Approaches in Deciphering Zero-Point Energy Effects for Molecular Clusters Using Diffusion Monte Carlo

Victor Guang Ming Lee

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Anne B. McCoy, Chair

Munira Khalil

Xiaosong Li

Program Authorized to Offer Degree:
Department of Chemistry

University of Washington

Abstract

Approaches in Deciphering Zero-Point Energy Effects for Molecular Clusters Using Diffusion Monte Carlo

Victor Guang Ming Lee

Chair of the Supervisory Committee:
Professor Anne B. McCoy
Department of Chemistry

Deciphering the differences of the structures formed with bulk water and the effects of isotopic substitution has been of interest. Studies of water clusters provide a connection to the intermediate cases between a water monomer and bulk water. Water clusters are an attractive system to study due to the fact that experimental and theoretical methods can be used to study the same cluster. However, one main challenge are the amounts of low-lying minima that are exhibited as the size of the cluster increases. When only the electronic energy is considered, finding the minimum energy structure of these clusters is an active field of work. However, when zero-point energy is added, determining the minimum energy structure of a particular sized cluster becomes even more difficult. Deuteration can also change the structures that are relevant in the vibrational ground state. One prime example is the water hexamer, where the minimum energy structure is the prism. Once zero-point energy is considered, the minimum energy structure is the cage. However, the deuterated form of water hexamer has been assigned as the prism structure based on microwave spectroscopy. Diffusion Monte Carlo (DMC) is a method that has been used to study water clusters in the vibrational ground state. However, large ensemble sizes are needed to ensure reliable results. For these calculations, this is due to the nature of

the couplings between the high and low frequency vibrations. One method to reduce the ensemble sizes needed for these calculations is to introduce a guiding function. The guiding function, Ψ_T, is used to enhance sampling in the relevant regions of the potential energy surface. When a guiding function is introduced into the DMC algorithm, the local energy, $\hat{H}\Psi_T/\Psi_T$, is evaluated, rather than just the potential energy.

In this thesis, I first explore the choice of Ψ_T and use this approach to evaluate the overlap integral of Ψ_T and the wave function calculated using DMC. I show that even using guiding functions as simple as a product of harmonic oscillators can accurately compute the ground state energies of extremely anharmonic molecules, such as the partially deuterated analogues of H_3^+. Based on these results, I construct a simple guiding function that describes the intramolecular vibrations of each isolated water monomer within a water cluster. I find that the ensemble size needed to calculate ground state energies and ground state wave functions of various sized water clusters is reduced by at least an order of magnitude. This DMC approach can be extended beyond water clusters, and can be used to study a broad range of molecules that exhibit couplings in the high and low frequency vibrations. I use this guiding function to study the effects of isotopic substitution in water hexamer. I find that for $(H_2O)_6$, the ground state wave function is the cage. I also find that $(D_2O)_6$ is also a cage, but the ground state energy difference between the cage and prism structures is only 9 cm^{-1}. These studies provide the groundwork for further investigation of the strength of hydrogen bonds in water hexamer.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

It takes a village to raise a graduate student. I would like to thank my village, whom have supported me throughout the process of obtaining a Ph.D. While, I would love to acknowledge every person in my village, but that could possibly double the length of my thesis. Here is a small, but important subset of groups and people I would like to acknowledge.

First and foremost, I would like to thank God for the honor and privilege to study for a Ph.D. Without His provisions and guidance, I would not be where I am today. I am humbled and blessed by the opportunities in how He has opened so many doors for me in the past and has continued to do so today.

I would like to thank my Ph.D advisor, Professor Anne B. McCoy for putting up with me for the past five years. Thank you for being a caring advisor and continuing to invest in me and being extremely patient with my many shortcomings.

In terms of my research productivity, this work would have been impossible without the entirety of the McCoy group. I would like to thank in no particular order, Drs. Lindsey Madison, Laura Dzugan, Bernice Opoku-Agyeman, Meng Huang during their time here. I would also like to thank Meredith Fore, Sarah Bolt, Kyle McGrath, Mark Boyer, Ryan DiRisio, Jacob Finney, Rachel Huchmala, Luke Hatcher, Chloe Chiu, Mathew Joyner, and Nick Vetterli. I also want to thank Joe Heindel and Kristina Herman in the Xantheas group. All of these people have played a critical role in my maturation as a graduate student.

I would like to thank my friends in the department that have carried me through, in particular, Jessica Kong, Trevor Johnston, and Ben Figueroa. I would like to also

thank my friends at Marketplace/Common Good Church for their support as well as helping me develop a side hobby, skiing. In particular, I would like to thank David Lee, Mariko Sandico-Lee, and Bryce Bowers for going out of your way to care for me. I would like to thank all of my friends back home in California that have kept in touch with me over the course of my time in grad school. Thank you for reaching out to me and for many of you, thank you for visiting me in Seattle and reminding me that you care enough to say hi when you're in town.

I would have never considered obtaining a degree in chemistry without the motivation of my high school chemistry teacher, Mr. Jonathan Fong in Lowell High School. Likewise, I would have never considered going for a Ph.D without the inspiration from Professor Hrant Hratchian at UC Merced. To the both of you, thank you so much for inspiring me to pursue a Ph.D.

Last and most definitely not least, I would like to thank my family and all of their sacrifices that they have made to give me a better life. I want to thank my parents in particular for dealing with me during my time here in UW. Although we were miles apart, thank you for showing your love through the constant checkups and always willing to listen to me vent during my time as a graduate student. I would also like to thank my sister for always looking out for me and keep my head in check.

DEDICATION

To my parents, Benley Kim Hong Lee, and Li Chang Huang.

Chapter 1
INTRODUCTION

1.1 Models Describing Molecular Vibrations

Quantum mechanical calculations of vibrational energies and wave functions provide unique challenges for chemists.[1–12] While the potential energy for a molecular system can be described as a sum of two-body Coulombic interactions, once the Born-Oppenheimer separation of the electronic and nuclear degrees of freedom is introduced there is no longer a simple and general functional form for the potential energy for the nuclear degrees of freedom.[13] Within the Born-Oppenheimer approximation, the potential energy for the electronic degrees of freedom remains Coulombic, and the electronic wave function can be expanded as a product of functions that are expressed in terms of the three spherical coordinates of each electron. This expansion is based on the linear combination of atomic orbitals (LCAO) model and provides important insights into the electronic structure of a broad range of chemical systems.

Just as the potential energy surface for the nuclear degrees of freedom cannot be expressed in a simple generic form, it is difficult to identify a single set of coordinates and basis functions that are appropriate for all vibrational problems. If we could solve for the vibrational energies and wave functions exactly, the choice of coordinates would not affect the final results. Practically, numerical evaluations of the energies and wave functions require a choice of coordinates and basis functions. The convergence properties of most vibrational calculations depend on this choice.[14]

Once the coordinates are determined, the simplest model that can be used to describe molecular vibrations is the harmonic oscillator. This simple model provides an insight of the vibrational energies and the associated motions. This model works

well in describing vibrations that are considered to be harmonic, such as a stretching motion of a diatomic molecule. This simple model breaks down when the motions that are described are not harmonic, with one example as torsional motions. As a result, we need to turn to more advanced treatments to describe the vibrational motions of molecules.

There are a multitude of treatments that are available to provide a more accurate description of molecular vibrations than the harmonic oscillator. We can use variational methods to solve for the vibrational energies, which are exact, in the limit of a large basis set. This approach scales poorly with system size and limits the problems that can be solved to only to small molecular systems, such as CH_4.[15] On the other hand, 2nd Order Vibrational Perturbation Theory (VPT2) is a computationally viable option as long as the perturbation applied onto the model is small. However, this approach fails when there are degenerate vibrational motions in the molecule, or when the initial model is a poor description of the associated vibration.

One method that provides an accurate description of the vibrational motions of the ground state wave function and generally has favorable scaling as the size of the system increases is diffusion Monte Carlo (DMC).[16–19] In this method, Monte Carlo approaches are used to evaluate the ground state wave function and ground state energy of the system of interest. Due to the fact that DMC is less sensitive to the choice of coordinates, these calculations are typically performed in Cartesian coordinates, and the wave functions are expanded in a basis of localized functions of the $3N$ Cartesian coordinates that define the nuclear geometry.[20,21] As a result, DMC is used to calculate the ground state wave function of a molecule, which can provide insight in a molecule's vibrations.

There are several limitations of using DMC to analyze the vibrational motions of a molecule. The first issue is the ability to compare a wave function using DMC to wave functions that are obtained from different methods. In order to compare the accuracy of a wave function using a different method to a wave function calculation using DMC,

we evalute an overlap integral. There are methods that have been developed to use DMC to calculate expectation values, but evaluating matrix elements with DMC is non-trivial.[22,23] As we focus on the simplest case, where the operator is the identity, the matrix element is simply an overlap integral. Being able to evaluate an overlap integrals opens the door to use DMC to calculate matrix elements of multiplicative operators.

The second limitation of DMC is the ensemble size that is needed for the convergence of the ground state energies and the ground state wave function. DMC can be used to calculate accurate ground state energies and ground state wave function of any molecule in the limit of an infinitely large ensemble. Typically speaking, the ensemble sizes that are needed range on the order of 10^4-10^5 walkers. However, the ensemble sizes that are needed can be quite large for larger molecules, such as water clusters. For example, Mallory and Mandelshtam[24,25] have calculated ground state energies of water clusters and they have shown that the ensemble sizes needed to obtain reliable results dramatically increases as a function of the cluster size, for which they have used up to a million walkers to calculate the ground state energy of water hexamer. These calculations make studies of these molecular systems a computational challenge to calculate the ground state energy and the ground state wave function of a larger water cluster.

In order to address both limitations, we use a variant of DMC that involves the use of a guiding function. This variant also known as importance sampled DMC. Importance sampled DMC has been used in studies of H_5^+[26] and $H_5O_2^+$[27], where comparisons to the results of calculations that do not employ guiding functions can be made. In the first limitation, importance sampled DMC provides a Monte Carlo sampling of the product of the guiding function and the wave function calculated using DMC. This measure provides a method for using DMC to calculate the overlap between two wave functions. In the second limitation, we use a guiding function to accurately describe the intramolecular vibrations, such as the OH stretches and the

HOH bend in the individual water molecule. We show that using a guiding function that only describes the intramolecular vibrations drastically reduces the ensemble sizes needed for these DMC calculations.

1.2 Nuclear Quantum Effects in Water Clusters

As we are able to reduce the ensemble sizes needed to calculate accurate ground state energies and ground state wave functions, we can explore the structures of water clusters upon isotopic substitution. In the limit of bulk water, calculations of the radial distribution function indicate the differences in the structures sampled in the bulk water makeup of (H_2O) and (D_2O).[28] Water clusters provide a way to study the effects of hydrogen bonding in water molecules as seen in bulk water. Water clusters are also an attractive molecular system to study due to the direct comparisons that can be made with experimental spectra and theoretical calculations. Experimental work involving water clusters have led to multiple collaborations with computer simulations to further understand the underlying structures of a particularly-sized water cluster, and make predictions on the structures on larger clusters.[29,30]

From an electronic structure standpoint, finding the minimum energy structure of a water cluster is challenging, and is currently still an active field of research.[31] However, it is unclear if the minimum energy structure changes when quantum zero-point effects are taken into account. That is due to the fact that there are multiple low-lying minima on the potential energy surface. One particular case is the water hexamer, as it is the smallest cluster to exhibit three-dimensional hydrogen bonding as its minimum energy structure, as the hydrogen bonding for the minimum energy structures for the trimer to the pentamer are cyclic. When only the electronic energy is taken into account, the minimum energy structure is the prism. However, once zero-point energy is added, the minimum energy structure becomes the cage, which has been confirmed with experimental findings.[29,32] Upon isotopic substitution, when deuterium is substituted in place of hydrogen, the minimum energy structure has

been shown to be a prism.[33] With nuclear quantum effects playing a large role for the structures in water clusters, we seek to understand the role of isotopic substitution on the different structures of water hexamer. We will do so with partial deuteration and determine the effect of hydrogen bonding based on the location of the D_2O molecule. Before proceeding to the water hexamer, we analyze the effects of isotopic substitution for water trimer, as the trimer is a smallest cluster that exhibit multiple low-lying minima. Using the findings from the trimer, we extend our studies to the hexamer.

1.3 Outline of Thesis

In this thesis, we will use DMC to calculate the ground state energies and the ground state wave function for several molecules. We review the theory needed to be able to understand the two variants of DMC that are used in this thesis. In Chapter 2, an example of a one-dimensional model system is provided to illustrate the results that are obtained using DMC.

We test the use of different guiding functions in Chapter 3, where we use DMC to the calculate overlap integrals to compare the accuracy of wave functions calculated using different methods. We first compare the accuracy of the numerical accuracy of the calculation for the overlap integral by exploring the various parameters discussed in Chapter 2 for a one-dimensional model system. We extend our approach to calculate overlap integrals of the ground state wave function of a molecular system using several choices guiding functions that range from products of harmonic oscillators to variational approaches that are highly accurate.

We further explore the use of the guiding function in calculating accurate ground state energies of water clusters. We show in Chapter 4 that using simple guiding functions can decrease the ensemble size needed to obtain reliable results from DMC simulations with one to three water molecules. Once we are able to show the savings that can be achieved using guided DMC, we perform a detailed study by partially deuterating the water trimer to understand the differences in the structures that are

sampled upon isotopic substitution.

We extend the use of the guiding function defined in Chapter 4 for the water hexamer, which is detailed in Chapter 5. We show that that the reduction in ensemble sizes needed to achieve the convergence can be extended to larger clusters. We explore the various effects of isotopic substitution such as fully deuterating the hexamer to confirm what is the minimum energy structure. We performed studies of partially deuterating the hexamer to understand the effects of hydrogen bonding in partially deuterated analogues of the water hexamer.

Chapter 2

THE DIFFUSION MONTE CARLO ALGORITHM WITH AND WITHOUT IMPORTANCE SAMPLING

2.1 Diffusion Monte Carlo Without Importance Sampling

We begin with the simplest flavor of diffusion Monte Carlo (DMC). In this approach, we solve the time-dependent Schrödinger equation for a system of N atoms, which has been rotated to imaginary time through a Wick rotation

$$\frac{\partial \Phi}{\partial \tau} = \left(\sum_{i=1}^{3N} \frac{\hbar^2}{2m_i} \frac{\partial^2}{\partial x_i^2} - V(\mathbf{x}) \right) \Phi(\mathbf{x}, \tau) \tag{2.1}$$

As such, $\tau = it/\hbar$. The solution to Eq. 2.1 can be expanded as a linear superposition of the eigenstates of the Hamiltonian,

$$\Phi(\mathbf{x}, \tau) = \sum_n c_n e^{-E_n \tau} \phi_n(\mathbf{x}) \tag{2.2}$$

where $\phi_n(\mathbf{x})$ represents the eigenstate with energy E_n. Analysis of Eq. 2.2 shows that in the limit of large τ, the ground state, $\phi_0(\mathbf{x})$, provides the leading contribution to $\Phi(\mathbf{x}, \tau)$ as the ratios of the amplitudes of higher energy states to that of the ground state decay to zero as $\exp\left[-\left(E_n - E_0\right)\tau\right]$. In these calculations, the zero in energy used in the propagation is shifted by an amount E_{ref}. This is achieved by multiplying $\Phi(\mathbf{x}, \tau)$ in Eq. 2.2 by

$$S = \exp\left(+E_{\text{ref}}\tau\right) \tag{2.3}$$

When $E_{\text{ref}} = E_0$, the amplitude of the ground state will remain constant, while the

amplitudes of excited states will decay to zero. Based on this observation, the value of E_0 is determined by adjusting the value of E_{ref} during the simulation based on the requirement that the amplitude of $\Phi(\mathbf{x}, \tau)$ remains constant. The time-averaged value of E_{ref}, obtained in this manner, provides the ground state energy of the system of interest.

As noted by Fermi,[34] Eq. 2.1 has the same structure as the diffusion equation, where the diffusion constant for the ith atom is replaced by $\hbar^2/2m_i$, and the potential energy contribution to Eq. 2.1 corresponds to a coordinate-dependent rate constant for a first order rate process. As noted by Anderson and others,[21] the solution to Eq. 2.1 can be obtained using the machinery developed to study diffusion processes, where $\Phi(\mathbf{x}, \tau)$ is replaced by an ensemble of localized functions. These localized functions will be referred to as walkers in the discussion that follows. The ensemble of walkers is propagated over a series of time steps with imaginary time increments represented by $\Delta\tau$.

In each time step, the actions of the kinetic and potential contributions to the Hamiltonian are considered separately. This is analogous to the split operator approach used in quantum dynamics.[35] The kinetic contribution to the Hamiltonian results in the diffusion of each of the walkers in each of the $3N$ three Cartesian coordinates that define its position, x_i, based on a normal distribution,

$$g(x_i) \propto \exp\left[-\frac{x_i^2}{2\sigma_i^2}\right] \tag{2.4}$$

with

$$\sigma_i = \sqrt{\frac{\hbar^2 \Delta\tau}{m_i}} \tag{2.5}$$

The potential energy at the coordinates of the jth walker is used to evaluate the walker's survival probability

$$P_j(\tau) = \exp\left[\left(E_{\text{ref}} - V(\mathbf{x}_j(\tau))\right)\Delta\tau\right] \tag{2.6}$$

In the work presented in this thesis, two approaches are used to account for $P_j(\tau)$. In the first, which is referred to as discrete weighting, each walker contributes equally to the ensemble. In this case, the integer part of $P_j(\tau)$ provides the number of walkers that are placed at the coordinates $\mathbf{x}_j(\tau)$ at the start of the next iteration in the simulation, while the fractional part of $P_j(\tau)$ provides the probability that an additional walker will be added to the ensemble at that geometry. In the second approach, referred to as continuous weighting, the weight of the jth walker, w_j, evolves with τ as

$$w_j(\tau + \Delta\tau) = P_j(\tau) w_j(\tau) \tag{2.7}$$

In both approaches,[21]

$$E_{\text{ref}}(\tau) = \overline{V}(\tau) - \alpha \frac{W(\tau) - W(\tau = 0)}{W(\tau = 0)} \tag{2.8}$$

where α is a simulation parameter, and is equated to $(2\Delta\tau)^{-1}$,[36] $\overline{V}$ is the average of the potential energies of the walkers, while $W(\tau)$ is the sum of the weights of the individual walkers at imaginary time τ.

2.2 Importance Sampling

In many cases unbiased sampling of the wave function, as is described above, provides the most effective approach for implementing DMC.[17,37,38] However, we have found that using an importance sampled approach can improve the precision in the calculation of the ground state energy of a molecular system of interest, as is illustrated in the results plotted in Figure 2.1.

In importance sampled DMC, the function that is sampled in the DMC simulation is $f(\mathbf{x}, \tau)$, which is the product of $\Phi(\mathbf{x}, \tau)$ and $\Psi_T(\mathbf{x})$, rather than $\Phi(\mathbf{x}, \tau)$. In defining $f(\mathbf{x}, \tau)$, we choose a trial wave function, $\Psi_T(\mathbf{x})$, which we expect will provide a good approximation to $\Phi(\mathbf{x}, \tau)$. As has been described by Reynolds[39], the evaluation of

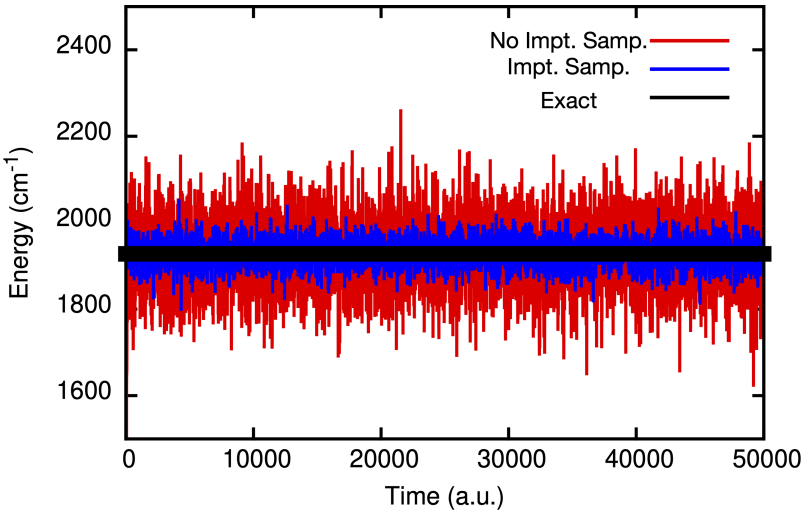

Figure 2.1: E_{ref} is plotted as a function of τ without importance sampling (red), and with importance sampling (blue), where Ψ_T is described as harmonic oscillator. The system of interest is a 1D model of an OH stretch using the potential shown in Figure 2.2. The exact ground state energy for the potential described is indicated by the black line.

$f(\mathbf{x}, \tau)$ is achieved by replacing $\Phi(\mathbf{x}, \tau)$ in Eq. 2.1 with

$$\Phi(\mathbf{x}, \tau) = \frac{f(\mathbf{x}, \tau)}{\Psi_T(\mathbf{x})} \tag{2.9}$$

Substituting Eq. 2.9 into Eq. 2.1 leads to[39]

$$\frac{\partial f}{\partial \tau} = \sum_{i=1}^{3N} \left[\frac{\hbar^2}{2m_i} \frac{\partial^2 f(\mathbf{x}, \tau)}{\partial x_i^2} - \frac{\partial \left(D_i(\mathbf{x}) f(\mathbf{x}, \tau) \right)}{\partial x_i} \right] - \left(E_{\text{L}}(\mathbf{x}) - E_{\text{ref}} \right) f(\mathbf{x}, \tau) \tag{2.10}$$

where

$$D_i(\mathbf{x}) = \frac{\hbar^2}{m_i} \cdot \frac{1}{\Psi_T(\mathbf{x})} \frac{\partial \Psi_T(\mathbf{x})}{\partial x_i} \tag{2.11}$$

is the drift term, and

$$E_{\mathrm{L}}(\mathbf{x}) = \frac{\hat{H}\Psi_T(\mathbf{x})}{\Psi_T(\mathbf{x})} \tag{2.12}$$

is the local energy. Additionally, the definition of E_{ref} is slightly modified

$$E_{\mathrm{ref}}(\tau) = \overline{E}_L(\tau) - \alpha \ln\left(\frac{W(\tau)}{W(\tau=0)}\right) \tag{2.13}$$

As in Eq. 2.8, $\alpha = (2\Delta\tau)^{-1}$ is a simulation parameter. In the limit where $W(\tau) - W(\tau = 0)$ is small, the value of E_{ref} obtained using Eq. 2.13 approaches the value obtained using Eq. 2.8.

Comparing Eq. 2.10 to Eq. 2.1, we find that the first and last terms take on similar roles. In the DMC simulation, these terms are accounted for using the approaches outlined in the previous section. The first term is the diffusion step, which is the same as the case without importance sampling. Next, we look at the drift term, which is the second term in Eq. 2.10. The drift term introduces an additional shift in the position of the ith coordinate of the jth walker by $D_i(\mathbf{x_j})\Delta\tau$.

