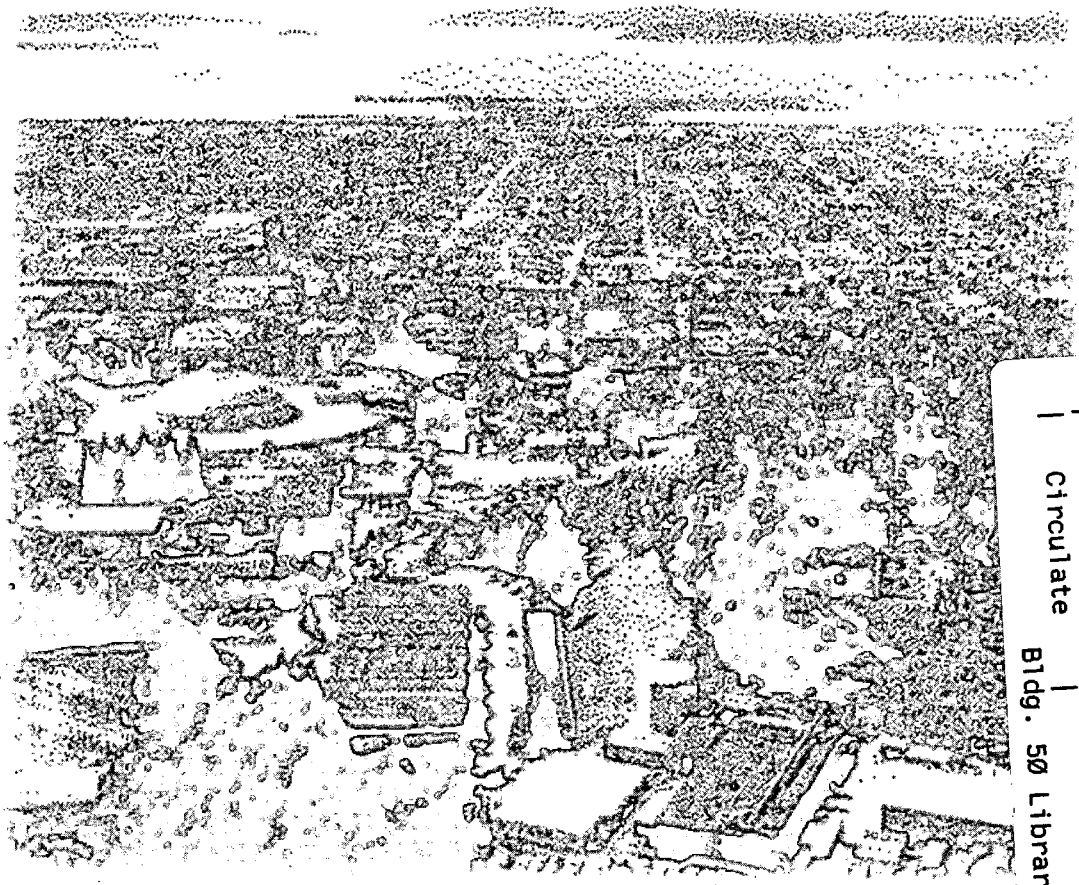
LBNL-39574
UC-401

ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

Quantum Monte Carlo for Vibrating Molecules

W.R. Brown
Chemical Sciences Division

August 1996
Ph.D. Thesis



DISCLAIMER

LBL-39574
UC-401

Quantum Monte Carlo for Vibrating Molecules

Willard Roger Brown
Ph.D. Thesis

Chemistry Department
University of California, Berkeley

and

Chemical Sciences Division
Ernest Orlando Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

August 1996

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 for Vibrating Molecules

by

Willard Roger Brown

Quantum Monte Carlo for Vibrating Molecules

by

Willard Roger Brown

B.S. (Michigan State University) 1990

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

GRADUATE DIVISION

of the

UNIVERSITY of CALIFORNIA, BERKELEY

Committee in charge:

Professor William A. Lester, Jr., Chair
Professor Herbert L. Strauss
Professor Steven G. Louie

1996

Abstract

Quantum Monte Carlo for Vibrating Molecules

by

Willard Roger Brown

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor William A. Lester, Jr., Chair

Quantum Monte Carlo (QMC) has successfully computed the total electronic energies of atoms and molecules. The main goal of this work is to use correlation function quantum Monte Carlo (CFQMC) to compute the vibrational state energies of molecules given a potential energy surface (PES). In CFQMC, an ensemble of random walkers simulate the diffusion and branching processes of the imaginary-time time dependent Schrödinger equation in order to evaluate the matrix elements $H_{mn} = \langle f_m | \widehat{\mathcal{H}} e^{-\widehat{\mathcal{H}}\tau} | f_n \rangle$ and $N_{mn} = \langle f_m | e^{-\widehat{\mathcal{H}}\tau} | f_n \rangle$ for a basis set of trial wavefunctions f_n. The solutions of the eigenvalue equation,

$$\sum_{m=1}^{M} [H_{mn}(\tau) - \Lambda_k(\tau) N_{mn}(\tau)] d_{km} = 0,$$

converge to the vibrational state energies of the molecule. Trial wavefunctions are constructed by linear combination of direct-product wavefunctions. Variational Monte Carlo (VMC) provides the basis function linear coefficients and a VMC energy optimization provides the non-linear basis function parameters.

The program QMCVIB was written to perform multi-state VMC and CFQMC calculations and employed for several calculations of the H_2O and C_3 vibrational states, using 7 PES's, 3 trial wavefunction forms, two methods of non-linear basis function parameter optimization, and on both serial and parallel computers. In order to construct accurate trial wavefunctions different wavefunctions forms were required

for H_2O and C_3. In order to construct accurate trial wavefunctions for C_3, the non-linear parameters were optimized with respect to the sum of the energies of several low-lying vibrational states. In order to stabilize the statistical error estimates for C_3 the Monte Carlo data was collected into blocks. Accurate vibrational state energies were computed using both serial and parallel QMCVIB programs. Comparison of vibrational state energies computed from the three C_3 PES's suggested that a non-linear equilibrium geometry PES is the most accurate and that discrete potential representations may be used to conveniently determine vibrational state energies.

The successful application of CFQMC to H_2O and C_3 molecules indicates the broad capability of quantum Monte Carlo methods and suggests that the utility of CFQMC and QMC methods will continue to grow in years to come.

Dedication

To Janet, my beloved wife, and Jesus, my Lord and my God.

Contents

List of Figures

List of Tables

Acknowledgments

Every good gift and perfect gift comes from above, and cometh down from the Father of lights, with whom is no variableness, neither shadow of turning.

James 2:17 (KJV)

According to the LORD'S mercy and loving-kindness, He has blessed me to write this thesis. Every faculty employed, every prospect explored, every help received, came from the LORD and the couriers of His grace have been many: my parents who always love me, my brothers, sisters and relatives who encourage me, the teachers at Marcus Whitman who taught me reading, writing and arithmetic, the faculty at Michigan State University who taught me chemistry, physics, mathematics and philosophy, the faculty at University of California Berkeley who taught me more chemistry, physics, and mathematics, Professor William Lester, my research advisor, who provided wise academic and scientific advice, Dr. Brian Hammond who introduced me to Quantum Monte Carlo, Dr. William Glauser whose collaboration began my research, the U.S. Department of Education which funded in part the first two years of my graduate studies, the Director, Office of Energy Research, Office of Basic Energy Sciences, Chemical Sciences Division, of the U.S. Department of Energy which supported this research in part under Contract No. DE-AC03-76SF00098, the National Energy Research Supercomputing Center (NERSC) which provided excellent computing support, Tim Pierce and Rich Zwakenburg at NERSC who helped optimize and port the CFQMC code for the CM-5, the Army High Performance Research Center (AHPRC) for a grant of computing time on their CM-5, Dr. Joy Andrews, Dr. Dario Bressinini, Prof. Ade Odutola, Dr. Maria Soto, and Dr. Zhiwei Sun. who helped me in my research, Dr. Robert Barnett and Dr. Carl Greeff whose thoughtful comments

refined this manuscript, David Anvar, Mark Moran and many dear friends who encouraged me to press on when all seemed dross, and my beloved wife Janet who is, GOD knows why, proud of me. I thank you all.

What part have I in this thesis? Look for mistakes and misconceptions. These are mine. All else is a good gift from above. LORD JESUS, I thank you for all.

Chapter 1

Molecules vibrate

For the invisible things of him from the creation of the world are clearly seen, being understood by the things that are made, even his eternal power and Godhead

Romans 1:20 (KJV)

Molecules vibrate and molecules rotate. The relative positions of atoms in a molecule and its orientation are not fixed. Indeed, they are not only dynamic, but also delocalized. They are both waves and particles. Therefore, the explanation and prediction of many chemical phenomena require a description of these molecular motions. The thermodynamics of bulk substances is to a significant degree an effect of molecular vibration and rotation because the partition function which describes the thermodynamic properties requires the vibrational and rotational energy levels of the molecule. The reactivity of chemicals in all phases is affected by the vibrations and rotations of the reactant and product molecules. The coupled channel equations that describe atom-molecule and molecule-molecule scattering, a microscopic description of chemical reactions, require the vibrational energies and wave functions of the molecules. Molecular spectroscopy is inseparable from molecular vibration and rotation. It is measurement of the properties of vibrating and rotating molecules. The value of the matrix element which describes the probability of a transition between two vibrational-electronic states of a molecule is greatly influenced by the Frank-Condon integral which is the overlap of the vibrational wave function of the two states.

The rotational and vibrational properties of many molecules, such as transi-

tion energies and intensities, may be measured spectroscopically, but the underlying physics requires the rotational and vibrational wave functions which may only be obtained by calculation. *Ab initio* and semi-empirical methods can also be used to compute some molecular properties which for practical reasons are not easily measured experimentally. Additionally, the experimentalist can use the calculated vibrational properties and wave function to predict band origins and line intensities, and assign the observed spectra. Therefore, *ab initio* and semi-empirical calculations of the vibrational and rotational properties of molecules are of great interest.

1.1 Vibrating molecule model

A vibrating and rotating molecule is composed of electrons and nuclei coupled by Coulomb interactions. Its properties can be obtained by solving the Schrödinger equation,

$$\widehat{\mathcal{H}}(\vec{\mathbf{R}},\vec{\mathbf{r}})\Psi(\vec{\mathbf{R}},\vec{\mathbf{r}}) = \mathcal{E}\Psi(\vec{\mathbf{R}},\vec{\mathbf{r}}). \tag{1.1}$$

The Schrödinger Hamiltonian $\widehat{\mathcal{H}}$, in atomic units ($\hbar = 1, e = 1$), is

$$\widehat{\mathcal{H}}(\vec{\mathbf{R}},\vec{\mathbf{r}}) = -\frac{1}{2}\sum_{\alpha=1}^{N} M_\alpha^{-1}\nabla^2_{\vec{\mathbf{R}}_\alpha} - \frac{1}{2}\sum_{i=1}^{n} m_i^{-1}\nabla^2_{\vec{\mathbf{r}}_i} + \sum_{i<j}^{n}\frac{1}{r_{ij}} + \sum_{\alpha<\beta}^{N}\frac{Z_\alpha Z_\beta}{r_{\alpha\beta}} - \sum_{\alpha=1}^{N}\sum_{i=1}^{n}\frac{Z_\alpha}{r_{\alpha i}}, \tag{1.2}$$

where $\vec{\mathbf{R}}_\alpha$ is the Cartesian coordinate vector for each of the N nuclei, $\vec{\mathbf{r}}_i$ is the Cartesian coordinate vector for each of the n electrons, $r_{\alpha\beta}$ is the separation between nuclei α and β, r_{ij} is the separation between electrons i and j, and $r_{\alpha i}$ is the separation between nucleus α and electron i. This model is rarely used to obtain molecular properties, but instead serves as a starting point for more approximate and widely used models.

The commonly used model of a vibrating molecule is composed of nuclei coupled by a smoothly varying, many-body potential. Its Hamiltonian is obtained from the Eq. (1.2) by application of the clamped nuclei approximation and the separation of the center of mass motion and vibrational-rotational motion[1].

1.1.1 Clamped nuclei approximation

In quantum mechanics the electrons and nuclei are explicitly coupled by a Coulomb potential. Therefore, the nuclear and electronic motion can be separated only approximately. The mass of the lightest nucleus is approximately 1820 times greater than the mass of an electron. The magnitude of the mass differences suggests that nuclei can be considered fixed relative to the motion of the electrons.[1] The behavior of the electrons in this model is described by the clamped nucleus Hamiltonian,

$$\widehat{\mathcal{H}}_{\rm el}(\vec{\mathbf{r}};\vec{\mathbf{R}}) = -\frac{1}{2}\sum_{i=1}^{n} m_i^{-1}\nabla_{\vec{\mathbf{r}}}^2 + \sum_{i<j}^{n}\frac{1}{r_{ij}} - \sum_{\alpha=1}^{N}\sum_{i=1}^{n}\frac{Z_\alpha}{r_{\alpha i}} + \sum_{\alpha<\beta}^{N}\frac{Z_\alpha Z_\beta}{r_{\alpha\beta}}, \tag{1.3}$$

in which the nuclear coordinates are parameters rather than variables. The solution to the Schrödinger equation for (1.3),

$$\widehat{\mathcal{H}}_{\rm el}(\vec{\mathbf{r}};\vec{\mathbf{R}})\Psi(\vec{\mathbf{r}};\vec{\mathbf{R}}) = \mathcal{E}_{\rm el}(\vec{\mathbf{R}})\Psi(\vec{\mathbf{r}};\vec{\mathbf{R}}), \tag{1.4}$$

is the focus of considerable research[2]. The eigenvalue of Eq. (1.4) $\mathcal{E}_{\rm el}(\vec{\mathbf{R}})$ is the total electronic energy of the molecule. It forms a smoothly varying, many-body potential energy which couples the nuclei. In this approximation the Hamiltonian,

$$\widehat{\mathcal{H}}_{\rm nuc}(\vec{\mathbf{R}}) = -\frac{1}{2}\sum_{\alpha=1}^{N} M_\alpha^{-1}\nabla_{\vec{\mathbf{R}}}^2 + \mathcal{E}_{\rm el}(\vec{\mathbf{R}}), \tag{1.5}$$

describes the vibration, rotation and translation of the molecule.

1.1.2 Center of mass separation

The the total electronic energy, $\mathcal{E}_{\rm el}(\vec{\mathbf{R}})$, is invariant under the uniform translations, $\mathbf{R}_i \rightarrow \mathbf{R}_i + \mathbf{a}$, and under orthogonal transformations, $\mathbf{R}_i \rightarrow \mathbf{S}\mathbf{R}_i$, where S is a 3×3 orthogonal matrix such that $\mathbf{S}^T\mathbf{S} = \mathbf{I}$, and $|\mathbf{S}| = \pm 1$. Therefore, $\mathcal{E}_{\rm el}(\vec{\mathbf{R}})$ can be transformed into the potential function $V(\vec{\mathbf{s}})$ where $\vec{\mathbf{s}}$ describes only relative positions of the nuclei. As a result, the Hamiltonian Eq. (1.5) can be separated into two parts; one that describes the translational dynamics and the other which describes the vibrational and rotational motion of the molecule.

[1] Born and Oppenheimer first proposed this approximation.[1].

The Hamiltonian describing the molecular translation is

$$\widehat{\mathcal{H}}_{\text{trans}}(\vec{\mathbf{Q}}) = -\frac{1}{2}\sum_{\alpha=1}^{N} M_T^{-1}\nabla^2_{\vec{\mathbf{Q}}}, \tag{1.6}$$

where M_T is the total mass of the molecule (or nuclei), and $\vec{\mathbf{Q}}$ is the center of mass coordinate vector of the molecule. The solution to the Schrödinger Eq. (1.6), the translational wave function, is the familiar plane wave,

$$\mathrm{T}(\vec{\mathbf{Q}}) = \exp(i\vec{\mathbf{k}}\cdot\vec{\mathbf{Q}}), \tag{1.7}$$

where $\vec{\mathbf{k}} \equiv (k_x, k_y, k_z)$.

The Hamiltonian describing the molecular vibration and rotation is

$$\widehat{\mathcal{H}}_{\text{ro-vib}}(\vec{\mathbf{q}}) = -\frac{1}{2}\sum_{i,j=1}^{N-1} \mu_{ij}^{-1}\nabla_{\vec{\mathbf{q}}_i}\cdot\nabla_{\vec{\mathbf{q}}_j} + V(\vec{\mathbf{s}}) \tag{1.8}$$

where μ is the generalized reduced mass, and $\vec{\mathbf{q}}$ is an internal-rotational coordinate vector.[2] The solutions to the Schrödinger equation for (1.8),

$$\widehat{\mathcal{H}}_{\text{ro-vib}}(\vec{\mathbf{q}})\Phi(\vec{\mathbf{q}}) = \mathcal{E}_{\text{ro-vib}}\Phi(\vec{\mathbf{q}}), \tag{1.9}$$

are the vibrational-rotational wave functions $\Phi(\vec{\mathbf{q}})$ and energies $\mathcal{E}_{\text{ro-vib}}$. The differences between the $\mathcal{E}_{\text{ro-vib}}$ form the vibrational-rotational spectrum which may also be measured experimentally.

1.2 Potential energy surface

The many-body potential $V(\vec{\mathbf{s}})$ which couples the atoms in the common model for the vibrating molecule forms a potential energy surface (PES). The harmonic or Morse PES's are model potentials meant to represent $V(\vec{\mathbf{s}})$ for diatomic molecules.[3] These model potentials are convenient because the Schrödinger Equations for the harmonic oscillator and for the Morse oscillator are exactly solved. For

[2] The vector $\vec{\mathbf{q}}$ is any set of internal coordinates. [1, 3].

[3] Harmonic and Morse Potential curves are one dimensional analogues of many body potential energy surfaces. A many-body potential is a hyper-surface in the $3N-6$ domain of atomic configuration space.

some diatomic molecules they accurately describe $V(\vec{s})$ near the equilibrium bond distance and may be used to obtain vibrational state energies. However, for other diatomic molecules and for polyatomic molecules, harmonic and Morse potentials poorly represent $V(\vec{s})$. More accurate many-body PES's are obtained from experimental measurement and from *ab initio* calculations[3].

1.2.1 Experimental potential energy surfaces

Conventionally, PES's have been obtained from spectroscopy in the region of the minimum. Force constants of diatomic molecules are obtained from spectroscopic parameters, ω_e, ω_e, χ_e, B_e, α_e, etc. of the expansion which relates the vibration-rotation energy levels to the quantum numbers ν and J. Force constants for polyatomic molecules may also be obtained from spectroscopic parameters, but the analysis is complicated by the multiple quantum numbers for the vibrational and rotational states[4]. The PES's are expressed as Taylor's series expansions in internal coordinates

$$V(\vec{s}) = \frac{1}{2!}\sum_i\sum_j f_{ij}s_is_j + \frac{1}{3!}\sum_i\sum_j\sum_k f_{ijk}s_is_js_k + \ldots, \qquad (1.10)$$

in which f are the force constants. This procedure is difficult and the convergence of the perturbation expansion, on which the method is based, is uncertain. Therefore, the PES's are accurate only in regions near minima.

The Sorbie-Murrell procedure may be implemented for triatomic molecules to construct a many-body PES which is constructed to be correct both in the region about the equilibrium configuration and at the dissociation limits[5]. The PES of a triatomic molecule may be represented as a sum of one-, two-, and three-body terms. The three-body interaction term in the many-body PES is deduced from the total potential and the one- and two-body terms. This assumes that the PES for the tri-atomic molecule ABC is already determined (as described above) and that the PES's for the three constituent diatomic molecules (AB, AC and BC) are known. The three-body interaction term is the product of a quartic polynomial that reproduces the force field PES and a range function that vanishes at the dissociation limits.

Both force field and Sorbie-Murrell PES's are useful specifically for molecules having a single minimum. Furthermore, experimental spectra may also be used to refine PES's. The PES parameters can be adjusted to have the calculated spectrum agree with the experimental spectrum. The quality of experimentally refined PES's relies on the accuracy of the experimental spectra, the quantity of experimental data to which the PES is fit, the flexibility of the analytic form of the PES and convergence of the optimization of the PES parameters.

1.2.2 *Ab initio* potential energy surfaces

In recent years *ab initio* electronic structure calculations have been used to produce accurate PES's for small molecules.[4] This is in part the result of the availability of more powerful methods of solving the Schrödinger equation for the clamped nucleus Hamiltonian and the availability of faster computers on which the calculations are performed.

Ab initio force field PES's may be constructed from calculated derivatives of $\mathcal{E}_{\rm el}(\vec{\rm s})$ with respect to $\vec{\rm s}$. The energy derivatives replace the experimental force constants in Eq. 1.10. The quality of the *ab initio* force field PES's relies on the accuracy, number and degree of calculated energy derivatives and on the choice of internal coordinates used in the analytical expression for $V(\vec{\rm s})$.

Ab initio PES's may also be constructed for molecules requiring a more global description, such as floppy molecules or molecules having multiple-minima in their PES. Typically, an approximate $\mathcal{E}_{\rm el}(\vec{\rm s}_{\vec{\rm R}})$ is obtained for a grid of points in $\vec{\rm s}$ distributed widely around the minimum of the total electronic energy. These points are fit to an analytical function of $\vec{\rm s}$. The quality of the PES is determined by the accuracy of the *ab initio* electronic structure calculation, the number and distribution of points at which $\mathcal{E}_{\rm el}(\vec{\rm s}_{\vec{\rm R}})$ is obtained, and the choice of the analytical form and fitting method.

The labor required for the construction of *ab initio* PES's for polyatomic

[4] Accurate *Ab initio* PES are available for several triatomic and a few tetra-atomic molecules. See, for example, references[6, 7, 8, 9, 10].

molecules increases significantly as the number of electrons and atoms increase. The computational and scientific burden of obtaining energies and energy derivatives is greater for molecules having more electrons. Additionally, the number of energy points required to adequately map the $3N-6$ dimensional PES's grows exponentially with the number of the atoms N. Therefore, there are presently few global *ab initio* PES's for tetra-atomic and larger molecules.

1.3 Solutions to the Schrödinger equation for vibrating molecules

Unlike the Morse and harmonic oscillators, molecules modeled by realistic PES's lack known exact solutions for the Schrödinger equation. This is a ubiquitous problem of quantum chemistry; the Schrödinger Equation is often written but rarely solved. Approximate solutions are sought by perturbative and variational methods. A complete explanation of these approaches is beyond the scope of this work. However, brief introductions for each follow.[5]

1.3.1 Perturbation theory

Perturbation theory begins with a known exact solution to the Schrödinger equation for an approximate Hamiltonian, called the *unperturbed or zeroth order Hamiltonian* and by successive corrections approaches the exact solution to the full Hamiltonian. A common unperturbed Hamiltonian for the vibrating-rotating molecule is,

$$\widehat{\mathcal{H}}_0 = \widehat{\mathcal{H}}_{\mathrm{RR}} + \widehat{\mathcal{H}}_{\mathrm{HO}}, \tag{1.11}$$

in which $\widehat{\mathcal{H}}_{\mathrm{RR}}$ is the Hamiltonian for a rigid rotor and $\widehat{\mathcal{H}}_{\mathrm{HO}}$ is the Hamiltonian for a harmonic oscillator. The unperturbed zeroth-order Hamiltonian neglects the rotational-vibrational coupling and all anharmonic terms in the PES. Solutions for

[5]For further discussion of the perturbation theory approach for vibrating molecules see [11] and for further discussion of the variational method for vibrating molecules see [3].

the ground and excited states of $\widehat{\mathcal{H}}_0$ are products of rigid-rotor and harmonic oscillator wave functions,

$$\widehat{\mathcal{H}}_0 \Phi_{RR} \Psi_{HO} = (\mathcal{E}_{\mathrm{RR}} + \mathcal{E}_{\mathrm{HO}}) \Phi_{RR} \Psi_{HO} \tag{1.12}$$

in which $\mathcal{E}_{\mathrm{RR}}$ is the energy of the rigid rotor and $\mathcal{E}_{\mathrm{HO}}$ is the energy of the harmonic oscillator. The anharmonicity of the potential and rotational-vibrational coupling are described to the first order by $\widehat{\mathcal{H}}_1$ and included as a perturbation to the zeroth-order Hamiltonian,

$$\widehat{\mathcal{H}}_{\mathrm{ro-vib}} \approx \widehat{\mathcal{H}}_0 + \widehat{\mathcal{H}}_1. \tag{1.13}$$

The first order corrections to the zeroth order energy are

$$\langle \Phi_{RR} \Psi_{HO} | \, \widehat{\mathcal{H}}_1 \, | \Phi_{RR} \Psi_{HO} \rangle \, . \tag{1.14}$$

The corrections to wave functions and higher order corrections to the energy may be made, but are increasingly complex.

In addition the series of corrections made by perturbation theory is not guaranteed to converge for all molecules and choices of unperturbed Hamiltonians and the convergence may be very slow with respect to the order of the correction. In particular, molecules having multiple minima or large anharmonicity in the PES, and floppy molecules are poorly described by the O^{th} order models and are unsuitable for perturbation theory approaches.

1.3.2 Variational method

The Rayleigh-Ritz variation method begins with a trial wave function for the vibrating molecule having variable parameters. The energy of the trial wave function is the expectation value of the Hamiltonian,

$$\mathcal{E}_{\mathrm{trial}} = \frac{\langle \Psi_T | \, \widehat{\mathcal{H}} \, | \Psi_T \rangle}{\langle \Psi_T \mid \Psi_T \rangle}, \tag{1.15}$$

where $\widehat{\mathcal{H}}$ is the Hamiltonian for the system and Ψ_T is the trial function. When the trial wave function is expanded in terms of the exact eigenstates of the Hamiltonian

Φ_i, the trial wave function energy is given by

$$\begin{aligned}
\mathcal{E}_{\text{trial}} &= \frac{\langle \sum_{i=0}^{\infty} c_i \Phi_i | \widehat{\mathcal{H}} | \sum_{j=0}^{\infty} c_j \Phi_j \rangle}{\langle \sum_{i=0}^{\infty} c_i \Phi_i \mid \sum_{j=0}^{\infty} c_j \Phi_j \rangle} \\
&= \frac{\sum_{i=0}^{\infty} c_i^2 \langle \Phi_i | \widehat{\mathcal{H}} | \Phi_i \rangle}{\sum_{i=0}^{\infty} c_i^2 \langle \Phi_i \mid \Phi_i \rangle} \\
&= \frac{\sum_{i=0}^{\infty} c_i^2 \mathcal{E}_i}{\sum_{i=0}^{\infty} c_i^2} \\
&= \sum_{i=0}^{\infty} C_i \mathcal{E}_i. \qquad (1.16)
\end{aligned}$$

Since all $C_i \left(\frac{c_i^2}{\sum_{i=0}^{\infty} c_i^2} \right)$ are non-negative and $\mathcal{E}_i \geq \mathcal{E}_0$ for all i, then

$$\mathcal{E}_0 \leq \sum_{i=0}^{\infty} C_i \mathcal{E}_i \qquad (1.17)$$

or

$$\mathcal{E}_0 \leq \mathcal{E}_{\text{trial}}. \qquad (1.18)$$

The energy expectation value of a trial wave function is always an upper bound to the exact energy of the ground state. As a result, minimization of the energy with respect to the wavefunction parameters yields the best possible approximation to the exact energy.

The general problem of minimizing the energy expectation value with respect to the variational parameters is difficult for trial wave function forms having many non-linear variational parameters. However, when only linear parameters are varied, the process of minimizing the trial wave function energy reduces to the problem of solving a set of linear algebraic equations. A trial function may be written as a linear combination of basis functions,

$$\Psi_T = \sum_{n=1}^{N} c_n f_n, \qquad (1.19)$$

where f_n are arbitrary known functions. The energy expectation value of the trial wave function is

$$\mathcal{E}_{\text{trial}} = \frac{\langle \sum_{i=1}^{N} c_i f_i | \widehat{\mathcal{H}} | \sum_{j=1}^{N} c_j f_j \rangle}{\langle \sum_{i=1}^{N} c_i f_i \mid \sum_{j=1}^{N} c_j f_j \rangle}$$

$$= \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} c_i c_j \langle f_i | \widehat{\mathcal{H}} | f_j \rangle}{\sum_{i=1}^{N}\sum_{j=1}^{N} c_i c_j \langle f_i \mid f_j \rangle}$$
$$= \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} c_i c_j H_{ij}}{\sum_{i=1}^{N}\sum_{j=1}^{N} c_i c_j N_{ij}}, \tag{1.20}$$

where $H_{ij} \equiv \langle f_i | \widehat{\mathcal{H}} | f_j \rangle$ is the *Hamiltonian matrix element* and $N_{ij} \equiv \langle f_i \mid f_j \rangle$ is the *overlap matrix element.* The minima may be found by differentiating the energy expectation value with respect to the linear variational parameters, *i.e.*,

$$\frac{\partial \mathcal{E}_{\text{trial}}}{\partial c_i}\left(\sum_{i=1}^{N}\sum_{j=1}^{N} c_i c_j N_{ij}\right) = \sum_{j=1}^{N}\left(2c_j H_{ij} - 2c_j N_{ij}\mathcal{E}_{\text{trial}}\right), \tag{1.21}$$

and solving for $\partial \mathcal{E}_{\text{trial}}/\partial c_i = 0$, thereby obtaining for each c_i the linear equation,

$$\sum_{j=1}^{N} 2c_j H_{ij} - 2c_j N_{ij}\mathcal{E}_{\text{trial}} = 0. \tag{1.22}$$

The linear equations 1.22 have non-trivial solutions only when the *secular determinant* vanishes,

$$\begin{vmatrix} H_{11} - \mathcal{E}_{\text{trial}}N_{11} & H_{12} - \mathcal{E}_{\text{trial}}N_{12} & \cdots & H_{1N} - \mathcal{E}_{\text{trial}}N_{1N} \\ H_{21} - \mathcal{E}_{\text{trial}}N_{21} & H_{22} - \mathcal{E}_{\text{trial}}N_{22} & \cdots & H_{2N} - \mathcal{E}_{\text{trial}}N_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1} - \mathcal{E}_{\text{trial}}N_{N1} & H_{N2} - \mathcal{E}_{\text{trial}}N_{N2} & \cdots & H_{NN} - \mathcal{E}_{\text{trial}}N_{NN} \end{vmatrix} = 0. \tag{1.23}$$

When the secular determinant is expanded it results in a polynomial of order N called the *secular equation.* The lowest energy root of the secular equation is an upper bound to the exact ground state energy, as shown in Eq. (1.18).

A corollary to the variational principle shows that the roots of successive calculations in which additional basis functions $f_{(N+j)}$ are added to the set of N basis functions, have equal or lower energy than the corresponding roots of the previous calculation,

$$\mathcal{E}_i^N \geq \mathcal{E}_i^{N+1} \geq \mathcal{E}_i^{N+2} \ldots \tag{1.24}$$

where the superscript indicates the successive calculations in terms of additional basis functions. From this observation, MacDonald demonstrated that the approximate

energies, $\mathcal{E}_i^{N+j}$, obtained from the Rayleigh-Ritz variational method (*i.e.*, the roots of the secular equation) are upper bounds to the exact energies $\mathcal{E}_i^{\text{exact}}$[12].

The upper bound limit is a strength and weakness of the Rayleigh-Ritz variation method. The accuracy of all the roots can be improved simply by enlarging the basis set. When the change in the approximate solutions (*i.e.* the roots of the secular equation) is small relative to the desired accuracy, then the variational calculation is considered converged for that vibrational state. However, the addition of any basis function may not lower the energy of any root. Therefore, it is unclear how many and which basis functions should be included. Furthermore, the higher energy roots typically are less well converged than the lower energy roots. The number of converged states (*i.e.* states having accurate energies measured by the change in energy with respect to the additional basis function) is only a small fraction of the N basis functions used in the calculation.

