Path integral ground state approaches for the study of weakly bound clusters and confined molecules

by

Matthew D.G. Schmidt

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Chemistry

Waterloo, Ontario, Canada, 2018

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Dr. Gilles Peslherbe Professor of Chemistry and Biochemistry; Director for Research in Molecular Modeling (CERMM) Concodia University, Montreal, QC, Canada
Supervisor	Dr. Pierre-Nicholas Roy Professor; Canada Research Chair in Quantum Molecular Dynamics University of Waterloo, Waterloo, ON, Canada
Internal Member	Dr. Marcel Nooijen Professor University of Waterloo, Waterloo, ON, Canada

Internal-external Member	Dr. Roger Melko
	Professor; Canada Research Chair in Computational
	Many-Body Physics
	University of Waterloo, Waterloo, ON, Canada
Other Member	Dr. Scott Hopkins
	Associate Professor
	University of Waterloo, Waterloo, ON, Canada

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.

First author Manuscript: M. Schmidt, J. M. Fernández, N. Faruk, M. Nooijen, R. J. Le Roy, J. H. Morilla, G. Tejeda, S. Montero, and P.-N. Roy, "Raman Vibrational Shifts of Small Clusters of Hydrogen Isotopologues", J. Phys. Chem. A 119, 12551 (2015). Copyright 2015 American Chemical Society.

The material used in this thesis relating to this manuscript is my own, except for the experimental work done to measure the Raman spectra for the orthodeuterium clusters which was done by my collaborators.

First author Manuscript: M.Schmidt and P.-N. Roy, "Path integral Molecular dynamic simulation of flexible molecular systems in their ground state: application to the water dimer", J. Chem. Phys. 148, 124116 (2018). Copyright 2018 American Institute Physics.

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

First author Manuscript: M. Schmidt, D. Fillion, J. Millar and P.-N. Roy, "Quantum molecular dynamics simulations of contained parahydrogen molecules within clathrate hydrates: merging low temperature dynamics with the zero temperature limit"

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

First author Manuscript: M. Schmidt and P.-N. Roy, "Analysing the accuracy and efficiency of different methods to calculate Raman vibrational shifts of hydrogen clusters"

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

First author Manuscript: M. Schmidt and P.-N. Roy, "Langevin equation path integral ground state calculation of the chemical potential of parahydrogen"

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

Abstract

This thesis presents the study of weakly bound clusters in the ground state (or the zero-temperature limit) using path integral molecular dynamics. Specifically, we look individually at the quantum properties of small clusters of hydrogen and water molecules and confined hydrogen within water cages, known as clathrate hydrates, which serves as a more practical application. Clathrate hydrates have been extensively studied as a clean storage container for molecular hydrogen and there have been discrepancies on the hydrogen occupancy number between various theoretical and experimental studies. It has been shown that the occupancy number is sensitive to the potential energy surfaces and models of the hydrogen and water systems. A preliminary study of hydrogen contained in clathrates is performed using a traditional hydrogen pair potential and water-hydrogen interaction potential. Hydrogen occupancy and structural distributions are compared to literature values. Small clusters of the individual molecules themselves are then focused on. The molecular hydrogen pair potentials are evaluated by calculating the Raman vibrational shift, a property that is very sensitive to the interaction potential, and comparing to experimental measurements. These shifts are calculated using first order perturbation theory based on pair distribution functions generated from Langevin equation path integral ground state (LePIGS) simulations for all bosonic isotopologues. It is determined that the shifts calculated using the Hinde pair potential give better agreement to experimen-

tal results than the traditional hydrogen potentials that we have been using in the past. The perturbation theory approach is then compared with two exact methods to calculate the shifts. For the application of hydrogen clusters, it is determined that perturbation theory is the best choice when balancing accuracy and precision. In the literature, there has been a discrepancy in the shape of the chemical potential at low temperature and in the ground state. We calculate the ground state chemical potential using LePIGS and find agreement with other PIGS results. We then extend our LePIGS code to simulate flexible molecules by investigating the water dimer. Ground state energies, dissociation energies, and structural properties are calculated using two empirically based interaction potentials and one *ab initio* potential, MB-pol, that includes polarizability and many-body effects which has been shown to reproduce experimental dissociation energies. We further demonstrate that imaginary time correlation functions generated from LePIGS can be used to calculate accurate vibrational transition energies. This work serves as a demonstration of the effectiveness of the LePIGS method towards calculating ground state properties of small clusters and provides useful information on the interaction potentials that should be used for systems containing hydrogen or water, specifically hydrogen contained in a flexible clathrate hydrate. An analytic form of the 1-D Hinde pair potential for molecular hydrogen is also contained for general use.

Acknowledgements

I would first and foremost like to acknowledge my supervisor, Dr. Pierre-Nicholas Roy, for giving me the chance to join the Theoretical Chemistry group when I was an undergraduate and for his guidance and support ever since. He has taught me both academically, but also many life skills including how to have patience, perseverance, and confidence when working with data or communicating your results. His perpetual energy and boundless ideas are the keys to this outstanding work environment and the 8 years that I have been in his group have been a true pleasure. Drs. Marcel Nooijen and Robert J. Le Roy have also played a significant role in my growth as a student and a person. Dr. Nooijen was the advisor for my undergraduate research project in electronic structure and gave me the opportunity to attend my first conferences. Due to him, I am one of the only members of the entire group to use the Emacs text editor and am very familiar in Fortran, both of which I am very proud of. He has always been a resource I can count on and I appreciate his honesty, respect, and care he shows to his students. I enjoyed working with the late Dr. Le Roy and all of the discussions that we have had both academically and otherwise. His keen eye for presentation aesthetics and grammatical errors in reports has made me a much better author/presenter than before I began. His dedication to the students and the group was extremely admirable.

I also owe a lot of appreciation to Kevin Bishop and Dmitri Iouchtchenko for their

friendship and valuable support both in research, but also maintaining our computing cluster! It's been awesome working with them through our PhDs and both have very bright futures. I also owe a lot of gratitude towards Toby Zeng, who has provided career advice and guidance and is always available for any research question that I may have. Spencer Yim, Lindsay Orr, and Neil Raymond have also been outstanding colleagues and great friends.

Finally, I'd like to thank all of the undergraduate students that have assisted in the research over the years including Jayme Millar, Devon Fillion, Xichen Lou, and Liam McGrath. It's been a pleasure mentoring you, I hope you learned a lot, and I wish you the best of luck in the future.

Dedication

To my family, friends, and colleagues who have supported me throughout the entire process and in memoriam of Dr. Robert J. Le Roy (1943-2018)

Table of Contents

Examining Committee Membership ii

Author's Declaration iv

Statement of Contributions v

Abstract vii

Acknowledgements ix

Dedication xi

List of Tables xvi

List of Figures xx

List of Abbreviations xxiv

1 Introduction 1

1.1 Clathrate hydrates as a hydrogen storage container 2

1.2 Using quantum molecular dynamics to determine the accuracy of interaction potentials . . . 6
1.3 Thesis outline . . . 10

2 Theory **12**
2.1 Finite temperature path integral theory . . . 13
2.2 Path integral ground state theory . . . 16
2.3 Path integral Langevin equation thermostat . . . 21
2.4 Implementation . . . 25

3 A preliminary study of confined parahydrogen molecules within clathrate hydrates using quantum molecular dynamics **26**
3.1 Introduction . . . 27
3.2 System setup . . . 28
3.3 Results . . . 30
3.4 Concluding remarks . . . 38

4 Raman vibrational shifts of hydrogen isotopologues **40**
4.1 Introduction . . . 40
4.2 Theoretical methodology . . . 42
4.2.1 Calculating the Raman shifts for a cluster of identical particles . . . 42
4.2.2 Constructing the effective H1-Ps for the dimer . . . 44
4.2.3 Ground state simulation methodology . . . 47
4.3 Results and discussion . . . 49
4.3.1 Experimental Raman vibrational shifts . . . 49

4.3.2 Raman shifts of the dimers and trimers used to quantify methodological errors and determine scaling factors 52
4.3.3 Raman vibrational shifts for entire cluster range 62
4.3.4 Portable analytic MLR models for our isotope-dependent H1-Ps . . 67
4.4 Concluding remarks 73

5 Analysing the accuracy and efficiency of different methods to calculate Raman vibrational shifts of hydrogen clusters 77
5.1 Introduction 78
5.2 Theory and methodology 80
5.2.1 Methods to calculate the Raman vibrational frequency shift for a pure cluster 81
5.2.2 Ground state simulation parameters and treatment of statistical errors 84
5.3 Results 87
5.4 Concluding remarks 91

6 Ground state chemical potential of parahydrogen clusters of size N=21-40 94
6.1 Introduction 94
6.2 System setup 97
6.3 Results 98
6.4 Concluding Remarks 99

7 An extension to flexible molecular systems in their ground state: Application to the water dimer 101
7.1 Introduction 101

7.2 Theory and methodology . . . 103
7.2.1 System parameters and optimization . . . 103
7.2.2 Imaginary time correlation functions . . . 107
7.3 Results . . . 109
7.3.1 Ground state energies . . . 110
7.3.2 Ground state distributions . . . 118
7.3.3 Vibrational energies of the water monomer . . . 122
7.4 Concluding remarks . . . 129

8 Conclusions and future work **132**
8.1 Concluding remarks . . . 132
8.2 Further investigation of confined hydrogen in structure-II clathrate hydrates 136
8.3 Continued software development of the LePIGS and PIMD methods . . . 137

Letter of Copyright Permissions **139**

References **141**

Appendices **154**

A A simplified model for the many-body problem **155**

List of Tables

3.1 System parameters for our finite temperature PIMD calculations and zero-temperature limit LePIGS calculations 30

3.2 Ground state energies (kJ/mol) of systems of confined parahydrogen molecules occupying the small (5^{12}) and large ($5^{12}6^4$) cages calculated from LePIGS simulations 31

3.3 Chemical potential (kJ/mol) of parahydrogen molecules occupying the large ($5^{12}6^4$) cage 31

3.4 Energies (kJ/mol) of confined parahydrogen molecules within the small (5^{12}) and large ($5^{12}6^4$) cages at various temperatures (using PIMD) and in the zero-temperature limit (using LePIGS). The parameter $\tau = 0.002$ K^{-1} is constant for each simulation. 38

4.1 Raman vibrational shifts (in cm^{-1}) of $(pH_2)_N$ and $(oD_2)_N$ clusters up to $N = 11$ obtained from simulation [$\Delta\nu$(Sim|PT)] and new experimental values for $(oD_2)_N$. Estimated uncertainties shown in parentheses. Those reported from simulation are standard errors using 1σ. Quantities marked with a * are obtained from exact diagonalization [$\Delta\nu(\psi_0|PT)$]. Results obtained using the B-P are also shown for comparison. The scaling factors used are 1.036 for pH_2 and 1.030 for oD_2. Note: PT refers to first-order perturbation theory 52

4.2 Ground state energies and Raman vibrational shifts (in cm^{-1}) of $(pH_2)_2$, $(oD_2)_2$, and $(pT_2)_2$ as calculated using our H1-Ps. Comparison is made to Hinde's results using his H6-PES Hinde 6-D Exact [68], new experimental data for $(oD_2)_2$, and previous experimental results for $(pH_2)_2$[31]. $\Delta\nu(\psi)$ refers to the method of using direct diagonalization to calculate the shift by directly taking the difference between the ground state energy [E_0 (ψ_0)] and the energy of the excited vibrational state. Using exact diagonalization, the shifts are calculated using first order perturbation theory [$\Delta\nu(\psi_0|\text{PT})$] and second order perturbation theory [$\Delta\nu(\psi_0|\text{PT-2})$]. The shifts are also calculated using the pair distribution functions from simulation [$\Delta\nu(\text{Sim}|\text{PT})$]. Estimated uncertainties are shown in parentheses. Those reported from simulation are standard errors using 1σ. 59

4.3 Raman vibrational shifts (in cm^{-1}) of $(pH_2)_3$, $(oD_2)_3$, and $(pT_2)_3$ as calculated using our H1-Ps. Comparison is made to new experimental data for $(oD_2)_3$, and previous experimental results for $(pH_2)_3$[31]. $\Delta\nu(\psi)$ refers to the method of using direct diagonalization to calculate the shift by directly taking the difference between the ground state energy and the energy of the excited vibrational state. The shifts are also calculated using first order perturbation theory with the pair distribution functions determined by the exact wavefunctions [$\Delta\nu(\psi_0|\text{PT})$] or from the results of simulation [$\Delta\nu(\text{Sim}|\text{PT})$]. Estimated uncertainties are shown in parentheses. Those reported from simulation are standard errors using 1σ. 61

4.4 Raman vibrational shifts (in cm^{-1}) of $(pT_2)_N$ clusters up to $N = 11$ obtained from simulation [$\Delta\nu(\text{Sim}|\text{PT})$] using our numerical H1-Ps and difference potentials for (pT_2) dimers. Estimated uncertainties, reported as standard errors using 1σ, are shown in parentheses. Quantities marked with a * are obtained from exact diagonalization [$\Delta\nu(\psi_0|\text{PT})$] using the H1-Ps. Results obtained using the B-P potential are also shown for comparison. The scaling factor used is $f^T = 1.011$ for pT_2. 64

4.5 Selecting range and density of the tabulated H1-P data points to fit the analytic potential to and the choice of uncertainty percentages. 69

4.6 MLR fit parameters for the ground state H1-Ps of pH_2, oD_2, and pT_2. Note all energies are in cm^{-1} and lengths in Å. Parameters common for all three H1-MLRs are $p = 6$, $q = 1$, $r_{\text{ref}} = 4.60$, and $\rho^{AB} = 1.10$. 70

4.7 Rotationless bound state energies and vibrational shifts (in cm^{-1}) of $(pH_2)_2$, $(oD_2)_2$, and $(pT_2)_2$ dimers calculated using the Lanczos diagonalization method with our H1-P and the present H1-MLR. Energies are relative to the two free monomers. 72

4.8 Rotationless bound state energies E_n and vibrational shifts (in cm^{-1}) of $(pH_2)_3$, $(oD_2)_3$, and $(pT_2)_3$ trimers calculated using the Lanczos diagonalization method with our H1-P and the present H1-MLR. Energies are relative to the three free monomers. 72

5.1 Parameters for $(pH_2)_N$ and $(oD_2)_N$ simulations: Timestep dt, centroid friction γ_0, relaxation parameter β, and length of simulation t 85

5.2 Ground state energies of pH_2 in Kelvin 87

5.3 Ground state energies of oD_2 88

5.4 Calculated Raman vibrational shifts ($\Delta\nu_{0,N}$) of pH_2 in cm^{-1} using three different methods: simulation (Sim), ratio (Ratio), and perturbation theory (PT) with the experimental results shown as well 89

5.5 Calculated Raman vibrational shifts ($\Delta\nu_{0,N}$) of oD_2 in cm^{-1} using three different methods: simulation (Sim), ratio (Ratio), and perturbation theory (PT) with the experimental results shown as well 89

7.1 Simulation parameters for H_2O Water Models: Timestep dt, friction γ^0, decorrelation time T_d, and relaxation parameter β 105

7.2 Simulation Parameters for D_2O Water Models 105

7.3 H_2O ground state energies (kcal/mol) calculated for different water models. 116

7.4 D_2O ground state energies (kcal/mol) calculated for different water models 116

7.5 H_2O monomer average ground state properties calculated for different water models 118

7.6 D_2O monomer average ground state properties calculated for different water models 118

7.7 Calculation of H_2O pure vibrational transition energies (kcal/mol) for each vibrational mode using the MB-pol potential. 127

7.8 Calculation of D_2O pure vibrational transition energies (kcal/mol) for each vibrational mode using the MB-pol potential. 127

List of Figures

2.1 Pictorial representation of 2 particles in a finite temperature PIMD simulation (left panel), denoted by a closed path, and in a LePIGS simulation (right panel), denoted by an open path. In PIMD, all beads are indistinguishable and each bead contributes to the average of energetic and structural properties. In LePIGS, an estimate of the ground state wavefunction is affixed to the end beads (blue). These end beads contribute to energetic properties, using the mixed estimator, and the middle bead (black) contributes to properties diagonal in the position representation. 16

3.1 Structures of the small dodecahedral (5^{12}) (left) and large hexakaidecahedral (5^{12} 6^4) (right) clathrate hydrates which comprise the structure-II unit cell. The blue dotted lines represent hydrogen bonds. 29

3.2 Probability distribution functions for the distance between the parahydrogen molecules and the center of the small (5^{12}) cage (top) and large ($5^{12}6^4$) cage (bottom), $R_{CCM} - R_{HCM}$, for various parahydrogen occupancies. 33

3.3 Pair distribution functions for the distance between the parahydrogen molecules occupying the small (5^{12}) cage (top) and large ($5^{12}6^4$) cage (bottom), $R_{HCM} - R_{HCM}$, for various parahydrogen occupancies. 34

3.4 Probability distribution functions for the angle between the parahydrogen molecules and the center of the large ($5^{12}6^4$) cage, $\theta_{H_2-cc-H_2}$, for cluster size $N = 4$. 36

4.1 Vibrational Raman spectrum of small $(oD_2)_N$ clusters, as recorded in a cryogenic free jet of $oD_2(10\%)$+He mixture. The numbers 2-11 indicate the size N of the cluster. Experimental conditions (see text): $T_0 = 40$ K, $P_0 = 3$ bar, $D = 50$ μm, $z = 300$ μm. The estimated local temperature is 1.1K. 51

4.2 H1-Ps of pH_2, oD_2, and pT_2 dimers in the ground ($v = 0$) vibrational state. 54

4.3 Difference between the H1-Ps of pH_2, oD_2, and pT_2 dimers in the first excited ($v = 1$) and the ground ($v = 0$) vibrational states. 55

4.4 Pair distribution functions generated from LePIGS simulations of $(pH_2)_4$, $(oD_2)_4$, and $(pT_2)_4$ clusters using the H1-P for each isotopologue. 56

4.5 Predicted Raman vibrational shifts (with scaling) for (a) $(pH_2)_N$ and (b) $(oD_2)_N$ clusters up to size $N = 11$ using the H1-P for each isotopologue. Experimental results are also shown. 63

4.6 Solid lines: Integrand of the vibrational shift equation: the difference potential multiplied by the pair distribution function obtained from simulation, shown for the $(pH_2)_3$, $(oD_2)_3$, and $(pT_2)_3$ trimers. Dotted lines: The cumulative sums of the integrand (scaled to fit the diagram) reflecting how the vibrational shifts change with intermolecular distance. 66

4.7 Ground state wavefunction of the $(pH_2)_2$, $(oD_2)_2$, and $(pT_2)_2$ dimers calculated by exact diagonalization using the H1-Ps. 71

4.8 One-dimensional particle distribution for energy levels E_n in Pekeris coordinates of (a) $(pH_2)_3$, (b) $(oD_2)_3$, and (c) $(pT_2)_3$ trimers using Lanczos diagonalization scheme. 74

6.1 Ground state chemical potentials of $(pH_2)_N$ for the B-P[70] interaction potential (upper) and the SG-P interaction potential (lower). LePIGS results are in black squares, the DMC results from Ref. [116] are in green triangles, the PIGS-MC results from Ref. [51] are in red, and the VPIMD results from Ref. [63] are in blue 100

7.1 Average MB-pol potential energy along the imaginary time step for the H_2O monomer. In our simulations of this system, we choose $\beta = 0.002\ K^{-1}$, which is in the converged regime. 106

7.2 Energy convergence with τ for the H_2O monomer (upper panel) and dimer (lower panel) using the q-TIP4P/F model comparing LePIGS results to DMC. The $\tau = 0$ points are the extrapolated E_0 values from the fits. . . . 112

7.3 Energy convergence with τ for the D_2O monomer (upper panel) and dimer (lower panel) using the q-TIP4P/F model comparing LePIGS results to DMC. The $\tau = 0$ points are the extrapolated E_0 values from the fits. . . . 113

7.4 Energy convergence with τ for H_2O monomer (upper panel) and dimer (lower panel) using the q-SPC/Fw model. The $\tau = 0$ points are the extrapolated E_0 values from the fits. 114

7.5 Energy convergence with τ for H_2O (left column) and D_2O (right column) monomer (upper row) and dimer (lower row) using the MB-pol model comparing LePIGS results to DMC. The $\tau = 0$ points are the extrapolated E_0 values from the fits. 115

7.6 H_2O monomer (upper panel), dimer (middle panel), and dissociation energies (lower panel) calculated using the MB-pol intermolecular potential at various temperatures using PIMD and in the zero-temperature limit using LePIGS. 119

7.7 Probability distribution functions for the OH bond lengths, R_{OH} (upper panel), HH distances, R_{HH} (middle panel), and the HOH bond angle, θ_{HOH} (lower panel), for the water monomer described by each of the models. . . 120

7.8 Probability distribution functions for the distance between oxygen atoms R_{OO} (upper panel), oxygen and hydrogen atoms R_{OH} (middle panel), and distance between hydrogen atoms R_{HH} (lower panel) for the water dimer described by each of the models. 121

7.9 Sample imaginary time correlation function of the H_2O symmetric stretch operator calculated from a LePIGS simulation using the MB-pol potential. The data points are shown as black circles and are connected by lines as a visual aid to view the exponential nature of the correlation function. 125

7.10 Sample log plot of the imaginary time correlation function of the H_2O symmetric stretch operator calculated from a LePIGS simulation using the MB-pol potential. The data points are represented as black circles and the linear regression is represented by a red line. 126

List of Abbreviations

AHR	Adiabatic Hindered Rotor
B-P	Buck Potential
DCT	Discrete Cosine Transform
DFT	Density Functional Theory
DMC	Diffusion Monte Carlo
H1-MLR	Hinde 1-D Potential Energy Surface fit to Morse Long Range form
H1-PES	Hinde 1-D Potential Energy Surface
H6-PES	Hinde 6-D Potential Energy Surface
INS	Inelastic Neutron Scattering
ITCF	Imaginary Time Correlation Function
LePIGS	Langevin equation Path Integral Ground State
MLR	Morse Long Range
MMTK	Molecular Modelling Toolkit
oD_2	orthodeuterium
PES	Potential Energy Surface
pH_2	parahydrogen
PIGS	Path Integral Ground State
PILE	Path Integral Langevin Equation
PIMC	Path Integral Monte Carlo
PIMD	Path Integral Molecular Dynamics
PT	Perturbation Theory
PT-2	Second-Order Perturbation Theory
pT_2	paratritium
RMD	Reptile Molecular Dynamics
RPMD	Ring Polymer Molecular Dynamics
SG-P	Silvera-Goldmann Potential
VMD	Variational Molecular Dynamics
VPIMD	Variational Path Integral Molecular Dynamics

Chapter 1

Introduction

The work reported in this thesis aims at providing theoretical tools and computational methods for the description of complex molecular systems. Nuclear quantum effects in the ground state will be of primary focus here and systems of interest will range from hydrogen and water clusters to hydrogen molecules confined in water cages. The simulation tools developed during this PhD work can be applied to a broad range of problems from fundamental properties to practical applications. To illustrate the latter point, we describe below a situation where hydrogen and water molecules are combined to provide storage for an alternate fuel.

1.1 Clathrate hydrates as a hydrogen storage container

Molecular hydrogen as a fuel source has long been considered as a clean energy alternative. It has the highest energy per mass of any fuel[1] and the process to release energy, an electrochemical reaction with oxygen molecules, produces water, a clean exhaust. However, a difficulty towards the advancement of hydrogen fuel is it's storage. Hydrogen has a very low density at room temperature, which translates to a low energy per unit volume. Current methods to store hydrogen are physical containers including pressurized hydrogen gas cylinders and liquid hydrogen in cryogenic tanks, however these carry significant mass overhead and hydrogen stored in high pressure containers may lead to safety concerns.

Due to this, the United States Department of Energy (DOE) has invested heavily in alternative storage containers. By 2020, they have set a target of finding an alternative storage system with a 4.5% weight percentage of hydrogen, which is a metric that ensures that the storage container is compact enough and can hold enough fuel that it can be used feasibly for transportation vehicles. Some alternative sources are material-based "containers" which include adsorption of hydrogen on a surface or absorption of hydrogen within a solid. One molecular compound that has been well-studied as a potential clean hydrogen storage container is the clathrate hydrate, a hydrogen-bonded network of water molecules

that form a cage structure. Specifically, experiments have shown that hydrogen molecules can exist in sII clathrate hydrates[2], whose unit cell has 2 distinct cage types: a 'small' dodecahedral cage that is composed of 20 water molecules and has 12 pentagonal faces (also denoted 5^{12}) and a 'large' hexakaidecahedral cage composed of 28 water molecules and has 12 pentagonal faces and 4 hexagonal faces (also denoted $5^{12}6^4$). The properties of hydrogen molecules confined within this class of water molecules and their potential use as an alternative storage container is our main focus. There have been many reviews[3, 4, 5, 6] summarizing the experimental and theoretical research performed on clathrate hydrates so we will only provide a brief summary limited to the studies of occupancy of hydrogen molecules within clathrate hydrates since that is our focus.

Hydrogen-containing clathrate hydrates were first experimentally synthesized in 1999 at high pressure of 15 kbar and reported the possibility of a hydrogen to water ratio, $R = 1/3$[7, 8]. In 2002, Mao *et al.*[2] synthesized the clathrate hydrate at 2 kbar and 300 K and conducted high-pressure Raman, infrared, x-ray, and neutron studies. They confirmed synthesis of the structure-II clathrate hydrate with $R = 0.45 \pm 0.05$ and deduced that the large cage contained 4 hydrogen molecules and the small cage contained 2 hydrogen molecules through x-ray and neutron diffraction. The clathrates were also cooled to 78 K and the pressure removed to vacuum (10^{-4} kbar) demonstrating the stability of a clathrate at low temperature with no external pressure keeping it together. Further experimental groups

have synthesized deuterated clathrates[9, 10], clathrates of different structure type[11, 12], and clathrates containing a promoter molecule tetrahydrofuran to decrease the pressure required for synthesizing the sII hydrate but decreases the mass percentage as it takes a void away from the hydrogen molecules[13, 14]. Raman vibronic spectra of hydrogen in clathrates can also be used to distinguish the parahydrogen and orthohydrogen vibrational transitions[15].

Theoretical calculations have also been performed to investigate occupancy of hydrogen molecules within clathrate hydrates under a variety of conditions while comparing to the experimental results. Patchkovskii and Tse performed density functional theory (DFT) calculations[16] of up to 3 hydrogen molecules in the small cage and up to 5 hydrogen molecules in the large cage and calculate reaction enthalpies of hydrogen encapsulation. They conclude that at the experimental conditions of Mao *et al*, that the small cage and large cage could occupy 2 and 4 hydrogen molecules respectively under the assumptions that the individual hydrogen molecules undergo free rotation and the collective motions of the cluster of hydrogen molecules within an individual cage must be treated as rigid body rotations. Classical molecular dynamics simulations had also been performed, concluding single occupancy in the small cage and quadruple occupancy in the large cage, however these neglected quantum effects that could be very important at temperatures as low as 100 K to which they report[17].

A number of theoretical studies[18, 19, 20, 21, 22, 23, 24, 25, 26, 27] have been performed by a group led by Bacic which involve quantum 5-dimensional calculations of the coupled translation-rotation eigenstates of one hydrogen (or deuterium) molecule in the small or large cage. The group also uses diffusion Monte Carlo (DMC) to calculate ground state properties of cages with occupancy of more than one hydrogen. Initially, the group used the PA-D potential energy surface (PES) by Hodges *et al.*[28] for the H_2-cage potential and the 4-D H_2-H_2 PES by Diep and Johnson[29]. The cage is rigid and the hydrogen molecules are treated as rigid rotors. It was determined that the small cage could stably occupy up to 2 hydrogen molecules[19]. However, they changed[22] their H_2-cage potential to the SPC/E[17, 30], which more accurately represents the many-body induced polarization of the H_2-water interaction which in turn gave better agreement between their calculations and results from inelastic neutron scattering (INS) experiments. Using this potential, DMC simulations showed that the small cage could only stably support a single hydrogen molecule. In a similar method, they show that up to 4 hydrogens can occupy the large cage[23]. Along with Tuckerman[25], they perform path integral molecular dynamics (PIMD) simulations of 2 and 4 hydrogen molecules in the large cage and determine at 50 K, the hydrogen molecules devolve from a strictly tetrahedral configuration to a higher degree of delocalization as it has enough energy to explore the higher regions of the PES, consistent with results seen from neutron diffraction experiments. Other theoretical works

have involved looking at reproducing inelastic neutron scattering spectra[26, 27]. However, this demonstrates the necessity for an accurate PES to correctly model the hydrogen system within the clathrate hydrate, as simply changing the water-hydrogen interaction results in a different predicted occupancy number in the small cage.

1.2 Using quantum molecular dynamics to determine the accuracy of interaction potentials

Determining the accuracy of a PES is typically achieved by performing a quantum or classical calculation to determine a quantity of interest that can also be experimentally measured. One such quantity that can be calculated both theoretically and experimentally is the Raman vibrational frequency shift. This quantity is defined as the change of the energy difference between the $v = 0$ and $v = 1$ vibrational levels of a specific molecule from isolation when it is doped in a cluster of other atoms/molecules. Theoretically, this involves running quantum calculations and comparing the energy of the doped system in the vibrational ground state to the energy of the system that has one quantum of vibration and subtracting from the energy difference when the molecule is in isolation. There is also an alternative method based on first order perturbation theory (PT) for calculating the Raman vibrational shift for a cluster of N identical particles that relies on the pair

distribution function, $P_N(R)$, of the particles in a cluster.

Experimentally, this involves performing a Raman scattering experiment, done by supersonic free jet expansion, which has been demonstrated for pure hydrogen[31, 32, 33] and nitrogen clusters[34], but also other small chromophores doped in helium atoms[35, 36, 37, 38, 39, 40]. For molecules doped in small clusters of hydrogen molecules or helium atoms, the values of these frequency shifts are on the order of wavenumbers with experimental accuracy on the order of hundredths of wavenumbers, thus these quantities are very sensitive to the PES used and very accurate PESs must be used to reproduce this data.

To calculate properties of these systems, for example a structural property such as the radial distribution $P_N(R)$ or energies at low temperature or in the zero-temperature limit, one needs a method to account for quantum mechanical effects. For large systems that are too large for basis set calculations or electronic structure calculations, path integral calculations can be used. There are two variants of path integral calculations that can be used: path integral Monte Carlo (PIMC) which randomly samples configuration space or path integral molecular dynamics (PIMD) which uses Hamiltons equations of motion to guide the sampling. PIMC calculations tend to be computationally faster, but only if the algorithm which proposes successive moves is efficient. PIMD simulations do not suffer from this problem and have an important extension that allows for the calculation of approximate real time correlation-functions through the method of ring polymer molecular

dynamics (RPMD).