By introducing the drift term into the DMC algorithm, the simulation is no longer guaranteed to obey microscopic reversibility. In the absence of the coordinate-dependent drift term, $\mathbf{D}(\mathbf{x})$, the probability of a walker being displaced by an amount Δx along one of the coordinates is independent of the initial and final values of the coordinate. The drift term introduces an additional displacement of the walker. In regions where the drift term is large or where it has a large coordinate dependence, the probability of moving from a set of coordinates $\mathbf{x}$ to a new set of coordinates $\mathbf{x}'$ may deviate significantly from the probability of making the move from $\mathbf{x}'$ to $\mathbf{x}$. To account for this, an additional Metropolis step is introduced.[39,40] It is based on the

ratio of the probabilities for these two moves

$$\begin{aligned} a\left(\mathbf{x},\mathbf{x}'\right) &= \frac{P(\mathbf{x}'\rightarrow\mathbf{x},\Delta\tau)}{P(\mathbf{x}\rightarrow\mathbf{x}',\Delta\tau)} \\ &= \frac{|\Psi_T(\mathbf{x}')|^2 G(\mathbf{x}'\rightarrow\mathbf{x},\Delta\tau)}{|\Psi_T(\mathbf{x})|^2 G(\mathbf{x}\rightarrow\mathbf{x}',\Delta\tau)} \end{aligned} \tag{2.14}$$

where $G(\mathbf{x}\rightarrow\mathbf{x}',\Delta\tau)$ is the Green's Function based on Eq. 2.10. As Reynolds has shown,[39]

$$G(\mathbf{x}\rightarrow\mathbf{x}',\Delta\tau) \approx e^{-([E_L(\mathbf{x})+E_L(\mathbf{x}')]/2-E_{\mathrm{ref}})\Delta\tau} \prod_{i=1}^{3N} \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\left(x_i'-x_i-D_i(\mathbf{x})\Delta\tau\right)^2/2\sigma_i^2} \tag{2.15}$$

Substituting Eq. 2.15 into Eq. 2.14 leads to

$$a(\mathbf{x},\mathbf{x}') = \frac{|\Psi_T(\mathbf{x}')|^2}{|\Psi_T(\mathbf{x})|^2} \prod_{i=1}^{3N} e^{-\left[D_i(\mathbf{x})\left(x_i'-x_i-D_i(\mathbf{x})\Delta\tau/2\right)+D_i(\mathbf{x}')\left(x_i'-x_i+D_i(\mathbf{x}')\Delta\tau/2\right)\right]\Delta\tau/\sigma_i^2} \tag{2.16}$$

When $a(\mathbf{x},\mathbf{x}') \geq 1$, the probability of moving from $\mathbf{x}'$ to $\mathbf{x}$ is at least as large as the probability of moving from $\mathbf{x}$ to $\mathbf{x}'$. In this case the move is accepted. If $a(\mathbf{x},\mathbf{x}') < 1$, its value is compared to a random number taken from a uniform distribution between 0 and 1. The move is accepted only if $a(\mathbf{x},\mathbf{x}')$ is larger than this random number.

Lastly, we consider the local energy, E_{L}, which can reduce the coordinate dependence on where the walkers sample the potential energy surface, based on the choice of Ψ_T. If $\Psi_T = \Phi$, then $E_L = E_0$ and there is no coordinate dependence. This is indicated by the black curve in Figure 2.2. However, when $\Psi_T \neq \Phi$, the coordinate dependence of E_L is not removed, but has been drastically decreased compared to the potential, which is shown in Figure 2.2.

It is useful to note that when $\Psi_T(\mathbf{x}) = 1$, the drift term vanishes, and the local energy becomes equivalent to the value of the potential energy at the coordinates of the walker. As such, in this limit, Eq. 2.10 reduces to Eq. 2.1.

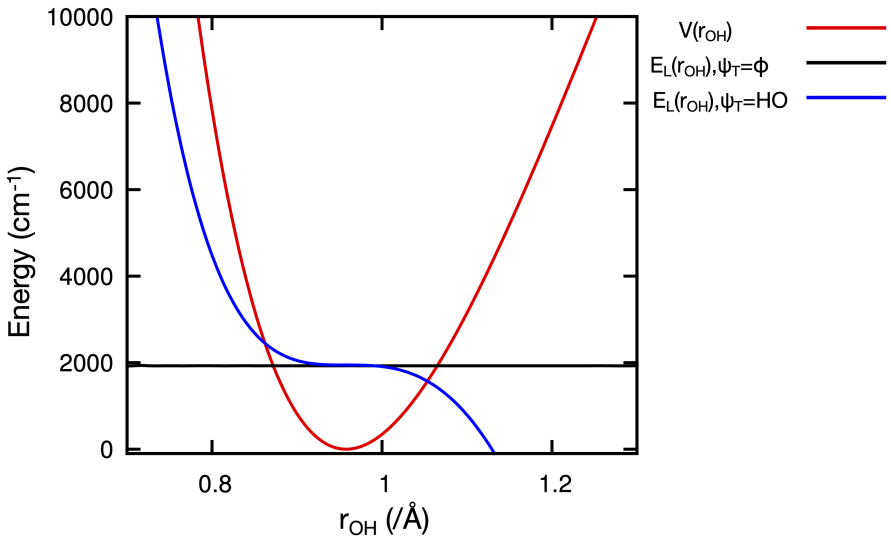

Figure 2.2: The potential energy of an OH stretch (red) based off of a cut along the potential energy surface of a water monomer developed by Partridge and Schwenke, and the local energy when $\Psi_T = \Phi$ (black), and Ψ_T is a harmonic oscillator based on the force constant of an OH stretch at the potential minimum (blue).[41]

2.3 Descendant Weighting

As described above, the ensemble of walkers that is propagated in a DMC simulation provides a Monte Carlo sampling of either $\Phi(\mathbf{x})$ or $f(\mathbf{x})$, while the average of E_{ref} over τ provides a measure of the ground state energy. As such, if the DMC simulation samples $f(\mathbf{x})$, we can evaluate integrals involving this function through Monte Carlo integration, as

$$\int f(\mathbf{x})g(\mathbf{x})d\mathbf{x} \approx \frac{\sum_j w_j(\tau_0)g(\mathbf{x}_j(\tau_0))}{W(\tau_0)} \tag{2.17}$$

where $\mathbf{x}_j(\tau_0)$ and $w_j(\tau_0)$ represent the coordinates and weight of the jth walker at a selected time in the simulation, τ_0. Finally, once the simulation has equilibrated the integral in Eq. 2.17 does not depend on the choice of τ_0. Consequently, we have dropped the τ-dependence of f and Φ in the discussion in this section.

Often other quantities are desired, for example expectation values of operators, which require the ability to evaluate $|\Phi(\mathbf{x})|^2$. When we use the Monte Carlo sampling based on $f(\mathbf{x})$, the summation on the right hand side of Eq. 2.17 provides a way to evaluate integrals involving $\Phi(\mathbf{x})\Psi_T(\mathbf{x})$. Obtaining expectation values of multiplicative operators, also requires a method for evaluating $\Phi(\mathbf{x})/\Psi_T(\mathbf{x})$ at the positions of the walkers. Barnett et al. and Suhm and Watts have shown[22,40] that at the coordinates of the jth walker at a time τ_0, $\mathbf{x}_j(\tau_0)$, $\Phi(\mathbf{x})/\Psi_T(\mathbf{x})$ is proportional to the number of descendants this walker has after a given number of time steps,

$$\frac{\Phi(\mathbf{x}_j(\tau_0))}{\Psi_T(\mathbf{x}_j(\tau_0))} = d_j(\tau_0) = \frac{w_j(\tau')}{w_j(\tau_0)} \tag{2.18}$$

Here $w_j(\tau_0)$ represents the weight associated with the jth walker at τ_0, and $w_j(\tau')$ represents the weight of this walker at a later time. When discrete weights are used $w_j(\tau_0) = 1$, while $w_j(\tau')$ represents the number of walkers at τ' that can be traced to the jth walker at τ_0. For the discussion that follows, $\tau_{\text{avg}} = \tau' - \tau_0$ provides the amount of time that the system is propagated to generate the descendant weights. Since at τ_0, $f(\mathbf{x})$ is represented by an ensemble of localized functions centered at $\mathbf{x}_j(\tau_0)$ with an associated weight $w_j(\tau_0)$,

$$\begin{aligned} \langle A \rangle &= \frac{\int A(\mathbf{x})|\Phi(\mathbf{x})|^2 d\mathbf{x}}{\int |\Phi(\mathbf{x})|^2 d\mathbf{x}} \\ &= \frac{\sum_j d_j(\tau_0) w_j(\tau_0) A(\mathbf{x}_j(\tau_0))}{\sum_j d_j(\tau_0) w_j(\tau_0)} \\ &= \frac{\sum_j w_j(\tau') A(\mathbf{x}_j(\tau_0))}{\sum_j w_j(\tau')} \end{aligned} \tag{2.19}$$

for an arbitrary multiplicative operator, A.

The descendant weights can also be used to obtain projections of $|\Phi^2(\mathbf{x})|$ onto any coordinate of interest, which is shown in Figure 2.3a). We have also calculated the

ground state probability amplitude using Discrete Variable Representation (DVR) and we use as our exact approach to compare with both flavors of DMC.[42] We find that the projections of $|\Phi^2(\mathbf{x})|$ for DMC with and without importance sampling show good agreement, as well as the comparisons with the exact approach. The differences in the comparison of the exact ground state probability amplitude and the two flavors of DMC are shown in Figure 2.3b) and both approaches show good agreement with the exact approach.

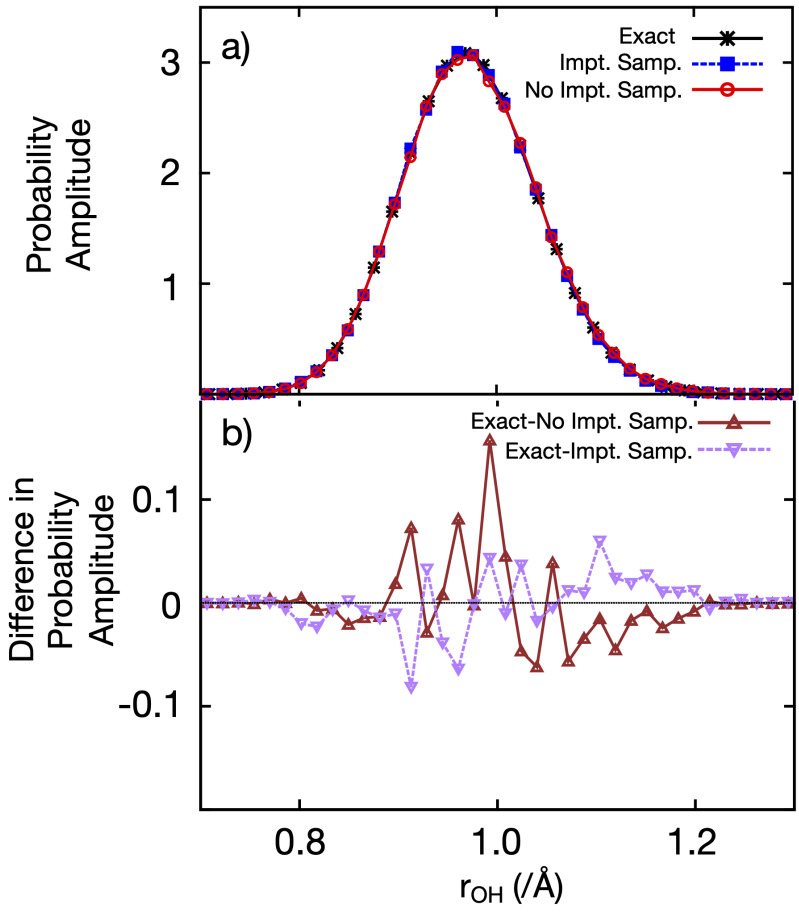

Figure 2.3: a) Comparisons of the 1D projections onto the OH stretch coordinate using DMC with importance sampling (blue dashed lines, filled square), and without importance sampling (red solid line, open circle). Both approaches are plotted on top of the exact representation of the ground state probability amplitude, which was calculated using Discrete Variable Representation.[42] b) The differences of the calculated ground state probability amplitude without importance sampling (brown solid lines, open triangle), and with importance sampling (dashed purple lines, open triangle).

Chapter 3

USING DIFFUSION MONTE CARLO TO EVALUATE OVERLAP INTEGRALS

3.1 Introduction

In this chapter, we will be exploring a methodological approach to calculate overlap integrals using the ground state wave function that was calculated using the As noted in chapter 2, DMC can be used to evaluate the ground state energy and wave function. Because the DMC wave function, Φ, is expressed in a basis of localized functions (called walkers), evaluating integrals that require Φ^2 is not entirely straightforward. Calculating integrals that involve products of two different wave functions is even more challenging.

Several approaches have been developed and explored for evaluating expectation values of multiplicative operators.[43–46] These fall into two general categories. The first is based on finite field methods from electronic structure theory along with ideas from perturbation theory. In this approach, a small perturbation, which is proportional to the operator for which the expectation value is desired, is added to the Hamiltonian being solved.[23,47] The finite field method, while accurate, can be expensive, as separate calculations must be performed for each operator of interest. Alternatively, since the ensemble of walkers provides a Monte Carlo sampling of the ground state wave

function, if we could evaluate the value of Φ at the locations of each of the walkers, we would have all the information that is needed to obtain expectation values. Barnett *et al.*[40] showed that this information can be obtained by propagating the wave function forward in time and equating the relative number of descendants of each walker to the value of Φ at the locations of these walkers. This approach is often referred to as descendant weighting.

Hornik *et al.*[48] and Barnett *et al.*[49] showed that one can recast the evaluation of an off-diagonal matrix element as an expectation value, which can be obtained using descendant weighting. That work formed the basis of several studies of properties of electronic wave functions.[50] In the present study, we explore applications of this approach to the evaluation of overlap integrals between approximate vibrational wave functions and those obtained by DMC.

The primary challenges in evaluating the overlap integrals using DMC come in normalizing the wave functions. This is particularly challenging when evaluating the overlaps between DMC and approximate wave functions, where an error that is no larger than 10^{-4} is desired. When we use the approaches that have been developed for evaluating expectation values to determine, $\langle\Phi|A|\Psi\rangle/\sqrt{\langle\Phi|\Phi\rangle\langle\Psi|\Psi\rangle}$, where the Monte Carlo-sampled wave function is Φ, the evaluation of the numerator and the first term in the denominator can be obtained in a relatively straightforward manner. To obtain $\langle\Psi|\Psi\rangle$, we equate it to $\langle\Phi|(\Psi/\Phi)^2|\Phi\rangle$, which can be evaluated using reciprocals of the normalized descendant weights. Because the values of the descendant weights are proportional to the corresponding value of Φ at the coordinates of the walker, when Φ is small the evaluation of the overlap integral becomes plagued by numerical instabilities. In this study, we explore how the accuracy of the evaluation of the calculated descendant weights affects our ability to obtain overlap integrals using DMC.

The remainder of this chapter is organized as follows. We describe the extensions that will be used to evaluate the overlap integrals with DMC. We explore the depen-

dence of the accuracy of this approach through a study of a Morse oscillator with varying degrees of anharmonicity and then for a study of formaldehyde and partially deuterated H_3^+ obtained using several approaches.[51,52]

3.2 Evaluating Overlap Integrals Using DMC

In Chapter 2, we have discussed that the descendant weighting approach is used to calculate expectation values of multiplicative operators. As we consider extending this approach to evaluating overlap integrals, the evaluation is complicated by the fact that we must obtain the normalization of a wave function, X, and the DMC wave function, Φ, which are both involved in the integral as

$$\langle X|\Phi\rangle = \frac{\int X(\mathbf{x})\Phi(\mathbf{x})d\mathbf{x}}{\sqrt{(\int |\Phi(\mathbf{x})|^2 d\mathbf{x})(\int |X(\mathbf{x})|^2 d\mathbf{x})}} \tag{3.1}$$

If X is an analytically determined wave function, which we require to be the same as Ψ_T, Eq. 3.1 becomes

$$\frac{\int \Psi_T(\mathbf{x})\Phi(\mathbf{x})d\mathbf{x}}{\sqrt{(\int |\Phi(\mathbf{x})|^2 d\mathbf{x})(\int |\Psi_T(\mathbf{x})|^2 d\mathbf{x})}} = \frac{\sum_j w_j(\tau_0)}{\sqrt{\left(\sum_j w_j(\tau')\right)\left(\sum_j \left(w_j(\tau_0)\right)^2 / w_j(\tau')\right)}} \tag{3.2}$$

To obtain the descendant weights for two states based on the same set of walkers, we first equilibrate the ensemble based on a ground state simulation. Starting from a distribution of walkers at τ_0, the ground state descendant weights are evaluated by continuing this propagation until τ', and evaluating the number of descendants each walker generates over $\tau_{\text{avg}} = \tau' - \tau_0$.

As we consider Eq. 3.2 we note that both contain a term that is proportional to the $(w_j)^{-1}$. When w_j is small, the calculation of the overlaps can suffer from numerical instabilities. To improve the accuracy of the calculations, rather than performing the

calculation of the weights a single time as is often done when expectation values are evaluated, weights are obtained several times and these values are averaged. In this way, w_j in Eq. 3.2 is replaced by

$$\overline{w}_j = \frac{\sum_m^{n_{\rm avg}} w_{m,j}}{n_{\rm avg}} \tag{3.3}$$

where $n_{\rm avg}$ reflects the number of evaluations of the descendant weights, $w_{m,j}$, that are obtained using a single snapshot of f. Each of the evaluations of $w_{m,j}$ uses a different series of random numbers in the DMC propagation, and this procedure will result in $n_{\rm avg}$ distinct sets of values of the w_j. In the discussion that follows, we will consider this averaging process and its influence on the accuracy of the final results.

3.2.1 Computational Details

The computational aspects of this study are divided into two parts. The first involves an exploration of the convergence properties of the evaluation of Eq. 3.2 for a one-dimensional Morse oscillator. We use a Morse oscillator potential for which the harmonic frequency is 913.65 $\rm cm^{-1}$, and the quadratic anharmonicity is 11.14 $\rm cm^{-1}$ when the reduced mass is 3.5 amu. For this part of the study, Ψ_T is the ground state of the harmonic oscillator that has the same mass and harmonic frequency used to define a Morse potential. In the DMC simulations, we perform five independent calculations where the wave functions are described by 20 000 walkers, which are propagated for 20 000 time steps with a time step of 10 a.u. using discrete weighting. After equilibrating for 12 000 time steps, the coordinates of the walkers are collected after every 2000 time steps. Descendant weights are evaluated using continuous weighting, and the reported results reflect the average of the results based on the analysis of these five wave functions.

For studies of the evaluation of overlap integrals for molecular systems, we focus on formaldehyde and partially deuterated H_3^+. For formaldehyde, we use the potential

that was developed by Yachmenev et al.[53], and a wave function that was obtained using the vibrational self-consistent field (VSCF) approach with optimized coordinates.[51] In the case of deuterated H_3^+, we use the potential developed by Aguando et. al.,[54] while the trial wave function was obtained using first order vibrational configuration interaction (VCI(1)) based on optimized coordinates.[52] For comparison, for all three systems we also perform simulations in which Ψ_T is the normal mode ground state, where the normal mode treatment was performed in either internal[1] or Cartesian coordinates. When the normal modes are based on Cartesian displacements an additional step is introduced in the DMC calculation in which the molecular coordinates are rotated into an Eckart frame based on the equilibrium geometry of the molecule.[55] This is done to ensure that these normal mode displacements represent vibrational motions of the system and not rotations or translations. In these calculations, we perform five independent simulations where 20 000 walkers are propagated using Eq. 2.10 for 200 000 time steps with a time step of 1 a.u., using discrete weighting. After propagating the wave function for 120 000 time steps, the coordinates are collected for every 20 000 time steps, and the $\overline{w}_j$ are evaluated using descendant weighting with $\tau_{\text{avg}} = 150$ a.u. This provides five wave functions that are used for the subsequent analysis.

3.3 Results and Discussion

3.3.1 Overlap Integrals for One-Dimensional Systems

To start, we focus on one-dimensional systems to better understand the convergence properties of the overlap integrals with respect to various parameters. These include the anharmonicity of the system, the amount of time the descendant weights are collected, τ_{avg}, and the number of independent evaluations of the descendant weights that are used in the evaluation of the average weights, n_{avg}. We focus on the one-dimensional Morse oscillator, described above. For these calculations, we use the

harmonic ground state wave function for the trial wave function, Ψ_T, and evaluate the normalized overlap integral between $\Psi_T(x)$ and the ground state wave function obtained from DMC, $\Phi(x)$. This overlap, which will be represented by

$$O = \langle \Psi_T | \Phi \rangle \tag{3.4}$$

is simply the overlap between the harmonic and anharmonic ground state solutions for this model system. For comparison, we also obtained the ground state wave function for the Morse oscillator described above variationally. Based on this evaluation $O(\text{Exact}) = 0.9963$.

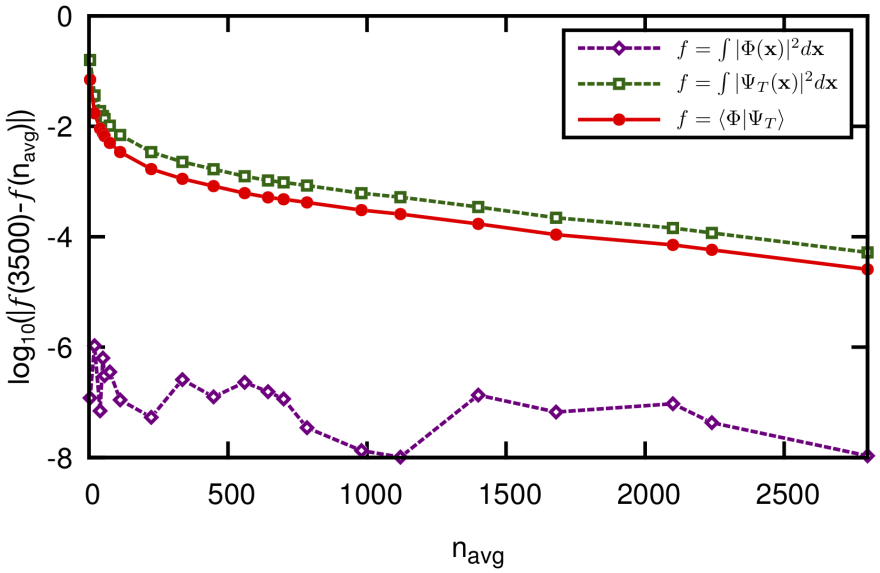

Figure 3.1: The convergence of the integrals that contribute to Eq. 3.1, f, for $f = \int \Phi^2(x)dx$ (purple diamonds/dashed line), $f = \int \Psi_T^2(x)dx$ (green squares/dashed line) and $f = O = \langle \Phi | \Psi_T \rangle$ (red filled circles/solid line) and plotted as functions of the number of independent evaluations of the weights (n_{avg}), which is used to determine the $\overline{w}_j$. The results are subtracted from $f(3500)$, where $\int \Phi^2(x)dx = 0.9999$, $\int \Psi_T^2(x)dx = 1.0073$ and $\langle \Phi | \Psi_T \rangle = 0.9964$. These calculations are based on the one-dimensional Morse oscillator potential described in the text, Ψ_T is the correspond harmonic ground state wave function, and weights are collected for τ_{avg} =8000 a.u.

As seen in Eqs. 3.1 and 3.2, calculating O requires the evaluation of three integrals. The numerator, $\int \Psi_T(x)\Phi(x)dx$, is trivially equated to the number of walkers that is used to represent $\Phi(x)$ at τ_0 since discrete weighting is used for this part of the propagation. In Figure 3.1, we explore the convergence properties of the terms in the denominator of O, scaled by the number of walkers at τ_0, with green squares/dashed line and purple diamonds/dashed line, as well as O itself, which is shown with the filled red squares/solid line. These three quantities are shown as functions of n_{avg}. To illustrate the convergence behavior, each of these integrals is subtracted from its value when $n_{\text{avg}} = 3500$, and the results are plotted on a $\log_{10}$ scale. Based on the results shown in Figure 3.1, the value of O shows fairly slow convergence behavior. Breaking O down into its three contributions, the slow convergence can be traced to the convergence properties of the evaluation of $\int \Psi_T^2(x)dx$. The relative insensitivity to the calculated value of $\int \Phi^2(x)dx$ to the value of n_{avg} (for this integral $|f(3500) - f(n_{\text{avg}})| < 10^{-6}$ independent of the value of n_{avg}) is consistent with earlier observations that a single evaluation of the descendant weights, w_j, is generally sufficient for calculations of expectation values. The reason for the higher sensitivity that is found for the evaluation of $\int \Psi_T^2(x)dx$ comes from its dependence on $(\overline{w}_j)^{-1}$ rather than $\overline{w}_j$. In geometries where $\overline{w}_j$ is small, a small absolute error can manifest itself as a large error in $(\overline{w}_j)^{-1}$.

The sensitivity of the accuracy of the calculated overlap integrals to the size of n_{avg} is further explored in Figure 3.2. In panel (a), the value of O is subtracted from the value obtained variationally. In these plots, the red circles/solid line provide the same results as were plotted in Figure 3.1, subtracted from O(Exact) rather than $O(n_{\text{avg}} = 3500)$. The blue diamonds/dashed line show the results obtained when we do not employ importance sampling for the evaluation of O. As is seen by comparing the magnitude of the errors represented by these two curves, the quality of the results improve greatly when importance sampling is used.

