# Path integral Langevin dynamics of complex molecular systems: from low-temperature quantum clusters to biomolecules

by

Christopher E. Ing

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

Waterloo, Ontario, Canada, 2011

© Christopher E. Ing 2011

I hereby declare that I am the sole author of this 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.

## Abstract

This thesis presents an implementation of path integral molecular dynamics (PIMD) for sampling equilibrium and dynamical properties within the molecular modelling toolkit (MMTK) [J. Comp. Chem. **21**, 79 (2000)], an open source Python package. Rigorous simulation using this code serves to benchmark this implementation as well as the robustness of the path integral Langevin equation as a thermostat [J. Chem. Phys. **133**, 124104 (2010)].

PIMD is used to calculate equilibrium properties for clusters of $He_N$-$CO_2$ at low-temperatures, with comparison to experimental and exact results. We characterize the convergence of structural and energetic properties as a function of path-integral discretization error. The radial and angular distribution of these clusters is studied as a function of size in the absence of rotation and bosonic exchange. These distributions are subsequently used to calculate vibrational shifts of $CO_2$. This result is compared to high-accuracy path integral Monte Carlo simulations which include rotational and exchange effects. These simulations indicate that the neglect of rotational degrees of freedom leads to an unphysical localization of helium atoms and incorrect vibrational shifts when compared to experiment.

Approximate real-time quantum dynamics is presented for doped helium clusters using the ring-polymer molecular dynamics (RPMD) method. The accuracy of RPMD is tested

for low-temperature simulations and compared to exact results. Preliminary calculation of the dynamics of the helium solvated $CO_2$ dopant with respect to the center of mass of the cluster is presented. The effect of a cartesian integrator versus a normal-mode integrator for quantum dynamics is addressed.

The path integral ground-state method is applied in order to calculate $T = 0K$ properties. A convergence study of the ground-state energy of the quantum harmonic oscillator with respect to sampling time and path discretization is shown. As a final application of this implementation, a sugar in a periodic water box is simulated at $T = 300K$. The calculation of rotamer populations and a dipole autocorrelation indicate negligible change with the inclusion of quantum effects.

This work offers a comprehensive foundation from which to base future PIMD centered research.

**Acknowledgements**

First and foremost I'd like to thank my advisor Dr. Pierre-Nicholas Roy for providing me with an ideal mixture of both software development and scientific challenges. My project evolved in a very natural way, and I couldn't be happier with how it turned out. Pierre's infinite patience with regards to my path-integral ignorance has been a blessing during my entire stay in his research group. Going beyond the line of duty of an advisor, Pierre managed to juggle the projects of an entire group, publish, outfit every desk with a mac, travel, build a huge computing cluster, write grants, and find time to debug my code on numerous occasions. Pierre's enthusiasm for science was a source of constant inspiration and nothing could beat the sense of satisfaction from him calling group members over to your computer to explain why your result is a big deal!

I owe Pierre extreme gratitude for encouraging me to apply to the France Consulate Scholarship to work with Konrad Hinsen in France on MMTK. I'd like to thank Konrad for revealing the glory of Python to me on numerous occasions, and for translating the names of many French food items for me. My degree would not be possible without my supportive co-supervisor Dr. Wing-Ki Liu. I'd like to thank Matthew Badali for introducing me to Pierre in the first place, and for suggesting that I speak to Wing-Ki about being a co-supervisor.

I'd like to thank my collaborators Jing Yang, Tao Zeng, Hui Li, and my future collaborators Steve Constable and Matthew Schmidt, all of whom I've had excellent discussions with in the office or at conferences. I'd like to thank all those in the theoretical chemistry group who have helped build an excellent environment to do computational research.

## Dedication

To my nephew Jack; may he be a super rockstar one day!

# Table of Contents

**Author's Declaration** ii

**Abstract** iii

**Acknowledgements** v

**Dedication** vii

**List of Tables** xi

**List of Figures** xii

**List of Abbreviations** xiv

**1 Introduction** 1

1.1 Molecular Dynamics . . . . . 3

1.2 Path integral Formulation of the Partition Function . . . . . 8

1.3 Path integral Molecular Dynamics . . . . . 11

1.3.1 Normal Mode Path Integral Molecular Dynamics . . . . . 14

1.4 Overview of Thesis . . . . . 17

**2 Equilibrium properties for doped helium clusters** **20**

2.1 Theoretical Approach . . . . . 22

2.1.1 Energy Evaluators . . . . . 22

2.1.2 Thermostats . . . . . 23

2.2 Computational Methods . . . . . 25

2.2.1 Simulation Parameters . . . . . 26

2.2.2 Analysis Parameters . . . . . 31

2.2.3 Exact basis set calculation . . . . . 32

2.3 Results and Discussion . . . . . 35

2.3.1 Energy Convergence . . . . . 35

2.3.2 Center of Mass Motion . . . . . 36

2.3.3 Structural Properties . . . . . 40

2.3.4 Vibrational Shifts . . . . . 47

2.4 Conclusion and Outlook . . . . . 49

**3 Dynamical properties for doped helium clusters** **51**

3.1 Theoretical Approach . . . . . 53

3.1.1 Ring Polymer Molecular Dynamics. . . . . 54

3.2 Computational Methods . . . . . 57

3.3 Results and Discussion . . . . . 59

3.3.1 Quartic Oscillator . . . . . 59

3.3.2 Helium Dynamics about $CO_2$ . . . . . 60

3.3.3 Cluster Dopant Dynamics . . . . . 62

3.4 Conclusion and Outlook . . . . . 64

**4 Properties for complex molecular systems** **66**

4.1 Ground State Equilibrium Properties . . . . . 68

4.1.1 Theoretical Approach . . . . . 68

4.1.2 Computational Methods . . . . . 72

4.1.3 Results and Discussion . . . . . 75

4.2 Biomolecular Simulation . . . . . 77

4.2.1 Computational Methods . . . . . 78

4.2.2 Results and Discussion . . . . . 80

4.3 Conclusion and Outlook . . . . . 83

**5 Conclusion** **85**

5.1 Future Directions . . . . . 88

**References** **91**

# List of Tables

2.1 Parameters of the MLR fit of the helium dimer potential. . . . . 29

4.1 Rotamer population distribution about the C4-C5 bond (*gt tg gg*) . . . . . 82

# List of Figures


1.1 Path integral isomorphism diagram for two atoms . . . . . 11

2.1 Isometric snapshot of a He-$CO_2$ molecular dynamics simulation . . . . . 26

2.2 Path integral discretization error convergence plot of the average energy of He-$CO_2$ . . . . . 37

2.3 Center of mass energy distribution for PILE/WNLE thermostats used on He-$CO_2$ . . . . . 39

2.4 Radial density convergence for increasing time slices of He-$CO_2$ . . . . . 41

2.5 Average $r$ and $r^2$ values for increasing time slices of He-$CO_2$ . . . . . 42

2.6 Chemical potential of for ($He_N$)-$CO_2$ . . . . . 43

2.7 Angular and radial distributions for $He_N$-$CO_2$ clusters . . . . . 45

2.8 Vibrational shifts of $He_N$-$CO_2$ as a function of cluster size . . . . . 46

2.9 Rotational distribution of $(He)_{17}$-$CO_2$ with and without rotation . . . . . 48

3.1 Ring polymer molecular dynamics simulation schematic . . . . . 56

3.2 Coordinate autocorrelations for the quantum quartic oscillator . . . . . 60

3.3 Coordinate autocorrelation for He-$CO_2$ . . . . . 61

3.4 Coordinate autocorrelation for $CO_2$ with respect to $(He)_{17}$-$CO_2$ center of mass for two integrators . . . . . 63

4.1 Path integral ground state computational methodology . . . . . 71
4.2 Ground-state wavefunction squared of quantum harmonic oscillator for two discretization errors . . . . . 73
4.3 Ground-state energy analysis of quantum harmonic oscillator . . . . . 76
4.4 methyl $\beta$-D-arabinofuranoside path integral molecular dynamics snapshot at 300K . . . . . 79
4.5 methyl $\beta$-D-arabinofuranoside C4-C5 dihedral distribution . . . . . 81
4.6 methyl $\beta$-D-arabinofuranoside point-charge dipole-dipole Kubo correlation function . . . . . 83

# List of Abbreviations

| | |
|---|---|
| PIMD | Path Integral Molecular Dynamics |
| PIMC | Path Integral Monte Carlo |
| PIGS-MD | Path Integral Ground State Molecular Dynamics |
| RPMD | Ring Polymer Molecular Dynamics |
| PILE | Path Integral Langevin Equation |
| WNLE | White Noise Langevin Equation |
| MMTK | Molecular Modelling Toolkit |

# Chapter 1

# Introduction

Since the dawn of numerical computation, scientists have been utilizing computational resources to model the profound complexities of nature. As technological discoveries lead to improvements in computing architecture, so does our mathematical grasp of simulating these natural models. In the field of chemistry as early as 1943, some twenty years after Schrödinger published the first paper on the theory of quantum mechanics [1], the first computational chemistry papers were being published [2] including the first chemistry application of the Monte Carlo method [3]. Thirty years after, similar computational methods are being used for microsecond-scale atomistic-resolution simulations of physiological significance [4]. Today's computational resources also allow scientists to probe small molecular systems with unprecedented accuracy, aiming to answer questions such as the

origin of the superfluid state of matter [5]. In both large and small molecular systems, there are deep scientific questions to be answered.

It is the aim of this research to address the challenges of simulating small molecular systems with varying degrees of quantum mechanical treatment. Using a specialized mathematical formalism, the incorporation of quantum effects can be adjusted in a systematic manner. The aim of this work is to achieve three primary objectives. The first is to benchmark the new implementations of these quantum mechanical simulation methods in an open-source molecular dynamics package. The second objective is to report on the computational requirements for the adequate and efficient quantum mechanical treatment of a number of interesting scientific systems, primarily that of doped helium clusters and a biomolecule. Lastly, this works aims to provide scientific insight from the equilibrium and dynamic measurements calculated for these systems.

To frame this research properly, we will discuss the mathematical principles behind simulating the behaviour of any molecule with standard molecular dynamics algorithms. The shortcomings of this method at low temperatures will then be addressed with the path integral molecular dynamics isomorphism.

## 1.1 Molecular Dynamics

Molecular dynamics is a computational tool used to examine many-body systems with atomic resolution. This technique is frequently used in the field of computational chemistry to obtain atomic trajectories from which one may extract properties comparable to experimental observables. Alternative methods such as Monte Carlo sampling may be used to calculate similar properties based on acceptance/rejection of random atomic displacements. What distinguishes molecular dynamics from other methods to sample atomic positions and momenta is that it physically guides the exploration of phase space using initial conditions and the force calculated by Newton's second law of motion, $F_i = m_i a_i$, where $m_i$ is the mass and $a_i$ is the acceleration for the $i$th atom in a system. Given a system containing many atoms, a set of coupled differential equations that exactly represents the system is unlikely to have an analytic solution. However, once equations of motion are formulated, numerical integrators can be used to propagate atoms in time by applying forces where the accuracy of the simulation can be improved by decreasing the time step, $\Delta t$.

By introducing a few concepts from statistical mechanics, one may understand how molecular dynamics can be used to estimate ensemble averages [6, 7]. For an $N$-particle system subject to only interparticle interactions, one can formulate Newton's equations of

motion in terms of a Hamiltonian consisting of a kinetic and a potential operator,

$$\hat{H} = \hat{K} + \hat{V} = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + V(q_1, \ldots, q_N). \tag{1.1}$$

Equations of motion can be obtained by differentiating the Hamiltonian $\hat{H}$ with respect to the phase-space variables, $p^N$ and $q^N$. In this thesis we describe the phase space variables of momentum and position as a vector of component momenta and positions, respectively for $N$ particles,

$$\begin{aligned} p^N &= (p_1, \ldots, p_N) \\ q^N &= (q_1, \ldots, q_N) \end{aligned} \tag{1.2}$$

A beneficial aspect to this formulation is that equations of motion conserve the Hamiltonian, allowing a constant value to be monitored through the simulation to ensure numerical error is not accumulating. This concept of conservation allows one to sample microstates on a particular phase-space surface, where in the current case, one samples from the micro-canonical statistical ensemble. To reproduce experimental results from a simulation, one is often concerned with maintaining a constant temperature by sampling from the canonical ensemble. By obtaining positions and momenta from this ensemble, one can calculate thermal expectation values for any physical observable $\hat{A}$. In quantum mechanical terms,

this thermal average can be obtained by computing

$$\langle A \rangle = \frac{\sum_i \exp[-E_i/k_B T]\langle i|A|i\rangle}{\sum_j \exp[-E_j/k_B T]}, \tag{1.3}$$

where $\langle i|A|i\rangle$ denotes the expectation value of the operator $A$ in the energy eigenstate $|i\rangle$ with energy $E_i$. In Dirac's notation, the members of this set of states and energies are related by the Schrödinger equation,

$$\hat{H}|i\rangle = E_i|i\rangle. \tag{1.4}$$

In practice, obtaining the energy eigenstates for a many-body system is computationally quite difficult. However, molecular dynamics as a computational tool may be used readily to calculate thermal expectation values. To illustrate how this can be done, Eq. 1.3 can be rewritten in the classical limit. Noting that $\exp[-E_i/k_B T] = \langle i|\exp[-\hat{H}/k_B T]|i\rangle$, the expectation value can be written independent of the energy basis as

$$\langle A \rangle = \frac{Tr\left[\exp[-\beta\hat{H}]\hat{A}\right]}{Tr\left[\exp[-\beta\hat{H}]\right]}, \tag{1.5}$$

where $\beta = 1/k_B T$. It would be convenient to split this exponential term and calculate this average in a basis such as the position or momentum eigenfunctions, but since the

Hamiltonian is the sum non-commuting operators, one must still consider the commutator term,

$$\exp[-\beta\hat{K}]\exp[-\beta\hat{V}] = \exp\left[-\beta(\hat{K}+\hat{V}+\mathcal{O}([\hat{K},\hat{V}]))\right]. \tag{1.6}$$

Given that the potential $V$ is a function of operator $\hat{q}$, the $\mathcal{O}([\hat{K},\hat{V}]) = \mathcal{O}([\hat{p}^2, V(\hat{q})])$ commutator is of order $\hbar$ by using canonical commutation relation. For a classical system in the limit of $\hbar \to 0$ this commutator term can be neglected, and each trace in Eq. 1.5 can be expressed by inserting a complete set of eigenstates of position and momenta as such,

$$Tr\left[\exp[-\beta\hat{H}\right] \approx \sum_{p,q}\langle q|\exp[-\beta\hat{V}]|q\rangle\langle q|p\rangle\langle p|\exp[-\beta\hat{K}]|p\rangle\langle p|q\rangle. \tag{1.7}$$

Evaluating these matrix elements, one obtains the following for a 3-dimensional system of $N$ particles, otherwise known as the classical partition function,

$$Tr\left[\exp[-\beta\hat{H}\right] \approx \frac{1}{h^{3N}N!}\int dp^N \int dq^N \exp\left[-\beta[\sum_i p_i^2/(2m_i) + V(q^N)]\right], \tag{1.8}$$

where $N!$ accounts for the indistinguishability of identical particles. Using this result, one

can rewrite the thermal expectation value in Eq. 1.3 for an observable $A$,

$$\langle A \rangle = \frac{\int dp^N dq^N \exp\left[-\beta[\sum_i p_i^2/(2m_i) + V(q^N)]\right] A(p^N, q^N)}{\int dp dq \exp\left[-\beta[\sum_i p_i^2/(2m_i) + V(q^N)]\right]} \tag{1.9}$$

Under what is referred to as the *ergodic hypothesis*, the simulation of a system obeying Hamilton's equations for infinite time allows for a complete exploration of position-momenta phase space, and is equivalent to a statistical ensemble average,

$$\langle A \rangle = \lim_{t\to\infty} \frac{1}{t} \int_0^t dt' A(p(t'), q(t')). \tag{1.10}$$

When taking into consideration fundamentally non-ergodic systems such as glasses or nearly harmonic solids, the ergodic hypothesis is a fundamental principle of molecular dynamics. In practice, care must be taken to ensure that phase-space is properly sampled for any system as there is no general technique that can be used to test if ergodicity is satisfied. Given this hypothesis and our knowledge of thermal expectation value calculation in Eq. 1.9, one can use molecular dynamics to compute these properties. However, when simulating low-temperature systems where it cannot be assumed that $\hbar \to 0$, a different theory must be utilized.