In much the same way, to study ground state properties of chemical systems, Quantum Monte Carlo methods can been used. Path Integral Ground State (PIGS)[41] and Diffusion Monte Carlo (DMC)[42, 43] are two such methods chosen to obtain ground state properties. DMC is typically chosen to obtain ground state energies due to it's high accuracy and efficiency. Recent work has analyzed sources of errors in both PIGS and DMC[44]. These authors have analyzed the so-called population size bias issue, a source of error not present in the PIGS approach. They observe that compared to DMC, PIGS and even finite temperature methods, have more favorable scaling[44]. They therefore conclude that PIGS is superior option for systems such as parahydrogen clusters. Also, calculating properties other than energies may be difficult to obtain due to the residual effect of the initial trial wavefunction in the resulting distribution $\phi_0\psi_T$[45, 44, 46, 47], whereas the square of the exact ground state wavefunction, $|\phi_0|^2$, is sampled directly using PIGS. This allows for the straight-forward calculation of ground state properties. As well, the errors in a PIGS simulation can each be removed systematically using fundamental power law scaling arguments such as the asymptotic Trotter error[48].

The PIGS method has primarily been used with Monte Carlo (MC)[49, 50, 51, 52] and a recent review[53] provides a detailed formulation of the PIGS-MC method and discusses the wide variety of applications including the ability to calculate of the Rényi entanglement

entropy for interacting bosons in a straight forward manner[54, 55, 56, 57]. However, there has been little effort in it's use in a molecular dynamics framework. Two related approaches termed Variational Path Integral Molecular Dynamics (VPIMD)[58, 59, 60, 61, 62, 63] and Langevin equation Path Integral Ground State (LePIGS)[64, 65] have been proposed. The use of thermostated molecular dynamics sampling means that the exploration of configuration space is guided by the gradient of the interaction potential and one does not have to develop MC updates. This aspect of the approach is important in systems that contain a range of timescales such as flexible molecular complexes as studied in the present work. Another advantage of using molecular dynamics sampling is the possible extension to a method similar to Ring Polymer Molecular Dynamics (RPMD)[66] to allow for the calculation of approximate real-time correlation functions.

In 2014, our group[67] calculated the Raman vibrational shifts of $(pH_2)_N$ clusters by generating the $P_N(R)$ using the LePIGS method in order to evaluate the accuracy of various PESs with regards to small hydrogen clusters. For these simulations, the 6-D Hinde potential (H6-PES)[68], a mostly *ab initio* potential, was reduced to 1-D both in the ground vibrational state (H1-P) and with one vibrational quantum using the Adiabatic Hindered Rotor (AHR) approximation[69]. The vibrational shifts were also calculated from simulations using the Buck potential (B-P)[70] and the Silvera-Goldmann potential (SG-P) [71], both empirically based potentials. Comparisons with experimental data showed

that our 1-D version of the H1-P potential outperforms the empirically based potentials for small cluster sizes ($N = 2 - 8, 13$), though the B-P appears to give better results for the last experimental point at cluster size $N \simeq 33$.

1.3 Thesis outline

This thesis consists of several parts. The details of our path integral and molecular dynamics methodology involved with our work are described in Chapter 2. In Chapter 3, we first look at a preliminary study of the occupancy of hydrogen molecules confined in clathrate hydrates. We then focus on small clusters of the individual molecules involved. In Chapter 4, we calculate the Raman vibrational shifts of hydrogen clusters and their isotopologues and compare them to experimental results. We also fit the H1-P potentials that we constructed to an analytic form for general use. In Chapter 5, we investigate three different methods of calculating the Raman vibrational shift, two exact methods, and a first order perturbation theory approach to determine which is best suited for small clusters, applied towards parahydrogen clusters. In Chapter 6, LePIGS is used to calculate the ground state chemical potential of parahydrogen clusters between the range of $N = 21 - 40$ as there has been discrepancy in the literature. In Chapter 7, we calculate the energetic and structural properties of the water monomer and dimer using three different PESs, two

empirically based and one *ab initio* that is polarizable and includes many-body effects. We demonstrate that LePIGS can be extended for flexible monomers and that it can be used to generate imaginary time correlation functions used to calculate vibrational transition energies. Concluding remarks and future applications of this work are discussed in Chapter 8.

Chapter 2

Theory

This chapter[1] will describe the finite temperature and ground state methodology that are used, as well as it's implementation into our molecular dynamics software.

[1]Sections of this chapter have been reprinted with permission from M. Schmidt and P.-N. Roy, "Path integral Molecular dynamic simulation of flexible molecular systems in their ground state: application to the water dimer", J. Chem. Phys. 148, 124116 (2018). Copyright 2018 American Institute Physics.

2.1 Finite temperature path integral theory

For a system of N translating particles of mass m, we can represent it's Hamiltonian $\hat{H}$ as a sum of the kinetic and potential energy operators, $\hat{K}$ and $\hat{V}$, respectively, as

$$\hat{H} = \hat{K} + \hat{V} = \sum_{i=1}^{N} \frac{\hat{p}_i^2}{2m_i} + V(\hat{q}_i, ..., \hat{q}_N) \ , \tag{2.1}$$

where $\hat{q}$ and $\hat{p}$ represent the position and momentum vectors respectively.

At finite temperature, the partition function is denoted as,

$$Z = \mathrm{Tr}\left[\exp(-\beta\hat{H})\right] \ , \tag{2.2}$$

where $\beta = 1/(k_B T)$. Expanding in the position representation results in the partition function taking the following form:

$$Z = \int dq \langle q| \exp[-\beta(\hat{K} + \hat{V})]|q\rangle \ . \tag{2.3}$$

Since in general, the kinetic and potential operators do not commute, the Trotter

factorization[72, 48, 73] can be used to redefine the density operator as:

$$\exp[-\beta(\hat{K}+\hat{V})] = \lim_{P\to\infty}\left[\exp(-\frac{\beta}{2P}\hat{V})\exp(-\frac{\beta}{P}\hat{K})\exp(-\frac{\beta}{2P}\hat{V})\right]^{P} \quad , \tag{2.4}$$

where P is the number of factors. The error of this factorization, known as the 'Trotter error' is $O[(\frac{\beta}{P})^3]$. The high-temperature density operator, $\hat{\rho}$, can now be defined, denoting $\tau = \beta/P$, as

$$\hat{\rho} = \left[\exp(-\frac{\tau}{2}\hat{V})\exp(-\tau\hat{K})\exp(-\frac{\tau}{2}\hat{V})\right] \quad , \tag{2.5}$$

such that the partition function can now be represented as:

$$Z = \lim_{P\to\infty}\int dq\langle q|\hat{\rho}^{P}|q\rangle \ . \tag{2.6}$$

Inserting $P-1$ resolutions of the identity operator of the position basis after each factor results in the discretization of the path,

$$Z = \lim_{P\to\infty}\int dq_1,...,dq_P\left[\prod_{i=1}^{P}\langle q_i|\hat{\rho}|q_{i+1}\rangle\right]_{q_{P+1}=q_1} \ . \tag{2.7}$$

It is important to note the circularity condition where $q_{P+1} = q_1$ as only $P-1$ resolutions of the identity were required due to the trace.

After evaluating the individual potential energy operators using the position representation and inserting a complete set of momentum eigenstates at each imaginary time slice to evaluate the kinetic energy operators, the final form of the partition function for a single particle in one dimension becomes:

$$Z = \lim_{P\to\infty} \left(\frac{m}{2\pi\tau\hbar^2}\right)^{P/2} \int dq_1, ..., dq_P \\ \exp\left\{-\sum_{i=1}^{P}\left[-\frac{m}{2\tau\hbar^2}(q_{i+1}-q_i)^2 + \tau V(q_i)\right]\right\}_{q_{P+1}=q_i}, \tag{2.8}$$

which can easily be extended to multiple particles in multiple dimensions.

Based on the expression, a pictorial view of finite temperature path integrals is shown in Fig. 2.1, which denotes each particle represented by a cyclic chain of P beads, which are connected by harmonic springs. Each bead on each particle then interacts with corresponding beads on the other particles via the potential energy. Each bead is identical and indistinguishable along the path and structural and energetic properties can be obtained from each bead during a simulation. The PIMD method is an approach to simulate finite temperature systems.

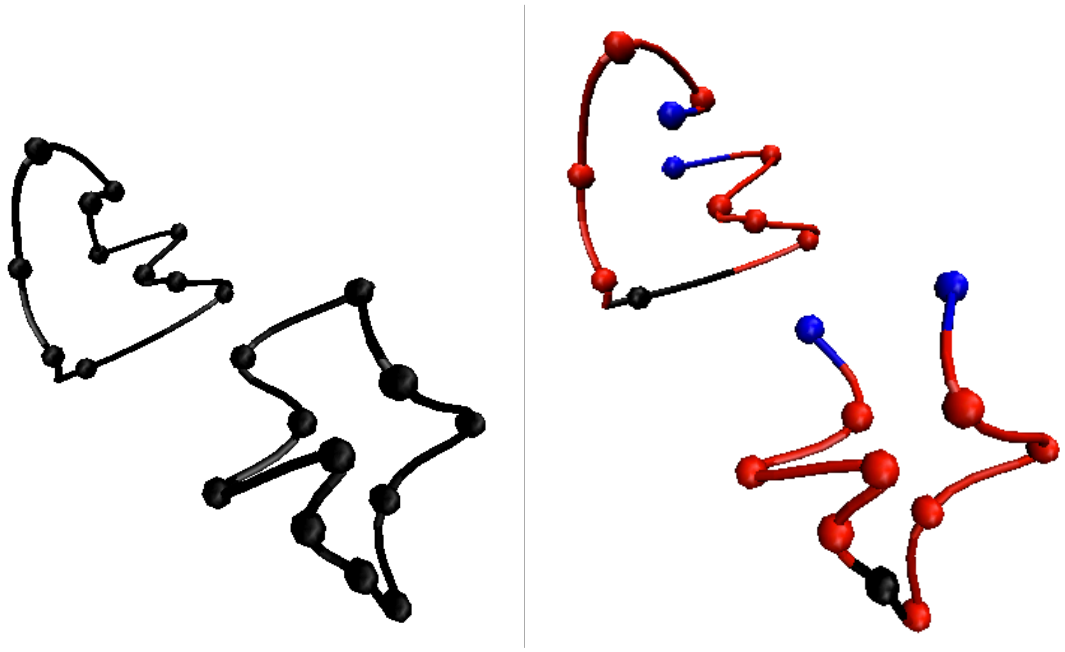


Figure 2.1: Pictorial representation of 2 particles in a finite temperature PIMD simulation (left panel), denoted by a closed path, and in a LePIGS simulation (right panel), denoted by an open path. In PIMD, all beads are indistinguishable and each bead contributes to the average of energetic and structural properties. In LePIGS, an estimate of the ground state wavefunction is affixed to the end beads (blue). These end beads contribute to energetic properties, using the mixed estimator, and the middle bead (black) contributes to properties diagonal in the position representation.

2.2 Path integral ground state theory

Once again, we start with a system of N translating particles with a Hamiltonian, $\hat{H}$, defined as a sum of the kinetic and potential energy operators, $\hat{K}$ and $\hat{V}$ respectively. This

time, we define a pseudo-partition function which only counts the ground state,

$$Z_0 = \langle \psi_T | \exp(-\beta \hat{H}) | \psi_T \rangle \ , \tag{2.9}$$

where ψ_T is a trial wavefunction and serves as an estimation or guess of the exact wavefunction ϕ_0. Should $\psi_T = \phi_0$, it is clear that the partition function simply reduces to the exact ground state partition function $\exp(-\beta E_0)$.

In the position representation,

$$Z_0 = \int dq \int dq' \psi_T(q) \langle q | \exp(-\beta \hat{H}) | q' \rangle \psi_T(q') \ . \tag{2.10}$$

The Trotter factorization[72, 48, 73] can once again be applied to obtain,

$$Z_0 = \psi_T(q_1) \left[\prod_{i=1}^{P-1} \langle q_i | \hat{\rho} | q_{i+1} \rangle \right] \psi_T(q_P) \ , \tag{2.11}$$

where $\tau = \beta/(P-1)$ and the $\hat{\rho}$ is the same high temperature density operator as in Eq. 2.5.

Evaluating the kinetic and potential operators as before, the final expression becomes:

$$Z_0 = \lim_{P \to \infty} \left(\frac{m}{2\pi\tau\hbar^2} \right)^{\frac{P-1}{2}} \int dq_1 ... \int dq_P \exp(-\tau V'(q)) \ , \tag{2.12}$$

where the effective potential $V'(q)$ is defined as

$$\begin{aligned} V'(q) = & \sum_{i=1}^{P-1} \left[\frac{1}{2} m\omega_{P-1}^2 (q_i - q_{i+1})^2 \right] \\ & + \sum_{i=1}^{P} \left[c_j V(q_j) \right] + V^{trial}(q_1) + V^{trial}(q_P) \ , \end{aligned} \tag{2.13}$$

the frequency, $\omega_{P-1} = (P-1)/(\beta\hbar)$, and

$$c_j = \begin{cases} \frac{1}{2} & \text{if } j = 1 \text{ or } j = P \\ \\ 1 & \text{otherwise.} \end{cases} \tag{2.14}$$

Since the ground state wavefunction has no nodes, typically our trial wavefunctions are also nodeless and thus the potential term contributed by the trial wavefunction can be written as:

$$V^{trial}(q) = -\frac{1}{\tau} \ln \left[\psi_T(q) \right] \tag{2.15}$$

This ground state pseudo-partition function looks very similar to the partition function in the finite temperature case, except for two major features. First, the ground state partition function is not cyclic, but an open path, and the second is that the potential energy term at the end beads is only half and includes information about the trial wavefunction. These differences are both pictorially reflected in Fig. 2.1.

Due to this, the way structural and energetic properties are calculated are also different. To calculate a ground state expectation value of a property A, we require

$$\langle A \rangle_0 = \frac{\langle \phi_0 | \hat{A} | \phi_0 \rangle}{\langle \phi_0 | \phi_0 \rangle} = \frac{1}{Z_0} \langle \phi_0 | \hat{A} | \phi_0 \rangle \ . \tag{2.16}$$

For a position-dependent property A, the primitive estimator can be written as:

$$\langle A \rangle_0 = \lim_{\beta, P \to \infty} \int dq_1 ... \int dq_P A(q_{(P+1)/2}) \exp\left[-\tau V'(q) \right] \ , \tag{2.17}$$

which results in only the middle bead contributing to any ground state structural distribution or other position-dependent properties like the potential energy. This also requires the number of beads in the simulation to be an odd number.

If an operator $\hat{A}$ commutes with exp(-$\beta \hat{H}$), then $\hat{A}$ can be shifted to the right hand side such that it acts directly on the trial wavefunction, ψ_T. The Hamiltonian operator, for example, commutes with itself and performing this operation leads to the mixed estimator for the energy:

$$E_0 = \left\langle \frac{\hat{H} \psi_T(q_P)}{\psi_T(q_P)} \right\rangle_0 \ , \tag{2.18}$$

where E_0 is the ground state energy and q_P is the position of the Pth bead. The Hamiltonian operator could also be shifted to the left and act on the 1st bead since $\hat{H}$ is Hermitian.

This allows for two contributions to the energy statistics (one at each end) along the path, instead of just one contribution at the middle for position-dependent properties. Since the same Trotter factorization is used in this formulation as in finite temperature path integrals, the same error exists.

An important feature of the PIGS formulation is that regardless of which trial wavefunction is used, the exact ground state wavefunction can be obtained as long as there is finite overlap between the two.

If we start with

$$\exp\left(-\frac{\beta}{2}\hat{H}\right)|\psi_T\rangle \ , \tag{2.19}$$

which appears to be Equation 2.9 split in half, we can decompose the trial wavefunction into it's exact energy eigenstates:

$$\exp\left(-\frac{\beta}{2}\hat{H}\right)|\psi_T\rangle = \exp\left(-\frac{\beta}{2}\hat{H}\right)\sum_{i=1}^{\infty} c_i|\phi_i\rangle \ , \tag{2.20}$$

where c_i is a coefficient representing the overlap between $|\psi_T\rangle$ and $|\phi_i\rangle$. Decomposing the Hamiltonian into it's energy levels leaves:

$$\begin{aligned}\exp\left(-\frac{\beta}{2}\hat{H}\right)|\psi_T\rangle &= c_0 \exp\left(-\frac{\beta}{2}E_0\right)|\phi_0\rangle \\ &+ c_1 \exp\left(-\frac{\beta}{2}E_1\right)|\phi_1\rangle + c_2 \exp\left(-\frac{\beta}{2}E_2\right)|\phi_2\rangle + \ldots \ .\end{aligned} \tag{2.21}$$

This can also be re-written as:

$$\exp\left(-\frac{\beta}{2}\hat{H}\right)|\psi_T\rangle = \exp\left(-\frac{\beta}{2}E_0\right) \times \left[c_0|\phi_0\rangle + c_1\exp\left(-\frac{\beta}{2}(E_1 - E_0)\right)|\phi_1\rangle + ...\right] . \tag{2.22}$$

Since each E_i is larger than E_0 for $i > 0$, all terms vanish in the limit as $\beta \to \infty$ and all that remains is the ground state:

$$\lim_{\beta\to\infty} \exp\left(-\frac{\beta}{2}\hat{H}\right)|\psi_T\rangle = c_0\exp\left(-\frac{\beta}{2}E_0\right)|\phi_0\rangle . \tag{2.23}$$

The above is true if c_0 is non-zero, implying the trial wavefunction we choose has finite overlap with the ground state wavefunction. The extent of the overlap will determine the convergence of the ground state wavefunction. As well, the error in the energy associated with β is exponential, as contributions of the excited states begin to factor in.

2.3 Path integral Langevin equation thermostat

In this section, the path integral Langevin equation (PILE) thermostat[74] will be described in the context of LePIGS[64] and modifications from the finite temperature PIMD will be explicitly noted.

In terms of phase space variables, the total Hamiltonian, $H_P(p,q)$, can be written as a sum of the free particle Hamiltonian, $H_P^0(p,q)$ and the potential, $V_P(q)$, as

$$H_P(p,q) = H_P^0(p,q) + V_P(q) \ , \tag{2.24}$$

where

$$V_P(q) = \sum_{i=1}^{N} \left(\left[\frac{1}{P} \sum_{j=1}^{P} V'(q_i^{(j)}) \right] + V^{trial}(q_i^{(1)}) + V^{trial}(q_i^{(P)}) \right) \tag{2.25}$$

and

$$H_P^0(p,q) = \sum_{i=1}^{N} \sum_{j=1}^{P-1} \left(\frac{\left[p_i^{(j)} \right]^2}{2m_i} + \frac{1}{2} m_i \omega_{P-1}^2 \left[q_i^{(j+1)} - q_i^{(j)} \right]^2 \right) \ . \tag{2.26}$$

The PILE thermostat requires the Hamiltonian to be expressed in terms of normal mode coordinates to efficiently sample the momenta. Since an open path exists in PIGS, normal mode coordinates can be written in terms of the Discrete Cosine Transform type II (DCT-II) instead of the Fourier matrix which is used for the closed path PIMD. The normal mode coordinates used in LePIGS are:

$$\tilde{p}_i^{(k)} = \sum_{j=1}^{P-1} C_{kj} p_i^{(j)} \tag{2.27}$$

and

$$\tilde{q}_i^{(k)} = \sum_{j=1}^{P-1} C_{kj} q_i^{(j)} \ , \tag{2.28}$$

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

$$C_{kj} = \begin{cases} \sqrt{\frac{1}{P}} \cos(\pi k(j-1/2)/P) & k = 0 \\ \sqrt{\frac{2}{P}} \cos(\pi k(j-1/2)/P) & k > 0. \end{cases} \tag{2.29}$$

The free particle ground state Hamiltonian, H_P^0 now takes a more desired form:

$$H_P^0(p,q) = \sum_{i=1}^{N} \sum_{k=0}^{P-1} \frac{\left[\widetilde{p}_i^{(k)}\right]^2}{2m_i} + \frac{1}{2} m_i \omega_k^2 \left[\widetilde{q}_i^{k}\right]^2 \tag{2.30}$$

where $\omega_k = 2\omega_{P-1} \sin(k\pi/(2P))$. In this form, it is possible to obtain an analytical solution to the equations of motion in terms of sines and cosines. At this point, the PILE thermostat is exactly the same for finite temperature PIMD as it is for LePIGS. The only differences are the modified potential energy, a slightly redefined ω_k value (changed from P links in the path to $P-1$ links) and the required use of the DCT-II instead of the Fourier transform.

The PILE thermostat then works by propagating first in the position space to calculate the potential energy and then converted into normal mode space to calculate the kinetic spring terms as independent normal modes. The momenta are then scaled prior and

following the propagation of the kinetic springs in normal mode space according to the Langevin equation:

$$\widetilde{p}_i^{(k)} \leftarrow c_1^{(k)} \widetilde{p}_i^{(k)} + \sqrt{\frac{m_i}{\beta}} c_2^{(k)} \xi_i^{(k)} \tag{2.31}$$

with the coefficients c_1 and c_2 defined as

$$c_1^{(k)} = \exp\left(-\frac{dt}{2}\gamma^{(k)}\right) \tag{2.32}$$

and

$$c_2^{(k)} = \sqrt{1 - \left[c_1^{(k)}\right]^2}\,. \tag{2.33}$$

The variable $\gamma^{(k)}$ is the Langevin equation friction on the kth normal mode and $\xi_i^{(k)}$ is a normally distributed random number respectively. The optimal Langevin frictions can be determined analytically for each of the k modes, except for the $k = 0$ normal mode which corresponds to the centroid of the path. These frictions take the form:

$$\gamma^{(k)} = \begin{cases} 1/\tau_0 & k = 0 \\ \\ 2\omega_k & k > 0, \end{cases} \tag{2.34}$$

where τ_0 is a parameter set by the user. Choosing a suboptimal τ_0 only affects the efficiency of our simulation, so the value of τ_0 that we choose is that which minimizes the simulation

decorrelation time for the total ground state energy since that is our main property of interest.

2.4 Implementation

Our finite temperature PIMD code is implemented in the open source Molecular Modelling Tool Kit (MMTK) [75] with the incorporation of the PILE thermostat by Ing*et al.*[76]. The Fourier transform is provided by the FFTW package[77]. The LePIGS method was also implemented by Constable *et al.*[64] in MMTK based on the implementation of the PILE thermostat. The discrete cosine transform code is provided by the same FFTW package.

Chapter 3

A preliminary study of confined parahydrogen molecules within clathrate hydrates using quantum molecular dynamics

3.1 Introduction

In this chapter, we use the PIMD and LePIGS methods for finite temperature and zero-temperature studies, respectively, to determine parahydrogen occupancy properties in the

small dodecahedral (5^{12}) and large hexakaidecahedral (5^{12} 6^4) sized cages which comprise the structure-II unit cell. We look at energetic and structural properties of small clusters of hydrogen molecules, treated as point-like particles, confined within each of the different sized clathrates, treated as rigid, to determine energetic and structural properties in the zero-temperature limit. We then calculate the energies as a function of temperature and merge the low temperature results, calculated using finite temperature PIMD, with the zero-temperature results, using LePIGS, demonstrating the two methods are compatible. The background theory and a description of the methods used are described in Section 3.2, followed by results in Section 3.3, and the chapter closes with conclusions in Section 3.4.

3.2 System setup

Since the rotational motion of parahydrogen is on a much faster timescale than it's nuclear translational motion of which we are simulating, it appears spherical and thus we are able to treat each parahydrogen molecule as a point-like particle. We use the Buck potential (B-P) [70] for intermolecular parahydrogen interactions. We use a 3-D water-hydrogen potential[78] based on the adiabatic hindered rotor (AHR) approximation of the Valiron potential[79]. Both of our water cages are rigid and non-interacting, so they solely provide an external potential for the parahydrogen molecules. The cage geometries are taken from

literature[80] and their structures shown in Figure 3.1.

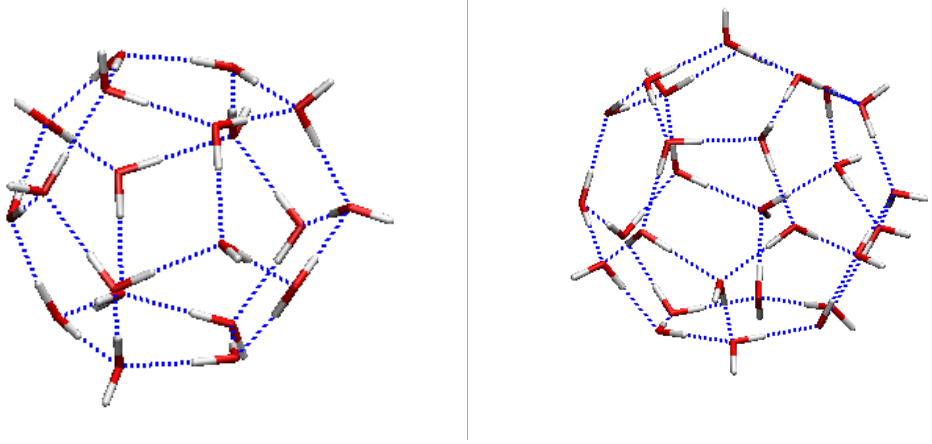


Figure 3.1: Structures of the small dodecahedral (5^{12}) (left) and large hexakaidecahedral (5^{12} 6^4) (right) clathrate hydrates which comprise the structure-II unit cell. The blue dotted lines represent hydrogen bonds.

There are 3 simulation parameters that must be optimized to perform our finite temperature PIMD simulations: the simulation time step, dt, which dictates the length of each particle move, the centroid friction, γ^0, which affects the efficiency of the Langevin equation thermostat and the decorrelation time, T_d, which allows us to calculate proper statistical errors for our energies. For a LePIGS simulation, an additional parameter, β, is required to relax our trial wavefunction to the exact ground state wavefunction. Our simulation parameters have been optimized and are reported in Table 3.1. The only exception is for the system of a single hydrogen occupying the large cage, our $T_d = 0.0875$ ps. For our LePIGS simulations, a unity trial wavefunction, $\psi_T = 1$ is used for our hydrogen

molecules.

Table 3.1: System parameters for our finite temperature PIMD calculations and zero-temperature limit LePIGS calculations

Method	dt(fs)	γ^0(ps^{-1})	T$_d$ (ps)	β (K^{-1})
PIMD	0.75 - 3	1/0.9	0.15	N/A
LePIGS	0.5	1/0.9	0.15	0.3

3.3 Results

The main objective of this study is to determine the stable low temperature parahydrogen occupancy within the small (5^{12}) and large ($5^{12}6^4$) cages of the structure-II clathrate hydrates and compare our findings to prior experimental and theoretical studies. We also wish to calculate these energies over a range of temperatures using finite-temperature PIMD and merge the results with the zero-temperature LePIGS calculation.

To determine the stable low-temperature parahydrogen occupancy, the criterion is used that the total energy of the hydrogen-cage system must be negative and that the chemical potential when adding an additional hydrogen must also be negative. If these conditions aren't met, a more stable configuration involves hydrogen molecules leaving the cage, as the zero of energy is defined as the hydrogen molecule at an infinite distance from the cage. The ground state energies from LePIGS simulations are reported below in Table 3.2 and

the chemical potential in Table 3.3 for the small (5^{12}) and large ($5^{12}6^4$) cage, defined as the change in total energy of the system when adding the Nth molecule. The energies, unless explicitly stated, are extrapolated to the $\tau = 0$ limit to remove the error due to the Trotter factorization.

Table 3.2: Ground state energies (kJ/mol) of systems of confined parahydrogen molecules occupying the small (5^{12}) and large ($5^{12}6^4$) cages calculated from LePIGS simulations

Cage Size	$N = 1$	$N = 2$	$N = 3$	$N = 4$	$N = 5$	$N = 6$
Small	-9.193(4)	-1.10(1)				
Large	-6.817(3)	-12.368(7)	-16.704(8)	-19.93(2)	-15.88(2)	-12.06(2)

Table 3.3: Chemical potential (kJ/mol) of parahydrogen molecules occupying the large ($5^{12}6^4$) cage

Cage Size	$N = 1$	$N = 2$	$N = 3$	$N = 4$	$N = 5$	$N = 6$
Small	-9.193(4)	8.09(1)				
Large	-6.817(3)	-5.551(8)	-4.34(1)	-3.22(2)	4.04(2)	3.82(2)

The results of our LePIGS calculations show stable occupancy of up to 1 hydrogen molecule in the small cage and up to 4 hydrogen molecules in the large cage using the Valiron H_2-cage potential and the B-P intermolecular hydrogen potential. This occupancy agrees with the most recent theoretical study done by Bacic[23].

To explain why these are stable occupancies, the distributions of the hydrogen within the cage can be looked at. The probability distribution functions of the distance between the parahydrogen molecule and the center of the small and large cages ($R_{CCM-HCM}$) are

shown in Figure 3.2 and the pair distribution functions of the distance between parahydrogen molecules ($R_HCM - HCM$) are shown in Figure 3.3. For reference, the equilibrium distance between hydrogen molecules is approximately 3.5 Å.

For the small cage, a single parahydrogen molecule can occupy the centre of the cage and has a broad distribution extending to approximately 1.25 Åand maximized at 0.5 Å(or approximately 1.5 Åaway from the edge of the cage). Adding an additional parahydrogen molecule pushes them away from each other and towards the edge of the cage resulting in a more narrow distribution. Contributions to the increase in energy would come from the repulsion of the two hydrogen molecules and either the decrease of or weaker hydrogen-cage interactions.

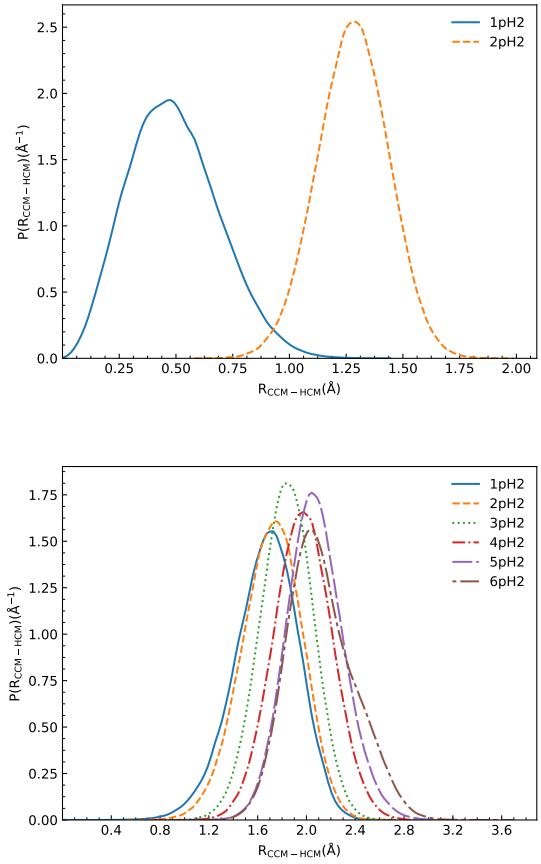


Figure 3.2: Probability distribution functions for the distance between the parahydrogen molecules and the center of the small (5^{12}) cage (top) and large ($5^{12}6^4$) cage (bottom), $R_{CCM} - R_{HCM}$, for various parahydrogen occupancies.

