Developing a Method to Study Ground State Properties of Hydrogen Clusters

by

Matthew D.G. Schmidt

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Chemistry

Waterloo, Ontario, Canada, 2014

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis contains work that has been published or submitted for publication. These works have been acknowledged in footnotes in the body of this thesis. My contributions, relating to this thesis, are summarized below.

Co-authored Manuscript: S. Constable, M. Schmidt, C. Ing, T. Zeng, and P.-N. Roy, "Langevin Equation Path Integral Ground State", J. Phys. Chem. A, 2013, 117 (32), pp 74617467. Copyright 2013 American Chemical Society.

The material used in this thesis relating to this manuscript is solely my own.

First author Manuscript: M. Schmidt, S. Constable, C. Ing, and P.-N. Roy, "Inclusion of trial functions in the Langevin equation Path Integral Ground State method: application to parahydrogen clusters and their isotopologues"

The material used in this thesis relating to this manuscript is solely my own.

Co-authored Manuscript: N. Faruk, M. Schmidt, H. Li, R. J. Le Roy, and P.-N. Roy, "First Principle Prediction of the Raman shifts of parahydrogen clusters"

Given the ground state interaction potential for hydrogen, I perform simulations of hydrogen clusters four molecules and larger to obtain pair distribution functions, which are used to calculate vibrational Raman shifts. My work does not include construction of

the interaction potentials, nor the difference potential which are also required to calculate vibrational Raman shifts.

Abstract

This thesis presents the benchmarking and development of a method to study ground state properties of hydrogen clusters using molecular dynamics. Benchmark studies are performed on our Path Integral Molecular Dynamics code using the Langevin equation for finite temperature studies and our Langevin equation Path Integral Ground State code to study systems in the zero-temperature limit when all particles occupy their nuclear ground state. A simulation is run on the first 'real' system using this method, a parahydrogen molecule interacting with a fixed water molecule using a trivial unity trial wavefunction. We further develop a systematic method of optimizing the necessary parameters required for our ground state simulations and introduce more complex trial wavefunctions to study parahydrogen clusters and their isotopologues orthodeuterium and paratritium. The effect of energy convergence with parameters is observed using the trivial unity trial wavefunction, a Jastrow-type wavefunction that represents a liquid-like system, and a normal mode wavefunction that represents a solid-like system. Using a unity wavefunction gives slower energy convergence and is inefficient compared to the other two. Using the Lindemann criterion, the normal mode wavefunction acting on floppy systems introduces an ergodicity problem in our simulation, while the Jastrow does not. However, even for the most solid-like clusters, the Jastrow and the normal mode wavefunctions are equally efficient, therefore we choose the Jastrow trial wavefunction to look at properties of a range of cluster

sizes. The energetic and structural properties obtained for parahydrogen and orthodeuterium clusters are consistent with previous studies, but to our knowledge, we may be the first to predict these properties for neutral paratritium clusters. The results of our ground state simulations of parahydrogen clusters, namely the distribution of pair distances, are used to calculate Raman vibrational shifts and compare to experiment. We investigate the accuracy of four interaction potentials over a range of cluster sizes and determine that, for the most part, the *ab initio* derived interaction potentials predict shifts more accurately than the empirically based potentials for cluster sizes smaller than the first solvation shell and the trend is reversed as the cluster size increases. This work can serve as a guide to simulate any system in the nuclear ground state using any trial wavefunction, in addition to providing several applications in using this ground state method.

Acknowledgements

I would like to acknowledge my supervisor, Dr. Pierre-Nicholas Roy, for his countless ideas that allowed me to achieve more out of my research. His energy and enthusiasm is contagious and makes it very enjoyable to come to work and learn each day and his gentle persistance allows me to maximize production. He is extraordinarily busy with meetings, conferences, and balancing his time with a growing number of group members, but he is always available when any of us need him and for that I am extremely grateful. He has helped me on numerous scholarship applications, proposals, presentations, and papers, regardless of how last-minute it is, and without his help I would not be where I am today.

I also owe acknowledgement to a number of past graduate students, namely Chris Ing and Steve Constable who took me under their wing when I was still in undergraduate studies. In addition to implementing the path integral Langevin equation thermostat into the molecular dynamics package we use, Chris walked me through all of the codes we run and made what could have been a very intimidating project very manageable. Steve incorporated the path integral ground state method in the Langevin equation molecular dynamics framework and his help in understanding the theory behind optimizing the simulation parameters was critical in me formulating this thesis. He also gave me the opportunity to collaborate with him on his paper, which really inspired me to follow up on his research. Tao Zeng, a past post doctoral student in the Roy group, is a true colleague and friend. I

had the privilege of attending many conferences with him and he was always helping me improve the quality of my posters/presentations and was able to answer any questions I had. His depth of knowledge over various subjects in theoretical chemistry astounded me. Nabil Faruk is a group member who joined about the same time I did and I have enjoyed the conversations we have had on issues with our path integral methods. I am glad we were able to collaborate on a paper, which is a critical application to the ground state method I have helped develop. Finally, I acknowledge the rest of the Roy group, namely Kevin Bishop, Dmitri Iouchtchenko, Spencer Yim, Lindsay Orr, and Lecheng Wang for the great research atmosphere and I hope we are able to collaborate in the near future!

Dedication

To my family, who have pushed me to be the best that I can be.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

PIMC	Path Integral Monte Carlo
DMC	Diffusion Monte Carlo
PIGS	Path Integral Ground State
VMD	Variational Molecular Dynamics
PIMD	Path Integral Molecular Dynamics
LePIGS	Langevin equation Path Integral Ground State
MMTK	Molecular Modelling Toolkit
PILE	Path Integral Langevin Equation
DCT	Discrete Cosine Transform
RPMD	Ring Polymer Molecular Dynamics
pH_2	parahydrogen
oD_2	orthodeuterium
pT_2	paratritium
RMD	Reptile Molecular Dynamics

Chapter 1

Introduction

1.1 Hydrogen Clusters

The hydrogen molecule has a mass of approximately 2.0 amu and consists of only two protons and two electrons[1]. Due to the coupling of the spins on the individual protons, hydrogen has two nuclear spin states. The triplet nuclear spin state is referred to as orthohydrogen and the singlet nuclear spin state is referred to as parahydrogen. Hydrogen has a melting point of 14 K and a boiling point of 20 K, both at very low temperatures. There are three main reasons why there is interest in studying hydrogen molecules. First, it has a simple structure, which enables computational chemists to perform very accurate calculations in a reasonable time frame. Also, the similar nuclear makeup of parahydrogen

to helium has suggested that parahydrogen may exhibit superfluidity[2]. For this reason, parahydrogen clusters have been the subject of important research both in theory[3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13] and in experiment[14, 15, 16, 17, 18]. Summaries of these studies have also been presented in reviews by Navarro and Guardiola[19] and Tao Zeng[20].

1.2 Recent Research on Parahydrogen Clusters in the Ground State

In the zero-temperature limit, where we expect parahydrogen clusters to exhibit full superfluidity, ground state studies are primarily performed using Path Integral Monte Carlo (PIMC) sampling[21] whether it be Diffusion Monte Carlo (DMC)[22] or Path Integral Ground State (PIGS)[23, 24]. Some groups have studied parahydrogen clusters using DMC[25, 26, 27], however it has been shown in a recent paper by Boninsegni and Moroni[28] that DMC has an inherent population size bias based on the choice of trial wavefunction that is used, which PIGS does not suffer from in principle. With PIGS, it can be shown that exact ground state properties can be obtained by using any trial wavefunction that overlaps with the exact ground state wavefunction, however these properties can be obtained most efficiently when the trial wavefunction has the greatest overlap. This PIGS method is also referred to as a variational path integral method[29, 30].

For several years, Cuervo et al used PIGS with Monte Carlo sampling to simulate ground state properties of helium[31] and then focused on small parahydrogen clusters the following year[32]. The group subsequently applied PIGS to pure orthodeuterium clusters[33] as well as parahydrogen clusters around an orthodeuterium trimer[34].

In 2009, Miura[35] developed a molecular dynamics analog to variational Monte Carlo called Variational Molecular Dynamics (VMD), using Nosé-Hoover chains[36] as the thermostat, which was shown to reproduce experimental results for liquid helium clusters in the ground state.

Our group has recently developed a different molecular dynamics variant to the PIGS method, using the Langevin equation thermostat, called LePIGS[37]. We will use this method in this thesis to gain further understanding of parahydrogen clusters and heavier bosonic isotopes and the results will be compared to previous research.

1.3 Thesis Outline

This thesis consists of several parts. The theory involved with our work is described in Chapter 2. Benchmark calculations are performed in Chapter 3 to test the validity of our new Path Integral Molecular Dynamics (PIMD) code which uses a recently upgraded version of the Molecular Modelling Toolkit (MMTK)[38] with an implementation of the

Path Integral Langevin Equation (PILE) thermostat[39]. The benchmark is the more traditionally used Monte Carlo simulations as well as a basis set calculation to provide 'exact' energy and structural distributions. This will be achieved using a simple system of one water and one parahydrogen. We further look at the low temperature properties of parahydrogen clusters with a non-rotating water dopant. The results are compared to PIMC simulations as before. In Chapter 4, we extend our focus to the zero temperature limit as we benchmark a new method, LePIGS. We compare ground state energies and other structural properties for a system of one parahydrogen interacting with a fixed water molecule and the structural and energetic properties are compared to the results of a basis set calculation. In Chapter 5, we develop a systematic method to run a LePIGS simulation and compare the effects on energetic and structural property convergence using different trial wavefunctions for systems of parahydrogen clusters and their isotopologues. In Chapter 6 we apply the results of these simulations to calculate Raman vibrational shifts and compare to experiment. Concluding remarks and future applications of this work are discussed in Chapter 7.

Chapter 2

Theory

2.1 Statistical Mechanics and Path Integral Formulation

Statistical mechanics is considered to be the science of determining macroscopic bulk thermodynamic properties from microscopic atomic/molecular properties by applying statistics. One reason we use statistical mechanics is to identify the accessible quantum states or energy levels of a particular system. Conceptually, many copies of a system are constructed and these copies, called an ensemble, are left to equilibrate over time. The fundamental postulate of statistical mechanics states that under equilibrium conditions, each copy can

occupy any one of its accessible states with equal probability. All of the information about this ensemble is stored in the partition function, which is a sum over all states. In the canonical ensemble, a system is held under the conditions of constant number of particles (N), volume (V), and temperature(T) over time. The partition function can be expressed as a trace[40]

$$Z = Tr(e^{-\beta \hat{H}}) \tag{2.1}$$

and β is defined as

$$\beta = \frac{1}{k_B T} \ , \tag{2.2}$$

where k_B is the Boltzmann constant and T is the temperature. The Hamiltonian is the sum of the kinetic and potential energy operators

$$\hat{H} = \hat{K} + \hat{V} \ . \tag{2.3}$$

If the coordinate representation of the partition function is used, the trace can be evaluated using an integral

$$Z = \int dR \, \langle R | \, e^{-\beta(\hat{K}+\hat{V})} \, | R \rangle \ . \tag{2.4}$$

In classical mechanics this expression can be solved directly, since it is assumed that the kinetic and potential energy operators commute, though this is not true in general for the

quantum case. This problem can be solved by separating the exponential into P discrete slices, subsequently called quantum beads, forming what will be known as the path integral. The limit as $P \to \infty$ will give the exact quantum mechanical definition. In order to solve the partition function in Equation 2.4, the following operator identity can be used [29]

$$e^{\tau(\hat{K}+\hat{V})+\tau^2\left[\hat{K},\hat{V}\right]} = e^{-\tau\hat{K}}e^{-\tau\hat{V}} \ , \tag{2.5}$$

where

$$\tau = \frac{\beta}{P} \tag{2.6}$$

and $[\hat{K},\hat{V}]$ is the commutator for kinetic and potential energy operators. At the low τ limit, the primitive approximation can be made

$$e^{-\tau(\hat{K}+\hat{V})} \approx e^{-\tau\hat{K}}e^{-\tau\hat{V}} \ . \tag{2.7}$$

Now, what is known as the Trotter factorization can be applied[41],

$$e^{\beta(\hat{K}+\hat{V})} = \lim_{P\to\infty}\left[e^{\tau\hat{K}}e^{\tau\hat{V}}\right]^P \ . \tag{2.8}$$

The partition function can now be written in a convenient form by substituting the variable Ω as the expression gained from a symmetrized version of the Trotter factorization[42]

$$\Omega = e^{-\frac{\tau}{2}\hat{V}} e^{-\tau\hat{K}} e^{-\frac{\tau}{2}\hat{V}} \tag{2.9}$$

$$Z = \int dR \, \langle R| \, \Omega^P \, |R\rangle \ . \tag{2.10}$$

To evaluate this partition function, it is convenient to split up the Ω^P term by inserting the resolution of identity operator $(P-1)$ times. The partition function can then be rewritten in the following cyclic manner

$$Z = \int dR_1 ... dR_P \, \langle R_1| \, \Omega \, |R_2\rangle \, ... \, \langle R_P| \, \Omega \, |R_1\rangle \ . \tag{2.11}$$

The potential operator in the position representation is easily evaluated as the potential energy at each R_i position

$$\langle R_i| \, \Omega \, |R_{i+1}\rangle = e^{-\frac{\tau}{2}V(R_i)} \, \langle R_i| \, e^{-\frac{\tau p^2}{2m}} \, |R_{i+1}\rangle \, e^{-\frac{\tau}{2}V(R_{i+1})} \ . \tag{2.12}$$

Inserting a complete set of momentum eigenstates allows the kinetic energy operator to be evaluated

$$\langle R_i| \, e^{-\frac{\tau p^2}{2m}} \, |R_{i+1}\rangle = \left(\frac{m}{2\pi\tau\hbar^2}\right)^{\frac{1}{2}} e^{-\frac{m}{2\tau\hbar^2}(R_{i+1}-R_i)^2} \ . \tag{2.13}$$

The parameter $\hbar$ is the reduced Planck's constant and m is the mass of the particle. Substituting this equation into the partition function results in the final expression

$$Z = \lim_{P\to\infty} \left(\frac{m}{2\pi\tau\hbar^2}\right)^{\frac{P}{2}} \int dR_1 ... dR_P \times \exp\left\{ -\sum_{i=1}^{P} \left[\frac{m}{2\tau\hbar^2} \left(R_{i+1} - R_i\right)^2 + \tau V\left(R_i\right) \right] \right\} . \tag{2.14}$$

There are many things to note about this equation. First, the partition function is the sum over all the different R values that the quantum beads occupy. Thus, the partition function is the path integral or the sum of all possible paths the particle can take. Second, the exponential contains two distinct parts: a potential energy term evaluated at each bead and also a term which resembles a harmonic oscillator between two neighbouring beads. This is analogous to the path being represented by quantum beads connected by harmonic springs. Also, with respect to the Hamiltonian, it is important to note that based on the primitive approximation made above that the energy will have an error of τ^2 associated with the commutation of the kinetic and potential energy. The final point, and one that is very important to path integral dynamics, is that the equation states that in the limit $P \to \infty$, the exact quantum definition of the partition function is satisfied. However, it can also be seen that in the limit of $P = 1$, the exact classical definition is satisfied.

2.2 Path Integral Ground State Formulation

When dealing with the zero-temperature limit, we require a different formulation[37]. To calculate the expectation value of a property $\hat{A}$ in the ground state, we require

$$\langle A\rangle_0 = \frac{\langle\Phi_0|\hat{A}|\Phi_0\rangle}{\langle\Phi_0|\Phi_0\rangle} = \frac{1}{Z_0}\langle\Phi_0|\hat{A}|\Phi_0\rangle \tag{2.15}$$

where Z_0 is some pseudo-partition function of the form

$$Z_0 = \int dR \int dR' \psi_T(R)\langle R|e^{-\beta\hat{H}}|R'\rangle\psi_T(R') \ . \tag{2.16}$$

The imaginary time propagator $\langle R|\exp(-\beta\hat{H})|R'\rangle$ can be factorized in the following way

$$\langle R_1|\exp(-\beta\hat{H})|R_P\rangle = \langle R_1|\exp(-\tau\hat{H})|R_2\rangle\langle R_2|\exp(-\tau\hat{H})|R_3\rangle ... \langle R_{P-1}|\exp(-\tau\hat{H})|R_P\rangle \tag{2.17}$$

$$= \prod_{j=1}^{P-1} \langle R_j|\hat{\rho}|R_{j+1}\rangle \ . \tag{2.18}$$

where $\tau = \beta/(P-1)$ and $\hat{\rho} = \exp(\text{-}\tau\hat{H})$ is the high temperature density operator.

The pseudo partition function Z_0 can now be written more compactly as:

$$Z_0 = \int dR_1 ... \int dR_P \psi_T(R_1) \prod_{j=1}^{P-1} \left[\langle R_j|\hat{\rho}|R_{j+1}\rangle\right] \psi_T(R_P) \ . \tag{2.19}$$

The Trotter factorization[41] taken to the second order as before, is now employed to obtain the following expression,

$$Z_0 = \lim_{P\to\infty} \sqrt{\frac{m\,(P-1)}{2\pi\beta\hbar^2}} \int dR_1 ... \int dR_P \exp\left\{-\tau V'(R)\right\} \tag{2.20}$$

where the effective potential V' is defined as

$$V'(R) = \sum_{j=1}^{P-1} \left[\frac{1}{2} m\omega_{P-1}^2 (R_j - R_{j+1})^2\right] + \sum_{j=1}^{P} \left[c_j V(R_j)\right] + V^{trial}\left(R_1\right) + V^{trial}\left(R_P\right) \ . \tag{2.21}$$

with $\omega_{P-1} = \frac{P-1}{\beta\hbar}$ and

$$c_j = \begin{cases} \frac{1}{2} & \text{if } j = 1 \text{ or } j = P \\ \\ 1 & \text{otherwise.} \end{cases} \tag{2.22}$$

Since the ground state wavefunction has no nodes, we can write the potential term contributed by the trial wavefunction as

$$V^{trial}(R) = -\frac{1}{\tau}\ln\left[\psi_T(R)\right] \ . \tag{2.23}$$

The form given in Eq. 2.20 looks very similar to that of the partition function for finite-temperature dynamics, although the path is not cyclical and the potential energy function includes information about the ground state wavefunction through the end beads.

Going back to our initial problem, for a position-dependant property A, the expectation value in the ground state $\langle A\rangle_0$ can be written as[37]

$$\langle A\rangle_0 = \lim_{\beta,P\to\infty} \frac{1}{Z_0}\int dR_1 ... \int dR_P A(R_{\frac{P+1}{2}}) \exp\left\{-\tau V'(R)\right\} \ , \tag{2.24}$$

representing the primitive estimator of $\hat{A}$.

It is interesting to note that the potential energy of the molecule and the structural distributions are obtained solely from the middle bead[31, 32]. This requires the number of beads in the simulation to be an odd number. Also, as long as the trial function has some overlap with the ground state, as $\tau\to 0$, any trial function will give the exact ground state results.

If an operator $\hat{A}$ commutes with $\exp(-\beta\hat{H})$, then $\hat{A}$ can be shifted towards the right hand side such that it acts on the trial wavefunction $|\psi_T\rangle$. The Hamiltonian is one such

operator and performing this operation leads to the mixed estimator for the energy:

$$E_0 = \left\langle \frac{\hat{H}\psi_T(R_P)}{\psi_T(R_P)} \right\rangle_0 \tag{2.25}$$

where E_0 is the ground state energy and $R(P)$ is the position of the end bead on the right. We can also move $\hat{H}$ to the left side so that it acts on $\langle\psi_T|$ since $\hat{H}$ is Hermitian. This means that although the potential energy and structural properties are obtained solely from the middle bead, the total energy can be obtained from both of the end beads, allowing us to double our sampling of the energy. With respect to error in the energy estimator, the convergence with τ is identical to that of the finite-temperature case since the same Trotter factorization is invoked. However, in the ground state formulation, the error in the energy also decays exponentially with β.

2.3 Langevin Equation Path Integral Ground State

It was observed above that the effective potential V' in the PIGS formulation reflects an open chain. However, the PILE was derived from the canonical Hamiltonian of PIMD that corresponds to a closed path[43]. We can re-write the Hamiltonian of the canonical ensemble in terms of the ground state effective potential V' that includes extra potential

terms that reflect the trial functions (from Eq. 2.21)[37]

$$H_P(p,q) = H_P^0(p,q) + V_P(q) \tag{2.26}$$

$$V_P(q) = \sum_{i=1}^{N} \left(\left[\frac{1}{P} \sum_{j=1}^{P} V'(q_i^{(j)}) \right] + V^{trial}(q_i^{(1)}) + V^{trial}(q_i^{(P)}) \right) \tag{2.27}$$

$$H_P^0(p,q) = \sum_{i=1}^{N} \sum_{j=1}^{P-1} \left(\frac{\left[p_i^{(j)} \right]^2}{2m_i} + \frac{1}{2} m_i \omega_{P-1}^2 \left[q_i^{(j+1)} - q_i^{(j)} \right]^2 \right) \tag{2.28}$$

where (p,q) are conjugate position and momentum states. We also have achieved an open chain since there is no circularity condition and there are $P-1$ kinetic spring terms as required. Also, note that an extra potential term is added to $V_P(q)$ due to the trial wavefunction $|\psi_T\rangle$ acting on the two end beads.

The PILE thermostat requires the Hamiltonian to be expressed in terms of normal mode coordinates. Since PIGS is isomorphic with an open path, we can write modified normal mode coordinates in terms of the Discrete Cosine Transform type II (DCT-II) matrix[37] instead of the Fourier matrix which we were able to use for the closed path PIMD[43]. The normal mode coordinates for PIGS are[37]:

$$\widetilde{p}_i^{(k)} = \sum_{j=1}^{P-1} C_{kj} p_i^{(j)} \ , \tag{2.29}$$

$$\widetilde{q}_i^{(k)} = \sum_{j=1}^{P-1} C_{kj} q_i^{(j)} \ , \tag{2.30}$$

where the coefficients C_{kj} for normal mode k are defined as:

$$C_{kj} = \begin{cases} \sqrt{\frac{1}{P}} \cos(\pi k(j-1/2)/P) & k = 0 \\ \sqrt{\frac{2}{P}} \cos(\pi k(j-1/2)/P) & k > 0. \end{cases} \tag{2.31}$$

The ground state Hamiltonian H_P^0 now takes the desired form:

$$H_P^0(p,q) = \sum_{i=1}^{N} \sum_{k=0}^{P-1} \frac{\left[\widetilde{p}_i^{(k)}\right]^2}{2m_i} + \frac{1}{2} m_i \omega_k^2 \left[\widetilde{q}_i^k\right]^2 \tag{2.32}$$

with $\omega_k = 2\omega_{P-1} \sin(k\pi/(2P))$. This Hamiltonian is much easier to deal with, since the kinetic spring terms are now decoupled to independent harmonic oscillators (normal modes), whose equations of motion are analytical in the form of sines and cosines. As well, it is important to note that the individual beads are propagated first in the position representation for the potential terms and then into the normal mode space for the kinetic spring terms. We use the PILE thermostat to perform this propagation efficiently. The requirement is that the momenta are scaled prior and following the propagation of the kinetic springs in

normal mode space in the following way:

$$\widetilde{p}_i^{(k)} \leftarrow c_1^{(k)} \widetilde{p}_i^{(k)} + \sqrt{\frac{m_i}{\beta}} c_2^{(k)} \xi_i^{(k)} \tag{2.33}$$

with

$$c_1^{(k)} = \exp\left(-\frac{\Delta t}{2}\gamma^{(k)}\right) , \tag{2.34}$$

$$c_2^{(k)} = \sqrt{1 - \left[c_1^{(k)}\right]^2} , \tag{2.35}$$

where $\gamma^{(k)}$ is the Langevin equation friction on the kth normal mode, and $\xi_i^{(k)}$ is a normally distributed random number. The optimal Langevin frictions are determined analytically for all modes, except for the zeroth normal mode corresponding to the centroid of the path. These frictions take the form:

$$\gamma^{(k)} = \begin{cases} 1/\tau_0 & k = 0 \\ \\ 2\omega_k & k > 0. \end{cases} \tag{2.36}$$

In this thesis, we calculate the autocorrelation time of the path centroid and use this as our τ_0 parameter.

2.4 Relationship Between Trial and Ground State Wavefunction

The formulation of the Path Integral Ground State method is derived to obtain nuclear ground state properties for a system described by this Hamiltonian:

$$\hat{H} = \sum_{i=1}^{N} \frac{\hat{p}_i^{\,2}}{2m_i} + \hat{V} \tag{2.37}$$

where N is the number of particles in the system, $\hat{p}_i$ is the momentum operator and m_i is the mass for particle i, and $\hat{V}$ is the potential energy operator that describes the interactions between all particles of the system.

Since the ground state wavefunction of a system is most often not known, we must use a so-called 'trial wavefunction' $\psi_T(R)$ in our simulation. Here R is a vector containing the $3N$ Cartesian coordinates of the N atoms in three dimensions. The exact ground state wavefunction in the position representation $\Phi_0(R) = \langle R|\Phi_0\rangle$ is proportional to the following integral:

$$\Phi_0(R) \propto \lim_{\beta\to\infty} \int \mathrm{d}R' \langle R| \exp\left(-\frac{\beta}{2}\hat{H}\right) |R'\rangle \psi_T(R') \ . \tag{2.38}$$

The parameter $\beta/2$ relaxes the trial wavefunction to the ground state wavefunction.