Variational methods can be applied to a wider range of molecules than perturbation methods. Nevertheless, the number of molecules for which ro-vibrational states have been computed is small. Recently the variational approach was used to compute the ro-vibrational states of the floppy triatomic molecule C_3 and the intermolecular vibrational modes of water dimer, $(H_2O)_2$. The variational approach has been used to compute the vibrational states of CH_3O and NH_3. However, the vibrational spectra of only a few tetra-atomic molecules have been computed variationally. No spectra have been computed variationally for molecules having more than four atoms that have included all vibrational modes.[6] In contrast, variational methods for electronic structure have been applied to atoms and molecules having hundreds of electrons. The objective of vibrational calculations includes the computations of excited states, whereas the objective of electronic structure calculations is the computation of single states. Nevertheless, the relatively small size of the molecules for which vibrational spectra have been computed indicates the difficulty of computing the vibrational states of diverse molecules. No coordinate systems or basis function forms are universally applicable as in electronic structure calculations. Therefore,

[6]Quantum Monte Carlo has been applied to 12-dimensional $(H_2O)_2$.

the kinetic energy operator has no universally applicable form; it is derived for each molecule and coordinate system. As a result matrix element integrals must be solved for each molecule and coordinate system. The complexity of these issues increases with the number of atoms in the molecule. In addition, the scarcity of accurate *ab initio* potential energy surfaces limits the number of molecules for which spectra may be computed. The need for methodological development is clear.

1.3.3 Recent developments

The methodology of computation of vibrational energies of molecules has received intense activity in recent years. New perturbation theory and variational approaches attempt to reduce the scale of the problem by introducing new basis functions or representations. Among these approaches are canonical Van Vleck perturbation theory in a super operator framework [13], Morse oscillator rigid bender internal dynamics (MORBID) [14], distributed Gaussian basis (DGB) [15], discrete variable representations (DVR) [16], the collocation method [17], a hybrid of the discrete variable and finite basis representations within the Lanczos approach (DVR/FBR Lanczos) [18], and the approach employed here, correlation function quantum Monte Carlo (CFQMC)[19, 20]. Sibert has used canonical Van Vleck perturbation theory to calculate the vibrational states of H_2O and H_2CO [21]. Jensen et al. have used MORBID to calculate vibrational states of C_3 [22, 7]. DGB-DVR has been used by Bačić and Light to calculate the vibrational states of LiCN/CNLi and HCN/CNH [23] and by Mladenovic et al. to calculate the vibrational states of C_3 [24]. Cohen and Saykally have used the collocation method to refine an intermolecular potential energy surface to the experimentally determined vibrational states of Ar-H_2O [25]. Recently, Bramley et al. introduced a hybrid DVR/FBR Lanczos method and calculated the band origins of C_3 up to $35,000\,\mathrm{cm}^{-1}$ with great efficiency [18]. What remains to be seen of these methods is their extensibility to larger polyatomic molecules.

Chapter 2

Quantum Monte Carlo methods for vibrating molecules

O the depth of the riches both of the wisdom and knowledge of God! how unsearchable are his judgments, and his ways past finding out!
Romans 11:33 (KJV)

Variational Monte Carlo (VMC) and Diffusion Monte Carlo (DMC) methods are used to solve the Clamped Nuclei Schrödinger equation for atoms and molecules, Eq. (1.3) [26, 27, 28]. In the present study, both methods are used to solve the vibrating molecule problem posed in Eq. (1.9). The program QMCVIB, written by, Dr. William Glauser and Willard Brown, employs the multi-state VMC to compute ground and excited state energies and wavefunctions, and Correlation Function quantum Monte Carlo (CFQMC) method, a diffusion Monte Carlo method, to compute very accurate ground and excited state energies.

2.1 Variational Monte Carlo

The VMC method applies Monte Carlo integration to evaluate the expectation value of a single wave function ψ,

$$\mathcal{E}_f = \frac{\langle \psi | \widehat{\mathcal{H}} | \psi \rangle}{\langle \psi \mid \psi \rangle}, \tag{2.1}$$

and the matrix elements of basis functions f_m, i.e., $\psi = \sum_m c_m f_m$,

$$\begin{aligned} H_{mn} &= \langle f_m | \widehat{\mathcal{H}} | f_n \rangle \\ N_{mn} &= \langle f_m \mid f_n \rangle . \end{aligned} \tag{2.2}$$

Monte Carlo integration is robust. Expectation values and matrix elements for any ψ or f_m may be evaluated provided that the first and second derivatives of ψ or f_m, respectively, are defined and have a finite variance. Trial wavefunctions are freely tailored to describe the molecule. VMC is a single or multiple state method. Single states are obtained by Eq. (2.1). Multiple states are obtained by evaluating the matrix elements, Eq. (2.2) for a set basis of functions and applying the Rayleigh-Ritz variational method.

2.1.1 Single state variational Monte Carlo

The operator expectation value, Eq. (2.1), may be expressed as an average over a probability density, that is,

$$\mathcal{E}_f = \int P_\psi(\vec{\mathrm{R}}) \mathcal{E}_{\mathcal{L}_\psi}(\vec{\mathrm{R}}) \mathrm{d}\vec{\mathrm{R}} \tag{2.3}$$

where $\mathcal{E}_{\mathcal{L}_\psi} \equiv \widehat{\mathcal{H}}\psi/\psi$ is the local energy and $P_\psi(\vec{\mathrm{R}}) \equiv \frac{\psi^2(\vec{\mathrm{R}})}{\int \psi^2(\vec{\mathrm{R}}) \mathrm{d}\vec{\mathrm{R}}}$ [27]. The Monte Carlo estimate of the expectation value, Eq. (2.1), is obtained by drawing sample points $\vec{\mathrm{R}}_i$ from the distribution $P_\psi(\vec{\mathrm{R}})$, is

$$\mathcal{E}_\psi = \left\langle \mathcal{E}_{\mathcal{L}_\psi} \right\rangle = \lim_{M \to \infty} \frac{1}{M} \sum_{i=1}^{M} \mathcal{E}_{\mathcal{L}_\psi}(\vec{\mathrm{R}}_i). \tag{2.4}$$

Each of the $\vec{\mathrm{R}}$ is a point in 3N-dimensional configuration space, specifying the coordinates of each atom in the molecules, and are often called *walkers*. The variance of the Monte Carlo estimate is

$$\sigma^2(\mathcal{E}_i) = \frac{\left\langle \mathcal{E}^2_{\mathcal{L}_\psi} \right\rangle - \left\langle \mathcal{E}_{\mathcal{L}_\psi} \right\rangle^2}{M - 1}. \tag{2.5}$$

when the sampling is uncorrelated.[1]

[1] The uncorrelated approximation for the variance is unneccessary. Further discusion of variance estimates for single state VMC may be found in [27].

The sample points $\vec{\mathbf{R}}_i$, distributed as $P_\psi(\vec{\mathbf{R}})$, are obtained by the Metropolis method. In the approach employed here, the walkers are moved by steps according to

$$\vec{\mathbf{R}}(\tau+\delta\tau) = \vec{\mathbf{R}}(\tau) + D\vec{\mathbf{F}}(\vec{\mathbf{R}}) + \vec{\chi}. \tag{2.6}$$

In Eq. (2.6) $D = 1/2m$ is the diffusion constant,[2] $\vec{\mathbf{F}}(\vec{\mathbf{R}}) \equiv (2\nabla\psi/\psi)$ is called the quantum force, $\vec{\chi}$ is a Gaussian random variable with a mean value of zero and a variance of $2D\delta\tau$, and $\delta\tau$ is the time-step of the random walk. The moves are accepted according to the distribution,

$$A\left(\vec{\mathbf{R}}(\tau+\delta\tau), \vec{\mathbf{R}}(\tau)\right) = \min\left[1, q\left(\vec{\mathbf{R}}(\tau+\delta\tau), \vec{\mathbf{R}}(\tau)\right)\right], \tag{2.7}$$

where

$$q\left(\vec{\mathbf{R}}(\tau+\delta\tau), \vec{\mathbf{R}}(\tau)\right) = \frac{\psi^2\left(\vec{\mathbf{R}}(\tau+\delta\tau)\right) G\left(\vec{\mathbf{R}}(\tau), \vec{\mathbf{R}}(\tau+\delta\tau); \delta\tau\right)}{\psi^2\left(\vec{\mathbf{R}}(\tau)\right) G\left(\vec{\mathbf{R}}(\tau+\delta\tau), \vec{\mathbf{R}}(\tau); \delta\tau\right)} \tag{2.8}$$

and $G(\vec{\mathbf{R}}(\tau+\delta\tau), \vec{\mathbf{R}}(\tau); \delta\tau)$ is the probability distribution function for $\vec{\mathbf{R}}(\tau)$ conditional on $\vec{\mathbf{R}}(\tau+\delta\tau)$. The probability distribution G is chosen to satisfy the Fokker-Planck equation and has the form

$$G\left(\vec{\mathbf{R}}(\tau+\delta\tau), \vec{\mathbf{R}}(\tau); \delta\tau\right) = \exp\left(-\frac{\left[\vec{\mathbf{R}}(\tau+\delta\tau) - \vec{\mathbf{R}}(\tau) - \vec{\mathbf{F}}(\vec{\mathbf{R}})\right]^2}{4D\delta\tau}\right). \tag{2.9}$$

The resulting walk samples from ψ^2 with reasonable efficiency and is shown in figure 2.1. At each step of the walk the local energy is evaluated for each walker. The average local energy of all points sampled is the Monte Carlo estimate of the energy expectation value of ψ, Eq. (2.4).

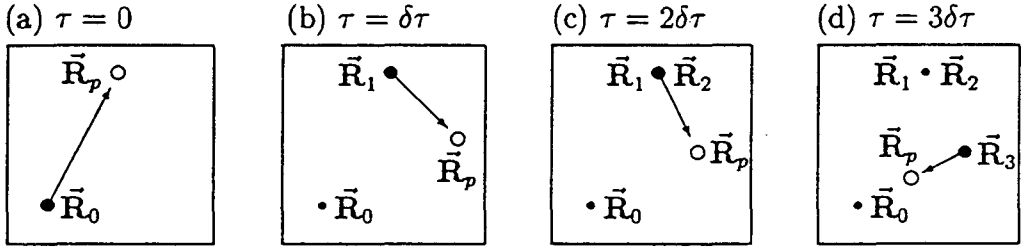
2.1.2 Multi-state variational Monte Carlo

The Hamiltonian and overlap matrix elements of Eq. (2.2) may be expressed as weighted averages,

$$H_{mn} = \int \psi_G^2 \left(\frac{f_m}{\psi_G}\right)\left(\frac{\widehat{\mathcal{H}} f_n}{\psi_G}\right) \mathrm{d}\vec{\mathbf{R}} \tag{2.10}$$

[2] In general the diffusion constant differs for atoms of inequivalent masses. Here the notationally simpler equivalent mass case is presented. The general case is implemented in the program QMCVIB.

Figure 2.1: Metropolis Random Walk in 2-dimensions: (a) at time $\tau = 0$, point $\vec{\mathbf{R}}_p$ is proposed, (b) at time $\tau = \delta\tau$, point $\vec{\mathbf{R}}_p$ is accepted as $\vec{\mathbf{R}}_1$ and a new point $\vec{\mathbf{R}}_p$ is proposed, (c) at time $\tau = 2\delta\tau$, $\vec{\mathbf{R}}_p$ is rejected and $\vec{\mathbf{R}}_1$ is retained as $\vec{\mathbf{R}}_2$; a new point $\vec{\mathbf{R}}_p$ is proposed, and (d) at time $\tau = 3\delta\tau$, point $\vec{\mathbf{R}}_p$ is accepted as $\vec{\mathbf{R}}_3$ and the walk continues as a new point $\vec{\mathbf{R}}_p$ is proposed.



and

$$S_{mn} = \int \psi_G^2 \left(\frac{f_m}{\psi_G}\right)\left(\frac{f_n}{\psi_G}\right) \mathrm{d}\vec{\mathbf{R}} \tag{2.11}$$

where the square of the guiding function ψ_G^2 is the weight [20]. The Monte Carlo estimate of these matrix elements is,

$$H_{mn} = \lim_{M\to\infty} \frac{1}{M} \sum_{i=1}^{M} \left(\frac{f_m}{\psi_G}\right)\left(\frac{f_n}{\psi_G}\right) \frac{1}{2}(\mathcal{E}_{L_m}(\vec{\mathbf{R}}_i) + \mathcal{E}_{L_n}(\vec{\mathbf{R}}_i)) \tag{2.12}$$

and

$$S_{mn} = \lim_{M\to\infty} \frac{1}{M} \sum_{i=1}^{M} \left(\frac{f_m}{\psi_G}\right)\left(\frac{f_n}{\psi_G}\right). \tag{2.13}$$

They are obtained by drawing sample points $\vec{\mathbf{R}}_i$ from the distribution $\psi_G^2(\vec{\mathbf{R}})$.[3] The sample points $\vec{\mathbf{R}}_i$ are generated by the Metropolis method, as for the single state VMC described above. The guiding function ψ_G replaces the trial wave functions in directing the Monte Carlo random walk. The guiding function differs from a trial wavefunction because it is chosen to have a large overlap with all the trial wavefunctions, whereas trial wavefunctions are constructed to approximate a particular vibrational state. The large overlap between the guiding function and each of the

[3]For clarity the summation in Eq. 2.12 and 2.13 is over sample points rather than random walkers and steps. The M sample points may be obtained from one random walker with M steps or with an ensemble of N random walkers with M/N steps each.

trial wavefunctions allows the simultaneous evaluation of all matrix elements from a single random walk.

The solution of the eigenvalue equation, the so called secular equations,

$$\sum_{i=1}^{m} [H_{mn} - \Lambda_k N_{mn}] \, d_{km} = 0, \tag{2.14}$$

provides eigenvalues Λ_k and eigenvectors $\vec{d_k}$ which are approximate vibrational state energies and wavefunctions, respectively.[4]

In order to measure errors resulting from the stochastic evaluation of the Hamiltonian and overlap matrix elements, the ensemble of N walkers is subdivided into blocks of walkers. Independent Monte Carlo evaluations of the Hamiltonian and overlap matrix elements are obtained from each block of walkers. These are averaged to obtain the ensemble estimate of the matrix elements. The secular equations are solved using matrix elements obtained from each block of walkers and the ensemble average. The eigenvalues obtained from the blocks are λ_k, and the eigenvalues obtained from the ensemble average are Λ_k. The statistical error of the eigenvalues is measured by

$$\sigma = (\overline{\lambda_k^2} - \overline{\lambda_k}^2)^{1/2} \tag{2.15}$$

where $\overline{\lambda_k}$ is the average of λ_k and $\overline{\lambda_k^2}$ is the average of λ_k^2. The non-linearity of Eq. (2.14) introduces an asymmetric bias into the λ_k. The biases in Λ_k are measured by

$$\tilde{b}_k = \Lambda_k - \overline{\lambda}. \tag{2.16}$$

As the size of the blocks approaches the size of the ensemble, $\tilde{b}_k \rightarrow 0$ is roughly $m^{-1/2}$ where m is the blocksize. For a large ensemble with many blocks, $\tilde{b}_k$ provides an upper bound on the remaining bias in Λ_k. As $M \rightarrow \infty$ the statistical error and bias both vanish. Therefore, Monte Carlo calculations can obtain arbitrary precision, notwithstanding limitations of computational resources. Vibrational state energies computed by VMC will be presented in the following chapter.

Tailoring of the basis functions for multiple states of a given molecule is an enormous task. Basis functions optimized for the ground state may not be optimal for

[4]QMCVIB employs a linear algebra routine developed by B. S. Garbow at Argonne National Laboratory. The package solves the real symmetric generalized eigenproblem $ax - \lambda bx = 0$.

excited states, for example, basis functions which contribute to bending states may be insignificant for stretching states. Basis functions suitable for linear molecules may be unsuitable for bent or tetrahedral molecules, etc. In order to satisfy the diversity of vibrational states, large expansions of basis functions are usually employed resulting in dramatic N_{basis}^2 growth in the number of matrix elements evaluated. The likelihood of near linear dependency increases with the number of basis functions while the efficiency of precisely estimating the matrix elements decreases with basis set size because the guiding function must encompass increasingly larger domains of configuration space. As a result, the likelihood of an apparent linear dependency causing the overlap matrix to be singular increases with basis set size. Therefore, the basis set cannot be increased arbitrarily and the calculations can not be made arbitrarily accurate by basis set exapnsion alone. Improved basis functions must be constructed. Despite these difficulties, VMC has great utility in constructing trial wavefunctions for DMC calculations which can, in turn, obtain very accurate results.

2.2 Diffusion Monte Carlo

Green's function Monte Carlo methods belong to two general formalisms, time-dependent and time independent. Time-dependent Green's function methods solve the time independent Schrödinger equation by solving an imaginary-time time-dependent Schrödinger equation. DMC and CFQMC, the multi-state generalization of DMC that is employed in QMCVIB, are based on the time-dependent formalism. The DMC and CFQMC methods are presented in this section. Time-independent Green's function methods solve the integral form of the time independent Schrödinger equation. These methods are labeled Green's function Monte Carlo (GFMC). Presentation of GFMC methods is available elsewhere [27].

Diffusion Monte Carlo begins with the time-dependent Schrödinger equation,

$$i\frac{\partial\phi(\vec{\mathbf{R}},t)}{\partial t} = \left(\widehat{\mathcal{H}} - \mathcal{E}_T\right)\phi(\vec{\mathbf{R}},t), \tag{2.17}$$

where $\mathcal{E}_T$ is an arbitrary energy shift and $\hbar = 1$ in atomic units, has the formal

solution,

$$\phi(\vec{\mathbf{R}},t)=\sum_{k=0}^{\infty}C_k\Phi_k(\vec{\mathbf{R}})e^{-i(\mathcal{E}_k-\mathcal{E}_T)t}, \tag{2.18}$$

where $C_k = \langle\Phi_k \mid \phi(t=0)\rangle$, possesses oscillatory behavior. When *imaginary time*, $\tau = it$, is substituted into Eqs. (2.17) and (2.18),

$$-\frac{\partial\phi(\vec{\mathbf{R}},\tau)}{\partial\tau}=-\left(\widehat{\mathcal{H}}-\mathcal{E}_T\right)\phi(\vec{\mathbf{R}},\tau), \tag{2.19}$$

and

$$\phi(\vec{\mathbf{R}},\delta\tau)=\sum_{k=0}^{\infty}C_k\Phi_k(\vec{\mathbf{R}})e^{-(\mathcal{E}_k-\mathcal{E}_T)\delta\tau}. \tag{2.20}$$

Now the time behavior is exponential and in the limit $\tau \to \infty$ the lowest energy eigenstate, the ground state dominates.

2.2.1 Diffusion-Rate analogy

Writing the Hamiltonian explicitly in Eq. (2.19) yields

$$\frac{\partial\phi(\vec{\mathbf{R}},\tau)}{\partial\tau}=-\frac{1}{2}\sum_{\alpha=1}^{N}M_\alpha^{-1}\nabla^2_{\vec{\mathbf{R}}_\alpha}\phi(\vec{\mathbf{R}},\tau)+(V(\vec{\mathbf{s}})-\mathcal{E}_T)\,\phi(\vec{\mathbf{R}},\tau). \tag{2.21}$$

This equation contains the processes of classical diffusion,

$$\frac{\partial C(x,y,z;t)}{\partial t}=D\nabla^2C(x,y,z;t), \tag{2.22}$$

and growth or decay,

$$\frac{\partial C(x,y,z;t)}{\partial t}=k(x,y,z)C(x,y,z;t). \tag{2.23}$$

By analogy with the classical processes, $C = \phi$ is the concentration distribution, $D = 1/2m$ is the diffusion constant, and $k = (V(\vec{\mathbf{s}}) - \mathcal{E}_T)$ is a position-dependent rate constant [29]. Since the classical diffusion equations and first order rate equations can individually be solved by simulation, it is expected that solutions for the imaginary-time time-dependent Schrödinger equation can also be obtained by simulation.

To obtain the time evolution as described by Eq. (2.19) the imaginary-time time-dependent Schrödinger equation is expressed in integral form,

$$\phi(\vec{\mathbf{R}}_2,\tau_2)=\int G(\vec{\mathbf{R}}_2,\tau_2;\vec{\mathbf{R}}_1;\tau_1)\phi(\vec{\mathbf{R}}_1,\tau_1)\mathrm{d}\vec{\mathbf{R}}, \tag{2.24}$$

where $G(\vec{\mathbf{R}}_2, \tau_2; \vec{\mathbf{R}}_1; \tau_1)$ is the Green's function. Information about $G(\vec{\mathbf{R}}_2, \tau_2; \vec{\mathbf{R}}_1; \tau_1)$ may be obtained from Eqs. (2.19) and (2.24) by operating on both sides of Eq. (2.24) with $\widehat{\mathcal{H}} - \mathcal{E}_T$,

$$\left(\widehat{\mathcal{H}} - \mathcal{E}_T\right) \phi(\vec{\mathbf{R}}_2, \tau_2) = \int \left(\widehat{\mathcal{H}} - \mathcal{E}_T\right) G(\vec{\mathbf{R}}_2, \tau_2; \vec{\mathbf{R}}_1; \tau_1) \phi(\vec{\mathbf{R}}_1, \tau_1) \mathrm{d}\vec{\mathbf{R}}, \tag{2.25}$$

and differentiating both sides of Eq. (2.24) with respect to τ_2,

$$\frac{\partial \phi(\vec{\mathbf{R}}_2, \tau_2)}{\partial \tau_2} = \int \frac{\partial G(\vec{\mathbf{R}}_2, \tau_2; \vec{\mathbf{R}}_1; \tau_1)}{\partial \tau_2} \phi(\vec{\mathbf{R}}_1, \tau_1) \mathrm{d}\vec{\mathbf{R}}. \tag{2.26}$$

The righthand sides of Eqs. (2.25) and (2.26) may be equated using Eq. (2.19), yielding,

$$\frac{\partial G(\vec{\mathbf{R}}_2, \tau_2; \vec{\mathbf{R}}_1; \tau_1)}{\partial \tau_2} = (\widehat{\mathcal{H}} - \mathcal{E}_T) G(\vec{\mathbf{R}}_2, \tau_2; \vec{\mathbf{R}}_1; \tau_1). \tag{2.27}$$

and demonstrating that the desired Green's frunction, like $\phi(\vec{\mathbf{R}}, \tau)$ satisfies the Schrödinger equation. Furthermore, from Eq. (2.19),

$$|\phi(\tau_2)\rangle = e^{-(\widehat{\mathcal{H}} - \mathcal{E}_T)(\tau_2 - \tau_1)} |\phi(\tau_1)\rangle . \tag{2.28}$$

Inserting a complete set of positions $\left|\vec{\mathbf{R}}_1\right\rangle \left\langle\vec{\mathbf{R}}_1\right|$ between the time evolution operator and the ket, and multiplying on the left by $\left\langle\vec{\mathbf{R}}_2\right|$ yields,

$$\phi(\vec{\mathbf{R}}_2, \tau_2) = \int \left\langle\vec{\mathbf{R}}_2\right| e^{-(\widehat{\mathcal{H}} - \mathcal{E}_T)(\tau_2 - \tau_1)} \left|\vec{\mathbf{R}}_1\right\rangle \phi(\vec{\mathbf{R}}_1, \tau_1) \mathrm{d}\Phi\vec{\mathbf{R}}. \tag{2.29}$$

Equating the right sides of Eqs. (2.24) and (2.29) shows that,

$$G(\vec{\mathbf{R}}_2, \tau_2; \vec{\mathbf{R}}_1; \tau_1) = \left\langle\vec{\mathbf{R}}_2\right| e^{-(\widehat{\mathcal{H}} - \mathcal{E}_T)(\tau_2 - \tau_1)} \left|\vec{\mathbf{R}}_1\right\rangle, \tag{2.30}$$

that is, the Green's function depends only on the difference $\tau_2 - \tau_1 \equiv \delta\tau$ with respect to imaginary time. Therefore, the integral form of the Schrödinger equation, Eq. (2.24), may be written,

$$\phi(\vec{\mathbf{R}}_2, \tau_2) = \int G(\vec{\mathbf{R}}_2, \vec{\mathbf{R}}_1; \delta\tau) \phi(\vec{\mathbf{R}}_1, \tau_1) \mathrm{d}\vec{\mathbf{R}}. \tag{2.31}$$

and solved iteratively.

2.2.2 Short-time approximation

Although the exact Green's function is unknown,[5] the fact that Eq. (2.21) may be represented as a diffusion and first-order rate process suggests that an approximate, yet analytical, Green's function may be constructed from the Green's functions of the classical diffusion and first order rate equations. Such a Green's function may be derived from the Trotter approximation; that is, in the *short-time approximation* (STA), the Green's function for small $\delta\tau$ is given by,

$$G_{\text{STA}} \equiv G_{\text{diff}} G_{\text{Rate}} = e^{\hat{T}\tau} e^{(\hat{V}-\varepsilon_T)\tau} \approx e^{-(\hat{T}+\hat{V}-\varepsilon_T)\tau}, \tag{2.32}$$

where $\hat{T}$ is the kinetic energy operator and $\hat{V}$ is the PES. Since the operators, $\hat{T}$ and $\hat{V}$, do not commute, G_{STA} is exact only in the limit $\delta\tau \to 0$[30, 31, 32, 33, 34]. The first correction term is,

$$G - G_{\text{STA}} = \frac{1}{2}\left[\hat{V}, \hat{T}\right](\delta\tau)^2 + O\left((\delta\tau)^3\right). \tag{2.33}$$

The long time limit $\tau \to \infty$, $\Phi \propto \phi_0$ can now be obtained by applying G_{STA} repeatedly using small timesteps $\delta\tau$.

The Green's function of the classical diffusion equation is[35]

$$G_{\text{diff}}(\vec{\mathbf{R}}_2, \vec{\mathbf{R}}_1; \delta\tau) = (4\pi D\delta\tau)^{3N/2} e^{-(\vec{\mathbf{R}}_2 - \vec{\mathbf{R}}_1)^2/4D\delta\tau}. \tag{2.34}$$

Diffusion is simulated by propagating an ensemble of walkers stepwise according to the dynamical equation,

$$\vec{\mathbf{R}}(\tau + \delta\tau) = \vec{\mathbf{R}}(\tau) + \vec{\chi}, \tag{2.35}$$

where $D = 1/2m$ is the diffusion constant, $\vec{\chi}$ is a Gaussian random variable with a mean value of zero and a variance of $2D\delta\tau$, and $\delta\tau$ is the time-step of the random walk. The Green's function of the first-order rate equation is

$$G_{\text{rate}}(\vec{\mathbf{R}}_2, \vec{\mathbf{R}}_1; \delta\tau) = e^{-\left(\frac{1}{2}\left[V(\vec{\mathbf{R}}_1)+V(\vec{\mathbf{R}}_2)\right]-\varepsilon_T\right)\delta\tau}. \tag{2.36}$$

The rate process may be simulated by a the creation or destruction of walkers depending on whether G_{rate} is greater or less that 1. The creation of walkers in regions

[5] This restates the ubiquitous problem of quantum chemistry; the Schrödinger Equation is often written but rarely solved exactly.

of low potential energy and destruction of walkers in regions of high potential energy gives shape to the Monte Carlo distribution. The combined effect of diffusion and rate processes produces the structure of the wavefunction.

2.2.3 Importance sampling

As in VMC importance sampling is used in DMC to increase the efficiency of the random walk. An analytical guiding function ψ_G is used to bias the random walk to produce the distribution $f(\vec{\mathbf{R}},\tau) \equiv \psi_G(\vec{\mathbf{R}})\phi_0(\vec{\mathbf{R}},\tau)$ rather than $\phi_0(\vec{\mathbf{R}},\tau)$. Importance sampling is introduced by multiplying Eq. (2.21) by ψ_G,

$$\frac{\partial f(\vec{\mathbf{R}},\tau)}{\partial \tau} = D\left[\nabla^2_{\vec{\mathbf{R}}} f(\vec{\mathbf{R}},\tau) - \nabla\cdot\left(f(\vec{\mathbf{R}},\tau)\vec{F}_Q(\vec{\mathbf{R}})\right)\right] + (\mathcal{E}_T - \mathcal{E}_{\mathcal{L}})f(\vec{\mathbf{R}},\tau), \qquad (2.37)$$

where $\vec{F}_Q \equiv \nabla \ln|\psi_G|^2 = 2\nabla\psi_G/\psi_G$ is a vector field called the *quantum force*. It directs walkers away from regions where ψ_G^2 is small.

Importance sampling changes the appearance of the Schrödinger equation. Nevertheless, it separates into rate terms, $(\mathcal{E}_T - \mathcal{E}_{\mathcal{L}})f(\vec{\mathbf{R}},\tau)$ and diffusion terms $D\left[\nabla^2_{\vec{\mathbf{R}}} f(\vec{\mathbf{R}},\tau) - \nabla\cdot\left(f(\vec{\mathbf{R}},\tau)\vec{F}_Q(\vec{\mathbf{R}})\right)\right]$. A Green's function may be constructed for each. The rate term is

$$\tilde{G}_{\text{rate}}(\vec{\mathbf{R}}_2,\vec{\mathbf{R}}_1;\delta\tau) = e^{-\left(\frac{1}{2}\left[\mathcal{E}_{\mathcal{L}}(\vec{\mathbf{R}}_1)+\mathcal{E}_{\mathcal{L}}(\vec{\mathbf{R}}_2)\right]-\mathcal{E}_T\right)\delta\tau}, \qquad (2.38)$$

and the diffusion term is

$$\tilde{G}_{\text{diff}}(\vec{\mathbf{R}}_2,\vec{\mathbf{R}}_1;\delta\tau) = e^{-\left(D\nabla^2+\nabla\cdot\vec{F}_Q+\vec{F}_Q\cdot\nabla\right)\delta\tau}. \qquad (2.39)$$

The quantum force $\vec{F}_Q$ is a function of $\vec{\mathbf{R}}$ and a functional of ψ_G, rendering an exact solution of the new diffusion equation unique for each guiding function. However, by assuming that $\vec{F}_Q$ is constant over the move from $\vec{\mathbf{R}}_1$ to $\vec{\mathbf{R}}_2$, the solution

$$\tilde{G}_{\text{diff}}(\vec{\mathbf{R}}_2,\vec{\mathbf{R}}_1;\delta\tau) = (4\pi D\tau)^{(-3N/2)} e^{-\left(\vec{\mathbf{R}}_2-\vec{\mathbf{R}}_1-D\delta\tau\vec{F}_Q(\vec{\mathbf{R}}_1)\right)^2/4D\delta\tau} \qquad (2.40)$$

may be obtained [27].