For the large cage, as was the case for the small cage, the equilibrium distance of a single hydrogen molecule is approximately 1.5 Åfrom the edge of the cage, but this time there is no density at the centre. As the cluster size increases, the hydrogen molecules are pushed further away from the centre of the cage but not as close to the edge as in the case

of the small cage. The hydrogen molecules are also not as confined in the larger cage and for almost all cluster sizes are separated by over 3 Å.

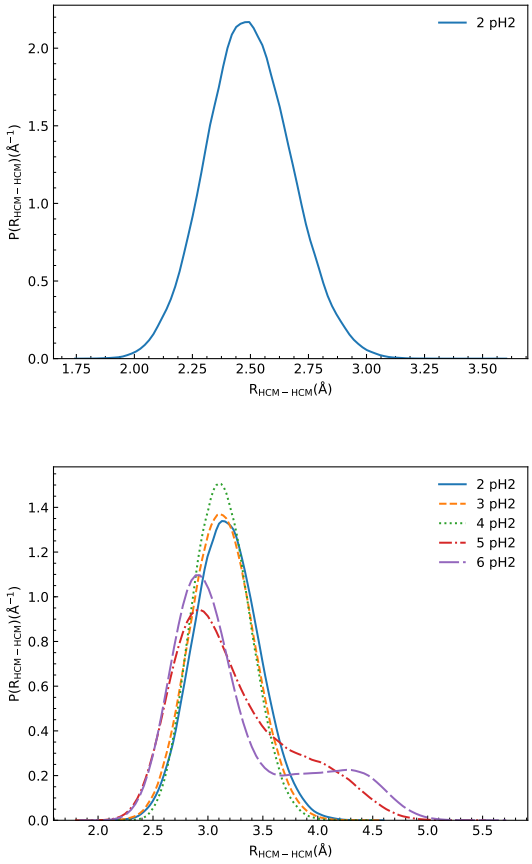


Figure 3.3: Pair distribution functions for the distance between the parahydrogen molecules occupying the small (5^{12}) cage (top) and large ($5^{12}6^4$) cage (bottom), $R_{HCM} - R_{HCM}$, for various parahydrogen occupancies.

However, the distribution overlaps strongest with the equilibrium geometry up to $N = 4$, the most stable hydrogen occupancy. After this, the distribution changes fairly

dramatically with less overlap in the attractive region, which could be the leading cause in the positive turnaround in the chemical potential. In comparison with DMC results[23], we agree with the trend that as cluster size increases the hydrogen molecules are distributed farther from the center of the cage. However, we find that the hydrogen molecules become slightly more closely packed when the cluster size increases, but DMC results show they are pushed slightly apart. As well, Bacic finds the tetrahedral formation of the hydrogen molecules and observed delocalization of this formation as temperature increases about 50 K[25]. We observe the same tetrahedral formation, as can be seen by the angular distribution between the hydrogens and the centre of the cage, shown in Figure 3.4. Each hydrogen forms an approximately 109.5 ° angle from the cage center, indicative of a tetrahedral formation. This delocalization of the tetrahedral formation at higher temperature is not observed in our calculations and thus these discrepancies could be due to the B-P hydrogen interaction potential being stronger than the Hodges potential used by Bacic.

The energies of these clusters are also calculated at various temperatures using finite-temperature PIMD for comparison to the zero-temperature results, using LePIGS, and reported in Table 3.4. For each simulation, the parameter $\tau = \beta/P$ remains constant at $\tau = 0.002\text{K}^{-1}$. The first observation is that the ground state energies are lower in energy than the energies at the lowest temperature, which is expected, although it is observed that the energies converge at temperature around 7 K, which could signify the

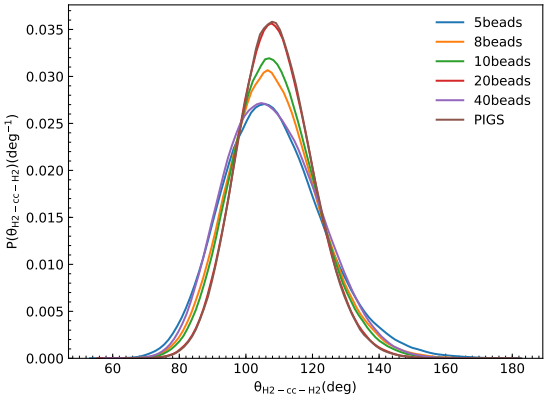


Figure 3.4: Probability distribution functions for the angle between the parahydrogen molecules and the center of the large ($5^{12}6^4$) cage, $\theta_{H_2-cc-H_2}$, for cluster size $N = 4$.

temperature at which the system reaches the ground state. The cause of this discrepancy could be twofold. First, we do not extrapolate our energies to the $\tau = 0$ limit and because the convergence of this PIGS parameter depends on our trial wavefunction (not present in the finite temperature PIMD method), the ground state energy may have a different convergence rate with respect to τ. The second issue could be ergodicity. There are some outlying energies in Table 3.4 (such as 3 pH_2 at $T = 3.7$ K) that are significantly different than the energies at surrounding temperatures, which should not be the case. The cause of this is ergodicity where the hydrogen molecules are trapped in a higher-energy configuration. The PIGS method (or open path sampling) has been demonstrated to have enhanced sampling (or reduced ergodicity issues) compared to closed path sampling methods such as PIMD[81]. Future work should investigate to determine if this discrepancy

is simply due to the difference in τ convergence or ergodicity issues.

Table 3.4: Energies (kJ/mol) of confined parahydrogen molecules within the small (5^{12}) and large ($5^{12}6^4$) cages at various temperatures (using PIMD) and in the zero-temperature limit (using LePIGS). The parameter $\tau = 0.002$ K^{-1} is constant for each simulation.

	Small Cage		Large Cage					
Temperature (K)	1 pH2	2 pH2	1 pH2	2 pH2	3 pH2	4 pH2	5 pH2	6 pH2
0	-9.257(5)	-2.12(2)	-6.856(4)	-12.543(8)	-17.10(1)	-20.58(2)	-17.22(2)	-14.19(2)
0.74	-9.214(1)	-1.985(2)	-6.836(1)	-12.471(2)	-16.973(2)	-20.298(3)	-16.673(3)	-13.772(3)
1.48	-9.211(2)	-1.986(3)	-6.836(2)	-12.469(3)	-16.952(3)	-20.337(4)	-16.671(5)	-13.768(6)
1.85	-9.214(2)	-1.983(4)	-6.832(2)	-12.468(3)	-16.967(4)	-20.123(4)	-16.666(5)	-13.771(6)
2.96	-9.209(2)	-1.984(5)	-6.831(2)	-12.465(4)	-16.964(5)	-20.119(6)	-16.661(7)	-13.764(8)
3.7	-9.213(3)	-1.987(5)	-6.838(2)	-12.472(4)	-16.546(5)	-20.118(6)	-16.823(8)	-13.753(9)
4.736	-9.213(3)	-1.984(6)	-6.833(3)	-12.467(4)	-16.987(5)	-20.116(7)	-16.664(9)	-13.77(1)
5.92	-9.217(3)	-1.988(7)	-6.837(3)	-12.461(5)	-16.731(6)	-20.127(8)	-16.687(9)	-13.76(1)
7.4	-9.211(4)	-1.989(7)	-6.832(3)	-12.446(5)	-16.974(7)	-20.117(8)	-16.71(1)	-13.78(1)
11.84	-9.217(4)	-2.004(9)	-6.814(4)	-12.410(6)	-16.801(8)	-20.13(1)	-16.67(1)	-13.72(1)
14.8	-9.206(4)	-1.98(1)	-6.794(4)	-12.364(6)	-16.914(9)	-20.10(1)	-16.75(1)	-13.71(2)
23.68	-9.196(5)	-1.98(1)	-6.668(5)	-12.205(7)	-16.76(1)	-20.31(1)	-16.67(2)	-13.62(2)
47.36	-8.996(6)	-1.91(2)	-6.236(6)	-11.623(9)	-15.97(1)	-19.71(2)	-15.97(2)	-12.98(3)
59.2	-8.841(7)	-1.79(2)	-6.006(6)	-11.30(1)	-15.46(2)	-19.15(2)	-15.45(2)	-12.45(3)
94.72	-8.304(9)	-1.08(2)	-5.431(7)	-10.22(1)	-13.85(2)	-16.67(2)	-15.22(3)	-13.49(3)
118.4	-7.89(1)	-0.30(2)	-5.049(8)	-9.49(1)	-12.67(2)	-14.83(3)	-13.37(3)	-8.96(3)
236.8	-5.80(2)	4.40(4)	-3.28(1)	2.25(4)	3.11(4)	7.19(4)	9.11(5)	13.41(3)
473.6	-1.71(3)	13.96(6)	5.787(8)	11.779(4)	17.60(1)	23.510(9)	29.36(1)	35.23(1)

3.4 Concluding remarks

The purpose of this chapter is to perform a preliminary study on confined hydrogen clusters in the small (5^{12}) and large ($5^{12}6^4$) cage of the structure-II clathrate hydrate. The zero temperature energies of various cluster sizes are calculated using the LePIGS method to determine the stable occupancy for each cage. It is determined that the most stable occupancies are 1 hydrogen molecule in the small cage and 4 molecules in the large cage. These are both consistent with Bacic's most recent results[23]. The structural properties

of the hydrogen molecules within the cages are also looked at and it is concluded that the B-P interaction between the hydrogen molecules may be too strong, as seen in the case for 4 hydrogen molecules occupying the large cage where delocalization of the tetrahedral formation is not observed at higher temperature, but had been seen by Bacic. Finally, the energies for all stable cluster sizes in both cages are calculated at finite temperature and compared to the ground state results. The energies seem to converge around 7 K, which may suggest that this is when the hydrogen clusters reach the true ground state in the cage confinement.

Chapter 4

Raman vibrational shifts of hydrogen isotopologues

4.1 Introduction

In this chapter,[1] we begin to evaluate the accuracy of the interaction potentials of the individual components, starting with hydrogen. Although the 6-D Hinde pair interaction potential[68] was reduced to 1-D for parahydrogen in previous work, in the present work the H6-PES is reduced to 1-D ground state pair potentials (H1-P) for the other two iso-

[1]Sections of this chapter have been reprinted with permission from M. Schmidt, J. M. Fernndez, N. Faruk, M. Nooijen, R. J. Le Roy, J. H. Morilla, G. Tejeda, S. Montero, and P.-N. Roy, "Raman Vibrational Shifts of Small Clusters of Hydrogen Isotopologues", J. Phys. Chem. A 119, 12551 (2015). Copyright 2015 American Chemical Society.

topologues, orthodeuterium (oD_2) and paratritium (pT_2), by the same method of using the AHR approximation[69] and appropriate internal vibrational averaging. The H1-Ps, in pointwise form, are used to calculate the Raman vibrational shift using the LePIGS method plus the perturbation theory approach described above. We report new experimental values of the Raman vibrational shifts of small $(oD_2)_N$ clusters for $N = 2-6, 9-11$. The predicted Raman vibrational shifts for $(oD_2)_N$ clusters are compared to the new experimental results, while for $(pH_2)_N$ clusters they are compared to earlier experimental results[31]. The shifts predicted here using our H1-Ps are also compared to those calculated using the empirically based, isotope-independent B-P.

It is important to note that our $(pH_2)_N$ shifts calculated here are improved with respect to our previous work[67]. We use a denser mesh for our difference potential and our path integral simulations are performed at smaller imaginary timestep values, however we only notice minor differences in the third decimal place between the values reported previously, so they are still valid. A larger basis set is used here for the dimer and trimer calculations. Also new is quantifying the error in using the AHR approximation and from vibrational averaging to reduce the 6-D potential to 1-D and quantifying the error in first-order perturbation theory. We also introduce a way to scale our results to account for these errors. Results not reported in the previous paper include Raman shifts calculated from simulation for the dimer and trimer, calculated using the $P(R)$ obtained from direct diagonalization

as opposed to simulation for the dimer, and calculated using direct diagonalization and second-order perturbation theory (PT-2) for the dimer.

Although all predicted shifts in this chapter are calculated using the H1-Ps as a tabulated set of points in both the diagonalization and the simulation methods, we also fit these potentials to an analytic Morse Long Range (MLR)[82, 83] functional form denoted H1-MLR. Describing our potentials with MLR functional forms is not only useful due to their compactness, but also convenient for simulations since they have analytic derivatives.

The remainder of this chapter is organized as follows: Theoretical and methodological details are described in Sec. 4.2, experimental vibrational shifts and their theoretical predictions are presented and discussed in Sec. 4.3, and we close with some concluding remarks and future direction in Sec. 4.4.

4.2 Theoretical methodology

4.2.1 Calculating the Raman shifts for a cluster of identical particles

The Raman vibrational shifts for a cluster of identical particles is defined as the energy difference between the cluster with one total vibrational quantum and with zero total

vibrational quanta compared to the energy difference of the free monomers is the Raman shift. In Appendix A, we use a simplified model of our system to calculate the energy levels for clusters of N identical particles by constructing and diagonalizing the many-body Hamiltonian for the $v = 0$ and $v = 1$ vibrational states. We emphasize that it is the symmetric nature of the transition operator, the mean molecular polarizability, which causes the Raman vibrational spectrum and this is largely independent of the bosonic character of the particles. What follows is if we assume the monomers to be identical and to be described by a pair distribution function $P_N(R)$ normalized as

$$\int_0^\infty P_N(R)dR = 1 \tag{4.1}$$

where R is the relative distance between monomers, we can use a first-order perturbation theory approach to calculate the Raman shift. This is done by sampling the difference potential $\Delta V(R)$, defined as the difference in energy of a pair of particles with 1 total quantum of vibration and 0 total quanta of vibration, over the ground state pair distribution

$$\Delta\nu_N = (N-1)\int_0^\infty \Delta V\,(R)\,P_N\,(R)\,dR\,. \tag{4.2}$$

The use of perturbation theory to calculate energy differences for a dopant (e.g. CO_2) in a cluster of monomers (e.g. H_2) has been widely used[84, 85, 86, 87, 88]. The specific case

when the dopant is identical to the monomers, leading to Eq. 4.2, was first considered in Ref. [31].

The quantity $\Delta V(R)$ that enters the theory can be determined explicitly by taking the difference between the ground state energy, E_0, and the symmetric excited vibrational state energy, E_+ and is derived in Appendix A. The actual calculation of E_+ and E_0 involves the construction of the full rovibrational Hamiltonian for the dimer. The energy E_+ is that of the lowest symmetric state with one quantum of vibration and is not necessarily the lowest energy state for the dimer. The full potential for a cluster of N monomers with one quantum of vibration can be expressed in terms of the ground state potential $V(R)$ and the difference potential $\Delta V(R)$ as

$$V_N^{\mathrm{v}=1}(\mathbf{R}) = \sum_{i<j} V(R_{ij}) + \frac{2}{N} \sum_{i<j} \Delta V(R_{ij}) \ , \tag{4.3}$$

where the $2/N$ factor comes from the restriction of counting distinct pairs. This provides a useful starting point for the discussion of perturbation theory later on.

4.2.2 Constructing the effective H1-Ps for the dimer

In practice the monomers within the cluster do not have a single internal degree of freedom. The exact space-fixed frame Hamiltonian for a cluster containing N vibrating and rotating

hydrogen molecules can be written as

$$\hat{H}_N = \sum_i (\hat{h}_i^{\text{rovib}} + \hat{T}_i^{\text{trans}}) + \sum_{i<j} \hat{V}^{\text{aniso}}(r_i, r_j, R_i, R_j) \tag{4.4}$$

where r_i is the relative vector between the two atoms forming hydrogen molecule i, and R_i is a vector representing the position of the centre of mass of that molecule. The diatomic rovibrational Hamiltonian and translational kinetic energy of monomer i are represented by $\hat{h}_i^{\text{rovib}}$ and $\hat{T}_i^{\text{trans}}$ respectively. The intermolecular interactions between molecules i and j are assumed to be pairwise additive and are represented by the $\hat{V}^{\text{aniso}}(r_i, r_j, R_i, R_j)$ term in the above expression. The superscript "aniso" is added to highlight the fact that the full anisotropy is accounted for at this stage. The construction of the H1-P potential for pH_2 was described in detail in a previous paper[67]. The same method is used here to construct the oD_2-oD_2 and pT_2-pT_2 interaction potentials. The Adiabatic Hindered Rotor (AHR) method was developed by Li *et al.*[69] in studying the vibrational shifts of CO_2 in $(pH_2)_N$ clusters and has since been used for other systems including H_2O-pH_2[78] and N_2O-pH_2[89]. The Hamiltonian of our system, composed of two identical monomers, in the space-fixed frame, with a 6-D interaction potential between them is:

$$\hat{H} = \hat{h}_1^{\text{rovib}}(r_1) + \hat{h}_2^{\text{rovib}}(r_2) + \hat{V}^{\text{6-D}}(R, r_1, r_2, \theta_1, \theta_2, \phi) \tag{4.5}$$

The AHR approach involves building the Hamiltonian at various distances, R, between molecules and diagonalizing over the 5 other coordinates. For the ground state potential, the lowest eigenvalues calculated at each R value are chosen resulting in a 1-D interaction potential. However, the individual monomer Hamiltonians $\hat{h}_1$ and $\hat{h}_2$ are not included as part of this AHR treatment, but instead the rovibrational energies (relative to the ground rotational state $J = 0$ for each vibrational level) are used directly. The rovibrational radial wavefunctions are used as a basis set of primitive functions for constructing the Hamiltonian matrix. The quantum numbers $J = 0, 2, 4$ are used for each monomer, reflecting the allowed rotational states for these molecules and symmetrization of the basis functions is incorporated as we are interested in the symmetric state, as discussed before. The basis functions are then converted into real spherical harmonics for convenience. For each system, one H1-P pair interaction potential is constructed with both monomers in their vibrational ground state and a second H1-P pair interaction potential is constructed with one quantum of vibration distributed between the two monomers. The so obtained energies for the ground and first excited vibrational states provide the potentials for the respective simulations.

The ground state Hamiltonian for the translational N-body problem then becomes

$$\hat{H}_g = \hat{T} + \sum_{i<j} E_0(R_{ij}) \ , \tag{4.6}$$

where the ground state $E_0(R_{ij})$ is composed of the individual intramolecular ground state energies h_0 and the intermolecular interaction potential $V(R)$

$$E_0(R) = 2h_0 + V(R) \ . \tag{4.7}$$

The Hamiltonian for the excited state then becomes

$$\hat{H}_{exc} = \hat{H}_g + \sum_{i<j} \Delta V(R_{ij}) \tag{4.8}$$

where

$$\Delta V(R_{ij}) = E_+(R_{ij}) - E_0(R_{ij}) \ . \tag{4.9}$$

and the energies for the excited state E_+ and the ground state E_0 are obtained from the 6-D calculations for the dimer. In the simulations below, $\hat{H}_g$ is used to obtain the ground state density, while the perturbative correction [Eq. 4.2] is used to calculate the vibrational shift from the monomer.

4.2.3 Ground state simulation methodology

Our LePIGS simulations use the parameters and trial wavefunctions from a previous paper that focused on ground state calculations of hydrogen clusters[65]. For simulations of

cluster sizes $N \geq 4$, the Jastrow trial wavefunction, of form

$$\psi_T = \exp\left\{-\frac{1}{2}\sum_{i<j}\left(\frac{b}{r_{ij}}\right)^5\right\} \tag{4.10}$$

with parameter $b = 3.65$ Å[50], is used along with the corresponding parameters from the $N = 4$ case from that paper[65]. For $N = 2 - 3$, the normal mode trial wavefunction, of form

$$\psi_T = \exp\left\{-\frac{1}{2\hbar}\sum_{k}\omega_k Q_k^2\right\} \tag{4.11}$$

is used along with the corresponding parameters for the $N = 4$ case, since it was determined that this trial wavefunction helps prevent the cluster from dissociating. The Q_k's represent the mass-weighted normal modes and the ω_k's represent the angular frequencies determined from the eigenvalues and eigenvectors after diagonalizing the mass-weighted force constant matrix (Hessian). For each cluster size, starting with the initial conditions from the Cambridge Cluster Database[90], a steepest descent minimization is performed using the specific interaction potential to obtain the equilibrium geometries.

For each simulation, regardless of which trial wavefunction is implemented, a simulation time step $dt = 5.0$ fs and an imaginary time step $\tau = 0.001$ K^{-1} is used. The initial conditions for all of our simulations come again from the Cambridge Cluster Database[90], and prior to our simulation that we sample from, a steepest descent minimization and an

equilibration simulation of 25 ps are performed.

4.3 Results and discussion

4.3.1 Experimental Raman vibrational shifts

Raman scattering of small $(oD_2)_N$ clusters has been recorded with the same experimental setup described in Ref. [91]. In brief, $(oD_2)_N$ clusters were formed in a supersonic free jet expansion by a continuous flow of a mixture of $oD_2(10\%)$ diluted in He, at a stagnation pressure of $P_0 =$ 3 bar into a very low pressure expansion chamber. The circular, thin-walled copper nozzle, 50 μm in diameter, was cooled to $T_0 = 40$ K by means of a closed cycle helium refrigerator. The temperature at the nozzle was actively controlled to ± 0.01 K. The $oD_2(10\%)$+He mixture was prepared in continuous flow by a pair of high-accuracy mass-flow controllers (Bronkhorst). The expansion chamber was evacuated by a 2 m^3/s turbomolecular pump, backed by a 400 m^3/hour roots blower and a 70 m^3/hour rotary pump. Background pressure in the expansion chamber was 0.006 mbar under the above given conditions.

oD_2 was prepared from commercial normal D_2 (PRAXAIR, nominal purity 99.8%) by continuous flow through a catalytic converter operated at 21 K and 5 bar. The residence

time under these conditions allowed the para-to-ortho conversion of the D_2 molecules to reach the equilibrium ratio of 97.5 % of $J = 0$ molecules.

Raman scattering was excited by 4 W of a single-mode Ar^+ laser at $\lambda = 514.5$ nm, sharply focused down to a 15 μm beam waist onto the gas jet. The size of the probed volume seen by the detector of the spectrometer, is $90 \times 15 \times 2$ μm^3. Raman scattering was collected at 90 degrees to both the gas jet and the exciting laser beam, by a high throughput ($f/1.8$) optical system with a total magnification $\times 10$, and was focused onto the 20 μm entrance slit of the spectrometer, providing a 2 μm spatial resolution along the jet axis. The spectrometer is equipped with a 2360 line/mm holographic 102×102 mm^2 grating and a back-illuminated CCD detector with 1340×400 pixels, cooled by liquid nitrogen to -120° C. The spectral resolution was 0.16 cm^{-1}.

The Raman spectrum in Fig. 4.1 was recorded at $z = 300$ μm from the nozzle orifice, at an estimated local temperature of 1.1 K in the jet. The peak at 2993.617 cm^{-1} is the Q(0) line[92] of the free oD_2 molecules, with an (off-scale) total signal of ~ 75 photons/s, while that at 2991.498 cm^{-1} is the Q(1) line due to the residual 2.5% pD_2 ($J = 1$) molecules. The red-shifted satellite peaks at the bottom of the monomer Q(0) line are assigned to the excitation of one vibrational quantum in the small $(oD_2)_N$ clusters, as was done for the small $(pH_2)_N$ clusters[31, 93, 91]. Peaks for $N = 2$ to 6 are clearly resolved in the spectrum of Fig. 4.1, while those for $N = 7$ and 8 are partially overlapped by the Q(1) line. The

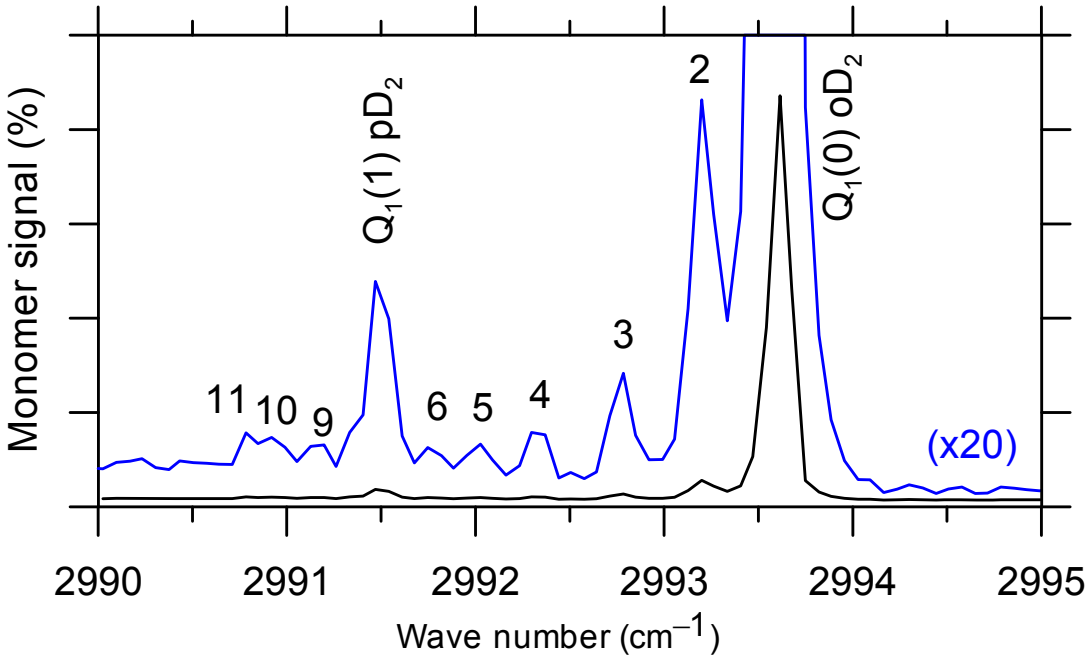


Figure 4.1: Vibrational Raman spectrum of small $(oD_2)_N$ clusters, as recorded in a cryogenic free jet of oD_2(10%)+He mixture. The numbers 2-11 indicate the size N of the cluster. Experimental conditions (see text): $T_0 = 40$ K, $P_0 = 3$ bar, $D = 50$ μm, $z = 300$ μm. The estimated local temperature is 1.1K.

features red-shifted from the Q(1) line are tentatively assigned to the $(oD_2)_N$ clusters with $N = 9$ to 11, although they can also have small contribution from mixed pD_2-oD_2 dimers.

The measured vibrational shifts of the $(oD_2)_N$ clusters with respect to the free oD_2 molecule are reported in Table 4.1.

Table 4.1: Raman vibrational shifts (in cm^{-1}) of $(pH_2)_N$ and $(oD_2)_N$ clusters up to $N = 11$ obtained from simulation [$\Delta\nu$(Sim|PT)] and new experimental values for $(oD_2)_N$. Estimated uncertainties shown in parentheses. Those reported from simulation are standard errors using 1σ. Quantities marked with a * are obtained from exact diagonalization [$\Delta\nu(\psi_0|PT)$]. Results obtained using the B-P are also shown for comparison. The scaling factors used are 1.036 for pH_2 and 1.030 for oD_2. Note: PT refers to first-order perturbation theory

	$(pH_2)_N$				$(oD_2)_N$			
N	Expt.[31]	Scaled	PT	B-P	Expt.	Scaled	PT	B-P
2	-0.400(20)	-0.40	-0.391*	-0.366(4)	-0.401(30)	-0.41	-0.394*	-0.360(3)
3	-0.822(20)	-0.83	-0.801*	-0.756(4)	-0.847(30)	-0.82	-0.795*	-0.731(3)
4	-1.251(20)	-1.25	-1.202(4)	-1.107(5)	-1.287(30)	-1.24	-1.202(3)	-1.096(3)
5	-1.594(20)	-1.57	-1.518(4)	-1.408(5)	-1.602(30)	-1.51	-1.467(3)	-1.341(3)
6	-1.910(20)	-1.86	-1.792(4)	-1.641(5)	-1.844(30)	-1.78	-1.724(3)	-1.572(3)
7	-2.136(20)	-2.09	-2.016(4)	-1.848(5)		-1.99	-1.927(3)	-1.756(3)
8	-2.350(20)	-2.28	-2.199(4)	-2.025(5)		-2.13	-2.063(3)	-1.882(3)
9		-2.46	-2.372(4)	-2.182(5)	-2.457(30)	-2.28	-2.214(3)	-2.021(3)
10		-2.62	-2.534(4)	-2.327(5)	-2.697(50)	-2.46	-2.383(3)	-2.174(3)
11		-2.83	-2.729(4)	-2.512(5)	-2.827(50)	-2.61	-2.532(3)	-2.307(3)

4.3.2 Raman shifts of the dimers and trimers used to quantify methodological errors and determine scaling factors

To determine the Raman vibrational shifts from theory, we reduce the H6-PES to 1-D using the AHR approximation for a pair of monomers with zero total quanta of vibration, our H1-Ps, and with one total quantum of vibration and use a first-order perturbation approach [Eq. 4.2] to calculate the shifts. There are two advantages to doing these steps. First, it is clear that reducing the number of dimensions in our PES leads to computational savings. Also, it has been shown that the convergence of statistical errors with simulation

time using our first-order perturbation approach is faster than simulating the cluster in each of it's vibrational states and directly calculating the difference in energies to get the shift[87]. For the dimer system of both pH_2 and oD_2, we directly calculate both the errors in the reduction from 6-D to 1-D and from using first-order perturbation theory instead of the direct method. We demonstrate the validity of using these two methods, as the cumulative error is smaller than that of experimental uncertainties. This allows us to propose a scaling factor applied to the entire cluster range to account for these two errors we introduce.

Our ground state 1-D pair interaction potentials (H1-Ps) are calculated for three $J = 0$ isotopologues, pH_2, oD_2, and pT_2 and are shown in Fig. 4.2. There are very small differences between them. As the mass increases, the minimum shifts towards shorter lengths and the well depth becomes shallower as the monomers become less polarizable. The difference potentials defined in the last section, are shown in Fig. 4.3. The difference potential becomes shallower as the mass increases since the vibrational energy levels become closer together. The ground state pair distribution functions from our simulations for cluster size $N = 4$ are shown in Fig. 4.4. As the mass increases, the distributions become more peaked reflecting the reduced quantum character of the system.