On the right side of Figure 3.2, the values of the average weights $\overline{w}_j(\tau_0)$ for 1000

individual walkers are plotted as functions of the corresponding $x_j(\tau_0)$ values for the three simulation conditions identified by arrows in Figure 3.2(a). The purple curves in panels (ii) and (iii) provide the expected values of $\Phi(x)/\Psi_T(x)$. When importance sampling is not used $\Psi_T(x)$ is a constant, and the purple curve in panel (i) provides the ground state wave function for the Morse potential. When importance sampling is used, the quality of the results improves with n_{avg}, and the average weights lie close to the purple curve when $n_{\text{avg}} = 1120$. On the other hand, using the same value of n_{avg} without importance sampling leads to a much greater spread in the values of the $\overline{w}_j$'s when compared to the value of the ground state wave function at the same location. It is for this reason that we will focus on the results obtained using importance sampling in the remainder of the discussion.

Before continuing, it should be noted that while the individual evaluations of the average weights do not generally lie along the purple curve, as is seen in Figure 3.3.1, their average values reproduce $\Psi^2(x)$ for this Morse oscillator. This is why the evaluation of $\int \Phi^2 dx$ in Figure 3.1 and the use of descendant weighting to obtain expectation values of multiplicative operators is not as sensitive to the accuracy of the individual weights as the overlap integrals are found to be.

Finally, the red curve in Figure 3.2(a) has a minimum when $n_{\text{avg}} \approx 1600$. This reflects the fact that when the weights are evaluated by propagating for $\tau_{\text{avg}} = 8000$ a.u., the sign of $O(n_{\text{avg}}) - O(\text{Exact})$ changes at this value of n_{avg}. This is seen in Figure 3.2(b). Looking at the dependence of O on the amount of time the DMC wave function is propagated to obtain the descendant weights, we find it to decrease, although the rate of this decrease becomes smaller for large values of n_{avg}. This reflects a buildup of the weight onto a small number of walkers when the descendant weighting is propagated to long times. This problem is even more dramatic when importance sampling is not employed. A propagation time of 8000 a.u. is chosen for the remainder of this part of the study as it is in the range of values of τ_{avg} over which the calculated value of O in Eq. 3.4 is numerically stable.[19,56]

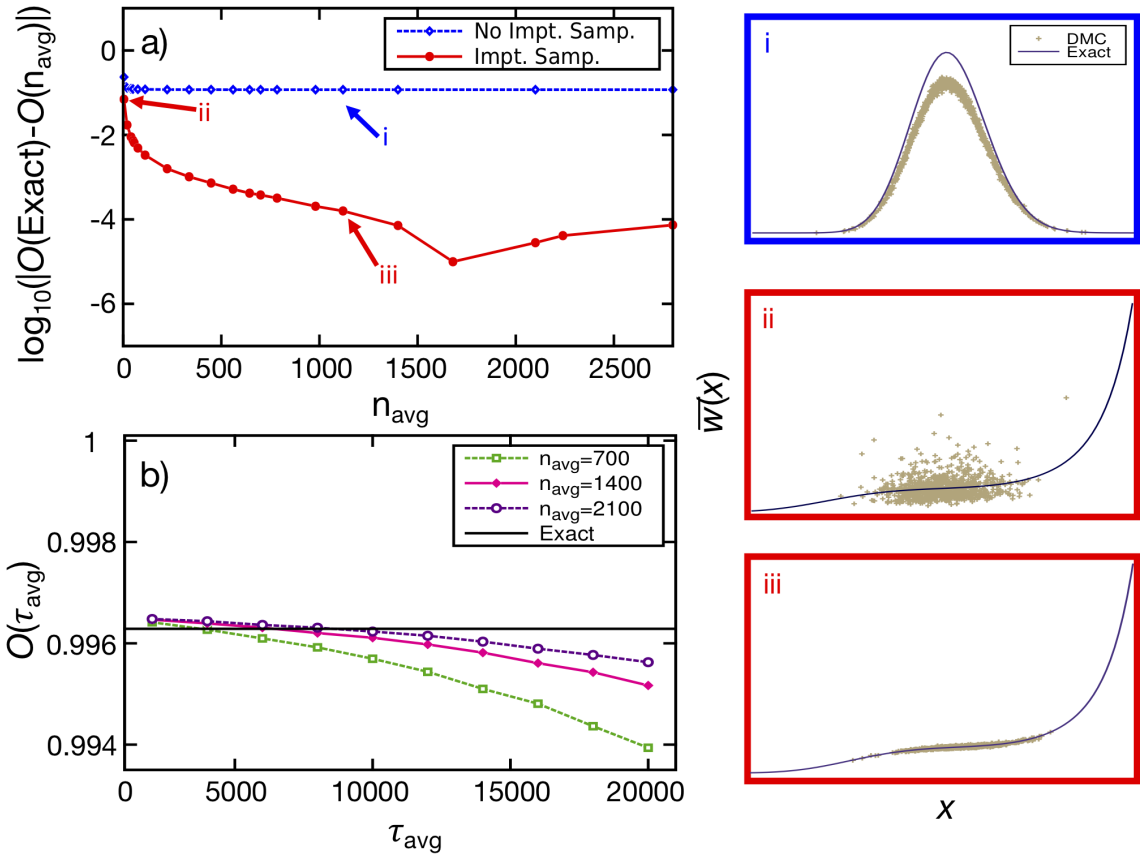

Figure 3.2: (a) The convergence of $O = \langle \Phi | \Psi_T \rangle$, evaluated using $\overline{w}_j$ values that are obtained from DMC simulations, which are propagated with importance sampling and $\tau_{\rm avg} = 8000$ a.u. (red filled circles/solid line), and without (blue diamonds/dashed line) importance sampling and $\tau_{\rm avg} = 600$ a.u. based on the Morse oscillator model described in the text. Panels (i)-(iii) provide the distribution of the average weights shown for 1000 walkers. These are obtained (i) without importance sampling and $n_{\rm avg} = 1120$ and with importance sampling and (ii) $n_{\rm avg} = 4$ and (iii) $n_{\rm avg} = 1120$, as indicated by corresponding colored arrows in panel (a). The solid purple curves in these plots show the expected value of (i) $\Phi(x)$ and (ii) and (iii) $\Phi(x)/\Psi_T(x)$. (b) The dependence of the error in O on $\tau_{\rm avg}$ is plotted as a function of $\tau_{\rm avg}$ for specified values of $n_{\rm avg}$.

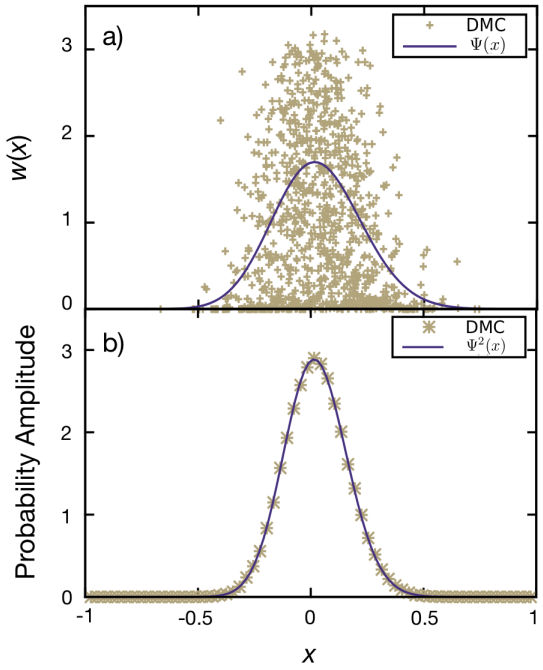

Figure 3.3: (a) The descendant weights (gold +), which have been calculated without importance sampling for 1000 walkers propagated based on the Morse oscillator described in the text, are plotted as functions of the coordinate of the walker, and compared to the value of the ground state wave function (solid purple line). (b) The distribution of walkers, weighted by their descendent weights, (gold *) is compared to the probability amplitude for the ground state of the Morse oscillator (purple line).

We have also explored the time step dependence on the accuracy of the evaluation of O. The results are plotted in Figure 3.4. Overall, convergence behavior of the overlaps with n_{avg} shown in Figure 3.1 does not depend on the size of the time step. While in the limit of large n_{avg} there is some dependence of the asymptotic value of O on the value of the time step, for time steps ranging from one to ten a.u., the error in the asymptotic value is on the order of 10^{-4} to 10^{-5}.

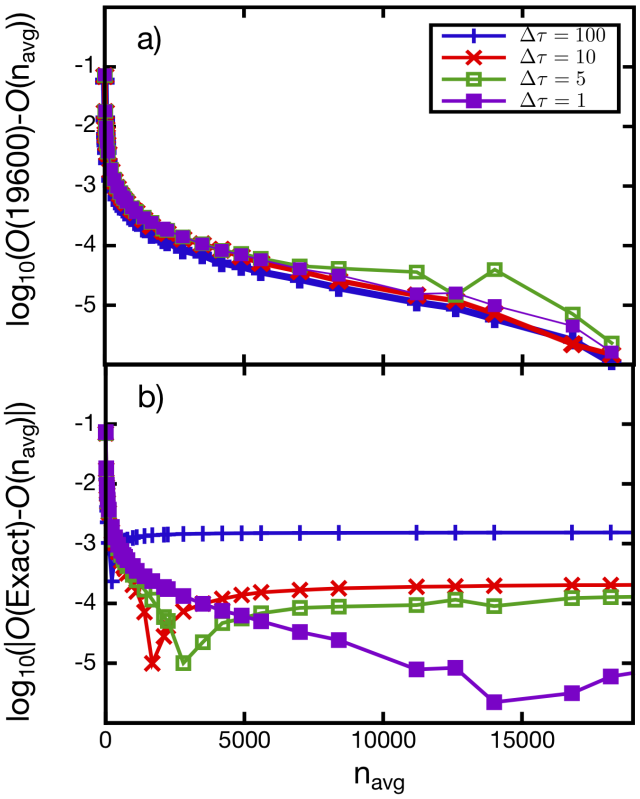

Figure 3.4: Evaluation of $O = \langle \Phi | \Psi_T \rangle$ for the Morse oscillator system described in the text as a function of the number of times w_j is evaluated to obtain $\overline{w}_j$, n_{avg}, and the size of the time step used when the system is propagated to obtain the w_j values, $\Delta\tau$. In all cases, the systems are propagated for τ_{avg} =8000 a.u.. In (a) the results are subtracted from the value obtained for $n_{\text{avg}} = 19\ 600$, where $O(19\ 600) = 0.9978$ for $\Delta\tau = 100$ a.u., 0.9965 for $\Delta\tau = 10$ a.u., 0.9964 for $\Delta\tau = 5$ a.u. and 0.9963 for $\Delta\tau = 1$ a.u. In (b) the values are subtracted from the independently calculated overlap of the Morse and harmonic wave functions, 0.9963.

The final quantity that is explored is the effect of the quality of the trial wave function as an approximation to the ground state wave function on the accuracy of the evaluation of both the overlap and the ground state energy. In particular, errors can be introduced if the trial wave function does not sample the range of configurations that are sampled by the true ground state. The results of this analysis are provided in Table 3.1. For this part of the study, we use the Morse oscillator described above. Adjusting the reduced mass has the effect of tuning the anharmonicity of the system, with smaller values of μ corresponding to the ground state wave function sampling more anharmonic regions of the potential. As is seen in the results reported in Table 3.1, when the mass is decreased, the overlap decreases, and the error in the calculated ground state energy increases. For overlaps exceeding roughly 0.995, the energies obtained using DMC with and without importance sampling differ by less than 1 cm^{-1}. When the overlap becomes smaller, the error in both the overlap and the energy obtained when importance sampling is employed increase, with the calculated overlap consistently providing an underestimate of the expected value. The reason for these trends can be found in panels (ii) and (iii) on the right side of Figure 3.2. When importance sampling is used, the walkers are localized in the region of configuration space where the trial wave function has amplitude. Regions where Ψ_T is small will be poorly sampled.[57–59] It is this loss of walker density in the tails of the trial wave function that leads to the underestimation of O and the errors in the energy obtained from a DMC simulation with importance sampling. As a result, the approach described in the previous section will provide a more accurate measure of the overlaps as they approach unity, while the calculated overlaps will become more qualitative when these values are smaller. The quality of the calculated value of O can be correlated to the agreement between the ground state energy obtained with and without importance sampling. In fact for the vibrational problems with nodeless ground states it is encouraging how well the simple unguided DMC calculation is able to reproduce the exact ground state energy.

Table 3.1: Ground State Properties for Morse Oscillators Using DMC With and Without Importance Sampling.[a]

μ[b] /amu	$\langle\Phi_{\rm Morse}\|\Psi_{\rm HO}\rangle$ Exact	$\langle\Phi_{\rm Morse}\|\Psi_{\rm HO}\rangle$[c] DMC,IS[d]	E_0/cm^{-1} DMC,IS	E_0/cm^{-1} DMC[e]	E_0/cm^{-1} Exact
28	0.9987	0.9987 (0.2)	161.17 (0.04)	161.17 (0.13)	161.16
14	0.9981	0.9982 (0.3)	227.77 (0.03)	227.66 (0.18)	227.72
7	0.9974	0.9974 (0.5)	321.70 (0.06)	321.87 (0.16)	321.63
3.5	0.9963	0.9963 (0.5)	454.37 (0.04)	454.13 (0.26)	454.04
1.725	0.9947	0.9947 (0.9)	641.28 (0.1)	640.31 (0.23)	640.48
0.8625	0.9925	0.9910 (1.6)	904.63 (0.04)	902.69 (0.32)	902.51
0.43125	0.9893	0.9802 (4.1)	1275.28 (0.11)	1269.81 (0.27)	1269.82

[a] Numbers in parentheses provide one standard deviation based on five calculations using statistically independent wave functions.
[b] The reduced mass.
[c] The standard deviations have been multiplied by 10^4.
[d] Evaluated using importance sampling with $n_{\rm avg} = 2100$ and $\tau_{\rm avg} = 8000$ a.u.
[e] Evaluated using DMC without importance sampling.

3.3.2 Overlap Integrals for Formaldehyde and H_3^+

While it is interesting to explore the accuracy of DMC for evaluating overlap integrals for model systems, the focus of this work is to use this approach as a tool for assessing the quality of approximate wave functions for polyatomic systems. Here we focus on H_2CO, a modestly anharmonic molecule, as well as H_2D^+ and D_2H^+, which sample much more anharmonic regions of the potential in their ground states than H_2CO. For all three molecules, we calculate overlaps between the DMC ground states and wave functions that were evaluated using recently developed VSCF and VSCF/CI approaches.[51,52]

We start by considering H_2CO, which is the least anharmonic of the three molecules being considered. A DMC calculation using the VSCF wave function for Ψ_T provides an energy that is nearly identical to those obtained using DMC without importance sampling and the reported ground state energy based on the potential surface used for the calculation.[51,53] In this case $\langle\Psi_T|\Phi\rangle$= 0.9996(2). Based on the analysis provided

above, this result is expected to be quantitatively accurate. Also, as is seen in the results reported in Table 3.2, $\langle\Psi_T|H|\Psi_T\rangle$ deviates from the literature value by 4 cm^{-1}, although when it is used as a guiding function for DMC, the ground state energy obtained by the two approaches differ by less than 1 cm^{-1}, which is the statistical uncertainty of these results.

While the ground state wave function for H_2CO is expected to be reasonably well-described by harmonic treatments, H_2D^+ and D_2H^+ are considerably more anharmonic. This is reflected in the larger deviation between the values of $\langle\Psi_T|H|\Psi_T\rangle$ obtained from harmonic and the VSCF or VCI(1) wave functions provided in Table 3.2. Despite the increased anharmonicity, the ground state energies for H_2D^+ and D_2H^+, calculated using all three trial wave functions, differ by less than 1.5 cm^{-1}. Additionally, energies evaluated using the normal mode wave functions based on linear combinations of internal displacement coordinates differ from the exact results for this potential by less than 0.1 cm^{-1}. Examining the calculated overlaps, we find that the overlaps between the VCI(1) wave trial wave function and the exact wave function obtained using DMC exceed 0.995. The overlaps are somewhat smaller when the harmonic ground state wave functions are used. When we compare the overlaps obtained using harmonic wave functions based on normal modes, which have been constructed as linear combinations of displacements of Cartesian and internal coordinates, we find that the overlap is larger when the harmonic trial wave function is based on Cartesian displacements. This was at first surprising as it is expected that normal modes based on internal displacement coordinates should provide a better zero-order description of bending vibrations in such anharmonic systems.[61,62] To better understand these trends, we also evaluated $\langle H\rangle$ for these trial wave functions numerically, and the results are reported in Table 3.2. Consistent with the above results, we find that $\langle H\rangle$ is smaller when the normal modes are expressed as linear combinations of Cartesian displacements. Based on this, we conclude that the vibrations in H_2D^+ (and D_2H^+) are unusual. In this ion, the natural vibrations are somewhere between those of a

triatomic molecule, where the vibrations are described as two stretching motions and a bend, and a complex of H_2 with D^+ where the bend is better described as a H-H stretch, the symmetric stretch becomes the D-H_2 stretch and the asymmetric stretch is now the hindered rotation of H_2 in the presence of D^+.

Table 3.2: Comparison of Overlaps and Ground State Energies For H_2CO, H_2D^+ and D_2H^+.[a]

Molecule	Ψ_T	$\langle\Psi_T\vert\hat{H}\vert\Psi_T\rangle$[b] /cm^{-1}	$\langle\Phi\vert\Psi_T\rangle$[c] DMC,IS[d]	E_0/cm^{-1} DMC,IS	E_0/cm^{-1} DMC[e]
CH_2O	VSCF[f]	5774.1	0.9996 (0.1)	5769.8 (0.1)	5770.0 (1.1)
	N.M. Cart.[g]	5854.9	0.9946 (1.5)	5770.9 (0.2)	
H_2D^+	VCI(1)[h]	4028.2	0.9967 (1.3)	3978.9 (0.1)	3978.4 (0.7)
	N.M. Cart.[g]	4217.6	0.9818 (1.3)	3978.4 (0.1)	
	N.M. Int.[i]	4259.5	0.9909 (3.2)	3979.9 (0.3)	
D_2H^+	VCI(1)[h]	3623.2	0.9957 (1.8)	3562.7 (0.2)	3561.7 (1.5)
	N.M. Cart.[g]	3751.3	0.9932 (2.9)	3562.0 (0.2)	
	N.M. Int.[i]	3816.8	0.9873 (5.1)	3563.2 (0.2)	

[a] Numbers in parentheses provide one standard deviation based on five calculations using statistically independent wave functions.
[b] Average energy based on the Ψ_T. The literature values for these zero-point energies are 5769.78 cm^{-1} (H_2CO)[53], 3979.905 cm^{-1} (H_2D^+) and 3563.086 cm^{-1} (D_2H^+).[60]
[c] The standard deviation has been multiplied by 10^4.
[d] Evaluated with DMC using importance sampling with n_{avg}=560 and τ_{avg}=150 a.u.
[e] Evaluated with DMC without importance sampling.
[f] Ref. 51
[g] Normal modes defined as a linear combination of Cartesian coordinates.
[h] Ref. 52
[i] Normal modes defined as a linear combination of internal coordinates.

As noted above, the introduction of a guiding function can introduce errors if the exact wave function has amplitude in regions of configuration space where the guiding function is vanishingly small.[56] This is why the forms of the guiding functions often focus on the long-range behavior (e.g. the use of Jastrow factors),[63–65] or good guesses to the exact wave function based on prior variational Monte Carlo calculations.[16,57,66]

This deterioration in the accuracy of the DMC results as $|O|$ in Eq. 3.4 decreases will also lead to challenges as we use DMC to assess the accuracy of the approximate ground state wave functions that are also being used as the guiding functions. As noted above, we can use the errors in the ground state energy as an additional measure of the robustness of the calculation of the overlaps. On the basis of the good agreement between the energies for these molecular systems, obtained with and without importance sampling, we expect that the reported overlaps are accurate within the uncertainties of the DMC results.

3.4 Conclusions

In this study, we have explored approaches for obtaining matrix elements involving two different wave functions using diffusion Monte Carlo. This enables us to evaluate the accuracy of wave functions obtained using other approximate approaches. We find that when the approximate wave function is used as a trial wave function, the combination of the accuracy of the calculated energy and the value of the overlap integral, $\langle\Phi|\Psi_T\rangle$, allow us to assess the accuracy of the wave function. In cases where the wave function is of high quality, such evaluations of the overlap integral may be made with accuracy of 10^{-4}. For less accurate wave functions, the accuracy of both the zero-point energy and the wave function deteriorate. High accuracy is achieved when the approach is tested for one-dimensional systems, and the approach is also demonstrated for a highly anharmonic molecular system.

Chapter 4

AN EFFICIENT METHOD FOR STUDIES OF WATER CLUSTERS USING DIFFUSION MONTE CARLO

4.1 Introduction

DMC has been shown to be a powerful approach for studying a broad range of molecules and clusters.[17,22,22,24–27,37,67–75] DMC works well for molecules and ions, including CH_5^+, $H_5O_2^+$ and H_5^+, in which the high and low frequency vibrations are strongly coupled. In hydrogen-bonded clusters where the couplings between inter- and intramolecular vibrations are weaker, the approach becomes much more demanding computationally.[24]

For the most part, early studies of hydrogen-bonded complexes using DMC focused on a rigid-body algorithm that was developed by Buch.[67] In this approach, the interacting molecules are treated as rigid bodies, and DMC is used to explore the rotational and translational motions of these interacting molecules. The introduction of this approximation was motivated by both the availability of rigid-monomer potential functions and results of spectroscopic studies of these complexes, which probed intermolecular vibrations through excitation in the microwave and terahertz regions of the electromagnetic spectrum. More recently, full dimensional water potentials have been developed, notably the MB-pol family of potentials developed by Paesani and coworkers.[76–79] These potentials have been used by Mallory and Mandelshtam in studies

of the ground state properties of water clusters with up to six water molecules.[25,75] In order to obtain converged results for the water hexamer, ensemble sizes as large as one million walkers were needed.[75] Additionally, the convergence was shown to be relatively slow with respect to the number of walkers, and the convergence became slower as the number of water molecules in the system was increased.[24] This slow convergence appears to place severe limitations on the size of the systems that can be studied using DMC.

A variant of DMC that is commonly used involves the introduction of a guiding function. Such an approach has been used in studies of H_5^+[26] and $H_5O_2^+$[27], where comparisons to the results of calculations that do not employ guiding functions can be made. In these cases, the use of the guiding function does not appear to provide a significant computational advantage. This is because the ground state wave function is highly delocalized, and finding a suitable guiding function, which localizes the wave function in relevant regions of the potential without over-constraining the system, is challenging. On the other hand, we have found that using guiding functions based on a harmonic description of the ground state wave function can lead to significant savings for molecules, like water, where the ground state wave function is localized in a single minimum on the potential surface.[80]

To address the above challenges, in the present work we explore a variant of DMC in which we use importance sampling to constrain the wave function in the intramolecular degrees of freedom while not imposing any constraints on the intermolecular degrees of freedom. Specifically, we employ a guiding function,

$$\Psi_T(\mathbf{q}) = \prod_{j=1}^{N_w} \prod_{k=1}^{3} \phi_k(q_k^{(j)}) \tag{4.1}$$

where $q_k^{(j)}$ represents one of the internal coordinates, r_1, r_2 or θ that defines the geometry of the jth water molecule, and N_w provides the number of water molecules in the system of interest. A similar approach was proposed by Suhm and Watts in their

studies of water clusters, although it appears that these ideas were not pursued beyond their initial description.[22] The focus of this chapter is to investigate the convergence behavior of this approach and apply it to a study of the various isomers of partially deuterated water trimer. Specifically, we explore the question of how the isomer populations are affected by partial deuteration.

4.2 Methods

In this study, we use DMC with importance sampling which is described in greater detail in Chapter 2. We will be referring this flavor of DMC as the guided DMC approach. In this approach, DMC is used to sample the product of the ground state wave function, Φ, and a trial wave function, Ψ_T. The details of this approach and our implementation are provided elsewhere.[17,80] The important point is that at the end of the simulation we have a measure of the zero-point energy of the system and an distribution of localized functions, walkers, which provide a Monte Carlo sampling of

$$f = \Phi\Psi_T \tag{4.2}$$

where Ψ_T is given by Eq. 4.1.