## 1.2 Path integral Formulation of the Partition Function

Although the previous expression for the ensemble average reduces the quantum picture to a classical one, it is useful to remain general and consider the case where $\hbar \neq 0$. The following derivation is considered in one dimension for simplicity in notation. One can begin with the expression for quantum mechanical partition function [8]. This is obtained by calculating the trace of the canonical density matrix, as in Eq. 1.7,

$$Z = Tr\left[e^{-\beta\hat{H}}\right] \tag{1.11}$$

When one evaluates this trace in the continuous coordinate basis $|q\rangle$, it can be calculated as an integral. In general, the kinetic and potential operators do not commute, so this exponential cannot be evaluated,

$$Z = \int dq \langle q|e^{-\beta(\hat{K}+\hat{V})}|q\rangle. \tag{1.12}$$

In a classical formulation, the commutator term was reasoned to go to zero, whereas in a quantum treatment, the Trotter theorem can be used [9]. This theorem states that for

any two operators, $\hat{A}$ and $\hat{B}$, that do not commute, the following expression is satisfied:

$$e^{\lambda(\hat{A}+\hat{B})} = \lim_{P\to\infty}\left[e^{\frac{\lambda}{2P}\hat{B}}e^{\frac{\lambda}{P}\hat{A}}e^{\frac{\lambda}{2P}\hat{B}}\right]^P. \tag{1.13}$$

Upon defining an operator $\Omega$ based on a Trotter factorization, the partition function can be written as

$$\Omega = e^{\frac{-\beta}{2P}\hat{V}}e^{\frac{-\beta}{P}\hat{K}}e^{\frac{-\beta}{2P}\hat{V}} \tag{1.14}$$

$$Z = \lim_{P\to\infty}\int dq\langle q|\Omega^P|q\rangle. \tag{1.15}$$

This product of $P$ factors can be expressed more simply by inserting $P-1$ resolution of the identity operators into the coordinate basis between each pair of factors. If one relabels the integration variables sequentially, the partition function becomes

$$Z = \lim_{P\to\infty}\int dq_1\ldots dq_P\left[\prod_{i=1}^{P}\langle q_i|\Omega|q_{i+1}\rangle\right]_{q_{P+1}=q_1} \tag{1.16}$$

To evaluate these coordinate space matrix elements, we may first evaluate the potential operator terms. Since the potential terms are functions of q, they act directly on the coordinate eigenstates, allowing the corresponding eigenvalues to be pulled out,

$$\langle q_i|\Omega|q_{i+1}\rangle = e^{\frac{-\beta}{2P}\hat{V}(q_i)}\langle q_i|e^{\frac{-\beta p^2}{2Pm}}|q_{i+1}\rangle e^{\frac{-\beta}{2P}\hat{V}(q_{i+1})} \tag{1.17}$$

The resulting matrix element of the kinetic energy operator can be evaluated by inserting a complete set of momentum eigenstates, utilizing the inner product of coordinate and momentum states, which results in the term

$$\langle q_i|e^{\frac{-\beta p^2}{2Pm}}|q_{i+1}\rangle = \left(\frac{mP}{2\pi\beta\hbar^2}\right)^{1/2} e^{-\frac{mP}{2\beta\hbar^2}(q_{i+1}-q_i)^2} \tag{1.18}$$

Using these coordinate-space matrix elements, the partition function for a 3-dimensional system can be written concisely fas

$$\begin{aligned} Z &= \lim_{P\to\infty}\left(\frac{mP}{2\pi\beta\hbar^2}\right)^{P/2}\int dq_1\ldots dq_P \\ &\times \exp\left\{-\sum_{i=1}^{P}\left[-\frac{mP}{2\beta\hbar^2}(q_{i+1}-q_i)^2+\frac{\beta}{P}V(q_i)\right]\right\}_{q_{P+1}=q_1} \end{aligned} \tag{1.19}$$

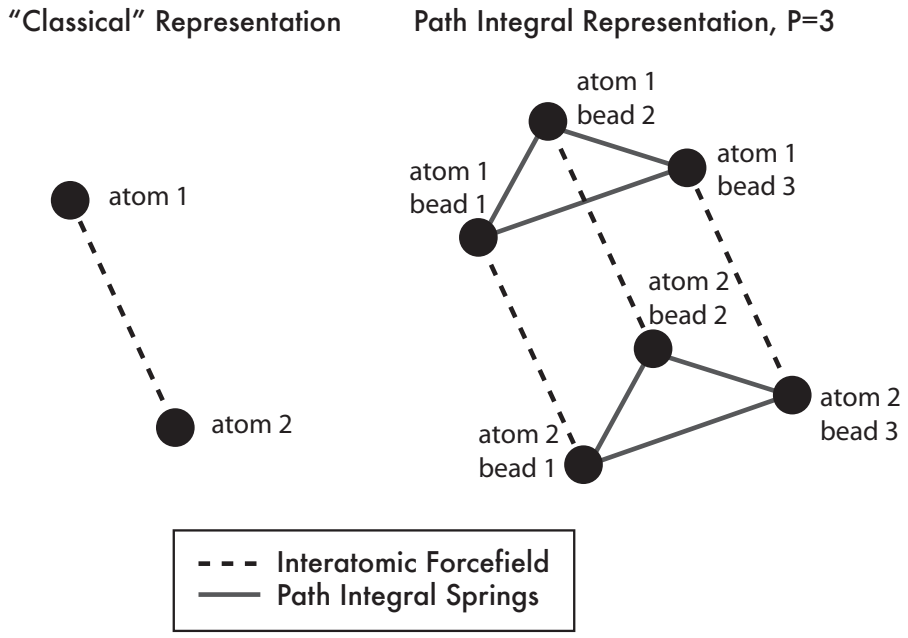


Figure 1.1: A representation of two atoms interacting with both a classical (left) and quantum path integral representation (right). In the quantum representation, the interatomic forcefield acts with a factor 1/P between all sets of beads, and path integral springs couple the beads of a single atom in a cyclic manner.

## 1.3 Path integral Molecular Dynamics

The next step after obtaining the discretized quantum partition function in Eq. 1.19 is to develop a methodology to sample states from this partition function. If one defines an effective potential,

$$V_{eff}(q_i, \ldots, q_P) = \sum_{i=1}^{P} \left[ -\frac{mP}{2\beta\hbar^2}(q_{i+1} - q_i)^2 + \frac{\beta}{P} V(q_i) \right]_{q_{P+1}=q_1} \tag{1.20}$$

the partition function can be rewritten in the form of a classical partition function for a $P$-particle system, where all particles are connected in a cycle by a harmonic interaction. In principle, this representation of the partition function is exact in the limit of infinite points P, henceforth referred to as beads, but in practice we simply truncate this integral while incurring a systematic error, resulting in an approximate partition function $Z_p$,

$$Z_p = \left( \frac{mP}{2\pi\beta\hbar^2} \right)^{P/2} \int dq_1 \ldots dq_P \exp\left( -\beta V_{eff}(q_i, \ldots, q_P) \right) \tag{1.21}$$

As was introduced by Chandler and Wolynes [10], one could refer to the path integral of a single quantum particle as being isomorphic to a classical system of $P$ harmonically connected beads with frequency proportional to both $P$, the number of beads, and $T$, the temperature of our simulation. This is depicted for the interaction of two atoms in Fig. 1.1. In this picture, each bead is subject to an effective potential that acts on it independently. Although path conformations can be readily sampled from this partition function using path integral Monte Carlo (PIMC) sampling, an additional modification must be made to do molecular dynamics. By supplementing the partition function with a set of $P$ Gaussian integrals,

$$Z_p = N \int dp_1 \dots dp_P \int dq_1 \dots dq_P \exp\left(-\beta \left[\sum_{i=1}^{P} \frac{p_i^2}{2m_i} + V_{eff}(q_i, \dots, q_P)\right]\right), \qquad (1.22)$$

one can include a fictitious classical momenta with the effective potential. The effect of these variables is negligible as they only contribute a constant factor from integration that can be chosen to reproduce the prefactor in Eq. 1.21. These fictitious momenta have an adjustable parameter $m_i$ of units mass that can be arbitrarily selected. This value of $m_i$ is chosen as the full mass of each atom for all path integral molecular dynamics simulations. This modified partition function amounts to a final fictitious classical Hamiltonian that is isomorphic to a quantum system from which we can obtain fully classical equations of motion,

$$H = \sum_{i=1}^{P} \left[\frac{p_i^2}{2m_i} + \frac{mP}{\beta^2\hbar^2}(q_{i+1} - q_i)^2 + \frac{1}{P}V(q_i)\right]_{q_{P+1}=q_1} \qquad (1.23)$$

To calculate thermal expectation values with quantum effects, one may now use the classical formulation of the thermal expectation value in Eq. 1.9 along with the equations of motion derived from the Hamiltonian above.

### 1.3.1 Normal Mode Path Integral Molecular Dynamics

This isomorphism is exact in the limit of $P \to \infty$, but several problems arise in numerical implementation with standard mathematical treatment. These inherent problems with path integral molecular dynamics were first addressed in the literature by Hall et al. [11]. Most significantly, as $P$ increases, the stiff harmonic springs, the second term in Eq. 1.23, dominate the interactions in the system. This may result in non-ergodic exploration of phase space or poor equipartition of energy in the case of a constant temperature simulation.

As the integration variables in Eq 1.22 are arbitrary (a trace is independent of a basis), it is beneficial to transform our the cartesian coordinates and masses to uncouple the quadratic term. Well cited transformations include the staging transformation[12] and the normal mode transformation [13, 14, 15]. For the purposes of this work, the normal mode transformation is utilized. This method will be presented in a manner similar to Ceriotti et al. [16]. One can perform this transformation by performing a Fourier expansion of the discrete cyclic path, resulting in the following transformation of both position and momenta,

$$\tilde{v}_i^k \leftarrow \sum_{j=1}^{P} v_i^j C_{jk}, \; \tilde{q}_i^k \leftarrow \sum_{j=1}^{P} q_i^j C_{jk} \tag{1.24}$$

where the coefficients $C_{jk}$ are defined by

$$C_{jk} = \begin{cases} \sqrt{1/P} & k = 0 \\ \sqrt{2/P}\cos 2\pi jk/P & 1 \leq k \leq P/2 - 1 \\ \sqrt{1/P}(-1)^j & k = P/2 \\ \sqrt{2/P}\sin 2\pi jk/P & P/2 + 1 \leq k \leq P - 1 \end{cases} .$$

This normal mode calculation has the benefit that the zeroth normal mode is the centroid of the path. This centroid variable will be used in Chapter 3 in order to calculate quantum dynamics, so it is convenient that it is already calculated by using this transformation. After converting to normal mode variables, we have an analogous expression to Eq. 1.23 that can be used to derive new normal mode equations of motion without a harmonic difference term,

$$H_{NM}(p,q) = \sum_{i=1}^{N}\sum_{k=0}^{P-1}\left(\frac{[p_i^k]^2}{2Pm_i} + \frac{Pm_i sin^2(k\pi/P)}{2\hbar^2\beta^2}[\tilde{q}_i^k]^2\right) \tag{1.25}$$

In the Liouville operator formalism, without going into depth, the Hamiltonian can be split into both a free ring path-integral propagation and a potential energy part, while incurring minimal error [17, 18]. This is done using the Trotter theorem as was done in the previous section. This splitting offers a very useful step-by-step method for performing

normal mode PIMD with a simple extension to constant temperature sampling.

The Velocity-Verlet algorithm is a popular tool for integrating equations of motion for molecular dynamics [19]. In terms of velocity $v$ and coordinate $q$, the following steps indicate the actual implementation for normal mode PIMD in the Molecular Modelling Toolkit (MMTK) integrator for a time-step $\Delta t$ with particles as was presented by Ceriotti et al. [16],

$$v_i^j \leftarrow v_i^j - \frac{\Delta t}{2m_i}\frac{dV(q_1^j, ..., q_N^j)}{dq_i^j} \tag{1.26}$$

$$\tilde{v}_i^k \leftarrow \sum_{j=1}^{P} v_i^j C_{jk},\ \tilde{q}_i^k \leftarrow \sum_{j=1}^{P} q_i^j C_{jk} \tag{1.27}$$

$$\begin{pmatrix} \tilde{v}_i^k \\ \\ \tilde{q}_i^k \end{pmatrix} \leftarrow \begin{pmatrix} cos(P\Delta t/\beta\hbar) & [-P/\beta\hbar]sin(P\Delta t/\beta\hbar) \\ \\ [\beta\hbar/P]sin(P\Delta t/\beta\hbar) & cos(P\Delta t/\beta\hbar) \end{pmatrix} \begin{pmatrix} \tilde{v}_i^k \\ \\ \tilde{q}_i^k \end{pmatrix} \tag{1.28}$$

$$v_i^k \leftarrow \sum_{j=1}^{P} \tilde{v}_i^j C_{jk},\ q_i^k \leftarrow \sum_{j=1}^{P} \tilde{q}_i^j C_{jk} \tag{1.29}$$

$$v_i^j \leftarrow v_i^j - \frac{\Delta t}{2m_i}\frac{dV(q_1^j, ..., q_N^j)}{dq_i^j} \tag{1.30}$$

Descriptively, these equations update the velocities based on a force calculation from the derivative of the potential $V$ for a half time-step, then transforms all variables to normal mode coordinates. The algorithm then propagates the free path analytically in Eq. 1.28 then transforms the variables back to cartesian coordinates and updates the velocities for another half time-step.

Though the normal mode coordinates are useful for a microcanonical constant temperature simulation, an effective new method was recently presented by Ceriotti et al. [16] that exploits properties of these coordinates for constant temperature simulation. Stochastic thermostatting methods like the path integral Langevin equation will be addressed in greater detail in Chapter 2, but for now, let it be stated that only a slight modification of the algorithm above is all that is needed to perform efficient constant temperature PIMD. Using the thermostatting method, path integral molecular dynamics may be used to efficiently calculate observables of any molecular system.