For the exact $G_{\text{diff}}(\vec{\mathbf{R}}_2,\vec{\mathbf{R}}_1;\delta\tau)$ the distribution ψ_G^2 is stationary,

$$\int G_{\text{diff}}(\vec{\mathbf{R}}_2,\vec{\mathbf{R}}_1;\delta\tau)\psi_G^2(\vec{\mathbf{R}}_1,\tau_1)\mathrm{d}\vec{\mathbf{R}}. = \psi_G^2(\vec{\mathbf{R}}_2,\tau_2) \qquad (2.41)$$

The importance sampled STA Green's function, Eq. (2.32), violates the symmetry property called *detailed balance*,

$$\tilde{G}_{\text{diff}}(\vec{\mathbf{R}}_2, \vec{\mathbf{R}}_1; \delta\tau) \neq \tilde{G}_{\text{diff}}(\vec{\mathbf{R}}_1, \vec{\mathbf{R}}_2; \delta\tau). \tag{2.42}$$

Therefore, the distribution ψ_G^2 is not stationary. Detailed balance may be imposed, as in VMC, by accepting the projected step according to the probability,

$$A(\vec{\mathbf{R}}_2, \vec{\mathbf{R}}_1; \delta\tau) \equiv \min\left(1, q(\vec{\mathbf{R}}_2, \vec{\mathbf{R}}_1; \delta\tau)\right), \tag{2.43}$$

where,

$$q(\vec{\mathbf{R}}_2, \vec{\mathbf{R}}_1; \delta\tau) \equiv \frac{|\psi_G(\vec{\mathbf{R}}_2)|^2 \tilde{G}_{\text{STA}}(\vec{\mathbf{R}}_1, \vec{\mathbf{R}}_2; \delta\tau)}{|\psi_G(\vec{\mathbf{R}}_1)|^2 \tilde{G}_{\text{STA}}(\vec{\mathbf{R}}_2, \vec{\mathbf{R}}_1; \delta\tau)}. \tag{2.44}$$

2.2.4 Energy estimator

The ground state is obtained from the weighted average of $\mathcal{E}_L$,

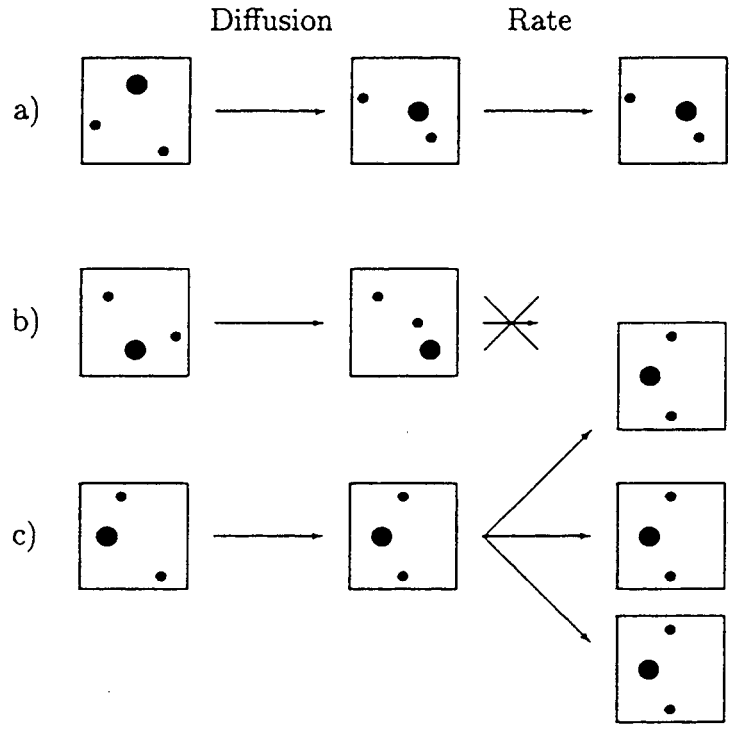
$$\begin{aligned}
\langle \mathcal{E} \rangle &= \frac{\int f_\infty(\vec{\mathbf{R}}) \mathcal{E}_L(\vec{\mathbf{R}}) \mathrm{d}\vec{\mathbf{R}}}{\int f_\infty(\vec{\mathbf{R}}) \mathrm{d}\vec{\mathbf{R}}} \\
&= \frac{\int \phi_0(\vec{\mathbf{R}}) \psi_G(\vec{\mathbf{R}}) \left[\frac{\hat{\mathcal{H}}\psi_G}{\psi_G}\right] \mathrm{d}\vec{\mathbf{R}}}{\int \psi_G(\vec{\mathbf{R}}) \phi_0(\vec{\mathbf{R}}) \mathrm{d}\vec{\mathbf{R}}} \\
&= \frac{\int \phi_0(\vec{\mathbf{R}}) \hat{\mathcal{H}} \psi_G(\vec{\mathbf{R}}) \mathrm{d}\vec{\mathbf{R}}}{\int \psi_G(\vec{\mathbf{R}}) \phi_0(\vec{\mathbf{R}}) \mathrm{d}\vec{\mathbf{R}}} \\
&= \frac{\int \mathcal{E}_0 \phi_0(\vec{\mathbf{R}}) \psi_G(\vec{\mathbf{R}}) \mathrm{d}\vec{\mathbf{R}}}{\int \psi_G(\vec{\mathbf{R}}) \phi_0(\vec{\mathbf{R}}) \mathrm{d}\vec{\mathbf{R}}} \\
&= \mathcal{E}_0,
\end{aligned} \tag{2.45}$$

which is estimated by

$$\mathcal{E}_0 = \lim_{M \to \infty} \frac{1}{M} \sum_{i=1}^{M} \mathcal{E}_L(\vec{\mathbf{R}}_i). \tag{2.46}$$

The accuracy of the energy estimate is independant of the guiding function but depends on the degree to which $f_\infty(\vec{\mathbf{R}})$ is exact. The resulting energy also has a time step bias because $\tilde{G}_{\text{STA}}(\vec{\mathbf{R}}_1, \vec{\mathbf{R}}_2; \delta\tau)$ is exact only as $\delta\tau \to 0$. The time step bias may be eliminated by evaluating $\mathcal{E}_0$ at different time steps and extrapolating to $\delta\tau = 0$.

Figure 2.2: One step of a diffusion Monte Carlo Random Walk. Each box represents a walker. The the large dot represents an oxygen atom and the two small dots represent hydrogen atoms. In a), the proposed step (diffusion) is accepted without replication or destruction (Rate). In b), the proposed step is rejected and the walker is deleted. In c), the proposed step is accepted with replication.



2.3 Correlation function quantum Monte Carlo

Diffusion Monte Carlo provides an improved accuracy over the VMC methods. Yet, it is primarily a ground state method and multiple vibrational states are desired. Excited states may be obtained by imposing wavefunction nodes by the guiding function. However, this requires prior knoweledge of the nodes of the excited state. The excited states are obtained more readily and naturally by CFQMC.

Correlation Function Quantum Monte Carlo is an extension of DMC which simultaneously computes the energies of the ground and excited states [19, 20]. Mathematically, it is based on the Rayleigh-Ritz variation and the sub-space iteration operator methods [36]. In the approach an operator $C(\widehat{\mathcal{H}})$, a function of the Hamiltonian, is applied to a set of basis functions f_m to project out the desired eigenfunctions.

$$C(\widehat{\mathcal{H}})f_m = \tilde{f}_m \tag{2.47}$$

Vibrational energies are obtained from the improved basis functions f_m. The multidimensional integrals of CFQMC are, from an analytical approach, intractable. For this reason, the choice of $C(\widehat{\mathcal{H}})$ is dictated by two considerations: ease of evaluation of the integrals for subspace iteration and capability of preferential projection of vibrational eigenfunctions. The imaginary-time evolution operator $e^{-\widehat{\mathcal{H}}\tau}$ satisfies both of these requirements for $C(\widehat{\mathcal{H}})$.

2.3.1 Preferential projection of eigenvalues

The imaginary-time evolution operator projects out the lower energy eigenfunctions from a set of trial functions. This effect is easily seen if we represent each trial wave function as a linear combination of eigenstates:

$$e^{-\widehat{\mathcal{H}}\tau}f_m = e^{-\widehat{\mathcal{H}}\tau}\sum_{i=0}^{N} c_i^m \Phi_i = c_i^m e^{-\varepsilon_i \tau}\Phi_i. \tag{2.48}$$

As $\tau \to \infty$ the higher energy components are projected out of the set of trial wave functions, thereby increasing the relative contribution of the lower energy components to the trial wave functions. The imaginary-time evolution operator evolves a set of M

approximate trial wave functions to the first M eigenstates of $\widehat{\mathcal{H}}$. With $C(\widehat{\mathcal{H}})$ given by $e^{-\widehat{\mathcal{H}}\tau}$ subspace iteration leads to the eigenvalue equation,

$$\sum_{m=1}^{M} \left[H_{mn}(\tau) - \Lambda_k(\tau) N_{mn}(\tau)\right] d_{km} = 0, \tag{2.49}$$

where

$$H_{mn}(\tau) = \langle f_m | \widehat{\mathcal{H}} e^{-\widehat{\mathcal{H}}\tau} | f_n \rangle = \sum_i c_i^m c_i^n \mathcal{E}_i e^{-\mathcal{E}_i \tau}, \tag{2.50}$$

$$N_{mn}(\tau) = \langle f_m | e^{-\widehat{\mathcal{H}}\tau} | f_n \rangle = \sum_i c_i^m c_i^n e^{-\mathcal{E}_i \tau}, \tag{2.51}$$

d_k is the k^{th} eigenvector, and $\Lambda_k(\tau)$ is the corresponding eigenvalue. The determinant in Eq. (2.49) may be written,

$$\det\left[H(\tau) - \Lambda(\tau) N(\tau)\right] = \sum_{i_1 < i_2 < \ldots} z_{i_1}(\tau) z_{i_2}(\tau) \ldots z_{i_n}(\tau) D^2_{i_0, i_1, \ldots, i_n} \tag{2.52}$$

where $z_i(\tau) = \left[\mathcal{E}_i - \Lambda(\tau)\right] e^{-\mathcal{E}_i \tau}$ and $D_{i_0, i_1, \ldots, i_n}$ is the determinant of coefficents c_i^m. The large time limit behavior of the determinant is,

$$\lim_{\tau \to \infty} e^{\mathcal{E}_0 \tau} e^{\mathcal{E}_1 \tau} \ldots e^{\mathcal{E}_n \tau} \det\left[H(\tau) - \Lambda(\tau) N(\tau)\right] = \left(\mathcal{E}_0 - \Lambda\right)\left(\mathcal{E}_1 - \Lambda\right) \ldots \left(\mathcal{E}_n - \Lambda\right) \quad D^2_{i_0, i_1, \ldots, i_n}. \tag{2.53}$$

If $D^2_{i_0, i_1, \ldots, i_n}$ does not vanish, then $\Lambda_k(\tau)$ converges to $\mathcal{E}_k$. In summary, as the trial functions approach the exact eigenstates, the $\Lambda_k(\tau)$ approach the exact eigenvalues of $\widehat{\mathcal{H}}$. Indeed, in the limit $\tau \to \infty$ each $\Lambda_k(\tau)$ decreases monotonically and converges exponentially fast to $\mathcal{E}_k$[19, 20].

2.3.2 Quantum Monte Carlo evaluation of correlation functions

The integrals Eqs. (2.50) and (2.51) may be computed stochastically by QMC. One rewrites them as integrals over a Green's function

$$\begin{aligned} H_{mn}(\tau) &= \int F_m(\vec{\mathrm{R}}_2) \mathcal{E}_{L_m}(\vec{\mathrm{R}}_2) G(\vec{\mathrm{R}}_2, \vec{\mathrm{R}}_1; \tau) F_n(\vec{\mathrm{R}}_1) P(\vec{\mathrm{R}}_1) \mathrm{d}\vec{\mathrm{R}}_1 \mathrm{d}\vec{\mathrm{R}}_2, \\ N_{mn}(\tau) &= \int F_m(\vec{\mathrm{R}}_2) G(\vec{\mathrm{R}}_2, \vec{\mathrm{R}}_1; \tau) F_n(\vec{\mathrm{R}}_1) P(\vec{\mathrm{R}}_1) \mathrm{d}\vec{\mathrm{R}}_1 \mathrm{d}\vec{\mathrm{R}}_2, \end{aligned} \tag{2.54}$$

where $G(\vec{\mathbf{R}}_2, \vec{\mathbf{R}}_1; \tau) \equiv \psi_G(\vec{\mathbf{R}}_2) \left\langle \vec{\mathbf{R}}_2 \right| e^{-\widehat{\mathcal{H}}\tau} \left| \vec{\mathbf{R}}_2 \right\rangle / \psi_G(\vec{\mathbf{R}}_1)$ is the Green's function of the time independent Schrödinger equation, $F_n(\vec{\mathbf{R}}) \equiv f_n(\vec{\mathbf{R}})/\psi_G(\vec{\mathbf{R}})$, $P(R) \equiv \psi_G^2(\vec{\mathbf{R}})$ is the probability that a rondom walk begins at $\vec{\mathbf{R}}_1$, $\mathcal{E}_{L_m}(\vec{\mathbf{R}}_2) \equiv f_m^{-1} \widehat{\mathcal{H}} f_m^{-1}$, and $\psi_G(\vec{\mathbf{R}})$ is the guiding function. $G(\vec{\mathbf{R}}_2, \vec{\mathbf{R}}_1; \tau)$ is the DMC Green's function of Eqs. (2.24) through (2.31) using the short time approximation, described in Eqs. (2.32) through (2.36). Diffusion is simulated by an importance sampled Metroplis Random walk as in DMC and VMC. However, the rate process is simulated by assigning a weight to each walker as it moves from $\vec{\mathbf{R}}_j$ to $\vec{\mathbf{R}}_{j+l}$,

$$W_{n,n+l} = \exp\left(-0.5\delta\tau \sum_{i=j}^{j+l-1} \left[\mathcal{E}_{L_G}(\vec{\mathbf{R}}_i) + \mathcal{E}_{L_G}(\vec{\mathbf{R}}_{i+1}) \right] \right), \tag{2.55}$$

where $\mathcal{E}_{L_G}(\vec{\mathbf{R}}) \equiv \psi_G^{-1}(\vec{\mathbf{R}}) \widehat{\mathcal{H}} \psi_G(\vec{\mathbf{R}})$

In the simplest form, the matrix elements H_{mn} and N_{mn}, Eqs. (2.54), could be obtained from:

$$\begin{aligned} h_{mn}(l\delta\tau) &= \frac{1}{4(p-l)} \sum_{i=1}^{p-l} F_m(\vec{\mathbf{R}}_i) W_{i,i+l} F_n(\vec{\mathbf{R}}_{i+l}) \mathcal{E}_{L_n}(\vec{\mathbf{R}}_{i+l}), \\ n_{mn}(l\delta\tau) &= \frac{1}{2(p-l)} \sum_{i=1}^{p-l} F_m(\vec{\mathbf{R}}_i) W_{i,i+l} F_n(\vec{\mathbf{R}}_{i+l}), \end{aligned} \tag{2.56}$$

where p is the total number of steps per walker in the random walk in which sampling occurs, l is the projection time in steps, and n and m index the trial wave functions. However, the matrix elements $h_{mn}(l\delta\tau)$ and $n_{mn}(l\delta\tau)$ are symmetrized to reduce fluctuations:

$$\begin{aligned} h_{mn}(l\delta\tau) &= \frac{1}{4(p-l)} \sum_{i=1}^{p-l} [\mathcal{E}_{L_m}(\vec{\mathbf{R}}_i) F_m(\vec{\mathbf{R}}_i) F_n(\vec{\mathbf{R}}_{i+l}) \\ &\quad + \mathcal{E}_{L_n}(\vec{\mathbf{R}}_i) F_n(\vec{\mathbf{R}}_i) F_m(\vec{\mathbf{R}}_{i+l}) + F_m(\vec{\mathbf{R}}_i) F_n(\vec{\mathbf{R}}_{i+l}) \mathcal{E}_{L_n}(\vec{\mathbf{R}}_{i+l}) \\ &\quad + F_n(\vec{\mathbf{R}}_i) F_m(\vec{\mathbf{R}}_{i+l}) \mathcal{E}_{L_m}(\vec{\mathbf{R}}_{i+l})] W_{i,i+l} \\ n_{mn}(l\delta\tau) &= \frac{1}{2(p-l)} \sum_{i=1}^{p-l} W_{i,i+l} \left[F_m(\vec{\mathbf{R}}_i) F_n(\vec{\mathbf{R}}_i) + F_n(\vec{\mathbf{R}}_i) F_m(\vec{\mathbf{R}}_{i+l}) \right], \end{aligned} \tag{2.57}$$

where the symmetry property, $h_{mn} = h_{nm}$, and Hermitian property of $\widehat{\mathcal{H}}$ were applied to obtain a four term sum for h_{mn}, and the symmetry property, $n_{mn} = n_{nm}$, was

applied to obtain a two term sum for n_{mn}. The average values of $h_{mn}(l\delta\tau)$ and $n_{mn}(l\delta\tau)$ from the the random walk are the integrals over the probality distribution of the path $P(\vec{\mathbf{R}})$. Therefore, as $p \rightarrow \infty$, $h_{mn}(l\delta\tau)$ and $n_{mn}(l\delta\tau)$ converge to $H_{mn}(l\delta\tau)$ and $N_{mn}(l\delta\tau)$.

In order to observe convergence in $\Lambda_k(l\delta\tau)$, the matrix elements $n_{mn}(k\delta\tau)$ and $h_{mn}(k\delta\tau)$ are evaluated for several different projection times $\tau = l\delta\tau$. Therefore, l takes values 0,1I, 2I,...,L, where where I in the projection time interval and L is the maximum calculated projection time. Calculating $n_{mn}(nl\delta\tau)$ and $h_{mn}(nl\delta\tau)$ at intervals $I\delta\tau$ rather than $\delta\tau$ reduces the computational burden by a factor I because between steps i and i +1 $\mathcal{E}_{L_m}$ and F_m are not computed. Typical values of $I\delta\tau$ lie in the range 10-15 H^{-1} where H is a Hartree. Efficiency of matrix element evaluation is important because it accounts for most of the computational cost of the approach.[6]

2.3.3 Multi-state energy estimator

After matrix element evaluation, Eq. (2.49) is solved at each projection time $nl\delta\tau$, resulting in a set of eigenvalues for each projection time. Ideally, each of the eigenvalues $\Lambda_k(nl\delta\tau)$ attains a plateau as $\tau = (nl\delta\tau)$ becomes large, in which case the exact eigenvalues are given by the plateau values. When this condition is not met, the slope of each $\Lambda_k(nl\delta\tau)$ with respect to τ is computed by finite difference. The minimum slope indicates convergence so long as the absolute value of the slope is vanishingly small or if the slope is smaller than the standard deviation of the eigenvalue. If the slope does not become vanishingly small prior to the maximum calculated projection time, only an upper bound to the exact eigenvalue is obtained.

2.3.4 Measurement of statistical error

In addition to computing $\Lambda_k(nl\delta\tau)$ which are obtained from the entire ensemble of random walkers, we compute eigenvalues $\lambda_k(nl\delta\tau)$ by diagonalizing matrices obtained from blocks of the random walkers, the average of these eigenvalues $\overline{\lambda}$, and

[6]For the calculations presented here the matrix element evaluation consumed more than 97% of computer time.

standard deviations of $\lambda_k(nl\delta\tau)$ about $\overline{\lambda}$ and $\Lambda_k(nl\delta\tau)$. The latter measure the bias that results from the non-linearity of Eq. (2.49).

2.3.5 Analysis

CFQMC has several of the desired attributes for a method of solving the vibrating molecule Schrödinger equation: flexibility in choice of basis set, multiple states, and higher accuracy than VMC. CFQMC has one obvious shortcoming. It does not provide analytic wavefunctions for the vibrational states. The utility of CFQMC for computing vibrational state energies is the focus of this study.

Chapter 3

Monte Carlo for vibrating H_2O

I returned, and saw under the sun, that the race is not to the swift, nor the battle to the strong, neither yet bread to the wise, nor yet riches to men of understanding, nor yet favour to men of skill; but time and chance happeneth to them all.

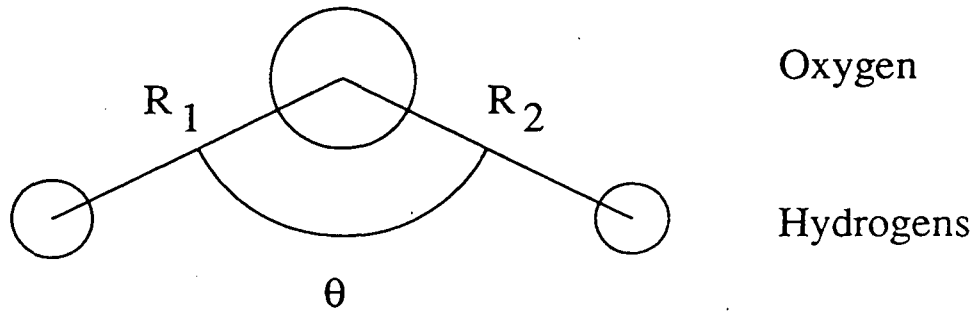
Ecclesiastes 9:11 (KJV)

3.1 Introduction

The evaluation of the utility of CFQMC for computing molecular vibrational states began with Bernu, Ceperley and Lester[19, 20]. They computed vibrational states for H_2O and H_2CO and examined the error sources in the computed eigenvalues. Following the initial CFQMC investigation, a second improved program, QMCVIB, was developed by Glauser, Brown and Lester. The vibrational energies of H_2O are re-computed by CFQMC to confirm the QMCVIB implementation and improve the earlier calculation.[1]

[1]The QMCVIB code, compiling instructions, sample input files, and execution instructions are included in the Appendix C.

Figure 3.1: Coordinate diagram of H_2O.



3.2 H_2O potentials

Water is an extensively studied molecule. Its vibrational spectra is well determined experimentally [37, 38, 39, 40]. Several PES's have been constructed for the water molecule [41, 42, 43, 44, 45] and the band origins have been computed many times by various methods [46, 44, 47, 45], including CFQMC [20]. In order to verify QMCVIB, the band origins of H_2O were computed using two potential energy surfaces, using the basis functions and guiding functions employed by Bernu et al. Pes I was developed by Hoy, Mills and Strey[41] and has the form,

$$V(\Delta R_1, \Delta R_2, \Delta R_3) = \sum_i \sum_j K_{ij} \Delta R_i \Delta R_j + \sum_i \sum_j \sum_k K_{ijk} \Delta R_i \Delta R_j \Delta R_k + \ldots, \quad (3.1)$$

where ΔR_1 and ΔR_2 are displacements of the bond lengths R_1 and R_2 from equilibrium, and $\Delta R_3 = \Delta\theta$ is the displacement of the bend angle θ from equilibrium. See Fig. 3.1. Pes II was developed by Carney, Curtis and Langhoff [44] to improve the potential description at large displacements and has the form,

$$V(\rho_1, \rho_2, \rho_3) = \sum_i \sum_j L_{ij} \rho_i \rho_j + \sum_i \sum_j \sum_k L_{ijk} \rho_i \rho_j \rho_k + \ldots, \quad (3.2)$$

where $\rho_1 = \Delta R_1/R_1$, $\rho_2 = \Delta R_2/R_2$ and $\rho_3 = \Delta\theta$, and was used by Bernu et al. in their H_2O calculations. The force constants for PES I and II are given in Table 3.1.

Table 3.1: Force constants for water PES I and II in units of 10^{-12} ergs/Å

PES I Internal Coordinates		PES II Simons-Parr-Finlan Coordinates	
Expansion Term	Force Constant	Expansion Term	Force Constant
$(\Delta r_1^2 + \Delta r_2^2)/2$	84.54	$r_e^2(\rho_1^2 + \rho_2^2)/2$	84.54
$\Delta r_1 \Delta r_2$	-1.01	$r_e^2 \rho_1 \rho_2$	-1.01
$(\Delta r_1 + \Delta r_2)(r_e \Delta\theta)$	2.288	$r_e(\rho_1 + \rho_2)(r_e \Delta\theta)$	2.288
$(r_e \Delta\theta)^2/2$	7.607	$(r_e \Delta\theta)^2/2$	7.607
$(\Delta r_1^3 + \Delta r_2^3)/r_e$	-94.708	$r_e^2(\rho_1^3 + \rho_2^3)$	-10.168
$\Delta r_1 \Delta r_2(\Delta r_1 + \Delta r_2)/r_e$	1.211	$r_e \rho_1 \rho_2(\rho_1 + \rho_2)$	0.201
$(\Delta r_1^2 + \Delta r_2^2)(r_e \Delta\theta)/r_e$	2.02	$r_e(\rho_1^2 + \rho_2^2)(r_e \Delta\theta)$	4.308
$\Delta r_1 \Delta r_2(r_e \Delta\theta)/r_e$	-4.02	$r_e \rho_1 \rho_2(r_e \Delta\theta)$	-4.02
$(\Delta r_1 + \Delta r_2)(r_e \Delta\theta)^2/r_e$	-1.175	$(\rho_1 + \rho_2)(r_e \Delta\theta)^2$	-1.175
$(r_e \Delta\theta)^3/r_e$	-1.595	$(r_e \Delta\theta)^3/r_e$	-1.595
$(\Delta r_1^4 + \Delta r_2^4)/r_e^2$	146.63	$r_e^2(\rho_1^4 + \rho_2^4)$	-10.684
$(\Delta r_1^2 + \Delta r_2^2)\Delta r_1 \Delta r_2/r_e^2$	-7.574	$r_e^2 \rho_1 \rho_2(\rho_1^2 + \rho_2^2)$	-6.162
$(\Delta r_1 \Delta r_2)^2/r_e^2$	1.305	$r_e^2 \rho_1^2 \rho_2^2$	2.717
$(\Delta r_1^3 + \Delta r_2^3)(r_e \Delta\theta)/r_e^2$	0.0	$r_e(\rho_1^3 + \rho_2^3)(r_e \Delta\theta)$	6.328
$(\Delta r_1 + \Delta r_2)\Delta r_1 \Delta r_2(r_e \Delta\theta)/r_e^2$	0.0	$r_e(\rho_1 + \rho_2)\rho_1 \rho_2(r_e \Delta\theta)$	-4.02
$(\Delta r_1^2 + \Delta r_2^2)(r_e \Delta\theta)^2/r_e^2$	-3.525	$(\rho_1^2 + \rho_2^2)(r_e \Delta\theta)^2$	-4.700
$\Delta r_1 \Delta r_2(r_e \Delta\theta)^2/r_e^2$	3.05	$\rho_1 \rho_2(r_e \Delta\theta)^2$	3.05
$(\Delta r_1 + \Delta r_2)(r_e \Delta\theta)^3/r_e^2$	0.0	$(\rho_1 + \rho_2)(r_e \Delta\theta)^3/r_e$	0.0
$(r_e \Delta\theta)^4/r_e^2$	-0.0318	$(r_e \Delta\theta)^4/r_e^2$	-0.0318

Table 3.2: Acceptance function parameters for H_2O in $1/bohr^2$

local modes		
μ	ν	$A_{\mu\nu}$
1	1	-19.611
1	2	-4.316
1	3	6.108
2	2	-19.611
2	3	6.108
3	3	8.599
$\rho = 0.41$		

3.3 Local mode basis functions

For the calculation of the vibrational state energies of H_2O, Bernu et al. used local mode (LM) basis functions of the form

$$f_m = \exp\left(-\sum_{\nu=1}^{3}\sum_{\mu=1}^{3}\Delta S_\nu A_{\mu\nu}\Delta S_\mu\right), \tag{3.3}$$

for zeroth order function

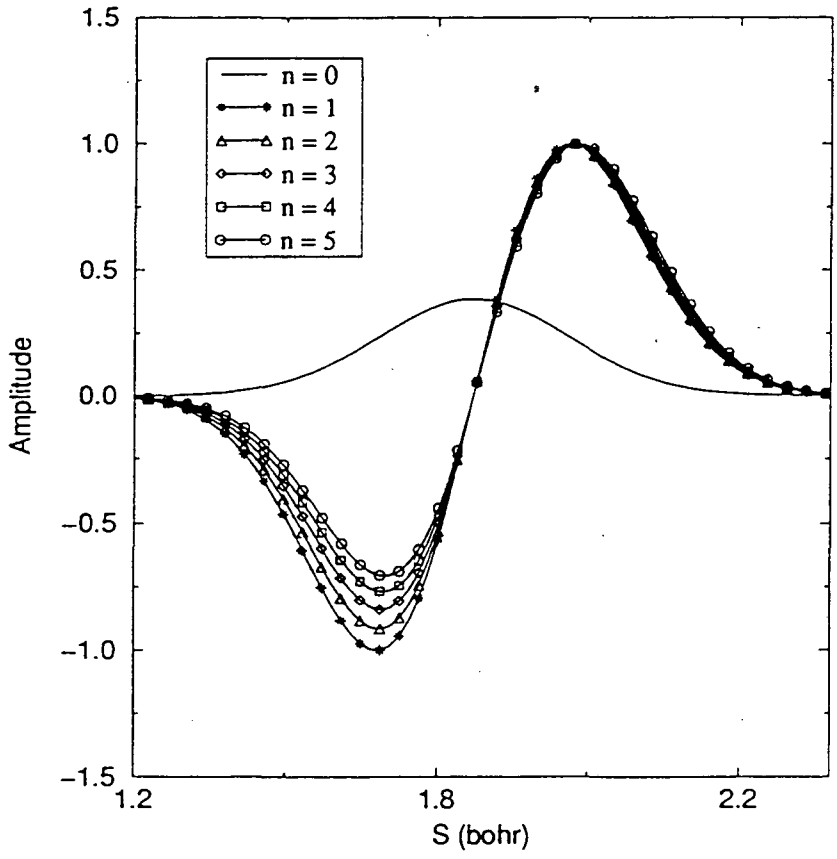
$$f_m = \left(\prod_{\nu=1}^{3} S_\nu^{n_\nu(m)} - \prod_{\nu=1}^{3}\left(S_\nu^0\right)^{n_\nu(m)}\right)\exp\left(-\sum_{\nu=1}^{3}\sum_{\mu=1}^{3}\Delta S_\nu A_{\mu\nu}\Delta S_\mu\right), \tag{3.4}$$

where S_ν are the inter-atomic distances, ΔS_ν are their displacements from from equilibrium, $A_{\mu\nu}$ are non-linear variational parameters. The zeroth order function approximates the nodeless ground state wavefunction. The higher order functions each have one node and do not individually approximate any excited state. Linear combinations of the zeroth and higher order functions are used to approximate the excited states. Ground and excited state basis functions are illustrated in Fig. 3.2. The values of $A_{\mu\nu}$ used by Bernu *et al.* are given in Table 3.2 [20].[2] Bernu et al. used a basis set of 35 functions with $n_1 + n + 2 + n_3 \leq 4$ to construct the VMC wavefunctions.

The guiding function is integral to the VMC and CFQMC methods, as explained in Chapter 2. Bernu et al. used a broadened ground state basis function

[2]Ref. [20] has a typographical error. The units of $A_{\mu\nu}$ for H_2O are $1/bohr^2$, not $1/\mathring{A}^2$.

Figure 3.2: Local Mode Basis Functions: The ground state has zero nodes and all higher order functions have one node each. See Eqs. (3.3) and (3.4).



in the H_2O calculations,

$$\psi_G = \exp\left(-\rho \sum_{\nu=1}^{3}\sum_{\mu=1}^{3} \Delta S_\nu A_{\mu\nu} \Delta S_\mu\right), \tag{3.5}$$

where ρ, the width parameter, is given in Table 3.2.

The present calculations use LM basis functions, guiding function and $A_{\mu\nu}$ parameters given in Table 3.2, but include all basis functions with $n_1 + n_2 + n_3 \leq 6$, resulting in 84 functions for the VMC basis set. The additional functions should allow a more accurate VMC calculation and, subsequently, an improved CFQMC calculation.

3.4 VMC and CFQMC calculations of PES I

The VMC ensemble of 16,384 walkers was propagated 20,000 steps with $\delta\tau = 160$ to reach an equilibrium distribution of ψ_G^2 and afterwards propagated 10,000 steps with $\delta\tau = 160$ to sample the Hamiltonian and overlap matrix elements, Eqs. (2.12) and (2.13). The total sampling of these matrix elements was $M = 163,840,000$.[3] The VMC wavefunctions from the twenty lowest energy states were used as trial wavefunctions for the CFQMC calculation. The CFQMC ensemble of 16,384 walkers was propagated 20,000 steps with $\delta\tau = 2$ to equilibrate and 10,000 steps with $\delta\tau = 2$ to sample the CFQMC matrix elements, Eq. (2.57). The matrix elements were computed at projection time intervals of 12 H^{-1} for $\tau = 0$ to $\tau = 288$ H^{-1}. The VMC and CFQMC vibrational state energies are given in Table 3.3.

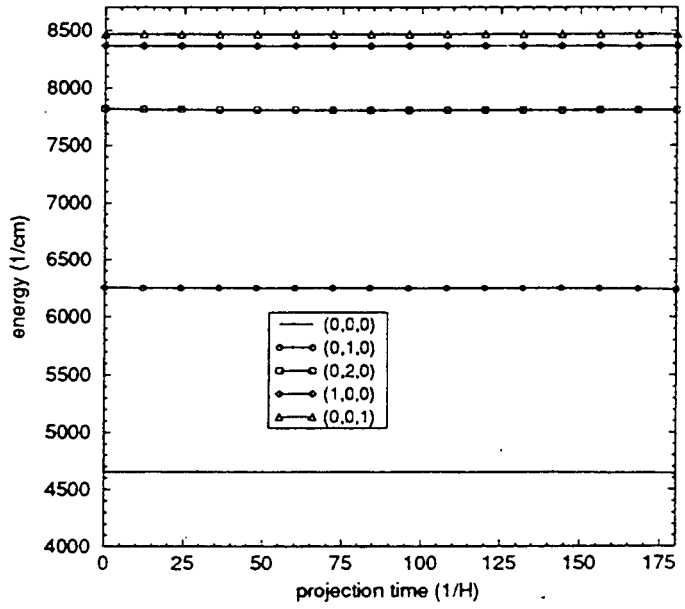
The variational wavefunction for the lowest vibrational states of H_2O are accurate. The differences between the VMC and converged CFQMC energies for states (0,0,0), (0,1,0), (1,0,0), (0,0,1), and (1,0,1) are less than 3 cm^{-1}. As the excitation of each mode increases, the variational wavefunctions are less accurate. The difference between the VMC and converged CFQMC energy is 38 cm^{-1} for (0,2,1) and for 92 cm^{-1} for (1,2,0). The $(\nu_1, \nu_2, \nu_3) - (0,0,0)$ energy differences computed

[3]This calculation was performed on a 512 node partition of a Thinking Machines CM5 parallel computer at the Army High Performance Supercomputer Research Center.