In the discussion that follows, we focus on two quantities: the ground state energy and the fraction of the ground state probability amplitude that samples each of the minima in the potential surface, e.g. the isomer fractions.

The ground state energy is evaluated as the time-average of E_{ref}, which is evaluated using Eq.2.13. In these calculations, we use the discrete weighting approach to obtain E_{ref} at every timestep, as well as the walkers, which represent a Monte-Carlo sampling of f. We use the continuous weighting scheme for the descendant weights for all of the walkers, for which both are described in Chapter 2. The descendant weights can be used to obtain projections of the probability amplitude of a coordinate of interest.

As noted in Chapter 2, one main difference of the guided DMC approach is that

rather than evaluating the potential energy of all the walkers, we evaluate the local energy, which is shown in Eq. 2.12.

In these calculations, we evaluate $\hat{H}\Psi_T$ as

$$\begin{aligned}\hat{H}\Psi_T &= \hat{T}\Psi_T + \hat{V}\Psi_T \\ &= \sum_{i=1}^{3N} \frac{-\hbar^2}{2m_i} \left[\sum_{j=1}^{N_w} \sum_{k=1}^{3} \left[\left(\frac{d^2\phi_k}{dq_k^{(j)2}} \left(\frac{dq_k^{(j)}}{dx_i} \right)^2 \prod_{l\neq j}^{N_w} \prod_{m\neq k}^{3} \phi_m\left(q_m^{(l)}\right) \right) + \right.\right. \\ &\qquad \left(\frac{d\phi_k}{dq_k^{(j)}} \frac{d^2 q_k^{(j)}}{dx_i} \prod_{l\neq j}^{N_w} \prod_{m\neq k}^{3} \phi_m\left(q_m^{(l)}\right) \right) + \\ &\left.\left. \left(\sum_{n\neq j}^{N_w} \sum_{p\neq k}^{3} \frac{d\phi_p}{dq_p^{(n)}} \frac{dq_p^{(n)}}{dx_i} \frac{d\phi_k}{dq_k^{(j)}} \frac{dq_k^{(j)}}{dx_i} \prod_{j\neq l\neq n}^{N_w} \prod_{p\neq k\neq m}^{3} \phi_m\left(q_m^{(l)}\right) \right) \right] \right] + V(\mathbf{x})\Psi_T \end{aligned} \quad (4.3)$$

Properties of the ground state wave function are obtained using descendent weighting which is described in Chapter 2,[22,40] which provides the weights that are required to perform a Monte Carlo integration over Φ^2. For example, the isomer fraction is evaluated as

$$F^{(\beta)} = \frac{\sum_i' w_i^{(\beta)}(\tau')}{W(\tau')} \quad (4.4)$$

where $w_i^{(\beta)}(\tau')$ represent the descendant weight of the ith walker's structural isomers of the system of interest (β), and $W(\tau')$ is the sum of the weights for all of the walkers.

4.3 Computational Details

In these calculations, the initial ensemble of walkers is obtained by randomly sampling geometries from a distribution that is based on the harmonic ground state wave function of the cluster of interest. Specifically, the ground state wave function for the cluster of interest is approximated by a product of harmonic ground state wave functions in each of the $3N-6$ normal modes as

$$\Psi_{\mathrm{HO}}(\mathbf{q}) = \prod_{n=1}^{3N-6} \exp\left[-\omega_n q_n^2/2\right] \tag{4.5}$$

The Cartesian coordinates of the walkers at $\tau = 0$ are chosen by randomly sampling from a $3N-6$-dimensional Gaussian distribution that has a width along each normal coordinate that is is twice the width of Ψ_{HO} in Eq. 4.5. The use of this broadened distribution ensures that relevant geometries of the cluster are sampled. In addition, the water molecules are are permuted so as to ensure that there are roughly equal numbers of walkers in configurations that correspond to each of the low-energy minima on the potential surface.

For the calculations in which we use a guiding function described in Eq 4.1, the one-dimensional wave functions in the two OH stretches and the HOH bend are evaluated based on one-dimensional cuts through the potential of Partridge and Schwenke.[41] For the OH stretches, we solve for the ground state wave function using a DVR approach that is based on 500 evenly spaced grid points ranging from 0.265 to 2.117 Å.[42] The bend is described as the ground state of a one-dimensional harmonic oscillator, with the values of the Wilson G-matrix elements[1] and harmonic frequencies are provided in Table 4.1. In the cases of unbiased sampling, $\Psi_T(\mathbf{q}) = 1$.

Table 4.1: Frequencies and Wilson G-matrix Elements Used to Define the Bend Contribution to Ψ_T in Eq. 4.1.

Angle	ω/cm^{-1}	G^a/amu^{-1} Å^{-2}
HOH	1655	2.338
HOD	1448	1.796
DOD	1207	1.255

[a] Wilson G-matrix element.[1].

In the DMC calculations of water clusters, we employ the the MB-pol potential energy surface developed by Paesani and coworkers.[76–79] The simulations are run with a variety of sizes of ensembles of water clusters, with $\Delta\tau = 1$ a.u.. The ground

state energies are obtained by equilibrating the simulations for 20 000 time steps, and averaging E_{ref} in Eq 2.13 over the subsequent 30 000 time steps. The reported energies and their standard deviations are based on the results of 5 independent calculations. Starting at 27 500 time steps, the coordinates of the ensemble of walkers are captured every 2500 time steps. This provides ten wave functions for each of the five simulations. For each of these wave functions, the walkers are propagated for an additional 150 time steps to obtain the descendant weights, which are used in evaluating the isomer fractions in Eq. 4.4.

4.4 Results and Discussion

4.4.1 Convergence properties

We start by considering the convergence behavior of the DMC approach described above. In Figure 4.1 we explore how the ground state energy of $(H_2O)_n$ with $n = 1$ to 3 is affected by the number of walkers that are used for the DMC calculation. The unfilled red squares, solid line provide the results from an unguided DMC simulation, while the filled blue circles, dashed line provide the results when the guiding function in Eq. 4.1 is used. The results are compared to the results of a large unguided DMC simulation ($\Psi_T = 1$), in which 60 000 walkers have been used.

For water monomer the ground state energies obtained using the guided DMC approach converges much more rapidly compared to energies obtained from unbiased DMC calculations using the same number of walkers. This can be understood from the plot of the one-dimensional potential and E_L for an OH stretch in water, shown in Figure 4.3(a). When the guiding function is introduced, the local energy for the OH stretch is constant. In contrast, the cut through the potential along an OH bond length in water has a strong coordinate dependence, and a larger number of walkers is required to sample the potential than are needed to sample the local energy. When the other OH stretch and the HOH bend are introduced, the value of E_L will depend

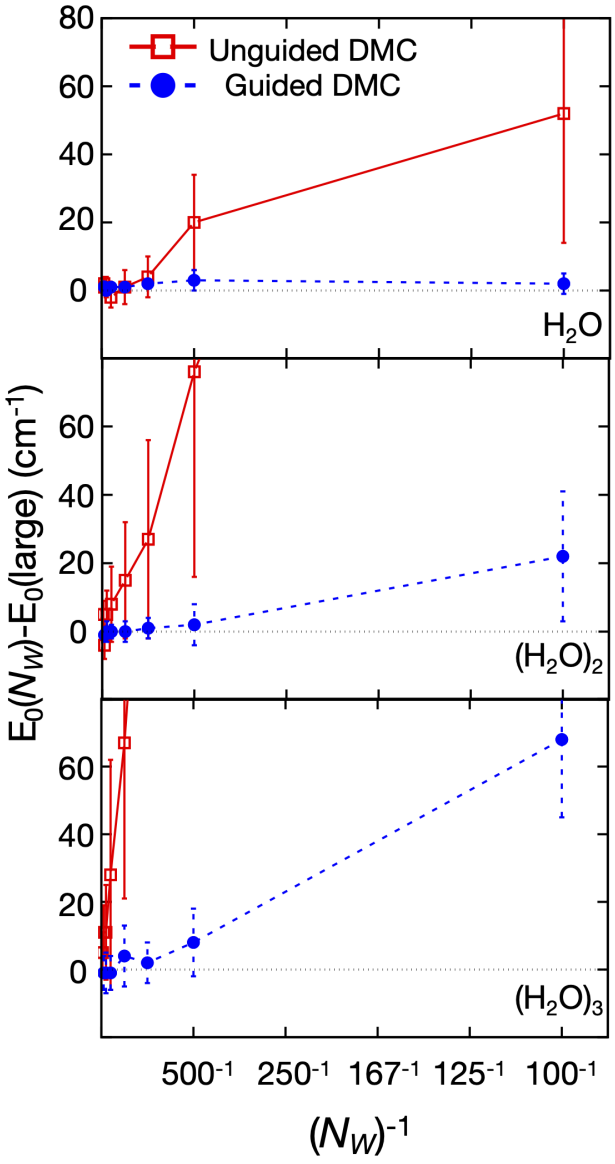

Figure 4.1: The convergence properties of DMC calculations of the ground state energies of $(H_2O)_n$ for $n = 1, 2, 3$ as a function of N_W^{-1}. The unfilled red squares provide the results of unguided DMC calculations in which $\Psi_T = 1$ while the filled blue circles give the results when Ψ_T in Eq. 4.1 is used. The error bars provide the standard deviation based on 5 independent calculations of the ground state energy obtained when N_W walkers are used in the simulation. The ground state energies are compared to the value obtained when 60 000 walkers are used in an unguided DMC calculation, E_0(large). These results are also provided in Table 4.2-4.4 and are plotted on an expanded scale in Figure 4.2.

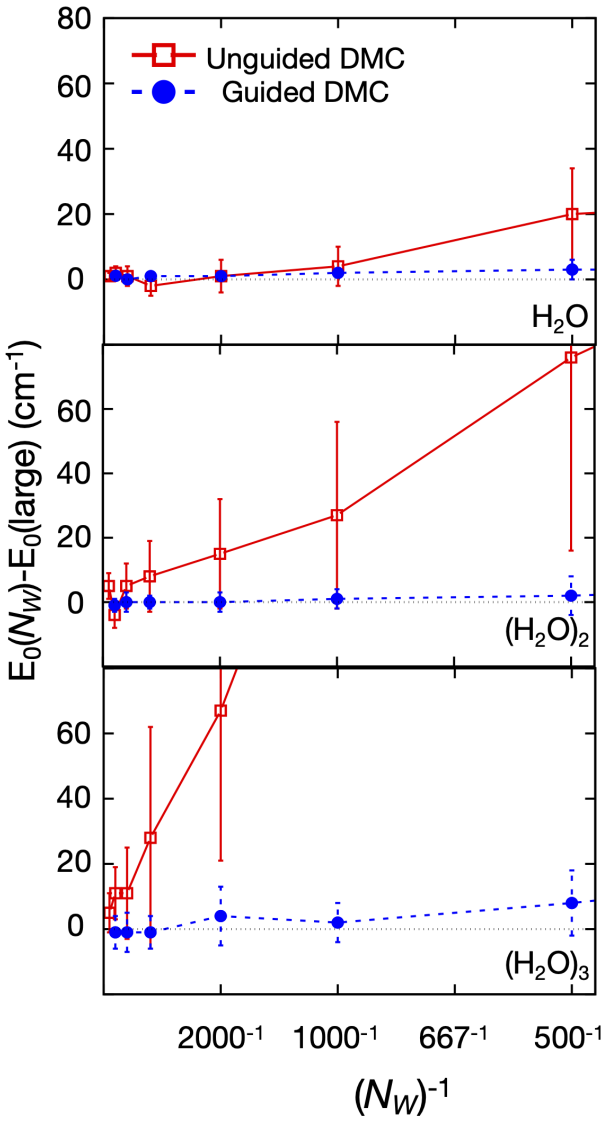

Figure 4.2: The same as Figure 4.1, plotted on an expanded scale.

Table 4.2: Values of the Difference in Ground State Energies (in cm^{-1}) Plotted in Figure 4.1 for (H_2O).

N_W^a	Unguided[b,c]	Guided[b,d]
100	52(38)	2(3)
500	20(14)	3(3)
1000	4(6)	2(1)
2000	1(5)	1(1)
5000	-2(3)	1(1)
10 000	1(3)	0(1)
20 000	2(2)	1(1)
40 000	1(1)	

[a] Number of walkers.

[b] Values are subtracted from the energy obtained from an unguided calculation with N_W = 60 000, where E_0=4637(1) cm^{-1}.

[c] The ground state energy and uncertainty are based on five independent simulations performed with $\Psi_T = 1$.

[d] The ground state energy and uncertainty are based on five independent simulations performed using Ψ_T in Eq. 4.1.

Table 4.3: Values of the Difference in Ground State Energies (in cm^{-1}) Plotted in Figure 4.1 for $(H_2O)_2$.

N_W^a	Unguided[b,c]	Guided[b,d]
100	347(132)	22(19)
500	76(60)	2(6)
1000	27(29)	1(3)
2000	15(17)	0(3)
5000	8(11)	0(2)
10 000	5(7)	0(3)
20 000	-4(4)	-1(2)
40 000	5(4)	

[a] Number of walkers.
[b] Values are subtracted from the energy obtained from an unguided calculation with N_W = 60 000, where E_0=8174(2) cm^{-1}.
[c] The ground state energy and uncertainty are based on five independent simulations performed with $\Psi_T = 1$.
[d] The ground state energy and uncertainty are based on five independent simulations performed using Ψ_T in Eq. 4.1.

Table 4.4: Values of the Difference in Ground State Energies (in cm^{-1}) Plotted in Figure 4.1 for $(H_2O)_3$.

N_W^a	Unguided[b,c]	Guided[b,d]
100	2884(399)	68(23)
500	354(70)	8(10)
1000	160(30)	2(6)
2000	67(46)	4(9)
5000	28(34)	-1(5)
10 000	11(14)	-1(6)
20 000	11(8)	-1(5)
40 000	5(6)	

[a] Number of walkers.
[b] Values are subtracted from the energy obtained from an unguided calculation with N_W = 60 000, where E_0=10 116(4) cm^{-1}.
[c] The ground state energy and uncertainty are based on five independent simulations performed with $\Psi_T = 1$.
[d] The ground state energy and uncertainty are based on five independent simulations performed using Ψ_T in Eq. 4.1.

on the values of the three coordinates, but its coordinate dependence will be much weaker than that of the underlying potential energy surface for water. A consequence of this weak coordinate dependence is the small number of walkers that are required to obtain an accurate ground state energy for water monomer from a guided DMC simulation. Specifically, when as few as 100 walkers are used, the calculated ground state energy differs from the fully converged value by 2 cm^{-1} while the statistical uncertainty based on five independent DMC simulations of the ground state energy is 6 cm^{-1}. The size of the error in the calculated energy and its statistical uncertainty both decrease as the size of the ensemble is increased. In contrast, when an unguided DMC simulation is performed, the error in the ground state energy is 52 cm^{-1} when 100 walkers are included in the ensemble. While the uncertainty in the energy is an

order of magnitude larger for the unguided simulation than for the guided one (60 compared to 6 cm^{-1}), the error in the calculated groud state energy is larger than this uncertainty. It is only when the ensemble has more than 5000 walkers that the error in the energy and its uncertainty are comparable to the values obtained from the guided simulation with 100 walkers. When a second water molecule is introduced,

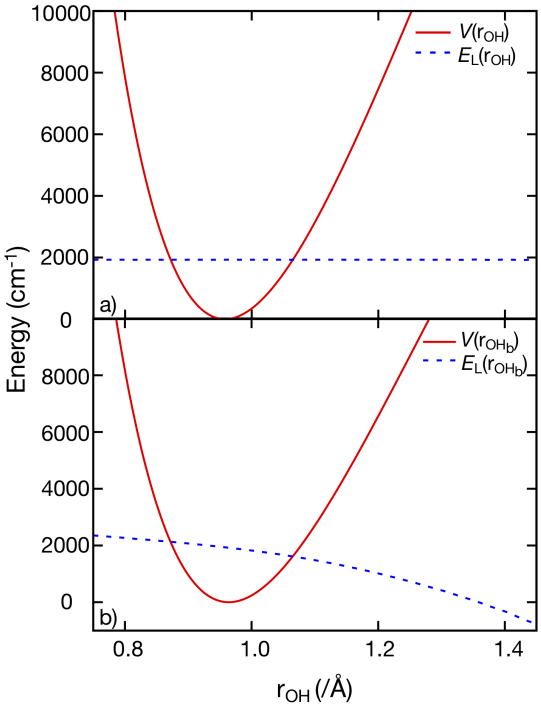

Figure 4.3: Comparison of the a cut through the potential (solid red lines) and local energies (dashed blue lines), plotted as functions of a) an OH bond length in H_2O, and b) the hydrogen-bonded OH bond length in $(H_2O)_2$. For both systems, Ψ_T is the ground state wave function for one of the OH bonds in water monomer.

at least 500 walkers are required to obtain an error below 5 cm^{-1}, and an uncertainty of roughly the same size when a guided DMC calculation of the ground state energy is performed. The increase in ensemble size compared to water momomer reflects the fact that Ψ_T in Eq. 4.1 is a function of only the intramolecular vibrational degrees of freedom. As a result, unlike the water monomer, not all of the vibrational degrees

of freedom are included in the definition of the guiding function. In an isolated water molecule, there are only three intramolecular degrees of freedom, while in the dimer, there are six intramolecular coordinates and six intermolecular coordinates. These intermolecular degrees of freedom are sampled without the benefit of a guiding function. Additionally, Ψ_T is based on cuts through the potential for water monomer. The equilibrium geometries, harmonic frequencies and anharmonicities of the water molecules that make up the water clusters deviate slightly from those of an isolated water molecule. As a result, when we consider the coordinate dependence of E_L for a one-dimensional cut through the potential along the hydrogen bonded OH bond length in $(H_2O)_2$, shown in Figure 4.3(b), E_L now shows a coordinate dependence, although it is still much weaker than the coordinate dependence of the potential function. It is interesting to note that the negative slope of E_L in Figure 4.3(b) has the effect of shifting walkers to longer OH bond lengths, as is required to model the larger anharmonicity of the hydrogen bonded OH bond in water dimer compared to an OH bond in water monomer.

Examination of the results for $(H_2O)_2$ reported in the middle panel of Figure 4.1 shows that the ground state energies obtained from the guided DMC simulation agree to within 5 cm^{-1} when 500 or more walkers are included in the ensemble. For the unguided simulation, ensemble sizes of at least 10 000 walkers are required to obtain a similar level of convergence to what we are able to achieve with 500 wakers in a guided simulation. As for the monomer, the uncertainties in the energies are larger for the unguided simulations than for the guided ones. Finally, similar behavior is found for water trimer where the errors in calculated ground state energies are smaller than 5 cm^{-1} when 1000 or more walkers are used in a guided simulation, while 40 0000 walkers are required to achieve similar convergence of the ground state energy from an unguided simulation. Based on the above results, decreases in the size of ensemble needed for DMC simulations of water clusters can be decreased by more than an order of magnitude through the introduction of well-chosen guiding functions.

In Table 4.5, the energies that have been calculated in this study are compared to those reported by Mallory and Mandelshtam, who performed an extensive study of the convergence of DMC approaches for water monomer and dimer using the same potential surface.[25,75] The energies reported in these studies were obtained using a larger value of $\Delta\tau$ (10 a.u. compared to 1 a.u.), and an extrapolation of the ground state energies to an infinite walker ensemble size. As is seen in the results reported in Table 4.5, these energies deviate somewhat from those reported in the present work. On the other hand, when we look at their reported ground state energies evaluated using 19 600 walkers and extrapolated to the zero time-step limit (given in the Extrap. column of Table 4.5), the agreement is excellent. In Table 4.5, we also compare our results to the reported binding energies for water dimer[81,82] and trimer[83] reported by Reisler and co-workers. As noted by Mandelshtam, these binding energies have a smaller time-step dependence than the ground state energies and, as a result, all the calculations provide very similar values for the binding energies of these water clusters. Additionally all of these values agree with the reported experimental values to within their uncertainties.

The energy is only one measure of the accuracy of DMC. In Figure 4.4, we compare projections of the ground state probability amplitude for water trimer onto one of the bound OH bonds (Figure 4.4(a)) and one of the OO distances (Figure 4.4(b)). As is seen, the two approaches give virtually indistinguishable results. To aid in the comparison of these two pairs of distributions, we have also plotted the difference between them below each of these plots in Figure 4.4(c) and (d).

Taken together, the above results demonstrate the accuracy of the partially importance sampled DMC approach proposed in the present study. We show that we are able to obtain accurate results with significantly smaller ensemble sizes compared to DMC simulations, which are performed with $\Psi_T = 1$. These smaller ensemble sizes translate to significantly lower computational costs compared to unbiased simulations. While the savings are useful for calculations of H_2O, $(H_2O)_2$ and $(H_2O)_3$

Table 4.5: Comparison of the Ground State Energies of Water Monomer and Dimer in cm^{-1} to Those Reported in Ref. 75.

System	$E_0^{UG\ a,b}$	$E_0^{G\ b,c}$	Ref. 25	Extrap.[d]	Expt.
H_2O[e]	4637(1)	4637(1)	4631	4637	
$(H_2O)_2$	8174(2)	8174(2)	8160[f]	8173	
D_0^g	1100(3)	1100(3)	1102	1101	1105 ± 10[h]
D_2O[e]	3391(1)	3390(1)	3389	3391	
$(D_2O)_2$	5541(2)	5541(1)	5540[f]	5540	
D_0^g	1241(3)	1239(2)	1238	1242	1244 ± 10[i]
$(H_2O)_3$	10 116(4)	10 115(5)	10 100[f]		
D_0^j	2695(5)	2696(5)	2693		2650 ± 150[k]
$(D_3O)_3$	6006(8)	6008(3)	6005[f]		
D_0^j	2926(8)	2923(3)	2924		

[a] This study, evaluated using an unguided DMC simulation with $N_W = 60\ 000$ walkers.

[b] Numbers in parentheses provide one standard deviation based on 5 independent DMC simulations.

[c] This study, evaluated using an guided DMC simulation with $N_W = 10\ 000$ walkers.

[d] Obtained from the data presented in Figure 1 of Ref. 75.

[e] The ground state energy for H_2O[41] is 4638.39 cm^{-1} and the ground state energy for D_2O[84] is 3390.26 cm^{-1}.

[f] Evaluated using the binding energies for water dimer or trimer.[25]

[g] Energy for $(H_2O)_2 \rightarrow 2\ H_2O$.

[h] Ref. 81.

[i] Ref. 82.

[j] Calculated energy for $(H_2O)_3 \rightarrow (H_2O)_2 + H_2O$.

[k] Ref. 83.

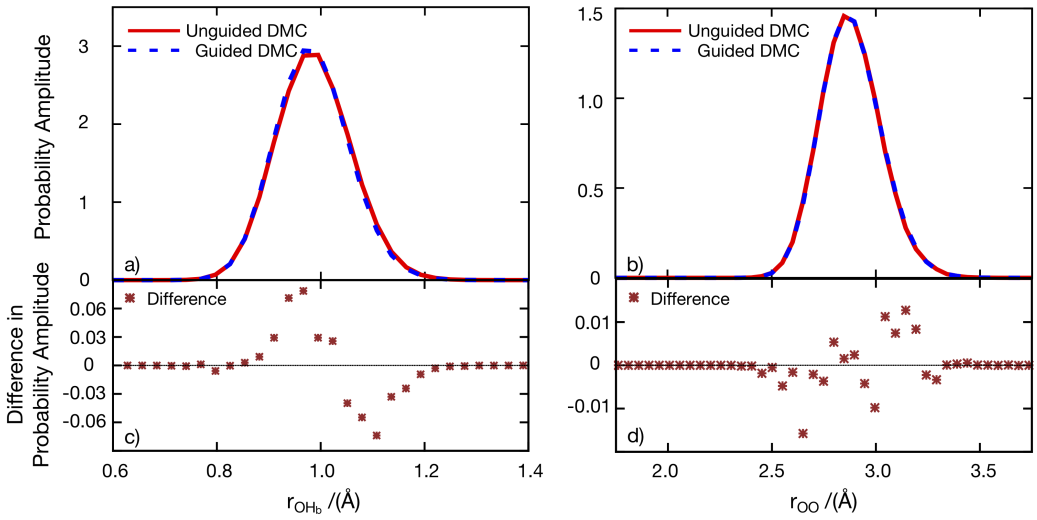

Figure 4.4: Projections of the ground state probability amplitude for water trimer onto (a) a hydrogen-bonded OH bond length (OH_b) and (b) an OO distance (R_{OO}). The differences between the two functions shown in panels (a) and (b) are displayed with brown *'s in panels (c) and (d), respectively. The plots in the upper panel are based on an unguided simulation with 60 000 walkers (red solid lines) and a DMC simulation with a guiding function and 5000 walkers (blue dashed lines).

the power of the approach comes in the ability to perform calculations of multiple conformers or isotopologues of a single cluster size and in the facilitation of studies of larger cluster sizes. As a demonstration of this, in the following section, we explore the ground state probability amplitude various isotoplogues of water trimer. Here we focus on the question of how partial deuteration affects how the cluster samples various low-energy minima on the potential energy surface.