For the dimer and trimer, in order to calculate the errors in the AHR and perturbation theory, we require the vibrational shifts to be calculated using diagonalization and simu-

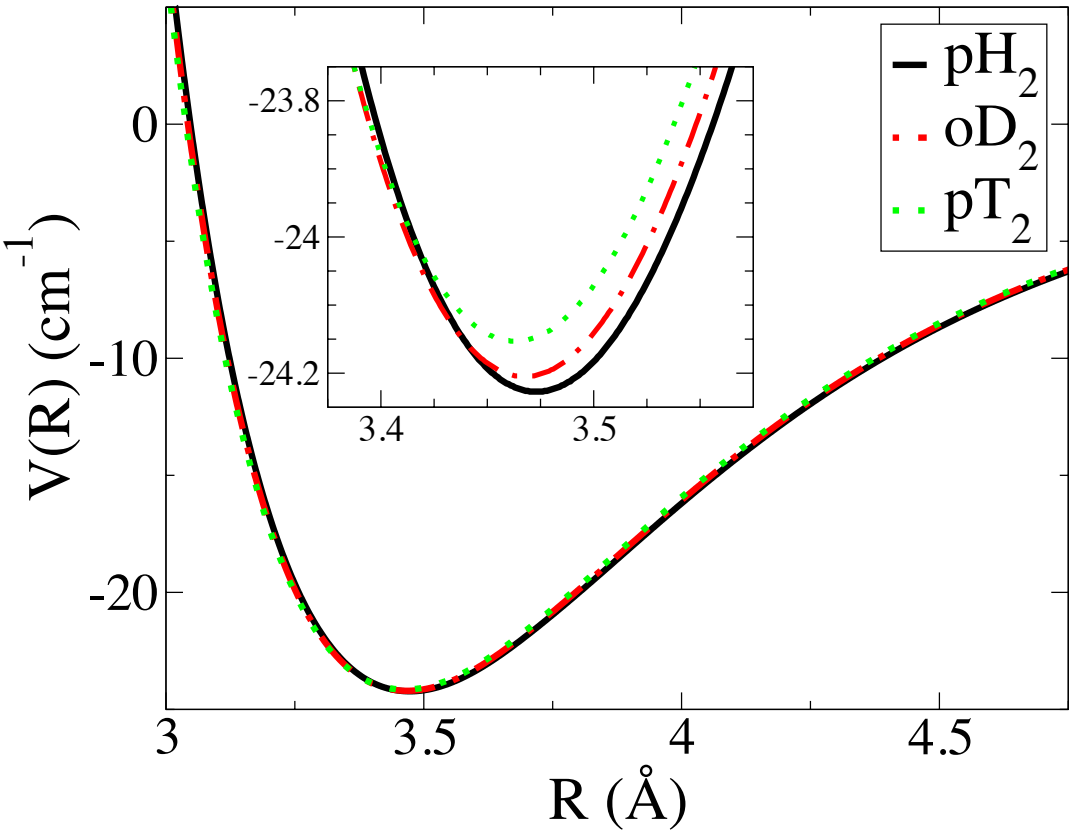


Figure 4.2: H1-Ps of pH_2, oD_2, and pT_2 dimers in the ground ($v = 0$) vibrational state.

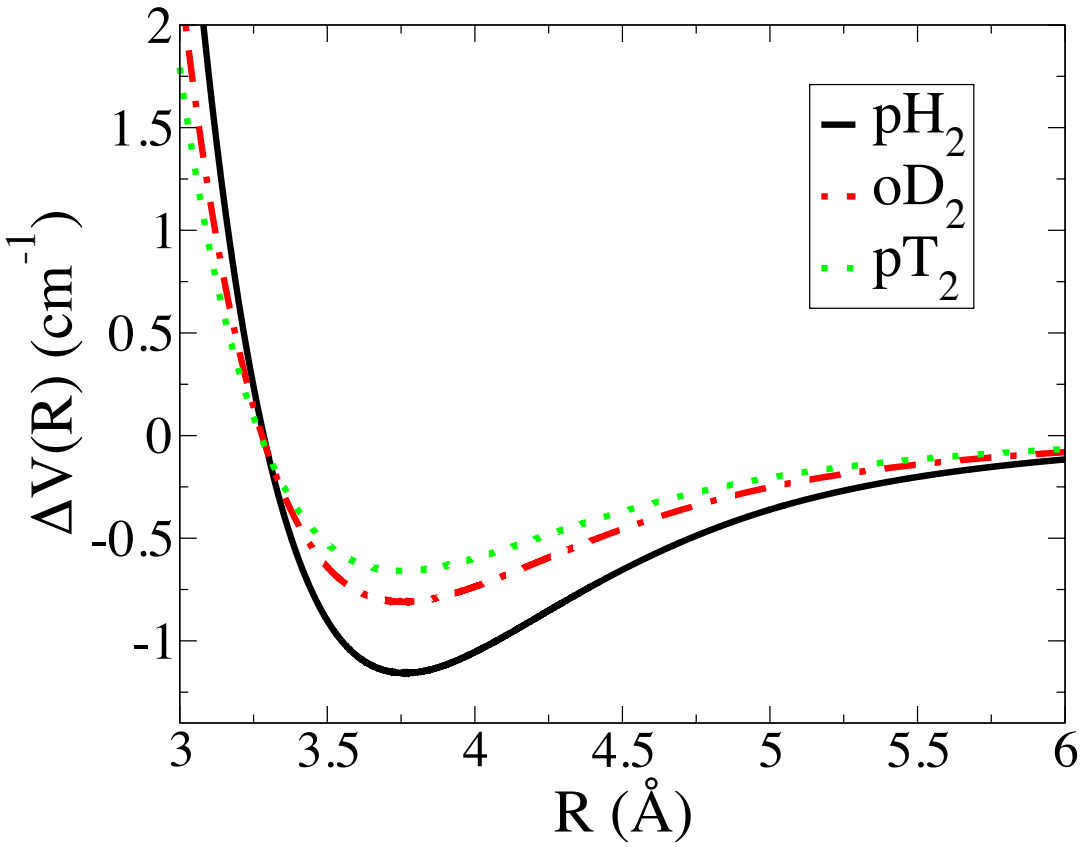


Figure 4.3: Difference between the H1-Ps of pH_2, oD_2, and pT_2 dimers in the first excited ($v = 1$) and the ground ($v = 0$) vibrational states.

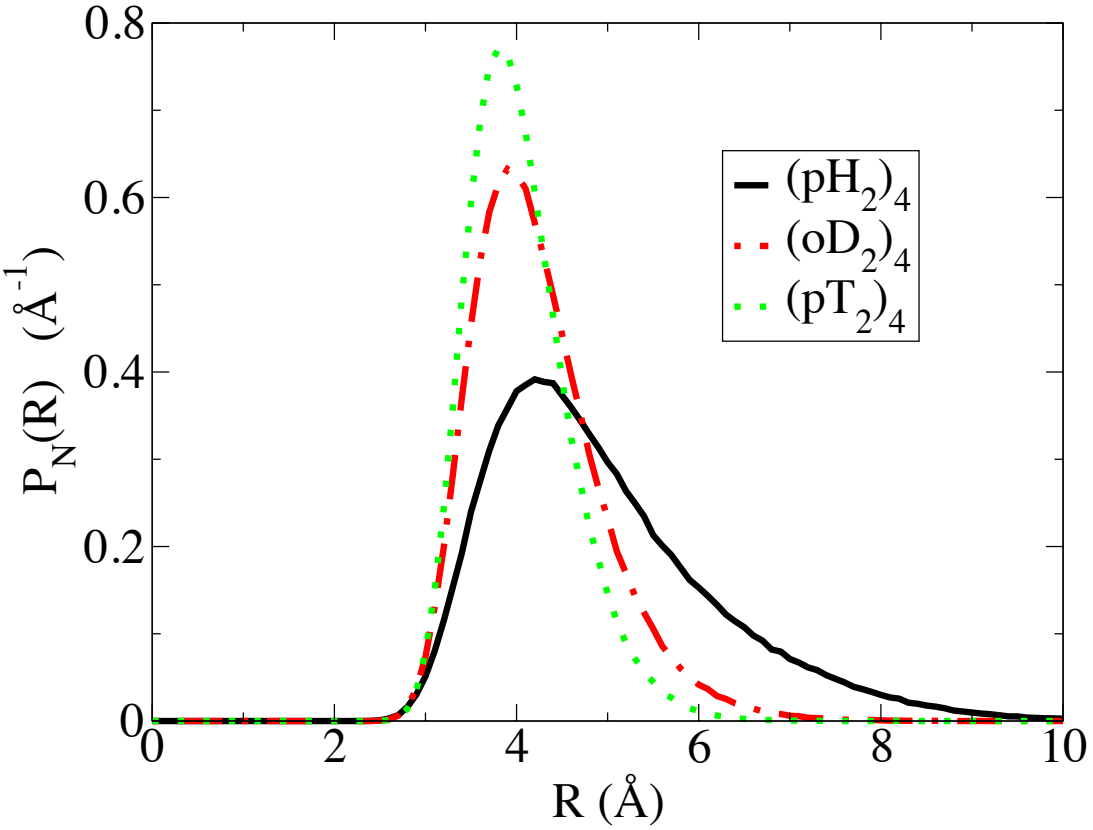


Figure 4.4: Pair distribution functions generated from LePIGS simulations of $(pH_2)_4$, $(oD_2)_4$, and $(pT_2)_4$ clusters using the H1-P for each isotopologue.

lations methods, using the H1-Ps. The most accurate method is to directly calculate the difference between the energy levels, denoted $\Delta\nu(\psi)$

$$\Delta\nu(\psi) = E_0^{\mathrm{v}=1} - E_0^{\mathrm{v}=0} \tag{4.12}$$

where E_0 represents the lowest eigenvalues obtained using the potential with one total quantum of vibration ($v = 1$) given by Eq. 4.3 or our H1-Ps with no excitation ($v = 0$). As mentioned in the previous section, perturbation theory can be applied to calculate the vibrational shifts by using Eq. 4.2. The ground state pair distribution function, $P_N(R)$ can be obtained through LePIGS simulations, and the shifts calculated using this method are denoted $\Delta\nu(\mathrm{Sim}|\mathrm{PT})$. However, from the diagonalizaton methods, we can calculate the pair distribution function directly from the ground state wavefunction, since $P_N(R) = |\psi_0|^2$. The vibrational shifts calculated this way are denoted $\Delta\nu(\psi_0|\mathrm{PT})$.

Second-order perturbation theory[94, 95] can also be used to estimate the leading correction to the first-order perturbation theory method described above. When the wavefunction of interest is that of the ground state, the contribution

$$\Delta\nu''_{0,N} = (N-1)^2 \sum_{k\neq 0} \frac{|\langle\Phi_k^{(0)}|\Delta\hat{V}|\Phi_0^{(0)}\rangle|^2}{E_0^{(0)} - E_k^{(0)}} \tag{4.13}$$

can be added to the $\Delta\nu$ of Eq. 4.2. The k's represent the excited vibrational states of the

system and the $\Phi^{(0)}$'s and $E^{(0)}$'s represent the eigenvectors (wavefunctions) and eigenvalues (energies) obtained by diagonalizing the Hamiltonian using the ground state interaction potential. It should be noted that this correction is negative or zero, meaning that calculating the Raman shifts using only first-order perturbation theory should give results that are smaller in magnitude than the true value of the shift. We use direct diagonalization to obtain the energies and wavefunctions for the dimer. The shifts calculated using this method are denoted $\Delta\nu(\psi_0|\text{PT-2})$.

Using these direct diagonalization methods, it is important to recognize that the reference frame is the body-fixed frame, the system is in the $J = 0$ state, and the energies we calculate are due to the intermolecular translational motion. In simulations, the reference frame is the space-fixed frame. In the analysis throughout this chapter, it should be noted that the shifts are compared based on their magnitude. Since only red shifts are observed, a larger shift means one that is more negative.

We first start with the dimer, which can be discussed in more detail. An exact method can be used to calculate the vibrational shift by diagonalizing the Hamiltonian for the v=0 and v=1 states using a Discrete Variable Representation (DVR) basis[96]. These results, together with the simulation results, experimental values, and Hinde's exact 6-D calculations are shown in Table 4.2. The parameters used in our diagonalization code are 10000 DVR grid points, r_{min}=0.1 Å r_{max}=20.0 Å for $(pH_2)_2$ and $(oD_2)_2$ and r_{max}=15.0 Å

for $(pT_2)_2$, and J_{max}=0 is used for all isotopologues. We note that our bound state energies are relative to that of the two free monomers.

Table 4.2: Ground state energies and Raman vibrational shifts (in cm^{-1}) of $(pH_2)_2$, $(oD_2)_2$, and $(pT_2)_2$ as calculated using our H1-Ps. Comparison is made to Hinde's results using his H6-PES Hinde 6-D Exact [68], new experimental data for $(oD_2)_2$, and previous experimental results for $(pH_2)_2$[31]. $\Delta\nu(\psi)$ refers to the method of using direct diagonalization to calculate the shift by directly taking the difference between the ground state energy [E_0 (ψ_0)] and the energy of the excited vibrational state. Using exact diagonalization, the shifts are calculated using first order perturbation theory [$\Delta\nu(\psi_0|\text{PT})$] and second order perturbation theory [$\Delta\nu(\psi_0|\text{PT-2})$]. The shifts are also calculated using the pair distribution functions from simulation [$\Delta\nu(\text{Sim}|\text{PT})$]. Estimated uncertainties are shown in parentheses. Those reported from simulation are standard errors using 1σ.

	$(pH_2)_2$	$(oD_2)_2$	$(pT_2)_2$
$E_0(\psi_0)$	-2.90213	-6.78567	-9.055
$\Delta\nu$(Hinde 6-D Exact)	-0.405	-0.406	N/A
$\Delta\nu(\psi)$	-0.401	-0.400	-0.363
$\Delta\nu(\psi_0\|\text{PT})$	-0.391	-0.394	-0.359
$\Delta\nu(\text{Sim}\|\text{PT})$	-0.388(3)	-0.400(2)	-0.363(2)
$\Delta\nu(\psi_0\|\text{PT-2})$	-0.401	-0.400	-0.363
$\Delta\nu$(Experiment)	-0.400(20)	-0.401(30)	N/A

The difference of our $\Delta\nu(\psi)$ from the $\Delta\nu$(Hinde 6-D Exact) can be directly attributed to the error in using the AHR approximation and vibrational averaging to reduce our potential from 6-D to 1-D. These errors lead to shifts that are 0.004 cm^{-1} smaller for $(pH_2)_2$ and 0.006 cm^{-1} smaller for $(oD_2)_2$ than Hinde's 6-D values. The difference in our $\Delta\nu(\psi_0|\text{PT})$ and $\Delta\nu(\psi)$ is the error in using perturbation theory for these predictions. We observe that this error leads to even smaller shifts by 0.01 cm^{-1} for $(pH_2)_2$, 0.006 cm^{-1} for

$(oD_2)_2$, and 0.004 cm^{-1} for $(pT_2)_2$. These two effects might explain why our $\Delta\nu$(Sim|PT) predictions for larger clusters are smaller than experimentally observed. This error is greatly eliminated with second order perturbation theory, $\Delta\nu(\psi_0|\text{PT-2})$, as it reproduces the shift using our most accurate method, $\Delta\nu(\psi)$. As expected, the shifts obtained from first order perturbation theory, either from simulation, $\Delta\nu$(Sim|PT), or from the exact calculation, $\Delta\nu(\psi_0|\text{PT})$, are in good agreement with each other.

Since we know the error due to the AHR approximation and to the perturbation theory and they are both in the same direction (cause smaller shifts), the predicted shifts calculated using the $P_N(R)$ from simulation [$\Delta\nu$(Sim|PT)] and from direct diagonalization [$\Delta\nu(\psi_0|\text{PT})$] can be scaled to correct for these errors. Using the shifts calculated for the dimer, $\Delta\nu_2$, the scaling factor f for $(pH_2)_N$ and $(oD_2)_N$, is calculated as the ratio:

$$f^{H/D} = \frac{\Delta\nu_2(\text{Hinde 6-D Exact})}{\Delta\nu_2(\text{PT})} \ . \tag{4.14}$$

For pT_2, there is no Hinde 6-D Exact calculation so we instead scale to our [$\Delta\nu(\psi)$] value using Lanczos diagonalization which accounts only for the error in perturbation theory but not the error due to the AHR approximation.

$$f^{T} = \frac{\Delta\nu_2(\psi)}{\Delta\nu_2(\text{PT})} \ . \tag{4.15}$$

The vibrational shift for the trimers can also be calculated exactly and perturbatively using the Lanczos diagonalization in Pekeris co-ordinates[97, 98]. These shifts are shown in Table 4.3, along with the simulation and experimental results. For each of the isotopologues, 90 dimensionless Jacobi DVR basis functions were used over 15000 iterations with a $V_{\max}$ of 1000 cm^{-1} for all cases except for the Lanczos-Exact calculation with $(pH_2)_3$ ($V_{\max}$=1100 cm^{-1}). For $(pH_2)_3$ and $(oD_2)_3$, $r_{\max}$ = 20 Å and for $(pT_2)_3$ $r_{\max}$ = 15 Å. The basis function parameters used for all isotopologues are $a = 5$ and $b = 0$ with symmetry label A_1. The first observation is the error in perturbation theory, or the deviation of

Table 4.3: Raman vibrational shifts (in cm^{-1}) of $(pH_2)_3$, $(oD_2)_3$, and $(pT_2)_3$ as calculated using our H1-Ps. Comparison is made to new experimental data for $(oD_2)_3$, and previous experimental results for $(pH_2)_3$[31]. $\Delta\nu(\psi)$ refers to the method of using direct diagonalization to calculate the shift by directly taking the difference between the ground state energy and the energy of the excited vibrational state. The shifts are also calculated using first order perturbation theory with the pair distribution functions determined by the exact wavefunctions [$\Delta\nu(\psi_0|\text{PT})$] or from the results of simulation [$\Delta\nu(\text{Sim}|\text{PT})$]. Estimated uncertainties are shown in parentheses. Those reported from simulation are standard errors using 1σ.

	$(pH_2)_3$	$(oD_2)_3$	$(pT_2)_3$
$\Delta\nu(\text{Sim}\|\text{PT})$	-0.798(4)	-0.792(3)	-0.722(2)
$\Delta\nu(\psi_0\|\text{PT})$	-0.801	-0.795	-0.723
$\Delta\nu(\psi)$	-0.814	-0.803	-0.728
$\Delta\nu$(Experiment)	-0.822(20)	-0.847(30)	N/A

$\Delta\nu(\psi_0|\text{PT})$ from the non-perturbative exact calculation, $\Delta\nu(\psi)$, is 0.013 cm^{-1} for $(pH_2)_3$, 0.008 cm^{-1} for $(oD_2)_3$, and 0.005 cm^{-1} for $(pT_2)_3$. We also confirm the same values of the

shifts are obtained using the two different perturbation theory methods, $\Delta\nu$(Sim|PT) and $\Delta\nu(\psi_0|\mathrm{PT})$.

4.3.3 Raman vibrational shifts for entire cluster range

We now extend our study to larger clusters, where direct diagonalization methods can no longer be used. Our predicted Raman vibrational shifts for $(\mathrm{pH_2})_N$ and $(\mathrm{oD_2})_N$ obtained using the H1-Ps over the range $N = 2 - 11$ are compared to experiment in Fig. 4.5. The numerical results are given in Table 4.1. The shifts predicted for $(\mathrm{pT_2})_N$ are given in Table 4.4, and they are smaller than those for both $(\mathrm{pH_2})_N$ and $(\mathrm{oD_2})_N$ for all cluster sizes.

The following discussion is based on the scaled values. The results using the B-P are also calculated for comparison. The shifts reported in Table 4.1, which were calculated with our H1-P for $(\mathrm{pH_2})_N$ are in good agreement with the experimental data and fall within the uncertainty for $N = 2 - 5$. For $N \geq 6$, the predicted shifts are distinctly smaller than the observed ones. This may be due to a number of possible factors. First, we introduce an *ad hoc* scaling factor for all cluster sizes based on the dimer to eliminate the effect of the AHR approximation and first order perturbation theory. Future work could be done to find a theoretically justified method to properly account for the errors due to those effects. Hinde's 6-D potential also uses the same long range coefficients for $\mathrm{D_2}$ as it does for $\mathrm{H_2}$. This may contribute to the deviation in the shifts of $(\mathrm{oD_2})_N$ clusters between

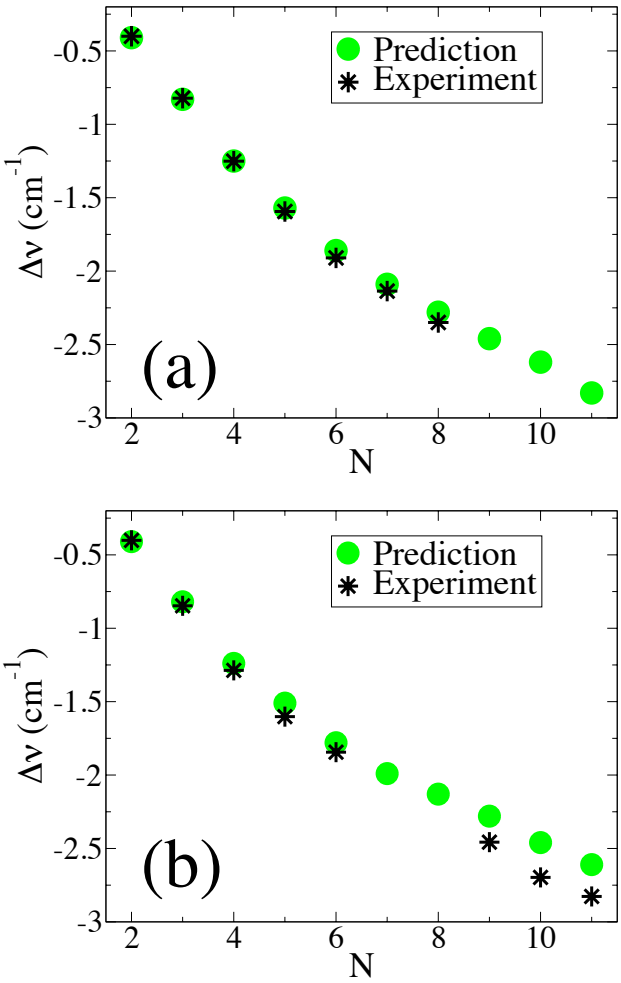


Figure 4.5: Predicted Raman vibrational shifts (with scaling) for (a) $(pH_2)_N$ and (b) $(oD_2)_N$ clusters up to size $N = 11$ using the H1-P for each isotopologue. Experimental results are also shown.

Table 4.4: Raman vibrational shifts (in cm^{-1}) of $(pT_2)_N$ clusters up to $N = 11$ obtained from simulation [$\Delta\nu$(Sim|PT)] using our numerical H1-Ps and difference potentials for (pT_2) dimers. Estimated uncertainties, reported as standard errors using 1σ, are shown in parentheses. Quantities marked with a * are obtained from exact diagonalization [$\Delta\nu(\psi_0|PT)$] using the H1-Ps. Results obtained using the B-P potential are also shown for comparison. The scaling factor used is $f^T = 1.011$ for pT_2.

	$(pT_2)_N$		
N	Scaled	PT	B-P
2	-0.36	-0.359*	-0.329(2)
3	-0.73	-0.723*	-0.670(2)
4	-1.10	-1.092(2)	-1.001(3)
5	-1.34	-1.327(2)	-1.223(3)
6	-1.57	-1.552(2)	-1.430(3)
7	-1.77	-1.755(2)	-1.611(3)
8	-1.86	-1.840(2)	-1.695(3)
9	-2.04	-2.013(2)	-1.842(2)
10	-2.18	-2.157(2)	-1.977(2)
11	-2.28	-2.254(2)	-2.069(2)

our calculated shifts and experiment. Also, we are using pair interaction potentials and perhaps many-body effects are important for the larger clusters.

It is also interesting to compare the relative magnitude of the vibrational shifts for $(pH_2)_N$ and $(oD_2)_N$. For $N \leq 5$, no difference in the experimental vibrational shift between $(pH_2)_N$ and $(oD_2)_N$ is observed within the estimated uncertainty, while for larger clusters, the experimental shift for $(pH_2)_N$ are larger than for $(oD_2)_N$. In turn, the ordering of the shifts for $(pH_2)_N$ and $(oD_2)_N$ agree with experiment; with perhaps the exception of $N = 5$, since $(pH_2)_5$ is predicted to have a larger shift than $(oD_2)_5$ by 0.06 cm^{-1} while the shifts

are experimentally observed to be almost equal.

For $N \leq 5$ clusters of $(oD_2)_N$ and $(pH_2)_N$ we notice a competing effect in the calculation of the vibrational shift from perturbation theory, shown in Fig. 4.6, due to the shape and location of the pair distribution function with respect to the difference potential. This figure plots the integrand of the vibrational shift equation [Eq. 4.2] for the $(pH_2)_3$, $(oD_2)_3$, and $(pT_2)_3$ trimers. The figure shows that $(oD_2)_3$ has a significantly, narrowly peaked pair distribution in the most negative region of the difference potential contributing a very negative value to the shift. The $(pH_2)_3$ pair distribution in the difference potential well is not as peaked and extends to longer distances, so although it does not contribute as negative a shift in the well region, since the distribution extends to longer distances it accumulates a more negative value. This leads to roughly equivalent values of the overall shift between $(pH_2)_N$ and $(oD_2)_N$.

For each isotopologue, the B-P gives consistently smaller vibrational shifts than our H1-Ps and as a result the H1-Ps with isotopologue-dependence give better agreement for all cluster sizes.

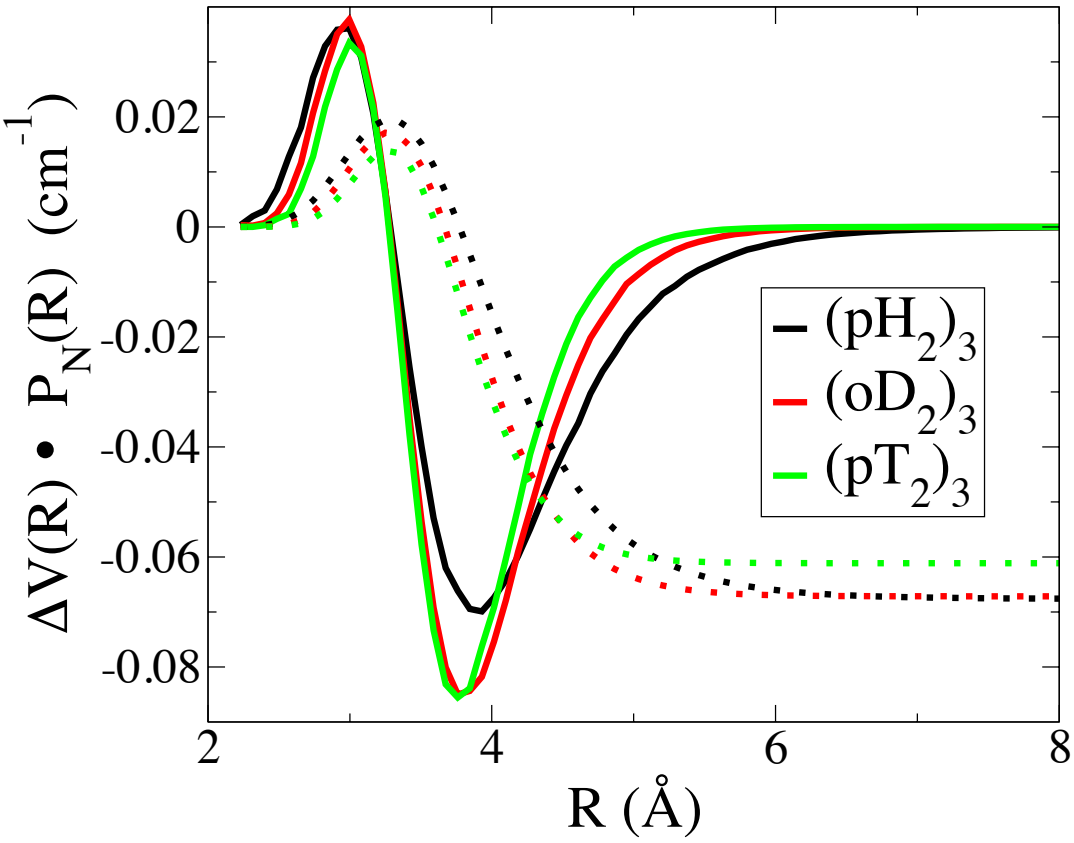


Figure 4.6: Solid lines: Integrand of the vibrational shift equation: the difference potential multiplied by the pair distribution function obtained from simulation, shown for the $(pH_2)_3$, $(oD_2)_3$, and $(pT_2)_3$ trimers. Dotted lines: The cumulative sums of the integrand (scaled to fit the diagram) reflecting how the vibrational shifts change with intermolecular distance.

4.3.4 Portable analytic MLR models for our isotope-dependent H1-Ps

Although the ground state H1-Ps used so far in this chapter, both in the diagonalization and simulation methods, are the set of points obtained directly from the AHR reduction and vibrational averaging procedure, it is convenient to express our potentials in an analytic form. The potentials are fitted to the modern version of the MLR functional form[99]:

$$V_{MLR}(r) = D_e \left\{ 1 - \frac{u_{LR}(r)}{u_{LR}(r_e)} \exp\left[-\beta(r) \cdot y_p^{\text{eq}}(r) \right] \right\}^2 \tag{4.16}$$

with D_e representing the well depth and r_e representing the equilibrium distance. The β parameters are defined as:

$$\beta_{MLR}(r) = y_p^{\text{ref}}(r)\,\beta_\infty + \left[1 - y_p^{ref}(r) \right] \sum_{i=0}^{N_\beta} \beta_i y_q^{\text{ref}}(r)^i \tag{4.17}$$

in which

$$\beta_\infty = \ln\left[\frac{2D_e}{u_{LR}(r_e)} \right] \tag{4.18}$$

and the y_a's are:

$$y_a^{\text{eq}}(r) = \frac{r^a - r_e^a}{r^a + r_e^a} \tag{4.19}$$

and

$$y_a^{\text{ref}}(r) = \frac{r^a - r_{ref}^a}{r^a + r_{ref}^a} \ . \tag{4.20}$$

The long range contribution has the form:

$$u_{LR}(r) = D_{m_1}(r)\frac{C_{m_1}}{r^{m_1}} + D_{m_2}(r)\frac{C_{m_2}}{r^{m_2}} + D_{m_3}(r)\frac{C_{m_3}}{r^{m_3}} + ... \tag{4.21}$$

in which the $D_m(r)$'s are "damping functions" from refs [99, 100].

$$D_{m_i}^{\text{ds}}(r) = \left\{1 - \exp\left(-\frac{3.30(\rho r)}{m_i} - \frac{0.423(\rho r)^2}{m_i^{1/2}}\right)\right\}^{m_i - 1} \ . \tag{4.22}$$

The programs betaFIT [101] and dpotFIT [102] are used to fit the H1-Ps to analytic MLR functions. To fit our potential in betaFIT, 756 points out of 17770 points were chosen from the original tabulated potential used in simulation. A required input to any sensible least squares fit program are "uncertainties" at each potential point. The uncertainties used here are a specific percentage of the value of the potential at each data point. The choice of points and uncertainties is listed in Table 4.5.

Our H1-MLRs are fit to ten beta parameters and to optimize the remaining parameters, they were fit using the betaFIT code program for a variety of different p (between 5-8), q (between 1-4), and r_{ref} parameters (between 3.00 Å - 6.00 Å) for each of the three interaction

Table 4.5: Selecting range and density of the tabulated H1-P data points to fit the analytic potential to and the choice of uncertainty percentages.