Table 3.3: VMC and CFQMC vibrational state energies using PES I for the ground state of H_2O in cm^{-1}. The numbers in parentheses are the estimated statistical errors (one standard deviation). The * denotes unconverged energies (no observed plateau), which are upper bounds. VMC and CFQMC refer to the present calculations. The experimental entries are from Ref. [38]. Variational refers to the calculation of Ref. [44].

(ν_1, ν_2, ν_3)	Expt.	VMC	CFQMC	Variational
ZPE		4652(0)	4651(1)	
(0,1,0)	1595	1600(0)	1598(1)	1597
(0,2,0)	3152	3169(1)	3158(3)	3159
(0,3,0)	4667	4735(4)	4683(3)	4683
(0,4,0)	6136	6644(3)	6205(3)*	6168
(1,0,0)	3657	3717(0)	3717(3)	3717
(1,1,0)	5235	5307(1)	5299(3)	5299
(1,2,0)	6775	6938(1)	6845(3)	6845
(2,0,0)	7201	7424(1)	7421(3)	7419
(0,0,1)	3756	3821(0)	3820(3)	3821
(0,1,1)	5331	5406(0)	5402(3)	5403
(0,2,1)	6872	6990(1)	6951(3)	6950
(0,0,2)	7445	7628(3)	7625(3)	7626
(1,0,1)	7250	7523(3)	7522(3)	7520

Figure 3.3: Λ_k for the 5 lowest energy states of H_2O from PES II. The labels (ν_1,ν_2,ν_3) indicate the vibrational state of the Λ_k



from converged CFQMC energies and the differences computed variationally agree within 1 CFQMC standard deviation.

3.5 VMC and CFQMC calculations of PES II

The VMC and CFQMC calculations for PES II were carried out using the same size ensemble, time steps, and number of trial functions as used in the PES I calculations. The matrix elements were computed at projection time intervals of 12 H^{-1} for $\tau = 0$ H^{-1} to $\tau = 180$ H^{-1}. The convergence behavior of the vibrational eigenvalues for 19 states of H_2O is shown in Figs. 3.3, 3.4, 3.5 and 3.6. The VMC and CFQMC vibrational state energies are given in Table 3.4.

As in the preceding calculation, the variational wavefunction for the lowest

Figure 3.4: Λ_k for 4 higher energy states of H_2O from PES II. The labels (ν_1,ν_2,ν_3) indicate the vibrational state of the Λ_k

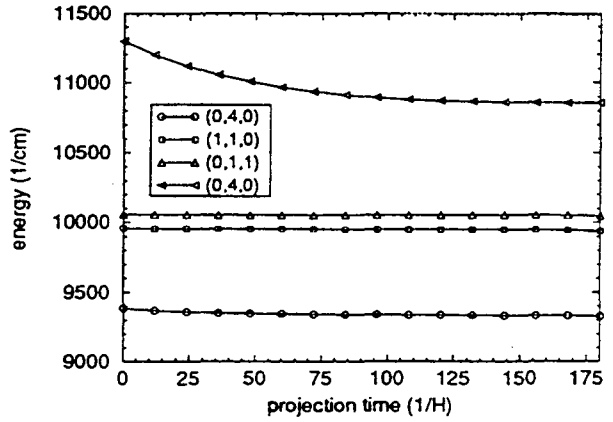


Figure 3.5: Λ_k for 5 higher energy states of H_2O from PES II. The labels (ν_1,ν_2,ν_3) indicate the vibrational state of the Λ_k

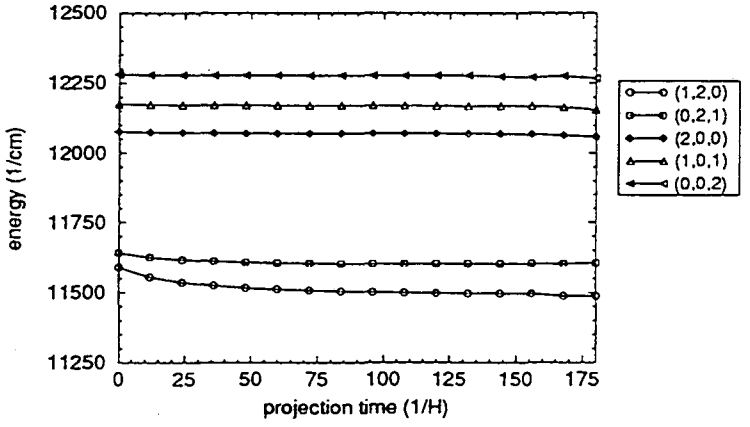
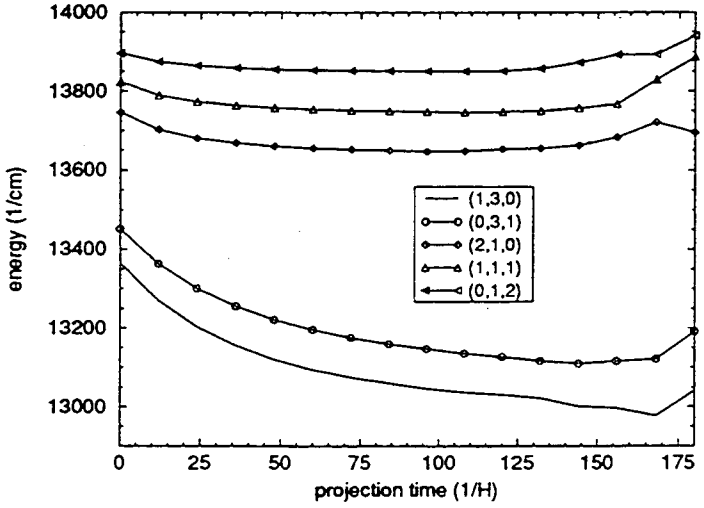


Table 3.4: VMC and CFQMC vibrational state energies using PES II for the ground state of H_2O in cm^{-1}. The numbers in parentheses are the estimated statistical errors (one standard deviation). The * denotes unconverged energies (no observed plateau), which are upper bounds. VMC and CFQMC refer to the present calculations. CFQMC* refers to the calculation of Ref. [20]. The experimental entries are from Ref. [38]. Variational refers to the calculation of Ref. [44].

(ν_1, ν_2, ν_3)	Expt.	VMC	CFQMC	CFQMC*	Variational
ZPE		4639(0)	4639(1)	4639(0)	
(0,1,0)	1595	1599(0)	1595(2)	1595(0)	1596
(0,2,0)	3152	3166(0)	3155(2)	3150(4)	3156
(0,3,0)	4667	4733(4)	4685(3)	4677(15)	4679
(0,4,0)	6136	6626(3)	6206(3)*	6230(20)*	6169
(1,0,0)	3657	3661(0)	3658(4)	3655(2)	3659
(1,1,0)	5235	5249(1)	5236(4)	5231(10)	5237
(1,2,0)	6775	6900(1)	6714(6)*	6799(15)	6777
(2,0,0)	7201	7242(1)	7208(3)	7221(15)	7211
(0,0,1)	3756	3760(0)	3756(3)	3756(1)	3758
(0,1,1)	5331	5344(0)	5335(2)	5321(10)	5335
(0,2,1)	6872	6933(1)	6879(4)	6886(15)	6877
(0,0,2)	7445	7466(3)	7451(3)	7451(15)	7453
(1,0,1)	7250	7292(3)	7257(4)	7256(15)	7250

Figure 3.6: Λ_k for 5 higher energy states of H_2O from PES II. The labels (ν_1,ν_2,ν_3) indicate the vibrational state of the Λ_k



vibrational states of H_2O are accurate. The differences between the VMC and converged CFQMC energies are less than or equal to 5 cm^{-1} for states (0,0,0), (0,1,0), (1,0,0), and (0,0,1), less than or equal to 15 cm^{-1} for states (0,1,1), (0,2,0), (1,1,0) and (0,0,2), and less than or equal to 54 cm^{-1} for (2,0,0), (1,0,1), (0,3,0) and (0,2,1). See Table 3.4. Each of these states reached a plateau before 180 H^{-1} projection time. See Figs. 3.3, 3.4, 3.5 and 3.6. States (0,4,0), (1,2,0), (1,3,0) and (0,3,1) did not reach a plateau. The Λ_k for states (1,3,0), (0,3,1), (2,1,0), (1,1,1) and (0,1,2) became unstable after 125 H^{-1} projection time. The statistical error in the CFQMC matrix elements generally increases with projection time and higher energy λ_k are most seriously effected by the statistical noise. See Fig. 3.6.

The $(\nu_1, \nu_2, \nu_3) - (0, 0, 0)$ energy differences computed from the two independent CFQMC calculations agree well. For most bands the measurements agree within 1 standard deviation and in all others the error bars overlap. The $(\nu_1, \nu_2, \nu_3) - (0, 0, 0)$ energy differences computed from the converged states of the CFQMC[1] calculation and the energy differences computed variationally agree within 2 CFQMC standard deviations. The CFQMC energy differences agree with the experimental measure-

ments of the bands within 6 cm^{-1}.

The present VMC calculation, despite the additional basis functions used, obtained slightly higher vibrational state energies than the VMC calculation of Ref. [20]. This may be accounted for by the more extensive sampling of the present calculations which resulted in smaller statistical error and bias in the vibrational energies compared to the prior calculations. Statistical bias in the lower states is typically negative, pushing the values of the vibrational energies down. Since the present calculations have generally smaller bias, the energies of the lower states tend to be higher than those of the prior calculation.

3.6 Analysis of H_2O calculations

The CFQMC algorithm works. The operation of the time-evolution operator on the trial function, by which the contribution of the higher energy eigenstates is removed, is seen indirectly in the converging Λ_k. This confirms that the short-time approximation to the Green's function, Eqs. (2.32) and (2.40), does not frustrate the improvement of the trial functions. This improvement is clearly limited by the statistical error in the CFQMC matrix elements. States with poor trial functions, such as (0,4,0) in the PES I calculation and (0,4,0) and (1,2,0) in the PES II calculation failed to converge.

The λ for the blocks of random walkers become unstable prior (in projection time) to the Λ_k for the whole ensemble of random walkers. The overlap matrices $N_{(mn)}$ for the blocks, and later for the whole ensemble, become singular and the corresponding eigenvalue equation, Eq. (2.49) becomes unsolvable.[4] Therefore, at large projection times the estimated statistical error in the Λ_k grows. This demonstrates that increased QMC sampling extends the projection time in which the Λ_k are stable. Also, accurate trial functions require less projection time for the Λ_k to converge than inaccurate trial functions as seen in states (0,4,0) and (2,0,0) of the second calculation, despite (2,0,0) being a higher energy state than (0,4,0). See Figs. 3.4 and

[4]This problem is discussed by Caffarel *et al.* paper discussing the application of Lanczos-type algorithm to QMC data.[48]

3.5.

The problem of singularities can occur in VMC calculations as well. In addition to the 84 basis function VMC calculations for PES II, a 120 basis function calculation was conducted. All basis functions with $n_1 + n_2 + n_3 \leq 7$ were included in the basis set. The $A_{\mu\nu}$ parameters, as given in Table 3.2, were used. A VMC ensemble of 128 walkers was propagated 50,000 steps with $\delta\tau = 160$ to reach an equilibrium distribution of ψ_G^2 and afterwards propagated 100,000 steps with $\delta\tau = 160$ to sample the Hamiltonian and overlap matrix elements, Eqs. (2.12) and (2.13). The total sampling of the matrix element was $M = 12$ million.[5] The overlap matrix, Eq. (2.13), was singular due to an apparent linear dependency. Formally, the basis functions used are linearly independent. However, the differences between LM basis functions decreases with the higher order of the function. See Fig. 3.2. The VMC evaluation of the matrix elements did not distinguish the basis functions at the level of sampling attempted.

[5]These calculations were conducted on a Cray-2 at the National Energy Research Supercomputer Center at Lawrence Livermore National Laboratory.

Chapter 4

Monte Carlo for floppy C_3

For I know the thoughts that I think toward you, saith the LORD, thoughts of peace, and not of evil, to give you an expected end.
Jeremiah 29:11 (KJV)

4.1 Introduction

CFQMC is capable of calculating the vibrational states of relatively rigid H_2O. The capacity of CFQMC to calculate the vibrational states of floppy molecules must be be explored, given the present high interest in determining the internal motion of floppy molecules. C_3 is ideal to test the approach because it is a well studied system. Weltner and Van Zee have reviewed C_3 and carbon cluster investigations prior to 1989[49]. Prior to the present work, the vibrational states of C_3 had not been computed by QMC.

During the last several years the interest in pure carbon molecules has grown. New experimental methods allow the production of large carbon clusters. These have fascinating complexities in their molecular properties and spectroscopy, such as the highly stable C_{60} and the Renner-Teller effects in C_3. Carbon clusters have astrophysical significance as reactants in the formation of long-chain cyanopolyynes, carbon dust, and polycyclic aromatic hydrocarbons. The clusters have combustion significance. They are involved in the nucleation of carbon particles and the formation

of soot in hydrocarbon flames.

C_3 has particularly interesting spectroscopy. The $^1\Sigma_g^+$ ground state and the $^1\Pi_u$ excited state are degenerate for C_3 in the linear configuration. Bending from linearity resolves the degeneracy, so that the vibronic interaction is involved and the Born-Oppenheimer approximation breaks down. Additionally, C_3 is an extremely floppy molecule, having large amplitude vibrations with bending angles up to 60 degrees.

There is some controversy concerning the linearity of C_3. Recent work by Vala indicates a non-linear equilibrium bond angle in agreement with the potential energy surface calculated by Kraemer et al., and refined by Jensen[50, 51, 52, 22]. The large complete active space self consistent field calculations of Rohlfing et al. [7] and recent coupled cluster calculations of Mladenovic et al. indicate a linear equilibrium geometry[8].

4.2 C_3 potentials

In this study of C_3 three potential energy surfaces were used. Pes I is a fit of configuration interaction[1] energies to the MORBID form (see below) which was subsequently refined via MORBID calculations to make the calculated spectrum match the experimental spectrum[22]. It has a non-linear equilibrium geometry and a small barrier to linearity of 16.5 cm^{-1}. Pes II is the fit of CI energies from which PES I was refined. It has a non-linear equilibrium geometry and a small barrier to linearity of 21 cm^{-1}. Pes III is a fit of complete active space self-consistent field[2] energies to the MORBID form without the introduction of experimental data.[7] It has a linear equilibrium geometry. Pes I, II and III each have a central carbon attached to two identical terminal carbon atoms and a single minimum. Because C_3 belongs to the S^3 permutation group the C_3 PES is symmetric with respect to the interchange of

[1] Configuration interaction (CI) refers to the ab initio method of calculating the total electronic energy of atoms or molecules in which the N-electron Hamiltonian is diagonalized in a basis set of Slater determinants.[2]

[2] Complete active space self-consistent field (CASSCF) refers to the *ab initio* method of calculating the total electronic energy of atoms or molecules in which ...

any two carbon atoms and have three identical minima However, since no splittings have been observed experimentally, we have chosen not to symmetrize the potential energy surfaces nor the trial wave functions.

The MORBID PES form is designed to treat tri-atomic molecules having a single floppy mode developed by Jensen.[53] It is a polynomial expansion,

$$\begin{aligned} V(\Delta r_1, \Delta r_2, \bar{\rho}) &= V_0(\bar{\rho}) + \sum_j F_j(\bar{\rho}) y_j + \sum_{j<k} F_{jk}(\bar{\rho}) y_j y_k \\ &+ \sum_{j<k<m} F_{jkm}(\bar{\rho}) y_j y_k y_m + \sum_{j<k<m<n} F_{jkmn}(\bar{\rho}) y_j y_k y_m y_n, \end{aligned} \quad (4.1)$$

where $y_j = 1 - \exp(-a_j \Delta r_j)$ describes the stretching part of the potential, $F(\bar{\rho})$ describe the bending part of the potential, $\Delta r \equiv r - r_0$ are deviations from the equilibrium bond length and $\bar{\rho}$ is the complement of the internal bond angle.

The function y_j is related to the form of the Morse potential,

$$V_{\text{Morse}}(\delta r) = D[1 - exp(-a\delta r)]^2, \quad (4.2)$$

which accurately describes the stretching potential of many molecules. Typically, polynomials in the internal bond angle θ or $\bar{\rho}$, are chosen to describe the bending part of the PES. Jensen chose to use expansions in cosines of $\bar{\rho}$ which satisfy the zero slope boundary conditions at $\bar{\rho} = 0$ and $\bar{\rho} = \pi$ require by symmetry. The angular part of the potential is described by the functions,

$$F_j(\bar{\rho}) = \sum_{i=1}^{4} f_j^{(i)} (\cos \bar{\rho}_e - \cos \bar{\rho})^i \quad (4.3)$$

and

$$F_{jk\ldots}(\bar{\rho}) = f_{jk\ldots}^{(0)} + \sum_{i=1}^{N} f_{jk\ldots}^{(i)} (\cos \bar{\rho}_e - \cos \bar{\rho})^i . \quad (4.4)$$

Additionally, the pure bending potential is described by

$$V_0(\bar{\rho}) = \sum_{i=1}^{8} f_0^{(i)} (cos\rho_e - \cos \bar{\rho})^i \quad (4.5)$$

for bent molecules and by

$$V_0(\bar{\rho}) = \sum_{i=1}^{8} f_0^{(i)} (1 - \cos \bar{\rho})^i \quad (4.6)$$

for either bent or linear molecules. The parameters for PES I, II and III are given in Table 4.1.

Table 4.1: Potential function parameters for C_3

	Units	Pes I	Pes II	Pes III
r_{12}	(Ang.)	1.289726	1.289726	1.30906
a_1	(Ang.$^{-1}$)	2.1	2.1	2.1
$f_0^{(1)}$	(cm^{-1})	-777	-913	
$f_0^{(2)}$	(cm^{-1})	10,759	10,966	2,531
$f_0^{(3)}$	(cm^{-1})	-38,889	-38,132	-2,298
$f_0^{(4)}$	(cm^{-1})	67,082	63,289	2,299
$f_0^{(5)}$	(cm^{-1})	-49,903	-42,461	
$f_0^{(6)}$	(cm^{-1})	19,146	17,844	
$f_0^{(7)}$	(cm^{-1})	-12,519		
$f_0^{(8)}$	(cm^{-1})	7,252		
$f_1^{(1)}$	(cm^{-1})	-902	-2,086	-1,996
$f_1^{(2)}$	(cm^{-1})	-5,024		-2,444
$f_1^{(3)}$	(cm^{-1})			5,073
$f_1^{(4)}$	(cm^{-1})			-3,548
$f_{11}^{(0)}$	(cm^{-1})	57,391	62,332	56,195
$f_{11}^{(1)}$	(cm^{-1})	-1,857	-1,857	1,171
$f_{11}^{(2)}$	(cm^{-1})			-3,920
$f_{13}^{(0)}$	(cm^{-1})	-1,503	179	-1,596
$f_{13}^{(1)}$	(cm^{-1})			-8,058
$f_{13}^{(2)}$	(cm^{-1})			7,267
$f_{111}^{(0)}$	(cm^{-1})	8,073	8,036	3,370
$f_{113}^{(0)}$	(cm^{-1})			-855
$f_{113}^{(1)}$	(cm^{-1})			1,293
$f_{1111}^{(0)}$	(cm^{-1})	30,239	30,239	3,029
$\bar{\rho}_e$	(deg)	17.5	18.18	0
barrier	(cm^{-1})	16.5	21	0

Table 4.2: Acceptance function parameters for C_3 LM basis set in $1/bohr^2$

local modes		
μ	ν	$A_{\mu\nu}$
1	1	-19.6
1	2	-3.3
1	3	2.1
2	2	-19.6
2	3	2.1
3	3	8.6
$\rho = 0.41$		

4.3 Local mode basis function calculations for PES I

The successful application of linear combinations of LM basis functions, Eqs. (3.3) and (3.4), for H_2O trial wavefunctions suggested that accurate trial functions for C_3 could be constructed from linear combinations of these functions. Significant differences between the nearly rigid H_2O molecule and the floppy C_3 suggested that different basis functions could be required. In order to determine the utility of the LM basis functions for floppy C_3, VMC and CFQMC calculations were conducted using PES I.

4.3.1 Guiding and basis functions optimization

The guiding functions parameters $A_{\mu\nu}$ were optimized by minimizing the energy of the the ground state basis function, Eq. (3.3). These are given in Table 4.2. Table 4.3 contains the list of basis functions used. The set of functions used is quite unusual, The orders of the basis functions used, particularly for mode 3, are not incremental. Rather certain orders are skipped. Although the basis functions are not linearly dependent, the choice of basis functions and the size of the basis set was limited in preliminary VMC calculations by the onset of apparent linear dependencies

Table 4.3: Local Mode Basis set for C_3

basis function	n_1,n_2,n_3	basis function	n_1,n_2,n_3	basis function	n_1,n_2,n_3
1	0 0 0	16	1 0 4	31	0 4 0
2	1 0 0	17	0 3 0	32	0 3 2
3	0 1 0	18	0 2 2	33	0 2 4
4	0 0 2	19	0 1 4	34	0 1 6
5	2 0 0	20	0 0 6	35	0 0 8
6	1 1 0	21	4 0 0	36	0 1 8
7	1 0 2	22	3 1 0	37	1 0 8
8	0 2 0	23	3 0 2	38	0 0 10
9	0 1 2	24	2 2 0	39	0 0 12
10	0 0 4	25	2 1 2	40	2 0 5
11	3 0 0	26	2 0 4	41	0 2 5
12	2 1 0	27	1 3 0	42	2 2 2
13	2 0 2	28	1 2 2	43	4 0 4
14	1 2 0	29	1 1 4	44	0 4 4
15	1 1 2	30	1 0,6	45	2 2 4

in preliminary as more basis functions were included in the basis set. By skipping certain order functions the overall size of the basis set was maximized.

4.3.2 VMC and CFQMC calculations

The VMC ensemble of 16,384 walkers was propagated 20,000 steps with $\delta\tau = 1600$ to reach an equilibrium distribution of ψ_G^2. Afterwards the ensemble was propagated 10,000 steps with $\delta\tau = 1600$. The total sampling of the Hamiltonian and overlap matrix elements, Eqs. (2.12) and (2.13) was $M = 163,840,000$.[3] The VMC wavefunctions from the fourteen lowest energy states were used as trial wavefunctions for the CFQMC calculation. An ensemble of 16,384 walkers was propagated 20,000 steps with $\delta\tau = 2$ to equilibrate the ensemble and 10,000 steps with $\delta\tau = 2$ to sample the CFQMC matrix elements, Eq. (2.57). The matrix elements were computed at

[3]This calculation was performed also on a 512 node partition of a Thinking Machines CM5 parallel computer at the Army High Performance Supercomputer Research Center.

Table 4.4: VMC and CFQMC vibrational state energies using PES I and LM basis functions for C_3 $\left(^1\Sigma_g^+\right)$ in cm^{-1}. The experimental entries from Ref. [54] are not marked. The experimental entries from Ref. [55] are marked with †. VMC and CFQMC refer to the present calculations. The numbers in parentheses are the estimated statistical errors (one standard deviation). The * denotes unconverged energies (no observed plateau), which are upperbounds. MORBID refers to the calculations from Ref. [22]

(ν_1, ν_2, ν_3)	Expt.	VMC	CFQMC	MORBID
ZPE		1699.7(0.2)	1672.0(0.2)	
(0, 2, 0)	132.7993†	170(0)	138.9(0.3)	138
(0, 4, 0)	286.11	409(0)	284.4(0.3)	290
(0, 6, 0)	461.09	790(0)	458.7(0.3)	453
(0, 8, 0)	647.59	1360(0)	657.5(0.4)*	642
(0, 10, 0)	848.40	1735(0)	920.4(0.5)*	846
(1, 0, 0)	1226.6	1299(0)	1221.5(0.5)	1220
(1, 2, 0)	1406.5	2052(1)	1418.1(0.5)	1406
(1, 4, 0)	1592.5	2267(1)	1588.3(0.5)	1587
(1, 6, 0)	1787.5	2446(1)	1793(1)*	1780

projection time intervals of 12 H^{-1} for $\tau = 0$ H^{-1} to $\tau = 900$ H^{-1}.[4] The VMC and CFQMC vibrational state energies are given in Table 4.4, using the typical linear molecule notation $(\nu_1, \nu_2^{l_2}, \nu_3)$ where ν_1 is the symmetric stretch mode, $\nu_2^{l_2}$ is the bend mode, and ν_3 is the asymmetric stretch mode. The quantum number, superscript l_2, shall be omitted, since $l_2 = 0$ for all calculations reported here.

The trial wavefunctions for this C_3 calculation are relatively inaccurate representations of the vibrational wavefunctions. The ground state trial wavefunction energy is at least 27 cm^{-1} higher than the true ground state energy. The (0,2,0) state trial wavefunction energy is at least 32 cm^{-1} higher than the exact (0,2,0) state energy. In total 7 states converged in the CFQMC calculation. This suggests that either

[4] After 684 H^{-1} all block and ensemble overlap matrices became singular and the CFQMC eigenvalue equations became insoluble.

the choice of the Local mode basis functions for C_3 or the the method of optimizing the parameters is unsuitable.

4.3.3 Improved guiding and basis functions optimization

An alternative method of optimizing the non-linear parameters of the basis functions was implemented. The sum of the energies of the several lowest states resulting from a multistate VMC calculation with a minimal basis set,

$$E(A_{\mu\nu}) = \Lambda_0(A_{\mu\nu}) + \Lambda_1(A_{\mu\nu}) + \Lambda_2(A_{\mu\nu}) + \Lambda_3(A_{\mu\nu}) + \Lambda_4(A_{\mu\nu}), \qquad (4.7)$$

are minimized with respect to the non-linear basis set parameters. This energy functional resembles the ideal energy functional which is the sum of the energy of the eesired states resulting from a large basis set VMC calculation. A simple procedure for locating the minimum in $E(A_{\mu\nu})$ was implemented using a basis set of 20 functions such that $n_1 + n_2 + n_3 \leq 3$. Individual parameters were varied until a minimum was found in $E(A_{\mu\nu})$ with respect to the varied parameter. This process was repeated until the multi-dimensional minimum was located. Perturbations in each parameter were made to double check the final minimum. Using this optimization procedure, nearly exact VMC energies were obtained for states (0,0,0) and (0,2,0), and much improved energies for many other states. The values of $A_{\mu\nu}$ are given in Table 4.5. The expansion of the basis set was again hindered by the onset of apparent linear dependencies. The basis set was expanded to include 35 functions basis with $n_1 + n_2 + n_3 \leq 4$. Inclusion of further functions or redistributing the functions to provide higher order functions in the ν_3, the floppy bending mode, led to apparent linear dependencies.

4.3.4 VMC calculation with re-optimized basis functions

The VMC ensemble of 64 walkers was propagated 8,000 steps with $\delta\tau = 1600$ to reach an equilibrium distribution of ψ_G^2. Afterwards the ensemble was propagated 8,000 steps with $\delta\tau = 1200$. The total sampling of the Hamiltonian and overlap matrix elements, Eqs. (2.12) and (2.13), of the re-optimized basis set, was $M = 512,000$. The VMC vibrational state energies are given in Table 4.6.

Table 4.5: Acceptance function parameters for C_3 SMOL basis set in $1/bohr^2$ optimized with respect to $E(A_{\mu\nu})$

local modes		
μ	ν	$A_{\mu\nu}$
1	1	-43.0
1	2	-7.0
1	3	4.0
2	2	-43.0
2	3	4.0
3	3	6.0
$\rho = 0.15$		

Table 4.6: VMC vibrational state energies using PES I and re-optimized Local Mode basis functions for C_3 $\left(^1\Sigma_g^+\right)$ in cm^{-1}. The experimental entries from Ref. [54] are not marked. The experimental entries from Ref. [55] are marked with †. VMC refers to the present calculations. The numbers in parentheses are the estimated statistical errors (one standard deviation). MORBID refers to the calculations from Ref. [22]

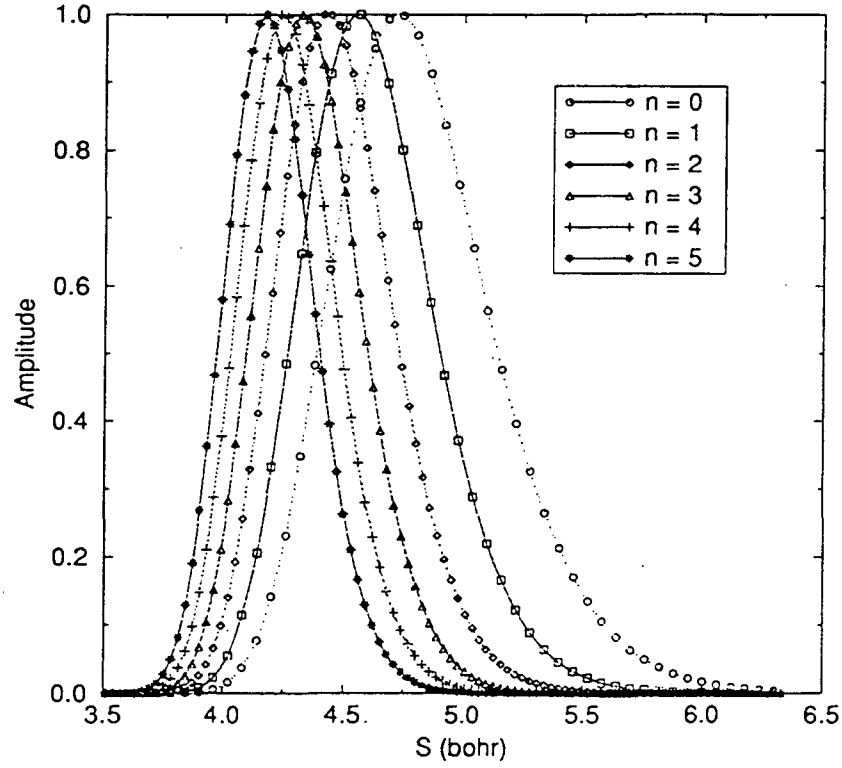
(ν_1, ν_2, ν_3)	Expt.	VMC	MORBID
ZPE		1667.7(0.3)	
(0, 2, 0)	132.7993†	140(0)	138
(0, 4, 0)	286.11	292(1)	290
(0, 6, 0)	461.09	514(3)	453
(0, 8, 0)	647.59	885(8)	642
(0, 10, 0)	848.40	1525(?)	846
(1, 0, 0)	1226.6	1217(1)	1220
(1, 2, 0)	1406.5	1425(2)	1406
(1, 4, 0)	1592.5	1670(5)	1587
(1, 6, 0)	1787.5	2026(1)	1780
(0, 0, 1)	2040.0192	2065(14)	2037

The VMC ground state energy obtained from the re-optimized non-linear parameters is slightly lower than the preceding CFQMC energies, which suggests that the CFQMC energies were upper bounds to the exact energy. The VMC energies for states (0,2,0), (0,4,0), and (1,0,0) are less than or equivalent to the converged variational energies reported by Jensen[22]. However, the accuracy of the VMC energies is poor for higher excited states because few high order basis functions could be included in the basis set (see above). Furthermore, several energetically low-lying, yet highly excited, bending states were over-looked in the latter VMC calculations. Therefore, further VMC and CFQMC calculations were not conducted. Local mode basis functions are not fully suitable for C_3.