## 1.4 Overview of Thesis

As a technical contribution by the author, PIMD in the normal mode coordinate representation with a path integral Langevin equation (PILE) thermostat along with cartesian PIMD using the white-noise Langevin equation were implemented in the open-source

python molecular dynamics package, Molecular Modelling Toolkit [20]. Prior to this work, no quantum simulation capabilities were supported in this software. This was done in a generalized manner to facilitate the use of both existing molecular modelling forcefields and custom-designed potentials such as the helium potentials used in the present work. Additional functionality in MMTK for trajectories, subspace manipulation, vibrational-modes calculations, and energy minimization remain available for use with path integral simulations.

In Chapter 2, I present benchmarks for this implementation by efficiently calculating equilibrium properties for doped helium clusters of different sizes. Computation of these properties required careful optimization of parameters, something rarely addressed in PIMD studies in the literature, providing an extremely useful resource for future PIMD studies.

Chapter 3 pertains to the calculation of approximations to real-time quantum dynamics using ring-polymer molecular dynamics (RPMD). We test the quality of RPMD for a model system and for low-temperature simulations of doped helium clusters. Low temperature dynamics calculations using RPMD has not previously been published.

In Chapter 4, more challenging applications of PIMD for the calculation of equilibrium and dynamic properties are presented. Path integral ground-state ($T = 0$) energies are shown for the harmonic oscillator, as well as a challenging biomolecular simulation including

quantum effects.

Finally, in Chapter 5, the overall conclusions of this research are presented. Numerous future studies are proposed which build on this work.

# Chapter 2

# Equilibrium properties for doped helium clusters

Clusters and droplets of light chemical species like helium and molecular hydrogen have unique physical properties. Due to their weak molecular interactions and low mass, delocalized quantum effects are prevalent at finite temperature and the ground state. Finite clusters of these atoms have also exhibited superfluid behavior at low temperatures [5]. Recent theoretical and experimental work has supported the discovery of molecular superfluids which persist at large cluster sizes. In the case of simulating low temperature helium in the bulk phase and in clusters, quantum Monte Carlo, in particular, the path integral Monte Carlo (PIMC) approach is most frequently used [21].

We present an implementation of the path integral molecular dynamics (PIMD) approach with a path integral Langevin equation (PILE) thermostat, [22] with applications to sampling properties in the ultra low temperature regime, in particular, the study of doped helium clusters. In this work we will test the effectiveness of this approach and confirm that our results are consistent with previously published theoretical [23] and, to a certain extent, experimental [24] results. In the use of the path integral formalism at low temperatures, efficient sampling is of great importance. We report a path integral error convergence study using a centroid virial energy estimator with a systematic method to obtain simulation parameters for efficient simulation. We present structural properties, the chemical potential and vibrational shifts of doped helium clusters of sizes ranging from $N = 1 - 20$. All the equilibrium PIMD results presented in this work neglect the rotational effects of the dopant molecule as well as bosonic exchange. Using the sampling technique discussed here, one can gain insight into the size dependence of cluster observables. The effective use of these techniques for small clusters, may provide a foundation for insight into the bulk properties of larger droplets. It is our hope that the method presented in this thesis will serve as a model for future low temperature path integral simulations on related systems. Miura and co-workers [25, 26, 27, 28, 29] have contributed to the development of PIMD techniques for low temperature and ground state doped helium clusters. Mizumoto and Ohtsuki [30] have applied similar ideas to study doped hydrogen clusters.

To our knowledge, however, the PILE approach has not yet been applied to these low temperature weakly-bound systems. The work of this chapter has been submitted to the Journal of Chemical Physics (Manuscript A11.07.0309).

## 2.1 Theoretical Approach

### 2.1.1 Energy Evaluators

To correctly measure observables for path integral simulations, corresponding estimators must be used. A primitive or Barker estimator [31] for the internal energy $E$ in three dimensions is derived by differentiating the partition function to obtain,

$$E_T = \frac{3NP}{2\beta} - \frac{Pm}{2\hbar^2\beta^2}\sum_{j=1}^{P}(r_j - r_{j+1})^2 + \frac{1}{P}\sum_{j=1}^{P}V_{CL}(r_j), \tag{2.1}$$

where $r_j$ is a vector of size 3N representing the cartesian position of an atom, where N is the total number of atoms in our system, and the squared term is interpreted as a dot product. $V_{CL}(r_j)$ is the classical potential and the first two terms estimate the kinetic energy. This estimator can be described qualitatively as the kinetic energy contribution of $\frac{3}{2\beta}$ per bead and subtracting the energy introduced by the path integral springs. However, the primitive energy estimator is problematic due to the increase of the mean square fluctuations with

number of beads proportional to $\frac{P}{2\beta^2}$. Conveniently, an estimator with better variance properties, called the virial estimator, can be reformulated with the virial theorem [21],

$$E_{CV} = \frac{3N}{2\beta} - \frac{1}{2P}\sum_{j=1}^{P}(r_j - \bar{r}) \cdot V'(r_j) + \frac{1}{P}\sum_{j=1}^{P} V_{CL}(r_j) \tag{2.2}$$

where the reference point is selected as the centroid of the path, $\bar{r} = \frac{1}{P}\sum_{j=1}^{P} r_j$, and N is the number of atoms. Energies calculated using path integral molecular dynamics are reported using either of the two estimators described above and exclusively the primitive estimator for PIMC.

### 2.1.2 Thermostats

Obtaining quantum mechanical properties using path integral simulations comes at a computational expense. At low temperatures, when one must include hundreds of additional beads to a simulation, high frequency path integral modes make it challenging to sample the slower degrees of freedom of the underlying system. Recent advances in the use of the stochastic Langevin equation thermostat have been published with the purpose of resolving these issues [22].

Langevin dynamics is an integration method that serves to ensure canonical sampling [32, 33]. This thermostat works by pairing the strength of a Gaussian distributed random

force with a velocity dependant frictional term in the equations of motion of our system. A relationship between the friction and stochastic fluctuations is enforced such that the equipartition theorem is satisfied for each degree of freedom on average, thus resulting in a constant temperature simulation. Once the characteristic timescale of the underlying system is established, a friction value is chosen so as not to disturb the characteristic motions of the system.

Langevin Dynamics is suitable for path integrals using a standard white-noise Langevin equation (WNLE) [17]. In the cartesian representation of path integrals, this thermostat is implemented by generating a Gaussian random number, $\xi_j$, for each $j = 1...P$ bead, each of mass $m_j$. A stochastic term is then calculated and added to the system velocities with a frictional parameter $\gamma$. In a Liouville operator formalism, this amounts to an additional velocity update before and after propagating atoms in position by a timestep $\Delta t$, shown here after a Trotter splitting for the velocity of bead j, [17, 18]

$$v_j = e^{(-\Delta t/2)\gamma} v_j + \sqrt{\frac{P}{\beta m_j}} \sqrt{1 - [e^{(-\Delta t)\gamma}]^2} \xi_j \tag{2.3}$$

Note that this step can only be made in normal mode coordinates, so one must first transform into this form and back to cartesian coordinates afterwards during each molecular dynamics loop. If one refers back to the normal-mode PIMD algorithm presented in

Chapter 1, this stochastic velocity change occurs before Eq. 1.26 and after the last step in Eq. 1.30. This thermostat stands to be the simplest method for constant temperature path integral simulations, but even still, additional optimizations can be made using the information known about the harmonic behaviour of a free path integral ring polymer. As presented by Ceriotti *et al.* [22], in normal mode coordinates, the internal modes of a path integral atom can be determined analytically and thus thermostatted explicitly. In particular, the Langevin frictional values applied to the path integral springs can be chosen to minimize the autocorrelation of the Hamiltonian. This works for all path integral modes with the exception of the friction on the centroid of the path, which must be determined through a sample calculation. In practice, multiple path integral rings are interacting so it remains to be shown if this approximation is valid, especially at low temperatures. This stands to be a highly effective method of thermostatting since the high frequency path integral springs make it very difficult to thermostat with a single friction value using white-noise Langevin dynamics.

## 2.2 Computational Methods

All simulations and analysis are performed with the MMTK package [20] with autocorrelations calculated using the Fast Correlation Algorithm implemented in the nMOLDYN

analysis package [34].

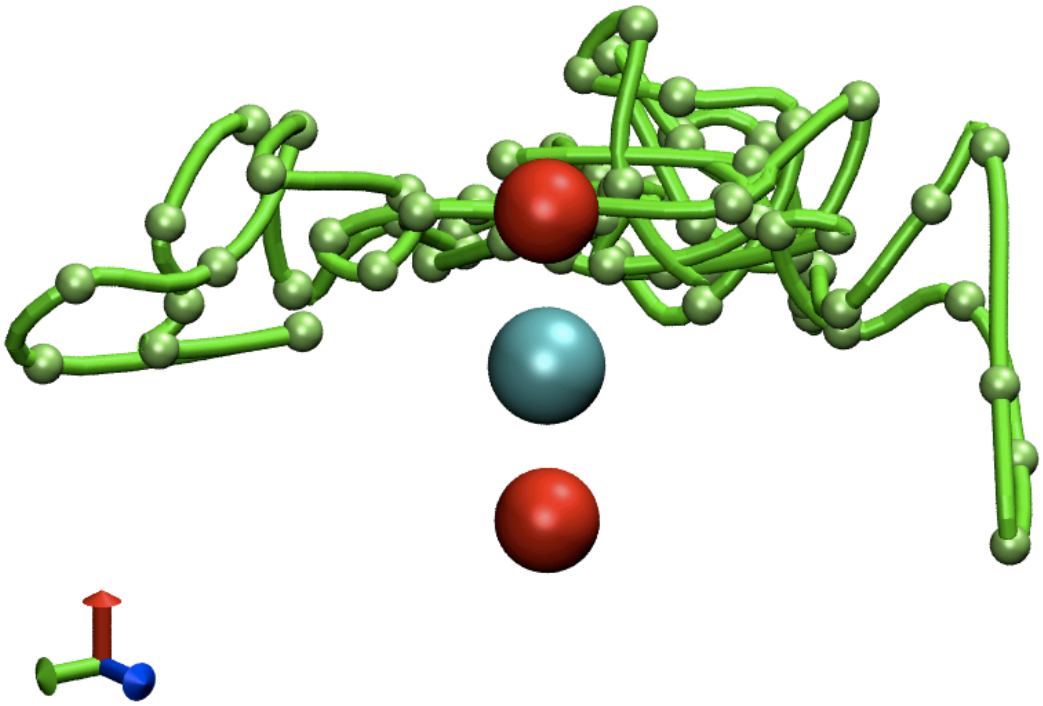

Figure 2.1: Isometric snapshot of a molecular dynamics trajectory consisting of a P=64 helium atom about a P=1 He-$CO_2$ impurity at T=0.37K. $CO_2$ is depicted in full molecular form, as opposed to a point particle, for the purpose of visualization. Image generated using Visual Molecular Dynamics [35].

### 2.2.1 Simulation Parameters

All simulations are performed at the experimental temperature of 0.37K. PIMD and PIMC results neglect the contribution of bosonic exchange in quantum statistics. Both are sim-

ulated with an identical number of path integral beads. All MD and MC simulations are then performed with a point $CO_2$ particle aligned about the x-axis with a 2D anisotropic He-$CO_2$ potential (Jacobi coordinates $r$ and $\theta$), [23] thus neglecting rotational degrees of freedom of the $CO_2$ molecule. A molecular dynamics snapshot depicts this for a He-$CO_2$ system in Fig. 2.1. This potential is obtained by fitting high-quality ab initio data points to 2D Morse/Long-Range (MLR) functions [36]. The analytical MLR potential form is obtained from the fitting program betaFIT[37] and is based on the Aziz He-He potential [38]. A smooth analytical form is necessary due to a discontinuity in the derivative of this potential which could result in problematic forces during molecular dynamics. In our case, the Helium pair potential is fitted to the 1D MLR form below,

$$V_{\mathrm{MLR}}(r) = D_e\left\{1 - \frac{u_{\mathrm{LR}}(r)}{u_{\mathrm{LR}}(r_e)}e^{-\beta(r)\cdot y_p^{\mathrm{eq}}(r)}\right\}^2, \tag{2.4}$$

in which $D_e$ is the potential well depth, $r_e$ is the equilibrium internuclear distance, $\beta(r)$ is a function controlling the shape of potential, and $u_{\mathrm{LR}}(r_e)$ is the value of the potential in the long-range tail at equilibrium position $r_e$. The long-range tail, $u_{\mathrm{LR}}$, is given by equation,

$$u_{\mathrm{LR}}(r) = \frac{C_{m_1}}{r^{m_1}} + \frac{C_{m_2}}{r^{m_2}} + \cdots, \tag{2.5}$$

where $C_{m_i}$ is a long range coefficient. Eq.(2.4) is further modified by adding a $V_{min}$ term which corresponds to the minimum position of the potential. The exponential term $\beta(r)$ in Eq.(2.4) is written in

$$\beta(r) = \beta_{\mathrm{MLR}}(y_p^{\mathrm{ref}}(r)) = y_p^{\mathrm{ref}}(r)\beta_\infty + \left[1 - y_p^{\mathrm{ref}}(r)\right] \sum_{i=0}^{N} \beta_i y_q^{\mathrm{ref}}(r)^i, \tag{2.6}$$

in which

$$y_p^{\mathrm{eq}}(r) = \frac{r^p - r_e^p}{r^p + r_e^p} \tag{2.7}$$

and

$$y_p^{\mathrm{ref}}(r) = \frac{r^p - r_{ref}^p}{r^p + r_{ref}^p}. \tag{2.8}$$

A detailed discussion of the parameters $p$ and $q$ can be found in the fitting manual [37]. The parameter set chosen for the helium pair potential is presented in Table 2.1. Cartesian force components are calculated by simply taking the derivative of the analytical potential form and using the chain rule. For the He-$CO_2$ potential, we used the same 2D MLR form as in Ref. [36].

To determine the time-step of our simulation, a vibrational modes calculation is performed on the entire system of all atoms in the path integral representation after a short microcanonical initialization. By selecting a timestep of 0.05 multiplied by the timescale

Table 2.1: Parameters of the MLR fit of the helium dimer potential.

| | | | |
|---|---|---|---|
| $\beta_0$ | -0.24721 | $\beta_5$ | 1.36044 |
| $\beta_1$ | 0.2109 | $\beta_6$ | 0.73 |
| $\beta_2$ | -0.3444 | $\beta_7$ | -1.1 |
| $\beta_3$ | 0.336 | $\beta_8$ | -2.2063 |
| $\beta_4$ | 0.02 | $\beta_\infty$ | 0.031316 |
| $C_6/(\mathrm{cm}^{-1}\mathrm{\AA}^6)$ | $7.04108 \times 10^3$ | $C_8/(\mathrm{cm}^{-1}\mathrm{\AA}^8)$ | $1.90422 \times 10^4$ |
| $C_{10}/(\mathrm{cm}^{-1}\mathrm{\AA}^{10})$ | $6.93473 \times 10^4$ | | |
| $r_e$/ Å | 2.96830 | $D_e$ / $\mathrm{cm}^{-1}$ | 7.614810 |

of the fastest vibrational mode, we ensure sampling of all the path integral degrees of freedom as the number of beads increases. This time-step is verified by confirming the energy conservation of a microcanonical simulation.