R (Å)	ΔR (Å)	"Uncertainty" (%)
2.23 - 3.08	0.05	10%
3.10 - 8.08	0.01	1%
8.13 - 19.98	0.05	10%

potential surfaces. The parameters were chosen such that they minimize the dimensionless root mean squared deviation (so-called '*dd*') between the fitted H1-MLRs and actual H1-Ps. The potFIT code was then used to further fine-tune the H1-MLRs, requiring that the dissociation limit be zero, which was not possible in betaFIT. The resulting '*dd*' values for each potential were approximately 0.02. The resulting potential surfaces obtained from potFIT are benchmarked by comparing the bound state energies of $(pH_2)_3$, $(oD_2)_3$, and $(pT_2)_3$ trimers obtained using our fitted H1-MLRs to those obtained using the true tabulated H1-P. It is most convenient to have the same MLR form (p, q, r_{ref} parameters) for each isotopologue and the parameters $p = 6, q = 1$, and $r_{\text{ref}} = 4.20$ Å were chosen since these gave us the best overall results for each of the potential surfaces. The other MLR fitted parameters for each isotopologue are listed in Table 4.6. Comparing the H1-P to the H1-MLRs of the dimers in Table 4.7, the accuracy of the energies are on the order of the 0.0001 cm^{-1} and the errors in our calculated vibrational shifts are 0.002 cm^{-1}, which are smaller than our simulated error bars. We find that the calculated ground state energies are the only bound states found using our H1-Ps and H1MLR-Ps for all three isotopologues. As

Table 4.6: MLR fit parameters for the ground state H1-Ps of pH_2, oD_2, and pT_2. Note all energies are in cm^{-1} and lengths in Å. Parameters common for all three H1-MLRs are $p = 6$, $q = 1$, $r_{\rm ref} = 4.60$, and $\rho^{AB} = 1.10$.

Parameter	pH_2	oD_2	pT_2
D_e	24.2288	24.2090	24.1595
r_e	3.47005	3.46468	3.46043
$V_{\rm min}$	-24.2288	-24.2090	-24.1595
C_6		5.82036400E+04	
C_8		2.87052154E+05	
C_{10}		1.80757343E+06	
β_0	-6.631E-02	-6.396E-02	-6.202E-02
β_1	1.346E-01	2.109E-01	2.795E-01
β_2	-3.300E-02	4.970E-01	9.750E-01
β_3	6.000E+00	7.410E+00	8.810E+00
β_4	-1.400E+01	-1.570E+01	-1.730E+01
β_5	-1.193E+02	-1.020E+02	-8.840E+01
β_6	2.290E+02	4.110E+02	5.880E+02
β_7	1.110E+03	1.160E+03	1.230E+03
β_8	-1.850E+03	-3.150E+03	-4.460E+03
β_9	-3.500E+03	-3.300E+03	-3.200E+03
β_{10}	6.000E+03	1.100E+04	1.600E+04

mentioned before, the bound states refer only to the intermolecular translational degrees of freedom. Using his H6-PES, Hinde obtained for $(oD_2)_2$ four bound states[68] considering the end-over-end rotation of the dimer. The ground state wavefunctions for the $(pH_2)_2$, $(oD_2)_2$, and $(pT_2)_2$ dimers calculated from the exact code using our H1-Ps are shown in Fig. 4.7. It should be noted that the $(pH_2)_2$ wavefunction is much broader than the other two. This reflects the greater quantum nature of $(pH_2)_2$ compared to it's heavier isotopologues.

The results of the trimer bound state energy calculations[97, 98], performed using the Lanczos diagonalization scheme, are shown in Table 4.8 using both the full pointwise H1-P and our fitted H1-MLR. We find two bound states for $(pH_2)_3$, three for $(oD_2)_3$ and four

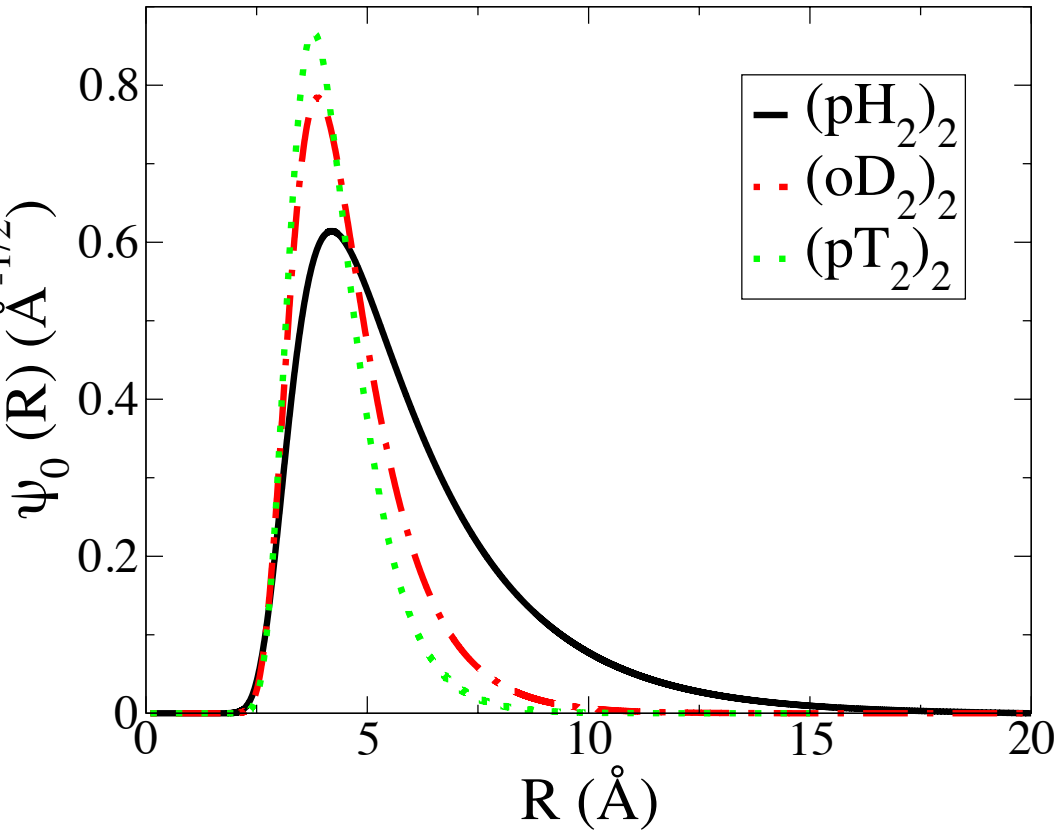


Figure 4.7: Ground state wavefunction of the $(pH_2)_2$, $(oD_2)_2$, and $(pT_2)_2$ dimers calculated by exact diagonalization using the H1-Ps.

Table 4.7: Rotationless bound state energies and vibrational shifts (in cm^{-1}) of $(pH_2)_2$, $(oD_2)_2$, and $(pT_2)_2$ dimers calculated using the Lanczos diagonalization method with our H1-P and the present H1-MLR. Energies are relative to the two free monomers.

	$(pH_2)_2$		$(oD_2)_2$		$(pT_2)_2$	
	H1-P	H1-MLR	H1-P	H1-MLR	H1-P	H1-MLR
$E_0(\psi_0)$	-2.9021	-2.9016	-6.7857	-6.7847	-9.055	-9.054
$\Delta\nu(\psi)$	-0.4013	-0.4018	-0.3997	-0.4006	-0.363	-0.365

Table 4.8: Rotationless bound state energies E_n and vibrational shifts (in cm^{-1}) of $(pH_2)_3$, $(oD_2)_3$, and $(pT_2)_3$ trimers calculated using the Lanczos diagonalization method with our H1-P and the present H1-MLR. Energies are relative to the three free monomers.

	$(pH_2)_3$		$(oD_2)_3$		$(pT_2)_3$	
Bound State Energies	H1-P	H1-MLR	H1-P	H1-MLR	H1-P	H1-MLR
E_0	-9.8863	-9.8864	-21.2386	-21.2388	-27.905	-27.904
E_1	-3.65104	-3.65098	-12.7640	-12.7634	-18.076	-18.075
E_2			-7.2730	-7.2727	-13.923	-13.922
E_3					-9.483	-9.485
$\Delta\nu(\psi_0\|PT)$	-0.8010	-0.8011	-0.7954	-0.7956	-0.72272	-0.72270

for $(pT_2)_3$. We note that all bound states of the trimers lie below those of the dimers. Our fitted H1-MLRs give good agreement with the original H1-Ps for the bound state calculations and the vibrational shifts. It is worth mentioning that similar calculations of the bound states of $(pH_2)_3$ in the literature[103, 104] found only one bound state. This could be due to the interaction potential and future work should be done to determine the source of this discrepancy.

The one-dimensional particle distributions representing the wavefunctions in Pekeris

co-ordinates are shown in Fig. 4.8 for the trimer of each isotopologue. The wavefunction for the bound state $n = 1$ level for the $(pH_2)_3$ trimer is almost twice as wide as that for the ground state wavefunction. This extended wavefunction is not observed in the $(oD_2)_3$ trimer, however a diffuse wavefunction is also observed in the $n = 3$ level of the $(pT_2)_3$ trimer that is much broader than the other three lower states.

4.4 Concluding remarks

In this chapter, effective 1-D pair interaction potentials for pH_2, oD_2, and pT_2 have been constructed based on the H6-PES[68]. Using these reduced pointwise H1-Ps, the Raman vibrational shifts were calculated using ground state molecular dynamics (LePIGS) simulations and applying first order perturbation theory. For $(pH_2)_N$ and $(oD_2)_N$, the predictions using the H1-Ps are in good agreement with experimental results and outperform the traditional B-P. This suggests that the hydrogen interaction potential that should be used in the future for small clusters should be the H1-Ps. We also present new experimental values for the Raman vibrational shifts for $(oD_2)_N$ clusters of size $N = 2 - 6, 9 - 11$.

For $(pH_2)_2$ and $(oD_2)_2$ dimers we calculate the error in reducing the H6-PES to our H1-Ps and the error in using perturbation theory. Both errors cause smaller predicted shifts. We find that second order perturbation theory corrects for the error of first order pertur-

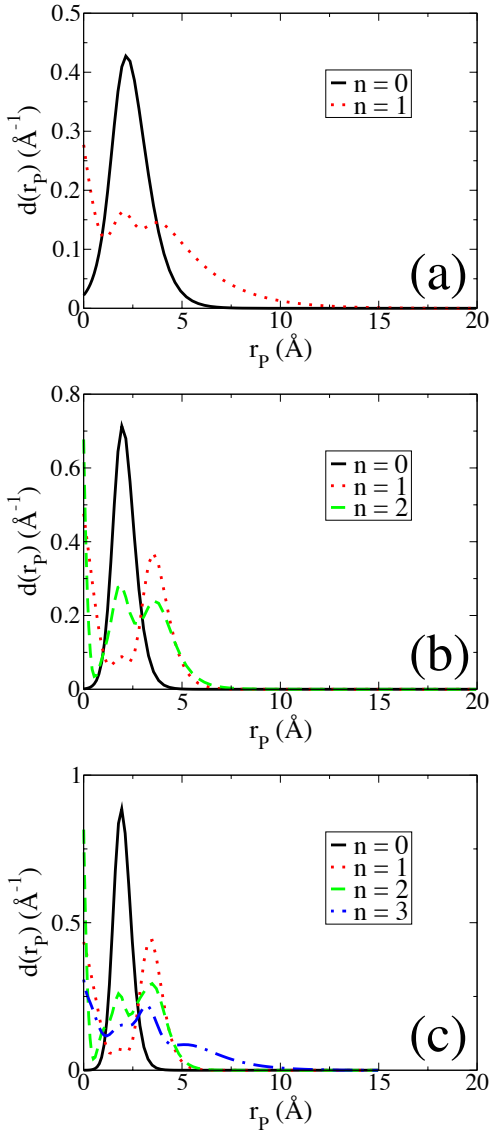


Figure 4.8: One-dimensional particle distribution for energy levels E_n in Pekeris coordinates of (a) $(pH_2)_3$, (b) $(oD_2)_3$, and (c) $(pT_2)_3$ trimers using Lanczos diagonalization scheme.

bation theory. A potential source of error for larger clusters is that the AHR reduction does not necessarily lead to a pairwise additive potential. Determining the quantitative effect of this error or if this assumption introduces many-body effects is a future research direction. For the trimers, we observe differing numbers of rotationless bound states of the trimer isotopologues: two for $(pH_2)_3$, three for $(oD_2)_3$, and four for $(pT_2)_3$, however only one bound state for $(pH_2)_3$ has been reported in the literature[103, 104] using similar methods to us, and future work should be used to determine the source of the discrepancy. We also observe some of the bound state wavefunctions for $(pH_2)_3$ and $(pT_2)_3$ extend to distances over twice that of the ground state wavefunctions. This is not seen for any of the $(oD_2)_3$ bound state wavefunctions.

Since our errors in the AHR approximation and perturbation theory are known, our simulation results can be scaled by a factor to account for the discrepancy. Our calculated shifts of small clusters, $N \leq 5$, of $(pH_2)_N$ and $(oD_2)_N$ agree with experiment within estimated uncertainty, but are smaller than experiment for larger clusters. This could be due to our naïve method of using a scaling factor to account for the errors in using the AHR approximation and first order perturbation theory. Also, the same long range coefficients are used for D_2 as they are for H_2 and this could lead to additional errors for the $(oD_2)_N$ shifts. We also use a pair interaction potential and the deviations observed for the larger clusters could be due to many-body effects. Future work could look at a theoretical method

to properly account for the errors introduced by the AHR approximation and first-order perturbation theory, as well as trying to quantify the three-body or many-body effects to determine their impact on the vibrational shifts. For $(pT_2)_N$, there are no experimental results, however a similar pattern is observed as for the other isotopologues that the H1-P gives larger shifts than the B-P for all cluster sizes. The predicted shifts for $(pT_2)_N$ clusters are all smaller than the corresponding shifts for both $(pH_2)_N$ and $(oD_2)_N$.

Although the analysis above used the ground state H1-P as a set of points directly obtained from the AHR reduction, our H1-Ps are also fitted to a Morse Long Range (MLR)[82, 83] functional form denoted H1-MLR for later convenience. The dimer and trimer bound state energies and vibrational shifts with first order perturbation theory using the fitted H1-MLRs are compared with those using the H1-Ps directly and the results are in excellent agreement. Having an analytic form for our interaction potentials is very useful for molecular dynamics simulations because they allow the derivatives (forces) to be calculated analytically. It also provides a compact way to represent the potential for dissemination purposes, instead of a tabulated form.

Chapter 5

Analysing the accuracy and efficiency of different methods to calculate Raman vibrational shifts of hydrogen clusters

5.1 Introduction

In this chapter, we continue to look at the theoretical calculations of the Raman vibrational frequency shifts of pure parahydrogen and orthodeuterium clusters of $N = 4-9$, expanding

on the previous chapter. Here, the shifts are calculated using three different methods and the values are compared to experiment and the variances are compared to each other. The first method requires the calculation of energies from two simulations, one when the cluster is in the $v = 0$ vibrational state and one when the cluster has $v = 1$ total quantum of vibration. The shift is calculated from the difference in those two energies. The second method requires only a $v = 0$ simulation to be performed. The ground state energy is calculated as usual and the excited state energy is calculated by using the distribution of the $v = 0$ simulation and the ratio of the density matrices between the $v = 1$ state and the $v = 0$ state. As before, the shift is calculated from the difference in those two energies. These first two are both exact methods of calculating the vibrational shift. The final method is perturbation theory, which has been used extensively in the past, where the shift is calculated by averaging the pairwise difference potential over the pair distribution function. However, this is an approximate method. This is the first extensive study comparing these three methods and it is found that, despite the approximations, perturbation theory method has the strongest balance between accuracy and precision when weighing against computational cost.

These methods of calculating the Raman vibrational frequency shifts are directly compared over a range of system sizes. The quantum calculations are molecular dynamics simulations using the PIGS method[105] and the systems of choice are small clusters of

pure parahydrogen and orthodeuterium clusters up to $N = 9$ which have been accurately calculated in the past and compared to experimental results[106, 107]. We also introduce an additional method which calculates the energy of the $v = 1$ state based off of the $v = 0$ simulation, which eliminates the need to run two separate simulations and halves the computational cost. We carefully determine the statistical errors of all of our reported quantities using the "binning analysis." We also use a $v = 1$ PES that is improved upon from the previous chapter. The main goals are to determine the extent of the perturbation theory approximation by comparing the vibrational shift to those calculated from exact methods over a range of cluster sizes and to determine the relative variance of each method and how they scale with cluster size.

The remainder of this chapter is organized as follows: theoretical and methodological details are introduced in Sec. 5.2, results are presented and discussed in Sec. 5.3, and we close with concluding remarks in Sec. 5.4.

5.2 Theory and methodology

The focus of this section is to describe the methodological techniques used to calculate the Raman vibrational shifts. Described in detail are the three different methods to calculate the Raman shifts and the associated approximations. We also highlight our simulation

method, the LePIGS technique and the optimized parameters used. The treatment and calculation of the statistical errors associated with our energies and shift results are also discussed.

5.2.1 Methods to calculate the Raman vibrational frequency shift for a pure cluster

The definition of the Raman vibrational frequency shift, $\Delta\nu_{0,N}$ is the difference between the energy difference between the $v = 0$ and $v = 1$ vibrational energy levels when the particle is in a vacuum (ΔE_{free}) and when the particle is a part of a cluster of N molecules ($\Delta E_{cluster,N}$),

$$\Delta\nu_{0,N} = \Delta E_{cluster,N} - \Delta E_{free} = E_{v=1,N} - E_{v=0,N} \tag{5.1}$$

where $E_{v,N}$ is the energy of the system relative to the dissocation limit for a cluster of N particles where the system has a total quantum number of vibration $v = 0$ or $v = 1$.

There are three main methods that can be used to calculate this Raman vibrational shift using path integral ground state. The first is by performing two separate simulations, using the PESs which correspond to the system being in the $v = 0$ and $v = 1$ state to calculate directly $\Delta\nu_{0,N}$ from both $E_{v=0,N}$ and $E_{v=1,N}$.

The second method is to perform one simulation using the PES corresponding to the

system in the $v = 0$ state. This allows for the direct calculation of $E_{v=0,N}$ as before. However, the energy of the system being in the $v = 1$ vibrational state, $E_{v=1,N}$, can be calculated by taking the ratio of the density matrices to effectively sample the $v = 1$ distribution and then calculating the corresponding energy as

$$\langle E_{v=1,N} \rangle = \left\langle \frac{\rho_{v=1}}{\rho_{v=0}} E_{v=1,N} \right\rangle_{\rho_{v=0}} \Bigg/ \left\langle \frac{\rho_{v=1}}{\rho_{v=0}} \right\rangle_{\rho_{v=0}} . \tag{5.2}$$

Since all molecules in our system are identical and thus have the same mass, the kinetic energy operators cancel out leaving the ratio of the densities trivially taking the form of:

$$\langle E_{v=1,N} \rangle = \left\langle \frac{\prod_{p=1}^{P} \exp\left[-\tau V_{v=1}(R_p)\right]}{\prod_{p=1}^{P} \exp\left[-\tau V_{v=0}(R_p)\right]} E_{v=1,N} \right\rangle_{\rho_{v=0}} \Bigg/ \left\langle \frac{\prod_{p=1}^{P} \exp\left[-\tau V_{v=1}(R_p)\right]}{\prod_{p=1}^{P} \exp\left[-\tau V_{v=0}(R_p)\right]} \right\rangle_{\rho_{v=0}} , \tag{5.3}$$

which can simplify to

$$\langle E_{v=1,N} \rangle = \left\langle \prod_{p=1}^{P} \exp\left[-\tau (V_{v=1} - V_{v=0})(R_p)\right] E_{v=1,N} \right\rangle_{\rho_{v=0}} \Bigg/ \left\langle \prod_{p=1}^{P} \exp\left[-\tau (V_{v=1} - V_{v=0})(R_p)\right] \right\rangle_{\rho_{v=0}} . \tag{5.4}$$

Both of these methods are considered "exact" methods as they do not introduce any approximations in the calculation of the shift $\Delta\nu_{0,N}$. However, the third method and an approach that has been commonly used in the past involves calculating the shift using first order perturbation theory. Exactly, the shift can be calculated as

$$\Delta\nu_{0,N} = E_{v=1,N} - E_{v=0,N} = \hat{H}_{v=1}\psi_{v=1} - \hat{H}_{v=0}\psi_{v=0} \ . \tag{5.5}$$

However, if one assumes that the two Hamiltonians corresponding to the $v = 0$ state and the $v = 1$ state are nearly identical, the excited vibrational Hamiltonian can be approximated as:

$$\hat{H}_{v=1} \approx \hat{H}_{v=0} + \Delta\hat{V}_N \tag{5.6}$$

where the difference of the PESs for a cluster of N molecules is defined as

$$\Delta\hat{V}_N = \hat{V}_{v=1,N} - \hat{V}_{v=0,N} \ . \tag{5.7}$$

First order perturbation theory can then be used to define the shift as:

$$\Delta\nu_{0,N} = \left\langle \psi_{v=0} | \Delta\hat{V}_N | \psi_{v=0} \right\rangle \ . \tag{5.8}$$

In our case, our PESs are pairwise additive based on the distance between the pairs of

point-like particles, therefore this quantity can be calculated by the following expression:

$$\Delta\nu_{0,N} = (N-1)\int_0^\infty dR\ \Delta V_2(R)\ \rho_{v=0,N}(R)\ , \tag{5.9}$$

where $\Delta V_2(R)$ represents the pairwise difference potential. Through the use of our path integral ground state simulations, this can be trivially calculated as a sum over independent simulation steps and over all pairs of identical particles.

$$\Delta\nu_{0,N} = (N-1)\frac{1}{nN_{pairs}}\sum_{i=1}^{n}\sum_{j=1}^{N_{pairs}}\Delta V_2(R_{i,j}) \tag{5.10}$$

5.2.2 Ground state simulation parameters and treatment of statistical errors

The pairwise additive Hinde PESs (H1-Ps)[68, 106, 107] are once again used as before for both the $v = 0$ and $v = 1$ surfaces. For simulations of N hydrogen molecules that require a total quantum of vibration of $v = 0$, the pairwise additive $v = 0$ potential can be used directly. However, for simulations of N hydrogen molecules require $v = 1$, the pairwise PES that is used takes the following form,

$$V_{v=1,N}(R) = V_{v=0,2}(R) + \frac{2}{N}\Delta V_2(R)\ , \tag{5.11}$$

where ΔV_2 represents the difference between the pairwise additive potentials with $v = 1$ and $v = 0$ quantum of vibrations denoted $V_{v=1,2} - V_{v=0,2}$. Using this potential allows each hydrogen molecule to remain indistinguishable as the quantum of vibration is shared amongst all of the molecules.

The simulation parameters common to all cluster sizes are reported in the following Table 5.1. We do not report the parameter $\tau = \beta/P$, which typically introduces a systematic error due to the Trotter factorization of the PIGS method, since for all cases we extrapolate to the $\tau = 0$ limit which removes such error. The Jastrow-type trial wavefunction used in the previous work also remains the same[106, 107].

Table 5.1: Parameters for $(pH_2)_N$ and $(oD_2)_N$ simulations: Timestep dt, centroid friction γ_0, relaxation parameter β, and length of simulation t

System	dt (fs)	γ_0 (ps^{-1})	β (K^{-1})	t (ns)
pH_2	5.00	0.005	1.125	400
oD_2	5.00	0.05	1.000	300

Since we know the form of the Trotter error for energies, we can perform simulations at multiple τ values and extrapolate to the $\tau = 0$ limit. Since the form only has even powers, we choose to include both the quadratic and quartic term for accuracy

$$E(\tau) = E_0 + a\tau^2 + b\tau^4 \ , \tag{5.12}$$

where $E(\tau)$ is the energy from our simulation at a specific τ value and E_0 is the exact ground state energy. When discussing errors, we use standard errors from our simulation energies, $E(\tau)$, and the asymptotic standard errors for our extrapolated or exact ground state energies E_0.

For each simulation, we can not save the entire trajectory as this would take up too much memory. For parahydrogen clusters we save the configuration at every 400 fs (for orthodeuterium, every 300 fs), but the energies of the systems at these configurations are still correlated and our errors would be underestimated if we took normal standard error statistics. To determine the integrated autocorrelation time for a specific data set, we use a binning analysis approach[108], so the true standard error can be accurately determined. Therefore, for each system and cluster size, we use the binning method to calculate the integrated autocorrelation time for each average property from Eqs. 5.1 and 5.2 for both the smallest and largest τ values. For each property, we then choose the largest autocorrelation time of the two and use that value over the entire τ range. For the perturbation theory calculations, from which the shifts are calculated from the ground state distribution of pair distances, the largest autocorrelation time from the ground state energy is used.

5.3 Results

Raman vibrational shifts can be calculated in three different methods. The first method, denoted the "simulation method," requires calculating the energy of the hydrogen cluster from two separate simulations when the total quantum of vibration is $v = 0$ and $v = 1$, denoted $E(v = 0)$ and $E_{sim}(v = 1)$. The second method, denoted the "ratio method," requires just the simulation when the cluster is in the $v = 0$ state, however it requires the average of two other properties from Eq. 5.2 and the ratio between them to calculate the energy for the $v = 1$ state, denoted $E_{ratio}(v = 1)$. No approximations are made for either of these two methods, so the energies corresponding to the $v = 1$ state should be identical within error bars. The results for the ground state energy of the pure $(pH_2)_N$ and $(oD_2)_N$ clusters are shown in Tables 5.2 and 5.3 below. All energies reported are extrapolated to $\tau = 0$.

Table 5.2: Ground state energies of pH_2 in Kelvin

N	E($v = 0$)	E$_{sim}$($v = 1$)	E$_{ratio}$($v = 1$)
4	-29.58(2)	-31.40(2)	-31.35(4)
5	-48.86(3)	-51.12(3)	-51.08(5)
6	-71.20(4)	-73.84(4)	-73.84(8)
7	-95.75(4)	-98.78(4)	-98.71(9)
8	-122.37(4)	-125.50(5)	-125.60(10)
9	-150.67(5)	-154.13(5)	-154.15(12)

As expected, the $E(v = 1)$ energies are identical within error bars for each cluster size,

Table 5.3: Ground state energies of oD_2

N	$E_0(v=0)$	$E_{0,sim}(v=1)$	$E_{0,ratio}(v=1)$
4	-61.44(3)	-63.29(3)	-63.21(5)
5	-96.61(4)	-98.74(4)	-98.77(6)
6	-137.69(5)	-140.30(5)	-140.22(9)
7	-181.10(5)	-183.93(5)	-183.95(13)
8	-225.92(6)	-228.92(6)	-228.96(11)
9	-274.78(6)	-278.01(6)	-278.04(16)

however the error bars of E_{ratio} are approximately twice as large as E_{sim}. As well, the magnitude of the error bars of the energies calculated by ratio increase at a much faster rate with respect to cluster size than the energies calculated by a separate simulation. This makes sense because it relies on taking the ratio of two average quantities, each with their own error bars. It is also clear that the errors of the $E(v=0)$ and $E(v=1)$ using the simulation methods are approximately the same for each cluster size, which is expected since the PESs are very similar.

The Raman vibrational shifts can now be compared using the three methods, introducing the perturbation theory results coming from the ground state pair distribution function from the $v=0$ PES and Eq. 5.8. The results are shown in Tables 5.4 and 5.5 below.

When looking at the reported error bars, the variance of the results using perturbation theory are on the order of 10^{-3} cm^{-1} and do not increase with system size, however the values themselves are always lower in magnitude than the experimental results. There is an

Table 5.4: Calculated Raman vibrational shifts ($\Delta\nu_{0,N}$) of pH_2 in cm^{-1} using three different methods: simulation (Sim), ratio (Ratio), and perturbation theory (PT) with the experimental results shown as well

N	Sim	Ratio	PT	Experiment[31]
4	-1.263(35)	-1.228(49)	-1.208(2)	-1.251(20)
5	-1.570(44)	-1.545(64)	-1.535(2)	-1.594(20)
6	-1.830(53)	-1.830(84)	-1.815(2)	-1.910(20)
7	-2.110(60)	-2.058(98)	-2.039(2)	-2.136(20)
8	-2.172(66)	-2.242(113)	-2.232(2)	-2.350(20)
9	-2.403(72)	-2.416(128)	-2.403(2)	

Table 5.5: Calculated Raman vibrational shifts ($\Delta\nu_{0,N}$) of oD_2 in cm^{-1} using three different methods: simulation (Sim), ratio (Ratio), and perturbation theory (PT) with the experimental results shown as well

N	Sim	Ratio	PT	Experiment[107]
4	-1.289(45)	-1.233(59)	-1.225(1)	-1.287(30)
5	-1.479(55)	-1.501(73)	-1.495(1)	-1.602(30)
6	-1.812(66)	-1.760(101)	-1.757(1)	-1.844(30)
7	-1.964(76)	-1.979(144)	-1.965(1)	
8	-2.085(83)	-2.111(128)	-2.103(1)	
9	-2.245(91)	-2.263(175)	-2.265(1)	-2.457(30)

error using this method due to the approximations made in the perturbation theory. The simulation and ratio methods are exact, as no approximations have been introduced, and the results are, as expected, within error of each other since they are based on the ground state energies that are common and excited state energies that themselves agree with each other within error. It is determined that, compared to experiment, the shifts calculated from the "simulation" for $(pH_2)_N$ agree within error to the experimental results up until

$N = 8$ where it seems to deviate significantly. For the ratio method, the shift values agree within error of the experimental results consistently (including $N = 8$) but this is because the errors are significantly larger for the shifts reported from the ratio method than the simulation method. This is the same trend as seen in the $(\mathrm{oD}_2)_N$ case.

We propose that using the Hinde potential to calculate the Raman vibrational shifts is very good for small clusters, but will underestimate (in magnitude) the value of the shift as the cluster size increases. Incidentally, with the increased errors of using the "ratio method," the experimental results can be reproduced, but it is believed this is just coincidence as the values using both the simulation and perturbation theory methods, with smaller error bars, differ significantly from the experimental results. We can also confirm that this underestimation is not due to the perturbation theory and, in fact, the perturbation theory results lie within error bars of the simulation results for cluster sizes $N > 4$.

5.4 Concluding remarks

In this chapter, we calculate the Raman vibrational shifts of small parahydrogen and orthodeuterium clusters using three different methods. The first method, so-called the "simulation method" is directly calculating the energy of the cluster both with $v = 0$ quantum of vibration and $v = 1$ from two separate simulations. The shift is then taken as the difference between those energy values. The second method, so-called the "ratio method" is directly calculating the energy of the cluster in the $v = 0$ state, but using the distribution of the ground state simulation and the ratio of density matrices to calculate the energy in the $v = 1$ state. The shift is again taken as the difference between those values, but only one ground state simulation is required. The final method is to calculate the shift through first order perturbation theory by multiplying the difference potential (the difference between the $v = 1$ and $v = 0$ pair potentials) over the distribution of the ground state pair distances.