4.4 Simplified Morse oscillator basis function calculations for PES I

Morse oscillator functions are often used to describe anharmonic motion. The Morse potential has only a finite number of bound states. Thus, Morse oscillator functions do not form a complete set without the inclusion of continuum states. Moreover, these functions are defined in terms of the associated Laguerre polynomials that have indices coupled in an way which makes these functions hard to use [56]; the recurrence relations are not straightforward. Tennyson and Sudcliffe defined the Morse oscillator-like basis functions to circumvent the completeness problem and avoid the use of associated Laguerre polynomials while retaining the anharmonicity of the Morse oscillator functions[47, 56, 57]. These function are based on the Morse oscillator functions from which they receive the name, Morse oscillator-like. From studies using the LM basis function form (3.4), we found that single-term basis functions are more readily sampled with Monte Carlo than polynomials. Therefore, we simplified the Morse oscillator-like basis functions by replacing the Laguerre polynomials in X_ν with single-term functions of X_ν^n where $X_\nu = \alpha \exp -\beta_\nu \Delta S_\nu$; α, and β are the familiar Morse parameters. No flexibility is lost by this simplification because Laguerre polynomials of order n can be represented by a linear combination

Figure 4.1: Simplified Morse oscillator-like basis Functions; Eq. (4.8)



of single-term functions of order n. Further, we retained the harmonic couplings of the Bernu et al. form yielding simplified Morse oscillator-like basis functions with harmonic couplings (SMOL-HC) given by

$$f_m = \left(\prod_{\nu=1}^{3} X_\nu^{(\alpha_\nu+1+2n_\nu(m))/2} \right) \exp\left(\sum_{\nu=1}^{3} -\frac{1}{2} X_\nu - \sum_{\nu=1}^{3} \sum_{\mu<\nu} \Delta S_\nu A_{\mu\nu} \Delta S_\mu \right) \tag{4.8}$$

where $A_{\mu,nu}$ are harmonic coupling coefficients.[5] SMOL functions for a single mode molecule are shown in Fig. 4.1. The functions continue to be distinct from one

[5]Ref. [58] has a typographical error. The leading minus sign is missing in the argument of the exponent in Eq. (2) in the paper. Eq. 4.8 is correct.

another as the order increases. This suggests that large basis sets may be used without apparent linear dependencies.

4.4.1 Guiding and basis functions optimization

For C_3, the SMOL-HC basis function form (4.8) leads to 9 non-linear variational parameters which can be reduced by the symmetry of modes S_1 and S_2 to 6. We optimized the non-linear parameters using small basis set VMC calculations. The variational procedure we employed is as follows. The first step in the optimization is the initial parameter specification. The relationship between the Morse parameters and the well depth and width determines the best initial value of the α and β parameters. A reasonable initial value of the harmonic coupling parameters is $A_{\mu\nu} = 0$, i.e., the uncoupled limit. Second, the VMC energies and standard deviations are obtained for a basis set of 20 functions with $n_1(\alpha) + n_2(\alpha) + n_3(\alpha) \leq 3$. Modes 1 and 2 were evenly represented in the basis set because they are symmetrical with respect to the PES. The basis set was limited to 20 functions total for computational efficiency. Third, one varies a non-linear parameter associated with the high energy modes (e.g. stretching in the present case), and calculates the VMC energies and standard deviations until a minimum in the sum of VMC energies of the first five states is found. Fourth, one varies the remaining parameters iteratively, beginning with the remaining high-energy modes and progressing to the low-energy modes, and evaluates the VMC energies and standard deviations until the minima are found as in step three. In varying the non-linear basis set parameters, the region where the basis functions are large typically changes. Therefore, a guiding function with broad spatial extent is desirable. Once the optimal values of the non-linear variational parameters are determined, the basis set is expanded to obtain accurate trial wave functions for the several lowest energy eigenstates. The optimized acceptance function parameters are given in 4.7 and the optimized SMOL parameters are given in table 4.8.

Table 4.7: Acceptance function parameters for C_3 in $1/(\text{bohr})^2$

local modes		
μ	ν	$A_{\mu\nu}$
1	1	-13.0
1	2	0.0
1	3	0.0
2	2	-13.0
2	3	0.0
3	3	2.8
$\rho = 0.25$		

Table 4.8: SMOL parameters for C_3. α are unitless and β are in $1/(\text{bohr})$

Mode	α	β
1	10	4
2	10	4
3	4	2

4.4.2 VMC calculation

As suggested by the mutual distinctiveness of the SMOL basis functions, apparent linear dependencies did not arise in the SMOL representation until the basis was expanded beyond 130 functions. The SMOL basis set was expanded to include 130 functions such that $n_1 + n_2 \leq 4$ for $n_3 \leq 5$, $n_1 + n_2 \leq 3$ for $6 < n_3 < 8$, $n_1 + n_2 \leq 2$ for $n_3 = 9$ and $n_1 \leq 1$, $n_2 \leq 1$ for $n_3 = 10$. The VMC ensemble of 80 walkers was propagated 20,000 steps with $\delta\tau = 1600$ to reach an equilibrium distribution of ψ_G^2 and afterwards propagated 36,000 steps with $\delta\tau = 1600$ to sample the Hamiltonian and overlap matrix elements, Eqs. (2.12) and (2.13). The total sampling of the matrix elements was $M = 2,880,000$. The wavefunctions were assigned state labels (ν_1, ν_2, ν_3) graphically, according to the number of nodes observed in the wavefunctions.

Figures 4.2 and 4.3 show the amplitude of the wavefunctions with respect to the symmetric stretch coordinate $\Delta S_{sym} = \frac{1}{2}(\Delta S_1 + \Delta S_2)$ with $\Delta S_1 = \Delta S_2$ and with the bond angle fixed at π for the lowest ten and highest seven VMC states. In Fig. 4.2 three wavefunctions have a single node each and seven wavefunctions have no nodes in the symmetric stretch mode. The maximum of the amplitude tends toward shorter bond lengths with increasing excitation of the bending mode. In Fig. 4.3 all seven wavefunction seem to have nodes. States (1,6,0) and (1,8,0) have a very distinct single node in the symmetric stretch coordinate. States (0,14,0) and (0,16,0) have fluctuations about zero amplitude in the contacted and extended regions of the symmetric stretch coordinates. These fluctuations show that some higher order stretching states are mixed into these VMC wavefunctions. However, the amplitude of the fluctuations is small and the VMC energies are too low to conclude that these states have doubly excited symmetric stretches. Therefore, the apparent nodes are spurious. States (0,0,1), (0,2,1) and (0,4,1) also have fluctuations in the contracted and extended regions of the symmetric stretch coordinates. These wavefunction plots are taken near a node in the asymmetric coordinate. The maximum amplitude in these wavefunctions is very small, relative to the amplitudes for pure bending and symmetric-stretch with bending states.

Figure 4.2: Symmetric stretch of the first 10 SMOL trial wavefunctions for C_3. Each wavefunction is scaled by $\log|\max\Psi_i|$.

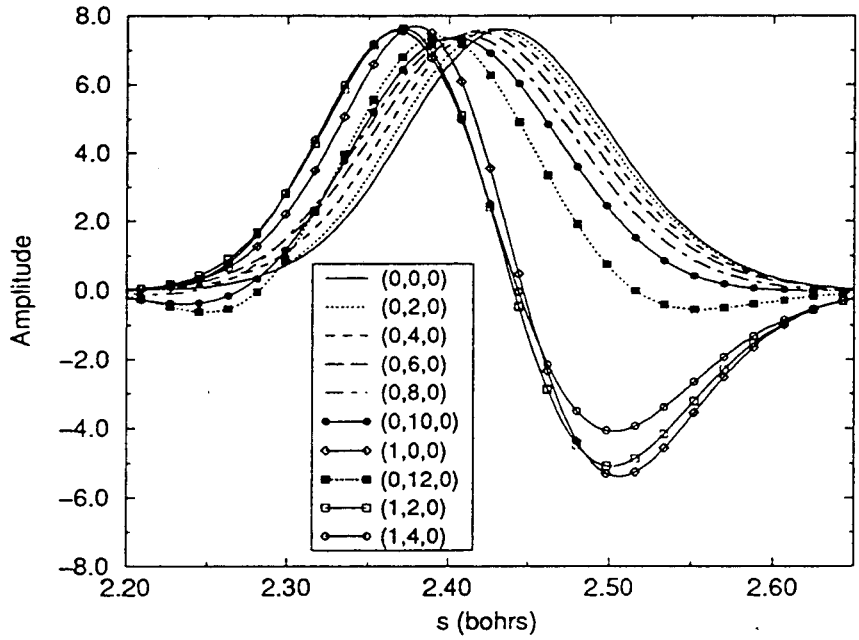


Figure 4.3: Symmetric stretch of the last 7 SMOL trial wavefunctions for C_3. Each wavefunction is scaled by $\log|\max\Psi_i|$.

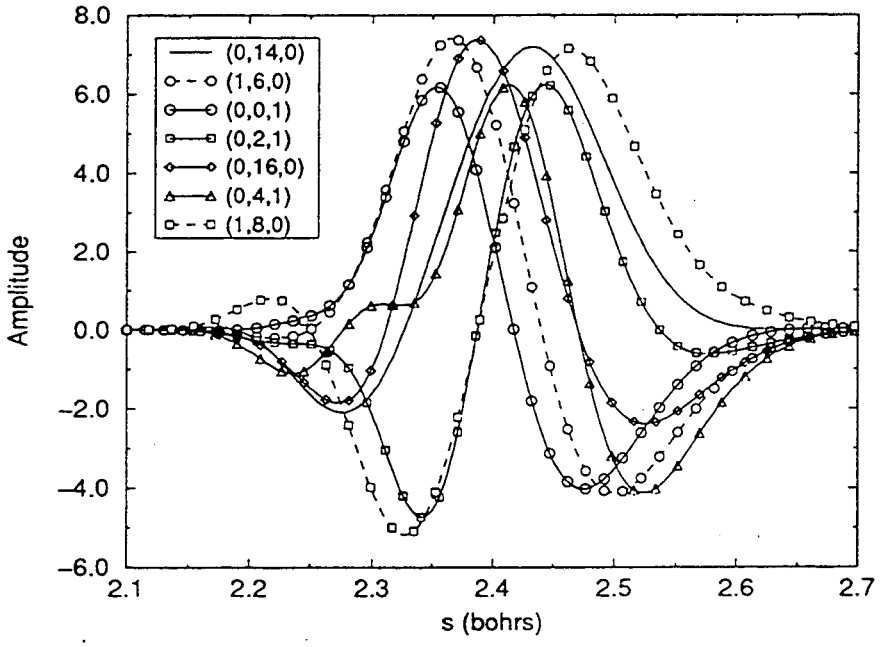
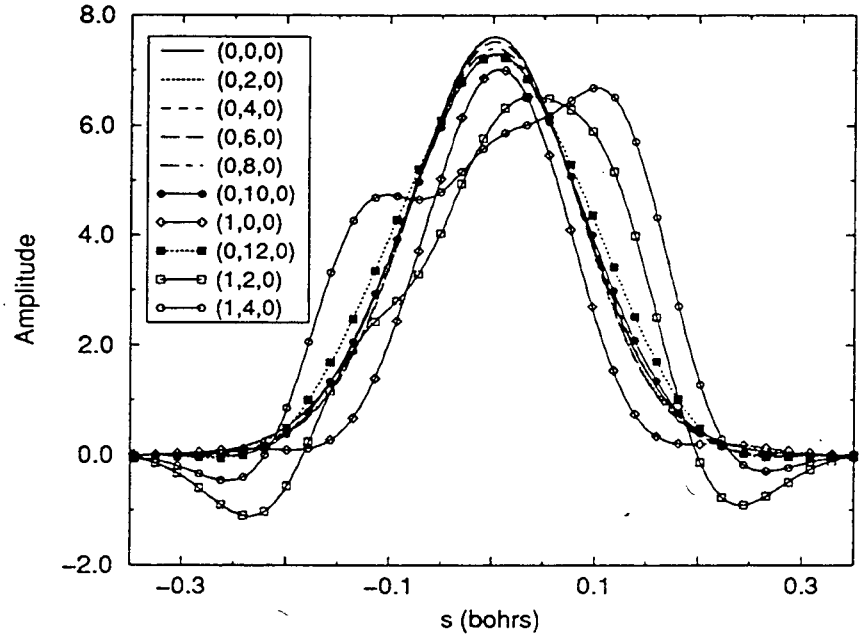


Figure 4.4: Asymmetric stretch of the first 10 SMOL trial wavefunctions for C_3 Each wavefunction is scaled by $\log|\max\Psi_i|$.



Figs. 4.4 and 4.5 show the amplitude of the wavefunctions with respect to the asymmetric stretch coordinate $\Delta S_{sym} = \frac{1}{2}(\Delta S_1 - \Delta S_2)$ with $\Delta S_1 = -\Delta S_2$ and with the bond angle fixed at π for the lowest ten and highest seven VMC states. In Fig. 4.4 eight of the 10 wavefunctions plotted have zero nodes in the asymmetric stretch mode. The remaining two wavefunctions seem to have nodes in the asymmetric stretch mode near ± 0.2. bohrs. These nodes are spurious, resulting from the finite basis approximation. Furthermore, both states have a node in the symmetric stretch mode. The asymmetric stretch wavefunction plots for these states are, in the symmetric stretch coordinate, near the node. All ten wavefunctions belong to states having zero excitations in the asymmetric stretch mode.

Figures 4.6 and 4.7 show the amplitude of the wavefunctions with respect to the bending angle with bond lengths fixed at 2.44 bohrs and 2.38 bohrs, respectively.

Figure 4.5: Asymmetric stretch of the last 7 SMOL trial wavefunctions for C_3 Each wavefunction is scaled by $\log|\max \Psi_i|$.

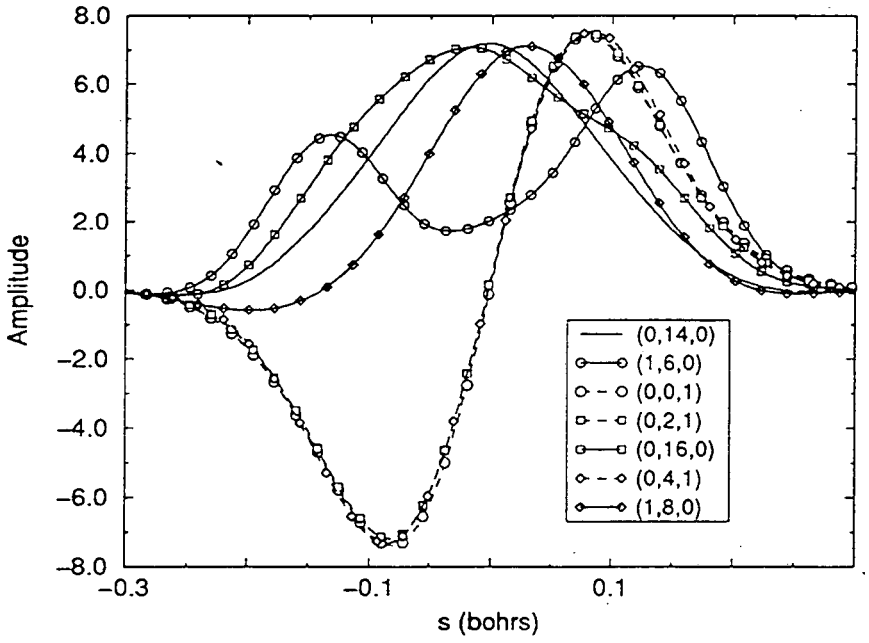
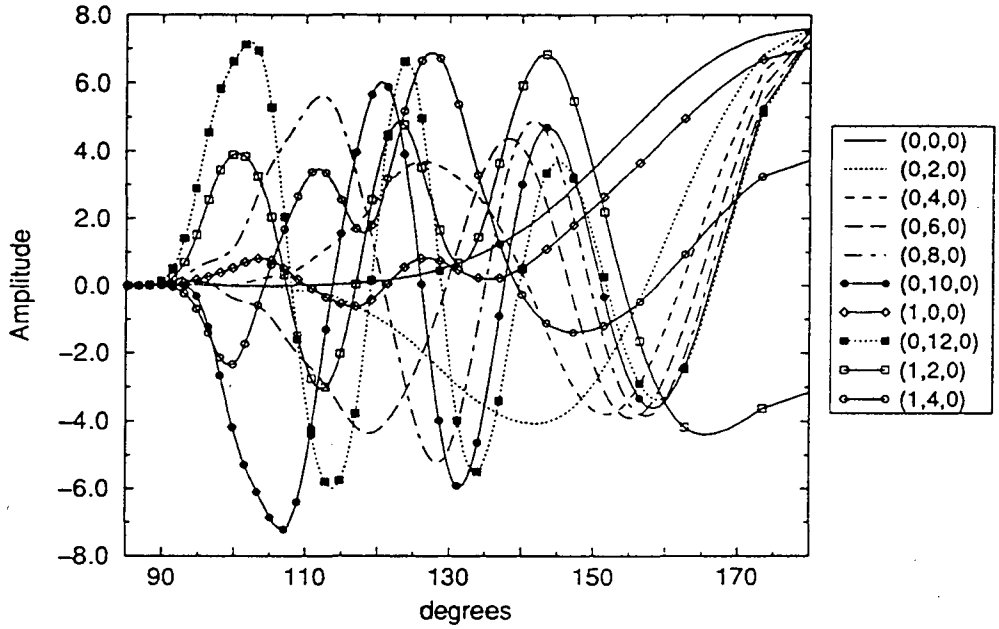


Figure 4.6: Bend of the first 10 SMOL trial wavefunctions for C_3 with $s_1 = s_2 = 2.44$ bohrs. Each wavefunction is scaled by $\log|\max \Psi_i|$.



Two of the wavefunctions have zero nodes in the bending mode, two have one node each, two have two nodes each, one has three nodes, one has four nodes, one has five nodes and one has six nodes. The wavefunctions of the three states having an excitation in the symmetric stretch are narrower, tending closer to the linear configuration, than their analogues in the manifold of states having no excitation in the symmetric stretch. This agrees with the observation of Northrup, *et al.* from their semi-rigid bender analysis of C_3 that excitations in the symmetric stretch make the effective bending potential narrower[59].[6] In Fig. 4.6 the wavefunctions plots, (1,2,0) and (1,4,0), are, in symmetric stretch coordinate, near nodes. They have spurious nodes resulting from the finite basis approximation. In Fig. 4.7 the plots of two

[6]Ref. [58] has a typographical error. The stated effect of the excitation of the symmetric and asymmetric stretches on the effective bending potential observed by Northrup, *et al.* are reversed. The effect observed by Northrup, *et al.* in the semi-rigid bender analysis and Brown, Glauser and Lester in the VMC calculations agree as reported here.

Figure 4.7: Bend of the first 10 SMOL trial wavefunctions for C_3 with $s_1 = s_2 = 2.38$ bohrs. Each wavefunction is scaled by $\log|\max \Psi_i|$.

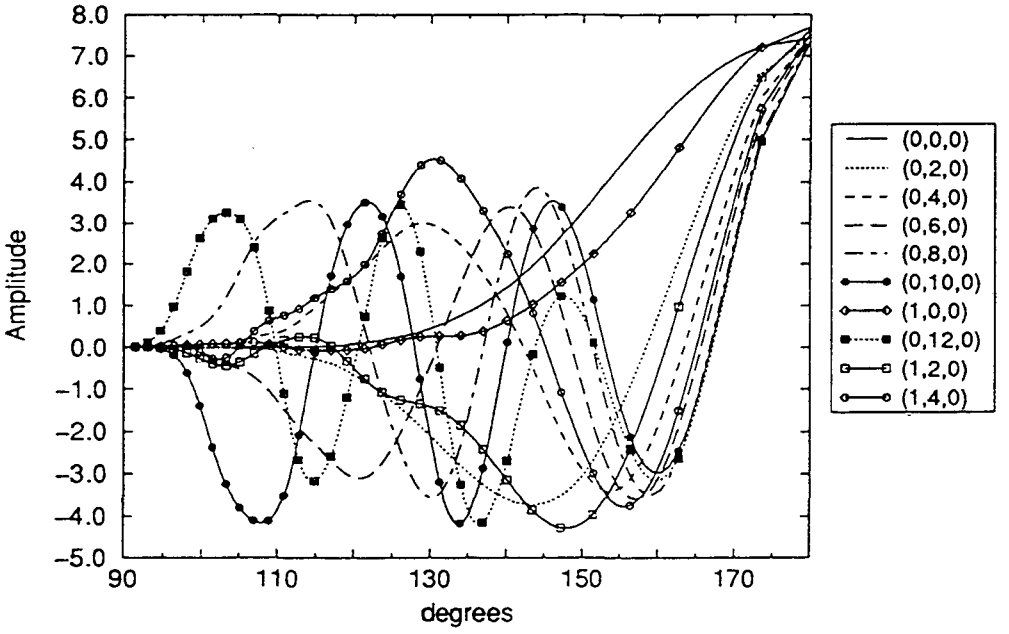
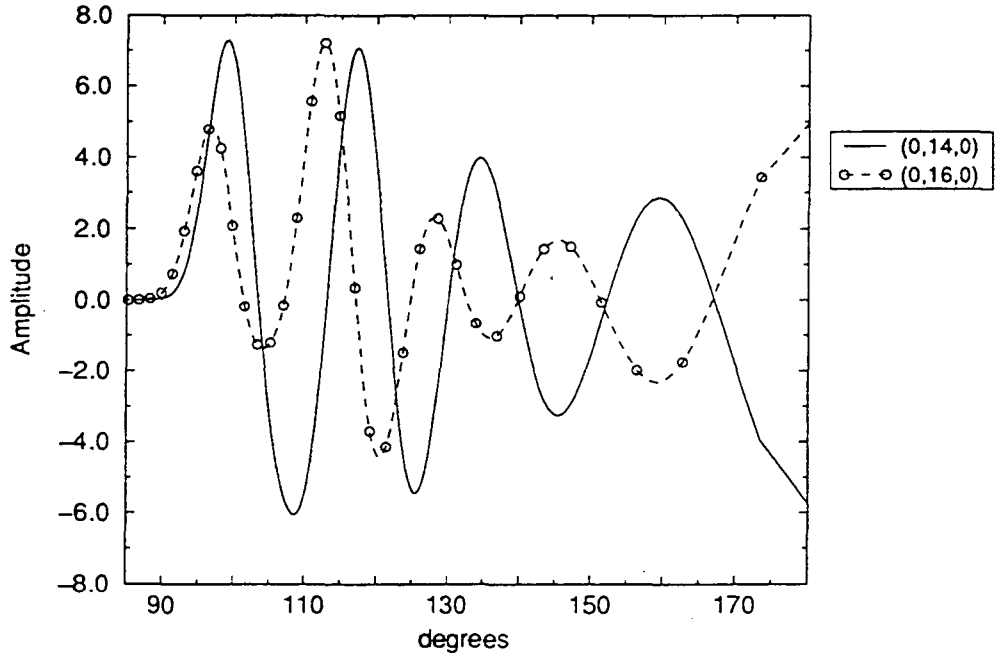


Figure 4.8: Bend of 2 of the last 7 SMOL trial wavefunctions for C_3 with $s_1 = s_2 = 2.44$ bohrs. Each wavefunction is scaled by $\log|\max \Psi_i|$.



wavefunctions are near a maximum in the symmetric-stretch node. Therefore, the indicative nodal structure in the bending mode is well described.

The SMOL VMC vibrational state energies are given in Table 4.9. The SMOL representation is slightly less accurate than the LM representation for the lowest lying vibrational states. The LM VMC energies for states (0,0,0), (0,2,0) and (0,4,0) lower than the analogous SMOL VMC energies. The SMOL representation is more accurate for more highly excited vibrational states. The SMOL VMC energies for states (0,10,0), (1,6,0), (0,0,1), and (0,2,1) are lower than the corresponding LM VMC energies. Since the SMOL basis set are not prone to have near linear dependencies, a more complete basis set may be used. Hence, the SMOL representation provides trial functions for more vibrational states than the LM representation. Therefore, the SMOL representation provided estimates for states (0,12,0), (0,0,2),

Figure 4.9: Bend of 1 of the last 7 SMOL trial wavefunctions for C_3 with $s_1 = s_2 = 2.38$ bohrs. Each wavefunction is scaled by $\log|\max \Psi_i|$.

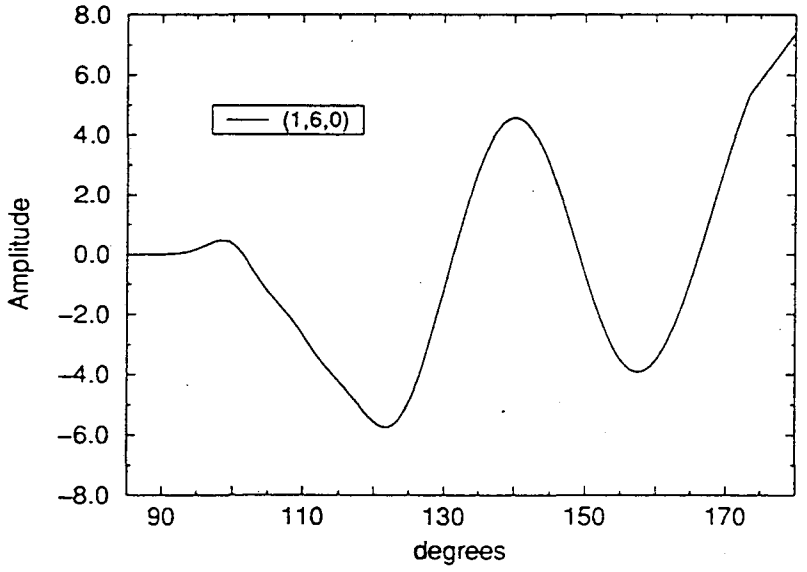
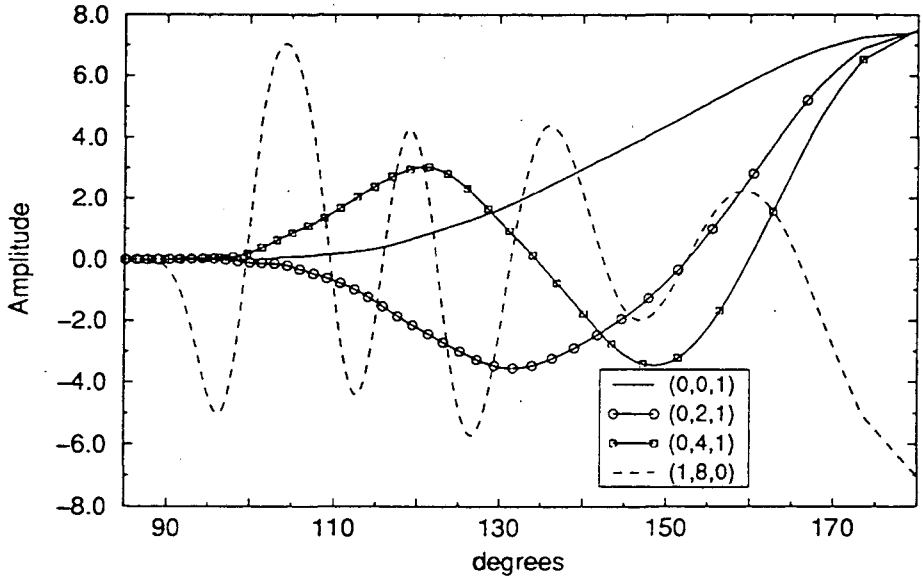


Figure 4.10: Bend of 4 of the last 7 SMOL trial wavefunctions for C_3 with $\Delta s_1 = -\Delta s_2 = 0.76$ bohrs. Each wavefunction is scaled by $\log|\max \Psi_i|$.



(0,4,1) and (0,6,1) which were lacking in the LM VMC calculations due to the small basis set size.

4.4.3 CFQMC calculation

The VMC wavefunctions of the twenty-two lowest energy states were used as trial wavefunctions for the CFQMC calculation. An ensemble of 80 walkers, 4 blocks of twenty walkers each, positioned at the final coordinates of the VMC walkers, was propagated 40,000 steps with $\delta\tau = 2$ to ensure the equilibration of the ensemble and 100,000 steps with $\delta\tau = 2$ to sample the CFQMC matrix elements, Eq. (2.57). The total sampling of the matrix elements was $M = 8,000,000$. The matrix elements were computed at projection time intervals of 12 H^{-1} for $\tau = 0$ H^{-1} to $\tau = 640$ H^{-1}. After 360 H^{-1} all block and ensemble overlap matrices became singular and the CFQMC eigenvalue equations became insoluble. Because the statistical error in the matrix elements of the highest energy trial functions dominates the error in the vibrational energies, we eliminated the 5 highest energy trial functions and repeated the CFQMC calculation An ensemble of 100 walkers, 5 blocks of 20 walkers each, positioned at the final coordinates of VMC calculation, was propagated 20,000 steps with $\delta\tau = 2$ to ensure the equilibration of the ensemble and 100,000 steps with $\delta\tau = 2$ to sample the CFQMC matrix elements. The total sampling of the matrix elements was $M = 10,000,000$.[7] The matrix elements were computed at projection time intervals of 12 H^{-1} for $\tau = 0$ H^{-1} to $\tau = 640$ H^{-1}. The block and ensemble overlap matrices became singular and the CFQMC eigenvalue equations became insoluble after 468 H^{-1}, 108 H^{-1} later than in the 22 trial wavefunction calculation.

The Λ_k for six lowest energy vibrational states from the 17 trial wavefunction CFQMC calculation are shown in Fig. 4.11. The identities of the states are clearly known from the state labels assigned to the VMC wavefunctions, since the eigenvalues do not re-order while converging. Improvement over the VMC energies is found for each state. Convergence is shown for states (0,0,0) to (0,8,0), but not for state (0,10,0).

[7]These calculations were conducted on a Cray-XMP at the National Energy Super Computer Research Center at Lawrence Livermore National Laboratory.

Table 4.9: VMC vibrational state energies for C_3 using PES I and SMOL basis functions for C_3 $\left(^1\Sigma_g^+\right)$ in cm^{-1}. The experimental entries from Ref. [54] are not marked. The experimental entries from Ref. [55] are marked with †. The experimental entries from Ref. [60] are marked with ‡. LM-VMC and SMOL-VMC refer to the present calculations. The numbers in parentheses are the estimated statistical errors (one standard deviation). MORBID refers to the calculations from Ref. [22]

(ν_1, ν_2, ν_3)	Expt.	LM-VMC	SMOL-VMC	MORBID
ZPE		1667.7(0.3)	1669.5(0.7)	
(0, 2, 0)	132.7993†	140(0)	147(1)	138
(0, 4, 0)	286.11	292(1)	306(1)	290
(0, 6, 0)	461.09	514(3)	506(2)	453
(0, 8, 0)	647.59	885(8)	751(2)	642
(0, 10, 0)	848.40		1048(3)	846
(0, 12, 0)	1061.96		1366(5)	1060
(1, 0, 0)	1226.6	1217(1)	1222(2)	1220
(1, 2, 0)	1406.5	1425(2)	1450(2)	1406
(1, 4, 0)	1592.5	1670(5)	1676(2)	1587
(1, 6, 0)	1787.5	2026(1)	1848(8)	1780
(2, 0, 0)	2434.9‡		4138(10)	2438
(0, 0, 1)	2040.0192	2065(14)	2037(2)	2037
(0, 2, 1)	2133.8876†		2140(2)	
(0, 4, 1)			3966(5)	
(0, 6, 1)			4151(10)	

The Λ_k for the eleven higher energy vibrational states from the 17 trial wavefunction CFQMC calculation are shown in Fig. 4.12. The identities of (1,2,0) and (1,6,0) are clearly known from the state labels assigned to the VMC wavefunctions, since their Λ_k do not re-order while converging. The identity of the remaining 9 states is generally less clearly known from VMC state labels since the converging eigenvalues re-order. States (1,0,0), (0,12,0), (1,4,0), (0,14,0), (0,0,1) and (0,2,1) each re-order once. The convergence properties allow these states to be positively identified. The highly excited bend states (0,12,0) and (0,14,0) have a large reduction in Λ_k with respect to projection time because the trial functions were relatively poor approximations to the exact eigenstates having large contributions from the quickly decaying, high-energy eigenstates; see Eq. 2.48. The trial wavefunctions for states (1,0,0) and (1,4,0) are more accurate and have less reduction in their associated Λ_k with respect to projection time. States (0,0,1) and (0,2,1) re-order with (0,16,0) which has the greatest reduction in its associated Λ_k with respect to projection time. The values of $\Lambda_k(0)$ for states (0,4,1), (1,8,0) and (0,16,0) are very similar. Within the first 96 H^{-1} of projection time these states underwent four crossings. Again, the assignments are certain because of the convergence properties of the state are very distinct. State (0,4,1) has large reduction in its associated Λ_k. State (1,8,0) has a larger reduction in its associated Λ_k and, as stated above, (0,16,0) has the greatest reduction. The convergence behavior depends most on the excitation of the bend. States with only a few excitations of the bend converged faster and with a smaller reduction in the value of $\Lambda_k(0)$ than states with higher number of excitations of the bend without regard to the total energy or the excitation of the stretches. This is a consequence of the trial functions being less accurate with the greater bend excitations.