The following analysis of Equation 2.38 will illustrate why the chosen trial wavefunction must overlap with the ground state wavefunction and why a trial wavefunction with more overlap is ideal. We begin with the right hand side

$$\exp\left(-\frac{\beta}{2}\hat{H}\right)|\psi_T\rangle \ . \tag{2.39}$$

The trial wavefunction $|\psi_T\rangle$ can be written as a linear combination of all orthonormal nuclear ground and excited state wavefunctions, $|\Phi_i\rangle$:

$$\exp\left(-\frac{\beta}{2}\hat{H}\right)|\psi_T\rangle = \exp\left(-\frac{\beta}{2}\hat{H}\right)\sum_i c_i|\Phi_i\rangle \tag{2.40}$$

where c_i is a coefficient representing the overlap between $|\psi_T\rangle$ and $|\Phi_i\rangle$. Now, using the fact that Φ_i has an associated energy eigenstate of the Hamiltonian, E_i, we get:

$$\exp\left(-\frac{\beta}{2}\hat{H}\right)|\psi_T\rangle = c_0\exp\left(-\frac{\beta}{2}E_0\right)|\Phi_0\rangle + c_1\exp\left(-\frac{\beta}{2}E_1\right)|\Phi_1\rangle + c_2\exp\left(-\frac{\beta}{2}E_2\right)|\Phi_2\rangle + \ldots \tag{2.41}$$

which can be re-written as:

$$\exp\left(-\frac{\beta}{2}\hat{H}\right)|\psi_T\rangle = \exp\left(-\frac{\beta}{2}E_0\right)\times$$
$$\left\{c_0|\Phi_0\rangle + c_1\exp\left[-\frac{\beta}{2}\left(E_1-E_0\right)\right]|\Phi_1\rangle + c_2\exp\left[-\frac{\beta}{2}\left(E_2-E_0\right)\right]|\Phi_2\rangle + \ldots\right\}\ . \quad (2.42)$$

We know that each E_i is larger than E_0 for $i > 0$, so in the $\beta \to \infty$ limit, the ground state term is all that remains

$$\lim_{\beta\to\infty}\exp\left(-\frac{\beta}{2}\hat{H}\right)|\psi_T\rangle = \ c_0\exp\left(-\frac{\beta}{2}E_0\right)|\Phi_0\rangle\ . \quad (2.43)$$

The above is true unless the ground state is degenerate which is not the case for the systems we are focusing on. Rearranging, and using the position representation, we arrive at an expression that is equivalent to Eq. (2.38) above up to a proportionality constant

$$\langle R|\Phi_0\rangle \propto \lim_{\beta\to\infty}\exp\left(-\frac{\beta}{2}\hat{H}\right)\langle R|\psi_T\rangle\ . \quad (2.44)$$

It is important to note that any trial wavefunction with finite overlap with the trial wavefunction will allow for the calculation of exact ground state properties given large enough sampling, however, the extent of the overlap affects the convergence of ground state properties. A trial wavefunction that has small overlap will require a longer projection time,

β, in order to obtain exact results, whereas one with larger overlap will require a much smaller β. In the limiting case where our trial wavefunction is orthogonal to the ground state wavefunction and thus has no overlap, we will not be able to reproduce any ground state properties. In the other case where our trial wavefunction is exactly the ground state wavefunction, any value β can be chosen to reproduce ground state properties. It is interesting to note that the error in β being a finite number in our simulation includes information about excited states through the overlap of our trial wavefunction and the excited state wavefunctions. It is also important to select a trial wavefunction with a large overlap with the ground state wavefunction to further minimize the effect of the excited states.

A significant difference between finite temperature PIMD and our LePIGS method is how nuclear exchange is treated. In finite temperature PIMD, we require separate algorithms (such as WORM[44]), to sample exchange. However, in LePIGS, we sample the ground state wavefunction which already accounts for nuclear exchange! This is important for bosonic systems, such as the hydrogen clusters that are studied in this thesis. Due to bosonic symmetry, the ground state wavefunction must account for all N! permutational isomers. This can be represented in the following way:

$$\psi_0(\vec{R}) = \frac{1}{N!}\sum_{i=1}^{N!}\psi_0\left(P_i(\vec{R})\right) \tag{2.45}$$

where P represents the N! configurational permutations.

2.5 Molecular Dynamics vs Monte Carlo Sampling

There are two main methods used to evaluate these path integrals, both at finite temperature and in the zero temperature limit (ground state). This research will focus on using molecular dynamics as the tool to simulate the systems of interest, although Monte Carlo sampling has traditionally been used[43]. The goal of Monte Carlo is to randomly sample path configurations. One such algorithm used under canonical ensemble conditions is Metropolis sampling[45], where N particles are initially placed in the system in any configuration. Beads of each particle are then moved one at a time, where each suggested move takes the form

$$X \rightarrow X + \alpha\xi_1 \ , \tag{2.46}$$

where X is the configuration of the particle prior to the move, α is the maximum specified distance of the move, and ξ_1 is a random variable chosen between (-1,1) specifying the actual distance of the move. The move is then accepted or rejected based on a probability meant to conserve the distribution of the canonical ensemble. If the energy of the system is lower after the move is made, the move is accepted. However, if the energy is higher after the move is made, the move is accepted only if the following condition is satisfied

$$\xi_2 < e^{(-\frac{\Delta E}{k_B T})} \ , \tag{2.47}$$

where ξ_2 is a different random number between (0,1), ΔE is the change in energy of the system, k_B is the Boltzmann constant, and T is the temperature of the system. If the move is not accepted, the configuration of the particle remains at its initial configuration prior to the suggested move. The next bead is then suggested a move and this sequence repeats for each particle over a specified number of steps. Regardless whether the configuration has changed or not, the configuration at each step is considered for all averages and energy calculations.

In contrast to the purely random walk sampling method, molecular dynamics[46, 47] uses generalized forcefields and equations of motion to move the atoms. As mentioned, we use the Langevin equation or Langevin thermostat[43], as it ensures that the canonical ensemble detailed balance is satisfied. It has been demonstrated with low temperature quantum systems and has shown to be extremely accurate and efficient[39]. The Langevin thermostat introduces a friction and Gaussian random force within the velocity term. This ensures that for each degree of freedom, the equipartition theorem is on average satisfied allowing the simulation to maintain a constant temperature. This friction value is chosen in such a way to optimize the efficiency of the simulation, so it can not significantly impede

the particle motions. The algorithm for updating the configurations of particles in the Langevin equation is through the velocity at a time-step of Δt. The velocity of a particular bead j with mass m_j of an atom is expressed as

$$v_j = e^{(-\frac{\Delta t}{2})\gamma} v_j + \sqrt{\frac{1}{\beta m_j}[1 - e^{(-\Delta t)\gamma}]\xi_j} \tag{2.48}$$

where γ is the friction chosen to dampen the velocities and ξ_j is the bead specific random number drawn from a Gaussian distribution of the form $\exp\left[-\frac{1}{2}(\Delta t)^2\right]$.

Molecular dynamics has multiple advantages over Monte Carlo. The first advantage is that a molecular dynamics simulation is performed in real time, using velocities and time-steps, and in Monte Carlo the length of the simulation is determined only by number of steps. This gives a more physical picture and also allows for the extension to Ring Polymer Molecular Dynamics (RPMD)[48], which provides information on dynamic properties and spectroscopic information. Another advantage is that the number of system-specific parameters needed to design MC moves grows with the number of particles. In molecular dynamics using the Langevin equation thermostat, the only parameters needed are the friction value and a correlation time which apply to the whole system regardless of the number of particles within it. Finally, Ing *et al.* showed that the efficiency between the two methods is on the same time scale, where Monte Carlo is approximately 25% faster

than dynamics[39].

Chapter 3

Benchmark: Finite Temperature Path Integral Langevin Equation

3.1 PIMD proof of principle : (Fixed H_2O)-pH_2

The first step is to confirm that the PIMD code works for a system that consists of a fixed water molecule and one parahydrogen molecule at a temperature of 0.37 K. The goal is to confirm that a simulation that uses a large number of beads over a long enough time period provides structural distributions and an energy (within error) of the exact energy and distributions calculated with an appropriate basis set[12]. As mentioned in the path integral theory section, the expected quadratic convergence of the energy with

τ must also be confirmed. Finally, since calculation costs increase with the number of beads, the minimum number required to give accurate energies and structural properties must be determined. This will give insight into the number of beads necessary for larger water-parahydrogen systems.

3.1.1 System Setup

The PILE thermostat[39] will be used to perform dynamics using the source code MMTK[38]. Due to the fact that the rotation of parahydrogen is much faster than that of water, the water molecule only feels the spherically averaged potential of the parahydrogen. This means that the parahydrogen molecule (pH_2) can be considered a point-like particle[49], as is discussed in a paper by Zeng *et al.*[12], which provides a 3D water-parahydrogen potential condensed from 5D using the adiabatic hindered rotor approximation. The adiabatic hindered rotor approximation averages over the internuclear distance between the two hydrogen atoms in pH_2 and an angle describing the relative orientation of the pH_2 bond axis with respect to the water molecule. It is this potential that will be used to construct the forcefield. The water-parahydrogen system can be described by three coordinates:

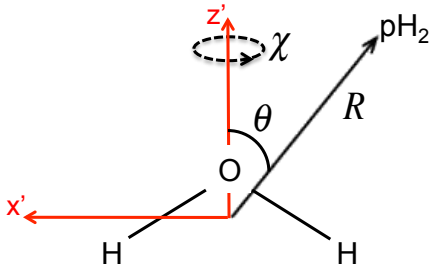

Figure 3.1: Co-ordinate system between water and a parahydrogen molecule (represented as a pointlike particle).

(1) the distance, R, between the pH_2 and the water centre of mass, (2) the polar angle θ, and (3) the azimuthal angle χ. The energy estimator that will be used for the dynamics is the centroid virial estimator, which is considered the most robust in the previous low temperature study by Ing *et al.*[39]. The primitive estimator will be used to calculate the PIMC energy.

The two parameters necessary to run an efficient simulation that gives independent data points are the centroid friction and the correlation time. The centroid friction is determined by running a microcanonical simulation and calculating the correlation times of the structural properties of the centroid. The friction is the inverse of the longest centroid correlation time. The simulation correlation time is determined by running a canonical simulation and calculating the correlation times of the structural properties using any bead. The longest correlation time is the one chosen for our simulation. The production

run consists of running the simulation for 10 000 independent time steps. The total energy and structural properties are averaged over all the beads at each independent step. A parahydrogen molecule is placed in the system at an arbitrarily chosen position close to the fixed water molecule and a steepest descent energy minimization is performed using the source code MMTK [38]. The beads of the parahydrogen molecule are initially placed on top of each other, so a short 0.1 fs equilibration is done to separate the beads. A large enough timestep must be used so the parahydrogen can sample many different orientations around the water molecule, but not large enough to allow for unphysical behaviour. A basis set calculation, used previously by a member of our group and described in detail in Ref. [12], will be the benchmark for our structural and energetic properties.

3.1.2 Results

The simulation is run using various numbers of beads ranging from $P = 128$ to $P = 4096$. The energy convergence with τ is shown below in Figure 3.2, including a quadratic fit, and the exact basis set result placed on the τ-intercept. It is important to note that all error bars shown for the energy of a simulation represent standard errors.

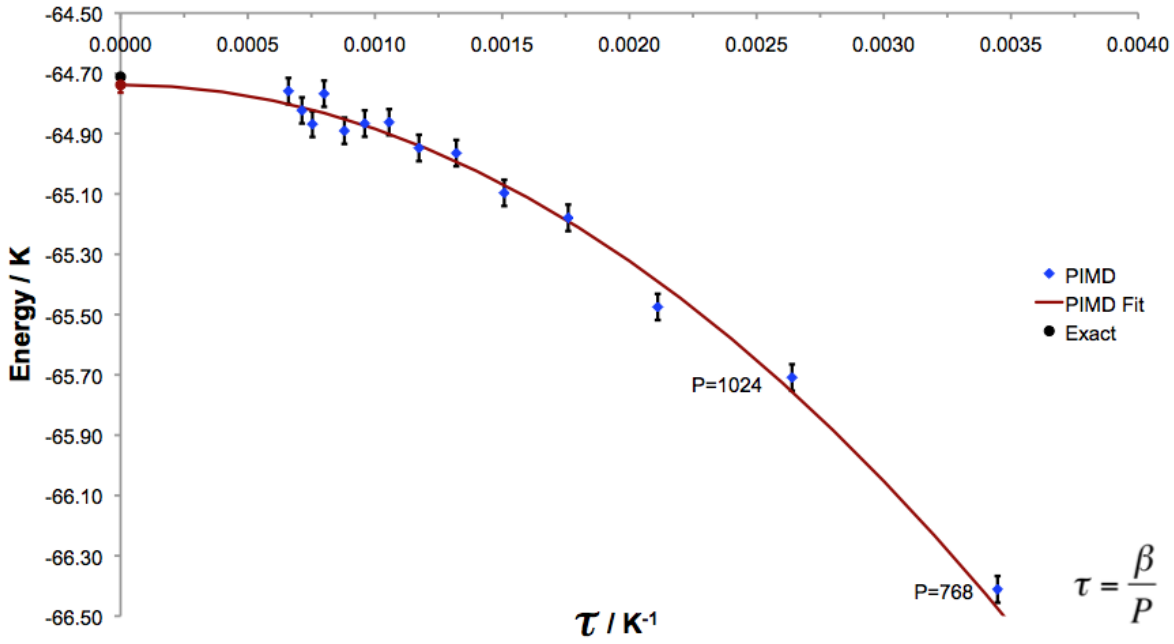

Figure 3.2: PIMD: Energy convergence with τ. Shown in blue are the energies obtained from PIMD simulation of varying number of beads, P. The red curve is a quadratic fit through the PIMD points, and the black point is the exact energy obtained from the basis set calculation.

It is clear that there is a quadratic relationship between the energy and τ and that in the limit as $\tau \to 0$, the exact energy is reached within error. It must also be noted that at the τ values corresponding to $P = 1024$ beads, the energy obtained from the PIMD simulation is within 1 K or approximately 1.5% of the exact energy. This value will be used as the minimum number of beads required to reach approximately the exact energy. It is now important to look at the convergence between structural properties and the number

of beads, shown in Figure 3.3. It should be noted that for a structural property A, the distributions $\rho(A)$ take the form

$$\int dA\rho(A)A = 1 \ . \tag{3.1}$$

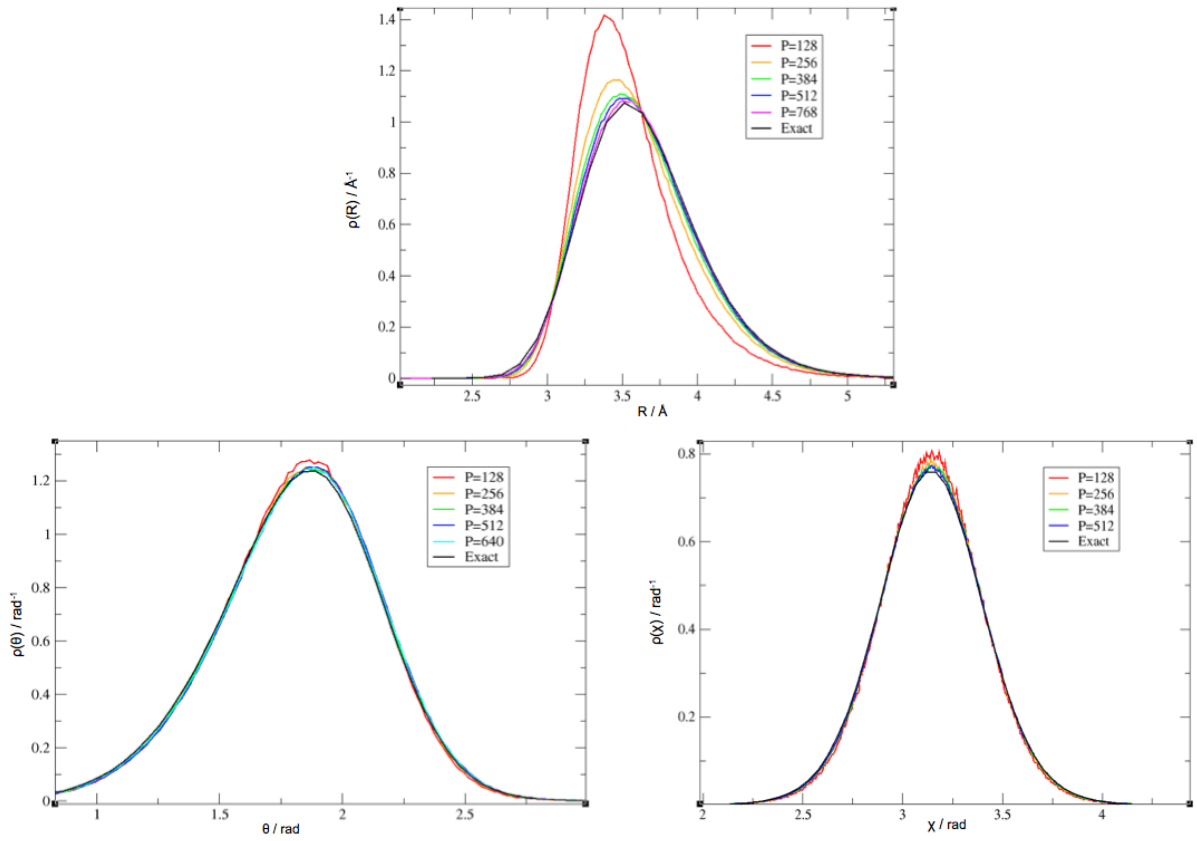

Figure 3.3: PIMD: Structural property convergence with number of beads, P. Shown in black are the exact distributions obtained from the basis set calculation.

The exact distributions of the structural properties are eventually obtained by increasing the number of beads. Also, it is important to note that the minimum number of beads required to reach the exact distributions is approximately $P = 768$. However, there is one small issue that must be clarified in regards to the distribution of the azimuthal angle χ. Based on the ergodic assumption which states that all orientations may be sampled over a sufficiently long time period, the parahydrogen molecule should be able to travel across the high energy barrier to the other side of the water molecule as it is symmetric and one side should not be favoured over the other. In reference to Figure 3.4 below, this is not the case.

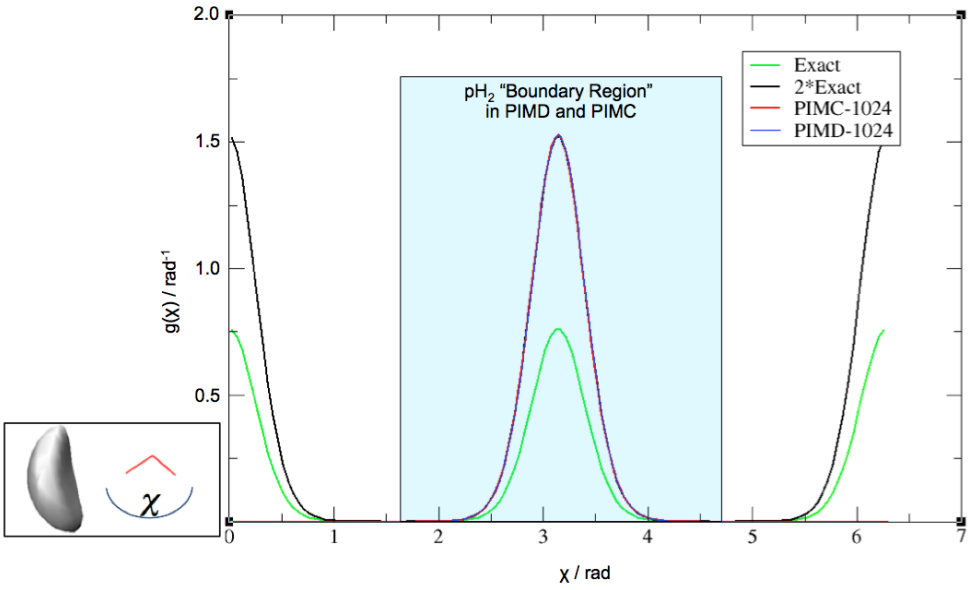

Figure 3.4: Ergodicity problem in PIMD for the angle χ. Isosurface (on left) shows the distribution of the parahydrogen molecule only on one side of the symmetric water molecule. Angular (χ) distribution (on right) shown for PIMD, PIMC, and basis set calculation.

In PIMD (and PIMC) the parahydrogen molecule appears to be 'stuck' on one side and is not able to cross the high energy barrier. Understanding that the water molecule has C_{2v} symmetry, the exact distribution of χ is multiplied by two since the density of the parahydrogen that should be evenly distributed on both sides is now only found on one side during the simulation. It is this modified χ distribution that agrees with those given by PIMD and PIMC. Although molecular dynamics has this so-called ergodicity problem for this system, it should be noted that the other simulation method, PIMC, also has this problem. To summarize the results, as expected the molecular dynamics code provides the exact structural and energetic properties benchmarked against a basis set calculation. The energy convergence with τ was quadratic as expected and in the limit as $\tau \to 0$, the exact energy was acquired. The minimum number of beads necessary to obtain the exact properties is $P = 1024$, since the energy convergence took more beads than the structural property convergence. Molecular dynamics for this system at low temperature introduced an ergodicity problem, however, this is not a big problem as PIMC sampling also introduces this problem.

3.2 PIMD benchmark vs PIMC : (H_2O)-$(pH_2)_n$

Having determined parameters, such as number of beads, friction, and autocorrelation time, the next step is to apply these parameters to systems involving larger clusters of up to 13 parahydrogen molecules. One set of parahydrogen clusters to be simulated with the water dopant will be treated as fixed and another set will have a translating, rigid water dopant. However, water rotation will not be accounted for since rigid-body rotation has not yet been implemented in our PIMD code. The energies will be benchmarked against the more traditional PIMC code. The expected result is that the energies for all cluster sizes will be within error of the benchmark PIMC code. The expected difference in the energy of a translating water should be higher than that of a fixed water since the energy increases when adding additional degrees of freedom.

3.2.1 System Setup

We must now introduce a parahydrogen intermolecular potential, so we choose the Buck potential[50]. The parameters determined in the previous section will be used for all cluster sizes. The translating water molecule will also be represented by the same number of beads, $P = 1024$, as the parahydrogen molecule.

3.2.2 Results

For cluster sizes of $n = 1 - 13$, the energy per parahydrogen molecule is shown in Figure 3.5 for both a fixed and translating water molecule.

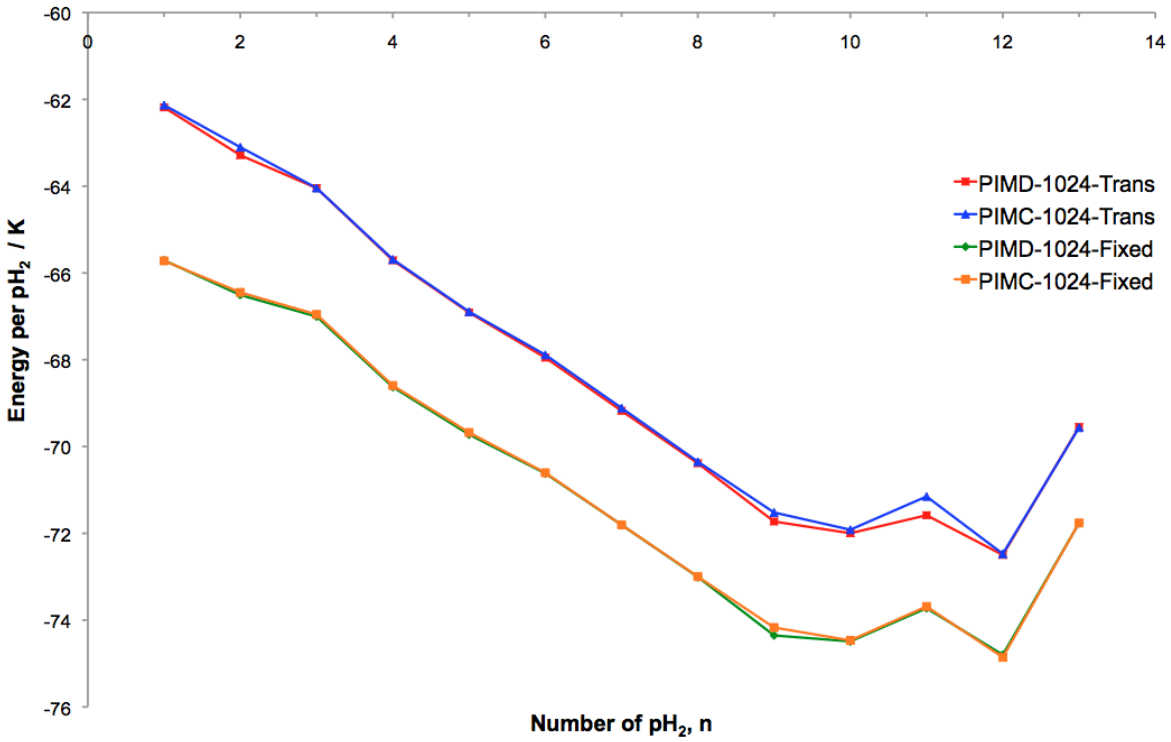

Figure 3.5: PIMD: Energy per pH_2 molecule of parahydrogen clusters of size $n = 1 - 13$ with a fixed and translating dopant water molecule. Errors are contained within the plotted points.

It can be seen that the energies obtained from PIMD match up very well against the benchmark of PIMC. However, there are some discrepancies at $n = 9, 11$. Since the energies from PIMC are always slightly higher than those from PIMD, this may be due

to an ergodicity problem in PIMC that is not present in PIMD at larger cluster sizes. However, more work is required to confirm this suggestion.

3.3 Conclusions

Overall, the objective of these benchmarks are to confirm that the PIMD code using the Langevin equation in MMTK gave the correct properties and convergence behaviour. We successfully simulate a parahydrogen interacting with a fixed water and we obtain the same energy and structural properties as our benchmark basis set calculation. We also confirm the energy convergence with τ is quadratic. Using these results, we extend our system to a cluster of parahydrogen molecules surrounding both a fixed water molecule and one that is translating, but remains rigid. Since our system is too large to perform the same basis set calculation, we choose the traditional PIMC method as our benchmark. The energies we obtain as a result of our MD simulation agree very well to those from PIMC, however there are a few cluster sizes where the energies differ, which may represent an ergodicity problem in our PIMC runs not present in our PIMD simulations. Future work is required to confirm this.

