRAM SHOULD ALSO ACKNOWLEDGE THE PRINCIPLE AUTHOR AND PROGRAM.
  ANY FURTHER DISTRIBUTION OF PROGRAM MUST INCLUDE THE ANNOTATED
  SOURCE CODE.

DATA $BANNER
------------

     One line title for first page of output

EXAMPLE:
$BANNER
    *************  Li dimer DZ STO basis + EEJ + ENJ ****************
$END
```

DATA $BASIS

Specification of geometry, basis set and E-N correlation parameters. Closely follows the HONDO style of input but has been freed of strict FORMAT rules and does not support symmetry generation of geometry.

```
EXAMPLE:                                                  | COMMENTS
$BASIS
   Carbon Hydride -- double-zeta basis                    | 'First  Card'
C         6.    0.        0.        0.    12.0  119.0    | 'Second Card'
CUSP  1S            7.969                                 | 'Third  Card'
CUSP  1S            5.231                                 | 'Third  Card'
CUSP  2S            1.820                                 | 'Third  Card'
CUSP  2SP           1.168                                 | 'Third  Card'
      2P            2.726                                 | 'Third  Card'
                                                          | 'Blank  Card'
H         1.    1.6738    0.        0.    0.05  11.0     | 'Second Card'
      1S            1.640                                 | 'Third  Card'
      1S            1.120                                 | 'Third  Card'
                                                          | 'Blank  Card'
$END

EXAMPLE ANNOTATED:
   Carbon Hydride -- double-zeta basis                    | 'First  Card'
   ^ BASIS TITLE

H         1.    1.6738    0.        0.    0.05  11.0     | 'Second Card'
^ ATOM LABEL (A3)                         ^ LAMBDA
          ^ NUCLEAR CHARGE                      ^ NU
               ^ X       ^ Y        ^ Z

CUSP  2SP           1.168                                 | 'Third  Card'
^ CUSP-less IF PRESENT
      ^ AO ORBITAL OR SHELL
                  ^ ZETA FOR AO
```

```
NAMELIST $CNTRL  IPRINT,VECFLG,PUNFLG,STATOL,REETOL,RENTOL,
---------------  PSITOL,DETOL,FQTOL

      IPRINT     PRINT FLAGS (Default=0) (Combine by addition)
                    0 ... Normal limited output

                    1 ... Averages or Opt. Coefs. & Energies printed
                          after each block
                   10 ... Info, MO's, ERROR messages,
                          Test results (unit7) printed
                  100 ... Print timing statistics for routines

                1 000 ... Scale factors printed out

               10 000 ... Routine headers

              100 000 ... Debug Info (walk parameters)

            1 000 000 ... Punch configurations after each printed walk
                          (see program PUNCOLL)
           10 000 000 ...

          100 000 000 ... DON'T use orbital shells (use general AO's)
                          (Good for comparison tests of AO's)
        1 000 000 000 ... RUN but DON'T 'WALK'

      ISCALE     Scale MO's to prevent possible Over-/Underflows
                    0 ... NO SCALING
                    1 ... SCALING (Default)

      PUNFLG     Printing option for electronic configurations
                   -1 ... DO NOT print PUN file (unit2)
                    0 ... Standard PUN file  (Default)
                    1 ... Configs written to PUN file after EACH block
                    2 ... Print density scatter plot data (unit7)

      STATOL     Tolerance for basis functions in MO's (Default=1.E-6)

      REETOL     Minimum allowed E-E distance in XXCHEK (Default=0.1)

      RENTOL     Minimum allowed E-N distance in XXCHEK (Default=0.1)

      PSITOL     Minimum allowed value of ABS(PsiT) in XXCHEK
                 *** NOT IMPLEMENTED ***

      FQTOL      Maximum allowed quantum force in XXCHEK
                 *** NOT IMPLEMENTED ***

      DETOL      Minimum allowed value of ratio of determinants
                 *** NOT IMPLEMENTED ***
```

NAMELIST $GUESS IGUESS,IFILL,ICHECK,EGUESS

IGUESS	Initial distribution of electronic configs -1 ... Generate RANDOMLY (Default) 1 ... Read from PUN file
IFILL	Fill ensemble if IGUESS=1 0 ... Go with given number of configs from PUN file 1 ... Copy configs as necessary (Default)
ICHECK	Check ensemble if IGUESS=1 0 ... DO NOT check ensemble (Default) 1 ... Check with tolerances in $CNTRL
EGUESS	Initial guess Energy, ONLY USED IF IGUESS=-1 and ETRIAL=0
RAN	Initial random number (Default=0.)