The CFQMC vibrational state energies are given in Table 4.10. The fourteen converged vibrational states agreed to within a few cm^{-1} of the MORBID energies and the experimental values.

Figure 4.11: Λ_k for the 6 lowest energy states of C_3. The labels (ν_1,ν_2,ν_3) indicate the vibrational state of the Λ_k

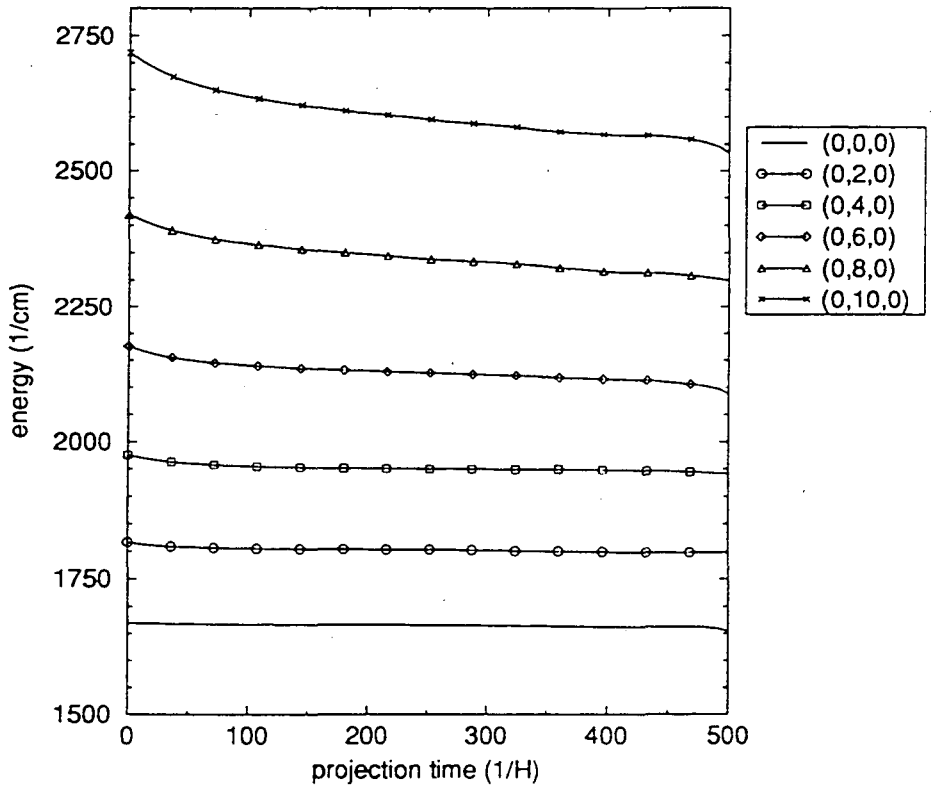


Figure 4.12: Λ_k for 11 higher energy states of C_3. The labels (ν_1,ν_2,ν_3) indicate the vibrational state of the Λ_k and x indicate an intersections of two λ_k. The label (ν_1,ν_2,ν_3)x(ν_1',ν_2',ν_3') indicates that the Λ_k associated with state (ν_1,ν_2,ν_3), after it intersects $\Lambda_{k'}$, is associated with state (ν_1',ν_2',ν_3').

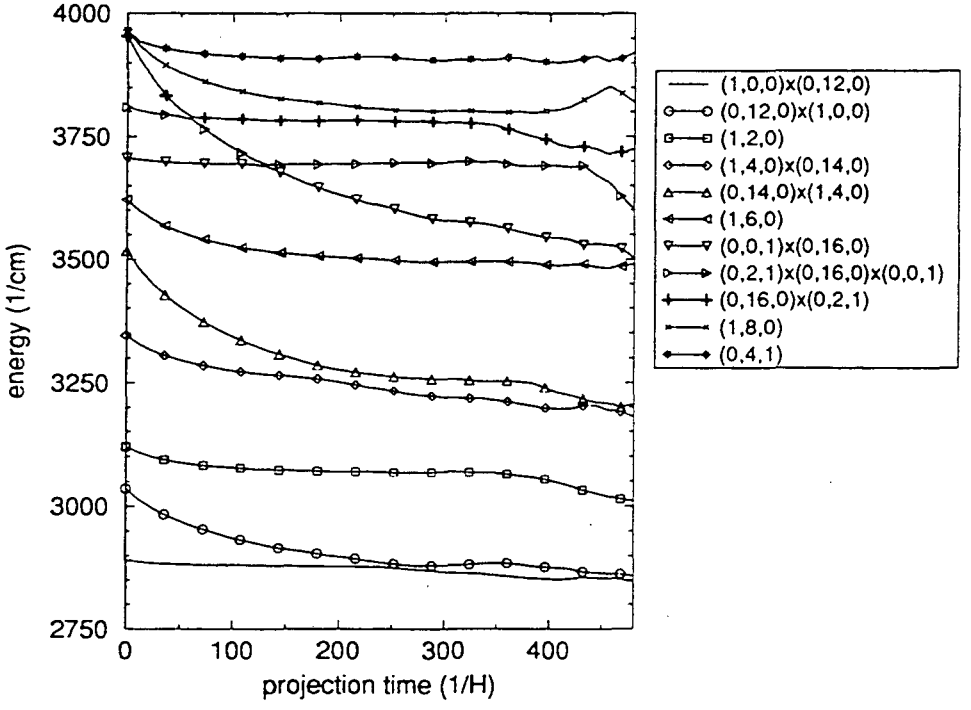


Table 4.10: CFQMC vibrational state energies using PES I and SMOL trial wavefunctions for C_3 $\left(^1\Sigma_g^+\right)$ in cm^{-1}. The experimental entries from Ref. [54] are not marked. The experimental entries from Ref. [55] are marked with †. The experimental entries from Ref. [60] are marked with ‡. VMC and CFQMC refer to the present calculations. The numbers in parentheses are the estimated statistical errors (one standard deviation). The * denotes unconverged energies (no observed plateau), which are upperbounds. MORBID refers to the calculations from Ref. [22]

(ν_1, ν_2, ν_3)	Expt.	VMC	CFQMC	MORBID
ZPE		1669.5(0.7)	1666.7(0.4)	
(0, 2, 0)	132.7993†	147(1)	136.9(0.9)	138
(0, 4, 0)	286.11	306(1)	283.4(0.9)	290
(0, 6, 0)	461.09	506(2)	457(3)	453
(0, 8, 0)	647.59	751(2)	647(11)	642
(0, 10, 0)	848.40	1048(3)	899(50)*	846
(0, 12, 0)	1061.96	1366(5)	1153(99)*	1060
(1, 0, 0)	1226.6	1222(2)	1210(2)	1220
(1, 2, 0)	1406.5	1450(2)	1401(1)	1406
(1, 4, 0)	1592.5	1676(2)	1585(4)	1587
(1, 6, 0)	1787.5	1848(8)	1796(25)	1780
(2, 0, 0)	2434.9‡	4138(10)	2393(13)	2438
(0, 0, 1)	2040.0192	2037(2)	2027(1)	2037
(0, 2, 1)	2133.8876†	2140(2)	2114(1)	
(0, 4, 1)		3966(5)	2238(4)	
(0, 6, 1)		4151(10)	2417(11)	

4.5 Simplified Morse oscillator basis function calculations for PES II

In order to gauge the semi-emperical refinement of PES I, VMC and CFQMC calculations were conducted using PES II and the VMC basis set from the PES I calculation. These parameters are shown in Tables 4.7 and 4.8; the basis set included 130 SMOL functions such that $n_1 + n_2 \leq 4$ for $n_3 \leq 5$, $n_1 + n_2 \leq 3$ for $6 < n_3 < 8$, $n_1 + n_2 \leq 2$ for $n_3 = 9$ and $n_1 \leq 1$, $n_2 \leq 1$ for $n_3 = 10$. A VMC ensemble of 80 walkers was propagated 20,000 steps with $\delta\tau = 1600$ to reach an equilibrium distribution of ψ_G^2 and afterwards propagated 50,000 steps with $\delta\tau = 1600$ to sample the Hamiltonian and overlap matrix elements, Eqs. (2.12) and (2.13). The total sampling of the matrix elements was $M = 4,000,000$. The VMC wavefunctions from the seventeen lowest energy states were used as trial wavefunctions for the CFQMC calculation. An ensemble of 80 walkers, 4 blocks of twenty walkers each, positioned at the final coordinates of the VMC walkers, was propagated 40,000 steps with $\delta\tau = 2$ to ensure the equilibration of the ensemble and 100,000 steps with $\delta\tau = 2$ to sample the CFQMC matrix elements, Eq. (2.57). The total sampling of the matrix elements was $M = 8,000,000$. The matrix elements were computed at projection time intervals of 12 H^{-1} for $\tau = 0$ H^{-1} to $\tau = 300$ H^{-1}. At 300 H^{-1} all block and ensemble CFQMC eigenvalue equations were soluble, but the statistical error was large for most states.

Only eight vibrational states converged and they were assigned state labels $(\nu_1, \nu_2^{l_2}, \nu_3)$ graphically, according to the number of nodes observed in the trial wavefunctions. The CFQMC vibrational state energies are given in Table 4.11. The ground state energy computed from PES II is 85 cm^{-1} higher energy that the ground state computed from PES I. Likewise, the computed energy spacings are greater for PES II than PES I.

Table 4.11: VMC and CFQMC vibrational state energies using PES II and SMOL basis functions for C_3 $\left(^1\Sigma_g^+\right)$ in cm^{-1}. The experimental entries from Ref. [54] are not marked. The experimental entries from Ref. [55] are marked with †. VMC and CFQMC refer to the present calculations using PES II. CFQMC* refers to calculations using PES I. The numbers in parentheses are the estimated statistical errors (one standard deviation). The * denotes unconverged energies (no observed plateau), which are only upperbounds. MORBID refers to the calculations from Ref. [7]

(ν_1, ν_2, ν_3)	Expt.	VMC	CFQMC	CFQMC*
ZPE		1752(1)	1751(1)	1666.7(0.4)
(0, 2, 0)	132.7993†	185(1)	173(2)	136.9(0.9)
(0, 4, 0)	286.11	425(2)	385(2)	283.4(0.9)
(0, 6, 0)	461.09	756(4)	643(6)*	457(3)
(0, 8, 0)	647.59	1223(4)	984(9)*	647(11)
(1, 0, 0)	1226.6	1281(2)	1272(3)	1210(2)
(1, 2, 0)	1406.5	1547(4)	1480(6)	1401(1)
(1, 4, 0)	1592.5	1840(4)	1762(15)	1585(4)
(0, 0, 1)	2040.0192†	2119(3)	2108(4)	2027(1)
(0, 2, 1)	2133.8876†	2294(5)	2251(3)	2114(1)

4.6 Simplified Morse oscillator basis function calculations for PES III

In order to compare the PES III, with PES I and PES II, calculations were conducted using PES III and the VMC basis set from PES the I calculation. These parameters are shown in Tables 4.7 and 4.8; the basis set included 130 SMOL functions such that $n_1 + n_2 \leq 4$ for $n_3 \leq 5$, $n_1 + n_2 \leq 3$ for $6 < n_3 < 8$, $n_1 + n_2 \leq 2$ for $n_3 = 9$ and $n_1 \leq 1$, $n_2 \leq 1$ for $n_3 = 10$. A VMC ensemble of 80 walkers was propagated 20,000 steps with $\delta\tau = 1600$ to reach an equilibrium distribution of ψ_G^2 and afterwards propagated 36,000 steps with $\delta\tau = 1600$ to sample the Hamiltonian and overlap matrix elements, Eqs. (2.12) and (2.13). The total sampling of the matrix elements was $M = 2,880,000$. The VMC wavefunctions from the twenty-two lowest energy states were used as trial wavefunctions for the CFQMC calculation. An ensemble of 80 walkers, 4 blocks of twenty walkers each, positioned at the final coordinates of the VMC walkers, was propagated 40,000 steps with $\delta\tau = 2$ to ensure the equilibration of the ensemble and 100,000 steps with $\delta\tau = 2$ to sample the CFQMC matrix elements, Eq. (2.57). The total sampling of the matrix elements was $M = 8,000,000$. The matrix elements were computed at projection time intervals of 12 H^{-1} for $\tau = 0$ H^{-1} to $\tau = 640$ H^{-1}. After 384 H^{-1} all block and ensemble overlap matrices became singular and the CFQMC eigenvalue equations became insoluble. Because the statistical error in the matrix elements of the highest energy trial functions dominates the error in the vibrational energies, we eliminated the 5 highest energy trial functions and repeated the CFQMC calculation An ensemble of 100 walkers, 5 blocks of 20 walkers each, positioned at the final coordinates of VMC calculation, was propagated 20,000 steps with $\delta\tau = 2$ to ensure the equilibration of the ensemble and 100,000 steps with $\delta\tau = 2$ to sample the CFQMC matrix elements. The total sampling of the matrix elements was $M = 10,000,000$.[8] The matrix elements were computed at projection time intervals of 12 H^{-1} for $\tau = 0$ H^{-1} to $\tau = 640$ H^{-1}. The block and ensemble overlap matrices became singular and the CFQMC eigenvalue equations

[8]These calculations were conducted on a Cray-XMP at the National Energy Super Computer Research Center at Lawrence Livermore National Laboratory.

became insoluble after 408 H^{-1}, only 24 H^{-1} later than in the 22 trial wavefunction calculation, but the statistical error was smaller for all $\Lambda_k(\tau)$. The wavefunctions were assigned state labels $(\nu_1, \nu_2^{l_2}, \nu_3)$ graphically, according to the number of nodes observed in the wavefunctions. The CFQMC vibrational state energies are given in Table 4.12. The thirteen converged vibrational states agreed to within a few cm^{-1} of the MORBID energies but tend to exceed the experimental values.

4.7 Discussion of C_3 calculations

Calculations of the J=0 vibrational states using PES I more accurately reproduce the experimental energies than the calculations using PES II or PES III; see Tables 4.10 and 4.12. For the manifold of pure bend states, calculations using PES II and PES III give energies which are greater than the experimental values, whereas the energies from the calculation using PES I are in close accord with experimental values. In the region where these states have large amplitude, $180° \geq \theta \geq 90°$, $-0.15a_0 \geq \Delta S_{sym} \geq 0.15a_0$, and $-0.15a_0 \geq \Delta S_{asym} \geq 0.15a_0$, pes I is broadest with repsect to θ and pes II is the most narrow. See Fig. 4.13. For the states (1,0,0) and (0,0,1), the symmetric and asymmetric stretches, respectively, the energies calculated using pes I are 13 cm^{-1} lower than the experimental values, but the energies calculated using pes III are lower by 30 cm^{-1}. The energies calculated using pes II are higher than the experimental values for states (1,0,0) and (0,0,1) by 46 cm^{-1} and 68 cm^{-1}, respectively. In the region where these states have large amplitude, $180° \geq \theta \geq 150°$, $-0.3a_0 \geq \Delta S_{sym} \geq 0.3a_0$, and $-0.3a_0 \geq \Delta S_{asym} \geq 0.3a_0$, pes II is the most narrow, and pes III is the broadest with repsect to the symmetric and asymmetric stretches. See Figs. 4.14 and 4.15. These considerations suggest that the overall width of the potential well is the predominate distinction between the three potential energy surfaces and the determining factor in the accuracy of the calculated vibrational states.

Jensen, *et al.* argue for the superiority of pes III which has a linear equilibrium geometry based on the size of $\Sigma - \Delta$ splittings[7]. They noted that the splittings observed from pes I were less accurate than those obtained from pes III. From these

Table 4.12: VMC and CFQMC vibrational state energies using PES III and SMOL basis functions for C_3 $\left(^1\Sigma_g^+\right)$ in cm^{-1}. The experimental entries from Ref. [54] are not marked. The experimental entries from Ref. [55] are marked with †. The experimental entries from Ref. [60] are marked with ‡. VMC and CFQMC refer to the present calculations. The numbers in parentheses are the estimated statistical errors (one standard deviation). The * denotes unconverged energies (no observed plateau), which are upperbounds. MORBID refers to the calculations from Ref. [7]

(ν_1, ν_2, ν_3)	Expt.	VMC	CFQMC	MORBID
ZPE		1684.5(0.8)	1681.6(0.5)	
(0, 2, 0)	132.7993†	162(1)	152(1)	154.3
(0, 4, 0)	286.11	360(1)	332(2)	333.2
(0, 6, 0)	461.09	594(1)	536(3)	527.8
(0, 8, 0)	647.59	873(2)	781(5)*	735.5
(0, 10, 0)	848.40	1200(2)	1028(37)*	955.2
(1, 0, 0)	1226.6	1189(1)	1186.9(0.9)	1191.9
(1, 2, 0)	1406.5	1422(1)	1372(4)	1378.3
(1, 4, 0)	1592.5	3383(3)	1600(2)	1592.5
(1, 6, 0)	1787.5	2047(5)	1886(15)*	1787.5
(2, 0, 0)	2434.9‡	2397(2)	2351(46)	2434
(2, 2, 0)	2656.3‡	2728(17)	2617(16)	2656
(0, 0, 1)	2040.0192†	2043(5)	2005(2)	2007.3
(0, 2, 1)	2133.8876†	2161(2)	2131(2)	2130.7
(0, 4, 1)		2415(4)	2289(3)	2283.2
(0, 6, 1)		2591(6)	2484(28)	2454.8

Figure 4.13: Bend dependence of C_3 potential energy surfaces at the equilibrium bond distances: for I and III $r_{12}^0 = 1.2897$ Å, for II $r_{12}^0 = 1.2906$ Å

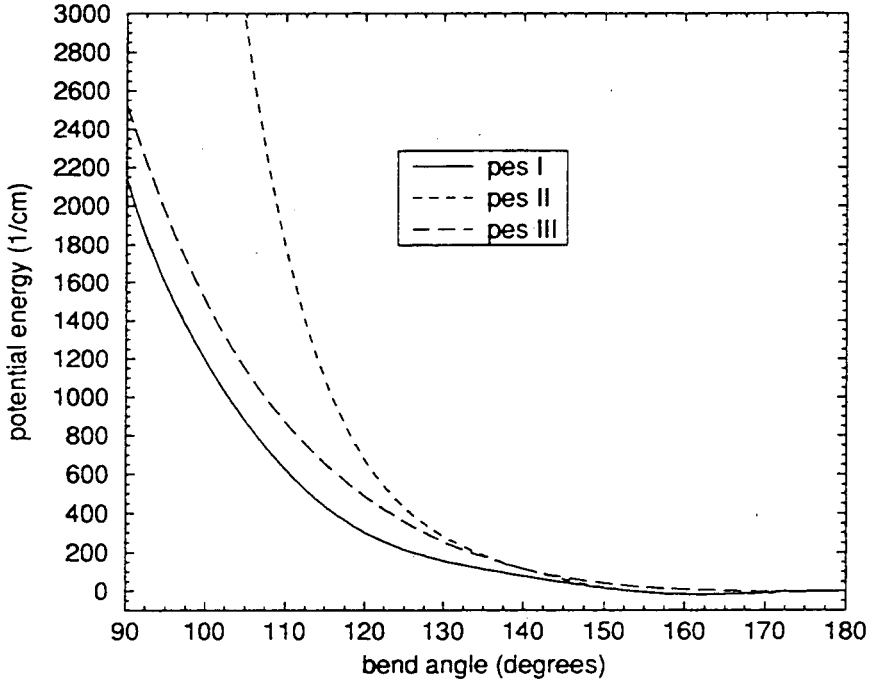


Figure 4.14: Symmetric stretch dependence of C_3 potential energy surfaces

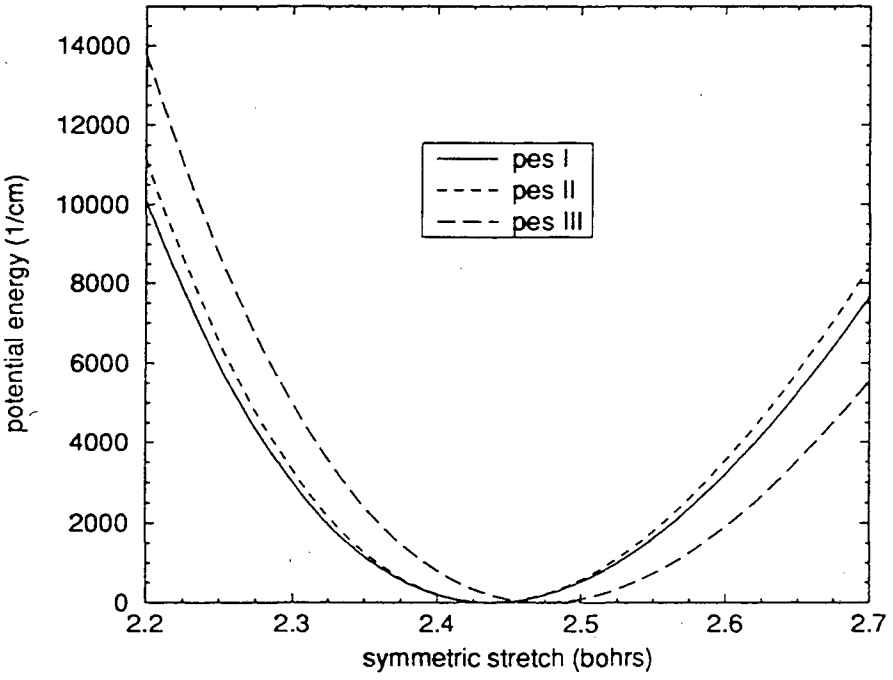
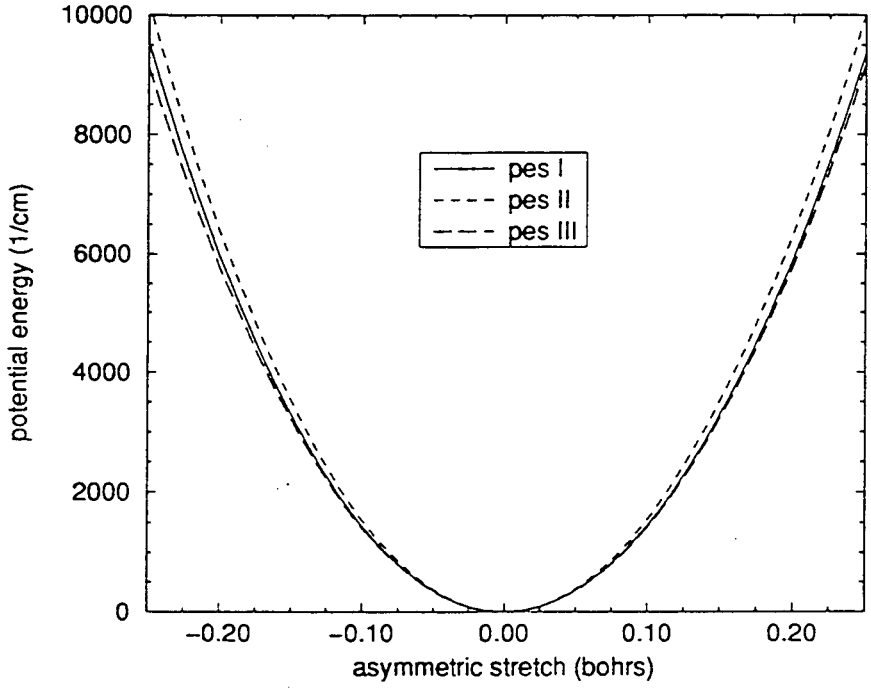


Figure 4.15: Asymmetric stretch dependence of C_3 potential energy surfaces



comparisons, they concluded that the pes of C_3 should not have a barrier to linearity. Since pes III is of higher quality than pes I for the $\Sigma - \Delta$ splittings but pes I is more accurate than pes III for the energies of the low lying vibrational states, fitting of pes III to the spectrum, as done to obtain pes I, might produce an extremely accurate as well as qualitatively correct pes. The recent pes constructed by Mladenovic et al. from large scale, single-, double- and triple-excitation coupled cluster calculations gives very accurate energies for the bending and stretching states, but the parameters of the pes are unpublished[24].

There has been considerable debate concerning the linearity of the equilibrium geometry of C_3. The differences in the potential energy surfaces favoring a linear structure and those favoring a bent equilibrium structure are very slight in the near linear region. The zero point energy of C_3 is approximately 1700 cm^{-1}. The predicted barrier to linearity in pes I is 16.5 cm^{-1} which is approximately 1% of the zero point energy. Furthermore, as discussed above, the width of the potential well determines the accuracy of the predicted vibrational states. For the potential energy surfaces investigated in this work, (pes I and II predicted a non-linear equilibrium geometry and pes III predicted a linear geometry) the wave function for the ground vibrational state has a single maximum located at the linear configuration. We note, however, that the wave function has significant amplitude at large deviations from linearity; see Figs. 4.6 and 4.7.

Excitation of the symmetric stretching mode tends to reduce the barrier to linearity and narrow the bending potential. Excitations in the asymmetric stretch mode tend to increase it and broaden the bending potential. A semirigid bender analysis by Northup, et al. showed that excitations of the symmetric stretch made the effective bending potential well narrower, whereas the excitations of the asymmetric stretch made the effective bending potential broader and induced a barrier to linearity[59]. The induced barrier to linearity increases with increasing excitation of the asymmetric stretching mode. Nevertheless, for pes I and pes II, the wave function for the states having one excitation in either the symmetric or asymmetric stretch has a single maximum at the linear configuration. The difference between the PES, a slight barrier to linearity, has no significant effect on the resulting wave functions

of the states examined.

4.8 Conclusions

The correlation function quantum Monte Carlo (CFQMC) method, a completely general approach, in principle, for computing excited state energies, has been used to study the low-lying vibrational states of floppy C_3 molecule and to test the applicability of the method to weakly bound systems. The CFQMC vibrational state energies are in excellent accord with results obtained using the MORBID method, with experimental measurements and comparable in accuracy to earlier results obtained with the method on the strongly bound H_2O and H_2CO systems.

In the course of this work, a novel basis function form based on the Morse oscillator like functions was introduced that makes possible reliable description of high bend excitations. In fact, three forms of basis functions, LM, SMOL and distributed Gaussian[9], were used to construct C_3 trial wavefunctions. The construction of several sets of trial wavefunctions highlights a liability that is shared by CFQMC and analytic methods. In order to converge very accurate vibrational state energies, CFQMC requires accurate trial wavefunctions which are constructed from analytical basis functions. The optimum choice of basis function form is a complex issue for any basis function dependent method. However, CFQMC was able to obtain accurate vibrational state energies for C_3 from trial wavefunctions constructed from LM basis functions and from trial wavefunctions constructed from SMOL basis functions. Furthermore, for QMC methods the implementation of diverse wavefunction forms, including explicitly correlated wavefunctions, is facile. Since the issue of basis function form is likely to grow in complexity with larger, floppier molecules, the convenient implementation of basis functions forms will be an increasingly important efficiency in computing the properties of vibrating molecules.

The great promise of QMC, the efficient scaling of Monte Carlo integration with respect to the dimensionality of the integral, remains largely untapped. As larger

[9]The application of distributed Gaussian basis functions to the construction of C_3 trial wavefunctions is reported in Appendix A

systems are investigated QMC methods are less likely to become compute-bound, than other high accuracy analytical methods. Therefore, the utility of CFQMC and QMC methods have potential for growth in the years to come.

Appendix A

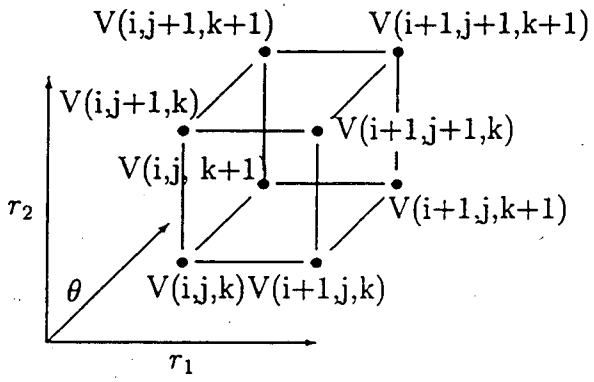
Further C_3 investigations with QMCVIB

A.1 Discrete potential representation

The many-body potential $V(\vec{s})$ that couples the atoms of the vibrating molecule is introduced in Chapter 1. It is called the potential energy surface (PES). Traditionally, the potential energy surface is represented by a function having parameters that are determined by a fit to empirical or *ab initio* data. Alternatively, the potential energy surface may be represented by a discrete potential representation (DPR). The potential energy is obtained for a set of points distributed in the 3N-6 dimensional space of internal coordinates. See Figure A.1. Interpolation and extrapolation schemes provide the potential for all other points in internal coordinate space when needed. Therefore, DPR provides all the information about the potential required by QMC methods without requiring raw potential data to be fit to an analytical form. This suggests that DPR-QMC may be a convenient and immediate method of computing vibrational properties from *ab initio* data.[61]

Interpolation schemes range from crude to elegant, with various guarantees regarding continuity of the function and derivatives. Bilinear interpolation is a simple, commonly used method of interpolating between points of a multi-dimensional gird. It provides continuous function values everywhere and continuous function gradients

Figure A.1: Discrete potential energy grid



everywhere except at the grid boundaries [62]. Taylor's series provides continuous function values around a point. These methods, used in conjunction, provide an interpolation-extrapolation scheme which is rigorously continuous except for small discontinuities at grid boundaries. As such, the hybrid method is suitable for testing the utility of DPR-VMC and DPR-CFQMC.

The DPR potential has the form of a weighted sum,

$$V(r_1, r_2, \theta) = \sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{k=0}^{1} c_{ijk} V^*_{ijl}(r_1, r_2, \theta) \tag{A.1}$$

where the coefficients

$$c_{ijk} = \left(1 - \frac{r_1 - r_1(i)}{r_1(i+1) - r_1(i)}\right) \tag{A.2}$$

$$\left(1 - \frac{r_2 - r_2(j)}{r_2(j+1) - r_2(j)}\right)\left(1 - \frac{\theta_1 - \theta_1(k)}{\theta_1(k+1) - \theta_1(k)}\right) \tag{A.3}$$

weight the contributions of each Taylor's series estimation

$$V^*_{ijl}(r_1, r_2, \theta) = V_{ijl} + \frac{[V_{i+1,jk} - V_{ijk}][r_1 - r_1(i)]}{r_1(i+1) - r_1(i)} \tag{A.4}$$

$$+\frac{[V_{i,j+1,k} - V_{ijk}][r_2 - r_2(j)]}{r_2(j+1) - r_2(j)} + \frac{[V_{ij,k+1} - V_{ijk}][\theta_1 - \theta_1(k)]}{\theta_1(k+1) - \theta_1(k)} \tag{A.5}$$

according to the distance from the grid points R_{ijk} and internal coordinates for which the potential is being computed. The requisite derivatives for the Taylor's series are approximated by finite difference estimations from the grid energies. Outside the grid domain the potential is computed by Taylor's series only.

A DPR of C_3 was constructed from a grid of C_3 total energies computed by Jensen, Rohlfing and Almlöf for the construction of PES III, for which parameters are shown in Table 4.1[7]. They computed CASSCF energies for C_3 at bond angles of 180° (linear), 162.5°, 150°, 130°, 120°, 110°, 90°, 80°, and 70° and bond lengths 1.16, 1.19, 1.22, 1.25, 1.2774, 1.31, 1.34, 1.37, 1.40, and 1.43 Å.[7] This grid consists of 100 points of C_{2V} symmetry and 450 points of C_S symmetry. Since PES III was obtained by fitting values for bond angles less than or equal to 90° and the standard deviation of the fit was 15.8 cm^{-1}, the DPR and PES III will not be equivalent even at the grid points. Furthermore, PES III will have continuous values and derivatives everywhere, whereas the DPR will not.