When selecting a friction for use with Langevin dynamics, one must be careful to maintain constant energy (on average) while ensuring that the friction doesn't restrict the sampling of the classical potential.[33] Using the harmonic limit sampling efficiency convention of Ceriotti et al. [16], the white-noise thermostat is considered efficient at sampling frequencies decreasing linearly for two orders of magnitude above and below a selected $\gamma$ value. For a sample simulation of 896 beads in a He-$CO_2$ simulation, the vibrational spectrum of the system had frequencies ranging between 0.0484 $\mathrm{ps}^{-1}$ and 13.81 $\mathrm{ps}^{-1}$, or four orders of magnitude. As suggested by these guidelines, overall, the most efficient white-noise friction is the median frequency of the non-zero vibrational modes

of our system, $\gamma_{central}$. The relaxation times of the classical potential and centroid virial kinetic energy are calculated using $\gamma_{central}$ as well the white-noise Langevin friction using Müser's rule of thumb where $\gamma_{muser} = 0.01/dt$. This latter value has been used in studies of quartz by Müser [39].

For a $P = 896$ He-$CO_2$ white-noise Langevin simulation, the relaxation time of the classical potential and centroid virial kinetic energy were both in the 1-2 ps range. In an overdamped system one may observe longer classical potential relaxation times but faster centroid virial kinetic energy relaxation times. An efficient Langevin friction is obtained when the relaxation times of the centroid virial kinetic energy and the classical potential are comparable. A rigorous white-noise friction optimization is not performed for this system.

The path integral Langevin equation thermostat is parameterized with a centroid friction variable [22]. Following the guidelines proposed by Ceriotti *et al.*, a microcanonical (50ps) simulation is used to determine the average relaxation time, $\tau$, of the centroid of a path integral atom. This is achieved by computing the autocorrelation of the centroid structural properties of our test system. In the present case, this is the centroid pair distance from the $CO_2$ dopant to the helium atoms denoted as $r$ as well as the pair angle with respect to the x-axis,

$$R_{centroid}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T (r(t+\tau) - < r >)(r(t) - < r >)\, \mathrm{d}t \tag{2.9}$$

For all cluster sizes presented here, the microcanonical centroid relaxation times are less than 2 ps. With this input parameter, the canonical relaxation time for the observables is equal to or shorter than that of the white-noise Langevin equation for all observables measured, making it an acceptable choice of friction for low-temperature simulations.

### 2.2.2 Analysis Parameters

To minimize statistical error and conserve disk space when calculating equilibrium properties, the longest thermostat relaxation time of a measured observable is used as the number of data points to skip when generating a PIMD trajectory. The error bars on the averages presented in this thesis are computed using the standard error of the mean, $e = s/\sqrt{n}$ where $s$ is the standard deviation and $n$ is the number of statistically independent values in the data set. We expect all our trajectory data to be uncorrelated on average due to the step skipping procedure.

Simulation length is limited by the time taken to converge the structural and energetic properties of our system. This varied for each cluster size, particularly the ones with non-ergodic behavior. The length of the simulation is adjusted based on the time-step

and number of steps skipped to obtain 10,000 data points for all equilibrium properties, resulting in total simulation times ranging from 1 to 10 ns.

In order to compare to previously published theoretical results for energy calculations, PIMC simulations were performed without rotational and bosonic exchange effects. The PIMC code employed here has been used in a number of previous studies [40, 41, 42, 43, 44, 45, 23, 46, 47]. In all simulations, 2000 blocks are used with 3000 passes per block in a block averaging scheme. These long simulations are then truncated to obtain energetic values with comparable error bars to that of the centroid virial estimator used with PIMD. The move-size is adjusted to ensure a 0.3-0.5 acceptance ratio throughout the simulation.

### 2.2.3 Exact basis set calculation

The exact results reported in this thesis come from a basis set calculation. The motion of the $CO_2$-He complex is separated into the translation of the centre of mass of the whole complex and the other internal motions, i.e., the radial stretch between the two moieties and the end-over-end rotation of the whole complex. The centre of mass translation is trivial and only contributes $\frac{3}{2}K_BT$ to the total thermal energy of the system and not other properties. Since the $CO_2$ has been constrained from any rotation in the space-fixed frame, with its bond axis pointing in the space-fixed $Z$-axis direction, the internal motion can actually be described as the motion of a particle with mass $\mu$, which is the reduced

mass of the $CO_2$ and He pair, under the action of the $CO_2$-He potential.

The internal motion Hamiltonian operator reads

$$\hat{H} = -\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial R^2} + \frac{\hbar^2}{2\mu R^2}\hat{\vec{l}}^2 + V(R,\theta), \tag{2.10}$$

where the first term is the radial stretching kinetic operator, the second is the centrifugal kinetic energy, with $\hat{\vec{l}}$ being the angular momentum operator, and the third is the 2-dimensional $CO_2$-He potential. Since the potential does not depend on the azimuthal angle $\phi$ about the $CO_2$ axis, the projection of the angular momentum of the point of mass $\mu$ on the $CO_2$ axis is conserved and serves as a good quantum number. Also, the Hamiltonian is conserved with respect to the inversion with respect to the centre of mass and parity is a good quantum label too. With these symmetry properties, we employ the spherical harmonics $\{|lm\rangle\}$ and the sine discrete variable representation (DVR) basis [48] $\{|R_\alpha\rangle\}$ for the bases of the internal rotation and radial stretching motions. With the direct product basis $|R_\alpha\rangle|lm\rangle$, the Hamiltonian matrix elements read

$$\begin{aligned}\langle l'm'|\langle R_\beta|\hat{H}|R_\alpha\rangle|lm\rangle &= T^R_{\beta,\alpha}\delta_{ll'}\delta_{mm'} + \frac{\hbar^2}{2\mu R_\alpha^2} l(l+1)\,\delta_{\alpha,\beta}\delta_{ll'}\delta_{mm'} + \\ &\quad \delta_{\alpha,\beta}\delta_{mm'}\int_0^\pi \sin\theta \\ &\quad \times \int_0^{2\pi} d\phi V(R_\alpha,\theta)\, Y^*_{l',m}(\theta,\phi)\, Y_{l,m}(\theta,\phi). \end{aligned} \tag{2.11}$$

One more condition imposed on the third term in Eq. 2.11 is that $l$ and $l'$ must be both even or odd, in order to have the same parity. The radial stretching kinetic matrix elements $\hat{T}_{\beta,\alpha}$ are taken from Eq. A6 of Ref. [48]. Diagonalization of the Hamiltonian matrix gives us eigenstates $\{|n\rangle\}$ and eigenenergies $\{E_n\}$. The partition function $Z$ and any thermal property $\hat{O}$ can be readily calculated as

$$Z = \sum_n \exp\left(-\beta(E_n - E_0)\right); \tag{2.12}$$

$$\left\langle \hat{O} \right\rangle = \frac{1}{Z} \sum_n \langle n| \hat{O} |n\rangle \exp\left(-\beta(E_n - E_0)\right). \tag{2.13}$$

Considering the low temperature employed in our simulation, the two summations above are truncated to include only bound states, and the unbound states have contributions of less than $10^{-33}$ to the partition function. 50 radial DVR basis functions for $2.2a_0 < R < 20a_0$ and spherical harmonics with maximum $l = 25$ converge the eigen energies to $10^{-4}$ cm$^{-1}$. These two values determine the basis set size used in this work.

## 2.3 Results and Discussion

### 2.3.1 Energy Convergence

In any path integral simulation, systematic error is incurred by the propagator through Trotter factorization or pair decomposition. This error can be reduced by decreasing $\tau$, with the *exact* result being recovered in the limit of $\tau \to 0$ where $\tau = \beta/P$. Faster error convergence can be achieved by higher-order propagators. [49, 50] In the most primitive path integral approximation, error in energy evaluation is on the order of $O(\tau^2)$ at large values of P. Using a quadratic fit, it's possible to extrapolate energy as well as structural properties, when $\tau$ is at 0 [15, 40]. This has the benefit of significantly lowering the overall computational cost of simulating highly quantum systems to obtain accurate results. In the case of He-$CO_2$ at 0.37K, simulations with $P < 384$ are not consistent with the quadratic behaviour and thus neglected. In Fig. 2.2, we plot the centroid virial energy estimator for the He-$CO_2$ complex as a function of $\tau$ for both the path integral Langevin equation and white-noise Langevin equation thermostatting methods and compare the results to the primitive estimator energy values used for PIMC. All the raw data points are fitted with a linear regression using the expression [15],

$$< E >_\beta = a + b\tau^2. \tag{2.14}$$

The error on the regression coefficient $a$ is calculated using a 95% confidence interval in order to compare to the exact quantum mechanical value at $\tau = 0$.

Cartesian path integral molecular dynamics, which inherently leads to non-ergodic trajectories [11], is used to obtain the white-noise Langevin equation averages in Fig. 2.2. Since the same number of independent data points were used to obtain both the WNLE and PILE averages, one could attribute the deviation from the PILE and PIMC results to either the sampling inefficiency of Cartesian path integral molecular dynamics or an an inefficient choice of friction. We observe a closer match to the higher accuracy PIMC results with the PILE thermostat.

### 2.3.2 Center of Mass Motion

In all of our simulations, the He-$CO_2$ complex is free to move, contributing $(3N/2) \times K_B T \times P$ to the total energy. In the energy convergence plot, Fig. 2.2, the exact calculations were corrected based on this assumption. For our two particle system, the total energy is written as the sum of 3 centre of mass degrees of freedom and 3 rovibrational degrees of freedom,

$$< E_{total} >=< E_{COM} > + < E_{Rovib} > . \tag{2.15}$$

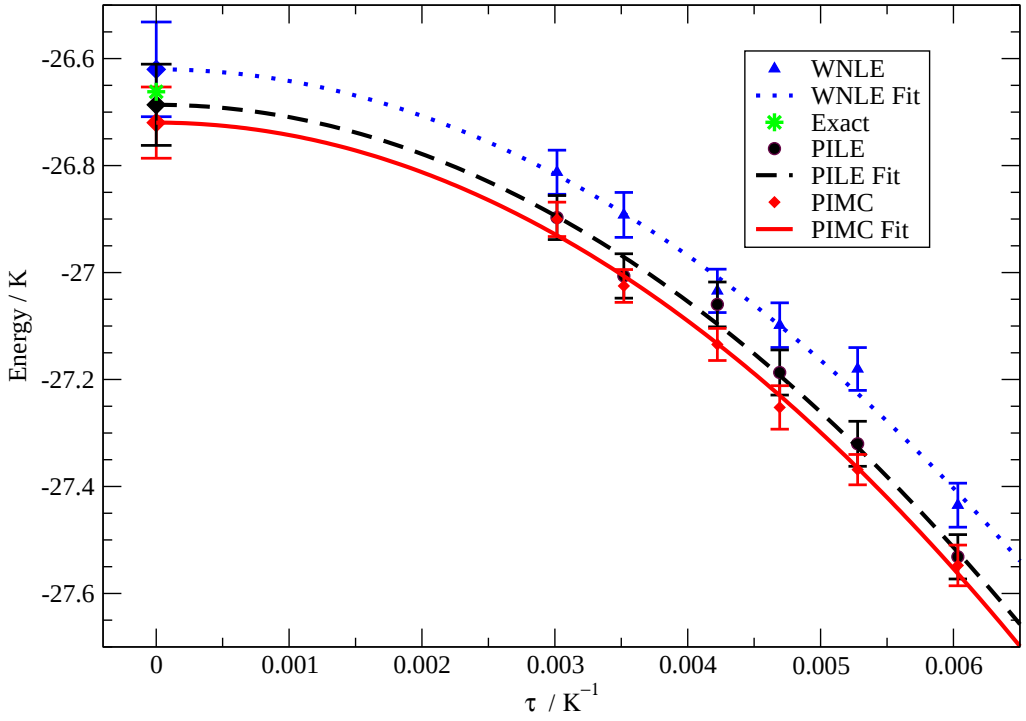


Figure 2.2: PIMD and PIMC (diamonds) results for the total energy of the He-$CO_2$ complex as a function of the imaginary time step $\tau = \beta/P$ at T=0.37K, where $\beta = 1/k_B T$ and the number of beads, P, varied from 256 to 896. Both PILE (circles) and WNLE (triangles) results are calculated for 10,000 data points. WNLE frictional value is selected at $\gamma_{central}$. Quadratic fits are performed using both PIMC and PILE, WNLE data represented by solid, dashed and dotted lines respectively. Extrapolated $\tau = 0$ data points are shown with error bars calculated with a 95% confidence interval. The exact value is represented by a star symbol (*).

The instantaneous centre of mass energy is calculated by transforming our atomic positions to centre of mass relative coordinates shown for the j'th set of beads with atom index $\alpha$,

$$R_j^{COM} = \frac{\sum_\alpha m_\alpha r_j^\alpha}{\sum_\alpha m_\alpha}. \tag{2.16}$$

A reformulated primitive energy estimator is used with the centre of mass coordinates to estimate the average centre of mass energy,

$$E_{COM} = \frac{3P}{2\beta} - \frac{P(m_{CO_2} + m_{He})}{2\hbar^2\beta^2} \sum_{j=1}^{P} (R_j^{COM} - R_{j+1}^{COM})^2. \tag{2.17}$$

In Fig. 2.3, we present the centre of mass energy distribution for simulations of He-$CO_2$ using the PILE and WNLE thermostats. We acknowledge the Cartesian integrator is used for the WNLE calculations, and cannot be directly compared to the normal mode integrator, but these shifted distributions may indicate that the WNLE thermostat at both frictional values insufficiently partitions the energy of our system to the centre of mass degrees of freedom. The PILE distribution of centre of mass energy is within error bars of the expected value of $(3/2) \times K_B T$. This result confirms that the PILE equations are exact for the free particle portion of the path integral action.