The trends for $(pH_2)_N$ and $(oD_2)_N$ were found to be the same in the analysis of both the energies and the vibrational shifts. The values of the energies in the $v = 1$ vibrational states, calculated directly from simulation or using the ratio of the density matrices, gave the same result as expected since both are exact methods. However, the variance in the energies was much higher (approximately twice as large) when using the ratio method than

the separate simulation method. This is due to the fact that the average of two properties (both with statistical errors) had to be taken and divided out. Since running a simulation twice as long only decreases the standard error by a factor of $\sqrt{2}$, between the two cases it is still not efficient to use this ratio method when comparing computational efficiency.

With respect to the vibrational frequency shifts, the simulation method and the ratio again give the same result within error bars, as expected since the calculation requires a common ground state energy and the excited state energies calculated from the two methods are the same within error bars. However, the shifts calculated from perturbation theory and the simulation method are the same within error bars for all cluster sizes except for $N = 4$. This could be due to the fact that the vibrational excitation is delocalized across fewer molecules, so it's a much larger perturbation than when it delocalized in larger clusters. With respect to the variance, the simulation method has errors larger than perturbation theory by an order of magnitude and increases with cluster size. However, the variance from perturbation theory remains consistently small (10^{-3} cm^{-1}, and since the values agree with those from the simulation method within error bars, the perturbation theory method is recommended despite the approximation.

In comparison to experimental values, the simulation results agree within error bars for $N = 4$ until $N = 7$. At $N = 8$, we notice that symmetry is reduced to the C_s group so this could be a contributing factor, but the shift from simulation also deviates from experiment

for $(D_2)_9$ even though the symmetry is enhanced for that cluster size. It could then be that this is a limitation of the pairwise additive Hinde potential, that it is very accurate up until $N = 7$ and past this, the potential may need to include more information or the pairwise potential is not sufficient and we require one that includes many-body effects.

Chapter 6

Ground state chemical potential of parahydrogen clusters of size N=21-40

6.1 Introduction

The shape of the ground state chemical potential of parahydrogen clusters between N=21-40 has been the subject of much debate. At low temperatures, down to 1 K, there is a general consensus that in this range of cluster sizes, the chemical potential of parahydrogen

is jagged indicating many stable structures and magic number clusters (clusters which have enhanced stability). However, below 1 K and in the zero-temperature limit, there has been a discrepancy between whether the magic numbers persist or whether so-called 'quantum-melting' occurs, where the chemical potential smooths out and monotonically increases. This has led to numerous studies summarized below and a review[109].

Path integral Monte Carlo calculations at finite temperature have been performed by many groups. Mezzacapo and Boninsegni calculated the chemical potential at 1 K[110] and found peaks at $N = 24, 26, 29, 32, 34, 37$, and 39. However, when performing calculations at a lower tempearture T=0.25 K[111], the distinct peaks disappear and the chemical potential becomes a smooth monotonically increasing curve suggesting some sort of quantum melting transition occurs between 0.25 K and 1 K. Khairallah, Ceperley, and Toennies also calculated the chemical potential at T=0.5 K[112], but observed distinct magic number peaks corresponding to $N = 23, 26, 29, 32, 34$, and 37 ($N = 38 - 40$ were not calculated for that temperature) which agree very well with the 1 K literature results. These magic number cluster sizes are analogous to those for alkaline earth metals and rare gases [113] indicating possible icosahedral or pseudo-closed packed structures.

Since T=0.25 K is in the ground state regime for parahydrogen clusters, the quantum ground state methods of diffusion Monte Carlo (DMC) and path integral ground state (PIGS) can be employed. Guardiola and Navarro have studied the chemical potential us-

ing DMC and found no magic number peaks[114, 115, 116] except for a study in which they investigate the chemical potential with respect to excitation energies[115] where they also see magic numbers of $N = 36$ and possibly $N = 31$. Sola and Boronat used DMC to calculate the energies of hydrogen clusters assuming they were either in solid-like configurations or in liquid-like configurations by varying the trial wavefunction. The lower energy of the two is the one that would be used for the chemical potential curve. For cluster sizes smaller than $N = 30$, the same smooth curve is formed, but for larger clusters, a jagged curve formed in stark contrast to the previous DMC results. Magic number cluster sizes were predicted for $N = 23, 25, 28, 30, 34, 37$, and 40.

However, Cuervo *et al.* used the Monte Carlo variant of PIGS to calculate the ground state chemical potential[117, 51] and found very distinct peaks in the curve corresponding to magic number cluster sizes of $N = 26, 29, 34$, and 39 with the Buck potential[70] which is in contrast to the smooth curve predicted by DMC. Just recently, Miura used variational path integral molecular dynamics (VPIMD), a molecular dynamics variant of PIGS to calculate[63] the ground state chemical potential and also see clear magic number cluster sizes for $N = 26, 29, 34$, and 39. Mezzacapo and Boninsegni predict[44] the difference between the DMC and PIGS results is due to population size bias that can arise in DMC if not properly accounted for. This bias is not found in PIGS.

The focus of this chapter is to use the LePIGS method to calculate the ground state

chemical potential and contribute to the literature review. LePIGS has been used previously[65] to study ground state energies of smaller hydrogen clusters up to $N = 19$ and $N = 33$ to explore their solid-like and liquid-like properties. Both the Buck (B-P) [70] and the Silvera-Goldmann (SG-P)[71] interaction potentials are used in the simulations to see if the magic numbers are potential-dependent. We will describe our system setup in Section 6.2, followed by results in Section 6.3, and we will close with conclusions and future directions in Section 6.4.

6.2 System setup

Both the B-P [70] and SG-P[71] interaction potentials for hydrogen are used in our LePIGS simulations. For the SG-P, we use the solid-effective pair potential, which includes the non-zero C_9 term and the potential minimum, r_m=3.44 Å. Traditionally, many groups who use the SG-P use r_m=3.41 Å, which results in the interaction potential having a slightly softer repulsive wall. This means that the energy values can not be directly compared, but qualitatively should have the same result. As before, the Jastrow trial wavefunction is used for the hydrogen molecules in our LePIGS simulations. Simulations for both potentials and all cluster sizes ($N = 20 - 40$) use the same optimized parameters: a simulation timestep, dt = 4.0 fs, a centroid friction, γ^0=0.5 ps^{-1}, a decorrelation time of 0.200 ps, an imaginary

time length, $\beta = 1.25$ K^{-1}, and each simulation is run for 10^6 independent steps.

6.3 Results

Here we report the ground state chemical potential for parahydrogen calculated from LeP-IGS simulations using both the B-P and SG-P interaction potentials. The chemical potential is defined as

$$\mu(N) = E_0(N-1) - E_0(N) \ , \tag{6.1}$$

where E_0 is the ground state energy extrapolated to the $\tau = 0$ limit. This notation means that magic number clusters are identified by the peaks in the chemical potential curve. Our LePIGS results are compared to previous literature results in Figure 6.1. For both interaction potentials, it is clear that the chemical potential curves calculated from DMC[116] are very smooth compared to the those generated from the other ground state methods. For the B-P, LePIGS have more magic number clusters and are more pronounced than PIGS-MC[51]. As well for the B-P, LePIGS predicts magic number clusters of $N = 23, 26, 29, 32, 34, 37$, and 39 which corresponds directly to those from the low temperature results of Khairallah[112]. For the SG-P, the LePIGS results agree very well with the VPIMD[63] results with the peaks and valleys aligning exactly over the entire chemical potential curve. Magic numbers are predicted at the exact same cluster sizes as

those calculated with the B-P, however the clear magic numbers are $N = 26, 29, 34$, and 39 with the other magic number cluster sizes much less peaked compared to the B-P results.

6.4 Concluding Remarks

In this chapter, the ground state chemical potential is calculated from LePIGS simulations of cluster sizes $N = 21 - 30$ using both the Buck (B-P)[70] and Silvera-Goldmann (SG-P)[71] interaction potentials and the results are compared to those from the literature. Our LePIGS results agree qualitatively with PIGS-MC results[51] and with recent VPIMD results[63] in that the chemical potential is not a smooth curve, but has magic number cluster sizes. This is in disagreement with DMC results[114, 115, 116] which had predicted a smooth curve with no distinguishable peaks. It has been speculated[44] that the discrepancy is due to population size bias in DMC that is not present in the PIGS formulation. Using LePIGS, magic number cluster sizes are predicted for $N = 23, 26, 29, 32, 34, 37$, and 39 which agree exactly with VPIMD results, although they are more pronounced when using the B-P than the SG-P.

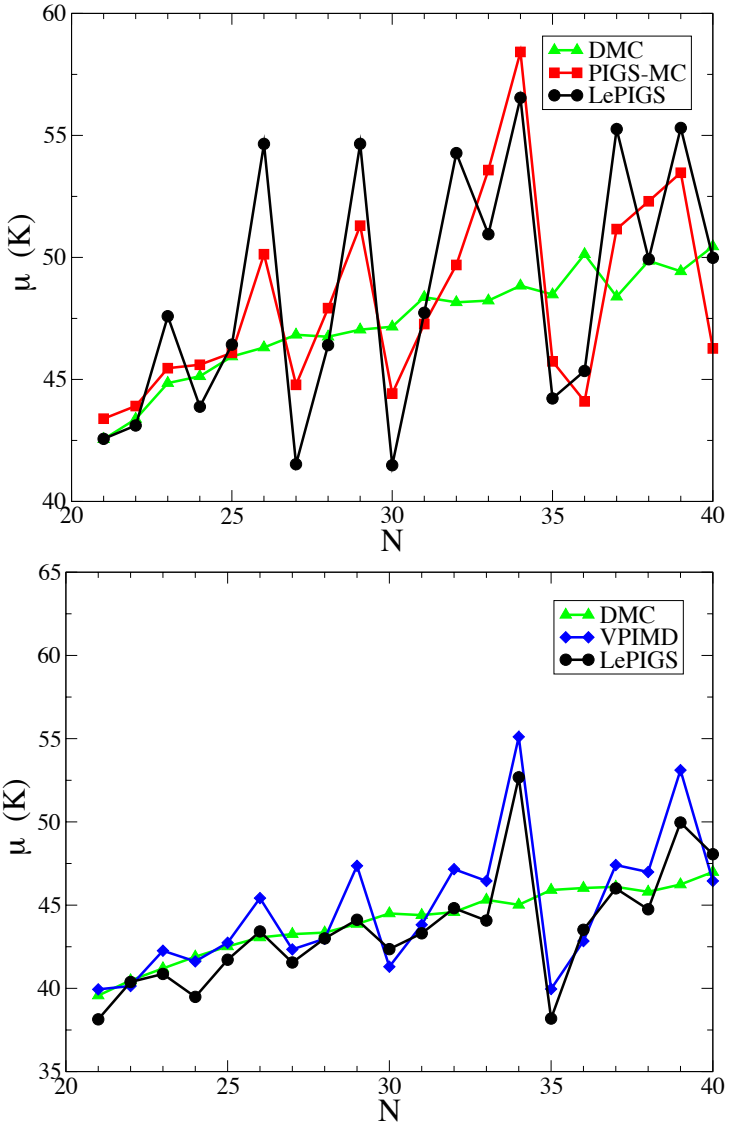


Figure 6.1: Ground state chemical potentials of $(pH_2)_N$ for the B-P[70] interaction potential (upper) and the SG-P interaction potential (lower). LePIGS results are in black squares, the DMC results from Ref. [116] are in green triangles, the PIGS-MC results from Ref. [51] are in red, and the VPIMD results from Ref. [63] are in blue

Chapter 7

An extension to flexible molecular systems in their ground state: Application to the water dimer

7.1 Introduction

In this chapter,[1] we illustrate the implementation of the LePIGS method for flexible molecular systems. We chose to use the water dimer system as an illustrative example. We first

[1]Sections of this chapter have been reprinted with permission from M. Schmidt and P.-N. Roy, "Path integral Molecular dynamic simulation of flexible molecular systems in their ground state: application to the water dimer", J. Chem. Phys. 148, 124116 (2018). Copyright 2018 American Institute Physics.

calculate ground state energies of the flexible water monomer and dimer system which are benchmarked against the DMC results[118, 119] and compared to experimentally measured dissociation energies[120]. All ground state properties are calculated for two quantum-parameterized models composed primarily of pair interactions between the monomers, q-TIP4P/F and q-SPC/Fw[121], and one *ab initio* model which includes many-body effects, MB-pol. The LePIGS method is then used further to calculate ground state structural properties and imaginary time correlation functions to obtain vibrational excited states, both in a straight-forward manner. The MB-pol potential was recently used to calculate temperature-dependent vibrational spectra using molecular dynamics[122], but this is the first known application of calculating vibrational transition energies from Path integral methods of the MB-pol water model in the ground state, as well as the first calculation of vibrational transition energies from a LePIGS simulation using imaginary time correlation functions.

Our work demonstrates that LePIGS can be extended to systems of flexible molecules and can be seamlessly merged with low temperature PIMD results. The remainder of this chapter is organized as follows: theoretical and methodological details are introduced in Sec. 7.2, results are presented and discussed in Sec. 7.3, and we close with concluding remarks in Sec. 7.4.

7.2 Theory and methodology

The focus in this section is to provide details of how the parameters required for a LePIGS simulation are optimized for the case of flexible molecular systems and the theory behind imaginary time correlation functions and how they can be used to calculate vibrational spectra from such a simulation.

7.2.1 System parameters and optimization

We use three water models denoted q-TIP4P/F[123], q-SPC/Fw[121], and MB-pol[124, 125, 126]. We note that for the case of the water monomer, the MB-pol potential reduces[124] to the spectroscopically accurate potential by Partridge and Schwenke[127].

For each model, the unity trial wavefunction, $\psi_T \propto 1$, is used due to it's simplistic form and that it is guaranteed to have some overlap with the true nodeless ground state wavefunction. In a previous paper[65], which determined the effect of trial wavefunctions on the ground state properties of hydrogen clusters, this unity trial wavefunction is found to be a suitable choice since it is uniform over configuration space. It does not prefer a specific geometry over another and thus introduces no additional ergodicity problems. However, since it includes no information about the system, more path integral beads are required to converge the energy than other trial wavefunctions that included hard-core

repulsion terms or a normal mode ansatz.

There are five parameters needed to run an efficient simulation: the simulation timestep dt, a centroid friction γ^0, a decorrelation time T_d, a relaxation parameter β, and an imaginary timestep τ and they are systematically converged one at a time following the same method used in the past by our group[65]. However, once the first set of "optimized" parameters were found, the decorrelation times were recalculated to ensure the data points used in the averaged properties were statistically independent. These optimizations are performed for each of our water models, except that the same friction value is used for all models based on analyses performed for the q-TIP4P/F model.

Here, it should be reinforced that errors associated with the simulation timestep dt, the relaxation parameter β, and the imaginary timestep τ are all systematic errors and are eliminated through convergence studies. We note that regardless of which trial wavefunction is used, assuming finite overlap with the true ground state wavefunction, it's systematic error is eliminated by choosing a large enough value of the relaxation parameter, β.

Our optimized parameters are shown in Tables 7.1 and 7.2 for H_2O and D_2O water models respectively.

The average potential energy was also calculated along the imaginary time path for an H_2O monomer using the MB-pol potential, shown in Fig. 7.1. Due to the fact that the

Table 7.1: Simulation parameters for H_2O Water Models: Timestep dt, friction γ^0, decorrelation time T_d, and relaxation parameter β

System	dt (fs)	γ^0 (ps^{-1})	T_d (ps)	β (K^{-1})
MB-pol (M)	0.10	1/0.17	0.030	0.002
MB-pol (D)	0.12	1/0.30	0.055	0.03
q-TIP4P/F (M)	0.12	1/0.17	0.084	0.005
q-TIP4P/F (D)	0.15	1/0.30	0.120	0.010
q-SPC/Fw (M)	0.12	1/0.17	0.080	0.005
q-SPC/Fw (D)	0.12	1/0.30	0.030	0.010

Table 7.2: Simulation Parameters for D_2O Water Models

System	dt (fs)	γ^0 (ps^{-1})	T_d (ps)	β (K^{-1})
MB-pol (M)	0.15	1/0.17	0.043	0.005
MB-pol (D)	0.12	1/0.30	0.024	0.02
q-TIP4P/F (M)	0.10	1/0.17	0.020	0.005
q-TIP4P/F (D)	0.10	1/0.30	0.120	0.05

potential energy operator does not commute with the Hamiltonian, the potential energy of the system must be calculated from the middle bead. Based on our total energy convergence, $\beta = 0.002$ K^{-1} was chosen and by observing the potential energy convergence along the path, it is clear that the potential energy at the halfway point, $\beta = 0.001$ K^{-1}, where the potential energy would be calculated from during the course of the simulation, is also converged. This is another property that justifies our choice in β as it is converged, but not over-converged as this would cause computational inefficiency.

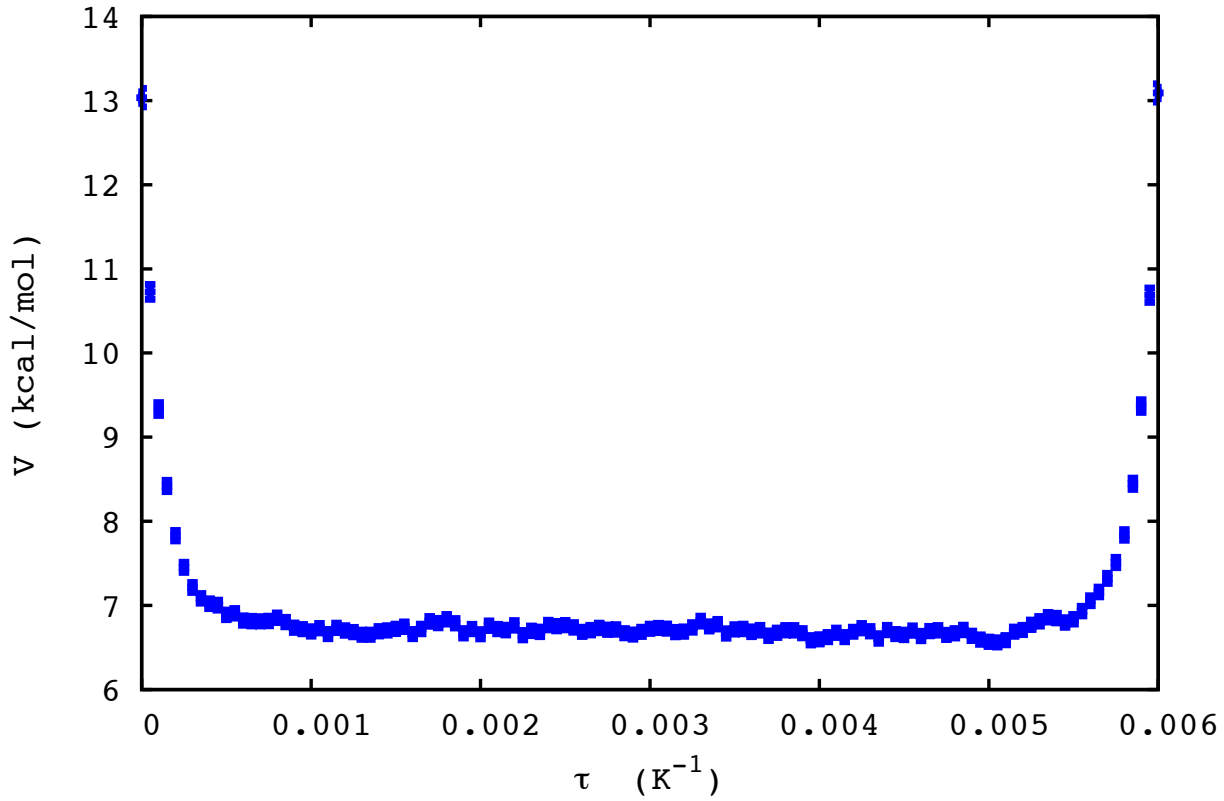


Figure 7.1: Average MB-pol potential energy along the imaginary time step for the H_2O monomer. In our simulations of this system, we choose $\beta = 0.002$ K^{-1}, which is in the converged regime.

As well, in order to reduce statistical error, we run 10^5 independent simulation steps. The total number of simulation steps required can be calculated as

$$N_{sim} = 10^5 \times T_d/dt \ . \tag{7.1}$$

The PILE thermostat does one potential (and gradient) call per simulation step, so the

number of potential calls is simply

$$N_{pot} = N_{sim} \times P \ , \tag{7.2}$$

since the potential is calculated for each set of beads. These statistics are the same for finite temperature PIMD. For example, it would require 6×10^9 potential calls to simulate the H_2O monomer with $P = 200$.

7.2.2 Imaginary time correlation functions

The PIGS method can be used to calculate vibrational energy transitions through the use of imaginary time correlation functions. Assuming a path of 2β, where the middle of the path is at β, the imaginary time correlation function of operator, $\hat{O}$, takes the form:

$$C(\tau) = \frac{1}{Z} \langle \psi_T | e^{-\beta \hat{H}} \hat{O} e^{-\tau \hat{H}} \hat{O} e^{\tau \hat{H}} e^{-\beta \hat{H}} | \psi_T \rangle \ , \tag{7.3}$$

where the normalization, Z, is the ground state pseudo-partition function defined by

$$Z = \langle \psi_T | e^{-2\beta \hat{H}} | \psi_T \rangle \ . \tag{7.4}$$

For simplicity, the energies of our Hamiltonian, E_n are chosen to be relative to that of the ground state, E_0. Next, three complete sets of states are inserted and the expression can be expanded as:

$$C(\tau) = \frac{1}{Z} \sum_{n,n',n''} \langle \psi_T | n \rangle \langle n | e^{-\beta \hat{H}} \hat{O} | n' \rangle \langle n' | e^{-\tau \hat{H}} \hat{O} | n'' \rangle \\ \langle n'' | e^{\tau \hat{H}} e^{-\beta \hat{H}} | \psi_T \rangle \ , \tag{7.5}$$

and the Hamiltonian operators can be evaluated and the terms regrouped to form:

$$C(\tau) = \frac{1}{Z} \sum_{n,n',n''} \langle \psi_T | n \rangle e^{-\beta(E_n - E_0)} \langle n | \hat{O} | n' \rangle e^{-\tau(E_{n'} - E_0)} \\ \langle n' | \hat{O} | n'' \rangle e^{-(\beta - \tau)(E_{n''} - E_0)} \langle n'' | \psi_T \rangle \ . \tag{7.6}$$

In the limit of $\beta \to \infty$, the ground state, $|0\rangle$, is the only surviving state and thus the expression can be simplified as:

$$C(\tau) = \frac{1}{Z} \sum_{n'} \langle \psi_T | 0 \rangle e^{-\beta(E_0 - E_0)} \langle 0 | \hat{O} | n' \rangle e^{-\tau(E_{n'} - E_0)} \\ \langle n' | \hat{O} | 0 \rangle e^{-\beta(E_0 - E_0)} \langle 0 | \psi_T \rangle \ . \tag{7.7}$$

By collecting the terms and relabelling the summation indices, we obtain the final expression:

$$C(\tau) = \frac{1}{Z}|\langle\psi_T|0\rangle|^2 \sum_n |\langle 0|\hat{O}|n\rangle|^2 e^{-\tau(E_n - E_0)} \ , \tag{7.8}$$

where the imaginary time correlation function is a sum of exponentials. However, if an operator can be found that has very good overlap with only one of the excited states, the correlation function should reduce to a single exponential, which can easily be modelled to determine the transition energy.

7.3 Results

Using our LePIGS simulations, we calculate ground state energies and structural properties for the flexible water monomer and dimer and report our results in this section. We compare our energies in the zero-temperature limit to previous DMC calculations and experimental results, as well as those calculated using finite temperature PIMD (from Ref. [128]) to demonstrate the seamless merging of the two methods. We also show that vibrational spectra can be calculated from a LePIGS simulation as we calculate the transition energies of the vibrational modes of the water monomer.

7.3.1 Ground state energies

The ground state energies of the H_2O and D_2O monomer and dimer are calculated using the LePIGS method for three different water models. Previous DMC studies[118, 119] have calculated these energies using two of the models, q-TIP4P/F and MB-pol, however the work in this section is simply to demonstrate that the LePIGS method can calculate energies as accurately as DMC and structural properties in a straight-forward manner.

In order to obtain small enough errors for the ground state energies, simulations were run for 10^5 independent steps. To achieve this, for each system involving a water monomer, one simulation was run for 25 independent steps and each of those would serve as initial conditions for simulations of 4000 independent steps. For the dimer systems, 100 parallel simulations of 1000 independent steps were produced. Our reported average properties are the average of all of those spawned simulations and we note the errors on our data points are given as standard errors.

The final parameter converged is τ, due to the systematic Trotter error when discretizing our path. As mentioned in Section 2.1, the order of the error of the partition function is $O(\tau^3)$ and since the energy is obtained by taking a derivative of the partition function with respect to β, the order of the error drops to $O(\tau^2)$. This is the same in both finite temperature PIMD and the PIGS formulation. In the 'low τ limit,' strictly a quadratic fit

should be used to fit the data. However not all data points may fall into this 'quadratic regime,' so a quartic term was added to the fit which corresponds to the next order of the Trotter error[72, 48],

$$E(\tau) = E_0 + b\tau^2 + c\tau^4 \ , \tag{7.9}$$

where E_0 is the exact ground state energy and b and c are the quadratic and quartic constants respectively. We eliminate the Trotter error by extrapolating our $E(\tau)$ fit to $\tau = 0$ and the fit value of E_0 is reflected, along with it's error bars given as the asymptotic standard error, by the data point at $\tau = 0$ on the subsequent Figures.

First shown in Figs. 7.2 and 7.3 are the energy convergences with respect to τ of the monomer and dimer systems of H_2O and D_2O, respectively, for the q-TIP4P/F model using LePIGS.

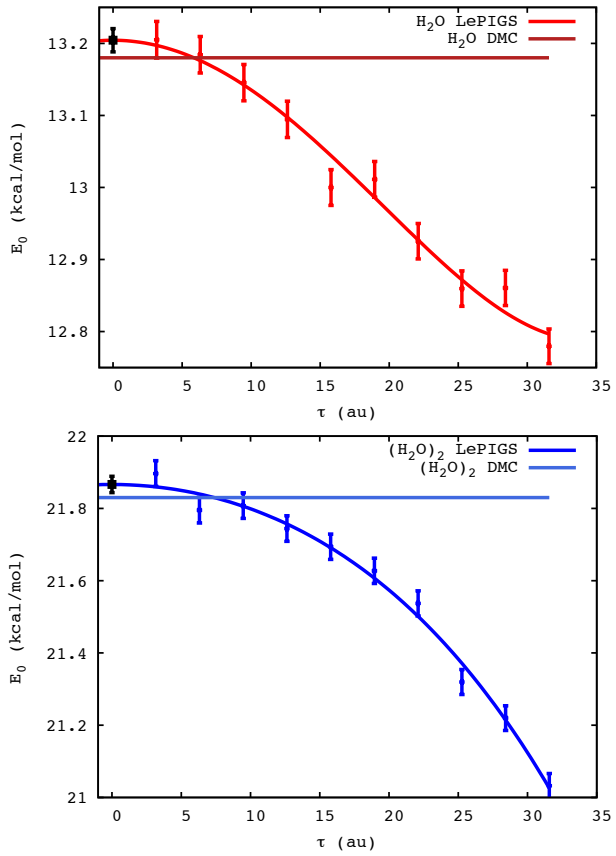


Figure 7.2: Energy convergence with τ for the H_2O monomer (upper panel) and dimer (lower panel) using the q-TIP4P/F model comparing LePIGS results to DMC. The $\tau = 0$ points are the extrapolated E_0 values from the fits.

The results are compared to those obtained from a DMC calculation using the same potential. We notice that our fit models the entire data set well and our extrapolated energies are in good agreement with the DMC results.

Also shown are the same convergence studies of our monomer and dimer energies with respect to τ for the q-SPC/Fw model in Fig. 7.4 and for the MB-pol model in Fig. 7.5,

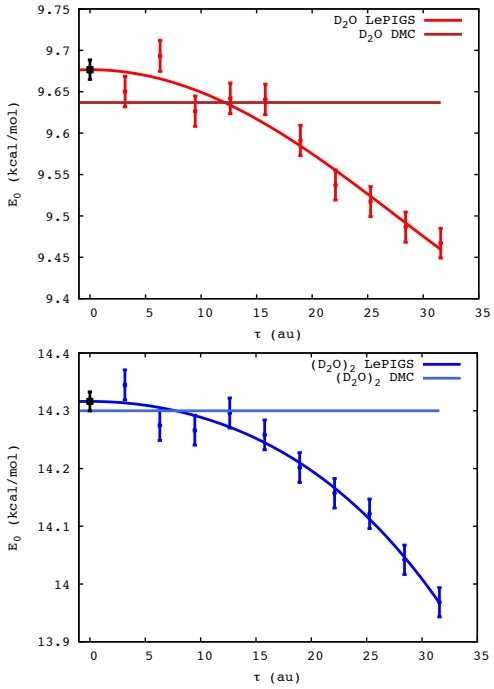


Figure 7.3: Energy convergence with τ for the D_2O monomer (upper panel) and dimer (lower panel) using the q-TIP4P/F model comparing LePIGS results to DMC. The $\tau = 0$ points are the extrapolated E_0 values from the fits.

plotted in a similar manner as the q-TIP4P/F model. As before, the entire data set fits the form well for both models.

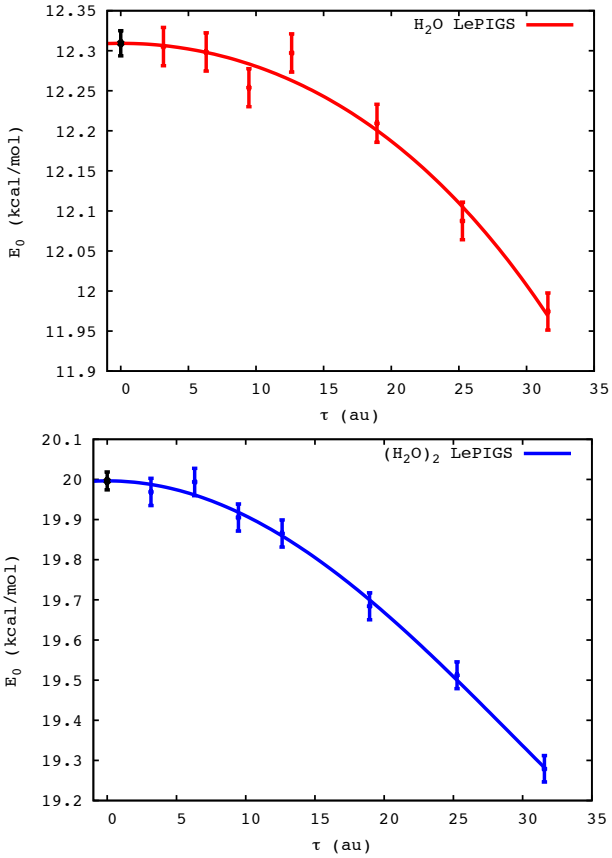


Figure 7.4: Energy convergence with τ for H_2O monomer (upper panel) and dimer (lower panel) using the q-SPC/Fw model. The $\tau = 0$ points are the extrapolated E_0 values from the fits.