4.4.2 Water trimer

Water trimer has two distinct isomers that are accessible by the zero-point level. They are characterized by the direction of the free OH bonds relative to the plane that contains the three oxygen atoms. As is seen in Figure 4.5, the lowest energy structure has two of the OH bonds above the plane of the oxygen atoms, while the third OH bond is below this plane. This isomer will be identified by UDU. The second isomer has all three of the OH bonds above the plane of the oxygen atoms. This isomer is denoted UUU, and is 248 cm^{-1} higher in energy than the UDU isomer. There are two more sets of isomers, which are denoted as the DDD and DUD. As these are equivalent to the corresponding UUU and UDU structures, we use the UUU and UDU notation to refer to these pairs of equivalent forms of the water trimer. In mixed isotopologues, H and D superscripts are added to the "U" and "D" to differentiate among the various isotpologues. Finally, while in HOD it is possible for either the OH or the OD bond to be involved in hydrogen bonding, when the ground state energy is accounted for, only those isotopologues where the hydrogen bond involves the OD bond in HOD are found to contribute to the ground state structure.[85,86].

Characterizing how the ground state probability amplitude samples these minima requires an approach for associating each walker in the ensemble with a specific minimum in the potential. This is achieved by using the approach described by Mallory and Mandelshtam[75] in which the structure of each walker is relaxed to a geometry that corresponds to a minimum on the potential using the conjugate gradient method.

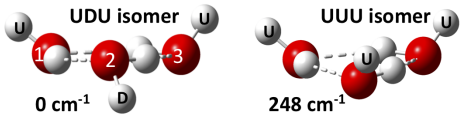

Figure 4.5: The two isomers of water trimer discussed in the text. They differ by the orientation of the free OH bonds. The isomer on the right has all three OH bonds above the plane containing the oxygen atoms, and is denoted UUU, while in the isomer on the left one of the free OH bonds points below the plane containing the three oxygen atoms and is denoted UDU.

The signs of the displacements of the free OH bonds from the plane that contains the three oxygen atoms are then used to label the minimum. As in the structure shown in Figure 4.5, the in the UDU structures, the water molecule labeled 1 makes a hydrogen bond with water 2, water 2 makes a hydrogen bond with water 3, and water 3 makes a hydrogen bond with water 1. The results of this analysis is reported in Table 4.6.

Table 4.6: Fractional Isomer Population of the Minima on the Potential for Water Trimer.

System		UDU		UUU
$(H_2O)_3$		0.979(0.005)		0.021(0.005)
$(D_2O)_3$		0.991(0.003)		0.009(0.003)
$(HOD)_3$		0.980(0.005)		0.020(0.005)
System	$U_2D_3U_1^c$	$U_1D_2U_3^c$	$U_3D_1U_2^c$	UUU
$(H_2O)_2(D_2O)^a$	0.35(0.04)	0.34(0.04)	0.30(0.05)	0.017(0.006)
$(H_2O)(D_2O)_2^b$	0.37(0.06)	0.35(0.06)	0.27(0.06)	0.012(0.004)
$(H_2O)_2(HOD)^a$	0.35(0.04)	0.31(0.05)	0.32(0.05)	0.021(0.006)
$(H_2O)(HOD)_2^b$	0.34(0.05)	0.33(0.05)	0.30(0.03)	0.022(0.005)
$(HOD)_2(D_2O)^a$	0.34(0.05)	0.32(0.05)	0.32(0.05)	0.017(0.005)
$(HOD)(D_2O)_2^b$	0.36(0.05)	0.36(0.06)	0.26(0.05)	0.012(0.004)

[a] Molecules 1 and 2 are H_2O or HOD, while molecule 3 is D_2O or HOD.
[b] Molecules 2 and 3 are D_2O or HOD, while molecule 1 is H_2O or HOD.
[c] See Figure 4.5

As noted by Mallory and Mandelstam, who performed unguided DMC simulations

Table 4.7: Harmonic Ground State Corrected Energies of the Minima on the Potential for Water Trimer and the Calculated Anharmonic Ground State Energy.

System		UDU[a]		UUU[a]	DMC[b]
$(H_2O)_3$		10 543		10 674	10 115(3)
$(D_2O)_3$		6240		6403	6006(5)
$(HOD)_3$		8314		8456	7989(6)
System	$U_2D_3U_1^a$	$U_1D_2U_3^a$	$U_3D_1U_2^a$	UUU	DMC[b]
$(H_2O)_2(D_2O)^c$	9106	9113	9110	9251	8746(4)
$(H_2O)(D_2O)_2^d$	7672	7675	7680	7827	7377(5)
$(H_2O)_2(HOD)^c$	9799	9801	9800	9935	9407(5)
$(H_2O)(HOD)_2^d$	9055	9058	9059	9196	8698(3)
$(HOD)_2(D_2O)^c$	7620	7625	7623	7772	7329(3)
$(HOD)(D_2O)_2^d$	6930	6931	6934	7087	6669(4)

[a] See Figure 4.5

[b] The number in parentheses provide one standard deviation based five independent simulations with 5000 walkers.

[c] Molecules 1 and 2 are H_2O or HOD, while molecule 3 is D_2O or HOD.

[d] Molecules 2 and 3 are D_2O or HOD, while molecule 1 is H_2O or HOD.

using the same potential surface, in $(H_2O)_3$ roughly 98% of the probability amplitude is assigned to the UDU conformer, while the remaining 2% reflects UUU geometries. This can be seen from projections of the ground state probability amplitude onto the displacements of two of the three OH bonds from the plane that contains the three oxygen atoms, shown in Figure 4.6(a). In the plots in Figure 4.6, the third OH bond is constrained to lie above the plane that contains the three oxygen atoms, and the structures that correspond to the regions of higher probability are identified in the corners of the plots.

As is seen, the amplitude is localized in the corners of the plots, with amplitude extending between the three corners that are associated with UDU structures. Compared to the other three regions, there is significantly less amplitude in the upper left corner, which corresponds to the UUU structure of water trimer. This is consistent with the results reported in Table 4.6. The amplitude between the corners that cor-

respond to UDU structures reflects the large amplitude motion of the OH bonds over the roughly 200 cm^{-1} barrier that separates this minimum. As is seen by comparing the panels (a) and (b) of Figure 4.6, the amplitude is larger when this motion corresponds to the displacement of an OH bond compared to displacements of OD bonds. Consistent with the results provided in Table 4.6, there is also less amplitude in the corner that corresponds to the UUU structure in $(D_2O)_3$ compared to $(H_2O)_3$. Finally, in $(H_2O)_3$ and $(D_2O)_3$ the three free OH bonds are equivalent, and of the UDU structures should have equal probability. The apparent differences between the amplitude in the lower left corner of these projections compared to the other two corners that correspond to UDU structures reflects the fact that the amplitude has been projected over the other degrees of freedom, particularly the large amplitude displacements of the third OH or OD bond off of the plane that contains the three oxygen atoms.

The situation becomes more interesting when one or two of the H_2O molecules in $(H_2O)_3$ are replaced by D_2O molecules. In these cases, there are three distinguishable forms of the UDU isomer based on which of the three monomers has a free OH bond that is in the down configuration. The results are presented in Table 4.6. We start by noting that the fraction of trimers that explore the UUU confirmation decreases as the number of D_2O molecules are increased. This is consistent with the increase in the energy difference between the two classes of isomers once the ground state energy is considered. In the absence of the ground state energy, the UUU structure is 248 cm^{-1} higher in energy than the UDU structure. In Table 4.7 we report the energies of the various structural isomers once ground state energy is introduced. As is seen, the energy difference between the various forms of the UDU isomer and the UUU isomer increases as the number of deuterium atoms is increased, while the differences among the ground state energies of the various UDU isomers of each isotopologue differ by between 4 and 11 cm^{-1}.

When we compare the fraction of the probability amplitude that is assigned to the

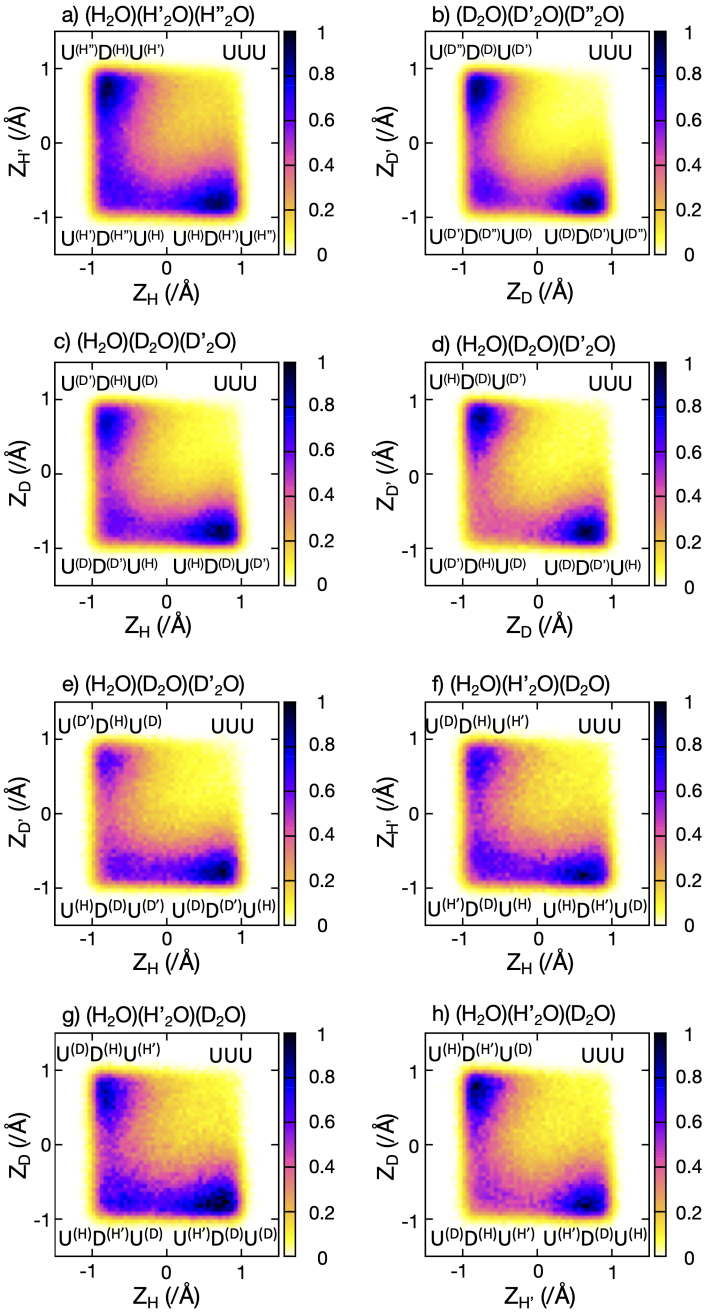

Figure 4.6: Projections of the ground state probability amplitude for water trimer onto the displacements of two of the free OH or OD bonds off of the plane that contains the three oxygen atoms for (a) $(H_2O)_3$, (b) $(D_2O)_3$, (c) the free OH and OD bonds in $(H_2O)(D_2O)_2$, (d) the two free OD bonds in $(H_2O)(D_2O)_2$, (e) the free OH and OD′ bonds in $(H_2O)(D_2O)_2$, (f) the free OH and OH′ bonds in $(H_2O)_2(D_2O)$, (g) the free OH and OD bonds in $(H_2O)_2(D_2O)$ and (h) the free OH′ and OD bonds in $(H_2O)_2(D_2O)$. In all cases, the third free OH or OD bond in the trimer is in a geometry that corresponds to a positive Z-value. These plots are based on the wave functions used to evaluate the isomer fractions reported in Table 4.6.

three minima that correspond to UDU structures, we find that the fractions are nearly equal in $(H_2O)_2(D_2O)$, and range from 0.30 to 0.35. These values are close to what one would expect if the minima were sampled based on simple statistics, and when the uncertainties are considered, the values are all roughly equal. Closer examination of the results of several simulations of the $(H_2O)_2(D_2O)$ cluster shows that the minima that correspond to the $U_2^{(H)}D_3^{(D)}U_1^{(H)}$ appears to have slightly more amplitude than the other two minima, which have comparable amplitude. This can be visualized through the projections of the probability amplitude onto the displacements of two of the OH or OD bonds from the plane that contains the three oxygen atoms provided in panels (f), (g) and (h) of Figure 4.6.

For $(H_2O)(D_2O)_2$ larger differences are found. The amplitude is similar for the $U_2^{(D)}D_3^{(D)}U_1^{(H)}$ (37%) and $U_1^{(H)}D_2^{(D)}U_3^{(D)}$ (35%) conformers, and smallest (27%) for the $U_3^{(D)}D_1^{(H)}U_2^{(D)}$ conformer. These trends can be rationalized by considering whether moving a H-atom from an U to a D position would lead to an UUU or an UDU structure. For the more populated minima, this motion would connect the minimum of interest to another UDU minimum. We focus on the H-atom because, as shown in Figure 4.6(a) and (b), the H displacements off of the OOO plane show much larger amplitude compared to the displacements of the D-atoms. Further examination of panels (c) and (d) of this figure provide a pictorial visualization of the changes in the size of the ground state probability amplitude for the various UDU minima in the potential for $(H_2O)(D_2O)_2$. This can be visualized through the projections of the probability amplitude onto the displacements of two of the OH or OD bonds from the plane that contains the three oxygen atoms provided in panels (c) and (d) of Figure 4.6. Comparison of these plots to those for $(H_2O)_3$ and $(D_2O)_3$ shows the lower amplitude in the $U^{(D)}D^{(H)}U^{(D)}$ structure compared to the other two UDU structures.

We also explore other five isotopomers of water trimer obtained by combinations of HOD, H_2O and D_2O. These results are also provided in Table 4.6. The trends are generally consistent with the discussion above. Specifically, in its ground state,

the HOD molecules assume configurations in which the OD bond is involved in the hydrogen bond, while the OH bond is unbound. In this way, $(H_2O)_n(HOD)_{3-n}$ is expected to appear similar to $(H_2O)_3$, while complexes of HOD and D_2O molecules should show similar trends to the $(H_2O)_n(D_2O)_{3-n}$ counterparts. Comparing the fraction of the isomers with the UUU configuration, we find that the larger the fraction of OH bonds in the free positions, the larger the fraction of the probability amplitude is sampling the UUU configurations. This is consistent with the amplitude of the motion of the free OH compared to free OD bonds seen in Figure 4.6. As with the sampling of thee various UDU isomers, the fraction of the amplitude that is in the UUU configurations is independent of whether the free OH bond is part of an HOD or an H_2O molecule.

We repeated the above analysis using a second approach for assigning walkers to minima in the potential. In this approach, we evaluate the sign of the displacements of the three free OH or OD bonds off of the plane that contains the three oxygen atoms based on the geometry of the walkers. The values of these displacements are used to deterime the minimum that the walker is assigned to. Use of this definition leads to a slightly larger fraction of the probability amplitude is assigned to the UUU structure, while the relative amplitude in the various UDU minima is unaffected. These results are provided in Table 4.8.

Finally, as we look at the results provided in Tables 4.6 and 4.8, we note that the uncertainties in the isomer fractions among the various UDU isomers are comparable to or larger than the differences among these fractions. For this reason, we repeated these calculations both with the same number of walkers (5000) and using twice as many walkers (10 000). These results are provided in Tables 4.9 and 4.10, respectively. While there is variability in the results, consistent with the stated uncertainties, the qualitative trends described above are consistent with these accumulated findings. The similarity of the UDU minima and the fact that the populations in these minima are nearly equal in all systems investigated makes this a more challenging calculation

Table 4.8: Fractional Population of the Minima on the Potential for Water Trimer Based on the Orientations of the Free OH Bonds at the Geometry of the Walkers

System		UDU[a]		UUU[a]
$(H_2O)_3$		0.91(0.02)		0.092(0.016)
$(D_2O)_3$		0.96(0.01)		0.043(0.010)
$(HOD)_3$		0.92(0.01)		0.082(0.014)
System	$U_2D_3U_1^a$	$U_1D_2U_3^a$	$U_3D_1U_2^a$	UUU
$(H_2O)_2(D_2O)^b$	0.33(0.04)	0.32(0.04)	0.28(0.05)	0.074(0.014)
$(H_2O)(D_2O)_2^c$	0.35(0.06)	0.33(0.06)	0.26(0.06)	0.059(0.012)
$(H_2O)_2(HOD)^b$	0.33(0.05)	0.28(0.05)	0.30(0.05)	0.092(0.018)
$(H_2O)(HOD)_2^c$	0.32(0.05)	0.31(0.05)	0.28(0.03)	0.089(0.015)
$(HOD)_2(D_2O)^b$	0.32(0.05)	0.31(0.05)	0.30(0.05)	0.074(0.014)
$(HOD)(D_2O)_2^c$	0.35(0.05)	0.35(0.06)	0.25(0.05)	0.059(0.011)

[a] See Figure 4.5.
[b] Molecules 1 and 2 are H_2O or HOD, while molecule 3 is D_2O or HOD.
[c] Molecules 2 and 3 are D_2O or HOD, while molecule 1 is H_2O or HOD.

compared to evaluating isomer fractions between structurally different isomers, as is the case for the comparison of the UUU and UDU fractions.

4.5 Conclusions

In this study, we presented an alternative approach for applying DMC for studies of molecular clusters. Motivated by the success of rigid-body DMC approaches for studies of such systems, we propose an algorithm in which the intramolecular degrees of freedom are sampled using guiding functions, while the intermolecular degrees of freedom are sampled using the unbiased form of DMC. This provides the advantage of removing the high-frequency contributions to E_{ref} by substituting the local energy for the potential energy in the evaluation of E_{ref}. This modification allows us to obtain converged energies and wave functions at a much lower computational expense compared to calculations performed with $\Psi_T = 1$.

We apply these approaches to studies of $(H_2O)_n$ with $n = 1$ to 3. We show that,

Table 4.9: Fractional Population of the Minima on the Potential for Water Trimer Obtained for a Second Set of Ensembles of 5000 Walkers.

System		UDU[a]		UUU[a]
$(H_2O)_3$		0.979(0.004)		0.021(0.004)
$(D_2O)_3$		0.991(0.004)		0.009(0.004)
$(HOD)_3$		0.978(0.004)		0.021(0.004)
System	$U_2D_3U_1^a$	$U_1D_2U_3^a$	$U_3D_1U_2^a$	UUU
$(H_2O)_2(D_2O)^b$	0.37(0.05)	0.30(0.06)	0.32(0.05)	0.017(0.005)
$(H_2O)(D_2O)_2^c$	0.37(0.06)	0.36(0.07)	0.25(0.07)	0.013(0.007)
$(H_2O)_2(HOD)^b$	0.33(0.05)	0.33(0.04)	0.32(0.04)	0.021(0.007)
$(H_2O)(HOD)_2^c$	0.32(0.04)	0.33(0.04)	0.32(0.04)	0.021(0.005)
$(HOD)_2(D_2O)^b$	0.32(0.06)	0.34(0.06)	0.32(0.04)	0.015(0.004)
$(HOD)(D_2O)_2^c$	0.36(0.05)	0.38(0.06)	0.25(0.05)	0.013(0.005)

[a] See Figure 4.5.
[b] Molecules 1 and 2 are H_2O or HOD, while molecule 3 is D_2O or HOD.
[c] Molecules 2 and 3 are D_2O or HOD, while molecule 1 is H_2O or HOD.

Table 4.10: Fractional Population of the Minima on the Potential for Water Trimer Obtained for Ensembles of 10 000 Walkers.

System		UDU[a]		UUU[a]
$(H_2O)_3$		0.980(0.004)		0.020(0.004)
$(D_2O)_3$		0.990(0.003)		0.010(0.003)
$(HOD)_3$		0.979(0.004)		0.020(0.004)
System	$U_2D_3U_1^a$	$U_1D_2U_3^a$	$U_3D_1U_2^a$	UUU
$(H_2O)_2(D_2O)^b$	0.33(0.04)	0.35(0.04)	0.30(0.04)	0.017(0.004)
$(H_2O)(D_2O)_2^c$	0.38(0.04)	0.36(0.03)	0.25(0.03)	0.012(0.003)
$(H_2O)_2(HOD)^b$	0.34(0.03)	0.32(0.03)	0.32(0.03)	0.020(0.004)
$(H_2O)(HOD)_2^c$	0.33(0.03)	0.32(0.04)	0.33(0.04)	0.020(0.004)
$(HOD)_2(D_2O)^b$	0.34(0.04)	0.34(0.04)	0.30(0.04)	0.016(0.003)
$(HOD)(D_2O)_2^c$	0.36(0.05)	0.33(0.05)	0.29(0.06)	0.013(0.003)

[a] See Figure 4.5.
[b] Molecules 1 and 2 are H_2O or HOD, while molecule 3 is D_2O or HOD.
[c] Molecules 2 and 3 are D_2O or HOD, while molecule 1 is H_2O or HOD.

in contrast to studies in which all of the vibrational degrees of freedom are treated without guiding functions, significantly smaller ensemble sizes can be used to obtain accurate zero-point energies. We also use this approach in a study of isomer fractions in water trimer and its deuterated analogues. Analysis of the ground state probability amplitudes shows that partial deuteration affects how this cluster samples the various minima on the potential. Further, these relatively small ensembles of walkers can be used to obtain accurate descriptions of these properties. We find that there is larger probability amplitude in the minimum that corresponds to the UUU configuration shown in Figure 4.5 as the number of OH bonds in the free position is increased. The trends in population of the three isomers that correspond to UDU configurations can be rationalized in terms of the larger amplitude of the motion of the OH bonds, compared to that for the OD bonds, as they move from above to below the plane that contains the oxygen atoms, shown in Figure 4.6.

The computational savings realized by the partial importance sampling approach developed in this study provides the opportunity to explore larger systems, where multiple structural isomers have similar zero-point energies, and their populations will be highly sensitive to partial deuteration.

Chapter 5

DIFFUSION MONTE CARLO STUDIES OF ISOTOPIC SUBSTITUTION IN WATER HEXAMER

5.1 Introduction

The structure and spectroscopy of water clusters have received significant attention as they represent molecular systems through which the role of hydrogen-bonding on the structure, stability and spectra of small assemblies of water molecules can be explored in detail. In addition, they provide a venue through which detailed comparison between experiment and theory can be made.[22,29,30,32,85,87–95] For the water clusters with five or fewer water molecules ($N_w \leq 5$), the most stable forms of are rings in which each water molecule is involved in two hydrogen bonds: one as a hydrogen-bond donor and one as a hydrogen-bond acceptor. Clusters with six water molecules are the smallest clusters for which the lowest energy structures are three-dimensional. In these structures, most, if not all, of the water molecules are involved in three hydrogen bonds, either two in which the water molecule donates OH bonds into the hydrogen bond and one as an acceptor (ADD) or donating to one hydrogen bond and taking the role as the accepting oxygen in two hydrogen bonds (AAD), with one OH that is not involved in a hydrogen bond. While attention has been paid to the role of such bonding environments on the spectrum of liquid water,[96–99] and water clusters, less attention has been paid to the implications on the low-temperature structures of these clusters or the effects of deuteration on such structures.