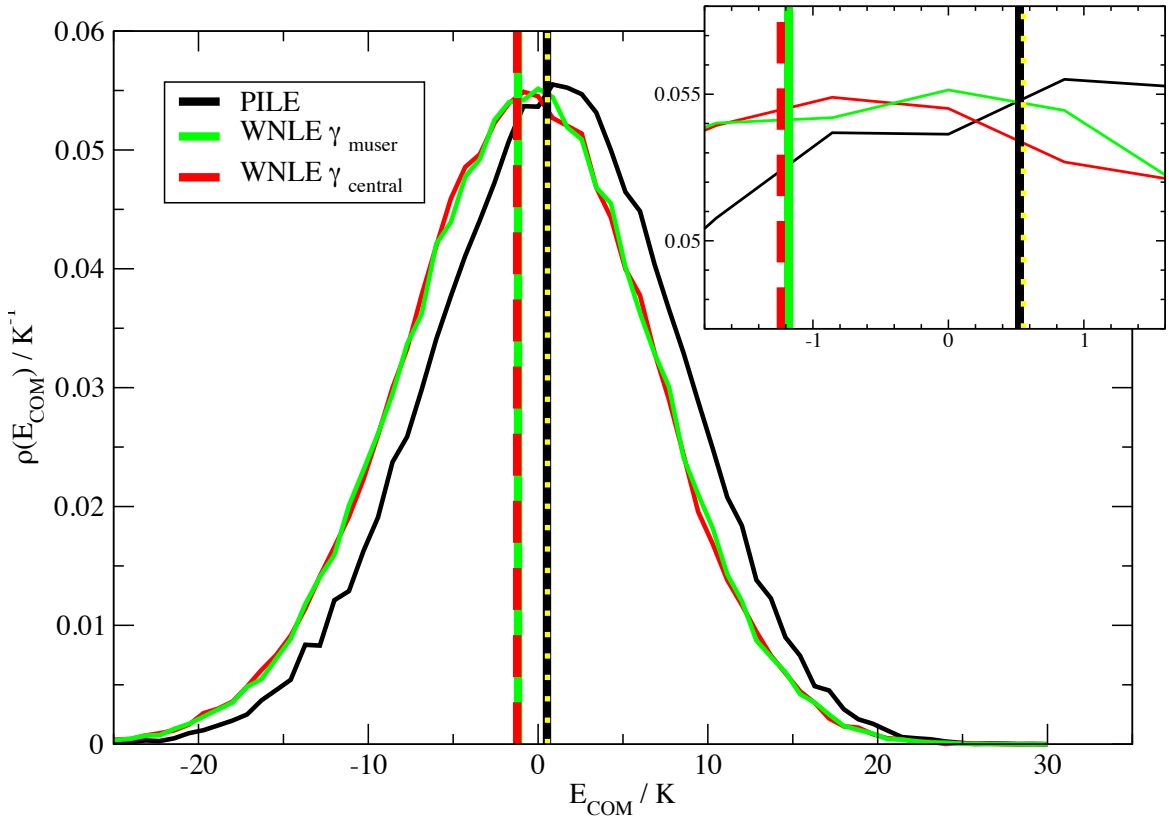


Figure 2.3: The distribution of centre of mass total energy for both PILE (black) and WNLE thermostats (red and green) at T=0.37K and P=256. From the equipartition theorem, the expected mean value of $(3/2) \times K_B T$ is indicated on the plot with a vertical dotted yellow line. The red, green, and black vertical lines indicate the mean value of the PILE, WNLE $\gamma_{muser}$, and WNLE $\gamma_{central}$ distributions respectively, where the width of each line indicates its standard error of the mean.

### 2.3.3 Structural Properties

The structural convergence of a He-$CO_2$ system is presented as a function of $\tau$ in Fig. 2.4 at T=0.37K. We observe that 512 beads are sufficient to reproduce the exact distribution obtained from a basis set calculation. In the following cluster size dependence studies, we expect a systematic error in the radial and angular distributions similar to those shown here for P=256 beads.

The error convergence of the average of the first and second moments of a He-$CO_2$ system are depicted in Fig. 2.5 as a function of $\tau$. These convergence plots depict a more rapid convergence of the $< r^2 >$ property.

The energetic change associated with increasing the number of helium atoms in a cluster can be studied by the chemical potential, defined as $\mu =< E_N > - < E_{N-1} >$ and depicted in Fig. 2.6 from $N = 1 - 20$. Qualitatively, the PIMD results match the PIMC results for non-rotating clusters with standard Boltzmann statistics. Some of the substantial energetic differences in the chemical potential can be explained by examining the evolution of structural probability densities as a function of cluster size. The averaged radial and angular density distribution of $N = 1 - 20$ helium atoms about a $CO_2$ impurity are presented in Fig. 2.7. For $N = 1$ to 5 helium atoms, our results support the existence of a ring-like structure at an average radius of $R \approx 3.3\mathring{A}$. A simulation of six helium atoms is

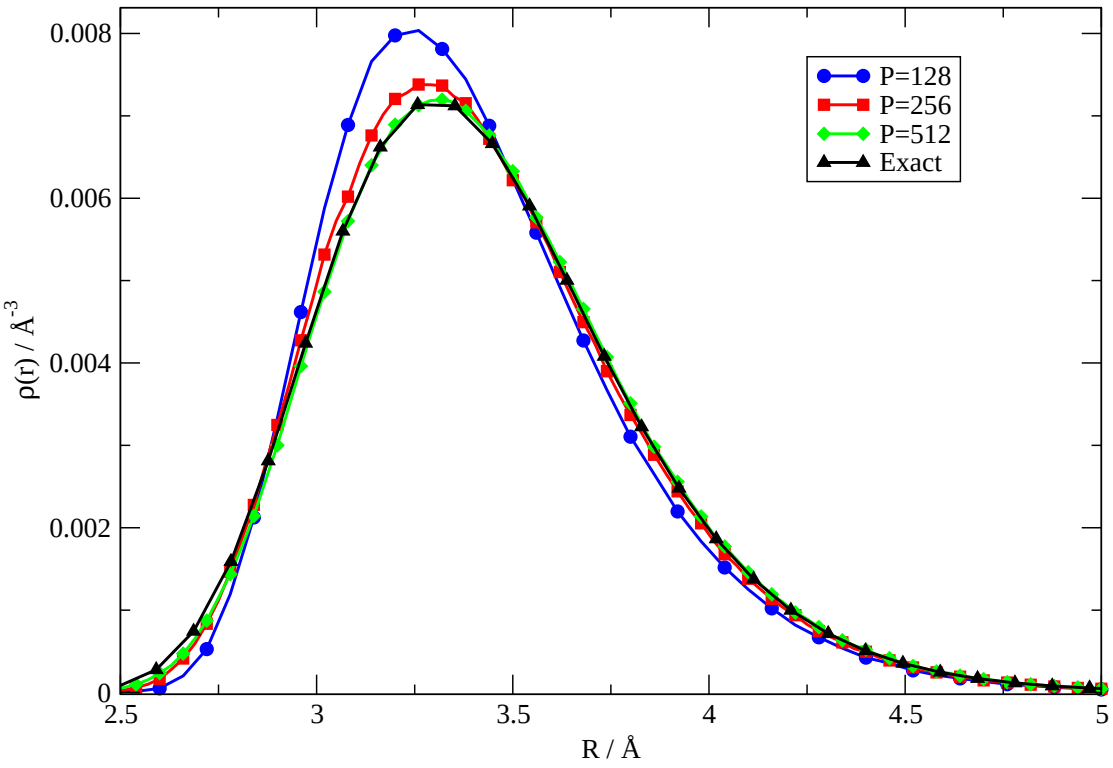


Figure 2.4: Radial distributions of a helium atom about a $CO_2$ impurity at T=0.37K for varying numbers of path integral beads; P=128 (blue circles), 256 (red squares), and 512 (green diamonds). Distributions are obtained using a path integral Langevin equation thermostat for 10,000 independent data points. A basis set calculation is used to obtain the exact result (black triangles).

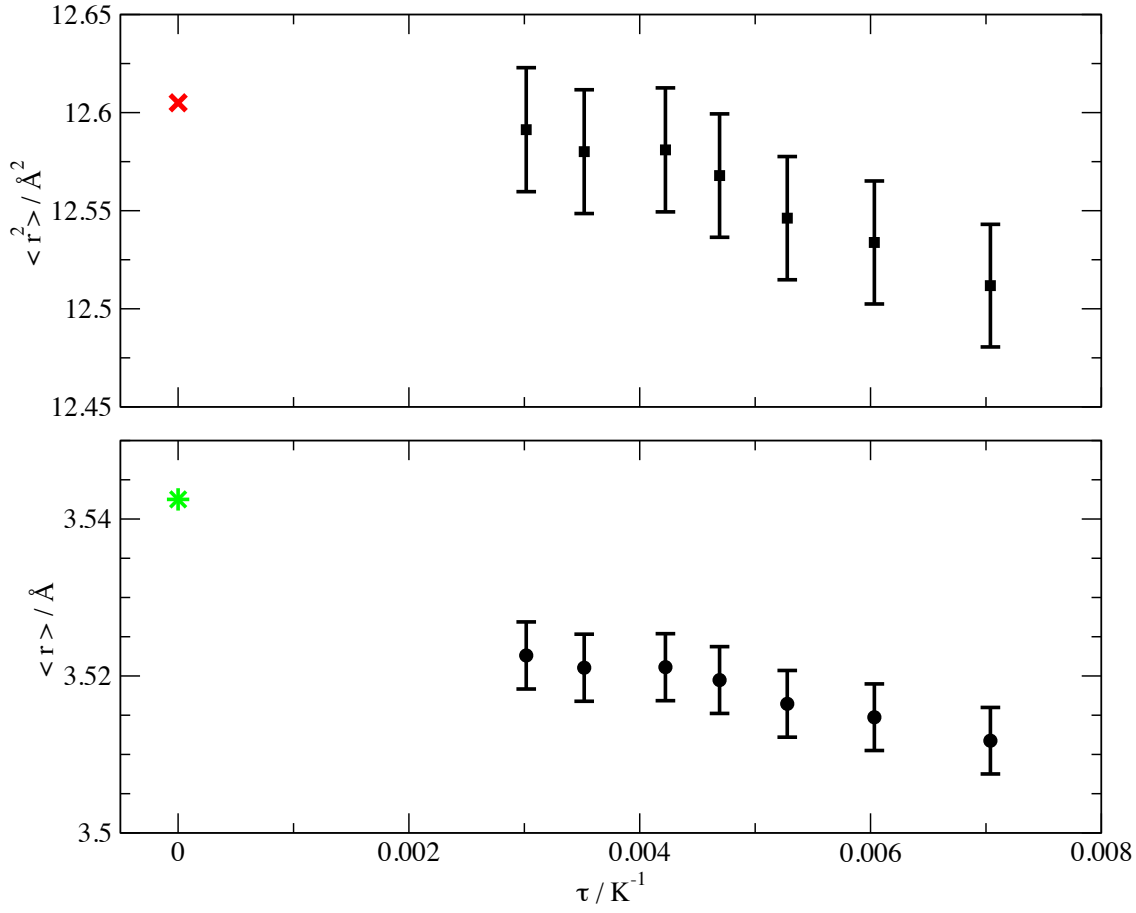


Figure 2.5: Average value of the radial position, $r$, and radial position squared, $r^2$, of a helium atom about a $CO_2$ impurity as a function of the imaginary time step $\tau = \beta/P$ at T=0.37K. Averages are calculated using all beads at $N = 10,000$ independent time steps with the PILE thermostat. Standard error bars are calculated based on the number of time steps, representing an upper limit on the error assuming there is a bead-bead correlation. The red X and green star represent exact data points that are calculated using a basis set calculation.

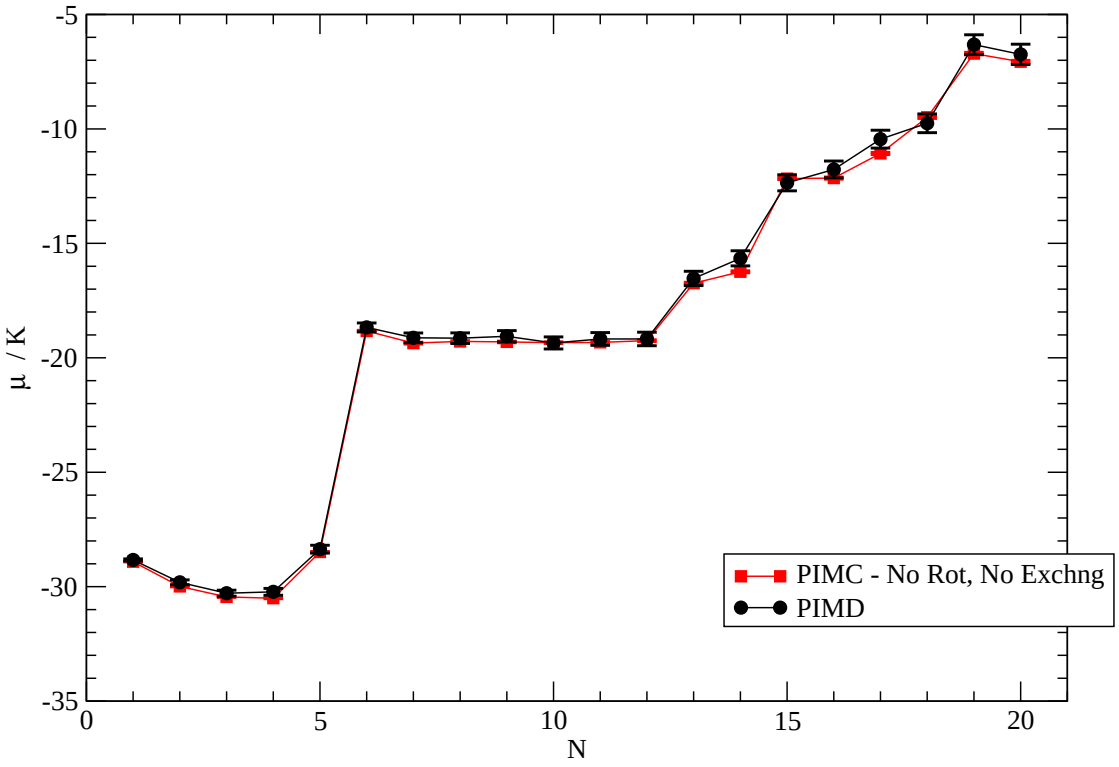


Figure 2.6: Chemical potential ($\mu$), the energy change resulting from the addition of a new helium atom to a cluster, of a $He_N$-$CO_2$ complex for both PIMD (circles) and PIMC (squares) methods. In both cases, the primitive energy evaluator is used.

the first indication of a less ergodic system when an atom is most stable at either $\theta \approx 35°$ or $\theta \approx 145°$. This is reflected by a large energetic jump at point $N = 6$ of the chemical potential. In this case, a longer simulation is required to converge structural properties.

In Fig. 2.7, for $N = 6 - 15$, the "wings" of angular density distribution becomes more pronounced and shifts closer to the central ring to accommodate more heliums. From the chemical potential, introducing new atoms requires constant energetic cost until 12 where we see the first significant indication of crowding effects from the $He - He$ interaction. This indicates that there is no stabilizing effect in a non-rotating helium cluster, even past a size of 17, the point of a complete solvation shell in both $CO_2$ doped helium[23] and parahydrogen clusters[51]. At 15 helium atoms, three separate rings and the 16th and 17th helium atoms "cap" the helium molecule which can be seen in the angular distribution by density around $\theta = 0°$ and $\theta = 180°$. In Fig. 2.7, we see a very localized distribution both in the angular and radial distribution compared to the previously published $He_N$-$CO_2$ clusters with rotation and exchange.[23] This effect is attributed to the lack of rotation in our calculations, which would introduce a centrifugal smoothing of the density as well as a shift to increased radial values.