```
NAMELIST $NOPT  OPTTOL,OPTDMP,OPTFLG,OPTCND,OSTEP,OPTCI,OPTZETA
--------------

   Coefficient optimization via V-QMC
   Only if RUNFLAG => 10.  Can do up to 16 different parameters
   by using 2-point auxiliary runs for each parameter
   and fitting a parabola to the energies.  If the minimum
   falls within the end-points then the new parameter is set to obtain
   that value for the next block.

   OPTTOL          Tolerance limit for changes in parameters
                   before stopping optimization
                   (Default = .005)

   OPTDMP          Damping factor to reduce the parameter step
                   size after each block
                   (Default = .90)

   OPTFLG(20)      What type of optimization for corresponding run
            1 ... E-E correlation parameters and E-N CRANGE
            2 ... CI coefficients
            3 ... Atomic orbital ZETA'S (can't use exact densities for
                                         density dependent correlation)
            4 ... E-N correlation parameters - LAMBDA, NU

   OPTCND(20)      Conditions for optimization of corresponding run
    (if OPTFLG = 1)  Select parameters for optimization
            1 ... E-N CRANGE
           10 ... E-E EECF(2)
          100 ... E-E EECF(3)
         1000 ... E-E EECF(4)
        10000 ... E-E EECF(0)=EECF(1)
    (if OPTFLG = 2)  Select which of first 16 CI coefs to hold fixed
                     EXAMPLE: OPTCND=000010  - hold the 2nd one fixed
    (if OPTFLG = 3)  NO conditions

    (if OPTFLG = 4)  NO conditions

   OSTEP(16,20)    Maximum step size for LINK-LIST
                   If first one for each run is < 0 then all are assigned a
                   value=abs(first)*corresponding parameter.
                   (Default OSTEP(1,*) = -.20)

   OPTCI(100)      Which of the CI coefs are linked
   OPTZETA(100)         ''      ZETA parameters ''
   OPTENC(100)          ''      E-N LAMBDA,NU parameters ''
                   such that the linked parameters are varied identically
                   it might be useful to do a short run first to
                   get an idea how the ZETA's, CI's, LAMBDA,NU's are noted.
                   -OR- just make the first one negative and the
                   program will automatically make a LINK-LIST of
                   ALL parameters that are the same (neglecting
                   minus signs in the case of CI coefs).

                   example: OPTCI=1,2,0,1,2,3,3,0,3
                   - links 1st & 4th, 2nd & 5th, 6th & 7th & 9th,
                     and the 3rd & 8th are free
```

```
NAMELIST $PSIT   EECF,EECFLG,ENCLFG,CRANGE,TFSCN,GTHETA,GRAT
--------------
      EECF(0:4)  E-E correlation function (EECF) parameters
         Note: Al|Au Like|Unlike spins only applies to single-det. wvfns

      EECFLG     Flag for which form of EECF to use
          (Default) 0 ... PADE-JASTROW
                           EECF(0)|EECF(1)=Al|Au*, EECF(2)=B
                                A*r     A        *Typically for HFSCF
                               ----- - ---              Al=.25
                               1+B*r    B               Au=.5
                          U = e

                    1 ... PADE-JASTROW B=Owen's Function of Density (n)
                           EECF(0)|EECF(1)=Al|Au*, EECF(2)=B1, EECF(3)=C
                                                          1/12
                           where B=B1*((n(r )+C)*(n(r )+C))
                                          i         j
                    2 ... SUN-REYNOLDS EXP(EXP)
                           EECF(0)|EECF(1)=Al|Au*, EECF(2)=B, EECF(3)=C,
                                                 2
                                   -A*r/B -C*r
                               -B*e
                          U = e

                    3 ... SUN-REYNOLDS EXP(EXP)
                          B=Owen's Function of density (n)
                           EECF(0)|EECF(1)=Al|Au*, EECF(2)=B, EECF(3)=C,
                                   -A*r/B
                               -B*e
                          U = e
                                              B1
                           where B=--------------------1/12
                                   ((n(r )+C)*(n(r )+C))
                                         i         j
      ENCFLG     FLAG FOR WHICH FORM OF ENCF TO USE
                    1 ... PADE-JASTROW
                    0 ... PADE-JASTROW determine parameters automatically
                          only first time then treat as constants
         (Default) -1 ... PADE-JASTROW determine parameters automatically
                          then vary as necessary for optimization runs

                              -lambda*r    lambda
                               -------- + --------
                                1+nu*r       nu
                          U = e

      CRANGE     RANGE OF E-N CUSP
                    = 0.  If not concerned with CUSP condition.
                    Typically >>1/MAX(Z) if solving for E-N CUSP.
         (Default) = .05

       TFSCN     Thomas-Fermi screening term for Yukawa potential
                 (if = 0.0 (default) then just ordinary Coulomb pot.)