Chapter 4

LePIGS proof of principle : (Fixed H_2O)-pH_2

4.1 Introduction

This chapter[1] will cover one of the first ground state tests of a complex system using molecular dynamics. Similar to the PIMD proof of principle, the goal will be to ensure that the energy and structural distributions agree with the exact ground state or zero-temperature energies and structural properties obtained from a basis set calculation[12].

[1]Sections of this paper have been reprinted with permission from S. Constable, M. Schmidt, C. Ing, T. Zeng, and P.-N. Roy, "Langevin Equation Path Integral Ground State", J. Phys. Chem. A, 2013, 117 (32), pp 74617467,http://dx.doi.org/10.1021/jp4015178. Copyright 2013 American Chemical Society.

A feature that is not present in PIMD, but is found in LePIGS[37] is the convergence of β. If β is too low, even in the limit of $\tau \to 0$, we will not obtain the correct energy and structural distributions. The final goal is to confirm that the energy convergence with the parameters β and τ agrees with theory. For simplicity, the most basic trial function is implemented, $\psi_T = 1$. Noted in the PIGS theory section is that the total energy is evaluated by the energy of the end beads only and the structural properties are determined from the middle bead only. The number of independent time steps for the simulation must be increased to 100 000 to reduce statistical error. The β convergence must be done first and once the correct β parameter has been chosen, the τ convergence can be completed. The smallest β and largest τ are optimal as these minimize the number of quantum beads required for the simulation. It is expected that the convergence of β and τ will agree with theory and using large enough β and small enough τ values, the structural distributions and energy from LePIGS will agree with the exact calculations obtained from the basis set.

4.2 System Setup

The LePIGS method[37] was implemented in the freely available MMTK[38] based on the implementation of PILE by Ing *et al*[39]. The DCT code was provided by the FFTW [51] package. The method for our basis set calculation is described in detail in Ref. [12] and

only differs from the one used in Section 3.1 by the temperature. Since the basis set code requires a temperature, we use $T = 0.0000001$ K to approximate zero temperature.

4.3 Results

The first application to a 'real' system using this technique is the water-parahydrogen dimer. As before, the water molecule is treated as a fixed molecule and provides an external potential to the parahydrogen molecule, treated as a point-like particle. The water-parahydrogen potential we use is from Ref. [12]. This potential has recently been used by our group study a parawater molecule embedded in superfluid hydrogen clusters[52].

The Hamiltonian for a system consisting of a parahydrogen particle of mass m and a fixed H_2O is,

$$H_{H_2} = \frac{p^2}{2m} + V_{H_2-H_2O}(\mathbf{R}) \ , \tag{4.1}$$

where $\mathbf{R}$ is the position of the point-like hydrogen molecule in the water molecule frame.

In order to perform the simulations, three parameters need to be determined, the time step dt, the centroid friction $\gamma^{(0)}$, and the correlation time of the system which is needed to obtain time-independent data points. A time step $dt = 2.0$ fs was chosen based on experimentation and previous low temperature path integral molecular dynamics

simulations[39].

Several steps are needed to determine the centroid friction value. First we perform an energy minimization to get our initial configuration, followed by a short LePIGS simulation without friction, but skipping enough steps to ensure uncorrelated configurations. Using each of these uncorrelated configurations as initial configurations, a microcanonical ensemble simulation is run (holding constant number of atoms, volume, energy). The friction value chosen is the averaged time-correlation of structural properties over all the microcanonical ensemble simulations. Once the centroid friction is determined, we run another short LePIGS simulation using that friction and not skipping any steps. The correlation times are then determined for each of the structural properties. The longest correlation time is the one used for the simulation.

The 'exact' energy, radial distribution and angular distributions were obtained using a basis set calculation for a parahydrogen interacting with a fixed water molecule using the approach of Ref. [12]. The definitions of these distributions are the same as in the last section using Figure 3.1. The energy convergence diagram of β is shown in Fig. 4.1. However, the energy will not converge to the 'exact' value since our τ value is not ideal. This is not a concern, since the energy will just converge to a lower energy value than the exact ground state energy.

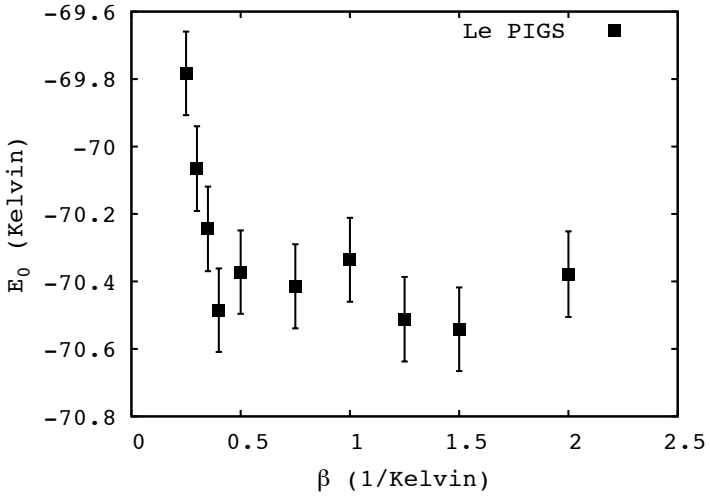

Figure 4.1: Exponential convergence of the LePIGS E_0 with respect to β for the water-parahydrogen dimer ($\tau = 0.007$ K^{-1}).

A β value of 1.00 K^{-1} is chosen, since it is safely in the region of the converged energy value. Using this β value, the energy convergence with τ is shown in Fig. 4.2. The black points are the LePIGS simulation energies and the red dashed line is the 'exact' energy. The blue curve is the quadratic fit with the blue point as the energy extrapolated to the $\tau = 0$ limit with it's associated error. The LePIGS simulation obtains the exact energy within error at $\tau = 0.001$ and in the limit as $\tau = 0$.

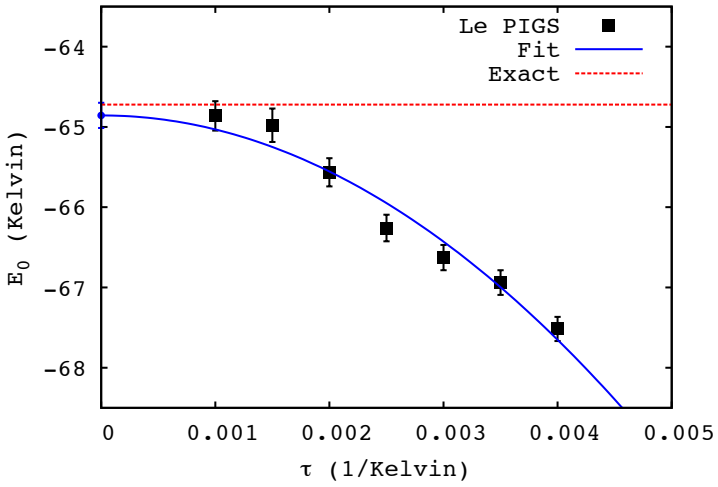

Figure 4.2: Convergence of the LePIGS E_0 of the water-parahydrogen dimer with respect to τ for $\beta = 1.00$ K^{-1}. A parabolic fit to the data is shown as a solid line. The exact value obtained from explicit diagonalization is shown as a dashed line.

The convergence of the radial and angular distributions with β and τ are also shown below and compared to the 'exact' distributions. Figures 4.3, 4.4, and 4.5 show the structural distribution dependence on τ. When varying β, there is only a significant difference in the θ distribution shown in Figure 4.6. In all cases, convergence is systematically achieved when β (τ) is sufficiently large (small).

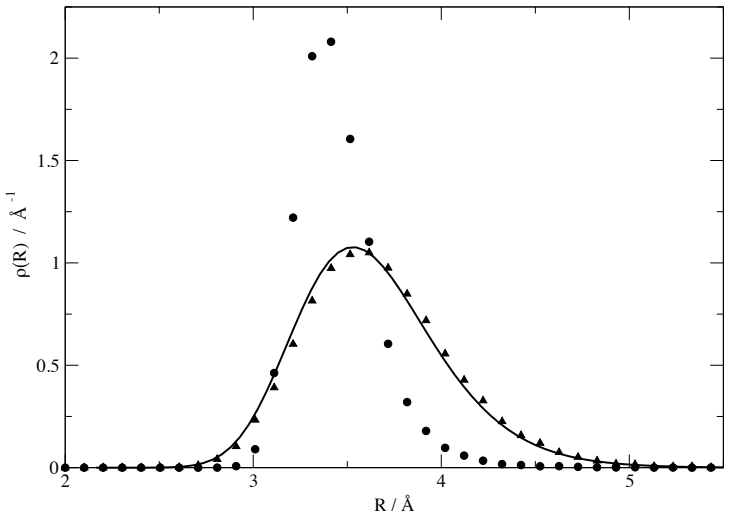

Figure 4.3: Radial distribution for β=1.00 K^{-1} at two different τ values. The circles represent τ=0.050 K^{-1}, the triangles represent τ=0.001 K^{-1}, and the solid line represents the exact distribution.

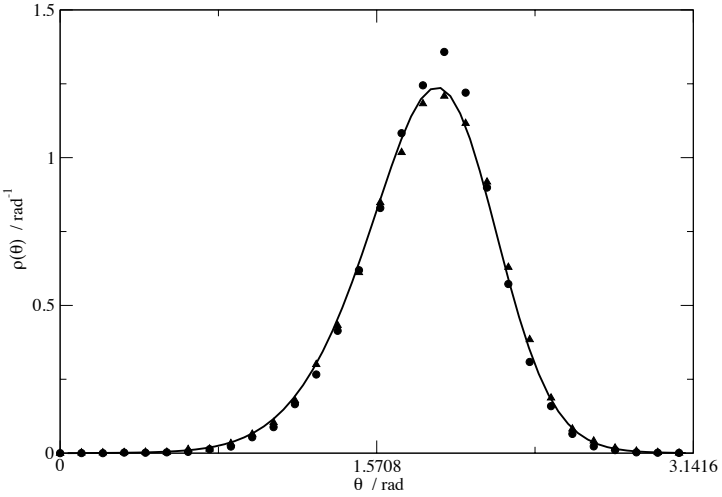

Figure 4.4: Angular distribution of θ for β=1.00 K^{-1} at two different τ values. The circles represent τ=0.050 K^{-1}, the triangles represent τ=0.001 K^{-1}, and the solid line represents the exact distribution.

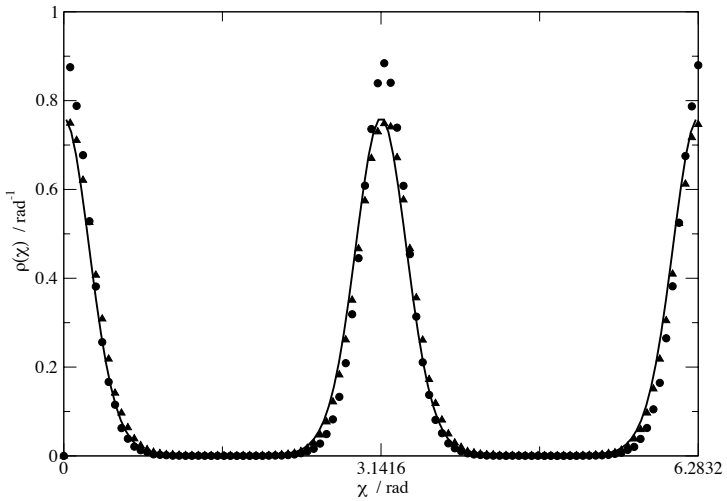

Figure 4.5: Angular distribution of χ for β=1.00 K^{-1} at two different τ values. The circles represent τ=0.050 K^{-1}, the triangles represent τ=0.001 K^{-1}, and the solid line represents the exact distribution.

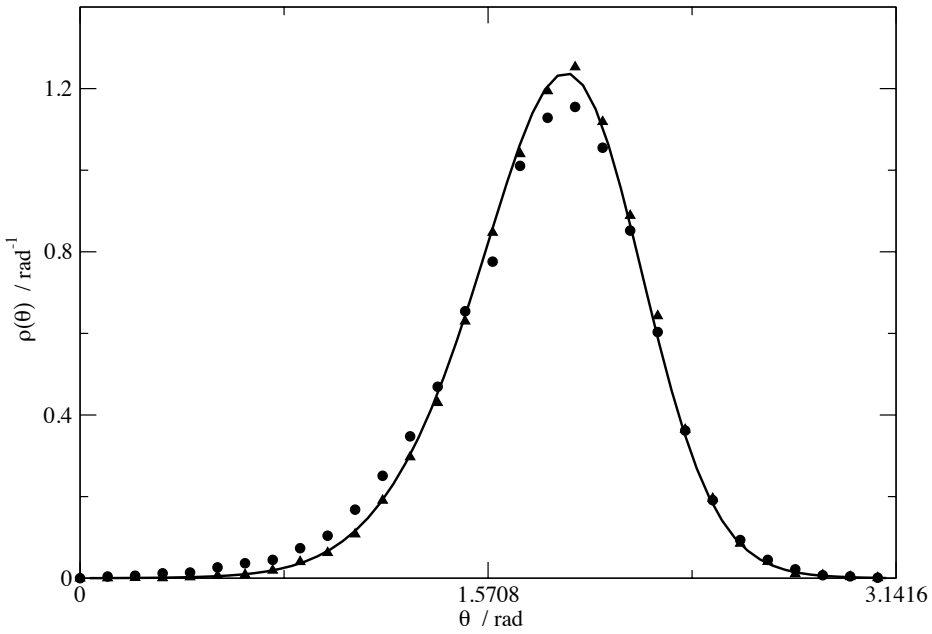

Figure 4.6: Angular distribution of θ for τ=0.007 K^{-1} at two different β values. The circles represent β=0.400 K^{-1}, the triangles represent β=1.0 K^{-1}, and the solid line represents the exact distribution.

4.4 Concluding Remarks

We have successfully calculated the ground state energy and structural properties of the water-parahydrogen dimer and our results agree with our basis set calculation benchmark. We have also successfully shown the quadratic behaviour of energy convergence with τ and the exponential behaviour of energy convergence with β. As a result, we conclude that the LePIGS method is practical and accurate.

Chapter 5

Inclusion of trial functions in the Langevin equation Path Integral Ground State method: application to parahydrogen clusters and their isotopologues

5.1 Introduction

In the last chapter, we only use a trivial trial function $\psi_T = 1$ to perform our LePIGS simulation. This chapter[1] will explore ground state properties of parahydrogen clusters and the effect of three different trial wavefunctions on the convergence of the parameters β (the imaginary projection time) and τ (the imaginary time step). We also thoroughly describe a systematic method to optimize the parameters of a LePIGS[37] simulation: the centroid friction, correlation time, and timestep dt. We extend this study to orthodeuterium (oD_2) and paratritium (pT_2) clusters and compare the effectiveness of using a liquid-like and solid-like trial wavefunction for systems involving heavier isotopes. As of today, we have not found previous literature values for ground state properties of tritium molecular clusters, though previous work has looked at clusters of spin-polarized tritium atoms.[53] This will be the first systematic study of the ground state properties of small paratritium clusters.

We closely examine the effect of two trial wavefunctions that represent a solid-like and liquid-like cluster on three cluster sizes $N = 4, 8, 13$ for parahydrogen, orthodeuterium, and paratritium. We then use these results to extend to the full cluster range from $N = 4 - 19, 33$.

[1]This chapter (including some of the figures) has been reprinted or adapted from a manuscript (M. Schmidt, S. Constable, C. Ing, and P.-N. Roy, "Inclusion of trial functions in the Langevin equation Path Integral Ground State method: application to parahydrogen clusters and their isotopologues") submitted for publication and currently under second revisions in The Journal of Chemical Physics

5.2 Lindemann Criterion

Whether hydrogen clusters are 'solid-like' or 'liquid-like' is a subject of much debate. As mentioned in the introduction, we intend to test the effect of different trial wavefunctions, one that favours a solid-like structure and one that favours a liquid-like structure, to see the effect on convergence of ground state properties. One such method in determining whether these clusters are solid-like or liquid-like and if the trial wavefunctions that we use bias towards a solid or liquid structure is the Lindemann melting criterion[54]. This criterion compares the relative root mean squared displacement of the bond lengths of a system to some threshold value. This threshold value may vary from 5-20% depending on the cluster size, although is generally assumed to be 15%[55]. In work by Cuervo *et al.*[33], the following definition of δ_L was used,

$$\delta_L = \frac{2}{N(N-1)} \sum_{i<j} \frac{\left(\langle r_{ij}^2 \rangle - \langle r_{ij} \rangle^2\right)^{1/2}}{\langle r_{ij} \rangle} \ , \tag{5.1}$$

where $\langle \cdots \rangle$ denotes the simulation average for each of the interparticle distances r_{ij} between particles i and j. It should be noted that the particles are distinguishable in this definition, therefore we call this quantity the "specific" Lindemann value. They also define a modified Lindemann quantity which treats the particles as indistinguishable, which they call the

"generic" Lindemann value,[33]

$$\delta_{L^G} = \frac{\left(\langle r^2 \rangle - \langle r \rangle^2\right)^{1/2}}{\langle r \rangle} \ , \tag{5.2}$$

where $\langle \cdots \rangle$ denotes the simulation average and the average over all pairs of particles. It was proposed that when these two quantities are the same, the cluster can be considered liquid-like and when the two quantities differ significantly[33], the cluster can be considered more rigid or even solid-like. This is one definition that we will use to determine the rigidity of our clusters and how we describe the effect of the trial wavefunction on the end beads. However, it should be noted that in the limit of a very long simulation, the generic and specific Lindemann measures should give the same values since all exchanges have been sampled. Therefore the difference of the Lindemann values is a measure of ergodicity of the simulation.

A Lindemann value for quantum clusters was proposed in 2011 by Guardiola and Navarro[56]. This new parameter, which we will refer to as δ_Q for 'quantum', is a true observable and is defined as

$$\delta_Q = \sqrt{N-1}\frac{\sqrt{\langle \left[\sum_{i<j} r_{ij}^2\right]^2 \rangle - \langle \sum_{i<j} r_{ij}^2 \rangle^2}}{\langle \sum_{i<j} r_{ij}^2 \rangle} \ . \tag{5.3}$$

Note, that it is the fluctuation of the observable

$$O = \frac{1}{N_P} \sum_{i<j} r_{ij}^2 \tag{5.4}$$

with a pre-factor of (N - 1)$^{1/2}$. They conclude that clusters are liquid-like for approximately $\delta_Q > 0.30$ and solid-like for approximately $\delta_Q < 0.20$, and in between both phases co-exist.

We have also defined a similar observable, although ours is not of the square of the pair distance, but the pair distance itself. We call it δ_F as it is a true fluctuation of pair distance defined as

$$\delta_F = \frac{\sqrt{\langle \left[\sum_{i<j} r_{ij}\right]^2 \rangle - \langle \sum_{i<j} r_{ij} \rangle^2}}{\langle \sum_{i<j} r_{ij} \rangle} \ . \tag{5.5}$$

With this definition, we can not claim which values represent liquid-like or solid-like clusters, but use this only as an alternative comparison to the other methods listed above and to view it's evolution with cluster size.

For the case study models, N=4, 8, and 13, only the Lindemann specific and generic will be used to determine ergodicity effects. However, the 'quantum' Lindemann and fluctuation of pair distances approach will be calculated for the full range of cluster sizes

and compared to the specific and generic Lindemann values.

5.3 System Setup

5.3.1 Trial Wavefunctions

We investigate the effect of 3 different trial wavefunctions, therefore varying the degree of overlap with the exact ground state wavefunction of these hydrogen clusters. The first is a unitary trial function taking the form:

$$\psi_T = 1 \tag{5.6}$$

This wavefunction is a very crude guess, however, it ensures some overlap with the ground state wavefunction. The energy using the mixed estimator results in,

$$\frac{\hat{H}\psi_T(R)}{\psi_T(R)} = \frac{\left(-\frac{\hbar^2}{2m}\nabla^2 + V(R)\right)\psi_T(R)}{\psi_T(R)} = V(R)\ , \tag{5.7}$$

where $\hat{H}$ is the Hamiltonian of the system, $\psi_T(R)$ is our trial wavefunction, $\hbar$ is the reduced Planck's constant, m is the mass of the particle, ∇ represents the gradient, and $V(R)$ is the potential energy of the system.

The second trial wavefunction is a Jastrow-type wavefunction that incorporates information on the hard-core repulsion between molecular pairs,

$$\psi_T = \exp\left\{-\frac{1}{2}\sum_{i<j}\left(\frac{b}{r_{ij}}\right)^5\right\} \tag{5.8}$$

This wavefunction was used by Cuervo *et al.* for parahydrogen clusters[32]. The parameter b is the inter-particle distance that reflects the onset of the hardcore repulsion and r_{ij} is the magnitude of the end bead displacements between a pair of hydrogen molecules. For hydrogen clusters, $b = 3.65$ Å. The energy for this wavefunction using the mixed estimator is:

$$\frac{\hat{H}\psi_T(R)}{\psi_T(R)} = \frac{\left(-\frac{\hbar^2}{2m}\nabla^2 + V(R)\right)\psi_T(R)}{\psi_T(R)} = \frac{\hbar^2}{2m}\sum_{i<j}\left\{\frac{10b^5}{r_{ij}^7} - \sum_{\alpha}\left[\frac{5b^5}{2r_{ij}^7}\alpha_{ij}\right]\right\} + V(R) \tag{5.9}$$

where α_{ij} is the Cartesian displacement of the end beads.

The third trial wavefunction that will be used is a normal mode trial wavefunction, which includes information about collective motions of the particles relative to a given minimum potential energy structure. In this study, the minimum structure will correspond to the classical Lennard-Jones global minimum structure[57] using $\sigma = 3.00$ Å followed by a steepest descent energy minimization with the Buck potential[50].

$$\psi_T = \exp\left\{-\frac{1}{2\hbar}\sum_k^{3N-6}\omega_k Q_k^2\right\} \tag{5.10}$$

The parameters ω_k are the frequencies associated with each of the mass-weighted normal modes Q_k and N is the number of particles in the system. The number of vibrational normal modes is $3N-6$ for non-linear systems. The energy using the mixed estimator for this normal mode wavefunction is

$$\frac{\hat{H}\psi_T(R)}{\psi_T(R)} = \frac{\left(-\frac{\hbar^2}{2m}\nabla^2 + V(R)\right)\psi_T(R)}{\psi_T(R)} = \left\{\sum_k^{3N-6}\frac{\hbar\omega_k}{2} - \sum_k^{3N-6}\frac{1}{2}\omega_k^2 Q_k^2\right\} + V(R)\ . \tag{5.11}$$

It is important to note that the first sum in the brackets is exactly the harmonic oscillator zero point energy.

5.3.2 Computational Details

The LePIGS method was implemented by Constable *et al.*[37] in the freely available MMTK[38], based on the implementation of PILE by Ing *et al.*[39]. The DCT code was provided by the FFTW package[51]. The hydrogen potential is from Buck *et al.*[50]. We use the following masses for pH_2: 2.015625035 amu, oD_2: 4.02820356 amu, and pT_2: 6.0320984 amu and a Boltzmann constant: k_B=0.0083144621 kJ·mol^{-1}·K^{-1}.

Our initial configurations for our case studies, N=4, 8, and 13, come from the Lennard Jones classical global minima of the Cambridge Cluster Database[57] and we use σ=3.00Å. For the range of cluster sizes, we only use the classical global minimum configurations for N=4,8,13,19,33. For the intermediary clusters, we choose the first N configurations. For example, the initial configuration for N=14 is chosen from the first 14 atoms from the N=19 classical global minimum co-ordinate set. We then perform a steepest descent energy minimization using the Buck potential and the trial wavefunction (as it adds an extra potential term to the end beads). The system then equilibrates for 5.0ps to separate the beads.

5.3.3 Parameter Optimization

There are two parameters that must be properly chosen in order to maximize the efficiency of the PILE thermostat. The first is known as the centroid friction. The PILE thermostat requires that the centroid friction γ^0=$1/\tau_0$, where τ_0 is the correlation time for the centroid mode. Ceriotti *et al.*[43] showed that choosing a friction too large will dampen the system too severely and will affect the true correlation times of energetic and structural properties. However, when no friction is placed on the centroid, this mode is effectively sampled without a thermostat and thus will not efficiently relax to the canonical ensemble equilibrium.

We borrow the techniques of RPMD[48] to determine this correlation time, although we are not interested at this point in extracting any dynamic data. We perform 2000 microcanonical ensemble simulations starting from independent configurations generated by a Boltzmann distribution through a canonical ensemble simulation. We then generate the normalized autocorrelation function of the interparticle centroid distance for each of those 2000 simulations and average them at each time step. The result of the averaged centroid pair distance autocorrelation functions for $(pH_2)_4$ and $(oD_2)_{13}$ are shown below in Fig. 5.1.

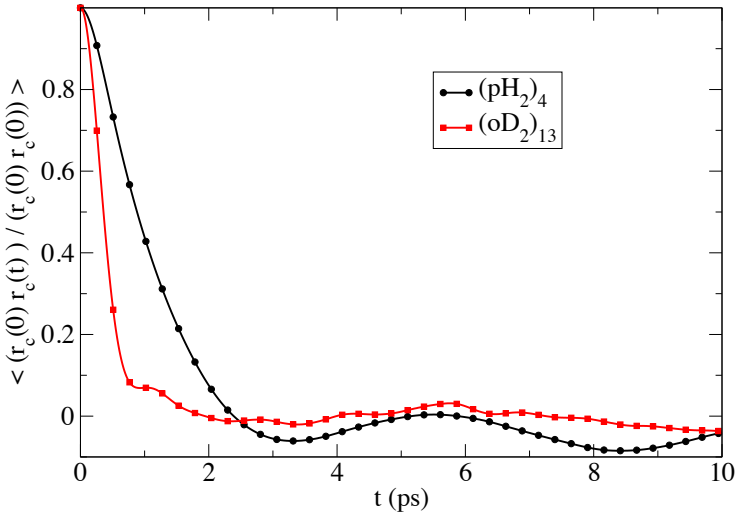

Figure 5.1: Normalized autocorrelation function of the centroid pair distance for $(pH_2)_4$ and $(oD_2)_{13}$.