In order to compare DPR and PES III, DPR-VMC and DPR-CFQMC C_3 vibrational state energies were computed. The basis set parameters used in the PES III calculations were used in the DPR calculations. These are given in Tables 4.7 and 4.8. The basis set included 130 SMOL functions such that $n_1 + n_2 \leq 4$ for $n_3 \leq 5$, $n_1 + n_2 \leq 3$ for $6 \leq n_3 \leq 8$, $n_1 + n_2 \leq 2$ for $n_3 = 9$ and $n_1 \leq 1$, $n_2 \leq 1$ for $n_3 = 10$. A VMC ensemble of 200 walkers was propagated 25,000 steps with $\delta\tau = 1500$ to reach an equilibrium distribution of ψ_G^2 and afterwards propagated 10,000 steps with $\delta\tau = 1600$ to sample the Hamiltonian and overlap matrix elements, Eqs. (2.12) and (2.13). The total sampling of the matrix elements was $M = 2,000,000$. The statistical uncertainty of VMC vibrational state energies was large, indicating that the variance of the local energy of the trial wavefunctions would be large. To minimize statistical uncertainty VMC wavefunctions from the 14 lowest energy states were used as trial wavefunctions for the CFQMC calculation, rather than 17 as in calculations using PES III. An ensemble of 200 walkers, 10 blocks of twenty walkers each, positioned at the final coordinates of the VMC walkers, was propagated 50,000 steps with $\delta\tau = 2$ to ensure the equilibration of the ensemble and 50,000 steps with $\delta\tau = 2$ to sample the CFQMC matrix elements, Eq. (2.57). The total sampling of the matrix elements was

Table A.1: VMC and CFQMC vibrational state energies using DPR and SMOL basis functions for $^1\Sigma_g^+$ C_3 in cm^{-1}. The experimental entries from Ref. [54] are not marked. The experimental entries from Ref. [55] are marked with †. VMC and CFQMC refer to the present calculations. The numbers in parentheses are the estimated statistical errors (one standard deviation). The * denotes unconverged energies (no observed plateau), which are upper bounds.

		DPR		Pes III
(ν_1, ν_2, ν_3)	Expt.	VMC	CFQMC	CFQMC
ZPE		1678(1)	1677(1)	1681.6(0.5)
(0, 2, 0)	132.7993†	167(2)	158(6)	152(1)
(0, 4, 0)	286.11	359(3)	335(6)	332(2)
(0, 6, 0)	461.09	590(3)	536(10)	536(3)
(0, 8, 0)	647.59	869(3)*	786(11)*	781(5)*
(0, 10, 0)	848.40	1194(6)*	1029(60)*	1028(37)*
(1, 0, 0)	1226.6	1178(2)	1176(2)	1186.9(0.9)
(1, 2, 0)	1406.5	1413(4)	1372(10)	1372(4)
(1, 4, 0)	1592.5	1678(4)	1599(40)	1600(2)
(1, 6, 0)	1787.5	1995(2)	1911(40)	1886(15)*
(0, 0, 1)	2040.0192†	2020(4)	1978(5)	2005(2)

$M = 10,000,000$. The matrix elements were computed at projection time intervals of 12 H^{-1} for $\tau = 0$ H^{-1} to $\tau = 480$ H^{-1}. After $\tau = 456$ H^{-1} all block overlap matrices became singular and their CFQMC eigenvalue equations became insoluble. The ensemble overlap matrices remained positive for all projection times sampled.[1] The vibrational state energies are given in Table A.1.

The DPR-VMC and DPR-CFQMC calculations agree with the PES III calculations for most converged eigenstates. The greatest differences were found in states (1,0,0) and (0,0,1). The more highly excited states agree more closely with PES III

[1]These calculations were conducted on a Cray-XMP at the National Energy Super Computer Research Center at Lawrence Livermore National Laboratory.

calculations than the lower states. Perhaps the lower vibrational states, which are the least broad, are more greatly perturbed by the differences between the DPR and PES III, whereas the more highly excited states, which sample more broadly, are perturbed less by small local differences between the the DPR and PES III.

Despite comparatively greater sampling of the matrix elements, the statistical error of the DPR calculations is greater than that of the Pes III calculation. This may indicate that the variance of the local energy is sensitive to the non-continuous derivatives at the grid boundaries in the DPR or the that the SMOL functions are most accurate for Morse type potentials.

Although the DPR-VMC and DPR-CFQMC lack the precision of the PES-type VMC and CFQMC calculations, it provided the same information about the vibrational states of C_3 without fitting the potential to a analytical form. DPR-QMC is therefore, a convenient tool for obtaining the vibrational states predicted from raw *ab intio* data.

A.2 Distributed Gaussian basis functions

Distributed Gaussian Bases (DGB) have been used in computing vibrational states of model dimers and van der Waals clusters.[63, 64]. They are are highly flexible functions. Both the Gaussian center and widths may be adjusted. By localizing them in regions of the multi-dimensional space where wave function amplitude is expected, basis sets for molecules with single or multiple potential minima may be constructed. DGB have the simple mathematical form:

$$f_m = \exp\left(\sum_{\nu=1}^{3} W_{\nu(m)}\left(S_\nu - S^0_{\nu(m)}\right)^2\right) \tag{A.6}$$

where $W_{\nu(m)}$ is the width parameter of the internal coordinate ν for basis function m and $S^0_{\nu(m)}$ is the Gaussian center with respect to the coordinate ν. This suggests that accurate trial wave functions for C_3 could be constructed from a DGB.

Being highly flexible functions, optimization of the DGB parameters is huge job. Each basis function has two independent parameters for each internal coordinate.

Symmetry constraints of C_3 limit the number of independent parameters for each basis function from 6 to 4. As for the LM and SMOL basis functions, the DGB parameters must be optimized with respect to the energies of the ground and several excited states. Most probably the optimal widths and centers will not be the same for different size DGB.

Since direct optimization of the multidimensional Gaussian functions is a huge job, external constraints were used to specify the parameters. In one approach the Gaussians were equally spaced and the single width parameter $W(m)$ used for each coordinate. The balance between distinct linear independence (large W) and low kinetic energy matrix elements (small W) constrains the choice of the width parameter.[63, 64] Basis sets constructed according to this prescription was problematic for C_3. Small basis sets provided very poor descriptions of the vibrational states with respect to energy. Apparent linear dependencies arose in small basis sets. Therefore, with the equally spaced prescription the DGB did not provide any vibrational states of C_3.

Another approach is suggested by the STO-nG basis used in electronic structure calculation in which Gaussians are used to approximate Slater type orbitals. Multidimensional Gaussian functions are topologically similar to SMOL but mathematically simpler. See Eqs. (smol-bf) and (A.6). The DGB parameters can be specified according to the SMOL basis function maxima and half-height widths. Since the SMOL basis function provided an accurate description of C_3 vibrational states, then the DGB fit to the SMOL function may also provide an accurate description of the C_3 vibrational states.

The Gaussian centers S_ν^0 are placed at the maxima of the SMOL basis functions which are given by:

$$S_\nu^{max} = -\frac{1}{\beta_\nu} \ln\left(\frac{\alpha + 1 + 2m}{\alpha}\right) + S_\nu^0 \tag{A.7}$$

where S_0 is the equilibrium value of the internal coordinate S, α and β are SMOL parameters, and m is the basis function order parameter. The width parameter $W_{\nu(m)}$

is related to the SMOL width at half-height by

$$W_{\nu(m)} = \frac{4 \ln 2}{\left(W_{1/2}^{SMOL}\right)^2} \tag{A.8}$$

The half-height widths of the SMOL functions are found numerically from the equation,

$$f_m\left(S_{1/2}\right) - \frac{1}{2} f_m\left(S^{max}\right) = 0 \tag{A.9}$$

where f_m are the SMOL basis functions given in Eq. (4.8).

The DGB parameters were constrained to be pseudo SMOL functions as described above. The optimal Morse parameters were obtained according to the prescription described in Chapter 4. The basis set was expanded to include 90 SMOL functions such that $n_1 + n_2 \leq 4$ and $n_3 \leq 5$. The larger basis set had apparent linear dependencies. A VMC ensemble of 100 walkers was propagated 10,000 steps with $\delta\tau = 1500$ to reach an equilibrium distribution of ψ_G^2 and afterwards propagated 50,000 steps with $\delta\tau = 1600$ to sample the Hamiltonian and Overlap matrix elements, Eqs. (2.12) and (2.13). The total sampling of the matrix elements was $M = 5,000,000$.[2] The VMC vibrational state energies are given in Table A.2.

In general the DGB basis function provided a less accurate description of the vibrational states than the LM and SMOL basis functions. The ground state was poorly described. Therefore, no CFQMC calculations were conducted. Alternative parameter optimization methods could overcome the observed inaccuracy of the DGB. The computational expedience of DGB and the possibility that DGB may provide relatively more accurate trial wave functions for other molecules suggests utility for DGB in VMC and CFQMC calculations.

[2]These calculations were conducted on a Cray-XMP at the National Energy Super Computer Research Center at Lawrence Livermore National Laboratory.

Table A.2: VMC vibrational state energies for C_3 using PES I and distributed Gaussian basis functions for $^1\Sigma_g^+$ C_3 in cm^{-1}. The experimental entries from Ref. [54] are not marked. The experimental entries from Ref. [55] are marked with †. The experimental entries from Ref. [60] are marked with ‡. LM-VMC, SMOL-VMC and DGB-VMC refer to the present calculations. The numbers in parentheses are the estimated statistical errors (one standard deviation).

(ν_1, ν_2, ν_3)	Expt.	LM-VMC	SMOL-VMC	DGB-VMC
ZPE		1667.7(0.3)	1669.5(0.7)	1750(3)
(0, 2, 0)	132.7993†	140(0)	147(1)	121(4)
(0, 4, 0)	286.11	292(1)	306(1)	392(4)
(0, 6, 0)	461.09	514(3)	506(2)	561(4)
(0, 8, 0)	647.59	885(8)	751(2)	674(4)
(0, 10, 0)	848.40		1048(3)	803(4)

Appendix B

Parallelization of QMCVIB

Massively Parallel computing offers much computational power. Tasks are divided and distributed among multiple CPU's increasing the quantity of operations performed per second with the number of CPU's available. Task interdependency necessitates scaling inefficiencies in parallel computing performance. The challenge of parallel computing is overcoming task interdependency by creative division and distribution of tasks among multiple CPU's.

Monte Carlo methods, particularly VMC and CFQMC, are well structured for parallel computation. They have long, independent chains of tasks which are exclusively linked between parent and child. Evaluation of the matrix elements, Eqs. (2.10) and (2.11) for multi-state VMC and, Eq. (2.57) for CFQMC, accounts for most of the computational cost. During the evaluation of the matrix elements the walkers exchange no information. Ensemble averages and statistics may be obtained serially following the evaluation of the matrix elements. According to Amdahl's Law, the maximum speedup of parallelism is inversely proportional to the fractional amount of the inherent serial computation. For CFQMC the serial fraction is typically less than 0.03 and the maximum speedup is, pessimistically, 33x. Amdahl's law is circumvented, if we measure performance by the the maximum work in a fixed amount of time rather than minimum time for a fixed amount of work. A large ensemble of walkers may be divided into sub-ensembles and these may be distributed among multiple processors to obtain performance gains which surpass 33x and scale approximately

with the number of processors. Unfortunately, the pre-VMC and pre-CFQMC equilibration stage, of which a small part is serial calculations, scales with the ensemble size. The effective serial portion of the code is increased and the maximum Amdahl's law speedup is decreased.

B.1 VMC and CFQMC on the C-M5

A CMF version of QMCVIB was constructed to run on a CM-5 parallel supercomputer using the single instruction multiple date (SIMD) program model. The CM-5 consists of a partition manager (PM) and 512 processing nodes (PN). The PM initiates all programs, coordinates the PN's and provides a UNIX interface for users. Each node consists of four virtual vector processors with 16 Mbytes local memory and a coordinating processor. The virtual vector units operate at 16 MHz and have a minimum vector length of 32. The processor unit communicates to the virtual vector units at 20 MBytes per sec. The nodes are interconnected by a hypertree network with a limiting speed of 20 Mbytes per second or slower.

QMC input data is read by the PM and distributed to every processor. During the QMC walk the FORTRAN loops over walkers which are distributed 32 walkers per processors meeting the minimum vector length. Therefore, each partition generated data from 1024 walkers. When all 16 partitions were employed, data was generated from 16,384 walkers. The QMC walk algorithm is performed simultaneously on every processor. After the walk, the matrix element data are collected and a serial eigensolver package solves the VMC and CFQMC linear equations.[1]

VMC and CFQMC calculations of the vibrational states of H_2O and C_3 were conducted on 256 and 512 node partitions of the CM5 at the Army High Performance Supercomputer Research Center in Minnepolis, Minnesota. See Tables 3.3, 3.4 and 4.4. The 512 node partition runs propagated 16,384 walkers 10,000 steps each for a total sampling of $M = 163,840,000$ steps. In typical Cray-XMP runs an ensemble of 100 walkers is propagated 100,000 steps for a total sampling of $M = 10,000,000$

[1]The parallel eigensolvers available were unable to solve the VMC and CFQMC linear equations reliably.

steps, only 1 sixteenth of the sampling obtained in the CM5 run.

Several calculations of the vibrational states of $H2O$ were performed on 32, 128, 256 and 512 node partitions of the CM5 in order to determine the performance of the CMF QMCVIB code.[2] One series of calculations conducted on 32, 128, 256 and 512 node partitions of the CM5 used a basis set of 36 functions and a total sampling M = 4,096,000 steps for VMC and for CFQMC. A second series of calculations conducted on 32, 128, 256 and 512 node partitions of the CM5 used a basis set of 84 functions and a total sampling M = 20,480,000 steps. One final calculation was made using the 84 function basis set, but having a total sampling M = 163,840,000 steps. This calculation was a production run and the pre-VMC and pre-CFQMC equilibration walks were longer than those used in the preceding calculations.

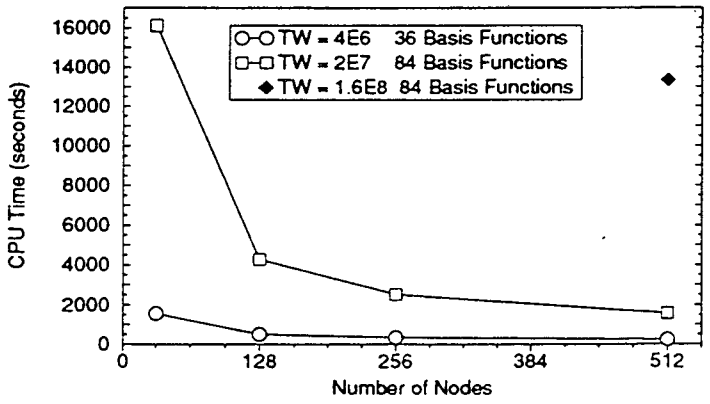
For each calculation a total CPU time was measured. It consisted of a the CPU time for a pair of VMC and CFQMC calculations: reading VMC input and initialization of the walkers, an equilibration of the VMC ensemble, the MC walk sampling the variational matrix elements, the solution of the secular equations, error analysis and result output, reading CFQMC input and initialization of the CFQMC ensemble, equilibration of the CFQMC ensemble, the MC walk sampling the CFQMC matrix elements, the solution of the secular equations, error analysis and result output. See Fig. B.1. Therefore, both serial computation and internode communication times are included in the measurement.

The total CPU times for the first series scales between $N^{-0.8}$ and $N^{-0.5}$. The total CPU time for the second series scales between $N^{-0.95}$ and $N^{-0.92}$. The increased sampling and increased vector length from the larger basis set diminish the influence of the serial and internode communication inefficiencies.

For each calculation the CM5 work rate was estimated. The CM5 work rate is (M/t), where M is the total work and t is the total CPU time. See Fig. B.1. As expected, the work rate depends heavily on the basis set size since this governs how many operations are needed for a walker to complete one MC step. For a given basis

[2]A CM5 disclaimer for Thinking Machines Corp.: These results are based upon a beta version of the software and, consequently, is not necessarily representative of the performance of the full version of the software.

Figure B.1: CM5 total CPU times for small, medium and large QMC calculations.



set size, the work rate should increase with M. However, the larger TW (M) 84 basis function calculation shows a decrease in efficiency relative to the the smaller TW 84 basis set calculation. Most likely, the longer pre-VMC and pre-CFQMC equilibration walk accounts for the slightly decreased the work rate. The work rate of first series scales between $N^{0.8}$ and $N^{0.5}$, and work rate of the second series scales between $N^{0.95}$ and $N^{0.92}$. Therefore, the larger runs use the machine more efficiently.

Finally, the node work rate, the quotient of the work rate and the number of nodes in the partition used for the calculation, was also estimated. See Fig. B.3. It provides a measurement of the scaling of the inefficiencies with respect to the number of nodes used. The first series shows an approximately 60% drop in node work rate between 32 node and 512 node partitions. The second series shows an 35% drop in efficiency between 32 node and 512 node partitions. If we assume little or no drop in efficiency between 1 and 32 nodes, the CM5 provided a 332x speed-up using all 512 processors.

Figure B.2: CM5 work rate for small, medium and large QMC calculations

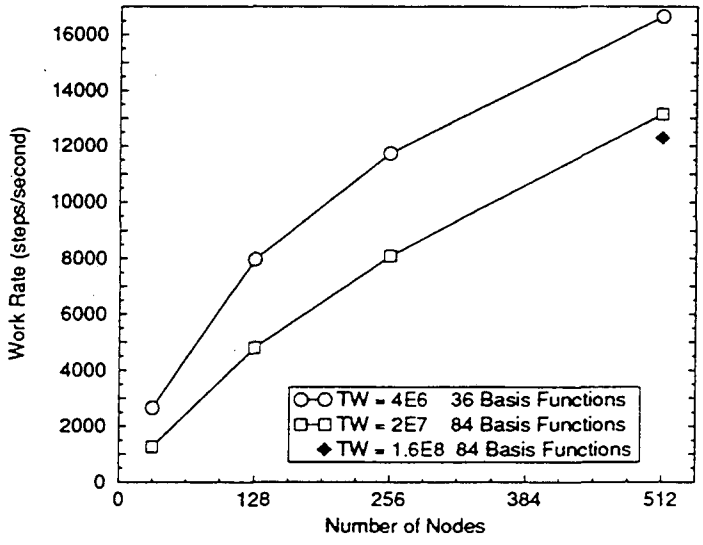
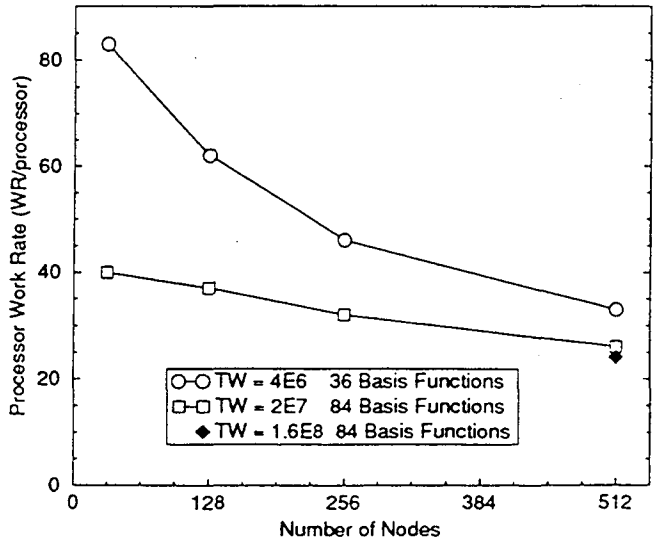


Figure B.3: CM5 processor efficiency of small, medium and large QMC calculations



B.2 Outlook for parallel QMC

The large speed-up provided by the CM5 indicate that parallelization offers great gains for QMC algorithms, but not necessarily with the CM5. The individual processors are slow relative to a CRAY-XMP processor. This results mainly from the large time cost for communication between the coordinating processor and the virtual vector units. A series two type calculation takes approximately 7200 seconds on the Cray-XMP. Therefore, the total speed-up of the 512 node partition on the CM5 over one processor on the CRAY-XMP is 5.14. Furthermore, programming in CMF FORTRAN is quite complicated. Hence, algorithmic or application development on the CM5 is extraordinarily problematic. Two platforms may need to be maintained for code development.

Appendix C

QMCVIB Manual

QMCVIB calculates ground and excited vibrational state energies of a molecule for a given potential energy surface. It employs multi-state VMC to construct trial wavefunctions and CFQMC to obtain highly accurate molecular vibrational state energies. QMCVIB contains three forms of basis functions for monomer systems.[1] QMCVIB contains two potential energy representations, discrete and analytic, and several potential energy surfaces for H_2O and C_3.

The fundamental QMC portion of the code requires little user specification, only the ensemble size, length of random walks, and projection times for which the CFQMC matrix is evaluated. However, the wavefunction and potential energy surfaces require considerable user specification. For VMC and CFQMC users supply the basis set non-linear parameters and specify which order basis functions are included in the basis set. The VMC calculation outputs the linear coefficients of the basis functions for the molecular vibrational wavefunctions used as trial functions in CFQMC. The user specifies the quantity of trial functions included in the CFQMC calculation. For VMC and CFQMC, users supply a subroutine which computes the potential energy for the molecule of interest given the cartesian coordinates of the atoms and atom-atom distances. Alternatively, the users supply the potential energy data for a grid of points in the internal coordinate space of the the molecule of interest

[1]Basis functions forms for dimers or trimers or larger polymers may be included as subroutines analogous to XLOCALBERNU.f and GETBFBERNU.f, or XLOCALMORSE.f and GETBFMORSE.f or XLOCALDGB.f and DGBASIS.f which are described below.

and QMCVIB interpolates the potential from these points.

C.1 Structure

QMCVIB is written in standard FORTRAN 77 for compilation and execution on Cray supercomputers. It has 46 subroutines and functions of which 17 are used in all calculations, 10 are used optionally depending on the type of QMC calculation, 8 are used alternatively depending on the type of basis functions used, 9 are used optionally depending on the PES employed, 2 are used optionally for numerical evaluation of wavefunction derivatives, and 2 are used optionally for dumping and restarting calculations. In addition to the subroutines, QMCVIB has an include file containing common block information.

The following is a list of all QMCVIB sub-routines:

BF The function which computes the value of Local Mode Basis functions, Eqs. (3.3) and (3.4). The function is called by FINITE in VMC or CFQMC calculations in which the LM basis function derivatives are computed numerically.

CORREL The subroutine which computes the CFQMC matrix elements, Eqs. (2.50) and (2.51). It is called by QMC in CFQMC calculations.

DEFAULT The subroutine that sets default parameters. It is called by QMCVIB in all calculations.

DGBASIS The subroutine that calculates the basis-function guiding-function ratio and local kinetic energy for distributed Gaussian basis functions, Eq. (A.6). It is called by XLOCALDGB in VMC and CFQMC calculations with DGB basis functions.

DIAGF The subroutine that solves the CFQMC eigenvalue equation, Eq. (2.49). It is called by QMC in CFQMC calculations.

DIAGV The subroutine that solves the secular equations, Eq. (2.14). It is called by QMC in VMC calculations.

EAVG The subroutine that computes energy averages and variances for the guiding function. It is called by QMC and used in all calculations.

FINITE The subroutine that computes derivatives of the guiding function or basis functions by finite differences. It is called by ONESTEP and XLOCALBERNU when the input parameter IFORM = 4. It is useful for verifing analytic derivatives, but alternative verification methods are preferred. Presently it may be used only for Local Mode basis functions. See BF.

FULLWIDTH The function that determines the fullwidth at half-height of simplified Morse oscillator-like basis functions. It is called by INITS in VMC and CFQMC calculations in which the distributed Gaussian basis function parameters are determined from simplified Morse oscillator-like basis function α and β parameters.

G The function that computes the value of the guiding function, Eq. (3.5). It is called by subroutine FINITE in calculations in which the guiding function derivatives are computed numerically.

GEOMET The subroutine that computes equilibrium interatomic distances of the molecule. It is called by QMCVIB in all calculations.

GETBFBERNU The subroutine that calculates the basis-function guiding-function ratio and local kinetic energy for LM basis functions, Eqs. (3.3) and (3.4). It is called by XLOCALBERNU in VMC and CFQMC calculations with local mode basis functions.

GETBFMORSE The subroutine that calculates the basis-function guiding-function ratio and local kinetic energy for simplified Morse oscillator-like basis functions, Eq. (4.8). It is called by XLOCALMORSE in VMC and CFQMC calculations with SMOL basis functions.

HEADER The include file containing global parameters and the common block. It is included in nearly all subroutines.

INIT The subroutine that initializes walkers for the Monte Carlo random walk. It is called by QMCVIB in all calculations.

INITS The subroutine that defines the VMC basis functions and CFQMC trial wavefunctions. It is called by QMCVIB in all calculations.

LREBAK, LREDUC, LTQL2, LTQLRA, LTRED1, LTRED2 Subroutines from the LRSG eigensystem subroutine package (See below).

LRSG The subroutine that calls the recommended sequence of subroutines from EISPACK to find the eigenvalues and eigenvectors for the real symmetric generalized eigenproblem.[2] It is called by DIAGV in VMC calculations and by DIAGF in CFQMC calculations.

ONESTEP The subroutine that conducts the importance sampled random walk. It is called by QMC in all calculations.

POTC3-1 The subroutine that evaluates the local potential energy of C_3 PES I. It is called from INIT and ONESTEP in calculations using C_3 PES I.

POTC3-2 The subroutine that evaluates the local potential energy of C_3 PES II. It is called from INIT and ONESTEP in calculations using C_3 PES II.

POTC3-3 The subroutine that evaluates the local potential energy of C_3 PES III. It is called from INIT and ONESTEP in calculations using C_3 PES III.

POTC3-3I The subroutine that evaluates the local potential energy of C_3 using a discrete potential representation. It is called from INIT and ONESTEP in DPR calculations of C_3. See QSORT.

POTH2O-1 The subroutine that evaluates the local potential energy of H_2O PES I. It is called from INIT and ONESTEP in calculations using H_2O PES I.

POTH2O-2 The subroutine that evaluates the local potential energy of H_2O PES II. It is called from INIT and ONESTEP in calculations using H_2O PES II.

[2]Eispack is maintained by B. S. Garbow at Argonne National Laboratory.

POTH2O-3 The subroutine that evaluates the local potential energy of H_2O PES III. It is called from INIT and ONESTEP in calculations using H_2O PES III.[3]

POTH2O-4 The subroutine that evaluates the local potential energy of H_2O pes IV. It is called from INIT and ONESTEP in calculations using H_2O pes IV.[4]

QDUMP The subroutine in which accumulated data is written out to the file fort.56 for use in restarting the calculation. It is called by QMC when IWALK(2) = 2,3 or 4. See RESTART.

QFORCE The subroutine that evaluates the amplitude and the quantum force of the guiding function, Eq. (3.5). It is called from INIT and ONESTEP in all calculations.

QMC The subroutine that executes the Metropolis random walk. It is called from QMCCTL in all calculations.

QMCCTL The subroutine that calls the phases of the calculation as specified by the ICTRL parameters. It is called by QMCVIB.

QREAD The subroutine that reads the input file. It is called by QMCVIB in all calculations.

QSORT The subroutine that reads files containing atomic coordinates and electronic energies, fort.25 and fort.26, and constructs the discrete potential representation arrays. See POTC3-3I.

QWRITE The subroutine that writes out the calculation specifications. It is called by QMCVIB.

RESTART The subroutine that reads file fort.55 containing data accumulated in previous calculation in order to continue the calculation. It is called by QMC-CTL when IWALK(3) = 2, 3 or 4.

---

[3] H_2O PES III is a SPF quartic pes based on CCSD(T) energies computed by Bartlett et al and fit by Ermler[65].

[4] H_2O pes IV is Morse Oscillator based pes developed by Coker[66].

SETRAN The subroutine that sets the initial random number seed. It is called by QWRITE in all calculations.

SYSCPU The subroutine that obtains total and elapsed CPU times. It is called by QMCVIB and QMC in all calculations.

SYSTIM The subroutine that obtains the date and time. It is called by QWRITE in all calculations.

TENERGY The subroutine that evaluates the local kinetic energy of the guiding function, Eq. (3.5). It is called from INIT and ONESTEP in all calculations.

XLOCALBERNU The subroutine that computes the Hamiltonian and overlap matrix elements for VMC calculations using LM basis functions, Eqs. (3.3) and (3.4).

XLOCALDGB The subroutine that computes the Hamiltonian and overlap matrix elements for VMC calculations using DGB basis functions, Eq. (A.6).

XLOCALMORSE The subroutine computes the Hamiltonian and overlap matrix elements for VMC calculations using SMOL basis functions, Eq. (4.8).

C.2 Compilation

C-shell makefiles are used to direct the compilation. The compilation command is 'make -f *makefile-filename*. QMCVIB must be compiled with the subroutine which evaluates the potential energy for the molecule of interest. Therefore, the makefile must be amended. The SRC and OBJ lists must include only the desired PES subroutine. QMCVIB has static memory. Therefore, the parameters NPAR (the number of atoms), NMOD (the number of modes), NBLOCK (number of blocks), MBLOCKSIZE (the number of random walkers per block, NBASIS (number of basis functions), NST (number of states), and NT (the number of CFQMC projection times) must be specified in the include file, header.f.

Figure C.1: An ordered subroutine tree for QMCVIB: Boxes are routines or subroutines. Lines are subroutines calls. Lines proceed from the bottom or sides of the parent routine into the top of the child routine. The subroutine calls are arranged from left to right according to order of execution.

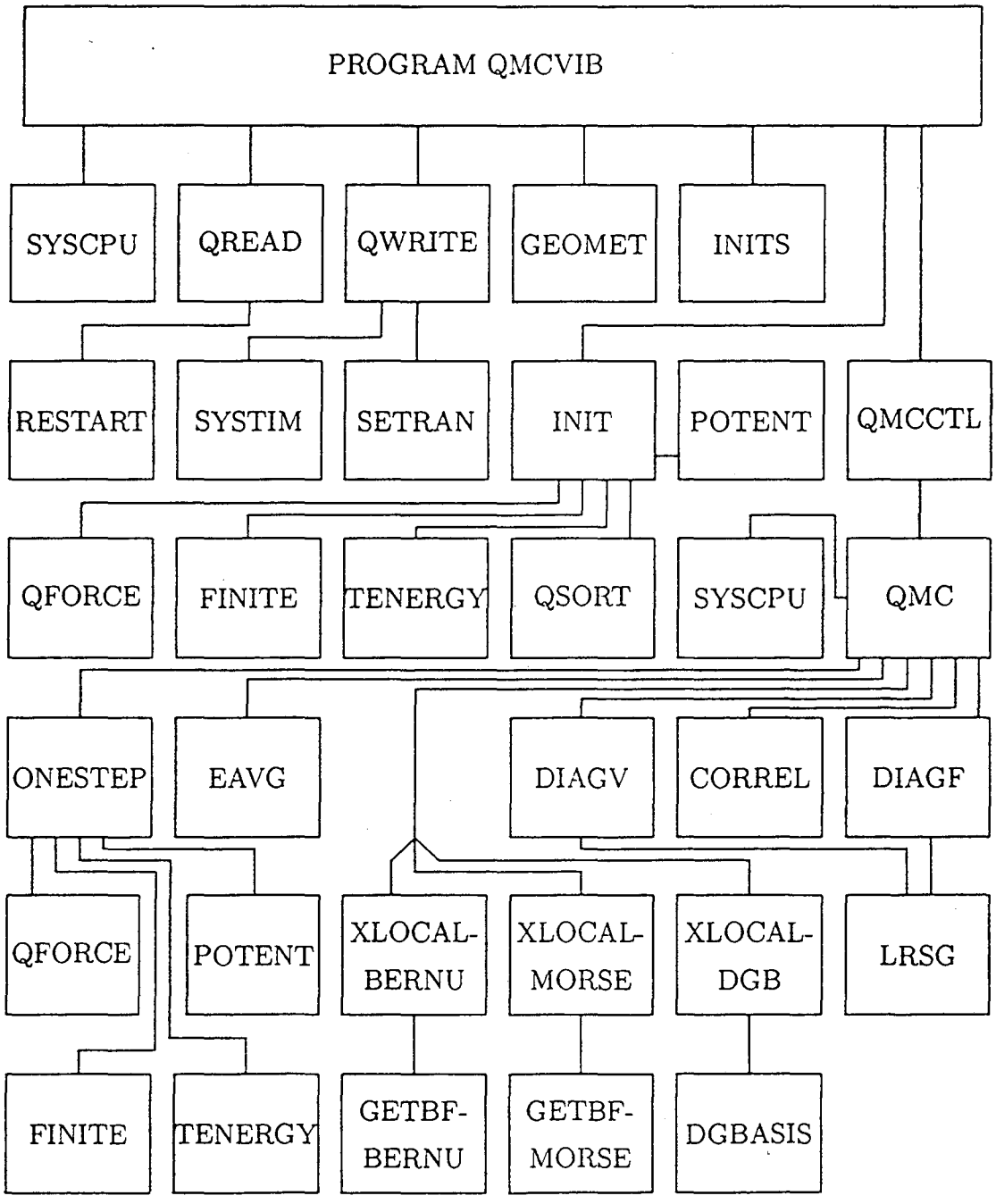


Figure C.2: QMC input file lines 1 to 14

```
---------------------------- QMC CONTROL ----------------------------
        PRE-MC EQULIBRATION  01          0000020000        MAXSTEP(1)
    VARIATIONAL MONTE CARLO  01          0000020000        MAXSTEP(2)
     SINGLE PROJECTION TIME  00          0000000000        MAXSTEP(3)
  MULTIPLE PROJECTION TIMES  00          0000000000        MAXSTEP(4)
INITIAL WALKER DISTRIBUTION  00                            IGUESS
      DUMP RESULTS OF PHASE  00                            IWALK(02)
           RESTART OF PHASE  00                            IWALK(03)
      TIME STEP (1/HARTREE)  01200.0000                    TAU
CORRELATION INTERVAL (N*TAU) 06                            LSTEP
  SAMPLING  INTERVAL (N*TAU) 05                            JTSP
       IVEC                  01
   DISCRETE POTENTIAL REP.   00                            IWALK(05)
PRINT FLAGS (20(I1,1X))ZZ    0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
```

C.3 Input files

Execution of QMCVIB requires one input file to specify the QMC calculation, random walk and wavefunctions. Auxiliary input files are used to specify initial walker positions, basis function parameters and discrete potential energy data. A sample input file is shown in Figs. C.2, C.3 and C.4. It specifies a multi-state VMC calculation of the vibrational states of C_3.[5] It has a fixed format; all lines must be present and strictly formatted.