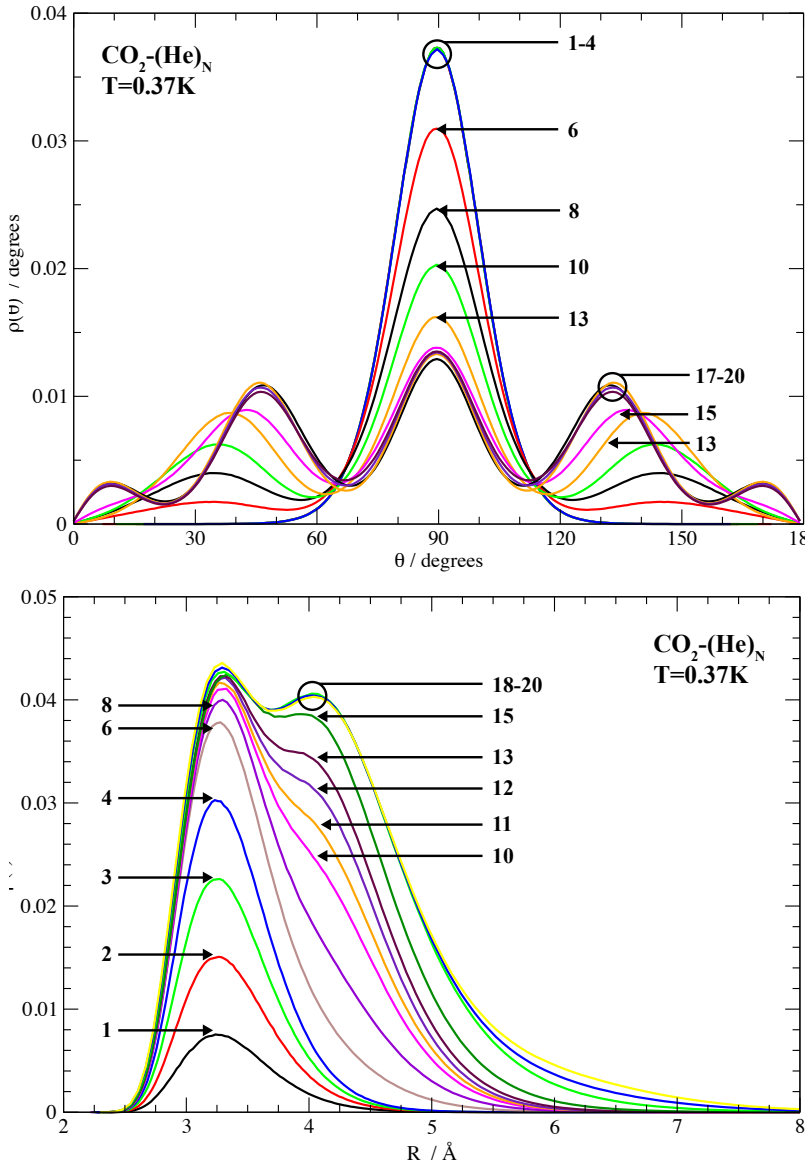


Figure 2.7: Angular (top panel) and radial (bottom panel) distributions of helium atoms about a $CO_2$ impurity in a $He_N$-$CO_2$ complex from $N = 1-20$. Angles are measured about the $CO_2$ aligned axis and the radial distance is measured as the pair distance from $CO_2$ to respective helium. The angular density is normalized so that $\int_0^\pi \rho(\theta)d\theta = 1$ and the radial distribution is averaged over all angles and normalized so that $4\pi \int_0^\infty \rho(r)r^2dr = N$. The angular distribution is symmetrized about $\theta = 90°$ due to the symmetry of the interaction potential.

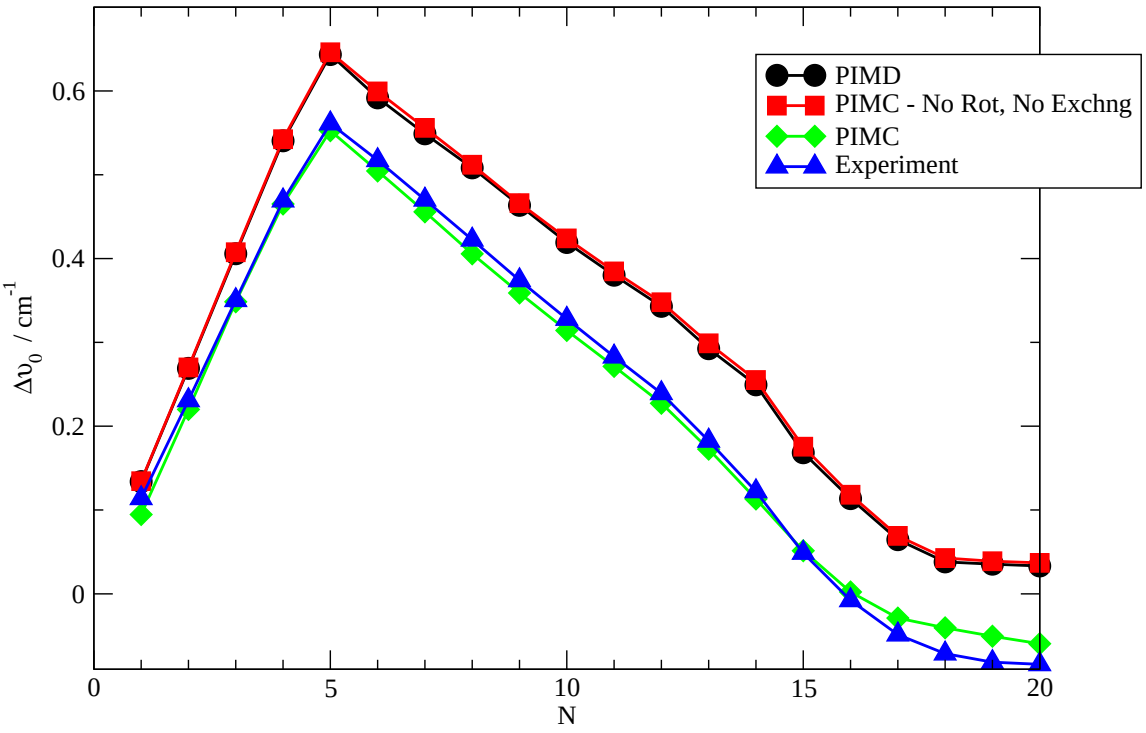


Figure 2.8: Experimental[24] (triangles) and theoretical[51] (diamonds) $CO_2(\nu_3)$ band origin shifts in $He_N - CO_2$ clusters from $N = 1 - 20$. PIMC results (diamonds) include rotation and bosonic exchange. Non-rotation and non-exchange results for PIMD and PIMC are depicted as circles and squares respectively. Theoretical values are obtained using the difference potential $\{V_1(r,\theta) - V_0(r,\theta)\}$ used in Ref. [23].

### 2.3.4 Vibrational Shifts

Though structural properties of PIMD and PIMC simulations are in agreement, we can quantify the error of neglecting rotation and exchange compared to experiment by calculating band origin shifts. Vibrational shifts are determined using an integration of the difference in potential between the ground state and first excited state for a summation and normalization of radial and angular pairs.[23] In Fig. 2.8, we present these shifts for three theoretical methods and experiment. The vibrational shifts are a reflection of the structure of the cluster. The agreement between our PIMD results and experiment for $N \leq 5$ is consistent with the expected rigid "donut" structure of the helium solvent around the $CO_2$ molecule which can be confirmed from the structural distributions in Fig. 2.7. In simulations with $N$ greater than 5, a difference of approximately $0.75 cm^{-1}$ is observed between PIMD results and experiment. The lack of molecular rotations leads to a more localized density distribution and the observed systematic error in the vibrational shift. It is interesting to note a deviation from this constant difference from experiment occurs at $N = 17$, the point at which one fills a solvation shell.

To quantify the effect of molecular rotations more accurately, PIMC results including these effects were simulated. A representative cluster of 18 helium atoms in Fig. 2.9 shows the effect of both rotation and exchange on the angular distribution. Our findings show

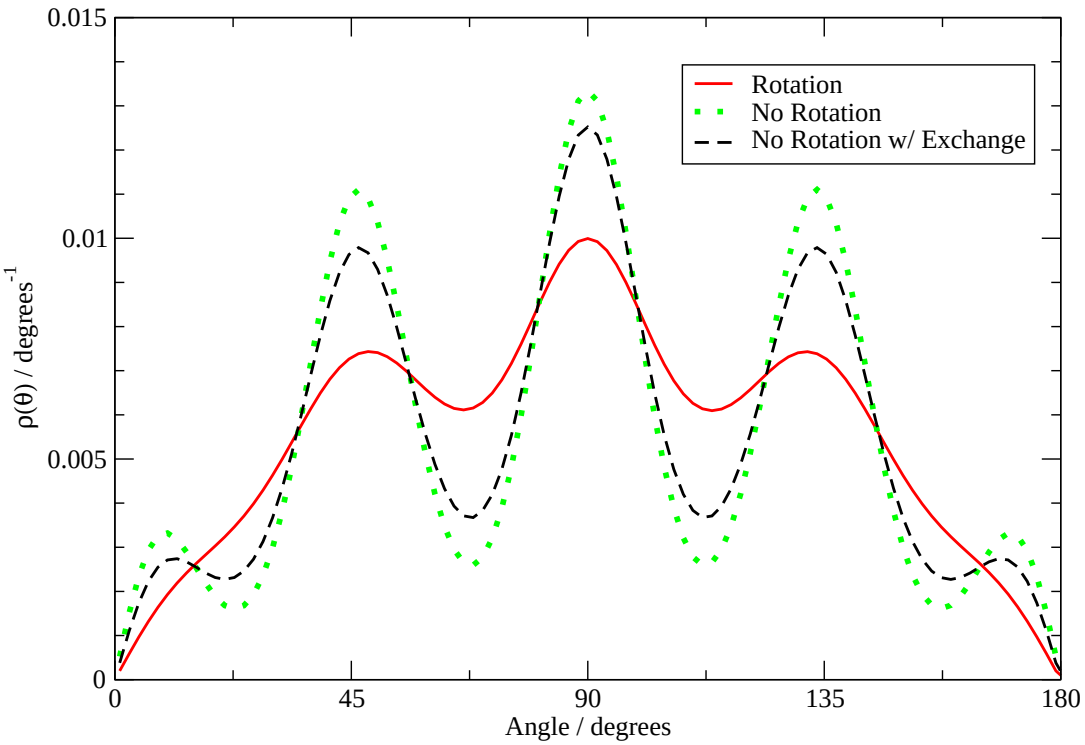


Figure 2.9: The rotational distribution of 18 helium atoms about a $CO_2$ impurity. The distribution with exchange (black) were calculated using PIMC with the worm algorithm for efficient sampling of bosonic exchanges. For the distribution including rotation (red) 128 rotational time slices were used to study the effects of rotation similar to the work in Ref. [23].

that the lack of rotation causes substantial localization of the angular peaks, resulting in the deviation from the experimental vibrational shift. The effect of bosonic exchange on doped helium clusters results in a relatively minor correct to the angular distribution in comparison. Molecular dynamics with rotation is a subject of interest for future work, and the importance of rotation in helium cluster simulations will be a subject for forthcoming calculations.

## 2.4 Conclusion and Outlook

Path integral molecular dynamics in combination with the path integral Langevin equation thermostat offers an efficient method for calculating equilibrium properties. This technique rivals the well adopted Monte Carlo sampling approach, while remaining straightforward to use with general potential models and forcefields. By studying the centre of mass energy distribution of our system, we observed the white-noise Langevin thermostat with the cartesian integrator insufficiently partitioned energy to translational degrees of freedom. Energy convergence studies confirm the need for a large number of beads when simulating $He - CO_2$ at low temperature. Energetic and structural properties are within statistical error of exact basis set calculations as well as path integral Monte Carlo results. Improvements on the convergence of both energetic and structural properties may be improved by

higher order propagators in future work. Further benchmarks on computational timings may be performed to assess the true advantages of PIMD over PIMC for low temperature systems. Compared to experiment and previously published theoretical results, doped helium clusters exhibited more pronounced localization (or more classical behavior) when exchange and rotations are absent in PIMD simulations. Future studies may be directed at quantifying the contribution of rotation and exchange in cluster simulations.

# Chapter 3

# Dynamical properties for doped helium clusters

Quantities obtained from classical molecular dynamics simulations at low temperatures and consisting mainly of light atoms incur substantial error due to the lack of nuclear quantum effects. Utilizing the path-integral isomorphism of the quantum partition function as presented in Chapter 1, one can recover these effects when calculating the equilibrium properties of chemical systems. Both molecular dynamics and Monte Carlo sampling methodologies may be used to obtain these average quantities. Although, due to the unphysical nature of Monte Carlo simulation dynamics, dynamical properties can only be obtained through molecular dynamics simulation, which relies on accurate interatomic force fields

to explore phase space.

The path integral formalism lends itself well to the calculation of thermodynamic averages, but obtaining dynamical properties is a much more challenging problem [52]. If one is interested in short-time quantum mechanical effects, perhaps those in condensed phase systems where interference effects are rapidly quenched, approximate methods may suffice. Both the centroid molecular dynamics (CMD) theory [53, 54, 55, 56, 57] and the ring polymer molecular dynamics theory [58] (RPMD) have been formulated for this purpose. The latter being more easily implemented and the subject of this research chapter.

In previous research, several publications present direct comparisons between both CMD and RPMD methodologies for identical systems, often including exact results [59, 60, 61, 62, 63]. With connection to the calculation of infrared spectra, Witt et al. [64] provide an excellent overview of the strengths and weaknesses of both methods. Further applications of centroid molecular dynamics includes rigid-body simulations of hexagonal ice [65], greater detail into the infrared spectrum of water [66], kinetic isotope effects in a decarboxylation and proton transfer reactions [67], rigorous comparison to model systems [68], large lithium para-hydrogen clusters [69], and the calculation of quantum reaction rate constants [70].

RPMD has been applied to study self-diffusion in quantum fluids [71, 72] and diffusion of a solute in liquid water and hexagonal ice [73], inelastic scattering of parahydrogen

[74], quantum effects of transition state theory [75, 76], spectra of hydrazine [77], the dynamics of an excess electron in supercritical fluid [78, 79], hybrid theoretical methods [80], proton transfer reactions in a polar solvent [59], comparison to linearized semiclassical initial value representation theory [81], gas-phase biomolecular reaction rates [82, 83] and quantum fluctuation in glass formation [84].

In this chapter, RPMD will be used with the PILE thermostat to obtain the quantum dynamics of the model quartic oscillator system and the He-$CO_2$ molecule. Autocorrelation functions are presented with comparison to exact calculations. Additionally, the dynamics of the dopant motion is presented using both the normal-mode and cartesian integrators, outlining the significance of the computational method. These results indicate that the RPMD method is applicable to compute the short-time quantum dynamics of molecular systems.

## 3.1 Theoretical Approach

Correlation functions are a essential tool for studying dynamical systems using molecular dynamics. In quantum mechanical representation, a correlation function is given as,

$$C_{AB}(t) = \frac{1}{Z} Tr \left[ e^{-\beta \hat{H}} \hat{B} e^{i\hat{H}t/\hbar} \hat{A} e^{-i\hat{H}t/\hbar} \right] \tag{3.1}$$

where Z is the partition function, $e^{-\beta\hat{H}}$ is a Boltzmann operator, and $e^{i\hat{H}t/\hbar}\hat{A}e^{-i\hat{H}t/\hbar}$ is the Heisenberg representation of the operator $\hat{A}$. When both operators $\hat{A}$ and $\hat{B}$ are identical, one can obtained an autocorrelation function as presented in this chapter.