We see that they decay smoothly to about the value of zero, which is what is expected. We choose the correlation time at which the autocorrelation function is equal to

Euler's number e. This is because we assume, all things being equal, the decay rate of the autocorrelation function $C(t)$ to be first order,

$$C(t) = \exp\left(-\frac{t}{\tau}\right) \tag{5.12}$$

so when we set $t = \tau$, our autocorrelation function becomes e^{-1}. Our friction value becomes $\gamma^0 = 1/\tau_0$.

Using a microcanonical ensemble simulation to determine the centroid friction has been done for the PILE thermostat for low temperature simulations by Ing*et al.*[39], however we perform an average over many microcanonical simulations to get a more accurate determination of the friction. Although in RPMD and low temperature PIMD, the polymer forms a closed ring compared to the open string polymer formed in a LePIGS simulation, the thermostat still acts on the centroid (or zeroth normal mode) so there should be no problem applying this method.

However, the 'best' friction is one that gives us the smallest correlation time. As in Ceriotti *et al*,[43] we plot the correlation times as a function of the log of the friction to determine the optimal friction. We use the $(pH_2)_4$ system and all of our trial wavefunctions for this study. We choose friction values of interest, including zero friction, a friction corresponding to Muser's rule of thumb[58] which is $\gamma^0 = 0.01/\mathrm{dt}$, the friction from the

LePIGS theory[37] which is the inverse of the centroid correlation time, and a number of 'guess' frictions to fill in the plot.

To calculate our correlation times, we perform a 1.0 ns simulation using centroid friction and calculate the correlation times for the properties of interest. We are interested in the total energy and the middle bead distances from which we determine structural properties. However, it is difficult to obtain a good correlation time of the middle bead distances since we can only do it for individual pairs of atoms and we wish to obtain a property of the whole system. Instead, we calculate the correlation time of the potential energy, which is a function of all of the middle bead pair distances in the cluster. The method we use to choose the correlation time from the normalized autocorrelation functions for these properties is the same way we choose the correlation time for the centroid friction using Eq. (5.12).

The correlation times for various friction values are shown below for the normal mode, the Jastrow, and the unity trial wavefunctions in Fig. 5.2. For each plot, it is interesting to note that the total energy correlation time is relatively constant for each trial wavefunction. This is an indication of the robustness of a Langevin equation based sampling approach.

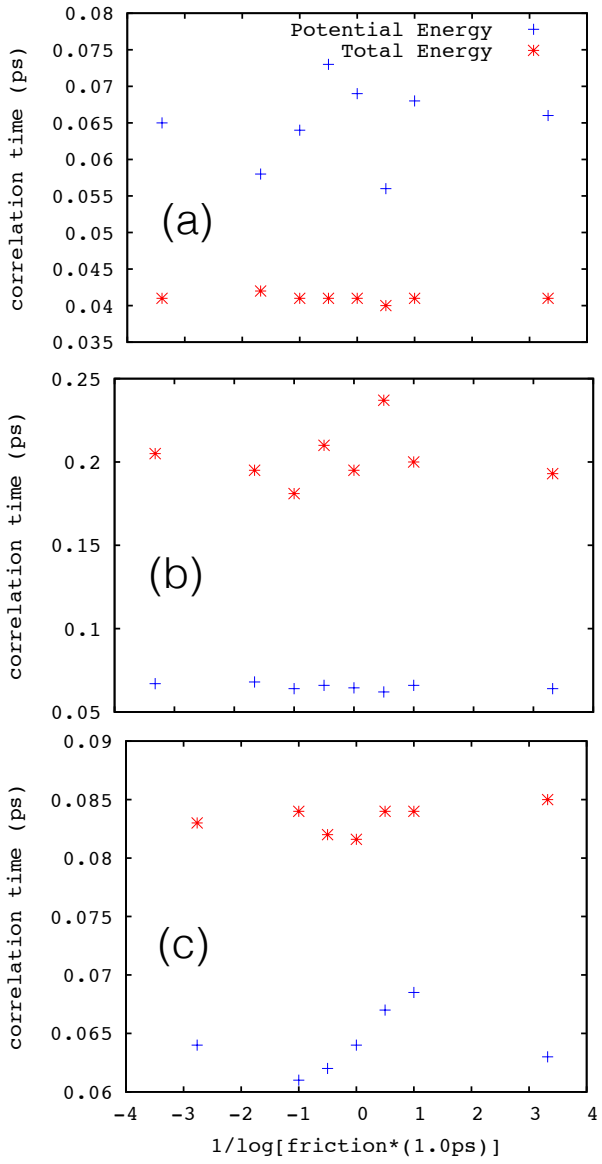

Figure 5.2: Correlation times of potential energy and total energy for the $(pH_2)_4$ system using the (a) unity, (b) Jastrow, and (c) normal mode trial wavefunctions at various centroid friction values.

Since there is nothing to indicate that the friction from LePIGS theory is not ideal, we

will use $\gamma^0 = 1/\tau_0$. The results of our friction and correlation times for both the middle bead and the end bead are shown in Table 5.1. The bold values represent the correlation time we used in our simulations as it is the longest of the two correlation times.

We also perform a full time step study for all trial wavefunctions and all hydrogen isotopes for cluster size $N = 4$ and for $(oD_2)_{13}$ using the normal mode trial wavefunction. We choose this one case to be a representative for the time step associated with all isotopes and trial wavefunctions for $N = 13$. For each study, we run a full simulation of 10^5 independent steps at various values of dt and plot the total energy vs. dt. We choose the largest value of dt at which the total energy converges. For illustration, refer to Fig. 5.3 for the total energy convergence for $(pH_2)_4$ using the various trial wavefunctions.

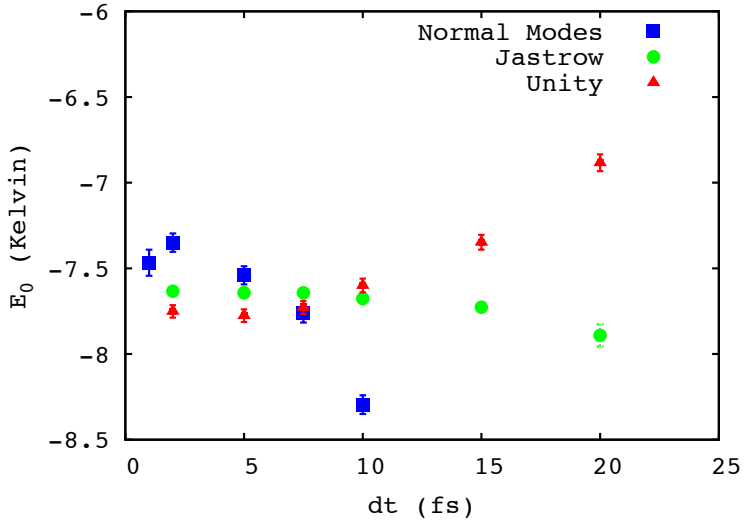

Figure 5.3: Total energy convergence of $(pH_2)_4$ with time step dt using various trial wavefunctions.

The results shown in Table 5.1 also include the time step, dt, that give us a converged value for the total energy. For $(oD_2)_{13}$, the total energy converged at a time step $dt = 3.0$fs for that case. This lower timestep is expected since as cluster size grows, it becomes more solid-like. We then choose dt=5.0fs for all isotopes and wavefunctions for $N = 8$, a reasonable estimate that falls between the time steps used for $N = 4$ and $N = 13$.

Table 5.1: Parameters used for our LePIGS simulations. Simulation values are in bold.

System	ψ_T	γ^0 (ps^{-1})	Correlation	Time (ps)	dt (fs)	β (K^{-1})
			r_{mid}	r_{end}		
$(pH_2)_4$	Unity	1.0/1.97	**0.065**	0.041	7.5	1.25
	Jastrow	1.0/2.00	0.067	**0.200**	7.5	1.25
	Normal Mode	1.0/1.15	0.074	**0.084**	5.0	2.00
$(oD_2)_4$	Jastrow	1.0/1.60	0.082	**0.309**	10.0	1.00
	Normal Mode	1.0/0.71	0.078	**0.101**	5.0	1.75
$(pT_2)_4$	Jastrow	1.0/1.10	0.093	**0.323**	5.0	1.25
	Normal Mode	1.0/0.60	0.095	**0.112**	5.0	1.25
$(pH_2)_8$	Jastrow	1.0/3.30	0.065	**0.193**	5.0	1.00
	Normal Mode	1.0/2.60	0.063	**0.076**	5.0	1.50
$(oD_2)_8$	Jastrow	1.0/2.86	0.075	**0.285**	5.0	1.00
	Normal Mode	1.0/2.10	0.077	**0.087**	5.0	1.50
$(pT_2)_8$	Jastrow	1.0/2.27	0.089	**0.292**	5.0	1.00
	Normal Mode	1.0/2.10	0.083	**0.094**	5.0	1.50
$(pH_2)_{13}$	Jastrow	1.0/2.40	0.067	**0.193**	3.0	0.80
	Normal Mode	1.0/0.49	**0.071**	0.064	3.0	1.00
$(oD_2)_{13}$	Jastrow	1.0/0.43	0.074	**0.269**	3.0	0.80
	Normal Mode	1.0/0.33	0.073	**0.074**	3.0	1.00
$(pT_2)_{13}$	Jastrow	1.0/0.39	0.080	**0.294**	3.0	0.80
	Normal Mode	1.0/0.33	0.081	**0.082**	3.0	1.00

Now that the centroid friction, correlation time, and timestep parameters have been determined, the energy convergence with the relaxation parameters β and τ can be performed for each system. We evaluate the effectiveness of our trial wavefunction by how the energy converges with these parameters β and τ.

5.4 Results and Discussion

5.4.1 A small equivalent molecule cluster: $N = 4$

We first look at a small system, $(pH_2)_4$, where all molecules are equivalent. We look at the energy convergence with β using all three trial wavefunctions. The top panel of Fig. 5.4 shows the convergence of total energy with respect to β for $\tau = 0.007\text{K}^{-1}$ for each wavefunction.

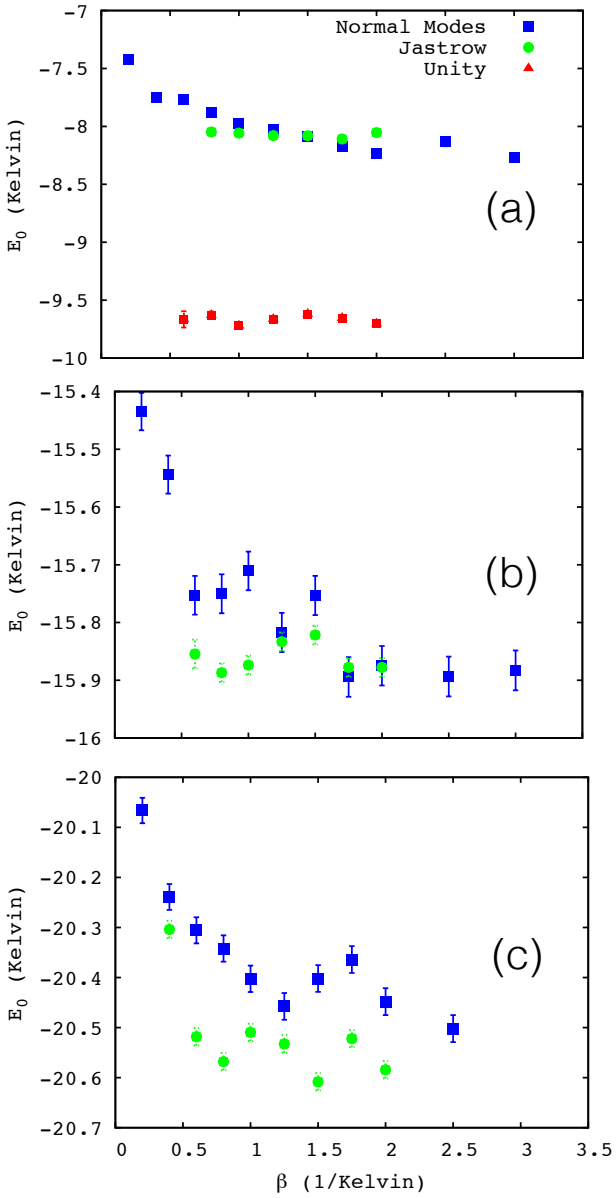

Figure 5.4: Total energy convergence of (a) $(pH_2)_4$, (b) $(oD_2)_4$, and (c) $(pT_2)_4$ with relaxation parameter β using various trial wavefunctions.

The first observation is that the energy does not converge to the same value for each

trial wavefunction. This is due to the τ systematic error, since some trial wavefunctions will take longer to converge than others. We choose a non-ideal τ value so that we capture the full effect of the β parameter and this means we converge to an energy that is not the exact ground state energy. It is also interesting that we have more data points for lower β values using the normal mode wavefunction. Since hydrogen is a very weakly bound system, it can dissociate during the simulation. However, the strength of applying a normal mode wavefunction on the end beads has the ability to keep the system together when hydrogen is represented as smaller chains. Table 5.1 shows our β values at which our energy safely converged.

Using these β values, we then perform our energy convergence with τ for each wavefunction. For $(pH_2)_4$, the results are shown in Fig. 5.5 (a). It is clear that all three wavefunctions converge to the same value within error at low τ. This is expected, since the PIGS theory states that any trial wavefunction with overlap with the ground state wavefunction should reproduce ground state properties. It is also clear that the Jastrow and normal mode trial wavefunctions converge much more quickly than that of the unity trial wavefunction. For this reason, in all further cases we will only compare the effects of using the Jastrow and normal mode trial wavefunctions.

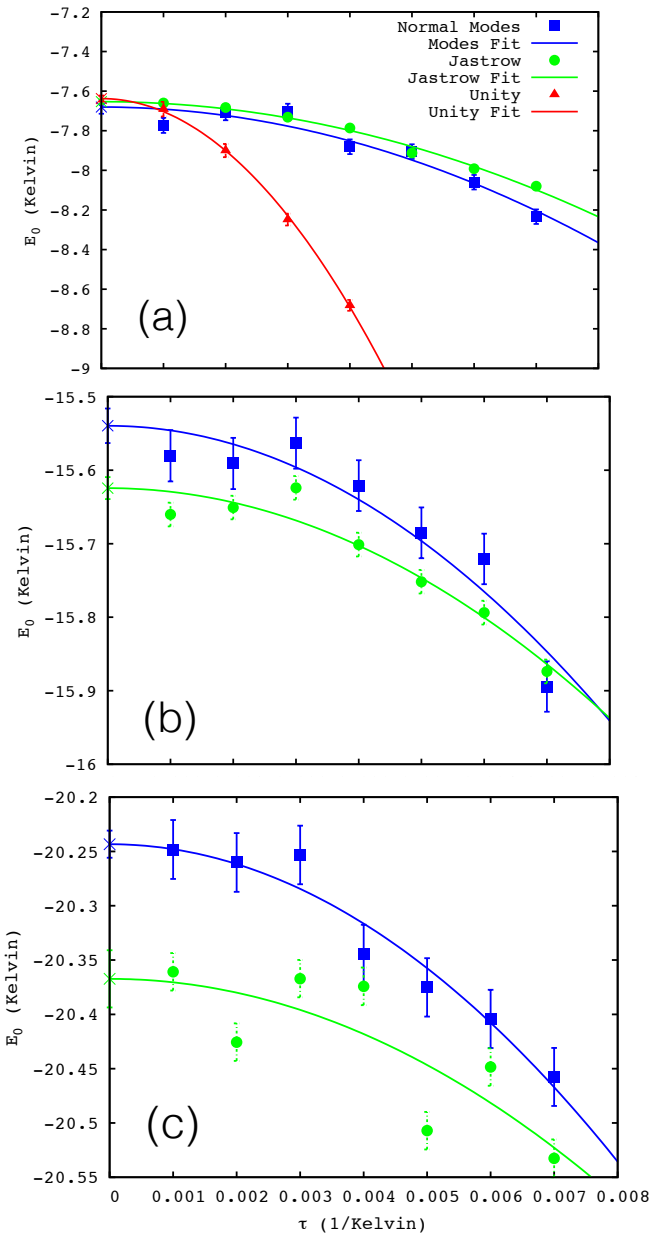

Figure 5.5: Total energy convergence of (a) $(pH_2)_4$, (b) $(oD_2)_4$, and (c) $(pT_2)_4$ with parameter τ using various trial wavefunctions.

To gain further insight, we compare the radial distributions of the end bead for each of the trial wavefunctions using their converged β value and $\tau = 0.001$ K^{-1} in Fig. 5.6 (a) for $(pH_2)_4$. We define our radial distribution, g(r), as:

$$4\pi \int dr g(r) r^2 = N - 1, \tag{5.13}$$

where N is the number of particles in the system.

Using the reference distribution of the unity trial wavefunction, we notice that the effects of adding the Jastrow wavefunction and introducing the hard-core repulsion term shifts the distribution to the right. However, using the normal mode trial wavefunction, the distribution of the end bead is very peaked and much less broad indicating the strength of this trial wavefunction to attempt to enforce a fixed geometry.

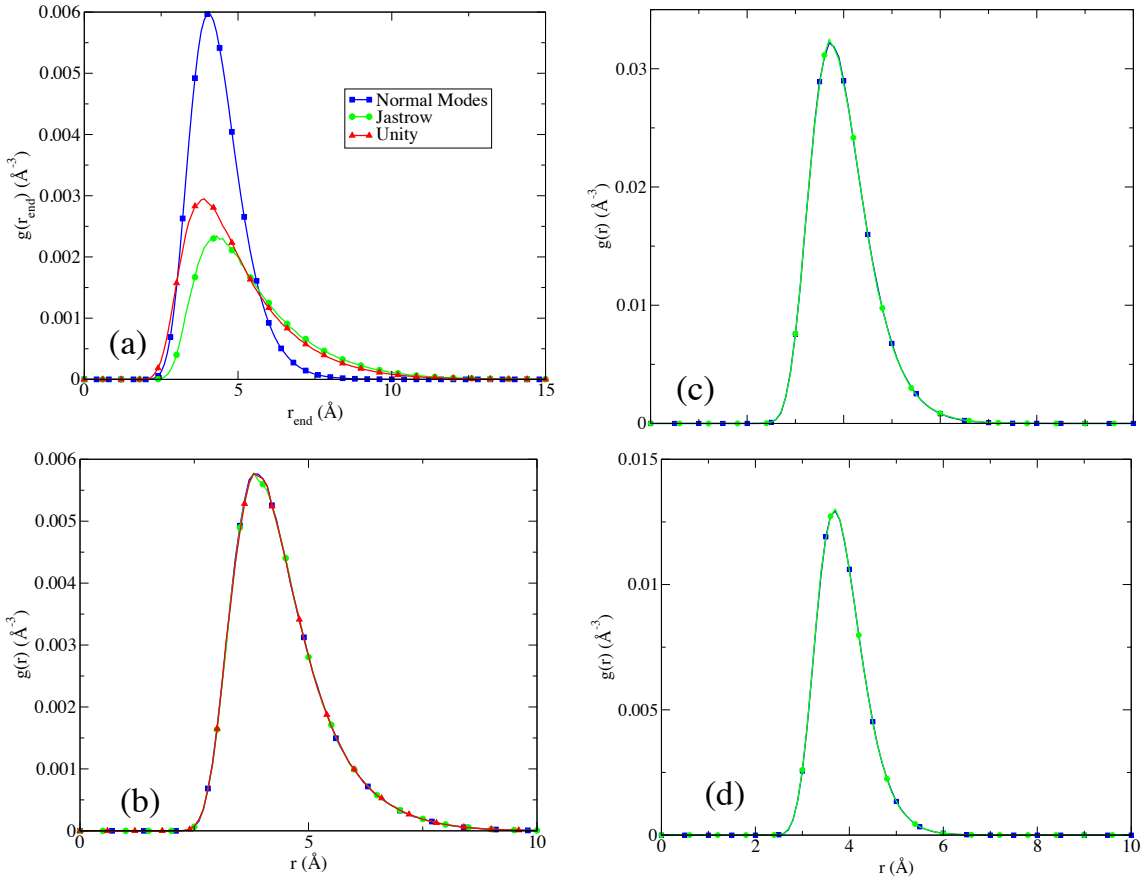

Figure 5.6: (a) Radial distribution of the end bead for $(pH_2)_4$ and (b)-(d) ground state g(r)s for $(pH_2)_4$, $(oD_2)_4$, and $(pT_2)_4$ respectively using various trial wavefunctions.

It is important to determine if the trial wavefunction has any biasing effects towards ergodicity of our simulation. Since the trial wavefunction is essentially an additional potential term or force acting on the end beads, it introduces additional ergodic properties on the system on top of those due to the interaction potential that could affect the efficiency of the sampling in our simulation. To ensure that no ergodicity problems occur in our

simulation, due to our trial wavefunction, we compare the middle bead radial distributions for each of the trial wavefunctions, since the middle bead radial distribution reflects the true ground state distribution of the system.

In reference to Fig 5.6 (b), the middle bead radial distributions using all three trial wavefunctions overlap, indicating that there is no effect on the distributions when varying the trial wavefunctions used in this study. We extend this study to two heavier isotopes, to determine if changing the mass affects the effectiveness of the trial wavefunctions. Once again, we perform our total energy convergence with the parameter β, represented in Figs. 5.4 (b) for $(oD_2)_4$ and (c) for $(pT_2)_4$. The values that we choose are summarized in Table 5.1. For both of these systems, it should be noted that there is a slight difference in the total energy of approximately 0.1 K^{-1} when using the different wavefunctions at the β values we choose. This could be due to the total energy not quite being converged with respect to β.

We now look at the total energy convergence with τ, shown for both the Jastrow and normal mode wavefunctions in Fig. 5.5 (b) for $(oD_2)_4$ and (c) for $(pT_2)_4$. The results are interesting. Unlike for $(pH_2)_4$, where the total energy converges to the same value for all wavefunctions in the limit of $\tau = 0$, we see a difference of up to 0.1 K^{-1} in the zero-tau limit for the heavy isotopes. This could be due to the total energy convergence with β, as mentioned above. We also compare the radial distribution functions for $(oD_2)_4$ and $(pT_2)_4$

in Fig 5.6 (b) and (c), respectively, and we obtain the same distribution regardless of which trial function is used.

Shown below in Table 5.2 are the specific and generic Lindemann criteria, where these values use the pair distance for the end beads, which represents the actual effect of the trial wavefunction, and the middle bead, which represents the true pair distance between molecules.

Table 5.2: Lindemann values for all isotopes and wavefunctions for cluster size $N = 4$

System	Property	Wavefunction	Lindemann Value (Specific)	Lindemann Value (Generic)
$(pH_2)_4$	End Bead	Unity	0.351	0.351
		Jastrow	0.337	0.337
		Normal Mode	0.215	0.215
	Middle Bead	Unity	0.255	0.255
		Jastrow	0.255	0.256
		Normal Mode	0.254	0.254
$(oD_2)_4$	End Bead	Jastrow	0.252	0.252
		Normal Mode	0.159	0.159
	Middle Bead	Jastrow	0.174	0.174
		Normal Mode	0.171	0.171
$(pT_2)_4$	End Bead	Jastrow	0.214	0.214
		Normal Mode	0.137	0.137
	Middle Bead	Jastrow	0.144	0.144
		Normal Mode	0.144	0.144

If we look at the Lindemann values for the middle bead of each system and wavefunction, the Lindemann specific agrees with the Lindemann generic for all systems. Based on the criterion above, all isotopes in the ground state are liquid-like. It is also interesting to look at the effect of the trial wavefunction on the system as well, reflected by the end bead pair distances. Looking at the Lindemann values for the end bead of each system and

wavefunction, the Lindemann specific values agree with the Lindemann generic values for all systems as well. This shows that all of our trial wavefunctions reflect the appropriate state of the cluster, allowing for relaxation to the ground state wavefunction.

5.4.2 Extending to a floppy structure : $N = 8$

Shown in the top panel of Fig. 5.7 is the total energy convergence with respect to the relaxation parameter β for the $(oD_2)_8$ system using both the Jastrow and normal mode wavefunctions. We can comfortably use these β values, shown in Table 5.1, for all three isotopes based on the analysis done in the previous section. We notice that the isotope effect on β was very small when using the Jastrow wavefunction. For the normal mode integrator, we notice that the isotope effect on β was small when moving from pH_2 to oD_2. When extending to pT_2, we notice that it takes a smaller β to converge the total energy, so the larger β obtained from the oD_2 study should also be converged.

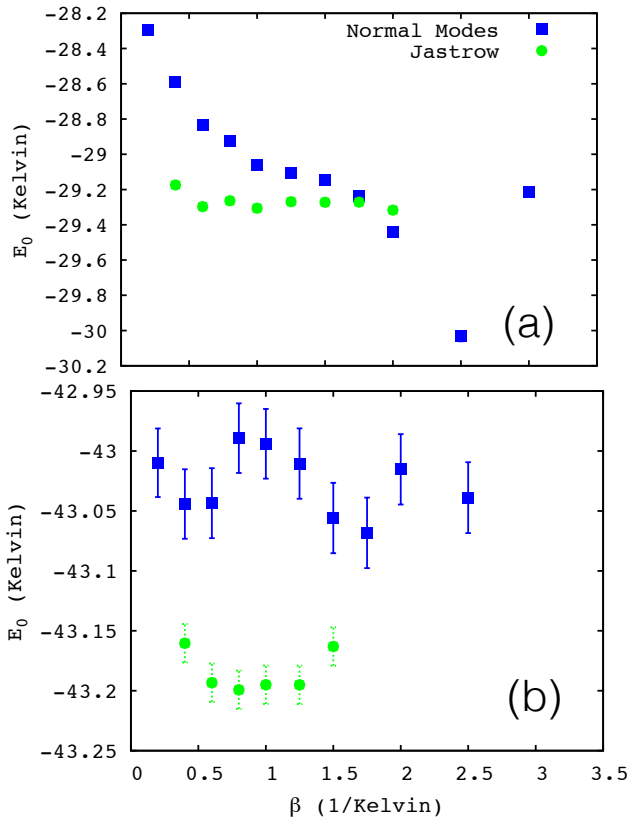

Figure 5.7: Total energy convergence of (a) $(oD_2)_8$ and (b) $(oD_2)_{13}$ with relaxation parameter β using various trial wavefunctions. The converged β values are used for parahydrogen and paratritium systems as well.