In the case of $(H_2O)_6$, electronic structure calculations have shown that the most stable form of this size cluster is the prism, which is shown on the left side in Figure 5.1. On the right side of Figure 5.1, we show the second stable structure, which is a cage

structure that has four isomers. These isomers have similar energies, and the most stable cage isomer is 87 cm^{-1} higher in energy than the prism structure based on the MB-pol potential used in this study. All four forms of the cage are shown in Figure 5.2, and they differ only in the orientation of the two free OH bonds on the left and right side of the structure. Electronic structure calculations using CCSD(T) extrapolated to the complete basis set limit show that the energy of the prism is 87 cm^{-1} lower than the cage.[100] The increased stability of the prism relative to the cage likely reflects the fact that the prism contains nine hydrogen-bonding interactions, while the cage contains only eight. On the other hand the hydrogen bonds are generally more strained in the prism than the cage, as indicated by slightly longer OO distances (on average) for the hydrogen-bonding interactions. Elegant experiments by Johnson and Pate and their co-workers demonstrated that when zero-point energy is introduced, the lowest energy structure of $(H_2O)_6$ becomes the cage,[32,95,101] while for $(D_2O)_6$, the prism is lower in energy.[33] This conclusion that the most stable structure of $(H_2O)_6$ is the cage structure was supported by recent Diffusion Monte Carlo calculations of Mallory and Mandelshtam, while for $(D_2O)_6$, their studies found the cage and prism structures are very close in energy[25]. Based on a single very large Diffusion Monte Carlo (DMC) simulation, they concluded that for $(D_2O)_6$, the prism is more stable than the cage.

In this chapter, we will use the same DMC approach and guiding function outlined in Chapter 4 to calculate the ground state energies and ground state wave functions of different isomers and isotopologues of water water hexamer. This allows us to obtain converged results using much smaller ensemble sizes as well as shorter time steps for the propagation compared to the previous study by Mallory and Mandelshtam.[24,25]. Using this approach we will explore the origins of the quantum effects that are responsible for these structural changes. In addition to exploring the structural properties of $(H_2O)_6$ (**6H**) and $(D_2O)_6$ (**6D**), we will also probe the role of of partial deuteration on the structures of the clusters and how they sample the potential surface by studying the zero-point properties of $(H_2O)_5(D_2O)$ (**5H1D**) and $(H_2O)(D_2O)_5$ (**1H5D**).

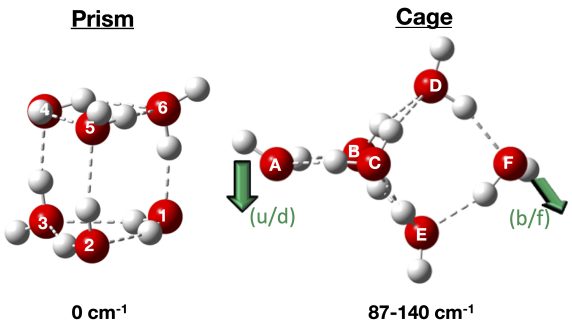

Figure 5.1: The minimum energy prism and cage structures for the water hexamer. The electronic energy shown here is the difference based off of the prism structure, which is the global minimum on the PES. The cage structure shown is the lowest energy structure out of the four cages that are considered in this study. The differences of each cage structure is based on the orientation of the free OH in the outer water which is denoted by the green arrows, and their electronic energies range from 87-140 cm^{-1} above the prism structure. All of the cage structures considered in this study are shown in Figure 5.2.

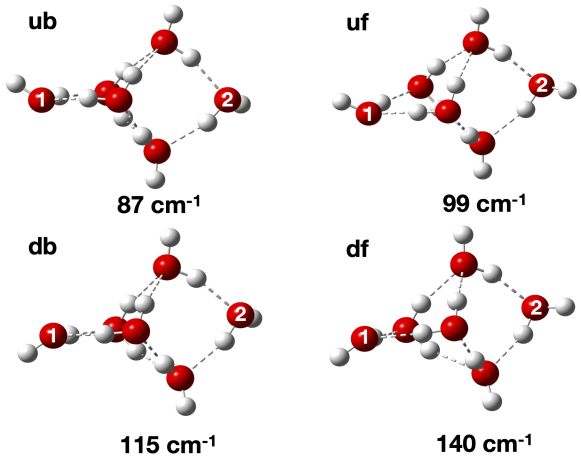

Figure 5.2: The four different cage structures that are considered in this study. The difference in each of the four structures is the orientation on either of the outer OH stretches where water 1 can flip up (u) or down (d), and water 2 can flip front (f) or back (b). The energy listed below each structure is the difference in electronic energy between that particular cage and the prism structure.

5.2 Methods and Computational Details

In this work, we use Diffusion Monte Carlo (DMC) to calculate the ground state energies and the vibrational ground state wave functions for several isomers in water hexamer, which has been outlined in Chapter 2 of this thesis. This approach has been used in the previous chapters to calculate the zero-point energies of various systems, such as one-dimensional Morse oscillators, CH_2O, isotopologues of H_3^+, and $(H_2O)_n$, where $n = 1-3$.[80,102] We will use the same approach to calculate the ground state energies and wave functions of water hexamer. The guiding function, Ψ_T, is the same guiding function that has been used in Chapter 4 of this thesis, where the OH stretches are based on the solutions to the one-dimensional Schrödinger equation based on cut of the OH stretch in the Partridge-Sckwenke potential.[41] The bend is described as a harmonic oscillator with the parameters listed in Table 5.1.

Table 5.1: Frequencies and Wilson G-matrix Elements Used to Define the Bend Contribution to Ψ_T in Eq. 4.1.

Molecule	Angle	ω/cm^{-1}	G^a/amu^{-1} Å^{-2}
6H	HOH	1655	2.338
5H1D	HOH	1668	2.338
	DOD	1222	1.255
1H5D	HOH	1668	2.338
	DOD	1222	1.255
6D	DOD	1207	1.255

[a] Wilson G-matrix element.[1].

Because these systems can sample several low-energy minima, care was taken to ensure that the choice of initial conditions would not affect the final conclusions. To this end, we started with the three lowest-energy structures - the prism, the cage (which has four isomers) and the prism-book (which has four conformers). For each of these nine structures, we generated a set of initial conditions based on the harmonic

ground state wave function in which the width was doubled to ensure that high energy configurations were being sampled, following the same procedure we used to sample the initial conditions in our study of the water trimer. In order to ensure that we are sampling all the relevant configurations, we choose the initial conditions such that they do not bias our results. For each geometry, we propagated an equal sized ensemble of walkers over a short time to equilibrate the ensembles for each of these minima in the potential surface. After 2000 time steps, these small ensembles were combined to generate a single larger ensemble. Initially, we generated our distribution of walkers using nine pre-equilibrated distributions, one prism (P), four cage (C) and four prism book (PB) structures. These initial conditions will be referred to as C/P/PB. We also performed simulations in which half of the ensemble was equilibrated based on the prism ground state and the other half from the four cage structures, and refer to these as C/P.

For the guided DMC simulations we ran two sets of five DMC calculations, one for each set of initial conditions, described above, while for the unguided simulations we ran five calculations based on the C/P/PB initial conditions. For all simulations a 1 a.u. time step was used and the potential energies are based on the MB-pol potential energy surface developed by Paesani and co-workers.[76,77,79,103] The combined ensembles were propagated for 98 000 time steps, or a total of 100 000 time steps including the initial equilibration of the smaller ensembles. Discrete weighting was used for this part of the calculations.

The ground state energy for each system was calculated by averaging E_{ref} starting at 60 000 time steps. After 60 000 time steps, the coordinates of the walkers were collected at every 2500 time steps, generating 17 wave functions used in the overall analysis. Those wave functions were propagated for an additional 150 time steps in order to obtain the descendant weights needed to analyze the probability amplitudes. For this part of the calculations, we used continuous weighting.

Finally for the calculations of the mixed clusters, we performed separate calcu-

lations for the prism and the cage with the parameters used to describe the OH stretches listed above and the parameters used for the bend listed in Table 5.1. The initial conditions were obtained by performing separate smaller calculations where the unique molecule is placed in each of the six possible positions in the cluster. This was done to minimize bias as to where this molecule ends up at the end of the simulation. Based on convergence studies for the **H6** and **D6** systems, these calculations employed 50 000 walkers, and were also run for 100 000 time steps. We use the same approach described for **6H** and **6D** to calculate the ground state energies and analyze the probability amplitudes of the ground state wave function for the mixed clusters.

5.3 Results and Discussion

5.3.1 Energies

Before considering the possible structures of the water hexamer, we start by exploring the convergence properties of the ground state energy of **6H**. For this part of the study we will focus on simulations of the cage structure. As mentioned in the introduction, the earlier work of Mallory and Mandelshtam[25] employed 1 000 000 walkers and propagated for 88 000 time steps of 10 a.u. using the same potential surface obtained in this study, found that the ground state energy of the cage structure is 16 240 cm^{-1}. The results of the present study are provided in Figure 5.3. Based on the results using the largest ensembles, 150 000 walkers for the unguided DMC calculations and 100 000 walkers for the guided one, the energy obtained from the guided simulation is nearly 100 cm^{-1} smaller than the energy obtained from the unguided simulation, and 20 cm^{-1} higher in energy compared to the ground state energy reported by Mallory and Mandelshtam. This is not surprising as earlier studies on the water hexamer showed that at least 480 000 walkers are needed to accurate accurate ground state energies using the q-TIP4P potential energy surface.[24] In contrast, all of the reported energies obtained from the guided simulations differ by no more than 25 cm^{-1} while the size

of the ensemble is increased from 10 000 to 100 000 walkers, and the energies agree to within the reported uncertainties for ensemble sizes of at least 30 000 walkers. The observed faster convergence with ensemble size when the guided DMC approach is used is consistent with what we previously reported for water trimer.[102] Although the results of the guided DMC calculation appear to be converged, the ground state energy is 18 cm^{-1} higher than the 16 240 cm^{-1} as stated by Mallory and Mandelshtam. This is likely due to the difference in the time steps used for the two studies – they used a time step of 10 a.u., while we used a time step of 1 a.u. Studies of the water monomer showed that the ground state energy calculated using the two time-steps differ by 6 cm^{-1} with the smaller time step providing an energy that is in good agreement with converged variational calculations. There is ongoing work in the understanding of the convergence properties of the energies for the unguided calculations.

Table 5.2: Values of the Ground State Energies for **6H** Using Guided and Unguided DMC Based on the Calculations Described in the Text.

N_W^a	Guided DMC				Unguided DMC			
	E_0 Cage [b] (/ cm^{-1})	$n_{\text{cage}}^{b,c}$	E_0 Prism [d] (/ cm^{-1})	$n_{\text{prism}}^{d,e}$	E_0 Cage [b] (/ cm^{-1})	$n_{\text{cage}}^{b,c}$	E_0 Prism [d] (/ cm^{-1})	$n_{\text{prism}}^{d,e}$
10 000	16 282 (15)	9	16 319	1				
20 000[f]	16 270 (12)	10		0	16 676 (92)	5	16 940 (14)	2
25 000	16 270 (9)	10		0				
30 000	16 266 (11)	10		0				
40 000	16 263 (12)	10		0				
50 000	16 261 (6)	9	16 313(5)[g]	4	16 508 (56)	4	16 698	1
100 000[f]	16 258 (4)	10		0	16 382 (25)	4	16 567 (3)	2
150 000					16 344 (37)	4	16 451 (65)	2

[a] The number of walkers.
[b] The ground state energy and uncertainty is based off of n_{cage} independent simulations.
[c] The number of independent simulations that collapsed to the cage structure.
[d] The ground state energy and uncertainty is based off of n_{prism} independent simulations.
[e] The number of independent simulations that collapsed to the prism structure.
[f] One of the unguided DMC simulations converged to a prism-book structure.
[g] Three additional calculations were initialized with only the prism structure.

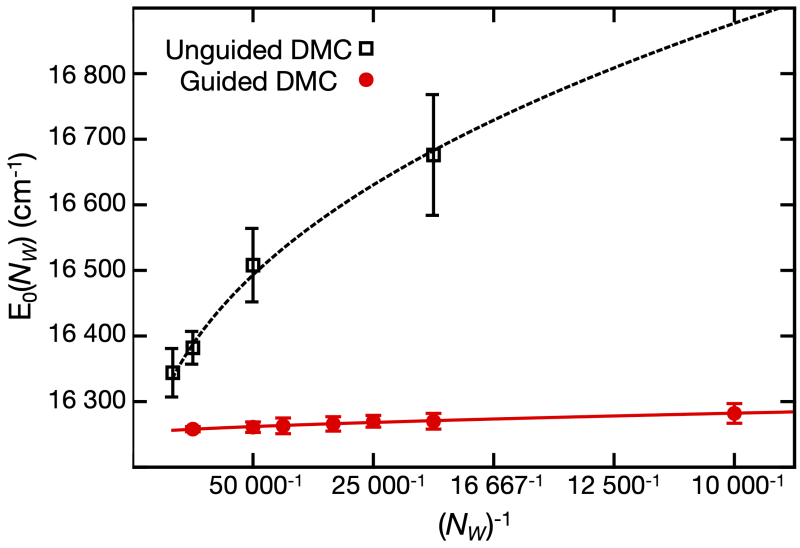

Figure 5.3: Convergence properties of the ground state energy calculations for **6H** simulations using an unguided approach ($\Psi_T = 1$) (open black squares), and with a guiding function described in the text (filled red circles) plotted as a function of $(N_W)^{-1}$, where N_W is the number of walkers. The energy difference between the guided DMC energies for 10 000 and 100 000 walkers is 24 cm^{-1}. The dashed black and solid red lines are shown to help guide the eye. The data used to generate this plot can be found in Table 5.3.

Table 5.3: Values of the Ground State Energies for **6H** Using Guided and Unguided DMC Plotted in Figure 5.3.

N_W^a	Guided DMC		Unguided DMC	
	E_0^b (/ cm^{-1})	$n_{\text{sim}}^{b,c}$	E_0 ^{b} (/ cm^{-1})	$n_{\text{sim}}^{b,c}$
10 000	16 282 (15)	9		
20 000	16 270 (12)	10	16 676 (92)	5
25 000	16 270 (9)	10		
30 000	16 266 (11)	10		
40 000	16 263 (12)	10		
50 000	16 261 (6)	9	16 508 (56)	4
100 000	16 258 (4)	10	16 382 (25)	4
150 000			16 344 (37)	4

[a] The number of walkers.
[b] The ground state energy and uncertainty is based off of n_{sim} independent simulations.
[c] The number of independent simulations.

5.3.2 Analysis of $|\Phi^2|$ for $(H_2O)_6$

In obtaining the energies for each ensemble size shown in Figure 5.3, we ran ten independent guided DMC simulations, and five independent unguided DMC simulations each of which started a set of smaller equilibrated ensembles in which the walkers were localized in different minima in the potential. As we considered the results of these simulations, we found that in many cases the resulting energies fell into two distinct sets, which differed in energy by at least 40 cm^{-1}. For the discussion above, we focused on the set of simulations that resulted in the lower energy, and the average energies and number of simulations in each group are provided in Table 5.3. For the discussion that follows, we focus on the ten guided simulations when 50 000 walkers were used. In this case nine of the calculations gave similar energies, 16 261(6) cm^{-1}, while the tenth calculation had an energy of 16 312 cm^{-1}.

A clue to the difference can be seen in the projections of the probability amplitude based on the wave functions obtained from each of these ten simulations onto the

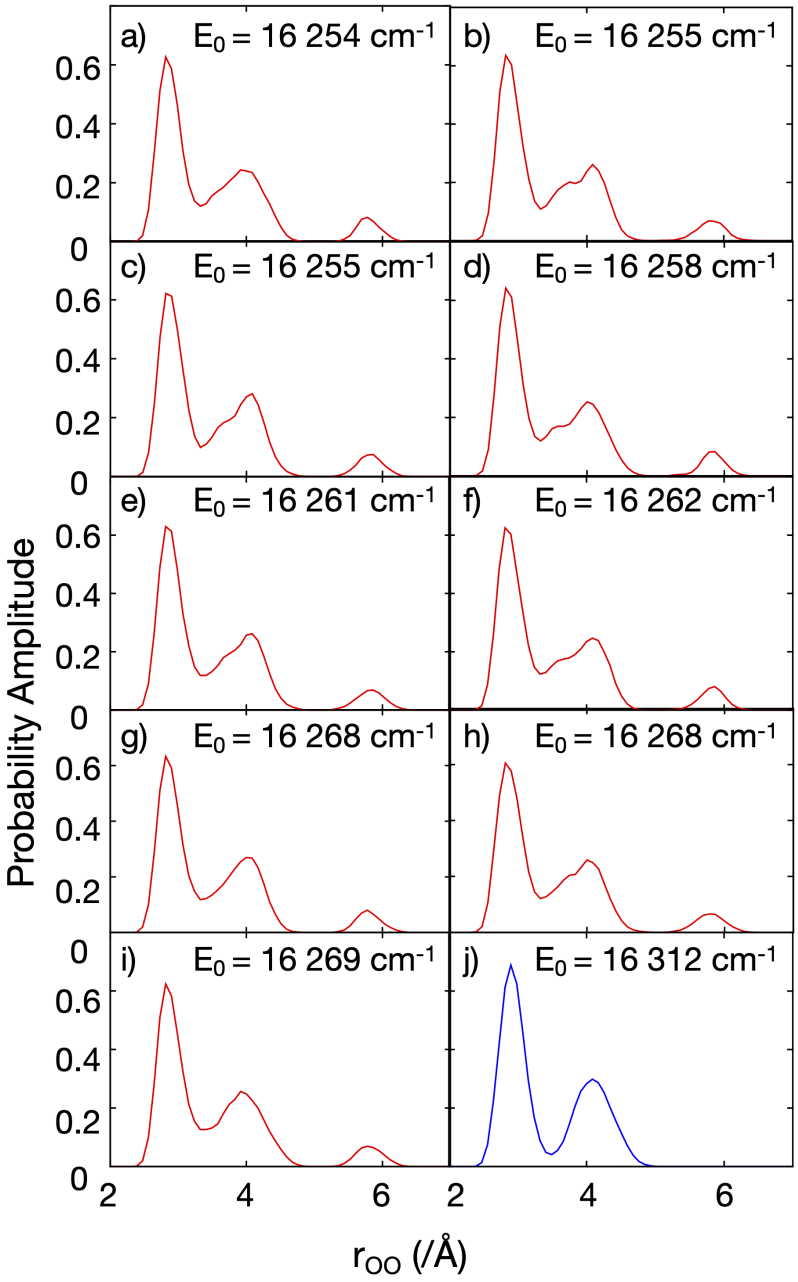

Figure 5.4: Projections of the ground state probability amplitude for all the possible OO distances in **6H** for each simulation. These projections were generated from simulations using 50 000 walkers where the Ψ_T used is described in the text. The ground state energy for each simulation is shown on the top right corner of each projection.

fifteen OO distances between the six oxygen atoms in **6H**. The results of this analysis is shown in Figure 5.4. In addition to the projections, we also report the calculated ground state energies for each simulation, and the results are reported in order of increasing ground state energy. Consistent with the above discussion, the distributions in panels a through i have energies that differ by only 15 cm^{-1}, and the distributions appear nearly identical. The tenth distribution, shown in Figure 5.4, has an energy that is 43 cm^{-1} larger than the one reported in panel i, and the distribution is clearly different. Notably the distributions shown in panels a through i all show three peaks, with the third peak being centered at slightly below 6 Å. The distribution in panel j has only two peaks. These two sets of distributions are consistent with the cage and prism structures, where the large OO distance reflects the distance between the oxygen atoms on the laft and right ends of the cage structure shown in Figure 5.1. While these results show a consistent story, and are consistent with the distribution of OO distances in **6H** described by Mallory *et al,*[25] such distributions are a fairly rough way to determine the structure, and do not preclude the possibility that the wave functions have amplitude in both minima in the potential.

To further explore the localization of the ground state probability amplitudes among these two minima, we have explored an approach for classification of individual walkers into minima on the potential surface. Mallory *et al* have used a similar metric to classify the structures from their calculations.[24,25] We use the ideas from Mallory *et al* that the OO distances provide a way to differentiate between the cage and prism structure. We compare all of the possible OO distances in **6H** with the minimum energy cage and prism structures. Specifically, for each walker, we calculate an array of fifteen OO distances, and permute the water molecules to minimize

$$\rho^{\alpha}(\mathbf{r}_{\mathrm{OO}}) = \sqrt{\sum_{j=1}^{n_{\mathrm{OO}}} \frac{(r_{\mathrm{OO},j} - r^{\alpha}_{\mathrm{OO},j})^2}{n_{\mathrm{OO}}}} \tag{5.1}$$

where $r^{\alpha}_{\mathrm{OO},j}$ represents the array of OO distances in the equilibrium structure for either α = cage or prism, and $n_{\mathrm{OO}} = 15$. In Figure 5.5, we plot the minimum value between ρ^{prism} and ρ^{cage} as a function of the difference between these two quantities for the distributions of walkers described above. As is seen in Figure 5.5, these distributions are localized at either positive (nine lower-energy simulations) or negative (higher energy simulation) values of

$$\Delta\rho(\mathbf{r}_{\mathrm{OO}}) = \rho^{\mathrm{prism}}(\mathbf{r}_{\mathrm{OO}}) - \rho^{\mathrm{cage}}(\mathbf{r}_{\mathrm{OO}}). \tag{5.2}$$

As is seen in these results, the sign of $\Delta\rho$ provides a way to differentiate between a prism or a cage structure and that the ensembles of walkers have clearly localized in a single minimum on the potential surface.

5.3.3 $(D_2O)_6$-Cage or Prism?

We next turn our attention to the fully deuterated cluster, **6D**. The calculated energies obtained using guided DMC simulations with ensemble sizes ranging from 10 000 to 100 000 walkers. Unlike **6H**, the energies appear to all fall in a relatively narrow energy window, raising the question of which isomer is being sampled. Based on the work of Pates and co-workers,[33] the prism should be the lower energy conformer at low-temperatures. On the other hand, previous studies of Mallory and Mandelshdam[25] showed that when ground state energy is considered the prism and cage are nearly degenerate.

To explore the structures that are sampled in the present DMC simulations, we again focus on the results based on 50 000 walkers, and project the probability amplitudes for the ten simulations onto $\Delta\rho$. The results are presented in Figure 5.6. The resulting distributions have been sorted by the sign of $\Delta\rho$ and reported in order of increasing ground state energy obtained from the simulation. In the left columns, $\Delta\rho$ is negative, and the simulation has collapsed into the minimum in energy that

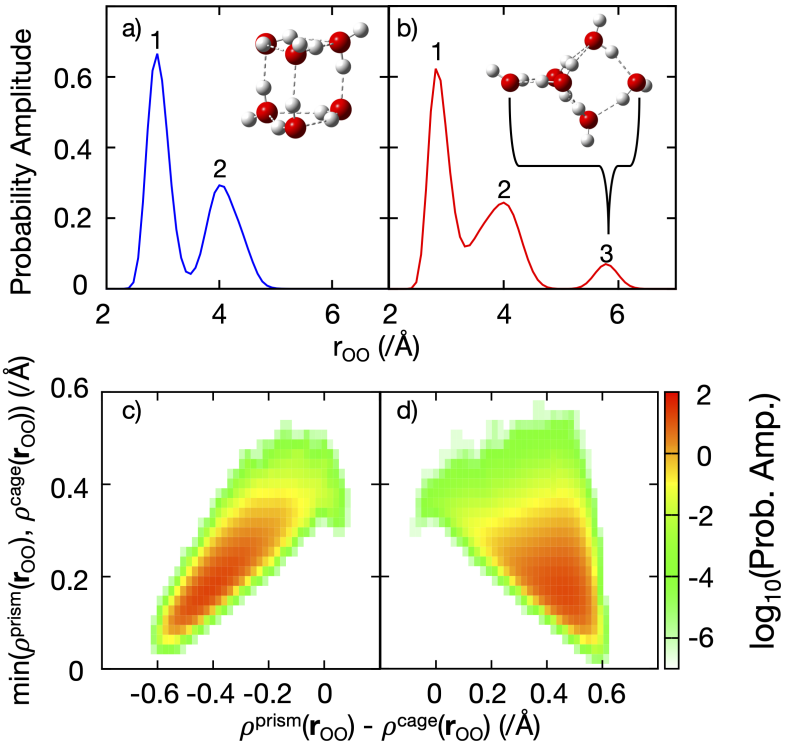

Figure 5.5: Projections of the ground state probability amplitude for **6H** onto (panels a and b) the OO distances and onto (panels c and d) $\rho^{\text{prism}}(\mathbf{r}_{\text{OO}}) - \rho^{\text{cage}}(\mathbf{r}_{\text{OO}})$ and the minimum value of $\rho^{\text{cage}}(\mathbf{r}_{\text{OO}})$ and $\rho^{\text{prism}}(\mathbf{r}_{\text{OO}})$, plotted on a $\log_{10}$ scale. These plots shown in panels a and c are based on a single DMC simulation of 50 000 walkers, which represents the ground state for the prism and panels b and d show the results of 9 DMC simulations of 50 000 walkers that correspond to the ground state of the cage.

corresponds to the prism. For these calculations, the ground state energy ranges from 7812 to 7838 cm^{-1}, while in the right column, $\Delta\rho$ is positive and the cage minimum is being sampled by the ensembles of walkers. In this case the ground state energies range from 7800 to 7821 cm^{-1}. These results show that the ranges of energies obtained for calculations of the ground state of the prism and cage overlap. Taken together, for this ensemble size, we find that the ground state energy of the cage is 7812 (7) cm^{-1} and the ground state energy of the prism is 7821 (9) cm^{-1}. Consistent with earlier studies,[25] we conclude that the two structures very close in energy, but based on the results of the ten simulation, we find that the cage to be slightly lower in energy than the prism. The small difference makes this result very sensitive to details of the potential surface and this 10 cm^{-1} difference is expected to be well within possible limitations of the electronic structure theory of fitting procedures used to obtain the potential surface. We also note, that the greater stability of the cage for **6D** is also found in RE-PIMD studies using the same potential surface for temperatures as low as 20 K.[104]

In order to understand the convergence properties of the energies for the cage and prism structures for **6D**, we calculated the ground state energy as a function of the ensemble size, for N_W ranging from 10 000 to 100 000 walkers. These results are presented in Figure 5.7, in which the ground state energies of the cage and prism structures are reported. Over this range of ensemble sizes, the ground state energies differ by less than 10 cm^{-1}, and in all cases, the ground state energy associated with the cage structure is is consistently smaller than the ground state energy of the prism.