### 3.1.1 Ring Polymer Molecular Dynamics.

In the manner first described by Craig et al.[58], we'll begin the justification for RPMD with the calculation of a thermal expectation of a coordinate dependent operator $A$ using a path integral discretization and a centroid variable,

$$\langle A\rangle_P = \frac{1}{(2\pi\hbar)^P Z_P}\int dp \int dq e^{-\beta H} A_P(q) \tag{3.2}$$

where $A_P(q)$ is the average over all beads 1...P of the $\hat{A}(q)$ operator. Both the path integral Hamiltonian $H$ in Eq. 1.23 and the discretized partition function $Z_P$ in Eq. 1.22 are described in Chapter 1. In principle, the above expression could be generalized to obtain the thermal expectation of the product of two coordinate dependent operators, $\hat{A}(q)$ and $\hat{B}(q)$,

$$\langle AB\rangle_P = \frac{1}{(2\pi\hbar)^P Z_P}\int dp \int dq e^{-\beta H(q,p)} A_P(q)B_P(q). \tag{3.3}$$

Instead of corresponding to the expectation of the operator product $\hat{A}\hat{B}$, this calculation gives us the exact real-time Kubo transformed correlation function at $t = 0$ for a Hamiltonian $\hat{H} = p^2/2m + V(q)$,

$$\tilde{C}_{AB}(t) = \frac{1}{\beta Z}\int_0^\beta d\lambda Tr\left[e^{-(\beta-\lambda)\hat{H}}Ae^{-\lambda\hat{H}}e^{i\hat{H}t/\hbar}\hat{B}e^{-i\hat{H}t/\hbar}\right] \tag{3.4}$$

By using the equations of motion derived from the Hamiltonian in Eq. 1.23, one can actually treat these $t = 0$ positions and momenta and initial conditions and evolve them in time to obtain approximate correlation functions that are valid for short times, and exact in the limit of a harmonic potential,

$$\langle A(0)B(t)\rangle_P = \frac{1}{(2\pi\hbar)^P Z_P}\int dp\int dq e^{-\beta H(q_0,p_0)}A_P(q_0)B_P(q_t). \tag{3.5}$$

Eq. 3.5 is displayed schematically in Fig. 3.1, in what is referred to as RPMD, where arrows represent molecular dynamics simulations. Statistically independent initial positions and momenta are sampled from a canonical simulation and real-time microcanonical simulations are run from these seeds. Autocorrelation calculations are computed for coordinate dependant observables, by averaging a quantity over all the beads in a atoms path, for the length of the simulation. Each initial condition's autocorrelation is then averaged together to obtain the approximate real-time Kubo correlation function. All the correla-

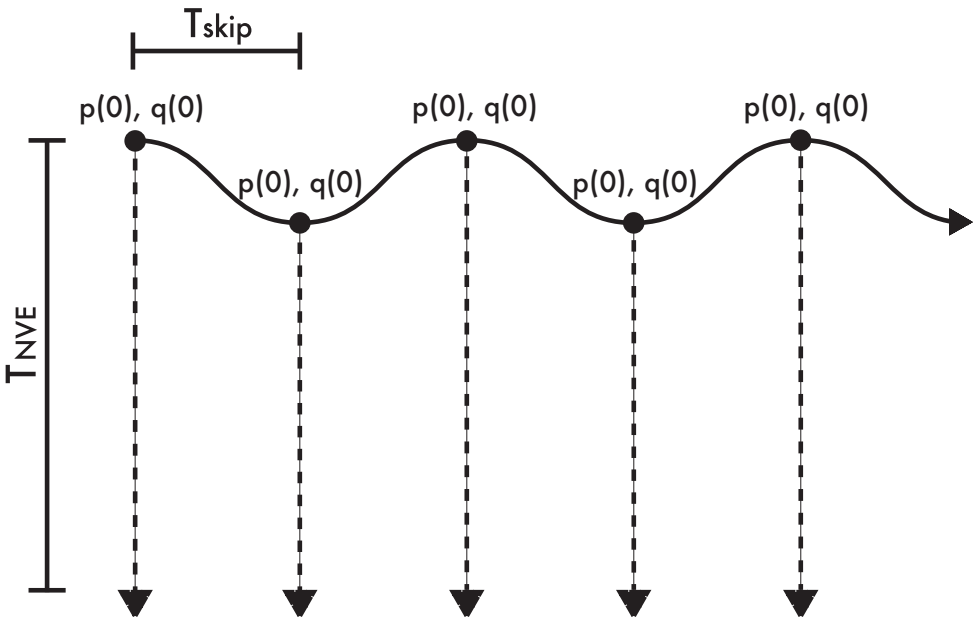


Figure 3.1: Diagrammatic representation of the RPMD method. The solid arrow indicates an NVT simulation where initial positions and momenta are sampled at a frequency of $1/Tskip$ over the length of a single trajectory. Dotted arrows indicate NVE simulations of length $T_{NVE}$ which are averaged over to obtain an approximation to quantum dynamics.

tion functions presented in this work are plotted after subtracting the mean value of the initial conditions.

Although the Kubo transformed real-time correlation function in Eq. 3.4 and the standard real-time correlation quantum mechanical correlation function in Eq. 3.1 differ, they are related through a Fourier transform relation,

$$G_{AB}(\omega) = (\hbar\beta\omega/2)[\coth(\hbar\beta\omega/2) + 1]\tilde{G}_{AB}(\omega), \tag{3.6}$$

where $\tilde{G}_{AB}(\omega)$ and $G_{AB}(\omega)$ are the Fourier transforms of $\tilde{C}_{AB}(t)$ and $C_{AB}(t)$ in Eq. 3.4 and Eq. 3.1 respectively [58].

This ad-hoc justification for the use of RPMD may be better appreciated in lieu of the formal derivation of centroid molecular dynamics [53, 54, 55, 56, 57] not presented here.

## 3.2 Computational Methods

Constant temperature simulations for the PIMD calculations of initial conditions is performed with the path integral Langevin equation thermostat.[22]. In an advantageous normal mode coordinate system or so-called Fourier path integral representation,[14, 15] the internal modes of a path integral atom can be determined analytically and thus ther-

mostatted explicitly with an optimal friction using this thermostat. The free parameter of this thermostatting technique is the centroid friction which must be obtained from a sample calculation. The value of $T_{skip}$ in Fig. 3.1, or the frequency with which to obtain initial conditions, is chosen as 10 times the relaxation time of the structural properties of x,y,z from a short NVT simulation and $T_{NVE}$ is obtained by overestimating the average dynamical relaxation times of initial RPMD results for our systems of study.

A vibrational modes calculation near the energy minimum of the system studied is performed to determine the smallest characteristic time-scale of the system. A value of approximately 10 times smaller than this time-scale is set as the time-step for both the microcanonical and canonical simulations.

As future studies will be used to study the effect of quantum statistics on the dynamics of doped helium clusters, path integral Monte Carlo methods are used to obtain initial conditions. Configurations are sampled every 2000 blocks with 2500 passes per block. As PIMC results are used for initial configurations only, random momenta is drawn from a Boltzmann distribution and applied to each bead.

Microcanonical simulations are performed with either the standard cartesian integrator or the normal mode integrator. The effect of bosonic exchange on paths amounts to the breaking and connecting of different paths throughout the simulation. This can be done efficiently with the worm algorithm implemented in our path integral Monte Carlo

code [85]. As the normal mode integrator must be reformulated to support arbitrary path connections, the cartesian integrator is benchmarked for quantum dynamics. The cartesian integrator is expected to introduce non-ergodicity to the exploration of phase space, but may still suffice for the short-time dynamics we chose to investigate.

## 3.3 Results and Discussion

### 3.3.1 Quartic Oscillator

In order to confirm the validity of the theoretical methodology, we first aimed to reproduce the correlation of a model system, the strongly anharmonic oscillator as was first calculated using RPMD by Craig et al. [58]. This oscillator is simulated using a potential of the form $V(x) = 0.25x^4$ about the simulation origin for each coordinate (3D) and $m = 1$ in atomic units. The PILE thermostat is used with $P = 32$ beads at $\beta = 8$ to sample initial conditions. The centroid friction of the thermostat is selected by measuring the autocorrelation time of the quartic oscillator from a short microcanonical energy simulation as was done in Chapter 2. Both Kubo and un-Kubo results are shown with comparison to the exact solution in Fig. 3.2. Exact results are obtained using a basis set calculation similar to what was presented Chapter 2, making use of Eq. 3.1. These RPMD results are consistent with previous calculations and support the study of the quantum dynamics of

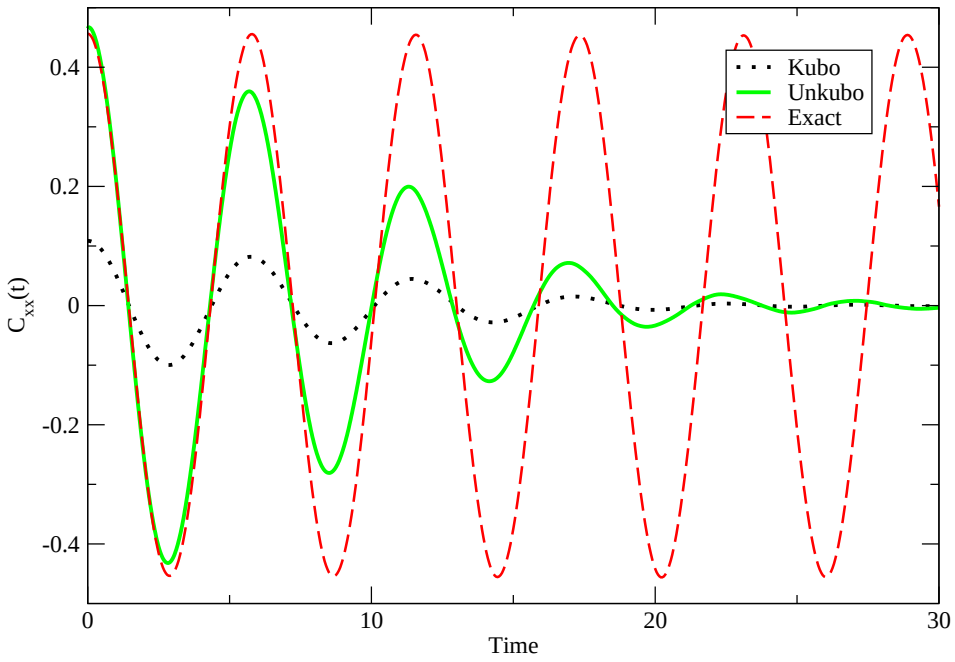


Figure 3.2: Coordinate autocorrelation for a strongly anharmonic 3D quartic potential $V = 0.25x^4$. This simulation is performed at $\beta = 8$ and averaged over 10000 trajectories with $P = 32$ beads. Kubo autocorrelation function (black dotted curve), which is obtained directly from RPMD averaging, and its transformed un-Kubo result (green solid), are compared to the exact calculation (red dashed). Both axis are in atomic units.

more complex systems.

### 3.3.2 Helium Dynamics about $CO_2$

The motion of helium about a $CO_2$ impurity aligned along the z-axis is depicted in Fig. 3.3.

This simulation is conducted at $T = 0.37K$ using $P = 256$ beads. Both the x and y

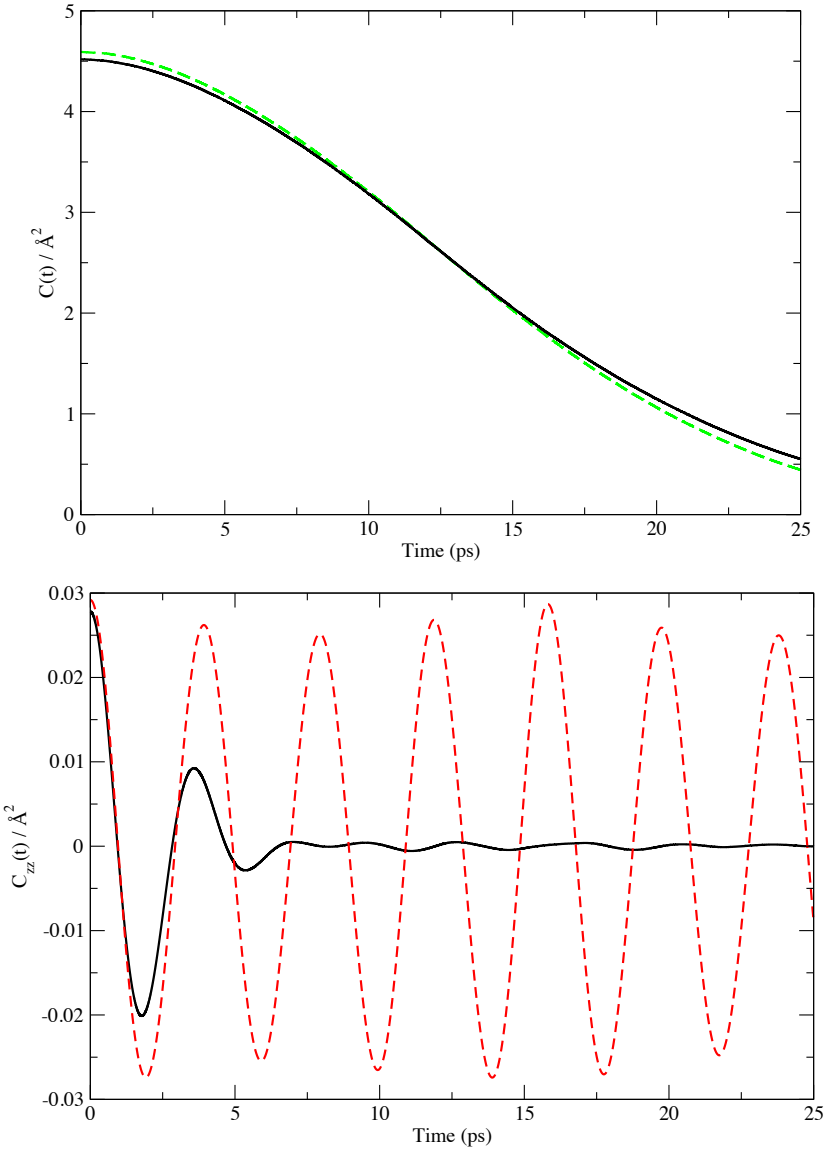


Figure 3.3: Coordinate Kubo autocorrelation functions for a helium atom about a $CO_2$ impurity at $T = 0.37K$ for $P = 256$. Symmetrical components x (black solid) and y (green dashed) are depicted in the top panel, while the bottom panel depicts the z component along the $CO_2$ aligned axis (black solid) with exact basis set calculation (dashed red). 30000 initial conditions are used followed by 25ps of microcanonical simulation.

coordinate Kubo autocorrelation functions are presented as a measure of the convergence of our simulation. A slight discrepancy at time zero indicates that statistical error is still present in these simulations and better sampled initial conditions could be obtained. The z coordinate Kubo autocorrelation is presented in the bottom panel of Fig. 3.3 with comparison to a basis set calculation. As was stated in Sec. 3.2, the time zero value for an RPMD calculation should converge to the exact quantum mechanical result in the limit of infinite beads, so a small systematic path integral error from $P = 256$ beads is also be present. These results indicate that RPMD is an effective tool for capturing the dynamics on the scale of 5ps at this temperature.