We now look at the total energy convergence with respect to τ. For the pH_2 cluster, shown in Fig. 5.8 (a), we see a significant difference in the total energy in the τ=0 limit between the Jastrow and normal mode wavefunctions.

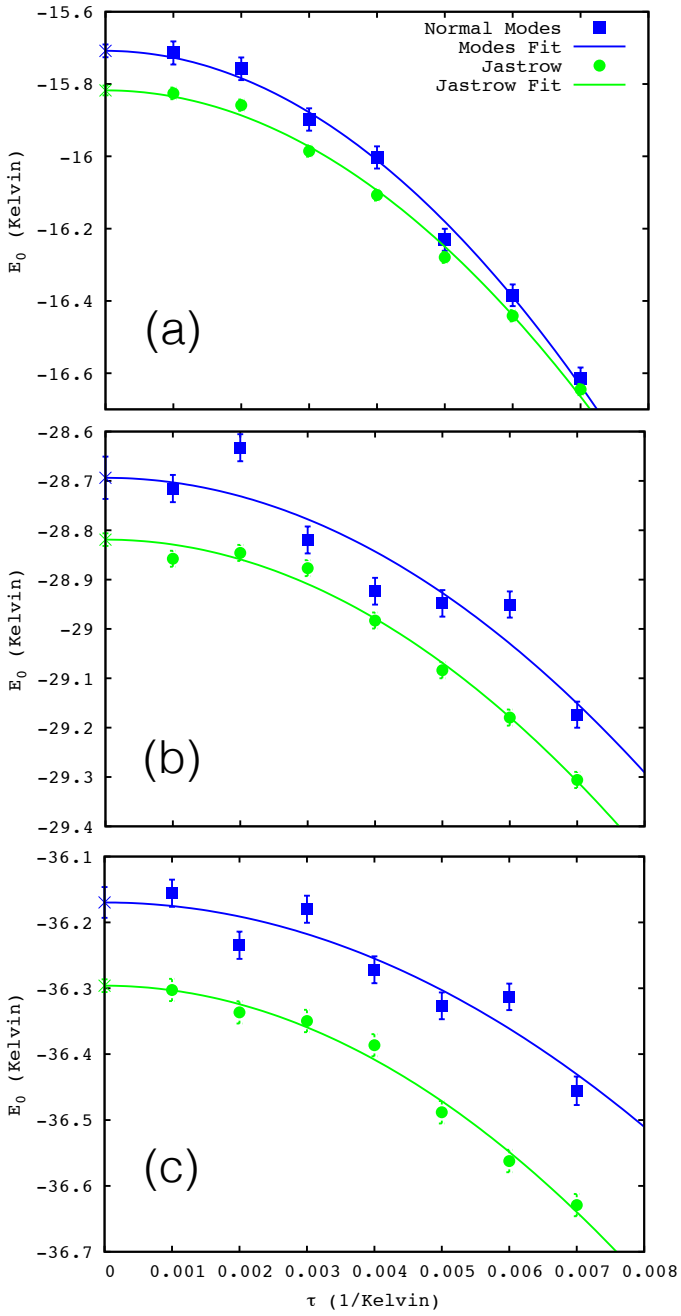

Figure 5.8: Total energy convergence of (a) $(pH_2)_8$, (b) $(oD_2)_8$, and (c) $(pT_2)_8$ with parameter τ using various trial wavefunctions.

As before, the normal mode wavefunction gives a higher total energy. This phenomenon is mimicked for the heavier isotopes as well, shown in (b) for oD_2 and (c) for pT_2. When using the normal mode wavefunction, we must specify an equilibrium geometry. These geometries were chosen from the classical Lennard-Jones structures for atoms. However, for the cluster size of $N = 8$, the global minimum geometry is a pentagonal bipyramid with a cap and it is clear that not every molecule is equivalent nor is it in an especially stable configuration (magic number cluster). Looking at the radial distribution functions, $g(r)$, for the three isotopes shown in Fig. 5.9, the radial distribution functions are slightly more peaked when using the normal mode wavefunctions than when using the Jastrow wavefunction. This again is a sign that the normal mode wavefunction is slightly biased towards the geometry we specified.

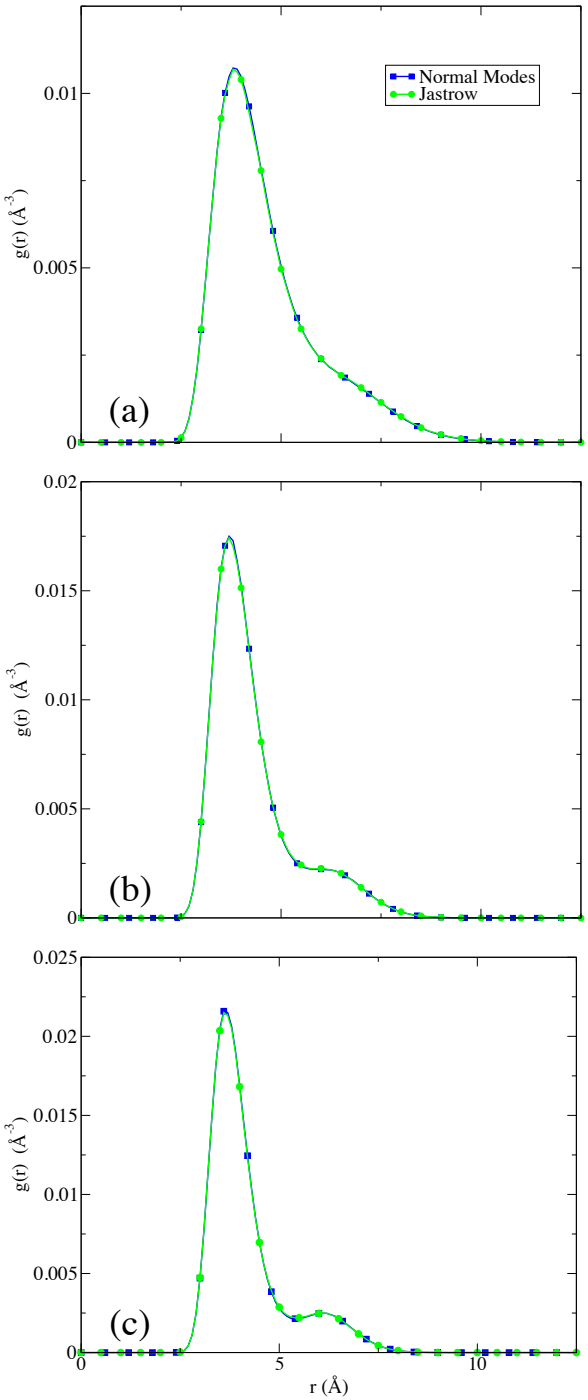

Figure 5.9: Ground state g(r) of (a) $(pH_2)_8$, (b) $(oD_2)_8$, and (c) $(pT_2)_8$ using different trial wavefunctions.

We can once again look to the Lindemann values for an explanation, shown in Table 5.3. The trend is the same for all isotopes.

Table 5.3: Lindemann values for all isotopes and wavefunctions for cluster size $N = 8$

System	Property	Wavefunction	Lindemann Value (Specific)	Lindemann Value (Generic)
$(pH_2)_8$	End Bead	Jastrow	0.335	0.335
		Normal Mode	0.153	0.290
	Middle Bead	Jastrow	0.298	0.298
		Normal Mode	0.271	0.296
$(oD_2)_8$	End Bead	Jastrow	0.299	0.300
		Normal Mode	0.122	0.280
	Middle Bead	Jastrow	0.275	0.275
		Normal Mode	0.204	0.275
$(pT_2)_8$	End Bead	Jastrow	0.284	0.287
		Normal Mode	0.108	0.276
	Middle Bead	Jastrow	0.264	0.268
		Normal Mode	0.164	0.269

By observing the Lindemann values of the end bead for the two wavefunctions, it is apparent that the specific and generic values are the same when using the Jastrow wavefunction, but different when using the normal mode wavefunction. This implies that the Jastrow wavefunction assumes a liquid-like system, but the normal mode wavefunction assumes a solid-like system. This explains why the radial distribution is slightly more peaked when using the normal mode trial wavefunction.

Since the total energy is significantly lower for each isotope when using the Jastrow wavefunction, we conclude that using the normal mode wavefunction presents an ergodicity problem in our simulation where the system is not exploring all of it's accessible states due to it's additional rigidity. The normal mode trial wavefunction tries to force a fixed

geometry on a floppy system and since our molecular dynamics simulation is of finite sampling time and finite projection times, the simulation does not efficiently sample the system when using that wavefunction. This is what is meant here as the ergodicity problem. However, since the normal mode wavefunction does have some overlap with the ground state wavefunction, with infinite sampling and projection time, we would be able to obtain the exact ground state properties. Thus, it is clear that for a floppy system, the Jastrow wavefunction must be used in favour of the normal mode wavefunction to obtain the correct estimate of energetic properties in a practical finite time simulation.

5.4.3 Extending to a magic number structure : $N = 13$

The final cluster size we choose is the magic number, $N = 13$, as the hydrogens tend to form a stable icosahedral shell. As in the previous section we look at the total energy convergence with β using deuterium for the two wavefunctions and use the converged β parameter for the other isotopes. The total energy convergence with β is shown in Fig. 5.7 (b) and the β value we choose can be found in Table 5.1.

We now look at the total energy convergence with respect to τ. For the pH_2 cluster, shown in Fig. 5.10 (a), we get very good agreement between the two wavefunctions in the limit of $\tau = 0$.

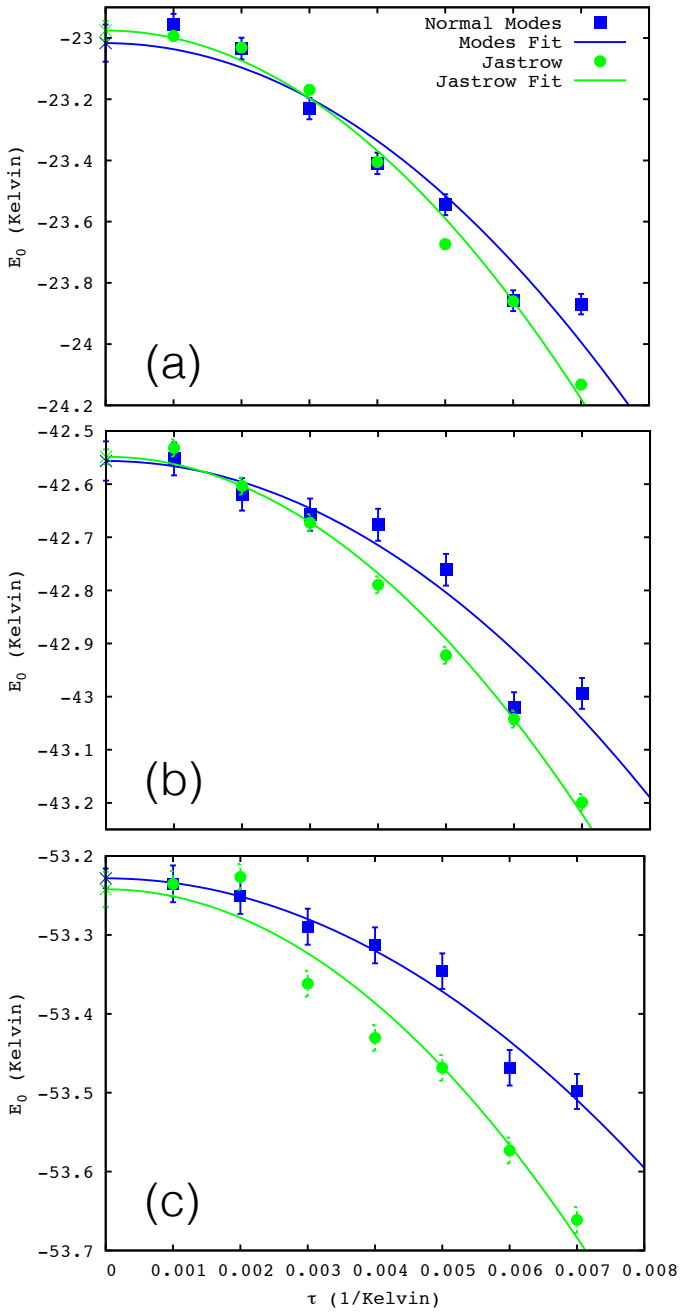

Figure 5.10: Total energy convergence of (a) $(pH_2)_{13}$, (b) $(oD_2)_{13}$, and (c) $(pT_2)_{13}$ with parameter τ using various trial wavefunctions.

We also get the same result when extending to the oD_2 cluster and pT_2 cluster, shown in (b) and (c) respectively. Also, for the heavier isotopes, the total energy converges much more quickly when using the normal mode trial wavefunction instead of the Jastrow wavefunction. This may be evidenced by the fact that for each value of τ, the total energy obtained using the normal mode trial wavefunction is closer to the exact ground state energy than the energy obtained using the Jastrow wavefunction. However, looking at the radial distribution function for $(pH_2)_{13}$ shown in Fig. 5.11 (a), we see that we obtain a more peaked radial distribution function when the normal mode trial wavefunction is used compared to when the Jastrow wavefunction is used.

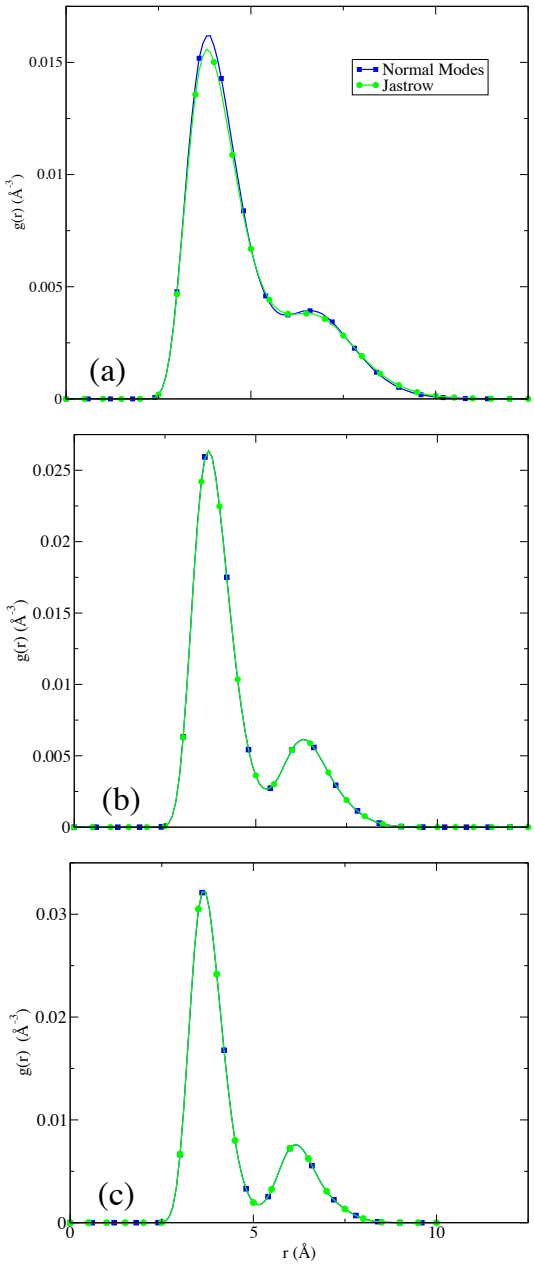

Figure 5.11: Ground state g(r) of (a) $(pH_2)_{13}$, (b) $(oD_2)_{13}$, and (c) $(pT_2)_{13}$ using different trial wavefunctions.

The more peaked g(r) obtained using the normal mode wavefunction indicates a more solid-like structure than the g(r) obtained using the Jastrow wavefunction. It is interesting that the energy does not seem to be affected by this. However, there is no trial wavefunction dependence in the radial distribution functions for $(oD_2)_{13}$ and $(pT_2)_{13}$ shown in (b) and (c) respectively. Both trial wavefunctions give the same radial distributions, which is expected.

However, to gain further insight let us compare the Lindemann values for these isotopes, shown below in Table 5.4.

Table 5.4: Lindemann values for all isotopes and wavefunctions for cluster size $N = 13$

System	Property	Wavefunction	Lindemann Value (Specific)	Lindemann Value (Generic)
$(pH_2)_{13}$	End Bead	Jastrow	0.334	0.336
		Normal Mode	0.127	0.287
	Middle Bead	Jastrow	0.299	0.303
		Normal Mode	0.169	0.297
$(oD_2)_{13}$	End Bead	Jastrow	0.185	0.301
		Normal Mode	0.103	0.280
	Middle Bead	Jastrow	0.120	0.282
		Normal Mode	0.119	0.282
$(pT_2)_{13}$	End Bead	Jastrow	0.153	0.291
		Normal Mode	0.093	0.277
	Middle Bead	Jastrow	0.103	0.278
		Normal Mode	0.103	0.278

Comparing the specific and generic end bead Lindemann values for $(pH_2)_{13}$ using the Jastrow wavefunction, they give approximately the same value, meaning that action of our wavefunction on the end bead assumes a liquid-like cluster. This is confirmed by the fact that the specific and generic middle bead Lindemann values give approximately the same value. However, when using the normal mode wavefunction, we see that the Lindemann

values differ between the specific and generic for both the end bead and the middle bead, indicating a more solid-like cluster. This confirms what we see in the radial distribution functions. In the case of $(pH_2)_8$, this leads to an ergodicity problem when using the normal mode wavefunction and to differences in the converged total energy. However, in this case, we recover the same ground state energy regardless of which wavefunction is used. For both of the heavier isotopes, the specific and generic Lindemann values for the end beads give different values, implying the action of both wavefunctions on the end bead predict a solid-like cluster. Like the cluster size of $N = 4$, using the Jastrow and normal mode wavefunctions give the same middle bead Lindemann values for both the specific and generic, though the Lindemann specific values differ from the generic values since the heavier isotopes for $N = 13$ are solid-like.

5.5 Expanding the Cluster Range

Now that we have determined a method to obtain all the parameters for our parahydrogen, orthodeuterium, and paratritium clusters, we can calculate energetic and structural properties for all cluster sizes between $N = 4 - 19$ and $N = 33$. Based on the knowledge gained from the previous sections, we will use the Jastrow trial wavefunction since we can avoid any additional ergodicity problems in our simulation. Our parameters for all cluster sizes

will be those optimized for $N = 4$ as shown in Table 5.1, except we choose a timestep $dt =$ 5.0 fs, $\beta = 1.00$ K^{-1}, and $\tau = 0.003$ K^{-1} for all isotopes and all cluster sizes. We benchmark our LePIGS results for parahydrogen and orthodeuterium clusters against previous studies by Cuervo *et al.*[32, 33] who used PIGS-MC using the same Buck potential[50]. We also predict energetic and structural properties for paratritium clusters, since to our knowledge there has been no previous systematic study of calculating ground state properties of small molecular paratritium clusters. In reference to Fig. 5.12 (a), we see that our LePIGS energies agree very well with those from Cuervo *et al.*. It should be noted that our values are consistently approximately 0.1 K smaller than our benchmark which could be due to a difference in our chosen τ parameter. When we compare our deuterium cluster energies to our benchmark in Fig. 5.12 (b), we observe an interesting result. Our energies are larger than those of our benchmark, although we are using a larger τ parameter. We then predict the tritium cluster energies in (c).

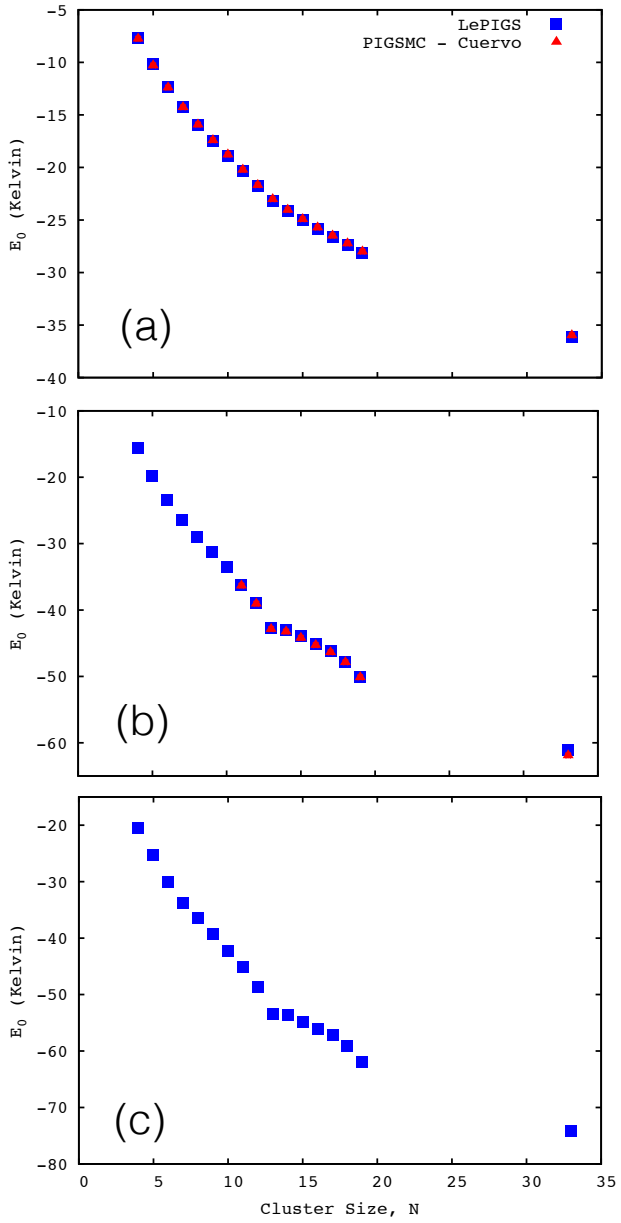

Figure 5.12: Comparison of our LePIGS energies for(a) $(pH_2)_N$, (b) $(oD_2)_N$, and (c) $(pT_2)_N$ for N = 4-19,33 benchmarked against those obtained from PIGS-MC.[32, 33]

We further investigate the ground state chemical potential curve for the parahydro-

gen, orthodeuterium, and paratritium clusters in Fig. 5.13, where we define our chemical potential

$$\mu(N) = E(N) - E(N-1) \tag{5.14}$$

where $E(N)$ is the total energy for a cluster of size N.

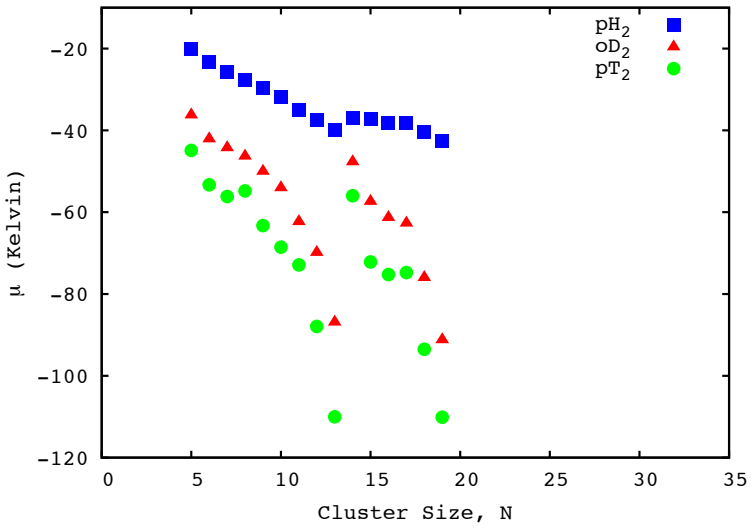

Figure 5.13: Chemical potential curves for $(pH_2)_N$, $(oD_2)_N$, and $(pT_2)_N$ for N = 5-19

It is interesting to note that the magic number clusters $N = 13$ and $N = 19$ become significantly more stable as you use heavier isotopes. We then compare the Lindemann values for parahydrogen in the top panel of Fig. 5.14.