At a first glance, we find the results for **6D** to be surprising. From prior experimental and theoretical findings, the prism has been reported to be the lowest energy structure for **6D**.[25,33] We find that in **6D**,the cage is lower in energy than the prism. However, as also shown in Figure 5.7, the error bars in the ground state energies for the simulations overlap one another. Mallory et. al have also calculated energies using DMC where they only allowed the walkers to sample only one minimum for

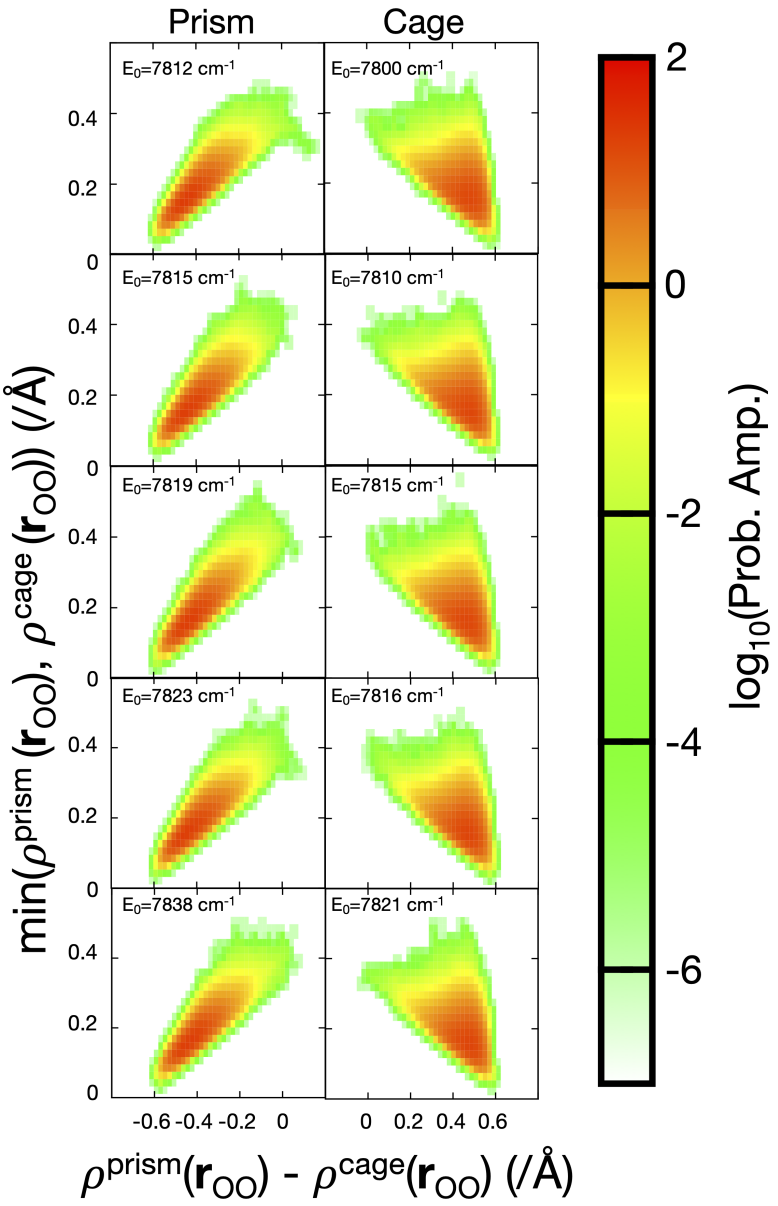

Figure 5.6: Projections of the ground state probability amplitude for **6D** onto $\rho^{\text{prism}}(\mathbf{r}_{\text{OO}}) - \rho^{\text{cage}}(\mathbf{r}_{\text{OO}})$ and the minimum value of $\rho^{\alpha}(\mathbf{r}_{\text{OO}})$, where α is either the minimum energy structure for the prism or cage, plotted on a $\log_{10}$ scale. The ground state energy of each simulation is shown on the top left corner of each projection.

Table 5.4: Values of the Ground State Energies for the Cage and Prism Structures for **6D** Plotted in Figure 5.7.

N_W^a	$E_0^{\text{cage}b}$ (/ cm^{-1})	${n_{\text{cage}}}^{b,c}$	$E_0^{\text{prism}d}$ (/ cm^{-1})	${n_{\text{prism}}}^{d,e}$
10 000	7835 (24)	6	7845 (13)	4
20 000	7828 (14)	7	7835 (13)	3
25 000	7826 (14)	8	7837 (1)	2
30 000	7819 (12)	7	7828 (4)	3
40 000	7820 (9)	7	7831 (10)	3
50 000	7812 (7)	5	7821 (9)	5
100 000	7818 (6)^f	9^f	7823	1

[a] The number of walkers.
[b] The ground state energy and uncertainty is based off of n_{cage} independent simulations.
[c] The number of independent simulations that collapsed to the cage structure.
[d] The ground state energy and uncertainty is based off of n_{prism} independent simulations.
[e] The number of independent simulations that collapsed to the prism structure.
[f] One simulation was averaged over the last 25 000 timesteps

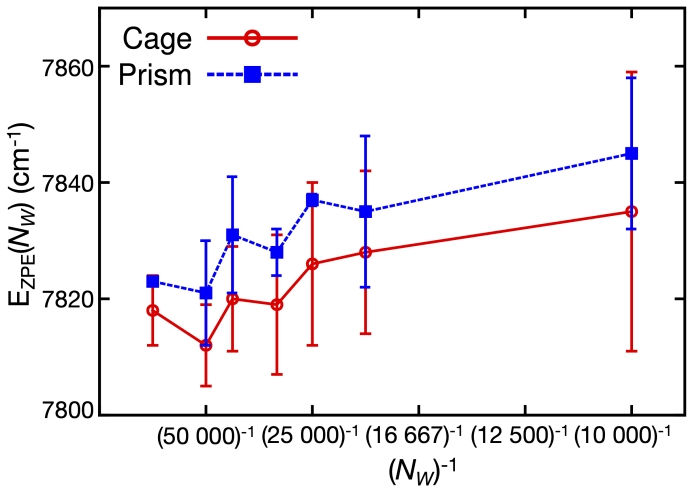

Figure 5.7: Comparison of the convergence of the ground state energy as a function of $(N_W)^{-1}$, where N_W are the number of walkers in the simulation for calculations that have collapsed into the cage isomer (solid lines, open red circles) or the prism isomer (dashed line, filled blue squares) for **6D**.

6H and **6D**. They find that while their calculated binding energy is lower for the prism in **6H**, the binding energy for **6D** for both the prism and cage is the same.[25] When the ground state energy is calculated from the reported binding energies, the cage is lower in ground state energy for **6H**, but both structures are reported to have identical ground state energies for **6D**. They have also reported the binding energies of various potential energy surfaces for both **6H** and **6D**. They found that for the deuterated analogues, the lowest energy structure changes depending on the choice of the potential energy surface.

5.3.4 Understanding the Nuclear Quantum Effects in the Cage Structure

Having shown that for the MB-pol potential energy surface cage structure is the lowest energy structure for **6H**, and is slightly lower in energy than the prism for **6D** although the energy of the minimum on the potential that corresponds to the prism is nearly 90 cm^{-1} lower in energy than the cage. The question naturally arises

as to what effects are at play that determine the lower energy structure and why it is altered by quantum effects including deuteration. The overall lower energy of the prism compared to the cage likely reflects the larger number of hydrogen bonds. The prism contains nine, while the cage contains only eight. On the other hand, the four member rings that characterize the cage structure are less strained compared to the three member rings in the prism structure. This difference can be seen when we compare the contribution of the OH stretches to the harmonic ground state energy for these two structures. Even though there is one more hydrogen bond in the prism structure, the average harmonic frequency of the OH stretches is lower in the cage. Additionally, as indicated by the arrows in Figure 5.1, the free OH bonds in the cage structure can undergo large-amplitude tunneling motions between pairs of minima that differ in energy by between 30 and 40 cm^{-1} (see Figure 5.2).

In Figure 5.8a), we project the probability amplitude onto these two large amplitude motions, which are defined by the displacement of these free OH bonds off of the plane that is defined by the oxygen atom contained in that water molecule and the oxygen atoms of the two water molecules that form hydrogen bonds to this water molecule, referred to as Z. The results are shown in Figure 5.8a) for **6H** and Figure 5.8b) for **6D**.

Table 5.5: Ratios of Cage Populations from Figures 5.8a) and 5.8b).

Quadrant	**6H**		**6D**
	Present Study	Ref. 25[a]	Present Study
ub[b]	0.31	0.330	0.43
uf[b]	0.27	0.288	0.20
db[b]	0.23	0.206	0.26
uf[b]	0.19	0.175	0.19

[a] A geometry minimization was performed to further classify the walkers in the calculation.
[b] Quadrant defined in Figures 5.8a) and 5.8b).

We first turn to the results for **6H**. In Figure 5.8, we find that the ground state

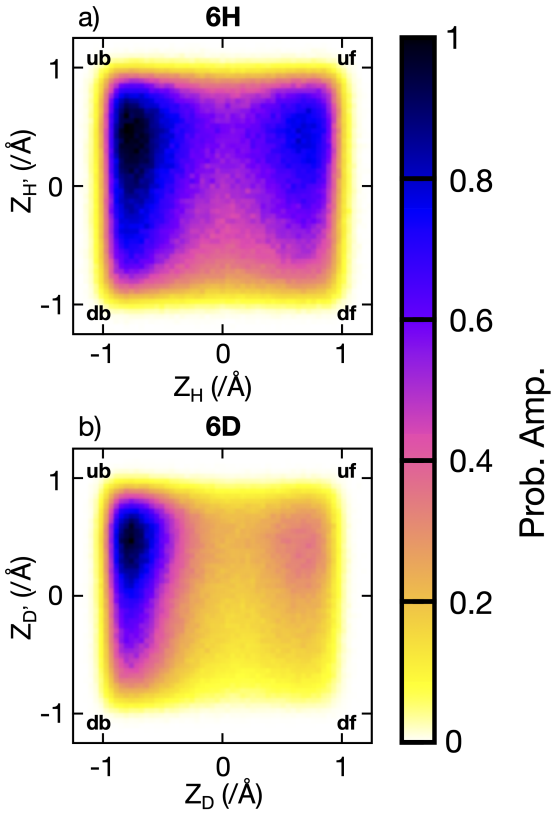

Figure 5.8: Projections of the ground state probability amplitude for the Z-component of the free a)OH or b) OD based off of a plane that contains 3 of the 6 neighboring oxygen atoms in the cage structure for a) **6H** and b) **6D**. The second Z-component of the other free a) OH or b) OD is based off of a plane defined by the oxygen atoms that were previously not used to define the latter free a) OH or b) OD. The labels on the corners correspond to the same labels that describes the same structures in Figure 5.2. The area of each quadrant shown in this figure can be found in Table 5.5.

probabilty amplitude is delocalized among all of the four minima, which are identified by the flipping motions seen for the cage structure in Figure 5.1. The largest amplitude is associated with the structure identified as "ub." This is the lowest energy structure. It is also the structure that is shown in Figure 5.1. We also find that the amplitude is delocalized vertically, into the region associated with the "db" form. Motion in this direction corresponds to motion of the free OH bond on the left side of the structure in Figure 5.1. Amplitude is also delocalized into the second set of minima. This delocalization is consistent with these vibrations being very low-frequency, and this motion plays an important role in lowering the ground state of the cage relative to the prism. We have also calculated the areas of each of the four minima by separating Figure 5.8 into four quadrants. These results are shown in Table 5.5. Mallory et. al have also reported fractions of the various cage structures after performing a energy minimization for all of the walkers in **6H**. We find reasonable agreement with the fractions of each structure reported for each minimum structure to the work of Mallory.[25] Comparing these results to those obtained for **6D**, we find that deuteration substantially diminishes the delocalization. This greater localization means that the frequency of this motion has not been decreased by the expected factor of $\sqrt{2}$, and the lowering of the ground state energy due to this large amplitude motion is not as significant in the deuterated version of the water hexamer. This is consistent with the smaller energy gap between the cage and prism forms of **6D** compared to **6H**.

5.3.5 Effects of Partial Deuteration

To further investigate the nuclear quantum effects in water hexamer, we perform studies of partial detueration, where we substitute a D_2O water in lieu of H_2O in the **6H** water, and we substitute a H_2O water in lieu of D_2O in the **6D** water. There have been many studies in partial deuteration of small water clusters, and they have concluded that placing the deuterium in the hydrogen-bonded positions lowers the overall ground state energy rather than placing the deuterium in the free

position.[85,102,105,106] While placing the deuterium in the free OD position would lower the free OD stretch by a factor of $\sqrt{2}$, placing the deuterium in the hydrogen-bonded position would play a larger role in lowering the ground state energy in the low frequency vibrations.

In the studies involving the various isopotomers of water hexamer, special care was taken to ensure we are able to accurately identify the various locations of the isotopic substitution. We use a graph theory approach to identify the hydrogen bonds within the cage or prism structures of water hexamer by describing the edge of a graph as a hydrogen bond. We describe the coordinates of the six water molecules and the unique molecule is placed last in the array of water molecules that describe the coordinates of the water hexamer.

For the cage structure, we first identify the water labeled as F in Figure 5.1. We follow the hydrogen bonded networks in the cage structure to classify all of the of the water molecules based on their hydrogen bonding network. Once we classify all of the waters based on their hydrogen bonds, we place a final check by checking that one of the hydrogen atoms in water D forms a hydrogen bond with water F. We perform a similar classification scheme for the prism structure, where we first identify the water labeled as 1 in Figure 5.1. Once we identify water 1 in the prism, we can classify the two other water molecules that help form the ring for the prism. With one of the 3-membered rings that has been properly classified, we find the hydrogen bond connecting water 3 and water 4, which is an ADD water. With water 4 identified, we can determine the positions of waters 5 and 6 based on their connections to water 4. Once we classify all of the waters in the hexamer, we can determine the location of the unique water in the hexamer, which is the last water in the array of coordinates that describe the geometry of the water hexamer. Using the classification scheme, we are able to distinguish each water in the hexamer and determine the location of the unique molecule.

In these calculations, we initialize our simulations by dividing up the ensemble

of walkers into all of the possible locations of the isotopic substitution and sample each minima for 2000 timesteps. After 2000 timesteps, the walkers are then combined from each minima and propagated for an additional 98 000 timesteps. These calculations can be found in Tables 5.6 and 5.7, where the cage and prism structures are considered. In these calculations, we find that over the course of the simulation, the walkers localize to a particular position where the isotopic substitution has taken place. As a result, we have also performed calculations where we only sample one minima to ensure that we have at least three independent calculations per isotopic substitution for each of **5H1D** and **1H5D** cage and prism calculations. We have combined those results for the cage in Table 5.8 and for the prism in Table 5.9. We have also plotted the difference in the ground state energies for all of the results of the isotopic substitution studies for the prism and cage in Figure 5.9. We determine where the isotopic substitution has occurred in water hexamer by classifying all of the hydrogen bond interactions of each water molecule in the hexamer, based on their donor-acceptor pairs. We start the calculation by equally sampling each location for isotopic substitution. However, when the simulation has equilibrated, we find that the isotopic substitution has localized within a given simulation for all of the walkers. In some of these calculations, there was only one simulation where the isotopic substitution has taken place in a particular water molecule. As a result, we performed additional calculations where we only perform an isotopic substitution in one of the six positions of water hexamer.

For the cage structure, we begin with **5H1D**, where we randomly substitute D_2O for H_2O in the hexamer, and these results are shown in Table 5.6. We focus on the results calculated using DMC based on ten calculations, there is population in positions B, C, and D. We find that in these calculations, eight out of the ten simulations have localized the D_2O in position C, whereas one simulation has localized to positions B and D. The ground state energies reported in Table 5.8 show that placing the D_2O in position C yields the most stable structure for the **5H1D** cage. We find that position

Table 5.6: Ground State Energies of the Cage Structure for **5H1D** and **1H5D** Water Hexamer Calculated Based on the Results of the Ten Calculations Described in the Text.

HOH/DOD Location[a]	**5H1D**		**1H5D**	
	DMC[b] ($/\mathrm{cm}^{-1}$)	n_{sim}^{c}	DMC[b] ($/\mathrm{cm}^{-1}$)	n_{sim}^{c}
A			9205(1)	2
B	14 853	1		
C	14 817(9)	8		
D	14 846	1	9220(6)[d]	3[d]
E			9219	1
F			9205(9)	4

[a] Refer to the labels in Figure 5.1.
[b] The ground state energy and uncertainty is based off of n_{sim} independent calculations.
[c] The number of simulations where the isotopically substituted water has localized.
[d] One of the simulations was averaged over the last 17 500 timesteps rather than over the last 40 000 timesteps.

Table 5.7: Ground State Energies of the Prism Structure for **5H1D** and **1H5D** Water Hexamer Calculated Based on the Results of the Ten Calculations Described in the Text.

HOH/DOD Location[a]	**5H1D**		**1H5D**	
	DMC[b] ($/\mathrm{cm}^{-1}$)	n_{sim}^{c}	DMC[b] ($/\mathrm{cm}^{-1}$)	n_{sim}^{c}
1	14 877(6)	3	9253	1
2	14 887(3)[d]	2[d]		
3	14 916	1	9232(4)	6
4	14 916	1	9249	1
5	14 924(24)	2	9226	1
6	14 909	1	9253	1

[a] Refer to the labels in Figure 5.1.
[b] The ground state energy and uncertainty is based off of n_{sim} independent calculations.
[c] The number of simulations where the isotopically substituted water has localized.
[d] One of the simulations was averaged over the last 25 000 timesteps rather than over the last 40 000 timesteps.

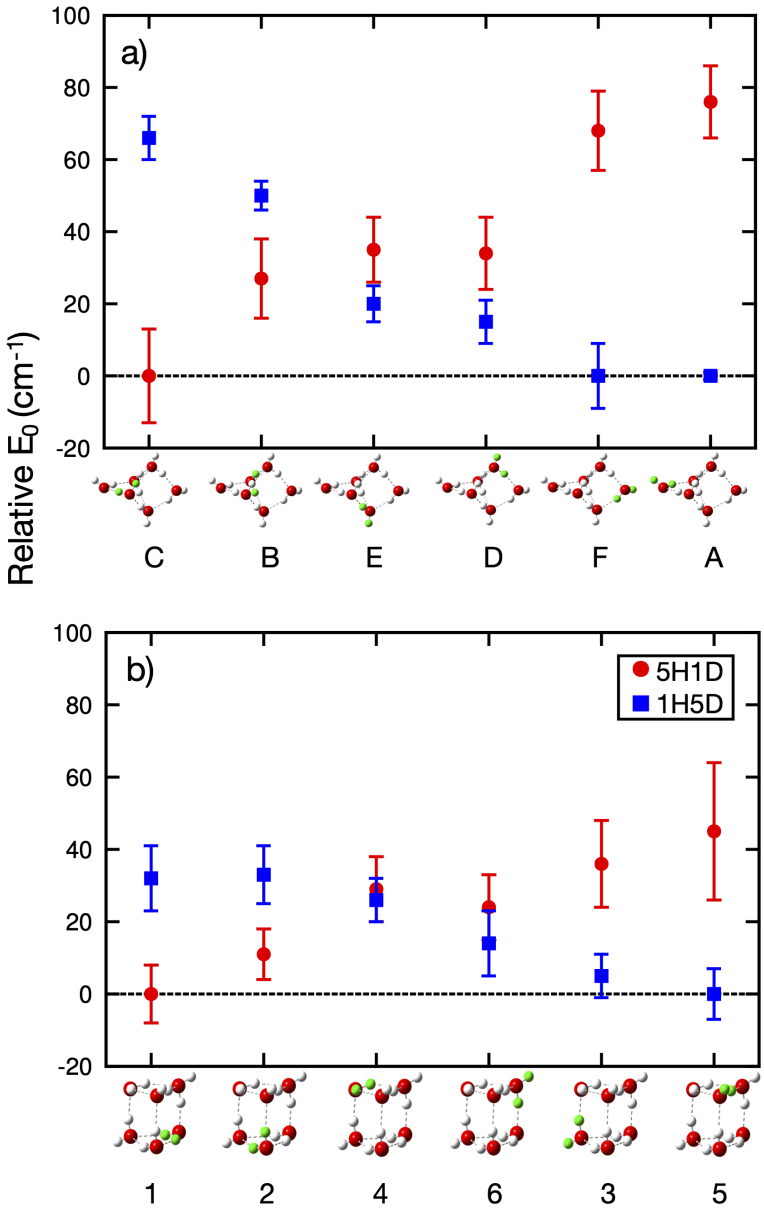

Figure 5.9: The relative difference in the ground state energies for the **5H1D** (red circles) and **1H5D** (blue squares) isotopomers of the a) cage, and b) prism structures. The location of the isotopic substitution is indicated with the green water in the figure. The ground state energies used to generate these plots are provided in Tables 5.8-5.9.

C is the lowest in energy as the water molecule in C is donating two hydrogen bonds, which is also classified as an ADD water. However, position B is also an ADD water, but within the analysis of the ground state energies, only one simulation landed in position B, and the ground state energy for position B is 27 cm^{-1} higher than the energies reported for C. We find that while both B and C are ADD waters, the water is position C is donating to an outer water, which is forming a stronger hydrogen bond. The water in position B is donating both to water molecules that accept two hydrogen bonds (AAD), which are forming weaker hydrogen bonds. We have also performed calculations for all of the possible isotopic substitutions and we find that placing the D_2O in the AAD waters is $\sim$ 35 cm^{-1} higher than placing it into the ADD water, and placing the the D_2O in the AD water is $\sim$ 70 cm^{-1} higher than the lowest ADD water. These results are consistent with our intuition that placing a D_2O the AAD or the AD water is higher in energy compared to placing the D_2O in the ADD water. Placing the D_2O in the ADD water has a larger effect on lowering the low frequencies than placing it in a position where there is a free OD stretch.