Line 1 is a title. Its format is A80.

Line 2 specifies whether a pre-equilibration phase random walk is conducted and the length of the pre-equilibration random walk. Its format is T30,I2,T40,I10. In Fig. C.2 the pre-equilibration phase of the program is turned on; 01 specifies on

[5]Some lines in the figures of the input file have been altered to meet margin requirements.

and 00 specifies off. The walklength is set to 20,000 steps per walker.

Line 3 specifies whether the multi-state VMC calculation is conducted and the length of the VMC random walk. Its format is T30,I2,T40,I10. In Fig. C.2 the multi-state VMC phase is turned on and the walklength is set to 20,000 steps per walker. QMCVIB is compiled with parameters NBASIS = NST and NT = 1.

Line 4 specifies whether a single time-projection CFQMC calculation is conducted and the length of the single time-projection random walk. Its format is T30,I2,T40,I10. In Fig. C.2 the single time-projection phase is turned off and QMCVIB should be compiled with parameter NT = 1. If single time projection is turned on, QMCVIB should be compiled with parameter NT = 2.

Line 5 specifies whether a multiple time-projection CFQMC calculation is conducted and the length of the multiple time-projection random walk. Its format is T30,I2,T40,I10. In Fig. C.2 the multiple time-projection phase is turned off and QMCVIB should be compiled with parameter NT = 1. If multiple time projection is turned on, QMCVIB should be compiled with parameter NT < 2.

Line 6 specifies the initial walker distribution. Its format is T30,I2. In Fig. C.2 the initial walker distribution is set to 00 specifying a random dispersion around the equilibrium geometry. Alternatively, the initial walker distribution may be set to 01 specifying the initial walker distribution to be read from input file fort.1. Each line in fort.1 specifies the cartesian coordinates of one atom and has the format 3F25.20. Each walker is specified in NPAR lines where NPAR is the number of atoms in the molecule. The file has, therefore, NCONF*NPAR lines where NCOF is the number of random walkers (configurations).

Line 7 specifies whether the matrix elements (and other QMC walk data) are written to a file to be used as input for a continuing QMC calculation. It format is T30,I2. In Fig. C.2 the dump results of phase option is turn off. By setting it to 02, 03 or 04 the raw data from the VMC, single projection CFQMC or multiple projection CFQMC phases would be written to file fort.56.

Line 8 specifies whether the matrix elements are read (and other QMC walk data) in order to continue a QMC calculation. Its format is T30,I2. In Fig. C.2 the restart phase option is turned off. By setting it to 02, 03 or 04 the raw data from

Figure C.3: QMC input file lines 15 to 20

```
------------------------- EQUILIBRIUM GEOMETRY --------------------
NUM  ---- X COORDINATE ----|--- Y COORDINATE ----|--- Z COORDINATE ---
ATOMIC
06  0.0000000000000000000 0.0000000000000000000 0.0000000000000000000
06  2.4372298870000000000 0.0000000000000000000 0.0000000000000000000
06 -2.3457246620000000000 0.6615677720000000000 0.0000000000000000000
```

the VMC, single projection CFQMC or multiple projection CFQMC phases would be read in from fort.55.

Line 9 specifies the timestep of the MC random walk (VMC or QMC). Its format is T30,F10.4. In Fig. C.2 the time step, $\delta\tau$, is set to 1500 H^{-1}.

Line 10 specifies the CFQMC correlation interval, the period of projection time between the evaluation of the CFQMC matrices. Its format is T30,I2. In Fig. C.2 the CFQMC correlation interval is set to a 6 step period. This option is relevant only when the multiple projection time phase option is selected. It is described in subsection 2.2.2.2.

Line 11 specifies the CFQMC correlation interval, the period in imaginary time of the random walk between the evaluation of all CFQMC matrices. Its format is T30,I2. In Fig. C.2 the CFQMC sampling interval is set to a 5 step period. This option is relevant only when the multiple projection time phase option is selected and is described in subsection 2.2.2.2.

Line 12 specifies whether the eigenvectors are computed and printed in VMC runs. Its format is T30,I2. Presently, the program always computes the eigenvectors.

Line 13 specifies whether the discrete potential option is used. Its format is T30,I2. In Fig. C.2 DPR is turned off.

Line 14 specifies whether the debugging print flags are turned on. Its format is T30,10(I1,X). Presently, they are commented out of the program for computational efficiency. Visual debugging tools are recommended.

Lines 15,16 and 17 are titles. The format is A80.

Line 18,19 and 20 specify the atomic number and equilibrium cartesian coordinates of the each atom in the molecule. The format is I2, 3F25.20. In Fig C.3 the atomic number for each atom is set to 6. The sample input file is for C_3. QMCVIB must be compiled with the parameters NPAR = 3 (the number of atoms) and NMOD = 3 (3*NPAR - 6).

Line 21 is a title. Its format is A80.

Line 22 specifies the guiding function power, ρ in Eq. (3.5)and in GPOW of QMCVIB. The parameter ρ specifies the relative breadth of the guiding function compared to the LM ground state wavefunction, Eq. (3.3). Its format is T30,F5.2. In Fig. C.4 GPOW is set at 0.15. Values of GPOW > 1 are not recommended. GPOW is needed only for LM basis set calculations.

Line 23 specifies the guiding function tolerance, the minimum value of the guiding function for an accepted step. Is format is T30,E9.2E2. In Fig. C.4 PSITOL is set to 10^{-35}.

Line 24 specifies the wavefunction form for multi-state VMC and CFQMC calculations. Its format is T30,I2. In Fig. C.4 IFORM is set to 00 specifying that Local Mode basis function will be used in the VMC calculation. Alternatively, IFORM may have been set to 00 or 02, specifying SMOL or DG basis functions respectively.

Line 25 specifies whether Harmonic couplings are used in conjunction with SMOL and DG Basis functions. If they are used they will be specified below. Its format is T30,I2. In Fig. C.4 the Harmonic couplings are turned on.

Line 26 specifies whether the basis function order parameters are read from fort.30. Its format is T30,I2. In Fig. C.4 IWALK(10) is 00 specifying that the IPOW array will be constructed from data specified in lines 29 and 34-46. If IWALK is set to 01, it would specify that the IPOW array would be read. The file fort.30 has one line for each basis function having the basis function order parameters listed with the format T6,NMOD(I2,2x) where NMOD is the number of modes in the molecule.

Line 27 specifies whether the auxiliary basis set parameters that specify the center and width of DG basis function are read in from fort.30. Its format is T30,I2. In Fig. C.4 IWALK(8) is 00 specifying that the auxiliary function will be constructed.

Figure C.4: QMC input file lines 21 to 46

```
-------------- WAVEFUNCTION PARAMETERS * GENERAL --------------------
    GUIDING FUNCTION POWER = 00.15                    GPOW     (T30,F5.2)
GUIDING FUNCTION TOLERANCE = 01.00E-35                PSITOL(T30,E9.2E2)
       BASIS FUNCTION FORM   00  (bernu,smol,dgbhc)  IFORM      (T30,I2)
        HARMONIC COUPLINGS   00  (00=OFF/01=ON)      IWALK(09) (T30,I2)
    READ IPOW OF BASIS SET   00  (from fort.30)      IWALK(10) (T30,I2)
      READ AUX. PARAMETERS   00  (00=NO/01=YES)      IWALK(8)  (T30,I2)
SYMMETRIZE BASIS FUNCTIONS   00                      IWALK(04) (T30,I2)
    MAX POWER OF ALL MODES = 04                      MAXPOW     (T30,I2)
        COEFFICIENT MATRIX = 01    (I,fort.35, 45, 55) IDIA    (T30,I2)
-------------- WAVEFUNCTION PARAMETERS * DIRECT PRODUCT ------------
------------- GUIDING FUNCTION --------------- BASIS FUNCTIONS -------
IPAR JPAR MODE    A-MATRIX       HARMONIC     A_MORSE  B_MORSE  MAXPWR
 001  002  001      -43.0000       0.0000       10.000   04.000      04
 001  003  002      -43.0000       0.0000       10.000   04.000      04
 002  003  003       -6.0000       0.0000        4.000   02.000      04
------------- WAVEFUNCTION PARAMETERS * HARMONIC COUPLINGS -----------
NUMBER OF MODE COUPLINGS  =  006                                (T30,I3)
-------------- GUIDING FUNCTION  BASIS FUNCTIONS --------------------
IMOD JMOD -------- A-MATRIX ----- HARMONIC (T2,I3,T7,I3,T20,
 001  002            -7.000        0.000   F10.4,T35,F10.4)
 001  003             4.000        0.000
 002  001            -7.000        0.000
 002  003             4.000        0.000
 003  001             4.000        0.000
 003  002             4.000        0.000
```

This option is relevant only when a DGBS is used (IFORM = 02). The auxiliary basis set parameters are listed one line per basis function with format T6,NMOD(F6.3,x) in fort.30 following the basis function order parameters.

Line 28 specifies whether the basis set will be symmetrized. The format is T30,I2. This option is designed for C_{2v} triatomic molecules. In Fig. C.4 IWALK(4) is 00 specifying that the basis function will not be symmetrized.

Line 29 specifies the maximum value of parameters in the IPOW array. Its format is T30,I2. In Fig. C.4 MAXPOW is set to 4. The option is relevant when the IPOW array is constructed.

Line 30 specifies whether the transformation matrix, the matrix of wave function coefficients, is the identity matrix (01), read in from fort.35 (03) or read in from fort.45. Its format is T30,I2 (04). This option is used in CFQMC calculations when VMC wavefunctions are used as trial wavefunctions. Each coefficient is listed on a separate line in file fort.26 and each line has the format E40.32E3.

Lines 31-33 are titles. Their format is A80.

Lines 34-36 specify several wavefunction parameters: the guiding function A-matrix parameters, the Harmonic parameters for SMOL and DG basis sets (usually 0), α and β parameters of the SMOL basis function, and the maximum values of IPOW for each mode.[6] Their format is T2,I3,T7,I3,T20,F10.4,T35,F10.4,-T50,F9.3,T60,F9.3,T75,I2.

Line 37 is a title. Its format is A80.

Line 38 specifies the number of non-zero harmonic coupling terms to be read. Its format is T30,I3.

Lines 39-40 are titles. Their format is A80.

Lines 41-46 specify the mode-mode harmonic coupling terms. Their format is T2,I3,T7,I3,T20,F10.4,T35,F10.4.

[6]MAXPOW = 4 and IPOW(IMOD) = 4 for all IMOD specifies 35 basis functions (7 choose 4), Therefore, QMCVIB must be compiled with parameters NBASIS = 35 and NST = 35.

C.4 Execution

The command line for executing the FORTRAN code QMCVIB is 'qmcvib < *input-filename* > *output-filename*'. Since execution times for production calculations are approximately 2-6 hours on a Cray, QMCVIB is typically run in the background or from a queue.

C.5 Output files

Execution of QMCVIB produces, at least, two output files. The standard output includes an echo of the standard input and reports the calculated vibrational state energies. The output file for the multi-state VMC calculation of the vibrational states of C_3 specified above is shown in Figs C.5, C.6, C.7, C.8, C.9.

The final positions of the random walkers are output to fort.12 which has the same format as the input file fort.1. The wavefunction order parameters and auxiliary parameters are output to fort.29 which has the same format as the input file fort.30. The VMC wavefunction coefficients are output to fort.35 which has the same format as the input file fort.26.

Figure C.5: QMCVIB output lines 1 to 29

```
**************************************************************************
                VIBRATIONAL QUANTUM MONTE CARLO
**************************************************************************
          START DATE  01/03/96        START TIME  02:43:35
          RANDOM NUMBER SEED              48131768981101
          ------------------------------------------------
                          RUN PARAMETERS
          ------------------------------------------------
               NUMBER OF BLOCKS             16
               BLOCK SIZE                   16    WALKERS
               ENSEMBLE SIZE               256    WALKERS
               TIME STEP                  1200.000 HARTREE-1
               WALK LENGTH      20000     20000        10        10
               NUM PROJ TIMES               1
               NO. OF STATES               35
               NO. BASIS FUNC.             35
          ------------------------------------------------
                   ADDITIONAL INPUT INFORMATION
          ------------------------------------------------
               PRINT FLAGS           00000000000000000000
               WALK OPTIONS                     0000000000
               QMC CTRL FLAGS        1100
               C MATRIX              UNIT MATRIX
               INITIAL COORD             0
               LSTEP                     6       STEPS
               EIGENVEC              CALC
               NUMBER OF POOLED WALKERS     0
               WAVEFUNCTION CUTOFF   0.10E-34
```

Figure C.6: QMCVIB output lines 30 to 57

```
-------------------------------------------

            EQUILIBRIUM GEOMETRY

-------------------------------------------

AT. NO.        X           Y           Z
   6        0.000000    0.000000    0.000000
   6        2.437230    0.000000    0.000000
   6       -2.345725    0.661568    0.000000

-------------------------------------------

         GUIDING FUNCTION INFORMATION

-------------------------------------------

     GUID FUNC EXP              0.15
     A( 1, 1) =        -43.0000
     A( 1, 2) =         -7.0000
     A( 1, 3) =          4.0000
     A( 2, 2) =        -43.0000
     A( 3, 3) =         -6.0000

-------------------------------------------

            BASIS SET INFORMATION

-------------------------------------------

WAVEFUNCTION FORM:  BERNU
            HARMONIC COUPLING CONSTANTS
     CH( 1, 1) =          0.0000
     CH( 1, 2) =          0.0000
     CH( 1, 3) =          0.0000
     CH( 2, 2) =          0.0000
     CH( 3, 3) =          0.0000
     MAX QUANTA/BASIS FUNC         4
     MAX QUANTA/MODE         4  4  4
```

Figure C.7: QMCVIB output lines 58 to 95

1	0	0	0
2	1	0	0
3	0	1	0
4	0	0	1
5	2	0	0
6	1	1	0
7	1	0	1
8	0	2	0
9	0	1	1
10	0	0	2
11	3	0	0
12	2	1	0
13	2	0	1
14	1	2	0
15	1	1	1
16	1	0	2
17	0	3	0
18	0	2	1
19	0	1	2
20	0	0	3
21	4	0	0
22	3	1	0
.	.	.	.
.	.	.	.
.	.	.	.
34	0	1	3
35	0	0	4

Figure C.8: QMCVIB output lines 96 to 122

```
------------------------------------------------

                EQUILIBRATION RUN

------------------------------------------------

ELAPSED TIME:          0.06      TOTAL TIME:          0.06
ELAPSED TIME:        388.19      TOTAL TIME:        388.25
                HISTOGRAM OF GUIDING FUNCTION DISTRIBUTION
 MODE =  1   0.0E+00   0.0E+00   0.4E+04   0.2E+07   0.3E+07   0.1E+05
 MODE =  2   0.0E+00   0.0E+00   0.8E+03   0.2E+07   0.3E+07   0.9E+04
 MODE =  3   0.2E+04   0.7E+05   0.9E+06   0.3E+07   0.1E+07   0.1E+05

          GF ENERGY 0.2765E-01   +/-  0.6532E-09 HARTREES
          GF ENERGY 0.6068E+04   +/-  0.1434E-03 CM**-1
          DELG**2     -0.03374   +/-     0.00205 HARTREES
          ACCEPTANCE RATIO       0.50089

------------------------------------------------

          VARIATIONAL DIAGONALIZATION RUN

------------------------------------------------

ELAPSED TIME:          0.11      TOTAL TIME:        388.36
ELAPSED TIME:       1817.48      TOTAL TIME:       2205.84
                ISUM = 16   IERR = 0
ELAPSED TIME:          0.44      TOTAL TIME:       2206.28
                HISTOGRAM OF GUIDING FUNCTION DISTRIBUTION
 MODE =  1   0.0E+00   0.0E+00   0.7E+04   0.4E+07   0.6E+07   0.2E+05
 MODE =  2   0.0E+00   0.0E+00   0.2E+04   0.4E+07   0.6E+07   0.2E+05
 MODE =  3   0.3E+04   0.1E+06   0.2E+07   0.6E+07   0.2E+07   0.2E+05
```

Figure C.9: QMCVIB output lines 123 to 167

GF ENERGY 0.2769E-01 +/- 0.6733E-09 HARTREES
GF ENERGY 0.6078E+04 +/- 0.1478E-03 CM**-1
DELG**2 -0.03374 +/- 0.00205 HARTREES
ACCEPTANCE RATIO 0.50091

STATE	ENERGY(H)	ENERGY(CM-1)	STAT.ERR(CM-1)	BIAS(CM-1)
0	0.00760	1667.622	0.386	0.618
1	0.00823	1807.191	0.475	0.538
2	0.00893	1960.012	0.778	0.547
3	0.00994	2182.194	2.290	0.560
4	0.01163	2552.854	7.987	0.900
5	0.01314	2884.641	0.847	0.401
6	0.01409	3093.151	1.768	1.005
7	0.01521	3338.334	4.435	2.439
8	0.01683	3693.607	1.466	0.669
9	0.01701	3733.455	13.507	11.718
10	0.01725	3786.294	0.598	0.289
11	0.01799	3947.654	1.696	0.260
12	0.01875	4115.078	2.064	1.017
13	0.01951	4281.881	4.165	1.672
14	0.02019	4430.881	10.515	3.747
15	0.02210	4851.476	32.823	12.473
16	0.02248	4934.224	0.792	0.035
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
33	0.04298	9433.925	8.698	-0.750
34	0.04605	10106.271	6.466	-0.807

ELAPSED TIME: 0.44 TOTAL TIME: 2206.72

Bibliography

[1] B. T. Sutcliffe. *Methods in Computational Chemistry: The Mathematics of Vibration-Rotation Calculations*, volume 4, pages 33–89. Plenum Press, New York, 1992.

[2] A. Szabo and N.S. Ostlund. *Modern Quantum Chemistry. Introduction to Advanced Electronic Structure Theory.* McGraw-Hill Publishing Company, second edition, 1989.

[3] S. Carter and N.C. Handy. The Variational Method for the Calculation of Ro-Vibrational Energy Levels. *Computer Physics Reports*, 5:115–172, 1986.

[4] J.N. Murrell, S.C. Farantos S. Carter and, P. Huxley, and A.J.C. Varandas. *Molecular Potential Energy Functions.* John Wiley and Sons, first edition, 1984.

[5] K. S. Sorbie and J. N. Murrell. Analytical potentials for triatomic molecules from spectroscopic data. *Molecular Physics*, 29:1387–1407, 1975.

[6] J. M. Bowman, A. Wierzbicki, and J. Zúniga. Exact vibrational energies of non-rotating H_2O and D_2O using an accurate ab initio potential. *Chemical Physics Letters*, 150:269–274, 1988.

[7] P. Jensen, C. M. Rohlfing, and J. Almlöf. Calculation of the complete-active-space self-consistent-field potential-energy surface, the dipole moment surfaces, the rotation-vibration energies, and the vibrational transition moments for $C_3(X^1\Sigma_g^+)$. *J. Chem. Phys.*, 97:3399–3411, 1992.

[8] M. Mladenović, S. Schmatz, and P. Botschwina. Large Scale *ab initio* calculations for C_3. *J. Chem. Phys.*, 101:5891–5899, 1994.

[9] H. Romanowski, J. M. Bowman, and L. B. Harding. Vibrational energy levels of formaldehyde. *J. Chem. Phys.*, 82:4155–4165, 1985.

[10] J. M. L. Martin, Timothy J. Lee, and Peter R. Taylor. An accurate *ab initio* quartic force field for ammonia. *J. Chem. Phys.*, 82:8361–8371, 1992.

[11] M. R. Aliev and J. K. G. Watson. Modern Research. In *Higher order effects in the vibration rotation spectra of semirigid molecules*, volume 3, chapter 1. Academic Press, New York, 1985.

[12] J. K. L. MacDonald. Successive Approximation by the Rayleigh-Ritz Variation Method. *Physical Review*, 43:830–833, 1933.

[13] E. L. Sibert III. Theoretical studies of vibrationally excited polyatomic molecules using canonical Van Vleck perturbation theory. *J. Chem. Phys.*, 88:4378–4390, 1988.

[14] P. J. Jensen. A new Morse oscillator-rigid bender internal dynamics (MORBID) Hamiltonian for Triatomic Molecules. *J. Mol. Spectrosc.*, 128:478–501, 1988.

[15] I. P. Hamilton and J. C. Light. On distributed gaussians basis for simple model multidimensional vibrational problems. *J. Chem. Phys.*, 84:306–317, 1986.

[16] Z. Bačić and J. C. Light. Highly excited vibrational levels of "floppy" triatomic molecules: A discrete variable representaion-distributed gaussian approach. *J. Chem. Phys.*, 85:4594–4604, 1986.

[17] W. Yang and A. C. Peet. The collaction method for bound solutions of the Schrödinger equation. *Chem. Phys. Lett.*, 153:98–104, 1988.

[18] M. J. Bramley, J. W. Tromp, T. Carrington, and G. C. Corey. Efficient calculation of highly exicted vibrational energy levels of floppy molecules: The band origins of H_3^+ up to $35,000\,cm^{-1}$. *J. Chem. Phys.*, 100:6175–6194, 1994.

[19] D. M. Ceperley and B. Bernu. The calculation of excited state properties with quantum Monte Carlo. *J. Chem. Phys.*, 89:6316–6328, 1988.

[20] B. Bernu, D. M. Ceperley, and W. A. Lester Jr. The calculation of excited state properties with quantum Monte Carlo. II. Vibrational excited states. *J. Chem. Phys.*, 100:552–561, 1990.

[21] E. L. Sibert III. Variational and perturbative descriptions of highly vibrationally excited molecules. *International Revies in Physical Chemistry*, 9:1–27, 1990.

[22] P. Jensen. The potential energy surface for the C_3 molecule determined from experimental data. Evidence for a bent equilibrium structure. *Collection of Czechoslovak Chemical Communications*, 54:1209–1218, 1989.

[23] Z. Bačić and J. C. Light. Accurate localized and delocalized vibrational states of HCN/HNC. *J. Chem. Phys.*, 86:3065–3077, 1987.

[24] M. Mladenovic, S. Schmatz, and P. Botschwina. Large-scale *ab initio* calculations for C_3. *J. Chem. Phys.*, 101:5891–5899, 1994.

[25] R. C. Cohen and R. J. Saykally. Extending the collocation method to multidimensional molecular dynamics: Direct determination of the intermolecular potential of $Ar - H_2O$ from a tunable far-infrared laser spectroscopy. *J. Chem. Phys.*, 94:7991, 1990.

[26] W. A. Lester Jr. and B.L. Hammond. Quntum Monte Carlo for the electronic structure of atoms and molecules. *Annual Review of Physical Chemistry*, 41:283–311, 1990.

[27] B.L. Hammond, W.A. Lester Jr., and P.J. Reynolds. *Monte Carlo Methods in Ab Initio Quantum Chemistry.* World Scientific, first edition, 1994.

[28] J. B. Anderson. Fixed-Node Quantum Monte Carlo. *International Reviews in Physical Chemistry*, 14:85–112, 1995.

[29] J. B. Anderson. Random-walk simulation of the Schrödinger equation. *J. Chem. Phys.*, 63:1499–1503, 1975.

[30] J.W. Moskowitz and K.E. Schmidt. The domain Green's function method. *J. Chem. Phys.*, 85:2868–2874, 1986.

[31] P.J. Reynolds, R.K. Owen, and W.A. Lester Jr. Is there a zeroth order time-step error in diffusion quantum Monte Carlo? *J. Chem. Phys.*, 87:1905–1906, 1987.

[32] J. B. Anderson and D.R. Garner. Vailidity of random-walk methods in the limit of small time steps. *J. Chem. Phys.*, 87:1903–1904, 1987.

[33] S.M. Rothstein and J. Vrbik. A Green's function used in diffusion Monte Carlo. *J. Chem. Phys.*, 87:1902–1903, 1987.

[34] J.W. Moskowitz and K.E. Schmidt. Erratum: the domain Green's function method. *J. Chem. Phys.*, 87:1906, 1987.

[35] L. E. Reichl. *A Modern Course in Statistical Physics*, pages 167–168. University of Texas Press, first edition, 1987.

[36] B. N. Parlett. *The Symmetric Eigenvalue Problem.* Prentice-Hall, 1980.

[37] D. F. Smith and J. Overend. Anharmonic force constants of water. *Spectrochim. Acta*, 28A:477–483, 1972.

[38] J.S. Garing and R. A. McClatchey. Atmospheric absorption line compilation. *Applied Optics*, 12:2545, 1973.

[39] J.-Y. Mandin, J.-P. Chevillard, C. Camy-Peyret, and J.-M. Flaud. High resolution spectrum of water vapor between 13200 and 16500 cm^{-1}. *J. Mol. Spectroscopy*, 116:167–190, 1986.

[40] C. Camy-Peyret, J.-M. Flaud, J.-Y. Mandin, J.-P. Chevillard, J. Brault, D. A. Ramsay, M Veroet, and J. Chauville. High resolution spectrum of water vapor between 16500 and 25250 cm^{-1}. *J. Mol. Spectroscopy*, 113:208–228, 1985.

[41] A.R. Hoy, I. M. Mills, and G. Strey. Anharmonic force constant calculations. *Mol. Phys.*, 24:1265–1290, 1972.

[42] K. S. Sorbie and J. N. Murrell. Analytical potentials for triatomic molecules from spectroscopic data. *Mol. Phys.*, 29:1387–1407, 1975.

[43] K. S. Sorbie and J. N. Murrell. Theoretical study of the $O(^1D) + H_2(^1\Sigma_g^+)$ reactive quenching process. *Mol. Phys.*, 31:905–920, 1976.

[44] G. D. Carney, L. A. Curtiss, and S. R. Langoff. Improved potential function for bent Ab_2 molecules: water and ozone. *J. Mol. Spectroscopy*, 61:371–381, 1976.

[45] S. Carter and N. C. Handy. A theoretical determination of the rovibrational energy levels of the water molecule. *J. Chem. Phys.*, 87:4294–4301, 1987.

[46] B. T. Darling and D. M. Dennison. The water vapor molecule. *Physical Review*, 57:128, 1940.

[47] B. R. Johnson and W. P. Reinhardt. Adiabatic separations of stretching and bending vibrations: Applications to H_2). *J. Chem. Phys.*, 85:4538–4556, 1986.

[48] M. Caffarel, F. X. Gadea, and D. M. Ceperley. Lanczos-type algorithm for quantum Monte Carlo Data. *Europhysics Letters*, 16:249–254, 1991.

[49] W. Weltner and R. J. Van Zee. Carbon Molecules, Ions and Clusters. *Chem. Rev.*, 89:1713–1747, 1989.

[50] J. Szczepanski and M. Vala. The $\nu_1 + \nu_3$ combination mode of C_3 in Ar and Kr matrices: Evidence for a bent structure. *J. Chem. Phys.*, 99:7371–7375, 1993.

[51] W. P. Kraemer, P. R. Bunker, and M. Yoshimine. A theoretical study of the rotation-vibration energy levels and dipole moment of CCN^+, CNC^+, and C_3. *Journal of Molecular Spectroscopy*, 107:191–207, 1984.

[52] P. Jensen and W. P. Kraemer. A comparison of Perturbative and variational rotation-vibration energies for HOC^+ and C_3 using the non-rigid bender and MORBID Hamiltonians. *Journal of Molecular Spectroscopy*, 129:172–185, 1988.

[53] P. Jensen. A new Morse oscillator-rigid bender internal dynamics (MORBID) Hamiltonian for triatomic molecules. *Journal of Molecular Spectroscopy*, 128:478–501, 1988.

[54] F. J. Northrup and T. J. Sears. Stimulated-emmission pumping spectroscopy study of jet-cooled C_3: pure bending levels and bend-symmentric stretch combination levels of $X^1\Sigma_g^+$. *Optical Society of America*, 7:1924–1934, 1990.

[55] K. Kawaguchi, K. Matsumura, H. Kanamori, and E. Hirota. Diode laser spectroscopy of C_3: the $\nu_2 + \nu_3 - \nu_2$, $2\nu_2 + \nu_3 - 2\nu_2$, and $2\nu_2 + \nu_3$ bands. *J. Chem. Phys.*, 91:1953–1957, 1989.

[56] J. Tennyson, S. Miller, and J. R. Henderson. The calculation of Highly excited vibration-rotation states of triatomic molecules. *Methods in Computational Chemistry*, 4: Molecular Vibrations:91, 1992.

[57] J. Tennyson and B. T. Sutcliffe. On the rovibrational levels of the H_3^+ and H_2D^+ molecules. *J. Mol. Spectroscopy*, 51:887–906, 1984.

[58] W. R. Brown, W. A. Lester Jr., and W. A. Glauser. Quantum Monte Carlo for floppy molecules: Vibrational states of C_3. *J. Chem. Phys.*, 91:9721–9725, 1995.

[59] F. J. Northrup, T. J. Sears, and E. Rohlfing. A semirigid bender analysis of an extensive set of rotation-vibration levels in $X^1\Sigma_g^+$ C_3. *Journal of Molecular Spectroscopy*, 145:74–88, 1991.

[60] A. E. Rohlfing. Laser-induced-fluoresence spectroscopy of jet-cooled C_3. *J. Chem. Phys.*, 91:4531–4542, 1989.

[61] Martin A. Suhm. Multidimensional vibrational quatum Monte Carlo technique using robust interpolation from static or growing sets of discrete potential points. *Chemical Physics Letters*, 214:373–380, 1993.

[62] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in FORTRAN: the art of scientific computing.* Cambridge University Press, second edition, 1992.

[63] I. P. Hamilton and J. C. Light. On distributed Gaussian bases for simple model multidimensional vibrational problems. *J. Chem. Phys.*, 84:306–317, 1986.

[64] J. Faeder. A distributed Gaussian approach to the vibrational dynamics of Ar-benzene. *J. Chem. Phys.*, 99:7764–7676, 1993.

[65] J. M. Bowman, A. Wierzbicki, and J. Zúniga. Exact vibrational energies of non-rotating H_2O and D_2O using an accurate *ab initio* potential. *Chemical Physics Letters*, 150:269–274, 1988.

[66] D. F. Coker and R. O. Watts. Structure and vibrational Spectroscopy of the water dimer using quantum simulation. *J. Chem. Phys.*, 91:2513–2518, 1987.

Ernest Orlando Lawrence Berkeley National Laboratory
One Cyclotron Road | Berkeley, California 94720

Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098