                          charge*exp(-TFSCN*r)
                   V(r) = --------------------
                                     r
```

```
GTHETA      Initial Guiding Wavefuntion parameter for inclusion
            of the 'one-electron density' to eliminate nodes for
            the released-node GFMC guiding wavefunction.
            (must be non-negative to be used)

             Psi = SQRT(Psi^2 + GTHETA*prod(n(r )))
                G          T             i     i

GRAT        Ratio of configurations such that they satisfy:
                 Psi^2 < GTHETA*prod(n(r )))
                    T             i     i
            If GRAT < 0 then GTHETA is kept fixed.
            (default = .10)
```

```
DATA $VEC
---------

      M.O. vectors in standard HONDO FORMAT
      M.O.'s must be listed in the same order as in $WFN
      coefficients must be listed in the same order as in $BASIS
      P -Shells have the order   Px,Py,Pz
      SP-Shells                S,Px,Py,Pz

EXAMPLE: 3 M.O's, 4 basis functions

----- ORBITALS FROM -RHFOP- Li atom Double Zeta Basis Set -----
$VEC
1  1 0.12257236E+00 0.88630102E+00 0.79684748E-02 0.49928299E-03
2  1 0.21106103E-01 0.13532553E+00 0.67402434E-01-0.10433066E+01
3  1 0.59592940E+00-0.19932147E+01 0.22001088E+01-0.70091995E+00
$END
```

```
NAMELIST $WALK   RUNFLG,NUMBLK,BLKTIM,TSTEP,RAN,
--------------   KONORM,KONMAX,KONMIN,ETRIAL,NSEMBL,ETW,DX,NCROSS

      RUNFLG(20) Run type (Default=0)
            0 ... Variational Monte Carlo
            1 ... Fixed-Node Diffusion Monte Carlo
            2 ... Released Node Green's function Monte Carlo
            3 ... Test Wavefunction by comparing analytic derivatives
                  with finite-difference derivatives.
           10 ... Variational Monte Carlo Optimization

      NUMBLK(20) Number of blocks in run (if <0 then run is performed, but
                 no output is generated...good for pre-randomization runs.)

      BLKTIM(20) Time for one block (In A.U.)

      TSTEP(20)  Time step (Average time step for GFMC) (In A.U.)

      KONORM     Initial size and size ensemble will be
                 renormalized to (Default=100)

      KONMAX     Maximum size before renormalization (Default=2*KONORM)

      KONMIN     Minimum size before renormalization (Default=KONORM/4)

      ETRIAL     Initial trial Energy (Default=NONE)
                    If no value supplied uses value in PUN file (IGUESS=1)
                    OR uses internal guess (IGUESS=-1) SEE $GUESS

      ETW        Trial Energy update weight (Default=0.5)
                    ETRIAL = (1 - ETW)*ETRIAL + ETW*EGRWTH

      DX(20)     Finite difference to use for each coordinate for
                    wavefunction test (Default = 1.0E-4)
                    IF DX < 0 THEN TEST SCALAR EVALUATIONS

      NCROSS     Released-Node maximum generations after node crossing
                 (NCROSS > 1)

      KFNMAX     GFMC Fixed-Node maximum ensemble size before renormalization
                 (default=3/2*KONORM)

      KFNMIN     GFMC Fixed-Node minimum ensemble size before renormalization
                 (default=2/3*KONORM)
```

```
DATA $WFN
---------

      Specification of M.O. occupancy for Single- and Multi-Determinant PsiT
      M.O.'s MUST be listed in the same order as in $VEC
      possible values are: DOC - DOUBLY OCCUPIED
                           ALP - ONE ELECTRON SPIN UP
                           BET - ONE ELECTRON SPIN DOWN
                           VAL - UNOCCUPIED MO