We also report our energies for each model in Tables 7.3 and 7.4 for H_2O and D_2O respectively and compare to the DMC results of Refs. [118, 119]. The dissociation energy is defined as

$$D_0 = 2E_{\text{mon}} - E_{\text{dim}} \tag{7.10}$$

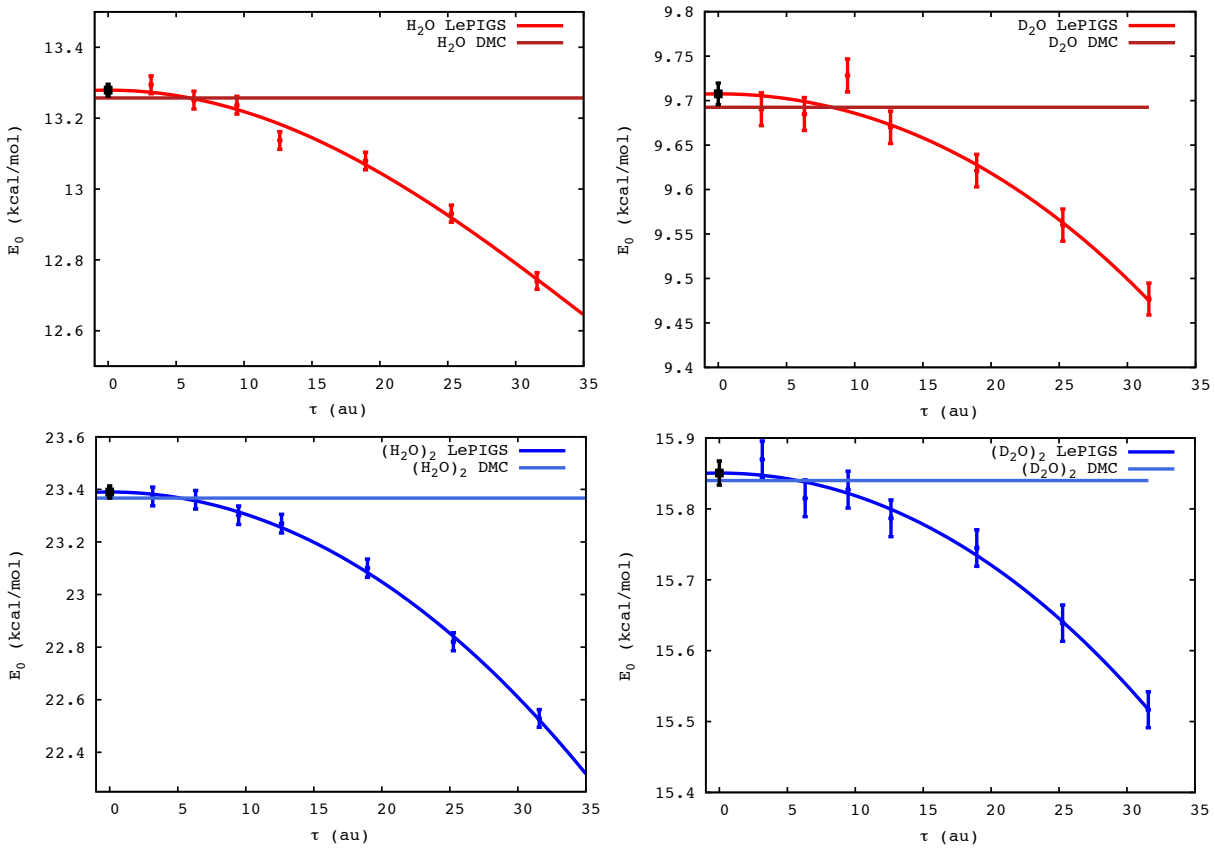


Figure 7.5: Energy convergence with τ for H_2O (left column) and D_2O (right column) monomer (upper row) and dimer (lower row) using the MB-pol model comparing LePIGS results to DMC. The $\tau = 0$ points are the extrapolated E_0 values from the fits.

and the calculated dissociation energies are compared to both the DMC results and experiment[120].

The energies for the monomer systems can also be calculated exactly using the Lanczos diagonalization in Pekeris co-ordinates, as it constitutes a trimer system[97, 129, 98]. For the H_2O monomer calculation, 150 dimensionless Jacobi Discrete Variable Representation (DVR)[97] basis functions were used over 15000 Lanczos iterations with a $V_{\max}$

of 100000 cm^{-1}. For the D_2O monomer calculation, 90 dimensionless Jacobi DVR basis functions were used over 10000 iterations with a V_{max} of 75000 cm^{-1}. The basis function parameters[97] used for all isotopologues are $r_{max} = 2.0$ Å, $a = 2$, and $b = 0$.

Table 7.3: H_2O ground state energies (kcal/mol) calculated for different water models.

	q-TIP4P/F		q-SPC/Fw	MB-pol	
	LePIGS	DMC[118]	LePIGS	LePIGS	DMC[119]
E_{mon}	13.20(2)	13.18	12.31(2)	13.28(2)	13.26
E_{dim}	21.87(2)	21.83	19.99(2)	23.39(2)	23.36
D_0	4.54(4)	4.53	4.63(4)	3.17(4)	3.15
D_0^{exp} (Ref. [120])					3.16(3)

Table 7.4: D_2O ground state energies (kcal/mol) calculated for different water models

	q-TIP4P/F		MB-pol	
	LePIGS	DMC[118]	LePIGS	DMC[119]
E_{mon}	9.68(1)	9.63	9.71(1)	9.69
E_{dim}	14.32(2)	14.30	15.85(2)	15.84
D_0	5.04(3)	4.96	3.56(3)	3.54
D_0^{exp} (Ref. [120])				3.56(3)

As expected, the LePIGS simulation results are in agreement with the benchmark DMC results from Mandelshtam *et al.* showing that the LePIGS method can calculate energies accurately and again observe that the MB-pol potential is superior to both the q-TIP4P/F and q-SPC/Fw when comparing to the experimental D_0 values. Both LePIGS and DMC exactly reproduce the exact diagonalization energies for the monomer using the MB-pol potential, 13.26 kcal/mol for the H_2O monomer and 9.69 kcal/mol for the D_2O monomer.

We also calculate the energies using the harmonic oscillator approximation using frequencies obtained by diagonalizing the mass-weighted hessian of the monomer and dimer systems with each water model. Although not shown in the Tables, it is a reasonable approximation, but always overestimates the energies. This results in smaller dissociations energies on the order of 0.1 kcal/mol.

We further compare our LePIGS calculations to finite temperature PIMD calculations for the H_2O monomer, dimer, and dissociation energy using the MB-pol potential[128]. The PIMD calculations were performed using the PILE thermostat of MMTK, completely analogous to the LePIGS calculations. The PIMD results are also extrapolated to $\tau = 0$ so as to eliminate the Trotter error. In reference to Fig. 7.6, it is clear that as the temperature decreases, the energies calculated from PIMD seamlessly converge to the ground state value calculated by LePIGS.

Table 7.5: H_2O monomer average ground state properties calculated for different water models

	q-TIP4P/F	q-SPC/Fw	MB-pol
R_{HH} (Å)	1.5408(4)	1.6553(4)	1.5382(4)
R_{OH} (Å)	0.9582(2)	1.0010(2)	0.9760(2)
θ (deg)	107.42(3)	112.03(3)	104.78(3)

Table 7.6: D_2O monomer average ground state properties calculated for different water models

	q-TIP4P/F	MB-pol
R_{HH} (Å)	1.5335(3)	1.5316(3)
R_{OH} (Å)	0.95351(13)	0.97110(13)
θ (deg)	107.33(2)	104.40(2)

7.3.2 Ground state distributions

The distribution of the bond lengths can also be reported for each of the water models. For the water monomers, the distributions are calculated at our lowest τ value (3.16 au). The distributions of the probability distribution functions for the OH bond lengths (R_{OH}), HH distances (R_{HH}), and HOH bond angle (θ_{HOH}) are shown in Fig. 7.7. Tables 7.5 and 7.6 also show the values of the average bond lengths and angle for the H_2O and D_2O monomers respectively.

For the water dimer, the probability distribution functions for the distance between hydrogen atoms (HH), distance between oxygen and hydrogen atoms (OH), and distance between oxygen atoms (OO) are shown in Fig. 7.8. We notice that the distributions of the

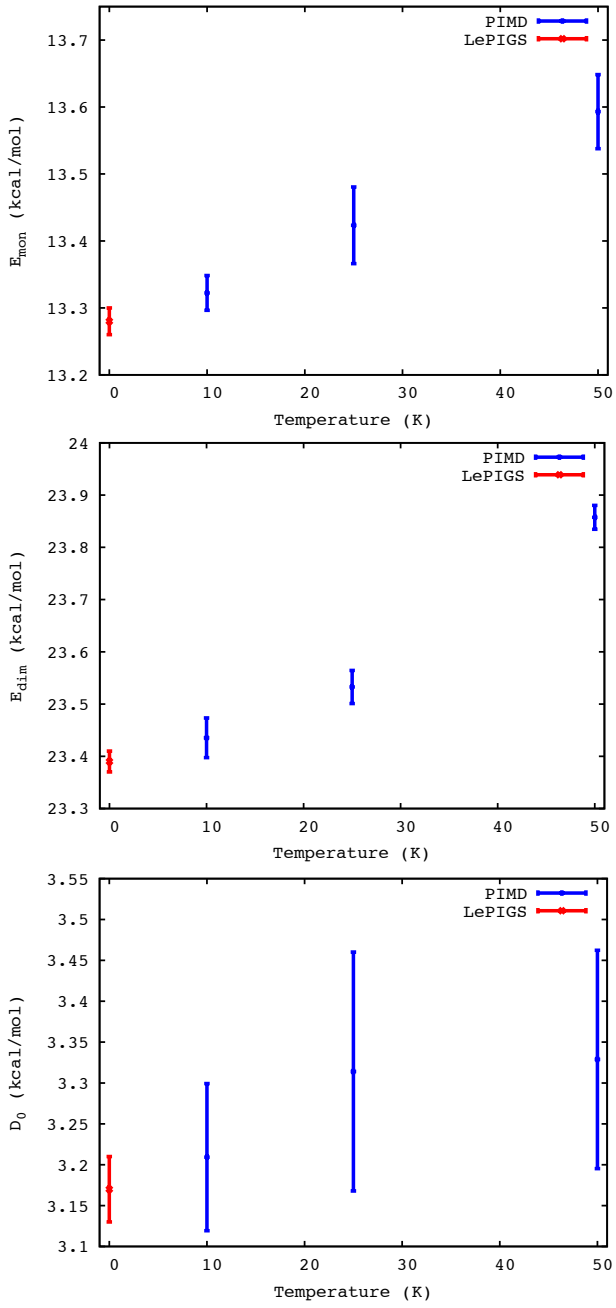


Figure 7.6: H_2O monomer (upper panel), dimer (middle panel), and dissociation energies (lower panel) calculated using the MB-pol intermolecular potential at various temperatures using PIMD and in the zero-temperature limit using LePIGS.

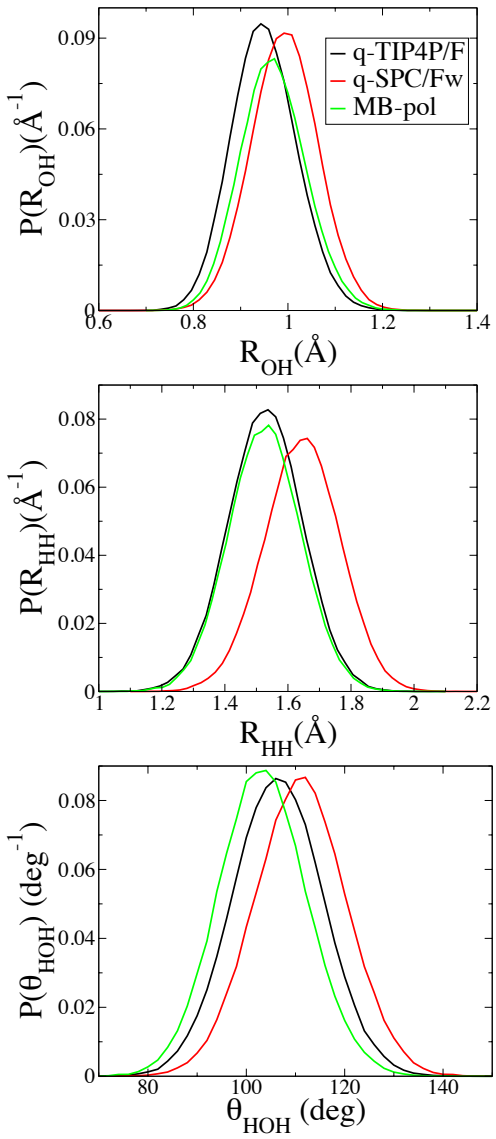


Figure 7.7: Probability distribution functions for the OH bond lengths, R_{OH} (upper panel), HH distances, R_{HH} (middle panel), and the HOH bond angle, θ_{HOH} (lower panel), for the water monomer described by each of the models.

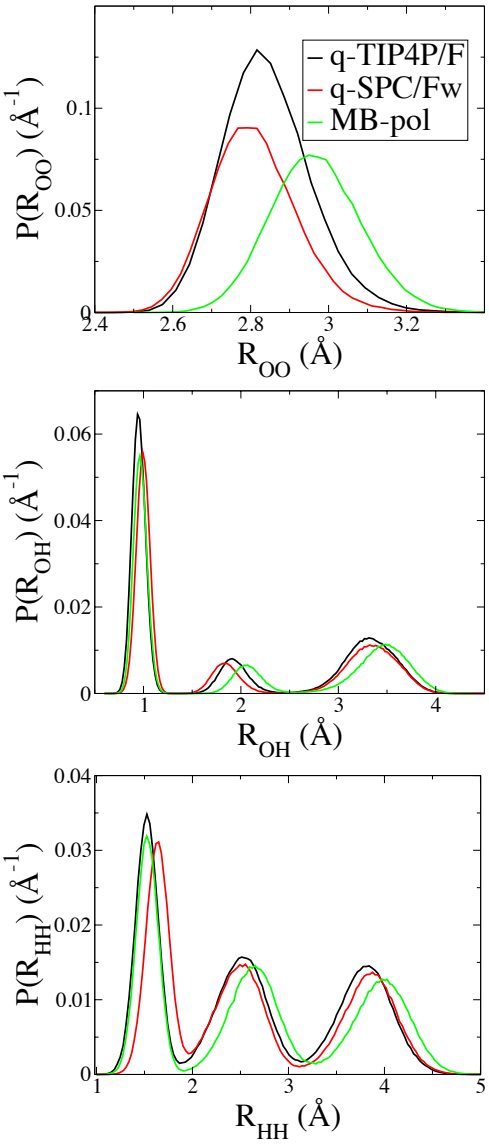


Figure 7.8: Probability distribution functions for the distance between oxygen atoms R_{OO} (upper panel), oxygen and hydrogen atoms R_{OH} (middle panel), and distance between hydrogen atoms R_{HH} (lower panel) for the water dimer described by each of the models.

three models are quite different, especially for the dimer case, except coincidentally the q-TIP4P/F and the MB-pol have very similar R_{HH} distributions for the water monomer.

7.3.3 Vibrational energies of the water monomer

As demonstrated in Section 7.2.2, vibrational transition energies can be calculated from LePIGS simulations in a straight-forward method through imaginary time correlation functions. DMC can also be used to calculate these through methods such as POITSE[130, 131] and correlation function quantum Monte Carlo (CFQMC)[132, 133, 134].

As proof of principle, the three vibrational excitation energies corresponding to the symmetric stretch, antisymmetric stretch, and symmetric bend for the H_2O and D_2O monomer systems have been calculated and compared to experimental values and energies calculated by exact diagonalization methods. It is important to note that the following analysis is done after a completed LePIGS simulation, just as the other structural and energetic properties are calculated, since each imaginary time correlation function just requires the positions of the beads at each simulation time step. Therefore, all three correlation functions required for the vibrational excitation energies can be calculated from the same simulation.

To calculate the vibrational energies using LePIGS simulations, an operator is required

to cause a transition from the ground state to the first excited state through the imaginary time correlation function. The operators chosen are exactly the symmetry coordinates for the water molecule for the symmetric and antisymmetric stretch modes and related to the symmetry coordinate of the bending mode. The operator for the symmetric stretch is the sum of the bond lengths, OH_1+OH_2, for the antisymmetric stretch is the difference of bond lengths, OH_1-OH_2, and for the bend is $\cos(\theta)$. As well, the theory dictates that the imaginary time correlation functions between the ground state and the nth excited state take the form

$$C(\tau) = |\langle 0|\hat{O}|n\rangle|^2 \exp(-\tau E_n) \tag{7.11}$$

where $\hat{O}$ is the transition operator and E_n is the difference in energy between the nth excited state and the ground state. This can also be rearranged and put in a linear form,

$$-\ln\left[C(\tau)\right] = -\ln[|\langle 0|\hat{O}|n\rangle|^2] - \tau E_n \ , \tag{7.12}$$

which allows us to perform a linear regression with respect to τ to calculate the vibrational energy level, E_n, as the slope. The number of data points used in the fit is that which minimizes the error in E_n. However, the error bars reported are the standard errors on the data points and not the error from the fit, as we feel these are a better reflection of the variance.

A sample imaginary time correlation function of the symmetric stretch of H_2O is shown in Fig. 7.9 along an exponential fit represented by Eq. 7.11. Reverting this to the linear form represented by Eq. 7.12 is also shown in Fig. 7.10 along with the linear fit. We note there is excellent agreement between the fit and the data points, especially as τ decreases, since there are more data points that contribute to the steep part of the curve.

The following protocol is used to calculate the imaginary time correlation functions (ITCF) from a LePIGS simulation. For each step of the trajectory, the value of the operator is calculated for the system at each path integral bead. Two ITCFs for each operator are generated at each simulation step, as the correlation functions are calculated by starting from the middle bead and going outwards by $(P - \beta_{opt}/\tau)/2$ beads towards both ends of the path so as to remain in the "ground-state regime," where β_{opt} is the optimized value shown in Tables 7.1 and 7.2 and τ must be systematically converged as before. The means are subtracted before correlating, so each has a smooth decay to zero. The individual ITCFs are then averaged, two from each step and over all simulation steps, to obtain a smooth, averaged ITCF from which the vibrational energies are determined through an exponential fit. The simulations are run for at least 10^5 independent simulation steps. For each simulation, to choose an optimal β parameter that allows for the full decorrelaton of our ITCF. We also choose a $\tau < 3$ a.u. which is enough to converge our energies using the mixed estimator and also ensures that we have a large number of data points for our

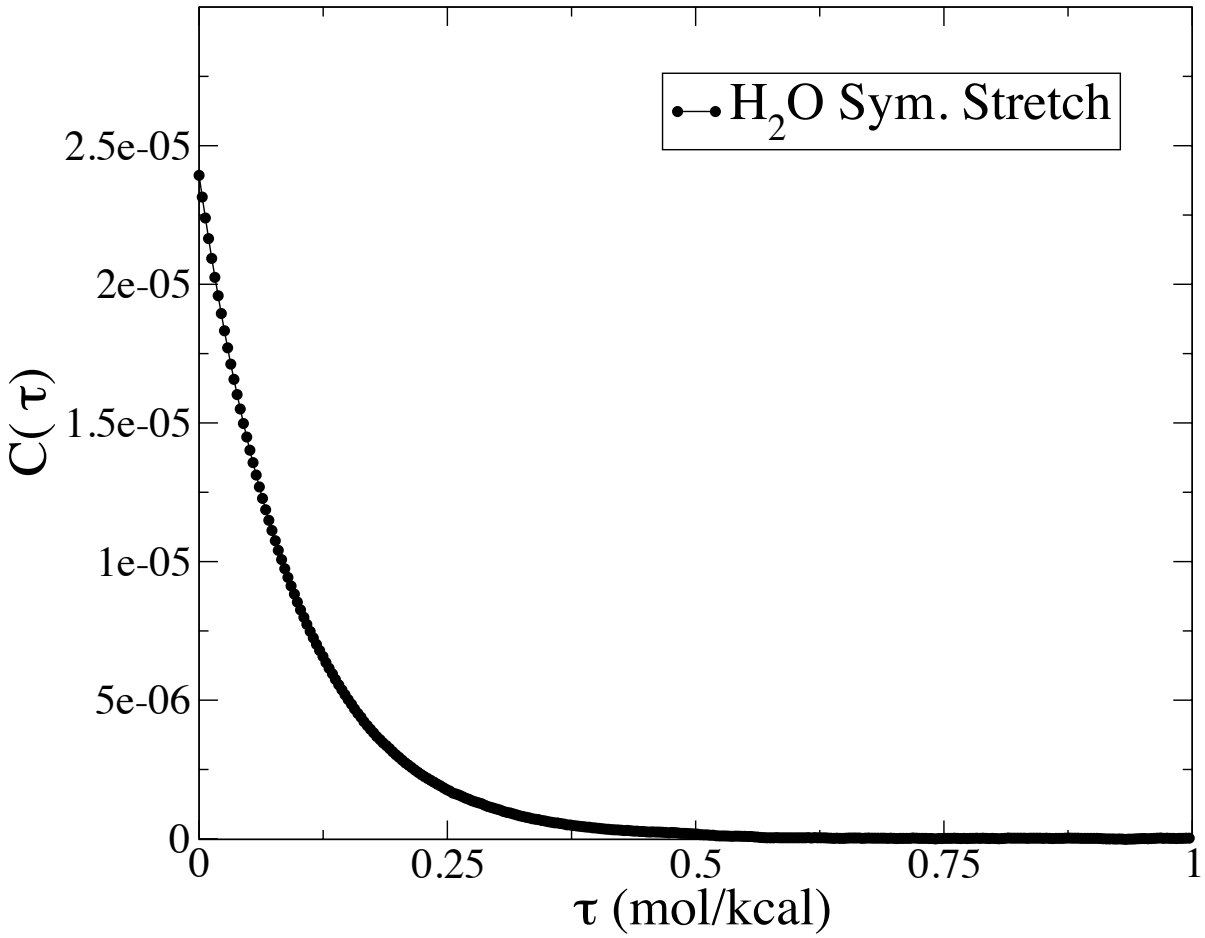


Figure 7.9: Sample imaginary time correlation function of the H_2O symmetric stretch operator calculated from a LePIGS simulation using the MB-pol potential. The data points are shown as black circles and are connected by lines as a visual aid to view the exponential nature of the correlation function.

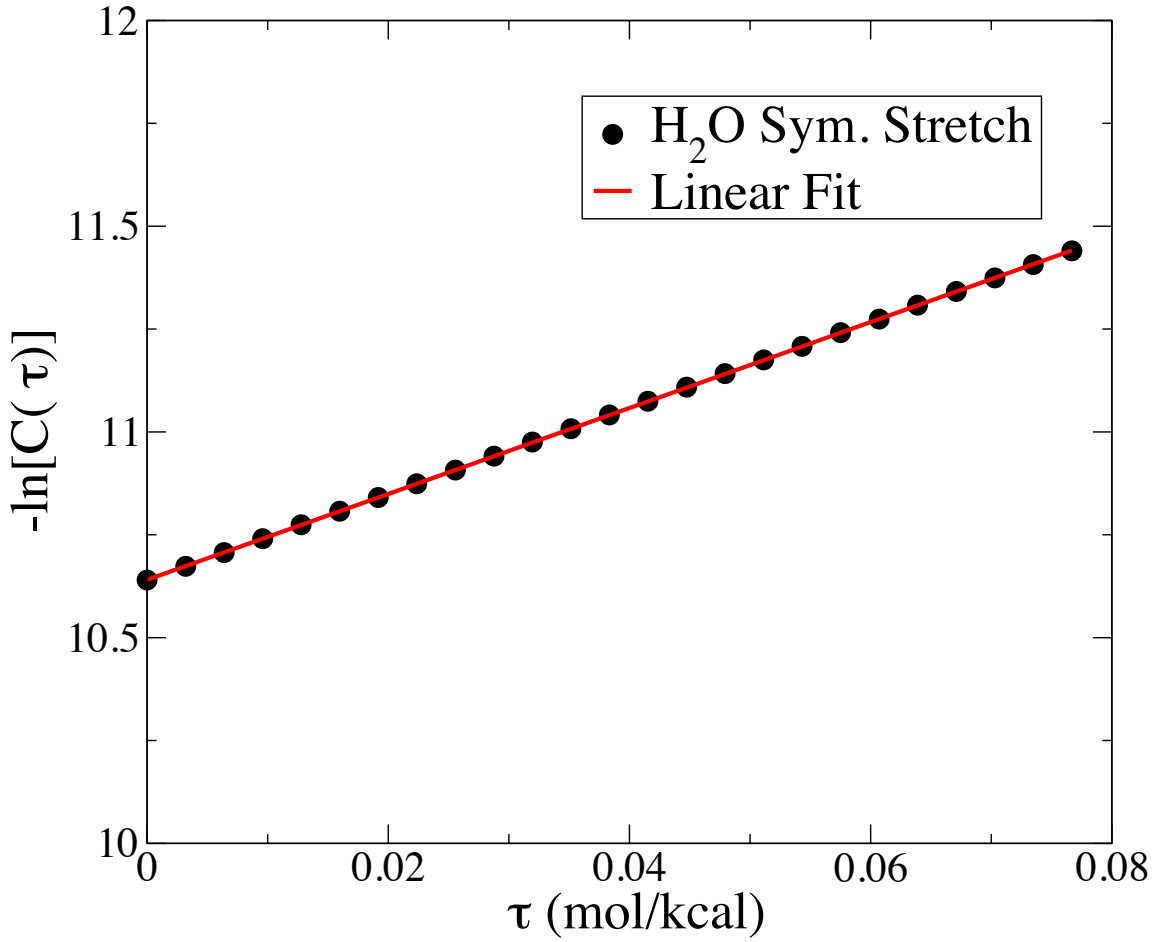


Figure 7.10: Sample log plot of the imaginary time correlation function of the H_2O symmetric stretch operator calculated from a LePIGS simulation using the MB-pol potential. The data points are represented as black circles and the linear regression is represented by a red line.

Table 7.7: Calculation of H_2O pure vibrational transition energies (kcal/mol) for each vibrational mode using the MB-pol potential.

	LePIGS	Exact	Exact (Ref. [127])	Experiment (Ref. [135])
Bend	4.56(3)	4.5590	4.5597	4.5592
Sym. Stretch	10.44(3)	10.454	10.456	10.455
Antisym. Stretch	10.83(5)	10.737	10.739	10.738

Table 7.8: Calculation of D_2O pure vibrational transition energies (kcal/mol) for each vibrational mode using the MB-pol potential.

	LePIGS	Exact	Exact(Ref. [127])	Experiment (Ref. [135])
Bend	3.38(1)	3.3687	3.3687	3.3690
Sym. Stretch	7.64(2)	7.6376	7.6377	7.6382
Antisym. Stretch	8.01(3)	7.970	7.970	7.971

ITCF.

As before with the ground state energies, the vibrational energies for the monomer systems can be calculated exactly using the Lanczos diagonalization in Pekeris co-ordinates. As mentioned, for the water monomer, the MB-pol potential reduces to the potential of Partridge and Schwenke [127]. This group had calculated the rovibrational energies to which we can also compare. These energies for the H_2O and D_2O monomer are shown in Tables 7.7 and 7.8 respectively, along with the simulation results and experimental results.

It should first be noted that, based on the exact calculations, the MB-pol potential gives excellent agreement to experimental results for the pure vibrational energies of both isotopologues. As well, the LePIGS results are in good agreement with the exact results

for the bend and symmetric stretch vibrational modes, but less so with the vibrational energy corresponding to the antisymmetric stretch.

A few things to mention are that despite using a minimum of 10^5 independent correlation functions for our statistics, the error bars are still quite large. Also, the calculated vibrational energy corresponding to the antisymmetric stretch is larger than exact calculation. The symmetry of the antisymmetric stretch is B_1, where the symmetry of the other two modes are A_1, so it could be that modes that are not fully symmetric take longer to converge. Also, even though the operator used for the antisymmetric stretch corresponds to the symmetry operator of the motion, it may be possible to use more complex operators to better represent the antisymmetric stretch. Using a trial wavefunction that includes more information about the system could help, even if for efficiency reasons to run longer simulations to decrease the error bars. Certainly, to calculate the vibrational frequencies of more complex systems, including those of the dimer where the individual water monomers can be labeled 'donor' and 'acceptor,' a better trial wavefunction is needed in order to make the computations more efficient. Such calculations will be the subject of future work.

7.4 Concluding remarks

We have demonstrated that the domain of applicability of the LePIGS method can be extended to flexible molecular systems. The water monomer and dimer have been successfully simulated in the ground state using the Langevin equation Path Integral Ground State method denoted LePIGS. Ground state energies, dissociation energies, and distributions have been calculated for the H_2O and D_2O monomer and dimer using three different water models: q-TIP4P/F and q-SPC/Fw, which are empirically based and MB-pol which is *ab initio*. We note that LePIGS samples the square of the ground state wave function, so structural and energetic properties can be obtained in a straight-forward manner. We calculate ground state energies and dissociation energies and benchmark against very accurate DMC calculations from Refs. [118, 119] which show the MB-pol potential reproduces experimental dissociation energies[120]. We note that treating the water molecule as a harmonic oscillator gives reasonable results, but overestimates the energies.

We also compare the ground state LePIGS method and the finite-temperature PIMD method. Although each method starts from a different partition function, there are only a few differences that result. In PIMD, each particle is represented by a closed path, but in LePIGS, a link is broken which results in an open path and an estimate of the ground state wavefunction is affixed to each end. In the PILE thermostat, the normal mode frequencies

are slightly redefined and the Fourier transform used in the finite-temperature case for a closed path is replaced by a cosine transform in LePIGS for an open path. We demonstrate that the LePIGS energies can be seamlessly merged with finite temperature PIMD results.

We also demonstrate that LePIGS can be used to calculate vibrational spectra of the three modes of the water monomer through imaginary time correlation functions. It is shown that the vibrational transition energies can be calculated for both the H_2O and D_2O monomer. Through "exact" methods, the MB-pol model is found to reproduce the experimental vibrational energies. Using LePIGS, the vibrational energies corresponding to the symmetric stretch and bend are reproduced nicely, however the calculated energy of the antisymmetric stretch is larger than expected. It could be that the fully symmetric modes are easier to converge than those that aren't or perhaps a more complex operator could be found that better represents the antisymmetric stretching motion.

This also lends itself towards future work to increase the complexity of our trial wavefunction in order to investigate the vibrational energies of the water dimer, which requires determining a way to distinguish between the donor and acceptor energies. Better adapted trial wavefunctions would also facilitate the simulation of larger clusters that possess several local minima such as the recently studied water hexamer[119]. As well, a recently developed path integral molecular dynamics method[136, 137] allows for the calculation of energy levels of different symmetries, more general than the transition from ground state

to first excited state from which we are calculating here. However, they only get excitation energies and not other ground state expectations, $\langle\psi_n|\hat{O}|\psi_n\rangle$, as the theory is based on the ratio of density matrices. Implementing this feature into LePIGS may assist in calculating some of the finer tunnelling splittings which occur in the water dimer.