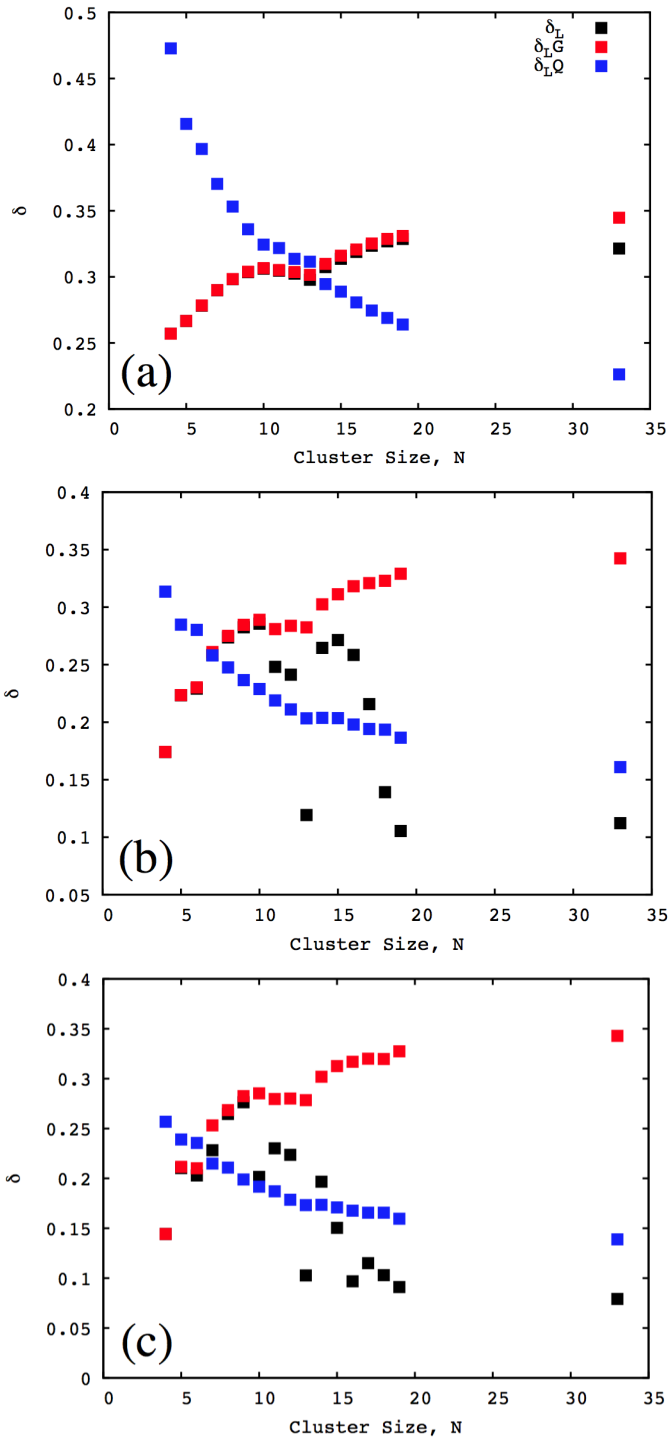

Figure 5.14: Specific, Generic, and Quantum Lindemann values for (a) $(pH_2)_N$, (b) $(oD_2)_N$, and (c) $(pT_2)_N$.

Based on the deviation of specific and generic Lindemann values, for cluster sizes up to $N = 19$, parahydrogen remains very liquid-like as the specific and generic Lindemann values do not differ, however they do appear to differ at $N = 33$. These Lindemann values are consistent with those of our benchmark[33]. We also observe the 'quantum' Lindemann decreases with cluster size, or that the relative deviation of the square of pair distances decreases. Based on the definition of this Lindemann value, the small clusters N <14 are solid-like as $\delta_Q > 0.3$ and all other cluster sizes co-exist as solid-like and liquid-like. This result is consistent with the DMC results of Guardiola and Navarro[56]. This is in complete contrast to the findings using our other Lindemann approach, where all parahydrogen clusters appear liquid-like.

When looking at the Lindemann values for orthodeuterium in the middle panel of that figure, we notice that differences between the specific and generic Lindemann values occur at a much smaller cluster size, $N = 11$, and cluster sizes $N = 13, 18, 19, 33$ are very solid-like. These values are also consistent with those from the published benchmark[33] except for a minor discrepancy in the specific Lindemann value for $N = 18$. When using the quantum Lindemann approach, we observe that the values of all of the small cluster sizes (except for N=4) correspond to the coexistence state and for N >15, the clusters appear solid-like. This result is also consistent with the DMC results of Guardiola and Navarro[56]. As before, this approach is inconsistent with that of the Lindemann specific and generic,

as the 'quantum' Lindemann does not predict a solid-like structure for the magic number cluster N=13.

We further show the Lindemann values for our paratritium clusters in the bottom panel. It is evident that deviations between the specific and generic Lindemann values occur at a small cluster size of $N = 7$ or $N = 10$. In addition, many more cluster sizes appear to be solid-like, including those $N > 14$. With the 'quantum' Lindemann approach, we observe the on-set of a solid-like cluster at N=9. Neither of the previous benchmarks did calculations for paratritium. In this case, it appears as though the two approaches yield the same conclusions on which cluster sizes are solid-like and liquid-like as opposed to the lighter isotopes where there is much discrepancy.

We also calculate the fluctuation of pair distances, a method similar to that of the 'quantum' Lindemann, shown in Figure 5.15. We note that the fluctuation of pair distances decrease monotonically, a similar trend to that of the 'quantum' Lindemann, but the two curves are not on the same scale. Future work is required to draw any true conclusions of solid-like and liquid-like behaviour.

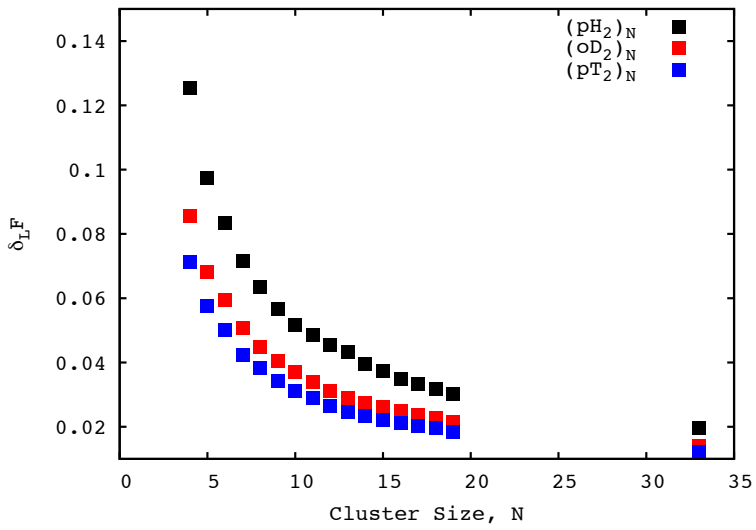

Figure 5.15: Fluctuation of pair distances for $(pH_2)_N$, $(oD_2)_N$, and $(pT_2)_N$.

We also calculate the Q_4 and Q_6 bond order parameters[59, 60], to potentially gain further insight into the phase or structure of the system. These parameters reflect the orientation of the nearest neighbours around each parahydrogen molecule, averaged over all molecules. Mandelshtam and co-workers[61, 62, 63, 64] have used such quantities to study structural transitions of Lennard-Jonnes clusters as a function of their temperature and size. Further, Chakravarty *et al.* have studied the correlation between these bond order parameters and the so called return distance of inherent structures[55]. They found that a solid behaviour is related to a negative correlation. Our results can be found in

the Appendix A.1. However, since our clusters are so small, the atoms are mostly on the surface and not fully connected in an ordered configuration. As a consequence, our calculated Q_4 and Q_6 values do not correspond to any of the theoretical values and our values monotonically decrease with increasing cluster size. Even for the $N = 13$ cluster size, we do not obtain the Q_4 and Q_6 values that correspond to an icosahedron, which is what we observe in our simulation. Due to this, we feel that our results do not provide additional insight to the phase or the structure of our clusters. Our results are consistent with previous work by Cuervo[65].

Finally, we show density profiles of all isotopes for many cluster sizes in Fig. 5.16. The definition of the density distribution, $\rho(r)$, is that of the $g(r)$ where r is the distance of each molecule to the centre of mass of the cluster, though normalized to the number of particles in the cluster, N. This is not the same definition used in the previous chapters.

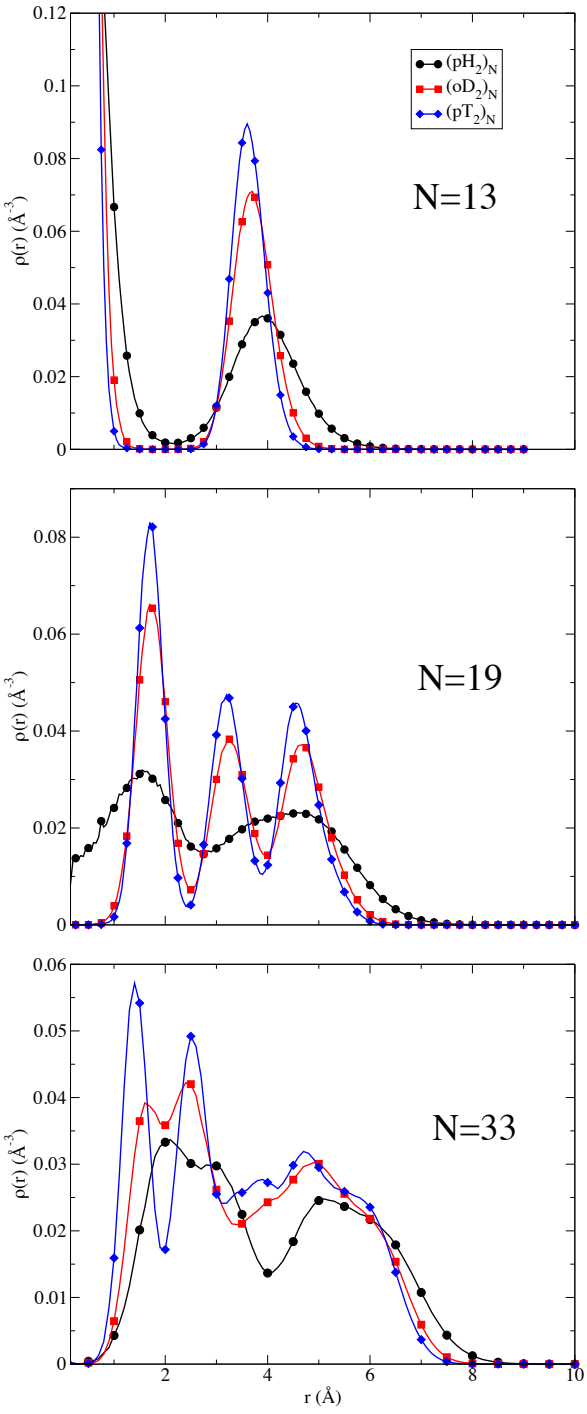

Figure 5.16: Density profile for hydrogen isotope clusters of size $N = 13, 19$, and 33.

We first present the density profile of cluster size $N = 13$. The results of parahydrogen and orthodeuterium agree very well with the benchmark[33] and it is clear that the density of molecules for orthodeuterium and paratritium clusters goes to 0 around $r = 2$ Å indicating distinct shells of molecules, whereby parahydrogen still remains fluid. For $N = 19$, our results also agree with our benchmark although we seem to have a slightly sharper first peak for orthodeuterium than their distribution. We also show the density profile for all isotopes for cluster size $N = 33$. We find that our density profile of parahydrogen agrees with Cuervo, however, for orthodeuterium we seem to have a shallow first peak whereas they predict a very sharp first peak. What is interesting is that paratritium catches this sharp first peak. We also present the radial distribution functions of all isotopes for specific cluster sizes in Fig. 5.17. It is clear that the radial distribution of deuterium and tritium are always much more peaked than that of hydrogen, further indicating they are consistently more ordered in structure.

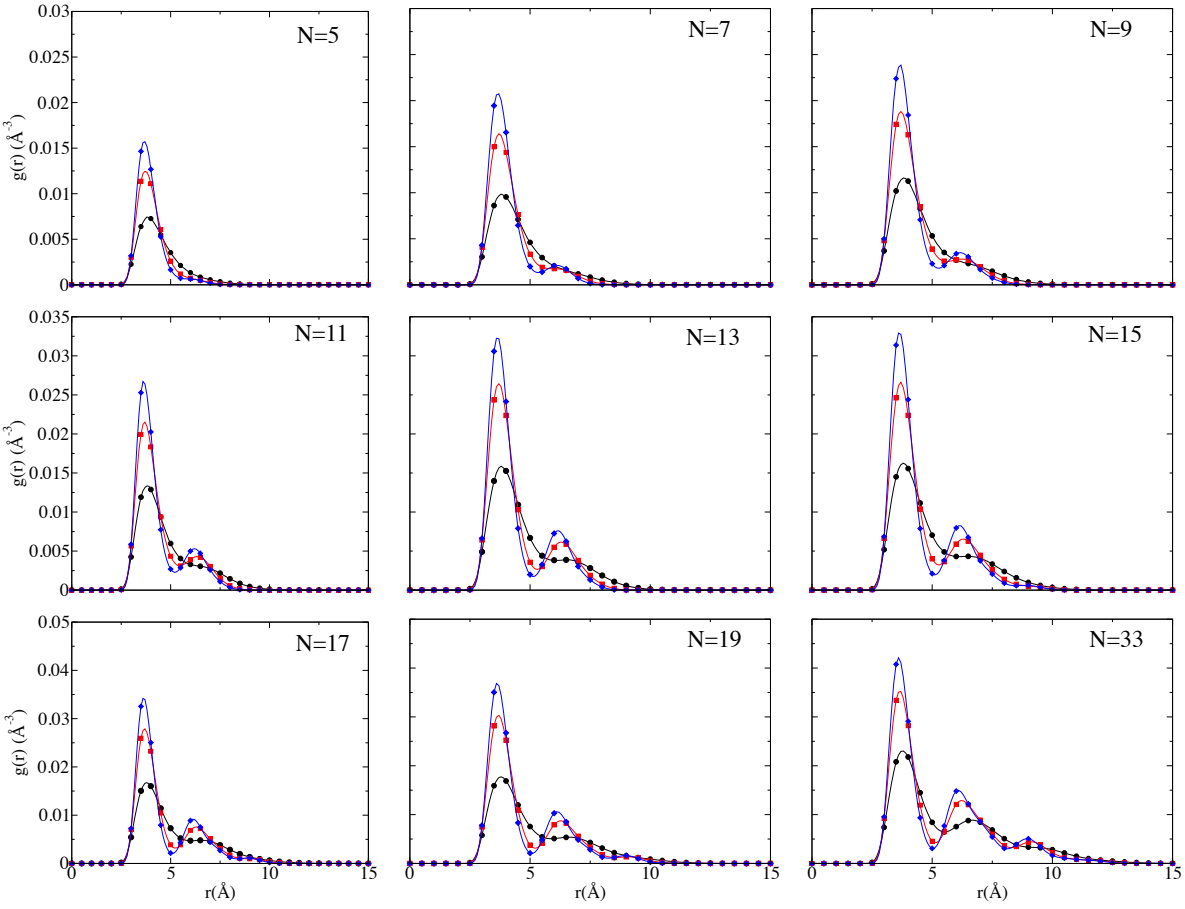

Figure 5.17: Radial Distribution for $N = 5, 7, 9, 11, 13, 15, 17, 19,$ and 33. Note the legend is the same as Figure 5.16

One final test we perform is looking at the validity of treating these systems in the harmonic approximation in Appendix A.2. We compare the harmonic energies, chemical potential, and radial distributions to our LePIGS results. We note that the absolute energies of the hydrogen clusters from the harmonic approximation are not valid, but as

we increase the mass, the harmonic chemical potential gives the correct trend and there is increased overlap between the true radial distributions and those generated using the harmonic approximation. We conclude that parahydrogen exhibits liquid-like behaviour, but as we increase the mass orthodeuterium and paratritium become much more solid-like.

5.6 Concluding Remarks

The outcome of this chapter is to demonstrate practical guidelines for optimizing parameters and adding a trial wavefunction to any system in LePIGS. The key steps are summarized below:

i) Choose your trial wavefunction. This requires adding a potential term to the Hamiltonian which acts on the end beads.

ii) Run a canonical ensemble simulation with no friction to obtain independent initial configurations.

iii) From each of those configurations, run a short microcanonical ensemble simulation to get a normalized autocorrelation function for the centroid pair distance. Obtain γ^0 by averaging those autocorrelation functions.

iv) Using this friction, run a short canonical simulation to obtain an autocorrelation function for the potential energy and total energy. Choose the longest autocorrelation time as

the time you must skip to obtain independent data points.

v) Using both the friction and correlation time, run a full canonical simulation using various time steps, dt. Plot the total energy vs. dt and choose the largest time step that results in the converged energy.

vi) Using all of the optimized parameters above, run canonical simulations using different β values choosing a reasonable τ value that is not too large. Plot the total energy vs. β and choose the smallest β that results in the converged energy.

vii) Using all of the optimized parameters above and the β value from the previous step, run canonical simulations using different τ values. Plot the total energy vs. τ and choose the largest τ that results in the converged energy. The exact ground state energy is the extrapolation of the total energy curve to $\tau = 0$, however, your converged energy should be very close to that value.

In this study, we determine which trial wavefunction is appropriate for certain systems, which allow us to maximize computational efficiency when calculating ground state properties. Our findings reflect that all of our trial wavefunctions have some finite overlap with the exact ground state wavefunction as we are able to obtain the same ground state properties for each trial wavefunction. However, we find that some trial wavefunctions are more suitable than others for different systems. Initially, we notice that the ground state energy converges with the imaginary time step τ much faster using the Jastrow and normal mode

wavefunctions than with the unity wavefunction for the $(pH_2)_4$ system, so the remainder of the studies only compare the Jastrow and normal mode wavefunctions. We note that all three wavefunctions gave the exact ground state properties for that system, as expected, but the trial wavefunctions only affect the convergence.

We note that for systems with a clear global minimum geometry (and high symmetry) like the $N = 4$ tetrahedron or a solid-like cluster like the heavier isotopes for the magic number $N = 13$ icosahedron, either the Jastrow or the normal mode trial wavefunction can be used. However, for floppy systems, like $N = 8$, only the Jastrow trial wavefunction should be used since it does not assume a rigid structure, the flaw of the normal mode trial wavefunction which leads to an ergodicity problem in our simulation. We also note, for systems that may be less quantum and increasingly solid-like, it would be tempting to use the normal mode wavefunction. However, knowledge of the global minimum is very important. If the true ground state wavefunction has little or no amplitude in the minimum chosen as the normal mode geometry, the ground state properties may not be obtained!

Furthermore, we use the Jastrow wavefunction and expand our system to a broad range of cluster sizes for all three isotopes and we benchmark our energies and structural properties for parahydrogen and orthodeuterium against previous studies done by Cuervo[32, 33]. We find that our energies and structural properties compare well for parahydrogen and orthodeuterium, although there are a few deviations. To our knowledge, there are no previous

ground state studies on clusters of molecular paratritium, so we predict the ground state energies and structural properties for these systems. Comparing the specific and generic Lindemann values, we find that paratritium exhibits solid-like behaviour at lower cluster sizes than that of parahydrogen and orthodeuterium. The density profiles and the ground state radial distribution functions are also much more peaked for paratritium than for the others, which is as expected since it is the heaviest isotope.

We wish to apply this work towards many different purposes. We wish to test other trial wavefunctions, including one that is a combination of the Jastrow and normal mode wavefunction. We feel this type of wavefunction would provide information on the structure from the normal mode portion, but the Jastrow portion will provide information as to whether or not the system is weakly bound. We also wish to look into methods to relieve ergodicity problems. One potential way of relieving the ergodicity problem is by increasing the projection time β, as we are able to obtain all ground state properties in the limit of $\beta \to \infty$. However, by increasing β, we increase the length of our path and the number of beads. In doing so, we introduce additional ergodicity and an increased computational performance problems. Another way is the use of enhanced sampling methods such as parallel tempering[66, 67] in order overcome some ergodicity issues.

We can also extend this methodology to perform ground state RPMD and extract approximate dynamical properties (note that the polymer would be an open chain reptile

rather than a ring). In fact, one group has already investigated calculating approximate real time correlation functions for systems in the ground state using the centroid formalism[68]. Additional formal work, however, needs to be done in order to obtain a ground state version of RPMD although Fig. 5.1 essentially contains the elements of a Reptile Molecular Dynamics (RMD) method.

Chapter 6

Application: First-Principles Prediction of the Raman shifts in parahydrogen clusters

6.1 Introduction

One method of testing the accuracy of an interaction potential between a chromophore with a solvent molecule is by calculating the vibrational Raman shift when a chromophore is surrounded by a cluster of these solvent molecules and comparing to experimental val-

ues[1]. For a parahydrogen chromophore, the vibrational shift corresponds to the change in vibrational energy of that molecule when it becomes part of a cluster of other parahydrogen molecules. The shift can also be defined as the difference between changes in the upper, $v_{\rm t} = 1$, and lower, $v_{\rm t} = 0$, vibrational energies when the chromophore is added to a cluster of size $N-1$ other molecules and represented in the following way:

$$\Delta\nu_{0,N} = \Delta E_{\rm cluster,N} - \Delta E_{\rm free} = E_{v_{\rm t}=1,N} - E_{v_{\rm t}=0,N} \ , \tag{6.1}$$

where $\Delta E_{\rm free}$ is the difference in vibrational energies of a free parahydrogen molecule, $\Delta E_{\rm cluster,N}$ is that difference when surrounded by a cluster of $N-1$ other pH_2 molecules. Since we deal with quantum mechanics and exchange is very important, $E_{v_{\rm t}=1,0,N}$ is the energy of the system when the *total* vibrational quanta of all the pH_2 monomers is either $v_{\rm t} = 1$ or 0.

Traditionally, for each cluster size, two simulations would be run: one with the chromophore in the ground vibrational state and the other with chromophore in the first vibrational excited state. By knowing the difference in energy of the free chromophore molecule, the vibrational shift can be calculated directly in the manner described in Equation 6.1.

[1]This chapter (including some of the figures) has been reprinted or adapted from a manuscript (N. Faruk, M. Schmidt, H. Li, R. J. Le Roy, and P.-N. Roy, "First Principle Prediction of the Raman shifts of parahydrogen clusters") submitted for publication and currently under revision in The Journal of Chemical Physics

However, a recent PIMC study[69] of CO_2 in a cluster of helium atoms showed very slow convergence of the statistical errors after increasing the number of simulation steps. Also, in the case of a pure parahydrogen cluster, when parahydrogen is the chromophore itself, there are exchange interactions which mean that treating one parahydrogen as 'distinct' is not quantum mechanically correct!

Using perturbation theory, the problem can be recast in a way that solves these problems. For our case, it requires only what will be known henceforth as the difference potential (ΔV), which is the difference between the potential energy surface of the ground vibrational state and the first vibrational excited state of the chromophore. We also only require one simulation where the interaction potential between parahydrogen molecules corresponds to the ground vibrational state[69]. From the simulation, we require what is known as a pair distribution function, $P(R)$ which is defined as :

$$\int dRP(R)R = 1 \ . \tag{6.2}$$

The equation to calculate the vibrational shift is below

$$\Delta\nu_{0,N} = (N-1)\int_0^\infty \Delta V\,(R)\,P_N\,(R)\,dR\,, \tag{6.3}$$

in which the scaling prefactor of $(N-1)$ accounts for the normalization. It is important to note that the Jacobian factor of $4\pi R^2$ is absorbed in our definition of the pair distribution function.

Since we run a simulation and obtain all pair distances for our cluster for each independent time, we can re-write the equation as :

$$\Delta\nu_{0,N} = \frac{1}{n}\sum_{i}^{n} \Delta V\left(R_{i,N}\right) \ , \tag{6.4}$$

in which n is the number of data points in the simulation trajectory for a particular cluster.

6.2 System Setup

To obtain our difference potential, we convert a 6-D potential by Hinde[70] to 1-D by first averaging over the ground state (or excited state) wavefunction and then using the adiabatic hindered rotor approximation; the details of which are not my work but can be found in our submitted paper[71].

We run LePIGS simulations and the pair distances we obtain were generated for comparison using a number of different $p\mathrm{H}_2$ interaction potentials. The Hinde[70] and Szalewicz[72] potentials are *ab initio* potential functions, while the Silvera-Goldman[73] and Buck[50] po-

tentials are empirical functions that were parameterized to conform to different types of experiments. For cluster sizes of $N = 4$ and above, a Jastrow-type trial wavefunction[32] is used in the simulations. Unfortunately, smaller clusters were too weakly bound and dissociated during simulation when this trial wave function was used. The simulation parameters were optimized using the $N = 4$ cluster, resulting in a β value of 1.00 K^{-1}, τ of 0.003 K^{-1}, time step of 5.0 fs, and a correlation time of 0.2 ps. The simulations were run for 20.0 ns.

6.3 Results and Discussion

Vibrational frequency shifts are calculated for clusters, generated from simulation using various interaction potentials and analyzed with two choices of the difference potential. These results are presented in Fig. 6.1 and in Table 6.1. The shifts in the top panel of the figure were obtained using the *ab initio* difference potential from Hinde[70] and those in the bottom panel were generated using an empirical Lennard-Jones difference-potential of Ref. [18].

Table 6.1: $\Delta\nu_0$ (cm^{-1}) obtained from pair distributions generated using different potentials and the two ΔV. Std. errors are ≤ 0.005 cm^{-1}. Experimental (Exp.) are provided for comparison purposes.