EXAMPLE: 4 M.O.'S, 6 Electrons, Triplet state
         where the 2nd & 3rd M.O.'s are nearly degenerate

$WFN
   0.950   DOC   DOC   ALP   ALP
   0.312   DOC   ALP   DOC   ALP
$END

      Note: The WVFN must strictly conform to the notation used.
            QMC will internally change the signs of the CI coefs
            since QMC uses a product of Slater determinants one for
            the up spins and one for the down spins.
            This is an artifact of the Fermi-Dirac statistics where
            the exchange of two operators in the 2nd quantization
            picture results in an overall change of signs of the
            wvfn.  The above case will have the effect:

                AB AB A  A  ->   -AAAABB

                AB A  AB A  ->   +AAAABB

            For a single-determinant (HFSCF) case you need not
            be concerned of the sign.

            The sum of the squared CI coefs should equal 1 ...
            if not QMC will tell you and renormalize them.
```

FILES

UNIT	PURPOSE	CTSS CRAY NAME	
1	INPUT PUNCH FILE	inxx	*must exist
2	OUTPUT PUNCH FILE	outxx	
5	MAIN INPUT FILE	cntrl	*must exist
6	MAIN OUTPUT FILE	qmcout	
7	OPTIMIZATION	iopt	
16	OUTPUT TAU, ENERGY, UNCERT.	idat	*must exist
4	CRAY TEXT TO SERIAL CONVERSION OF FOR005	tp04	*must exist

PROGRAM CONSIDERATIONS

The program is broken up into several subsections as follows:

Necessary Files:

```
   Q90MAIN.FOR  - main program and driver
   Q90RDXX.FOR  - configuration input/output routines
   Q90WALK.FOR  - Monte Carlo driver routines
   Q90GFMC.FOR  - Green's Function driver routines
   Q90WVFN.FOR  - wavefunction evaluation routines
   Q90LINE.FOR  - simple routines that can be inserted 'inline'
                  by the compiler to save 'call' overhead
```

Choose one of the following...or make your own:

```
   SYS90CD.VAX  - VAX VMS system dependent routines
   SYS90CD.CRA  - CTSS CRAY system dependent routines
   Q903OVL      - CTSS LDR overlay file
                   Compile: CFT77 i=q903cd,inline=q90line
                   Link:    LDR df=q903ovl
   If you make one of your own please send a documented copy to
   R.K.Owen so that it can be included in later releases.
```

Necessary for SYS90CD.VAX:

```
   RANF.FOR     - VAX VMS specific sequential congruential
                  pseudo-random number generator
```

General File for tailoring your own pseudo-random number generator:

```
   URAND.FOR    - machine independent sequential congruential
                  pseudo-random number generator given by
                  Forsythe, Malcolm, Moler "Computer Methods for
                  Mathematical Computations" which can be
                  tailored for your own system.

   QMC90CD.DOC  - This file

   Q90MAKE.COM  - VAX VMS DCL quasi-make/menu driver for the Q90 suite of
                  program files.
```

PROGRAM CONSIDERATIONS CONTINUED

Mass array storage in COMMON BLOCK /BIG/ can be reliably changed by globally changing the PARAMETER M0BIG, and M1BIG in each routine. They should be at least as large as NCORE0 and NCORE1, which is printed out at the beginning of the program.

AUXILIARY PROGRAMS

PUNCOLL.FOR	- ANSI FORTRAN-77 program to collect append PUNCH files (produced when IPRINT=0001000000) into a single PUN file.
PUNXX.FOR	- isolates each electron and outputs $QMCXX scatter plot data.
CDADD.FOR	- Statistically add UNIT16 data points.
CDFIT.FOR	- Determines best polynomial Least-Squares fit to collection of data points...generally useful for time step extrapolations of UNIT16 combined points.
CDSCALE.FOR	- Performs some simple math operations on data points.
CDSPLINE.FOR	- Finds cubic spline of data points and gives a new set of data points...interpolated from the original.
CDPACK.COM	- VAX VMS DCL front-end for the CD suite of programs.
LIST.FOR	- program prints out a nice listing of a FORTRAN program with page numbers, file line numbers, subroutine line numbers with subsection name at bottom of each page. prints out alphabetical table of contents at end.
ORBPLOT.FOR	- reads HONDO $BASIS data and produces a 1-electron density plot data.
ORBGRF.FOR	- reads 1-electron density plots, PUNXX $QMCXX scatter data, etc. and plots to screen or device.

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