Chapter 8

Conclusions and future work

8.1 Concluding remarks

Overall, this thesis accomplishes a number of goals. In Chapter 3, a preliminary study on confined hydrogens within the small (5^{12}) and large ($5^{12}6^4$) cages of the structure-II clathrate hydrate is performed using finite temperature path integral molecular dynamics (PIMD) and the Langevin equation path integral ground state (LePIGS). It is determined using the Buck model[70] to represent the hydrogen interactions and the Valiron potential[79, 78] to represent the hydrogen-water interactions, that the most stable ground state occupancy of hydrogens is 1 in the small cage and 4 in the large cage. This is in agreement with prior literature results[23]. The energies at finite temperature are merged

with ground state energies and it is found that the energies converge at approximately 7 K, suggesting this is when they may reach the ground state in this confined environment.

We then focus on the models of the individual components, starting with hydrogen in Chapter 4, where we theoretically calculate the Raman shifts of pure hydrogen clusters and their isotopologues and compare to experimental results to test the accuracy of the interaction potentials. A first-order perturbation theory approach is employed, which requires ground state pair distribution functions, which we generate from LePIGS simulations. For each isotopologue, 1-D potential energy surfaces (PESs) are constructed from the 6-D Hinde potential[68] using the adiabatic hindered rotor approach and are fit to analytic Morse Long Range (MLR) forms for general use. Using the H1-P interaction potential to generate our pair distribution functions gave Raman shifts that are in good agreement with experimental results for the parahydrogen and orthodeuterium clusters and outperformed those obtained using the Buck potential. Since the deviation between the predicted Raman shifts and the experimental results are on the order of fractions of wavenumbers, this is a very fine-tune test of the PES. In the future, we recommend the H1-P to be the interaction potential to model small parahydrogen clusters. In Chapter 5, we investigate the validity of the first-order perturbation theory approach to calculate the Raman shifts by comparing to two exact methods that do not introduce any approximations. The first is the direct calculation of the difference in energy of the system with no vibrational excitation and

when it has been excited and has an overall vibrational quantum number, $v = 1$. The second involves only performing a simulation of the cluster in $v = 0$. The energy when the system is in the $v = 0$ state is obtained by the energy estimator, but the energy of the system in the $v = 1$ state is obtained by taking a ratio of the density matrices. It is demonstrated that first order perturbation theory is the best approach for large cluster sizes $N > 4$. While the values calculated from all methods are within statistical errors of each other, the variance of the exact methods increases steeply with cluster size making it too difficult to draw conclusions, whereas the variance of first order perturbation theory actually decreases with cluster size. For small cluster sizes $N < 5$, it is recommended to perform two separate simulations, one when the cluster is in the $v = 0$ state and one when the cluster is in the $v = 1$ state as the variance is reasonable and no approximations are made. However, for cluster sizes $N = 5$ and beyond, it is recommended to use first order perturbation theory.

In Chapter 6, the ground state chemical potential for parahydrogen clusters between $N = 21 - 40$ is calculated using LePIGS as there has been discrepancy in the literature whether it is a smooth curve or if it's jagged, in which we can identify magic number clusters, which have enhanced stability. In general, the chemical potential calculated from diffusion Monte Carlo (DMC) was smooth, in contrast to those calculated from PIGS related methods. Our calculations using LePIGS are in agreement with recent VPIMD

results[63] (another molecular dynamics variant of PIGS) and magic number cluster sizes are predicted for $N = 23, 26, 29, 32, 34, 37$ and 39. It is speculated that this discrepancy is due to population size bias that may have been present in DMC and is not present in the PIGS methodology. We then extend our LePIGS code to simulate flexible molecular systems with the application towards the water monomer and dimer in Chapter 7. We reproduce accurate DMC calculations[118, 119] of the ground state energies and dissociation energies using LePIGS and note that the MB-pol model[124, 125, 126] of water reproduces experimental dissociation energies, thus it is recommended that for systems involving flexible water monomers, that the MB-pol model be used. We also demonstrate that LePIGS energies can be merged seamlessly with finite temperature PIMD results. Lastly, we use LePIGS to calculate vibrational transition energies of the water monomer through imaginary time correlation functions. Our results are in good agreement with our exact calculations for the symmetric stretch and bending motion, but our calculated energy for the antisymmetric stretch is larger than expected. This could be due to the different symmetry of the mode as it is not fully symmetric like the other two.

8.2 Further investigation of confined hydrogen in structure-II clathrate hydrates

There are a number of future directions that the results of this thesis lead to, but the primary goal should be to apply these results towards further understanding low temperature hydrogen and water systems, in particular systems of confined hydrogen in the structure-II clathrate hydrates due to the importance of clean energy storage in the current climate. The Hinde H1-P potential should be used for the hydrogen interaction potential as we demonstrate it outperforms the Buck potential for small hydrogen clusters based on the Raman vibrational frequency shifts. For a rigid-water cage, it would be of use to compare the 3-D Valiron PES that we use for the water-hydrogen interaction potential and the SPC/E model[17] that Bacic has used[23]. This could be done by calculating the Raman shifts of hydrogen clusters within the clathrates using the SPC/E potential or the Valiron potential as the hydrogen-cage interaction potential and the H1-P as the hydrogen pair interaction potential and comparing theoretical predictions[24, 138, 139, 140, 141] to experimental results[142, 15] As well, groups may wish to treat the clathrate hydrate as a system of flexible water monomers and we would recommend the MB-pol interaction potential between the water molecules. It would also be of use to compare the MB-pol model for water-water interactions to the SPC/E for small water clusters. Using the most

accurate interaction potentials should resolve the uncertainty in the occupancy of confined hydrogen in the structure-II clathrate hydrates and provide useful information on how to maximize the weight percentage of hydrogen to meet the standards of feasibility set by government agencies.

8.3 Continued software development of the LePIGS and PIMD methods

The LePIGS method has continued to be developed throughout this thesis, extending it to flexible monomers and using the generated imaginary time correlation functions to calculate vibrational transition energies. However, there are still developments that can be made to enhance the understanding of ground state properties of systems of interest. First, is the inclusion of rigid-body rotation. Our group is currently working on modifying our PIMD/LePIGS code to perform centre of mass translation and introducing rigid-body rotation in a path integral Monte Carlo (PIMC) framework. This hybrid code would enable users to obtain ground state properties of rigid-body systems that include rotation. This could be useful for removing the spherical approximation of hydrogen and treating it as a rigid rotor within the clathrate system. The second main application would be to use LePIGS to calculate approximate real time dynamic properties through the use

of correlation functions, similar to ring polymer molecular dynamics (RPMD)[66] but in the ground state. Work is underway to formulate the theory behind this reptile molecular dynamics (RMD) method. This could be useful for looking at tunnelling rates or obtaining spectra.

Letter of Copyright Permissions

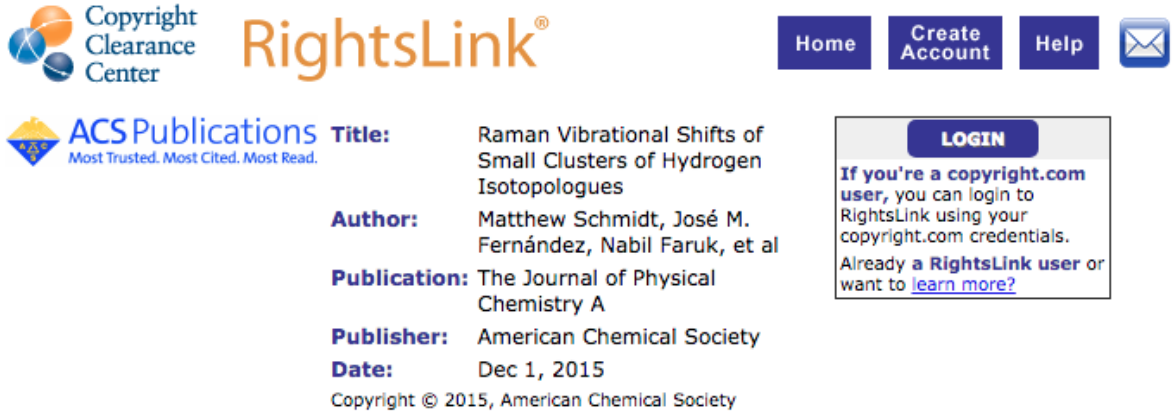
M. Schmidt and P.-N. Roy, "Path integral Molecular dynamic simulation of flexible molecular systems in their ground state: application to the water dimer", J. Chem. Phys. 148, 124116 (2018). Copyright 2018 American Institute Physics.

License Numbers: 4432600601129

4432600731331

4432600867036

M. Schmidt, J. M. Fernández, N. Faruk, M. Nooijen, R. J. Le Roy, J. H. Morilla, G. Tejeda, S. Montero, and P.-N. Roy, "Raman Vibrational Shifts of Small Clusters of Hydrogen Isotopologues," J. Phys. Chem. A 119, 12551 (2015). Copyright 2015 American Chemical Society.



PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

References

[1] United States Department of Energy Office of Energy Efficiency and Renewable Energy Guidelines for Hydrogen Storage. Available from: http://energy.gov/eere/fuelcells/hydrogen-storage.

[2] 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).

[3] V. V. Struzhkin, B. Militzer, W. L. Mao, H.-k. Mao, and R. J. Hemley, Chem. Rev. **107**, 4133 (2007).

[4] H. P. Veluswamy, R. Kumar, and P. Linga, Appl. Energy. **122**, 112 (2014).

[5] T. A. Strobel, K. C. Hester, C. A. Koh, A. K. Sum, and E. D. Sloan Jr, Chem. Phys. Lett. **478**, 97 (2009).

[6] S. Alavi and J. A. Ripmeester, Mol. Simul. **43**, 808 (2017).

[7] Y. A. Dyadin, E. G. Larionov, A. Y. Manakov, F. V. Zhurko, E. Y. Aladko, T. V. Mikina, and V. Y. Komarov, Mendeleev Commun. **9**, 209 (1999).

[8] Y. A. Dyadin, E. Larionov, E. Y. Aladko, A. Y. Manakov, F. Zhurko, T. Mikina, V. Y. Komarov, and E. Grachev, J. Struct. Chem. **40**, 790 (1999).

[9] 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).

[10] D. Colognesi, M. Celli, L. Ulivi, M. Xu, and Z. Bacic, J. Phys. Chem. A **117**, 7314 (2013).

[11] T. A. Strobel, C. A. Koh, and E. D. Sloan, J. Phys. Chem. B **112**, 1885 (2008).

[12] A. Martin and C. J. Peters, J. Phys. Chem. B **113**, 7558 (2009).

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

[14] J. Liu, J. Hou, J. Xu, H. Liu, G. Chen, and J. Zhang, Int. J. Hydrog. Energy **42**, 17136 (2017).

[15] A. Giannasi, M. Celli, L. Ulivi, and M. Zoppi, J. Chem. Phys. **129**, 084705 (2008).

[16] S. Patchkovskii and S. T. John, Proc Natl Acad Sci U S A **100**, 14645 (2003).

[17] S. Alavi, J. Ripmeester, and D. Klug, J. Chem. Phys. **123**, 024507 (2005).

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

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

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

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

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

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

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

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

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

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

[28] M. P. Hodges, R. J. Wheatley, G. K. Schenter, and A. H. Harvey, J. Chem. Phys. **120**, 710 (2004).

[29] P. Diep and J. K. Johnson, J. Chem. Phys. **112**, 4465 (2000).

[30] H. Berendsen, J. Grigera, and T. Straatsma, J. Phys. Chem. **91**, 6269 (1987).

[31] G. Tejeda, J. Fernández, S. Montero, D. Blume, and J. Toennies, Phys. Rev. Lett. **92**, 223401 (2004).

[32] R. D. Beck, M. F. Hineman, and J. W. Nibler, J. Chem. Phys. **92**, 7068 (1990).

[33] S. Bhatnagar, E. J. Allin, and H. Welsh, Can. J. Phys. **40**, 9 (1962).

[34] E. Knuth, F. Schünemann, and J. Toennies, J. Chem. Phys. **102**, 6258 (1995).

[35] D. Blume, M. Lewerenz, F. Huisken, and M. Kaloudis, J. Chem. Phys. **105**, 8666 (1996).

[36] J. Tang and A. McKellar, J. Chem. Phys. **119**, 754 (2003).

[37] J. Tang and A. McKellar, J. Chem. Phys. **119**, 5467 (2003).

[38] J. Tang, A. McKellar, F. Mezzacapo, and S. Moroni, Phys. Rev. Lett. **92**, 145503 (2004).

[39] J. Tang and A. McKellar, J. Chem. Phys. **121**, 181 (2004).

[40] A. McKellar, J. Chem. Phys. **128**, 044308 (2008).

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

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

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

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

[45] S. Moroni and M. Boninsegni, J. Low Temp. Phys. **136**, 129 (2004).

[46] M. Holzmann, B. Bernu, C. Pierleoni, J. McMinis, D. M. Ceperley, V. Olevano, and L. Delle Site, Phys. Rev. Lett. **107**, 110402 (2011).

[47] I. Bosá and S. M. Rothstein, J. Chem. Phys. **121**, 4486 (2004).

[48] S. A. Chin, Phys. Lett. A **226**, 344 (1997).

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

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

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

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

[53] Y. Yan and D. Blume, J. Phys. B: At. Mol. Opt. Phys. **50**, 223001 (2017).

[54] C. M. Herdman, P.-N. Roy, R. G. Melko, and A. D. Maestro, Phys. Rev. B **89**, 140501 (2014).

[55] C. Herdman, S. Inglis, P.-N. Roy, R. Melko, and A. Del Maestro, Phys. Rev. E **90**, 013308 (2014).

[56] D. Iouchtchenko and P.-N. Roy, Estimating ground state entanglement entropy using path integral molecular dynamics, in *Recent Progress in Quantum Monte Carlo*, pages 145–154, ACS Publications, 2016.

[57] C. M. Herdman, P.-N. Roy, R. G. Melko, and A. D. Maestro, Nature Phys. **13**, 556 (2017).

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

[59] S. Miura, Comput. Phys. Commun. **182**, 274 (2011).

[60] S. Miura, Mol. Sim. **38**, 378 (2012).

[61] S. Miura, J. Phys. Conf. Ser. **454**, 012023 (2013).

[62] Y. Kamibayashi and S. Miura, J. Chem. Phys. **145**, 074114 (2016).

[63] S. Miura, J. Chem. Phys. **148**, 102333 (2018).

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

[65] M. Schmidt, S. Constable, C. Ing, and P.-N. Roy, J. Chem. Phys. **140**, 234101 (2014).

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

[67] N. Faruk, M. Schmidt, H. Li, R. J. Le Roy, and P.-N. Roy, J. Chem. Phys. **141**, 014310 (2014).

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

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

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

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

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

[73] K. Schmidt and M. A. Lee, Phys. Rev. E **51**, 5495 (1995).

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

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

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

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

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

[79] P. Valiron, M. Wernli, A. Faure, L. Wiesenfeld, C. Rist, S. Kedžuch, and J. Noga, J. Chem. Phys. **129**, 134306 (2008).

[80] S. F. Dec, J. Phys. Chem. C **113**, 12355 (2009).

[81] G. Mazzola and M. Troyer, arXiv:1608.01262 (2016).

[82] R. J. Le Roy, Y. Huang, and C. Jary, J. Chem. Phys. **125**, 164310 (2006).

[83] R. J. Le Roy and R. D. E. Henderson, Mol. Phys. **105**, 663 (2007).

[84] D. Eichenauer and R. J. Le Roy, J. Chem. Phys. **88**, 2898 (1988).

[85] D. J. Chartrand, J. C. Shelley, and R. J. L. Roy, J. Phys. Chem. **95**, 8310 (1991).

[86] Y. Kwon and K. B. Whaley, Phys. Rev. Lett. **89**, 273401 (2002).

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

[88] T. Zeng, H. Li, and P.-N. Roy, J. Phys. Chem. Lett. **4**, 18 (2013), PMID: 26291205.

[89] L. Wang, D. Xie, R. J. L. Roy, and P.-N. Roy, J. Chem. Phys. **139**, 034312 (2013).

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

[91] J. H. Morilla, J. M. Fernández, G. Tejeda, and S. Montero, Phys. Chem. Chem. Phys. **12**, 12060 (2010).

[92] D. E. Jennings, A. Weber, and J. W. Brault, Appl. Opt. **25**, 284 (1986).

[93] S. Montero, J. H. Morilla, G. Tejeda, and J. M. Fernández, Eur. Phys. J. D **52**, 31 (2009).

[94] E. Schrödinger, Ann. Phys. **385**, 437 (1926).

[95] J. W. Raleigh, *Theory of Sound*, 2nd edition (London: Macmillan, 1894).

[96] D. T. Colbert and W. H. Miller, J. Chem. Phys. **96**, 1982 (1992).

[97] P.-N. Roy, J. Chem. Phys. **119**, 5437 (2003).

[98] M. Nightingale and P.-N. Roy, J. Phys. Chem. A **110**, 5391 (2006).

[99] R. J. Le Roy, C. C. Haugen, J. Tao, and H. Li, Mol. Phys. **109**, 435 (2011).

[100] C. Douketis, G. Scoles, S. Marchetti, M. Zen, and A. J. Thakkar, J. Chem. Phys. **76**, 3057 (1982).

[101] R. J. Le Roy, betaFIT 2.1: A computer program to fit potential function points to selected analytic functions, University of Waterloo Chemical Physics Research Report No. CP-665 (2009).

[102] R. Le Roy, J. Seto, and Y. Huang, DPotFit 2.0: A computer program for fitting diatomic molecule spectra to potential energy functions, University of Waterloo Chemical Physics Research Report CP-667 (2013).

[103] H.-G. Yu, J. Chem. Phys. **120**, 2270 (2004).

[104] L. S. Costa and D. C. Clary, J. Chem. Phys. **117**, 7512 (2002).

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

[106] N. Faruk, M. Schmidt, H. Li, and P.-N. Roy, J. Chem. Phys. **141**, (2014).

[107] M. Schmidt, J. M. Fernandez, N. Faruk, M. Nooijen, R. J. Le Roy, J. H. Morilla, G. Tejeda, S. Montero, and P.-N. Roy, J. Phys. Chem. A **119**, 12551 (2015).

[108] V. Ambegaokar and M. Troyer, Am. J. Phys. **78**, 150 (2010).

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

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

[111] F. Mezzacapo and M. Boninsegni, J. Phys. Condens. Matter **21**, 164205 (2009).

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

[113] D. Rayane, P. Melinon, B. Cabaud, A. Hoareau, B. Tribollet, and M. Broyer, Phys. Rev. A **39**, 6056 (1989).

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

[115] J. Navarro and R. Guardiola, J. Low Temp. Phys. **148**, 857 (2007).

[116] R. Guardiola and J. Navarro, Open Phys. **6**, 33 (2008).

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

[118] J. D. Mallory, S. E. Brown, and V. A. Mandelshtam, J. Phys. Chem. A **119**, 6504 (2015).

[119] J. D. Mallory and V. A. Mandelshtam, J. Chem. Phys. **145**, 064308 (2016).

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

[121] F. Paesani, W. Zhang, D. A. Case, T. E. Cheatham, and G. A. Voth, J. Chem. Phys. **125**, 184507 (2006).

[122] S. K. Reddy, D. R. Moberg, S. C. Straight, and F. Paesani, J. Chem. Phys. **147**, 244504 (2017).

[123] S. Habershon, T. E. Markland, and D. E. Manolopoulos, J. Chem. Phys. **131**, 024501 (2009).

[124] V. Babin, C. Leforestier, and F. Paesani, J. Chem. Theory Comput. **9**, 5395 (2013).

[125] V. Babin, G. R. Medders, and F. Paesani, J. Chem. Theory Comput. **10**, 1599 (2014).

[126] G. R. Medders, V. Babin, and F. Paesani, J. Chem. Theory Comput. **10**, 2906 (2014).

[127] H. Partridge and D. W. Schwenke, J. Chem. Phys. **106**, 4618 (1997).

[128] K. P. Bishop and P.-N. Roy, J. Chem. Phys. **148**, 102303 (2018).

[129] Y. D. Liu and P.-N. Roy, J. Chem. Phys. **121**, 6282 (2004).

[130] D. Blume, M. Lewerenz, P. Niyaz, and K. Whaley, Phys. Rev. E **55**, 3664 (1997).

[131] P. Huang, A. Viel, and K. B. Whaley, Efficient implementation of the projection operator imaginary time spectral evolution (poitse) method for excited states, in *Recent Advances In Quantum Monte Carlo Methods - Part II*, pages 111–126, World Scientific, 2002.

[132] D. Ceperley and B. Bernu, J. Chem. Phys. **89**, 6316 (1988).

[133] B. Bernu, D. Ceperley, and W. Lester Jr, J. Chem. Phys. **93**, 552 (1990).

[134] F. V. Prudente, L. Costa, and P. H. Acioli, J. Phys. B **33**, R285 (2000).

[135] T. Shimanouchi, J. Phys. Chem. Ref. Data **1**, 189 (1972).

[136] E. Mátyus, D. J. Wales, and S. C. Althorpe, J. Chem. Phys. **144**, 114108 (2016).

[137] E. Mátyus and S. C. Althorpe, J. Chem. Phys. **144**, 114109 (2016).

[138] K. Ramya and A. Venkatnathan, J. Chem. Phys. **138**, 124305 (2013).

[139] Z. Futera, M. Celli, L. del Rosso, C. J. Burnham, L. Ulivi, and N. J. English, J. Phys. Chem. C **121**, 3690 (2017).

[140] N. Plattner and M. Meuwly, J. Chem. Phys. **140**, 024311 (2014).

[141] A. Powers, Y. Scribano, D. Lauvergnat, E. Mebe, D. M. Benoit, and Z. Bačić, J. Chem. Phys. **148**, 144304 (2018).

[142] T. A. Strobel, E. D. Sloan, and C. A. Koh, J. Chem. Phys. **130**, 014506 (2009).

Appendices

Appendix A

A simplified model for the many-body problem

In this appendix,[1] we describe a simple model and in a general way determine how to calculate the vibrational shift. Throughout, we describe specifically how this theory relates to experimental observations. Let us consider a cluster formed by N moving H_2 molecules, whose vibrational wavefunctions will be denoted by $|v\rangle$. Each molecule will be either in the ground ($v = 0$) or in the first excited ($v = 1$) vibrational state and will move with respect to all the other molecules, but we will neglect any rotational motion. We can recognize a

[1]Sections of this appendix have been reprinted with permission from M. Schmidt, J. M. Fernndez, N. Faruk, M. Nooijen, R. J. Le Roy, J. H. Morilla, G. Tejeda, S. Montero, and P.-N. Roy, "Raman Vibrational Shifts of Small Clusters of Hydrogen Isotopologues", J. Phys. Chem. A 119, 12551 (2015). Copyright 2015 American Chemical Society.

timescale separation between the fast (intramolecular) vibrational degrees of freedom and the slow (intermolecular) translational ones. This suggests an adiabatic approximation, where the vibrational Hamiltonian will have a parametric dependence on the slow intermolecular coordinates. A basis set for the vibrational wavefunction of the cluster, having one quantum of vibrational excitation in the ith-monomer, can be represented as

$$|\mathbf{1_i}\rangle = |0\rangle|0\rangle...|1_i\rangle|0\rangle...|0\rangle \tag{A.1}$$

while the ground state is represented by

$$|\mathbf{0}\rangle = |\psi_{\mathrm{v}=0}\rangle = |0\rangle|0\rangle...|0\rangle \ . \tag{A.2}$$

Let us assume a field of strength ϵ, that promotes a local state $|0_i\rangle$ to $|1_i\rangle$, represented by

$$\hat{O} = \epsilon \sum_i \hat{O}_i = \epsilon \sum_i \left(|0_i\rangle\langle 1_i| + |1_i\rangle\langle 0_i|\right) \ , \tag{A.3}$$

which is a symmetric sum of local transition operators, implying that

$$\hat{O}|\psi_{\mathrm{v}=0}\rangle = \epsilon \sum_i |\mathbf{1_i}\rangle \ . \tag{A.4}$$

It follows that by nature of the transition operator, the final internal state is symmetric. This is true independent of the bosonic nature of the particles. Given the fact that the particles are bosons, there is a second reason for this form of the wavefunction. The exact translational ground state is totally symmetric under permutations. Since $\hat{O}$ only acts on the internal degrees of freedom, a totally symmetric wavefunction would be a product of a symmetric translational part and a symmetric internal part. Following either argument, the internal excited wavefunction is given by

$$|\psi_{\mathrm{v}=1}\rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} |\mathbf{1_i}\rangle \ . \tag{A.5}$$

In the present experiment, the transition operator responsible for the Raman spectrum can be described by the mean molecular polarizability, which is scalar in nature. This requires that the final internal state be symmetric and that effects of internal rotation of the molecules in the cluster be neglected.

We wish to evaluate the energy for such a state for a 2-body Hamiltonian

$$\hat{H} = \hat{h} + \hat{V} \ . \tag{A.6}$$

where $\hat{h}$ represents the collective one-body, non-interacting Hamiltonian and $\hat{V}$ is the in-

teraction potential. The relevant matrix elements are

$$h_0^i = \langle 0_i|\hat{h}|0_i\rangle \tag{A.7}$$

$$h_1^i = \langle 1_i|\hat{h}|1_i\rangle \tag{A.8}$$

$$V_{ij} = \langle 0_i 0_j|\hat{V}|0_i 0_j\rangle \tag{A.9}$$

$$D_{ij} = \langle 1_i 0_j|\hat{V}|1_i 0_j\rangle \tag{A.10}$$

$$X_{ij} = \langle 1_i 0_j|\hat{V}|0_i 1_j\rangle \tag{A.11}$$

The above model essentially assumes that the H_2 monomers are pinned in space and these matrix elements will depend on the positions of the monomers. The Hamiltonian matrix for the $v = 1$ excited states with overall one quantum of vibrational excitation can be written as

	$\|1,0,0,...\rangle$	$\|0,1,0,...\rangle$	$\|0,0,1,...\rangle$	$\cdots$
$\|1,0,0,...\rangle$	$H_d^{(1)}$	X_{12}	X_{13}	$X_{14}\cdots$
$\|0,1,0,...\rangle$	X_{21}	$H_d^{(2)}$	X_{23}	$X_{24}\cdots$
$\vdots$			$\ddots$	

with $H_d(i)$ given by

$$H_d^{(i)} = \langle \mathbf{1_i} | \hat{H} | \mathbf{1_i} \rangle = h_1^i + \sum_{j \neq i} h_0^j + \sum_{j \neq i} D_{ij} + \frac{1}{2} \sum_{k,l \neq i} V_{kl} \ , \tag{A.12}$$

where all sums in what follows run over all monomers. It then follows that the total energy of the N-body system in the $v = 1$ excited vibrational state is

$$E_N^{\mathrm{v}=1} = \langle \psi_{\mathrm{v}=1} | \hat{H} | \psi_{\mathrm{v}=1} \rangle = \frac{1}{N} \sum_{i,j} \langle \mathbf{1_i} | \hat{H} | \mathbf{1_j} \rangle = \frac{1}{N} \left(\sum_i H_d^{(i)} + \sum_i \sum_{j \neq i} X_{ij} \right) \tag{A.13}$$

while the energy of the system in the ground vibrational state ($v = 0$) is

$$E_N^{\mathrm{v}=0} = \langle \psi_{\mathrm{v}=0} | \hat{H} | \psi_{\mathrm{v}=0} \rangle = \sum_i h_0^{(i)} + \frac{1}{2} \sum_{k,l} V_{kl} \ . \tag{A.14}$$

The difference of Eq.s (A.13) and (A.14) gives the energy ν_N of the $v = 0 \rightarrow 1$ transition in the cluster

$$\nu_N = E_N^{\mathrm{v=1}} - E_N^{\mathrm{v=0}} = \frac{1}{N}\left(\sum_i h_1^{(i)} - \sum_i h_0^{(i)}\right) + \frac{1}{N}\sum_i \sum_{j\neq i} \left(D_{ij} + X_{ij} - V_{ij}\right) \ . \quad \text{(A.15)}$$

The $(h_1 - h_0)$ term gives the energy of the $v = 0 \rightarrow 1$ transition in the (free) non-interacting monomers, so if we remove these terms we get the vibrational shift of the cluster with respect to the free molecule as

$$\Delta\nu_N(\mathbf{R}) = \frac{1}{N}\sum_i \sum_{j\neq i} \left(D_{ij}(|\mathbf{R_j} - \mathbf{R_i}|) + X_{ij}(|\mathbf{R_j} - \mathbf{R_i}|) - V_{ij}(|\mathbf{R_j} - \mathbf{R_i}|)\right) \quad \text{(A.16)}$$

Here we have explicitly included position dependence into our theory and the matrix elements on the right-hand side of the equation only depend on the relative distance between monomers. Now, if we assume the monomers to be identical and to be described by a pair distribution function $P_N(R)$ normalized as

$$\int_0^\infty P_N(R)dR = 1 \quad \text{(A.17)}$$

where R is the relative distance between monomers and we use Eq. A.16 to form an average

difference potential

$$(N-1)(D+X-V)(R) = (N-1)\Delta V(R)\ , \tag{A.18}$$

we can use a first-order perturbation theory approach by sampling the difference potential using the ground state distribution

$$\Delta\nu_N = (N-1)\int_0^\infty \Delta V\,(R)\,P_N\,(R)\,dR\,. \tag{A.19}$$

The use of perturbation theory to calculate energy differences for a dopant (e.g. CO_2) in a cluster of monomers (e.g. H_2) has been widely used[84, 85, 86, 87, 88]. The specific case when the dopant is identical to the monomers, leading to Eq. A.19, was first considered in Ref. [31]. Here we derived the expression using a simplified model and emphasize that it is the symmetric nature of the transition operator that selects the symmetric state and this is largely independent of the bosonic character of the particles.

The quantity $\Delta V(R)$ that enters the theory can be determined explicitly by considering the Hamiltonian for the dimer, with energies of the excited state given by

$$E_\pm(R) = h_1 + h_0 + D(R) \pm X(R) \tag{A.20}$$

and of the ground state given by

$$E_0(R) = 2h_0 + V(R) \ . \tag{A.21}$$

The difference between the ground and excited vibrational state energy, using the symmetric state with energy E_+, is given by

$$E_+(R) - E_0 = (h_1 - h_0) + (D + X - V)(R) \ . \tag{A.22}$$

The actual calculation of E_+ and E_0 involves the construction of the full rovibrational Hamiltonian for the dimer. The energy E_+ is that of the lowest symmetric state with one quantum of vibration and is not necessarily the lowest energy state for the dimer. The full potential for a cluster of N monomers with one quantum of vibration can be expressed in terms of the ground state potential $V(R)$ and the difference potential $\Delta V(R)$ as

$$V_N^{\mathrm{v}=1}(\mathbf{R}) = \sum_{i<j} V(R_{ij}) + \frac{2}{N} \sum_{i<j} \Delta V(R_{ij}) \ , \tag{A.23}$$

where the $2/N$ factor comes from the restriction of counting distinct pairs.