	$\Delta\nu_0$, *ab initio* ΔV						$\Delta\nu_0$, empirical ΔV			
N	Exp.[18]	Hinde	Szalewicz	Buck	S-G	Buck DMC[18]	Hinde	Szalewicz	Buck	S-G
4	-1.251	-1.19	-1.203	-1.118	-1.092	-1.255	-1.197	-1.223	-1.261	-1.180
5	-1.594	-1.517	-1.527	-1.417	-1.385	-1.597	-1.527	-1.566	-1.622	-1.515
6	-1.910	-1.785	-1.800	-1.683	-1.648	-1.904	-1.812	-1.863	-1.923	-1.809
7	-2.136	-2.016	-2.025	-1.878	-1.852	-2.141	-2.045	-2.098	-2.175	-2.043
8	-2.350	-2.206	-2.217	-2.059	-2.028	-2.344	-2.248	-2.315	-2.392	-2.257
9		-2.369	-2.384	-2.219	-2.181		-2.431	-2.493	-2.588	-2.447
10		-2.542	-2.553	-2.367	-2.335		-2.611	-2.674	-2.774	-2.628
11		-2.729	-2.750	-2.545	-2.516		-2.800	-2.887	-2.990	-2.821
12		-2.965	-2.981	-2.752	-2.702		-3.028	-3.109	-3.222	-3.036
13	-3.140	-3.236	-3.276	-3.010	-2.922	-3.330	-3.297	-3.405	-3.507	-3.270
14		-3.279	-3.289	-3.043	-3.002		-3.352	-3.449	-3.573	-3.369
15		-3.334	-3.339	-3.111	-3.064		-3.425	-3.519	-3.657	-3.452
16		-3.404	-3.404	-3.160	-3.122		-3.495	-3.595	-3.736	-3.529
17		-3.453	-3.472	-3.209	-3.175		-3.566	-3.665	-3.802	-3.606
18		-3.532	-3.550	-3.284	-3.250		-3.651	-3.753	-3.899	-3.694
19		-3.668	-3.683	-3.378	-3.333		-3.787	-3.892	-4.024	-3.793
33	-4.390	-4.784	-4.811	-4.423	-4.169	-4.870	-4.941	-5.085	-5.183	-4.794

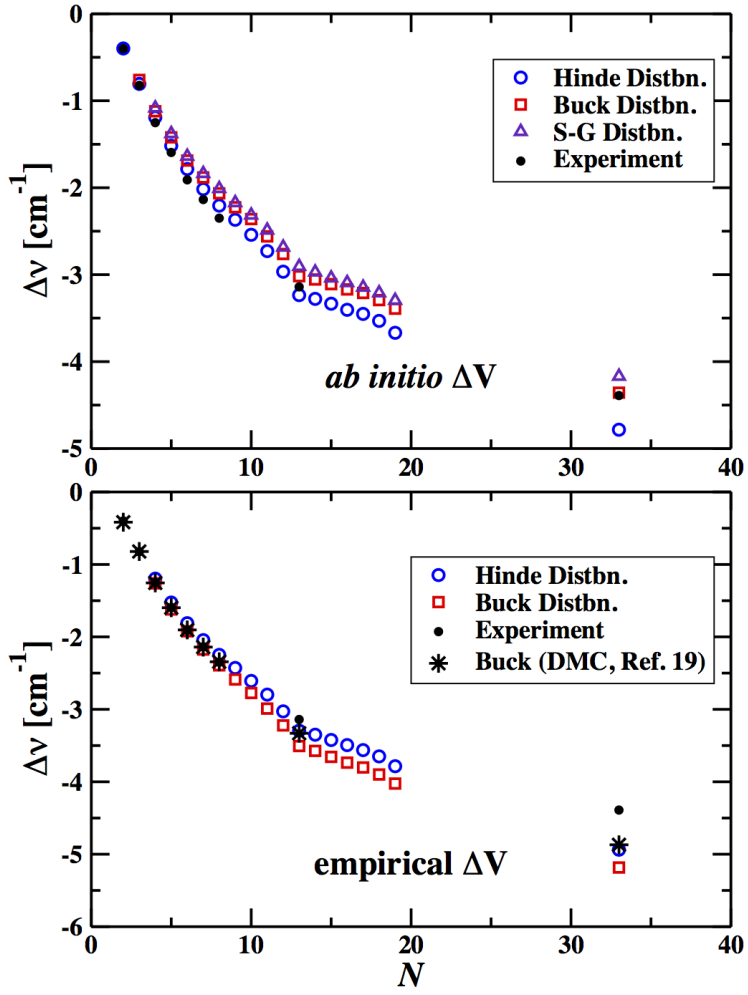

Figure 6.1: Vibrational frequency shifts as a function of cluster size calculated using pair distributions generated from different simulation potentials (see key). These shifts were calculated in two different ways: using an *ab initio* difference-potential (top figure) or an empirical difference-potential (bottom figure).

When comparing the effect of varying the difference potential, we notice for the small cluster sizes, $N < 9$, calculating the shifts using the empirical Lennard-Jones difference po-

tential gives closer results to experiment than using the *ab initio* Hinde difference potential. This is most likely due to the fact that the empirical difference potential was fitted to two parameters in order to reproduce the experimental shifts for $N < 9$. This can be observed by how closely their DMC calculations match up to experimental values. However, when we compare the shifts for the larger cluster size outside of their test set for $N = 13, 33$, the shifts obtained using the *ab initio* difference potential give better agreement than using the empirically based one.

We can now observe how the shifts using the various interaction potentials match up to experiment by looking at the top panel of Figure 6.1 or just the results when using the *ab initio* difference potential. For clarity, the Szalewicz potential is not plotted, however it follows the same trend as the Hinde potential. It is clear that for the small cluster sizes $N < 9$, the Hinde *ab initio* interaction potential gives much better agreement to experiment than the two empirical interaction potentials (Buck and Silvera-Goldman). This trend appears to continue to $N = 13$ as well, although the shifts obtained using empirical potentials do give better agreement to experiment than they had for the smaller clusters. What is interesting to note is that for the large cluster $N = 33$, using the Buck potential is very accurate in predicting the vibrational shift, followed by the Silvera-Goldmann potential. The shifts obtained from the *ab initio* potentials differ by much more. We speculate that this is due to the fact that the *ab initio* potentials are calculated pair

potentials and in larger clusters, many-body effects become much more important, which explains why the Silvera-Goldmann potential gives very good agreement. However, why the Buck potential gives the best agreement is unclear, as it is also a pure pair potential, albeit derived empirically from experimental scattering cross-section data.

We also note that the shift 'curve' has a kink at $N = 13$. This is due to the filling of the first solvation shell of the icosahedral geometry and the fact that the peak of the pair distribution function for the first solvation shell lies nearly directly over the potential minimum of all of the interaction potentials. This means that each molecule added that contributes to that first solvation shell makes a big impact in the vibrational shift, thus resulting in the 'curve' being steeply sloped. As the molecules are added farther away from the target molecule, that region of the pair distribution function will overlap with the less negative portion of the potential energy surface and thus the slope will be smaller.

In reference to the bottom figure of 6.1 when using the empirical Lennard-Jones difference potential, the same shifts are not obtained when using the pair distribution function of DMC (through Eq. 6.3) and the statistical calculation using the pair distances from the LePIGS simulations (through Eq. 6.1). There are many reasons why this may be occur. First, the pair distribution functions of Ref. [18] are noisy, which may not be representative of the true pair distribution function of an infinitely long simulation. Also, we noticed that the values we calculated for our shifts were strongly dependent on how we 'binned'

our data; factors such as 'binning' to the left, center, or right and also how many bins we chose. Additionally, calculating the uncertainty in the shift value is very difficult when integrating data over a histogram. For these reasons, we calculate statistically the average shift as it is less variable and statistical errors are straight forward to calculate.

6.4 Conclusions

In this chapter, we use the distribution of pair distances of a LePIGS simulation using various interaction potentials to predict vibrational Raman shifts of pure parahydrogen clusters and compare the results to experiment. We first comment on the use of an empirical, fitted difference potential compared to a purely *ab initio* difference potential. We find that only in the cluster size regime that the empirical difference potential was fitted, do their results give better agreement to experiment than using the *ab initio* Hinde difference potential. We also compare the shifts obtained from various interaction potentials using the Hinde difference potential. We find that for small cluster sizes below the first solvation shell, the Hinde and Szalewicz *ab initio* interaction potentials give better agreement to experiment than the Buck and Silvera-Goldmann empirical potentials. However for cluster size $N = 33$, the trend is reversed. We speculate this is due to the Silvera-Goldmann potential incorporating many-body effects. However, we are unsure why the Buck potential

is so accurate, as it is a pure pair potential.

The difference between our LePIGS result and a previous DMC result is discussed. We choose to calculate our vibrational shift using a purely statistical way, instead of integrating over a pair distribution function. We found that the way we 'bin' the histogram causes significant variation in results and the statistical errors associated with that are more difficult to calculate. We also note the trivial difference of changing the interaction potential in the LePIGS code can lead to a wealth of information in merging theory with experiment.

Chapter 7

Conclusions and Future Work

7.1 Concluding Remarks

Overall, this thesis accomplishes a number of goals. In Chapter 3, we benchmark our PIMD code against a basis set calculation for a system of one parahydrogen molecule interacting with a fixed water molecule. We confirm that the energy converges to the benchmark value within statistical error and the convergence behaviour is as expected. We also successfully benchmark our structural properties. We then benchmark our PIMD code against PIMC for a cluster of parahydrogen molecules interacting with either a fixed or a translating, rigid water molecule. We show that our PIMD energies agree very well with those obtained from PIMC. We then extend our study to the zero temperature limit, where our system is in the

nuclear ground state, in Chapter 4. We again benchmark our new method against a basis set calculation for a system of a parahydrogen molecule interacting with a fixed water, but this time for our Langevin equation Path Integral Ground State (LePIGS) code. We are again able to reproduce the energy and structural properties of the benchmark within statistical error. We also recover the correct convergence behaviour of our energy with both β and τ consistent with theory. However, we only use a trial wavefunction $\psi_T = 1$. In Chapter 5, we study the effect of three trial wavefunctions on parahydrogen clusters and their isotopologues. Most importantly, we outline a systematic method to optimize all parameters of a LePIGS simulation. We study very thoroughly the solid-like and liquid-like nature of three specific hydrogen cluster sizes $N = 4, 8$, and 13 using various Lindemann criteria and investigate the effect the trial wavefunctions have on the system. We find that the unity trial wavefunction is inefficient compared to the Jastrow and normal mode trial wavefunctions. For equivalent molecule structures or magic number clusters, both the Jastrow and normal mode trial wavefunctions are acceptable to use, but for a floppy system, the normal mode trial wavefunction can introduce an ergodicity problem and so it is preferable to use the Jastrow. We use the Jastrow wavefunction as we extend to larger cluster sizes and investigate the solid- and liquid-like behaviour. Our energies and structural properties of parahydrogen and orthodeuterium clusters agree with previous studies which used the PIMC analog of PIGS. We also predict energy and structural

properties of neutral paratritium clusters, as we have found no previous research regarding these systems. In Chapter 6, we predict vibrational Raman shifts using the distribution of pair distances obtained from LePIGS simulations, which can be compared directly to experiment. We calculate these shifts for a range of parahydrogen clusters using various interaction potentials. We find that for small cluster sizes less than the first solvation shell, the shifts obtained using *ab initio* interaction potentials of Hinde and Szalewicz give better agreement to experiment than empirical potentials of Buck and Silvera-Goldmann, however the trend is reversed for larger clusters.

There is much more research that can be done. We outline the two most pertinent directions we can take our research. We wish to investigate hydrogen storage in clathrate hydrates for low temperatures (using PIMD) and in the zero-temperature limit (using LePIGS). We also wish to further develop our codes by introducing nuclear exchange into PIMD and rotation into both PIMD and LePIGS, since PIMC currently has these features. Introducing nuclear exchange is necessary when studying the low temperature behaviour of bosons and we can attempt to merge low temperature results of PIMD with ground state results of LePIGS. Rotation is equally important, since we are eliminating a potentially important degree of freedom when performing our simulations.

7.2 Application to hydrogen storage in structure II clathrate hydrates

Hydrogen is used as a source of energy and due to the cost of cooling pure hydrogen to liquid temperature for storage purposes, water molecules have been suggested as a clean energy carrier in the form of clathrate hydrates or water cages[74]. Several experiments and ground state calculations have been performed on what is referred to as a structure II clathrate hydrate. The unit cell is made up of two different size clathrates (shown in Figure 7.1) and many experiments and calculations have shown that the smaller of the two may hold up to two hydrogen molecules and the larger clathrate may hold up to four hydrogen molecules with the theoretical maximum mass percent of hydrogen at 5%[75]. However, there have been conflicting results in both experiment[75, 76, 77, 78] and theory[79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89] as to the exact number of hydrogen molecules that can be stored.

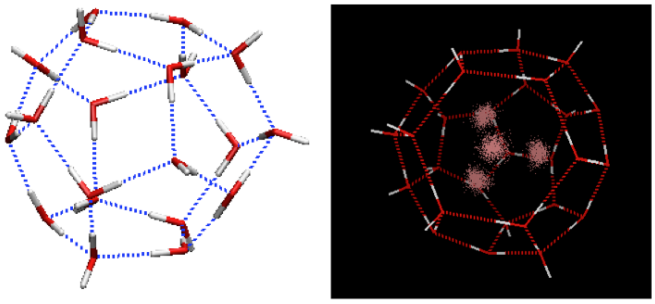

Figure 7.1: Empty small 5^{12} cage (left) and large $5^{12}6^4$ cage occupied by four pH_2 molecules (right).

Our goal would be to determine finite-temperature and ground state properties of parahydrogen stored in various sized clathrates through molecular dynamics simulations, and to compare these results to previous studies. By calculating the distributions of parahydrogen with the centre of the cage, we may determine where they lie and how localized they are inside the cage. For cases of multiple occupancy, we can compare the distribution of the hydrogens in the cage to that of a pure hydrogen cluster to determine if they are more compressed. In addition, the energy can give us information on how stable they are and, for multiple parahydrogens, the chemical potential by number could give insight as to how many hydrogens within a cage is the most energetically stable.

7.3 Other software development

There is also still a lot of work involved in the software side of molecular dynamics. First, rigid-body rotation is able to be sampled using PIMC, but constructing the Hamiltonian and using molecular dynamics to simulate is more difficult. Another important aspect of low-temperature dynamics is nuclear exchange. At low enough temperature, it has been suggested that pH_2 molecules might exhibit superfluidity. Including nuclear exchange is necessary to observe such behaviour in a simulation[90]. This will require another significant coding effort[29].

APPENDIX

Appendix A

Supporting Information on solid- vs liquid-like behaviour of hydrogen clusters

A.1 Bond Order Parameters

One method to potentially identify the structure and/or phase of a cluster is by calculating certain bond order parameters[60, 59]. Bond order parameters for a specific atom are quantified by the symmetry of the orientation of nearest neighbour bonds. The bond

length itself is not important, however there is a cutoff bond length in order to ensure an atom is a nearest neighbour. The cutoff values we choose are the minimum between the first two peaks of the radial distribution function.

The order parameter for one bond $\mathbf{r}$, or the local bond order parameter, is defined as

$$Q_{lm}(\mathbf{r}) = Y_{lm}(\theta(\mathbf{r}), \phi(\mathbf{r})) \ , \tag{A.1}$$

where Y_{lm} is a spherical harmonic, θ is the polar angle, and ϕ is the azimuthal angle.

The global bond order parameter is an average over all the bonds, N_b.

$$\bar{Q}_{lm} = \frac{1}{N_b} \sum_{bonds} Q_{lm}(\mathbf{r}) \tag{A.2}$$

The order parameters must be invariant with respect to rotations of the reference frame, so the following equation is required

$$Q_l = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |\bar{Q}_{lm}|^2} \tag{A.3}$$

The table below shows the Q_4 and Q_6 values that correspond to the liquid phase, icosahedral surface structure (Ih surface), and other specific geometries such as face-centered

cubic (fcc), hexagonal close-packed (hcp), simple cubic (sc), and body-centered cubic (bcc).

Table A.1: Geometries Associated with Bond Order Parameters. Data from Reference [60]

Geometry	Q_4	Q_6
fcc	0.19094	0.57452
hcp	0.09722	0.48476
sc	0.76376	0.35355
bcc	0.08202	0.50083
Ih	0	0.66332
liquid	0	0
Mackay Ih bulk	0	0.19961
Mackay Ih surface	0	0.20729

We calculate the Q_4 and Q_6 for each configuration sampled in our simulations. The cutoff values we use are shown below in Table A.2. Note that for $N = 4$, all atoms are nearest-neighbours to each other, so we effectively do not have a cutoff point. For all others, we use the minimum between the first two peaks of the radial distribution function.

Table A.2: Cutoff values (Å) used to calculate bond order parameters.

N	pH_2	oD_2	pT_2
4	20.0	20.0	20.0
5	6.0	5.5	5.3
6	6.0	5.5	5.3
7	6.0	5.5	5.3
8	6.0	5.5	5.3
9	6.0	5.4	5.2
10	5.8	5.3	5.2
11	5.8	5.3	5.2
12	5.7	5.2	5.1
13	5.6	5.2	5.1
14	5.6	5.2	5.1
15	5.6	5.2	5.1
16	5.5	5.2	5.1
17	5.5	5.2	5.1
18	5.5	5.1	5.1
19	5.5	5.1	5.0
33	5.3	5.0	5.0

The calculated bond order parameters from our ground state simulations for parahydrogen, orthodeuterium, and paratritium clusters from size $N = 4 - 19, 33$ are shown in Table A.3. The Q_4 and Q_6 distribution functions are also shown in Figures A.1, A.2, and A.3.

Table A.3: Calculated Q_4 and Q_6 bond order parameters for parahydrogen, orthodeuterium, and paratritium clusters.

	Q_4 values			Q_6 values		
N	pH_2	oD_2	pT_2	pH_2	oD_2	pT_2
4	0.279745	0.228479	0.21934	0.401305	0.463654	0.492322
5	0.243492	0.18591	0.161108	0.377559	0.39629	0.408952
6	0.19034	0.147323	0.134616	0.317411	0.340882	0.366932
7	0.159467	0.127727	0.107949	0.276265	0.283046	0.266181
8	0.137275	0.115342	0.108618	0.246298	0.271017	0.28832
9	0.12011	0.10163	0.0869631	0.223144	0.245063	0.222947
10	0.115756	0.0933536	0.0782435	0.216676	0.21813	0.194756
11	0.101584	0.0847545	0.0735975	0.19916	0.199425	0.186025
12	0.0934904	0.0727505	0.061206	0.19165	0.188749	0.171578
13	0.0845702	0.0579923	0.0489708	0.179277	0.149762	0.134353
14	0.0832213	0.0619453	0.0525888	0.176049	0.1533	0.132364
15	0.080185	0.0582131	0.0463512	0.169551	0.145344	0.122359
16	0.0811748	0.0544316	0.0462925	0.167446	0.133767	0.11809
17	0.0781166	0.0520348	0.0445418	0.161435	0.126949	0.116325
18	0.0741545	0.0489325	0.0405236	0.155644	0.122201	0.108875
19	0.0691853	0.0431381	0.0360582	0.148146	0.110648	0.0976483
33	0.0439497	0.0308463	0.0259603	0.100502	0.0803515	0.0728881

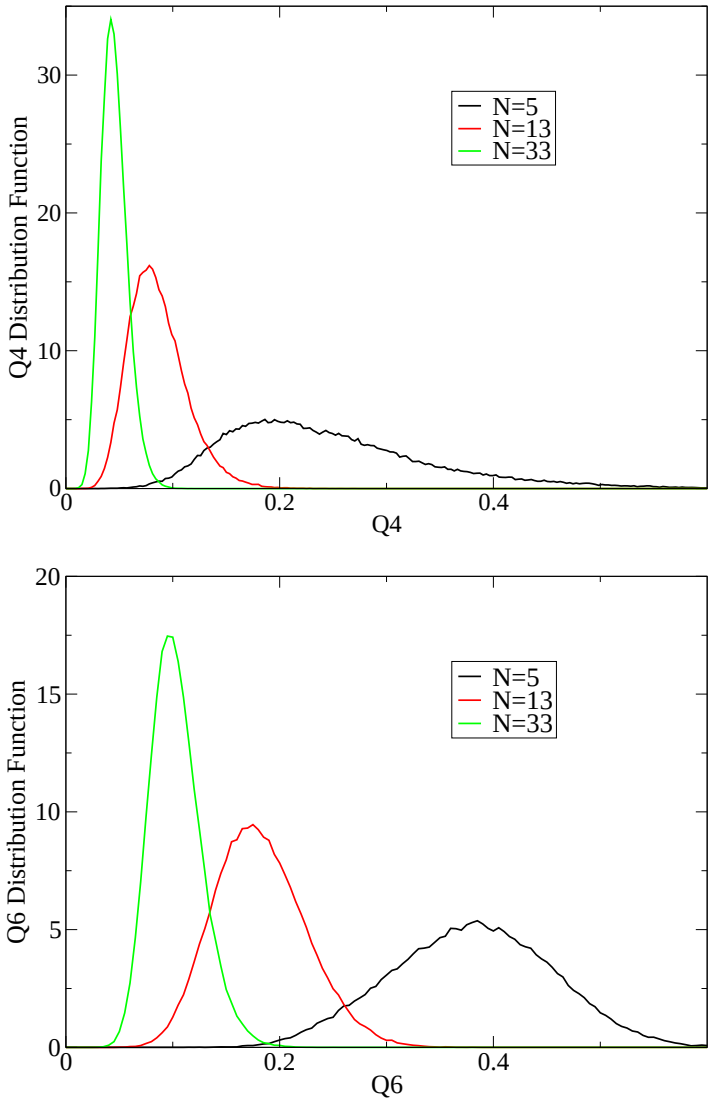

Figure A.1: Q_4 and Q_6 distribution functions for pH_2 clusters of size $N = 5, 13, 33$

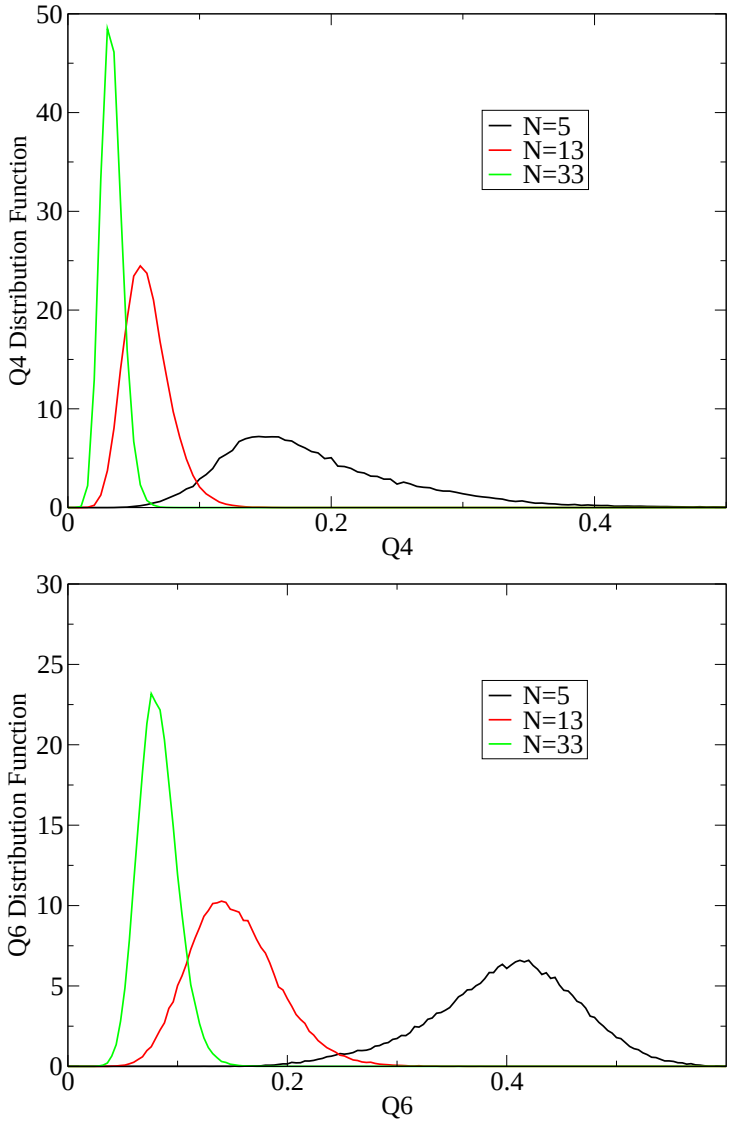

Figure A.2: Q_4 and Q_6 distribution functions for oD_2 clusters of size $N = 5, 13, 33$

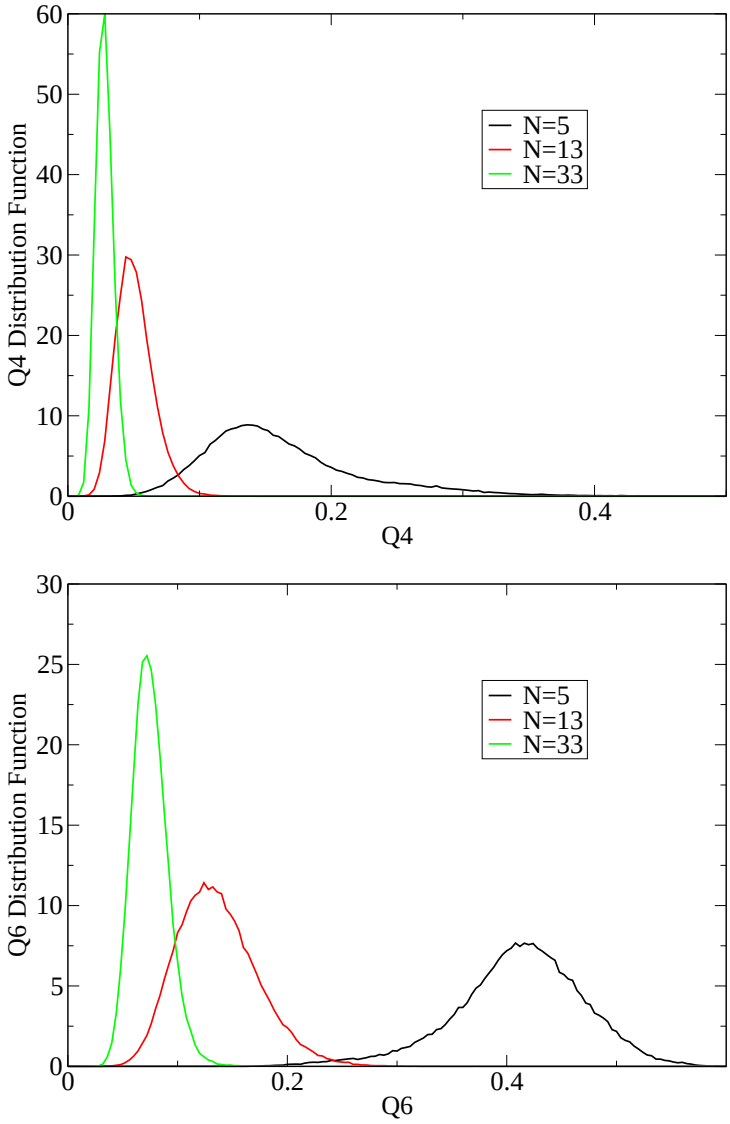

Figure A.3: Q_4 and Q_6 distribution functions for pT_2 clusters of size $N = 5, 13, 33$

It is clear that both the Q_4 and Q_6 bond order parameters monotonically decrease as cluster size increases and do not match any of the specific geometries listed above.