We turn to the case of the **1H5D** cage structure, with the results calculated in Table 5.8. We find that placing the H_2O in positions A and F, where the free OH in the water molecule is allowed to flip based on an OOO plane, differ by 1 cm^{-1}, which is the completely opposite trend of the case for the **5H1D** substitution. This is consistent with our initial intuition as placing the H_2O in positions where there can be a free OH lowers the overall ground state energy. We also find that the H_2O can also localize in positions D, and E, which are the AAD waters and the difference in energy between positions D and E is 5 cm^{-1}. We find that the difference in the ground state energies for sampling for the calculation to either localize in the AAD water or the outer water is $\sim$ 15 cm^{-1}. Finally, when considering the effecting of placing an H_2O in B, or C, the difference in energy is about $\sim$ 60 cm^{-1} as placing a H_2O in B, or C is energetically unfavorable due to raising the energies in the intermolecular vibrations.

With the energies of the cage calculations analyzed, we turn to the prism. Al-

Table 5.8: Ground State Energies of the Cage Structure for **5H1D** and **1H5D** Water Hexamer Calculated Using DMC

HOH/DOD Location[a]	**5H1D**		**1H5D**	
	DMC[b] ($/cm^{-1}$)	n^c_{sim}	DMC[b] ($/cm^{-1}$)	n^c_{sim}
A	14 893(4)	3	9205(1)	3
B	14 844(7)	6	9255(4)	3
C	14 817(9)	8	9271(6)	3
D	14 851(4)	3	9220(6)[d]	3[d]
E	14 852(2)	3	9225(5)	3
F	14 885(6)	3	9205(9)	4

[a] Refer to the labels in Figure 5.1.
[b] The ground state energy and uncertainty is based off of n_{sim} independent calculations.
[c] The number of simulations where the isotopically substituted water has localized.
[d] One of the simulations was averaged over the last 17 500 timesteps rather than over the last 40 000 timesteps.

though there is one more hydrogen bond in the prism, the structure forms two, three-membered rings, causing there to have more strain on the overall structure than the cage. As a result, although the electronic energy of the prism is lower than the cage, when ground state energies are taken into account, the cage is a much more stable structure. We use DMC to calculate the ground state energies of the **5H1D** prism structures. We repeat the same procedure that was used for the cage, and we present the results of the energies in Table 5.9 and the difference of the ground state energies are plotted in Figure 5.9b). We find similar trends that we have found in the cage for the prism. Placing the D_2O in the ADD positions (1,2, and 4) leads to a lower ground state energy than placing the D_2O in the AAD positions (3,5, and 6) which is consistent with the results shown for the **5H1D** cage. We also perform the same analysis for the **1H5D** prism. The energies calculated using DMC are shown in Table 5.9 and we find similar trends that is seen in the **1H5D** cage, as placing the H_2O in the AAD waters lower the ground state energy of the hexamer rather than in the

ADD waters.

Table 5.9: Ground State Energies of the Prism Structure for **5H1D** and **1H5D** Water Hexamer Calculated Using DMC.

HOH/DOD Location[a]	**5H1D**		**1H5D**	
	DMC[b] ($/\mathrm{cm}^{-1}$)	n_{sim}^{c}	DMC[b] ($/\mathrm{cm}^{-1}$)	n_{sim}^{c}
1	14 877(6)	3	9259(8)	3
2	14 888(3)[d]	3[d]	9260(6)	3
3	14 913(11)	3	9232(4)	6
4	14 906(7)	4	9253(4)	3
5	14 922(18)	5	9227(5)	5
6	14 901(7)	3	9241(7)	5

[a] Refer to the labels in Figure 5.1.
[b] The ground state energy and uncertainty is based off of n_{sim} independent calculations.
[c] The number of simulations where the isotopically substituted water has localized.
[d] One of the simulations was averaged over the last 25 000 timesteps rather than over the last 40 000 timesteps.

Based on the results for both the cage and prism, we are able to determine the ordering of the placement of the more stable structures based off of the trends on hydrogen bond donor and acceptor pairs. We are able to incorporate anharmonicity into the vibrational degrees of freedom, therefore lowering the calculated ground state energy. As a result, we are able to use DMC to make accurate predictions of the location where the isotopic substitution would take place based on ground state energies calculated for the cage and prism isomers of water hexamer.

5.4 Conclusions

In this chapter, we show by using a guided DMC approach, we can reduce the ensemble sizes needed in the calculations of water hexamer. Reducing the ensemble sizes for these calculations opens the door to perform detailed studies on the nuclear quantum effects in water hexamer as well as larger water clusters. We find that the result

of each simulation is dependent on the initial conditions of the calculation due to the multiple low-lying minima on the potential and the large barriers between each minima. We are able to confirm that for **6H**, the lowest energy structure is the cage which has been shown by both theory and experiment, but in **6D**, we find that the lowest energy structure is also the cage using the MB-pol potential energy surface, contrary to experimental findings.[25,32,33] We suspect that there are some sensitive details to the potential energy surface, as the difference in the ground state energy between both structures is small (9 cm^{-1} for **6D** compared to 51 cm^{-1} for **6H**) for 50 000 walkers. We have also performed the studies on isotopic substitution where we calculate the ground state energies of the isotopically substituted water hexamer. We show that when we substitute a D_2O for H_2O, the D_2O tends to be in the ADD position. When we perform the study for **1H5D**, the H_2O tends to be in a position where it makes only one hydrogen bond. In the cage structure, the H_2O tends to go in the outer water motion, as placing the H_2O in the outer water lowers the zero-point energy due to the flipping motion of the outer water.

Chapter 6
SUMMARY AND ONGOING WORK

In this thesis, we show the applications of using DMC with the implementation of a guiding function to calculate the ground state energies and ground state wave functions for several molecules. In Chapter 2, we present the theory and implementation of DMC for both the unguided and the guided approaches. In both formalisms, we define our basis as an ensemble of localized functions, which we refer to as walkers. We propagate the walkers over a period of time steps to obtain the ground state energy and ground state wave function. We show that with the guided DMC approach, rather than sampling just the DMC wave function, Φ, we sample a product of Φ and Ψ_T. With the guided DMC approach, the walkers are biased to sample regions where Ψ_T contains amplitude. We show that with the introduction of Ψ_T we are able to reduce the statistical fluctuations seen in the evaluation of the energies at each time step. This is due to the fact that in the guided DMC approach, we evaluate the local energy rather than the potential energy, which is less coordinate dependent based of the choice of Ψ_T. We show that with DMC, we are able to calculate the ground state energies of any system of interest, and the properties of the ground state wave function can be calculated using descendant weighting.

In Chapter 3, we use DMC to calculate overlap integrals between the wave function calculated using DMC and wave functions obtained from several methods. While calculating $\langle\Phi|\Phi\rangle$ can be easily evaluated using descendant weighting, we find that calculating the normalization of the overlap integral is challenging due to the need to evaluate of $\langle\Psi_T|\Psi_T\rangle$. We have shown that the results of evaluating $\langle\Psi_T|\Psi_T\rangle$ by using the descendant weights only once is numerically unstable. We find that averaging

multiple independent evaluations of the descendant weights is needed to calculate the overlap integral. We also find that using a guided DMC approach can drastically improve the convergence property of evaluating the overlap integral. We calculate the overlap of a wave function obtained using DMC with wave functions obtained using other approaches. We test our evaluation of the overlap integral with different model systems, ranging from one-dimensional model to molecular systems such as formaldehyde and deuterated analogues of H_3^+. Even with guiding functions as simple as products of harmonic oscillators, we can calculate accurate ground state energies and ground state wave functions for extremely anharmonic molecules such as the isotopologues of H_3^+.

Upon learning that simple guiding functions can be used to accurately calculate the ground state energy of a molecule, we develop a guiding function that can be used in the guided DMC approach to efficiently calculate the ground state energy and wave function of water clusters, ranging from 1-3 water molecules. We present these findings in Chapter 4. The guiding function that we have developed is based on the intramolecular vibrations of each water molecule (i.e. OH stretches, and HOH bends). We find that using a guided DMC approach with the Ψ_T described above dramatically reduces the ensemble size that are needed to obtain converged energies for $(H_2O)_n$, when $n = 1 - 3$. As a result of the savings obtained from the guided DMC approach, we performed a detailed study on the water trimer to understand the effects of isotopic substitution, primarily substituting hydrogen atoms for deuterium. We find that the deuterium atoms tend to be in the hydrogen-bound OH positions whereas the hydrogen atoms tend to be in the free OH position.

In Chapter 5, we use the Ψ_T defined in Chapter 4 for calculate the ground state energy and ground state wave functions for larger clusters, particularly, the water hexamer cage and prism structures. We find that with these calculations, the walkers in the simulation localize in a particular minimum based on the initial conditions used for the calculations. We present an algorithm to sort the different structures that are

sampled based on the OO distances in water hexamer and we use this classification to determine whether the minimum that the calculation has localized on. We find that the cage structure is lower in energy than the prism for both **6H** and **6D**. We also perform calculations for **5H1D** and **1H5D** for the cage and prism structures and we find that for **5H1D**, the D_2O will be in the ADD position, whereas for **1H5D**, the H_2O tends to be in the AAD position.

In summary, we show that DMC with a guiding function is a powerful tool that we can use to investigate the ground state wave function of a molecule. The first example is shown that we can use this approach to calculate overlap integrals. This approach for calculating overlap integrals has been extended to calculate the intensities to unravel features in the vibrational spectrum. We are also able to use DMC to reduce the ensemble sizes needed to calculate the ground state energies and ground state wave function for water clusters. Using this guiding function, we are able to calculate accurate ground state energies of water clusters, ranging from a monomer to a hexamer. This approach can be also extended to larger clusters to study the effects of isotopic substitution in larger clusters. The guiding function can also be generalized to defining the high frequency vibrations for any molecular system, with ongoing studies focused on protonated water clusters, and protonated methane. Using these new guiding functions will significantly reduce the ensemble sizes needed and allow such calculations to be feasible. Lastly, we have analyzed the effect of isotopic substitution for the water trimer and hexamer. The results from these isotopic substitution studies can help further understanding the nature of hydrogen bonding in water clusters.

BIBLIOGRAPHY

[1] Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular Vibrations*; Dover: New York, 1955.

[2] Scott, A. P.; Radom, L. *J. Chem. Phys.* **1996**, *100*, 16502–16513.

[3] Whitehead, R.; Handy, N. *J. Mol. Spectrosc.* **1975**, *55*, 356–373.

[4] Bowman, J. M. *J. Chem. Phys.* **1978**, *68*, 608–610.

[5] Willetts, A.; Handy, N. C.; Green, W. H.; Jayatilaka, D. *J. Phys. Chem.* **1990**, *94*, 5608–5616.

[6] Barone, V. *J. Chem. Phys.* **2005**, *122*, 014108.

[7] Bowman, J. M.; Carter, S.; Huang, X. *Int. Rev. Phys. Chem.* **2003**, *22*, 533–549.

[8] Christiansen, O. *Theor. Chem. Acc* **2006**, *116*, 106–123.

[9] Christiansen, O. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2942–2953.

[10] Bowman, J. M.; Carrington, T.; Meyer, H.-D. *Mol. Phys.* **2008**, *106*, 2145–2182.

[11] Roy, T. K.; Gerber, R. B. *Phys. Chem. Chem. Phys.* **2013**, *15*, 9468–9492.

[12] Bloino, J.; Baiardi, A.; Biczysko, M. *Int. J. Quantum Chem.* **2016**, *116*, 1543–1574.

[13] Born, M.; Oppenheimer, R. *Ann. Phys.* **1927**, *389*, 457–484.

[14] Podolsky, B. *Phys. Rev.* **1928**, *32*, 812.

[15] Nikitin, A. V.; Protasevich, A. E.; Rey, M.; Tyuterev, V. G. *J. Chem. Phys.* **2018**, *149*, 124305.

[16] Austin, B. M.; Zubarev, D. Y.; Lester, J., William A. *Chem. Rev.* **2012**, *112*, 263–288.

[17] McCoy, A. B. *Int. Rev. Phys. Chem.* **2006**, *25*, 77–107.

[18] Umrigar, C.; Nightingale, M.; Runge, K. *J. Chem. Phys.* **1993**, *99*, 2865–2890.

[19] Hammond, B. L.; Lester, W. A.; Reynolds, P. J. *Monte Carlo Methods in Ab Initio Quantum Chemistry*; World Scientific, 1994; Vol. 1.

[20] Anderson, J. B. *J. Chem. Phys.* **1975**, *63*, 1499–1503.

[21] Anderson, J. B. *J. Chem. Phys.* **1976**, *65*, 4121–4127.

[22] Suhm, M. A.; Watts, R. O. *Physics Reports* **1991**, *204*, 293 – 329.

[23] Lee, H.-S.; Herbert, J. M.; McCoy, A. B. *J. Chem. Phys.* **1999**, *110*, 5481–5484.

[24] Mallory, J. D.; Brown, S. E.; Mandelshtam, V. A. *J. Phys. Chem. A* **2015**, *119*, 6504–6515.

[25] Mallory, J. D.; Mandelshtam, V. A. *J. Chem. Phys.* **2016**, *145*, 064308.

[26] Acioli, P. H.; Xie, Z.; Braams, B. J.; Bowman, J. M. *J. Chem. Phys.* **2008**, *128*, 104318.

[27] Huang, X.; Cho, H. M.; Carter, S.; Ojamae, L.; Bowman, J. M.; Singer, S. J. *J. Phys. Chem. A* **2003**, *107*, 7142–51.

[28] Soper, A.; Benmore, C. *Physical review letters* **2008**, *101*, 065502.

[29] Liu, K.; Cruzan, J. D.; Saykally, R. J. *Science* **1996**, *271*, 929–933.

[30] Gregory, J. K.; Clary, D. C. *J. Phys. Chem. A* **1997**, *101*, 6813–6819.

[31] Rakshit, A.; Bandyopadhyay, P.; Heindel, J. P.; Xantheas, S. S. *The Journal of Chemical Physics* **2019**, *151*, 214307.

[32] Pérez, C.; Muckle, M. T.; Zaleski, D. P.; Seifert, N. A.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. *Science* **2012**, *336*, 897–901.

[33] Evangelisti, L.; Perez, C.; Lobsiger, S.; Seifert, N. A.; Zaleski, D. P.; Pate, B.; Kisiel, Z.; Temelso, B.; Shields, G. C. *Proceedings of the 69th International Symposium on Molecular Spectroscopy* International Symposium on Molecular Spectroscopy: Urbana–Champaign, IL, June 16–20, 2014.

[34] Metropolis, N.; Ulam, S. *J. Am. Stat. Assoc.* **1949**, *44*, 334–341.

[35] Feit, M. D.; Fleck, J. A. *J. Chem. Phys.* **1983**, 301–8.

[36] Petit, A. S.; Ford, J. E.; McCoy, A. B. *J. Phys. Chem. A* **2014**, *118*, 7206–7220.

[37] Lin, Z.; McCoy, A. B. *J. Phys. Chem. A* **2013**, *117*, 11725–11736.

[38] Petit, A. S.; McCoy, A. B. *J. Phys. Chem. A* **2013**, *117*, 7009–7018.

[39] Reynolds, P. J.; Ceperley, D. M.; Alder, B. J.; Lester, W. A. *J. Chem. Phys.* **1982**, *77*, 5593–5603.

[40] Barnett, R.; Reynolds, P.; W.A Lester, J. *J. Comput. Phys.* **1991**, *96*, 258 – 276.

[41] Partridge, H.; Schwenke, D. W. *J. Chem. Phys.* **1997**, *106*, 4618–4639.

[42] Colbert, D. T.; Miller, W. H. *J. Chem. Phys.* **1992**, *96*, 1982–1991.

[43] Liu, K.; Kalos, M.; Chester, G. *Phys. Rev. A* **1974**, *10*, 303.

[44] Reynolds, P.; Barnett, R.; Hammond, B.; Grimes, R.; Lester Jr, W. *Int. J. Quantum Chem.* **1986**, *29*, 589–596.

[45] Langfelder, P.; Rothstein, S. M.; Vrbik, J. *J. Chem. Phys.* **1997**, *107*, 8526–8535.

[46] Warren, G. L.; Hinde, R. J. *Phys. Rev. E* **2006**, *73*, 056706.

[47] Sandler, P.; Buch, V.; Clary, D. C. *J. Chem. Phys.* **1994**, *101*, 6353–5.

[48] Hornik, M.; Snajdr, M.; Rothstein, S. M. *J. Chem. Phys.* **2000**, *113*, 3496–3498.

[49] Barnett, R.; Reynolds, P.; Lester Jr, W. *J. Chem. Phys.* **1992**, *96*, 2141–2154.

[50] Barnett, R.; Johnson, E.; Lester Jr, W. *Phys. Rev. A* **1995**, *51*, 2049.

[51] Bulik, I. W.; Frisch, M. J.; Vaccaro, P. H. *J. Chem. Phys.* **2017**, *147*, 044110.

[52] Bulik, I. W.; Frisch, M. J.; Vaccaro, P. H. *J. Chem. Theory Comput.* **2018**, *14*, 1554–1563.

[53] Yachmenev, A.; Yurchenko, S. N.; Jensen, P.; Thiel, W. *J. Chem. Phys.* **2011**, *134*, 244307.

[54] Aguado, A.; Roncero, O.; Tablero, C.; Sanz, C.; Paniagua, M. *J. Chem. Phys.* **2000**, *112*, 1240–1254.

[55] Eckart, C. *Phys. Rev.* **1935**, *47*, 552–558.

[56] Blume, D. *Phys. Rev. A* **2002**, *66*, 053613.

[57] Umrigar, C.; Wilson, K.; Wilkins, J. *Phys. Rev. Lett.* **1988**, *60*, 1719.

[58] Bernu, B.; Ceperley, D.; Lester Jr, W. *J. Chem. Phys.* **1990**, *93*, 552–561.

[59] Bernu, B.; Ceperley, D.; Lester Jr, W. *J. Chem. Phys.* **1991**, *95*, 7782–7782.

[60] Miller, S.; Tennyson, J. *J. Mol. Spectrosc.* **1987**, *126*, 183–192.

[61] Sibert, E. L.; Hynes, J. T.; Reinhardt, W. P. *J. Phys. Chem.* **1983**, *87*, 2032–2037.

[62] McCoy, A. B. *J. Phys. Chem. B* **2014**, *118*, 8286–8294.

[63] Jastrow, R. *Phys. Rev.* **1955**, *98*, 1479.

[64] Drummond, N.; Towler, M.; Needs, R. *Phys. Rev. B* **2004**, *70*, 235119.

[65] Bouabça, T.; Braïda, B.; Caffarel, M. *J. Chem. Phys.* **2010**, *133*, 044111.

[66] Toulouse, J.; Umrigar, C. J. *J. Chem. Phys.* **2007**, *126*, 084102.

[67] Buch, V. *J. Chem. Phys.* **1992**, *97*, 726–729.

[68] Johnson, L. M.; McCoy, A. B. *J. Phys. Chem. A* **2006**, *110*, 8213–8220.

[69] Fore, M. E.; McCoy, A. B. *J. Phys. Chem. A* **2019**, *123*, 4623–4631.

[70] Mella, M.; Clary, D. C. *J. Chem. Phys.* **2003**, *119*, 10048–10062.

[71] McCunn, L. R.; Roscioli, J. R.; Johnson, M. A.; McCoy, A. B. *J. Phys. Chem. B* **2008**, *112*, 321–327.

[72] McCoy, A. B.; Dzugan, L. C.; DiRisio, R. J.; Madison, L. R. *Faraday Discuss.* **2018**, 443–466.

[73] Gregory, J. K.; Clary, D. C. *J. Chem. Phys.* **1995**, *102*, 7817–7829.

[74] Severson, M. W.; Buch, V. *J. Chem. Phys.* **1999**, *111*, 10866–10875.

[75] Mallory, J. D.; Mandelshtam, V. A. *J. Chem. Phys.* **2015**, *143*, 144303.

[76] Babin, V.; Medders, G. R.; Paesani, F. *J. Phys. Chem. Lett.* **2012**, *3*, 3765–3769.

[77] Babin, V.; Paesani, F. *Chem. Phys. Lett.* **2013**, *580*, 1 – 8.

[78] Babin, V.; Medders, G. R.; Paesani, F. *J. Chem. Theory. Comput.* **2014**, *10*, 1599–1607.

[79] Paesani, F. *Acc. Chem. Res.* **2016**, *49*, 1844–1851.

[80] Lee, V. G. M.; Madison, L. R.; McCoy, A. B. *J. Phys. Chem. A* **2019**, *123*, 4370–4378.

[81] Rocher-Casterline, B. E.; Ch'ng, L. C.; Mollner, A. K.; Reisler, H. *J. Chem. Phys.* **2011**, *134*, 211101.

[82] Ch'ng, L. C.; Samanta, A. K.; Czakó, G.; Bowman, J. M.; Reisler, H. *J. Am. Chem. Soc.* **2012**, *134*, 15430–15435.

[83] Ch'ng, L. C.; Samanta, A. K.; Wang, Y.; Bowman, J. M.; Reisler, H. *J. Phys. Chem. A* **2013**, *117*, 7207–7216.

[84] Miller, C. E.; Yung, Y. L. *J. Geophys. Res.* **2000**, *105*, 29039–29051.

[85] Sorenson, J. M.; Gregory, J. K.; Clary, D. C. *Chem. Phys. Lett.* **1996**, *263*, 680–686.

[86] Karyakin, E. N.; Fraser, G. T.; Lovas, F. J.; Suenram, R. D.; Fujitake, M. *J. Chem. Phys.* **1995**, *102*, 1114–1121.

[87] Keutsch, F. N.; Cruzan, J. D.; Saykally, R. J. *Chem. Rev.* **2003**, *103*, 2533–2578.

[88] Keutsch, F. N.; Saykally, R. J. *Proc. Natl. Acad. Sci. U.S.A* **2001**, *98*, 10533–10540.

[89] Liu, K.; Cruzan, J.; Saykally, R. *Science* **1996**, *271*, 929–933.

[90] Goldman, N.; Leforestier, C.; Saykally, R. *Philos. Trans. R. Soc. A* **2005**, *363*, 493–508.

[91] Gregory, J. K.; Clary, D. C. *Mol. Phys.* **1996**, *88*, 33–52.

[92] Richardson, J. O.; Pérez, C.; Lobsiger, S.; Reid, A. A.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Wales, D. J.; Pate, B. H.; Althorpe, S. C. *Science* **2016**, *351*, 1310–1313.

[93] Temelso, B.; Archer, K. A.; Shields, G. C. *J. Phys. Chem. A* **2011**, *115*, 12034–12046.

[94] Temelso, B.; Shields, G. C. *J. Chem. Theory Comput.* **2011**, *7*, 2804–2817.

[95] Diken, E. G.; Robertson, W. H.; Johnson, M. A. *J. Phys. Chem. A* **2004**, *108*, 64–68.

[96] Tainter, C.; Ni, Y.; Shi, L. a.; Skinner, J. *J. Phys. Chem. Lett.* **2013**, *4*, 12–17.

[97] Bakker, H.; Skinner, J. *Chem. Rev.* **2010**, *110*, 1498–1517.

[98] Ohno, K.; Okimura, M.; Akai, N.; Katsumoto, Y. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3005–3014.

[99] Yang, N.; Duong, C. H.; Kelleher, P. J.; McCoy, A. B.; Johnson, M. A. *Science* **2019**, *364*, 275–278.

[100] Bates, D. M.; Tschumper, G. S. *J. Phys. Chem. A* **2009**, *113*, 3555–3559.

[101] Kananenka, A. A.; Skinner, J. *J. Chem. Phys.* **2018**, *148*, 244107.

[102] Lee, V. G. M.; McCoy, A. B. *J. Phys. Chem. A* **2019**, *123*, 8063–8070.

[103] Babin, V.; Leforestier, C.; Paesani, F. *J. Chem. Theory. Comput.* **2013**, *9*, 5395–5403.

[104] Brown, S. E.; Götz, A. W.; Cheng, X.; Steele, R. P.; Mandelshtam, V. A.; Paesani, F. *J. Am. Chem. Soc.* **2017**, *139*, 7082–7088.

[105] Odutola, J.; Dyke, T. *J. Chem. Phys.* **1980**, *72*, 5062–5070.

[106] Sabo, D.; Bačić, Z.; Graf, S.; Leutwyler, S. *Chem. Phys. Lett.* **1996**, *261*, 318–328.