This may be due to the fact that our cluster sizes are so small that the majority of our atoms are 'surface atoms' and not fully connected 'core atoms' which most would be in a bulk system. We also note that the width of the Q_4 and Q_6 distributions, or standard deviation, also decreases with cluster size. This could be indicative of more atoms becoming more connected and thus the bond order parameters are better defined. However, further research is needed to more accurately explain these results. These results are consistent with previous research performed using a Monte Carlo version of PIGS by Cuervo *et al.* in 2008[65].

A.2 Validity of Harmonic Approximation

In an effort to determine how 'solid-like' our system is, we compare the ground state properties of our simulation to that using the harmonic approximation. It should be noted, that the configurations we use for the normal modes of our harmonic approximation are the classical Lennard Jones global minimum structures[57] using σ=3.00Å. We then perform a steepest descent energy minimization using the Buck potential to be consistent with our simulation[50].

We diagonalize the Hessian to obtain the eigenvalues (which correspond to harmonic frequencies) and eigenvectors (which correspond to the normal modes). We note that the

radial wavefunction is proportional to the square of the wavefunction

$$\psi^2 = \exp\left(\sum_{i=1}^{3N-6} \frac{-2\omega_i q_i^2}{2\hbar}\right) , \tag{A.4}$$

where N is the number of atoms in the cluster ($3N-6$ normal modes when in the minimum of the potential), q_i is the normal mode and ω_i is the corresponding frequency.

Written in this way, it is easy to see that we sample the normal modes as a Gaussian of zero mean and standard deviation of σ=sqrt[$\hbar/(2\omega)$]. Using the eigenvectors, we can convert back to Cartesian co-ordinates and obtain pair distances.

To demonstrate how the radial distribution functions compare between the harmonic approximation and our LePIGS simulation, we choose cluster size N=13 and N=33 for all isotopes. We observe, in Figure A.4, a large disagreement for N=13 parahydrogen and even as the cluster size increases to N=33, we see a similar disagreement.

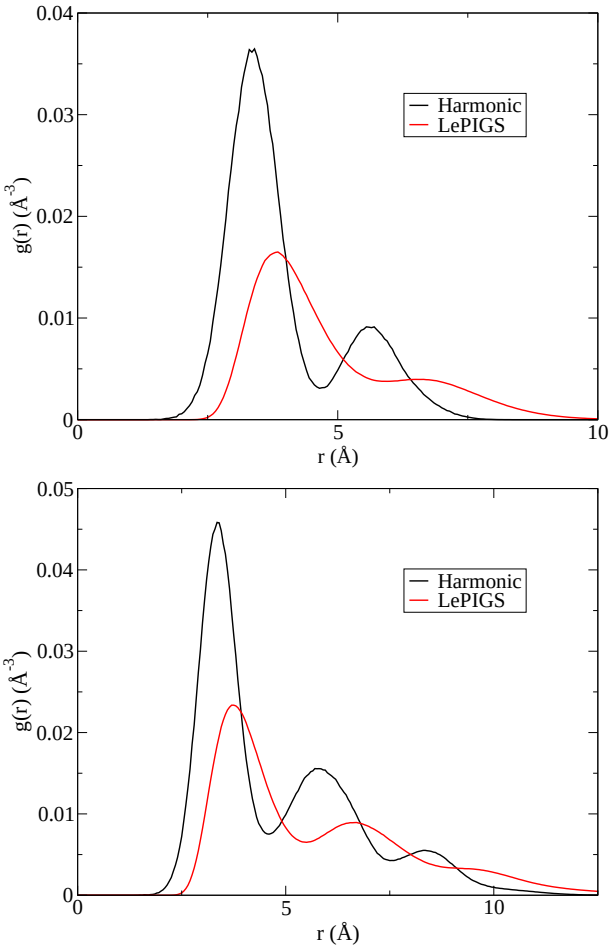

Figure A.4: Comparison of radial distribution functions for parahydrogen clusters of size $N = 13$ (top) and $N = 33$ (bottom) using the harmonic approximation and LePIGS simulation.

For N=13 orthodeuterium, we observe the radial distribution functions of the harmonic approximation and LePIGS simulation agree more than that of parahydrogen as shown in Figure A.5 and for N=33, there is significant overlap between the two distribution

functions.

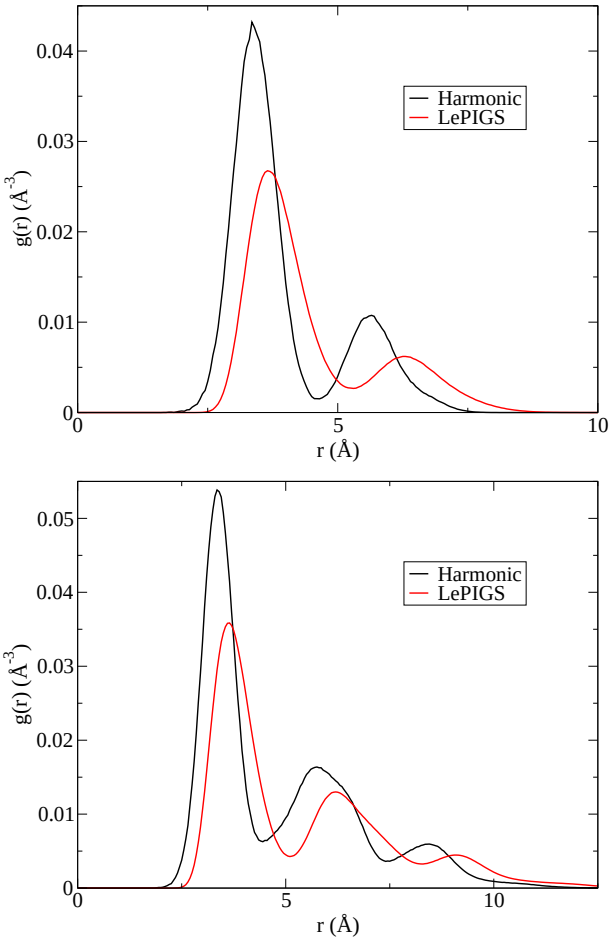

Figure A.5: Comparison of radial distribution functions for orthodeuterium clusters of size $N = 13$ (top) and $N = 33$ (bottom) using the harmonic approximation and LePIGS simulation.

We see the same trend for paratritium, as shown in Figure A.6, where the radial distributions agree more as mass increases and the distributions overlap more for N=33 than for N=13. This observation leads us to conclude that N=33 is a more solid-like cluster for

the heavier isotopes.

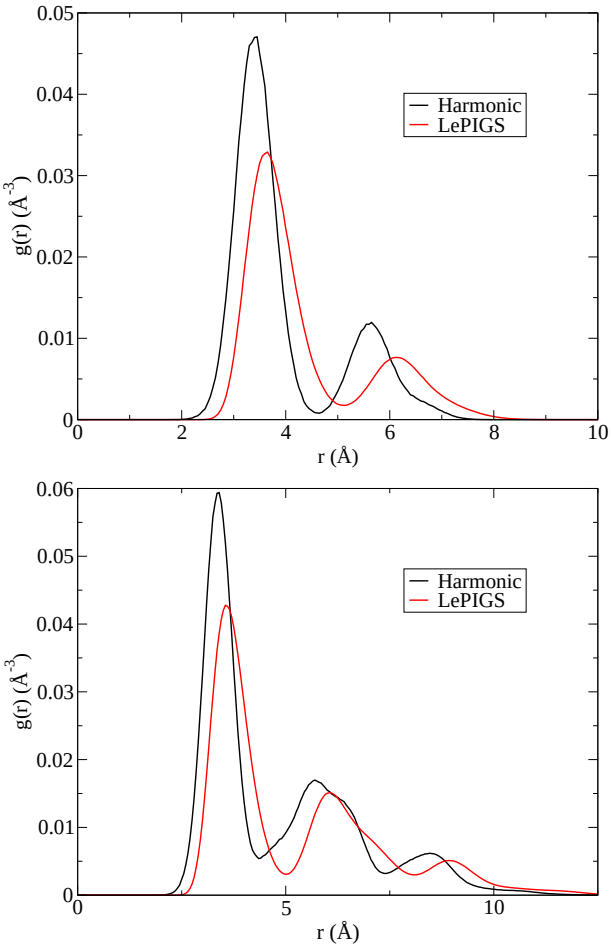

Figure A.6: Comparison of radial distribution functions for paratritium clusters of size $N = 13$ (top) and $N = 33$ (bottom) using the harmonic approximation and LePIGS simulation.

We also compare the energy per particle. For the harmonic approximation, the total ground state energy is defined as

$$E_0 = \sum_{i=0}^{3N-6} \left(\frac{\hbar\omega_i}{2} \right) + V_{min} \ , \tag{A.5}$$

where V_{min} is the minimum of the potential. Our V_{min} is obtained after a steepest descent energy minimization.

The comparison of energy per particle is shown in Figure A.7. We note that the absolute values are substantially different between the harmonic and LePIGS calculations for all three isotopes over the range of cluster sizes. For the parahydrogen clusters, the energies do not even follow the correct trend, but as the mass increases the energy values appear to agree more.

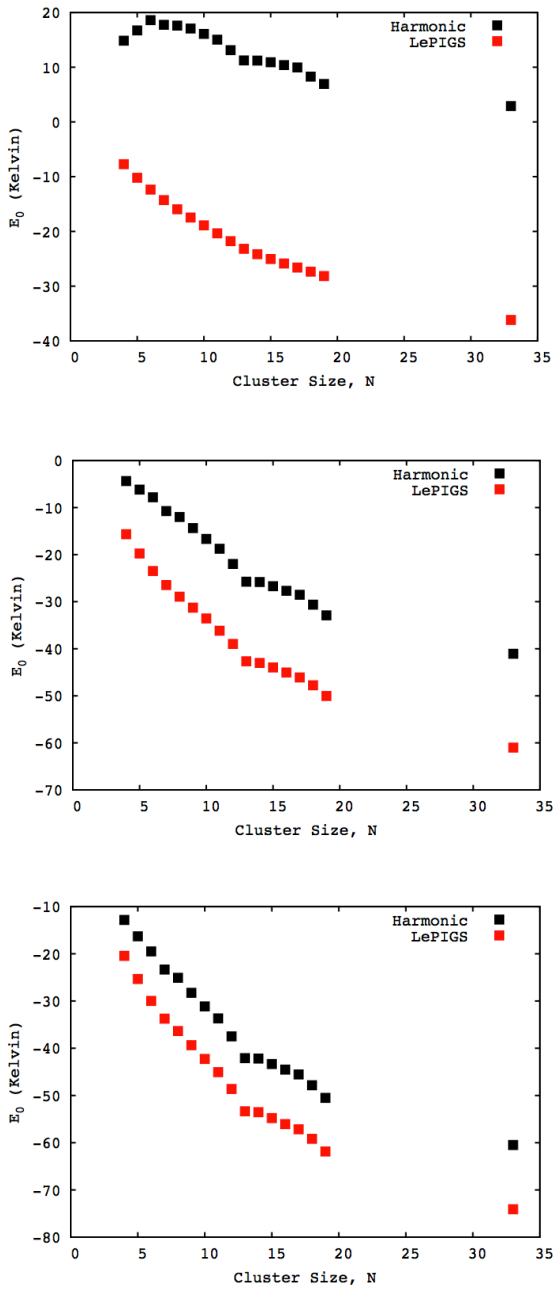

Figure A.7: Comparison of energy per particle for clusters of parahydrogen (top), orthodeuterium (middle), and paratritium (bottom) using the harmonic approximation and LePIGS simulation.

We also compare the total chemical potential, defined as

$$\mu = E_0(N) - E_0(N-1) \ . \tag{A.6}$$

The comparison of chemical potential, shown in Figure A.8, mimics that of the energy per particle. For parahydrogen clusters, the harmonic approximation does not give good agreement with the LePIGS calculation, though as the mass increases, the agreement becomes significantly better.

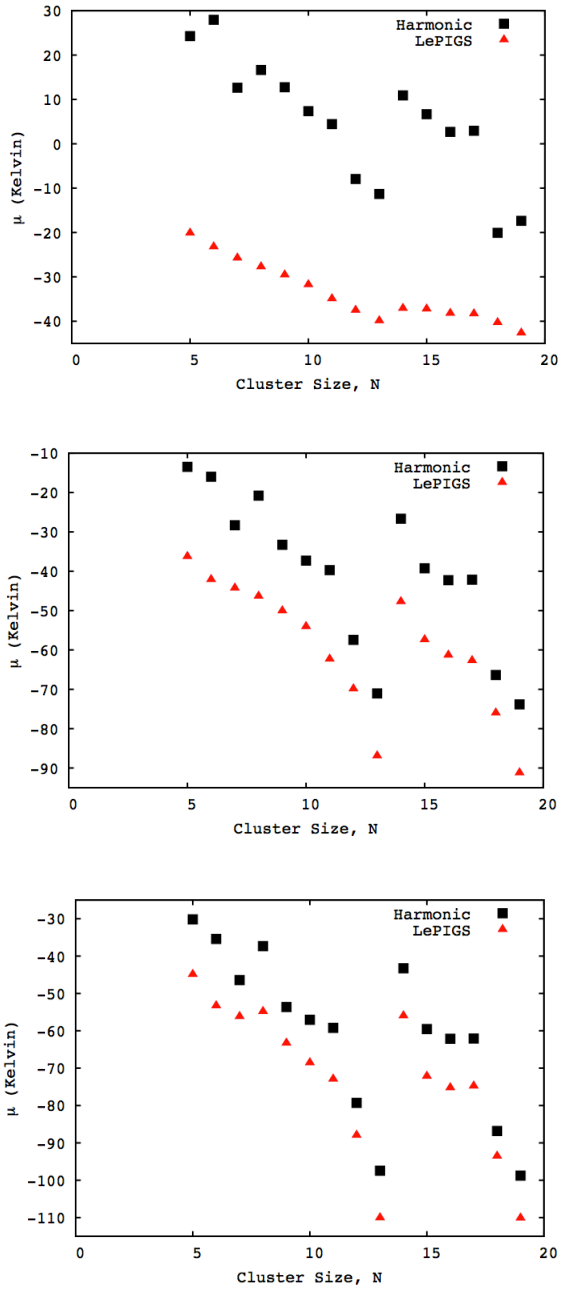

Figure A.8: Comparison of chemical potential for clusters of parahydrogen (top), orthodeuterium (middle), and paratritium (bottom) using the harmonic approximation and LePIGS simulation.

Overall, we notice parahydrogen is not represented well by the harmonic approximation and thus can be characterized as 'liquid-like', but there is evidence that other isotopologues show signs of 'solid-like' behaviour in the chemical potential trends. However, based on the radial distribution functions and the absolute value of energy per particle, the harmonic approximation is still a poor representation of these clusters.

References

[1] J. Macintyre, *Dictionary of inorganic compounds*, volume 3 (Chapman & Hall/CRC, 1992).

[2] V. L. Ginzburg and A. A. Sobyanin, Pis' ma Zh. Eksp. Teor. Fiz **15**, 343 (1972).

[3] P. Sindzingre, D. M. Ceperley, and M. L. Klein, Phys. Rev. Lett. **67**, 1871 (1991).

[4] M. C. Gordillo and D. M. Ceperley, Phys. Rev. B **65**, 174527 (2002).

[5] E. Rabani and J. Jortner, J. Phys. Chem. B **110**, 18893 (2006).

[6] F. Mezzacapo and M. Boninsegni, Phys. Rev. Lett. **97**, 045301 (2006).

[7] F. Mezzacapo and M. Boninsegni, Phys. Rev. A **75**, 033201 (2007).

[8] F. Mezzacapo and M. Boninsegni, Phys. Rev. A **76**, 021201 (2007).

[9] S. A. Khairallah, M. B. Sevryuk, D. M. Ceperley, and J. P. Toennies, Phys. Rev. Lett. **98**, 183401 (2007).

[10] S. Warnecke, M. B. Sevryuk, D. M. Ceperley, J. P. Toennies, R. Guardiola, and J. Navarro, Eur. Phys. J. D **56**, 353 (2010).

[11] M. B. Sevryuk, J. P. Toennies, and D. M. Ceperley, J. Chem. Phys. **133**, 064505 (2010).

[12] T. Zeng, H. Li, R. J. L. Roy, and P.-N. Roy, J. Chem. Phys. **135**, 094304 (2011).

[13] H. Li, R. J. Le Roy, P.-N. Roy, and A. R. W. McKellar, Phys. Rev. Lett. **105**, 133401 (2010).

[14] S. Grebenev, B. Sartakov, J. P. Toennies, and . A. F. Vilesov, Science **289**, 1532 (2000).

[15] S. Grebenev, B. Sartakov, J. P. Toennies, and A. Vilesov, Phys. Rev. Lett. **89**, 225301 (2002).

[16] S. Grebenev, B. G. Sartakov, J. P. Toennies, and A. F. Vilesov, Europhys. Lett. **83**, 66008 (2008).

[17] S. Grebenev, B. G. Sartakov, J. P. Toennies, and A. F. Vilesov, J. Chem. Phys. **132**, 064501 (2010).

[18] G. Tejeda, J. M. Fernndez, S. Montero, D. Blume, and J. P. Toennies, Phys. Rev. Lett. **92**, 223401 (2004).

[19] J. Navarro and R. Guardiola, Int. J. Quantum Chem. **111**, 463 (2011).

[20] T. Zeng and P.-N. Roy, Rep. Prog. Phys. **77**, 046601 (2014).

[21] S. M. Rothstein, Can. J. Chem. **91**, 505 (2013).

[22] J. B. Anderson, J. Chem. Phys. **63**, 1499 (1975).

[23] B. Hetenyi, E. Rabani, and B. J. Berne, J. Chem. Phys. **110**, 6143 (1999).

[24] A. Sarsa, K. E. Schmidt, and W. R. Magro, J. Chem. Phys. **113**, 1366 (2000).

[25] E. Sola and J. Boronat, J. Phys. Chem. A **115**, 7071 (2011).

[26] R. Guardiola and J. Navarro, Phys. Rev. A **74**, 025201 (2006).

[27] R. Guardiola and J. Navarro, Cent. Eur. J. Phys. **6**, 33 (2008).

[28] M. Boninsegni and S. Moroni, Phys. Rev. E **86**, 056712 (2012).

[29] D. M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).

[30] R. J. Hinde, Chem. Phys. Lett. **418**, 481 (2006).

[31] J. E. Cuervo, P.-N. Roy, and M. Boninsegni, J. Chem. Phys. **122**, 114504 (2005).

[32] J. E. Cuervo and P.-N. Roy, J. Chem. Phys. **125**, 124314 (2006).

[33] J. E. Cuervo and P.-N. Roy, J. Chem. Phys. **128**, 224509 (2008).

[34] J. E. Cuervo and P.-N. Roy, J. Chem. Phys. **131**, 114302 (2009).

[35] S. Miura, Chem. Phys. Lett. **482**, 165 (2009).

[36] G. J. Martyna, M. L. Klein, and M. Tuckerman, J. Chem. Phys. **97**, 2635 (1992).

[37] S. Constable, M. Schmidt, C. Ing, T. Zeng, and P.-N. Roy, J. Phys. Chem. A **117**, 7461 (2013).

[38] K. Hinsen, J. Comp. Chem. **21**, 79 (2000).

[39] C. Ing, K. Hinsen, J. Yang, T. Zeng, H. Li, and P.-N. Roy, J. Chem. Phys. **136**, 224309 (2012).

[40] C. Ing, M.Sc. Thesis, Department of Physics, University of Waterloo (2011).

[41] H. F. Trotter, Proc. Amer. Math. Soc. **10**, 545 (1959).

[42] M. Suzuki, Phys. Lett. A **180**, 232 (1993).

[43] M. Ceriotti, M. Parrinello, T. E. Markland, and D. E. Manolopoulos, J. Chem. Phys. **133**, 124104 (2010).

[44] M. Boninsegni, N. Prokof'ev, and B. Svistunov, Phys. Rev. Lett. **96**, 070601 (2006).

[45] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[46] M. Parrinello and A. Rahman, J. Chem. Phys. **80**, 860 (1984).

[47] M. E. Tuckerman, Path Integration via Molecular Dynamics, in *Quantum Simulations of Complex Many-Body Systems*, volume 10, pages 269–298, John von Neumann Institute for Computing, 2002.

[48] I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **121**, 3368 (2004).

[49] H. Li, P.-N. Roy, and R. J. L. Roy, J. Chem. Phys. **133**, 104305 (2010).

[50] U. Buck, F. Huisken, A. Kohlhase, D. Otten, and J. Schaefer, J. Chem. Phys. **78**, 4439 (1983).

[51] M. Frigo and S. Johnson, Proc. *IEEE* **93**, 216 (2005).

[52] T. Zeng, H. Li, and P.-N. Roy, J. Phys. Chem. Lett. **4**, 18 (2013).

[53] I. Bešlić, L. V. Markić, and J. Boronat, J. Chem. Phys. **128**, 064302 (2008).

[54] F. Lindemann, Physik. Z. **11**, 609 (1910).

[55] C. Chakravarty, P. G. Debenedetti, and F. H. Stillinger, J. Chem. Phys. **126**, 204508 (2007).

[56] R. Guardiola and J. Navarro, J. Phys. Chem. A **115**, 6843 (2011).

[57] D. J. Wales and J. P. K. Doye, J. Phys. Chem. A **101**, 5111 (1997).

[58] M. H. Müser, J. Chem. Phys. **114**, 6364 (2001).

[59] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, Phys. Rev. B **28**, 784 (1983).

[60] Y. T. Wang, S. Teitel, and C. Dellago, J. Chem. Phys. **122**, 214722 (2005).

[61] P. A. Frantsuzov and V. A. Mandelshtam, Phys. Rev. E **72**, 037102 (2005).

[62] P. A. Frantsuzov, D. Meluzzi, and V. A. Mandelshtam, Phys. Rev. Lett. **96**, 113401 (2006).

[63] V. A. Mandelshtam and P. A. Frantsuzov, J. Chem. Phys. **124**, 204511 (2006).

[64] L. Zhan, J. Z. Y. Chen, and W.-K. Liu, J. Chem. Phys. **127**, 141101 (2007).

[65] J. Cuervo and U. of Alberta (Canada)., *Ground State Properties of Pure and Doped Quantum Clusters*, Canadian theses (University of Alberta (Canada), 2008).

[66] Y. Sugita and Y. Okamoto, Chemical Physics Letters **314**, 141 (1999).

[67] U. H. Hansmann, Chemical Physics Letters **281**, 140 (1997).

[68] L. Hernandez de la Pena, Mol. Phys. **112**, 929 (2014).

[69] H. Li, N. Blinov, P.-N. Roy, and R. J. Le Roy, J. Chem. Phys. **130**, 144305 (2009).

[70] R. J. Hinde, J. Chem. Phys. **128**, 154308 (2008).

[71] N. Faruk, M. Schmidt, H. Li, R. J. LeRoy, and P.-N. Roy, J. Chem. Phys **submitted** (2013).

[72] K. Patkowski, W. Cencek, P. Jankowski, K. Szalewicz, J. B. Mehl, G. Garberoglio, and A. H. Harvey, J. Chem. Phys. **129**, 094304 (2008).

[73] I. F. Silvera and V. V. Goldman, J. Chem. Phys. **69**, 4209 (1978).

[74] V. Struzhkin et al., Chem. Rev. **107**, 4133 (2007).

[75] W. L. Mao, H.-k. Mao, A. F. Goncharov, V. V. Struzhkin, Q. Guo, J. Hu, J. Shu, R. J. Hemley, M. Somayazulu, and Y. Zhao, Science **297**, 2247 (2002).

[76] L. Florusse, C. Peters, J. Schoonman, K. Hester, C. Koh, S. Dec, K. Marsh, and E. Sloan, Science **306**, 469 (2004).

[77] T. A. Strobel, C. J. Taylor, K. C. Hester, S. F. Dec, C. A. Koh, K. T. Miller, and E. D. Sloan, J. Phys. Chem. B **110**, 17121 (2006).

[78] K. A. Lokshin, Y. Zhao, D. He, W. L. Mao, H.-K. Mao, R. J. Hemley, M. V. Lobanov, and M. Greenblatt, Phys. Rev. Lett. **93**, 125503 (2004).

[79] S. Patchkovskii and J. Tse, PNAS **100**, 14645 (2003).

[80] M. Xu, Y. S. Elmatad, F. Sebastianelli, J. W. Moskowitz, and Z. Bacic, J. Phys. Chem. B **110**, 24806 (2006).

[81] F. Sebastianelli, M. Xu, Y. S. Elmatad, J. W. Moskowitz, and Z. Bacic, J. Phys. Chem. C **111**, 2497 (2007).

[82] F. Sebastianelli, M. Xu, D. K. Kanan, and Z. Bacic, J. Phys. Chem. A **111**, 6115 (2007).

[83] M. Xu, F. Sebastianelli, and Z. Bacic, J. Phys. Chem. A **111**, 12763 (2007).

[84] M. Xu, F. Sebastianelli, and Z. Bacic, J. Chem. Phys. **128**, 244715 (2008).

[85] F. Sebastianelli, M. Xu, and Z. Bacic, J. Chem. Phys. **129**, 244706 (2008).

[86] M. Xu, F. Sebastianelli, and Z. Bacic, J. Phys. Chem. A **113**, 7601 (2009).

[87] A. Witt, F. Sebastianelli, M. E. Tuckerman, and Z. Bacic, J. Phys. Chem. C **114**, 20775 (2010).

[88] M. Xu, L. Ulivi, M. Celli, D. Colognesi, and Z. Bacic, Phys. Rev. B **83**, 241403 (2011).

[89] M. Xu and Z. Bacic, Phys. Rev. B **84**, 195445 (2011).

[90] P. L. Raston, W. Jäger, H. Li, R. J. Le Roy, and P.-N. Roy, Phys. Rev. Lett. **108**, 253402 (2012).