### 3.3.3 Cluster Dopant Dynamics

By previous cluster studies [23] and the work presented in Chapter 1, we know that the $CO_2$ dopant is fully solvated for this cluster size, giving one complete shell of heliums about the impurity. In this study, statistically independent initial conditions are sampled from PIMC simulations and inserted into the PIMD software. PIMC initial conditions are sampled to allow for the future comparison to two quantum statistical methods which are available in the PIMC code, namely Bose and Boltzmann statistics. By comparing quantum dynamics with proper quantum statistics with and without, one could gain insight into error incurred by neglecting bosonic exchange in helium cluster simualations.

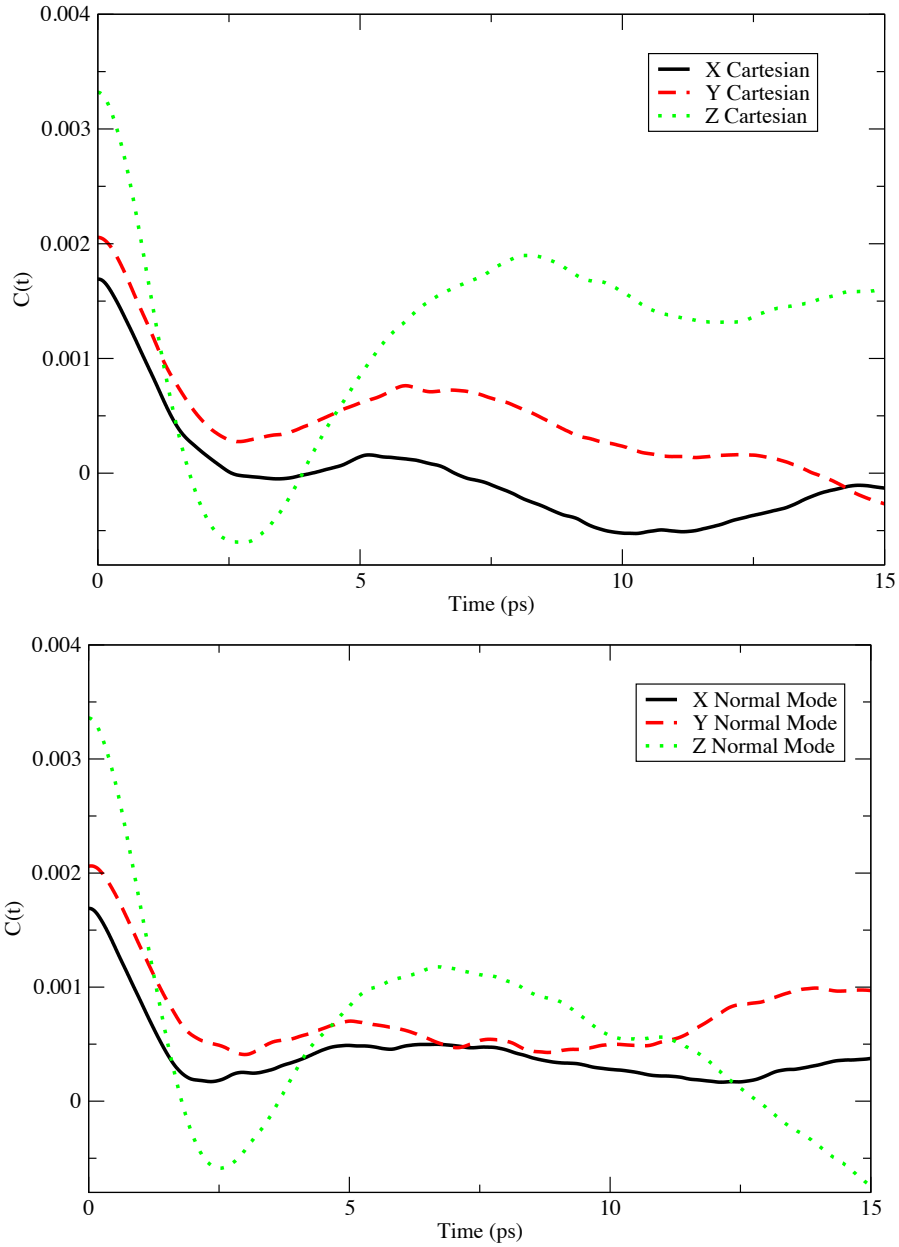


Figure 3.4: Coordinate Kubo autocorrelation for point $CO_2$ in a solvated helium cluster of $He_{17}$ for $T = 0.37K$ with $P = 256$. 1000 initial conditions are used followed by 15ps of microcanonical simulation using the cartesian integrator (top panel) and the normal mode integrator (bottom panel).

Standard Boltzmann statistics are used to study motion of a $CO_2$ dopant within a $(He)_{17}$-$CO_2$ cluster in Fig. 3.4. RPMD calculations are shown using both the cartesian and normal mode integrators when performing constant energy dynamics. It is well accepted that dynamics in either of these integration schemes will differ by a considerable amount for large numbers of beads, but it remains to be shown if short-time constant energy simulation on average will converge to the same result over a large number of trajectories. Both integration methods yield correlation functions that are similar for less than 5ps and diverge potentially only due to statistical error. We expect the time zero value for the X and Y autocorrelations to be the same due to symmetry in the interaction potential, so it's clear that 1000 initial conditions is statistically insufficient for our average. On this preliminary evidence, one may suppose that although the normal-mode integrator is considered the gold standard for correct PIMD simulation, the cartesian integrator may be used for short-time simulations.

## 3.4 Conclusion and Outlook

As traditionally used PIMC methods are incapable of obtaining the quantum dynamic properties such as those presented, PIMD remains an extremely valuable tool. It is reassuring to reproduce the exact results in the short-time limit where purely classical correla-

tions would be woefully incorrect. The short-time dynamics of He-$CO_2$ are reproduced for less than 5 ps. Although an exact result is not presented for the cluster dopant dynamics, we observe a correlation decay of less than 2ps.

Future studies should push these dynamic calculations to more complicated systems. As a follow up to these results, the dipole autocorrelation of a biomolecule is presented in Chapter 4 using identical simulation methods. RPMD itself offers a quick and reliable method for short-time quantum dynamics for the systems studied, but a more complete study should include a comparison calculation done with centroid molecular dynamics method to determine if RPMD gives a satisfactory approximation.

# Chapter 4

# Properties for complex molecular systems

Though weakly bound clusters are an excellent system to benchmark the efficacy of low-temperature path-integral sampling, generality of this method is of equal concern. We expect that the path integral Langevin equation should be effective for the simulation of any chemical system governed by any interaction potential for a range of temperatures. Though an exhaustive benchmark is not feasible, two key applications of this method will be presented.

In this chapter we will address the calculation of the ground state properties of the

quantum harmonic oscillator using the path integral ground state (PIGS) technique. In principle, this method allows one to calculate exact ground state expectation values for a quantum many-body system. This technique is a powerful modification of the traditional path integral method that rivals the techniques of diffusion Monte Carlo [86] and variational Monte Carlo [21] for studying ground state properties. For these path integral ground state simulations, it remains to be shown if the path integral Langevin equation sampling still suffices.

We'll also present the PIMD simulation of a biomolecule in solution, in particular, the sugar $\beta$-D-arabinofuranoside. In this study, the simplistic diatomic potentials of He-$CO_2$ and He-He used in prior studies are replaced with a highly parameterized AMBER [87] forcefield which is composed of a number of bonded and non-bonded interactions designed for biomolecules. With complex forcefields, it remains to be shown if the sampling method presented in the previous chapter suffices. Additionally, PIMD will be performed using periodic boundary conditions to better reproduce solvation effects. In the case of this biomolecule, exact solutions are not known and experimental results for observables are not available in the literature, so experimental comparisons must be drawn with caution.

## 4.1 Ground State Equilibrium Properties

Previous PIGS calculations for observables include initial applications to van der Waals complex $He_2Cl_2$, liquid and solid $^4$He [88] including fourth order propagator studies of solid $^4$He [50], para-$H_2$ and ortho-$D_2$ clusters [89, 90]. PIGS-MD, the ground state sampling technique used in the following study, was first published by Miura and applied to solid and liquid $^4$He [91, 29]. This following theoretical approach provides a background on PIGS-MD and closely follows the derivation by Miura [91, 29].

### 4.1.1 Theoretical Approach

The theoretical formulation of projection methods like path integral ground state have been frequently addressed in the literature. For a more complete derivation of the path integral ground state method, we refer the reader to Ceperely [21] or Sarsa [88]. At its core, given a system of N particles of mass $m_i$ where the coordinates of all atoms are represented by the vector $q^N$, and a Hamiltonian of the form,

$$\hat{H} = \hat{K} + \hat{V} = \frac{-\hbar^2}{2m_i} \sum_{i=1}^{N} \nabla_i^2 + V(q^N), \tag{4.1}$$

one can obtain a new wavefunction, $\psi$, which converges to the exact ground-state wave-

function in the limit of infinite imaginary time, $\beta$,

$$|\psi_0\rangle \propto \lim_{\beta\to\infty} exp[-\beta\hat{H}/2]|\Psi_T\rangle \tag{4.2}$$

where $\Psi_T$ is some trial wavefunction. By taking a scalar product of the above relation with itself, one can obtain a pseudo-partition function by inserting a complete set of states $q$ and $q'$,

$$Z_0 = \langle\Psi_T|exp[-\beta\hat{H}]|\Psi_T\rangle = \int\int dqdq'\langle\Psi_T|q\rangle\langle q|exp[-\beta\hat{H}]|q'\rangle\langle q'|\Psi_T\rangle \tag{4.3}$$

The matrix element in the right side of this expression may be familiar to the reader from Eq. 1.12. This matrix element could be referred to as the density matrix at temperature $\beta$. As was done in the Chapter 1, this term can rewritten in terms of a discretized path integral,

$$\langle q|[exp[-\tau\hat{H}]]^P|q'\rangle = \int\left[\prod_{s=1}^{P-1} dq_s\right]\prod_{s=0}^{P-1}\langle q_s|exp[-\tau\hat{H}]|q_{s+1}\rangle \tag{4.4}$$

where $\tau = \beta/P$ is the unit of discretized imaginary time and the matrix element in the right side of this representation is the discretized imaginary time action. An explicit expression for the action is dependant on a short-time approximation [21]. Using this approximation

and the discretization procedure, one can rewrite the pseudo-partition function as,

$$Z_0 \propto \int \left[\prod_{s=0}^{P} dq_s\right] \Psi_T(q_0) \left[\prod_{s=0}^{P-1} \langle q_s | exp[-\tau \hat{H}] | q_{s+1} \rangle\right] \Psi_T(q_P). \tag{4.5}$$

This partition function could be interpreted as a configurational integral over $N$ interacting ring polymers of $P$ beads in a similar manner to the path integral isomorphism. However, these paths are now open chains as opposed to cyclic rings. Additionally, end-point coordinates $q_0$ and $q_P$ for all atoms are influenced by the trial wavefunction $\Psi_T$. This model is depicted schematically in Fig. 4.1.

It can be shown that in the limit of a small imaginary time discretization, $\tau \to 0$, and the long imaginary time, $P\tau \to \infty$, that the middle bead of the path (assuming P is odd) is sampled from a probability distribution to the square of the exact ground state wavefunction regardless of the quality of the $\Psi_T$ value.

Quantities that are linear in coordinate representation, $\hat{A}(q)$, can be obtained by simply averaging over a set of $T$ path mid-point conformations,

$$\langle \psi_0 | \hat{A}(q) | \psi_0 \rangle \approx \frac{1}{T} \sum_{t=1}^{T} A(q_{P/2}(t)). \tag{4.6}$$

As with standard path integral molecular dynamics, energy estimators must be used

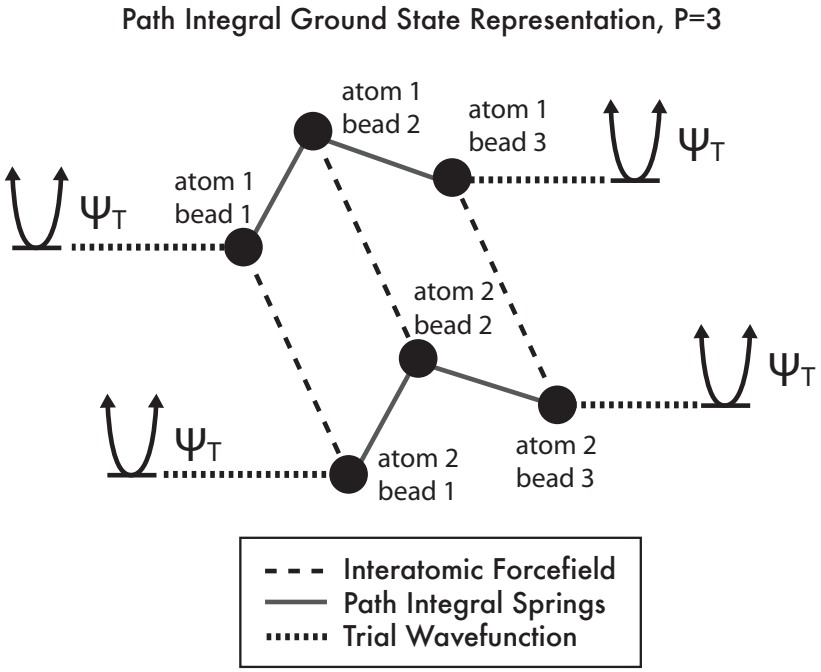


Figure 4.1: The smallest representation of two atoms interacting in the path integral ground state isomorphism. The interatomic forcefield acts with a factor 1/P between all sets of beads for all atoms, and path integral springs couple the beads of a single atom and end-point coordinates are affected by a trial wavefunction.

to obtain ground state energies. Though a more robust estimator can be formulated in a similar manner to the primitive estimator of path integral molecular dynamics, in its simplest form, the "mixed estimator" may be used, which works since the operator $\hat{H}$ commutes with the operator $exp[-\tau\hat{H}]$,

$$\langle\psi_0|\hat{H}|\psi_0\rangle \approx \frac{1}{T}\sum_{t=1}^{T}\frac{\hat{H}\Psi_T(q_1(t))}{\Psi_T(q_1)(t))} \approx \frac{1}{T}\sum_{t=1}^{T}\frac{\hat{H}\Psi_T(q_P(t))}{\Psi_T(q_P)(t))} \tag{4.7}$$

### 4.1.2 Computational Methods

The implementation of PIGS-MD in MMTK [20] required the removal of one path integral harmonic spring to open the path and the addition of two potential terms at the end beads. As the normal-mode integrator doesn't support path integrals with open paths, the white-noise Langevin equation is used with a friction selected at the geometric center of the vibrational frequencies from a vibrational modes calculation. For the results presented in the following section, a model system is studied, so a vibrational analysis is superfluous, but is performed as such for generality.

The harmonic oscillator model system we studied had the parameters $m = 1$ and $\omega = 1$ in atomic units. In a similar manner to the previous studies, the time-step is set to 0.05 times the fastest characteristic time-scale from a vibrational modes calculation.

Both energy and square ground-state wavefunction properties are sampled at a frequency of $0.0005/\Delta t$ number of steps for $2.5/\Delta t$ steps in total, resulting in 5000 data points per average. Using a convenient trial wavefunction, $\Psi_T = 1$ we can sample the ground state energy at the ends of the path at the same frequency. Both ends of the path

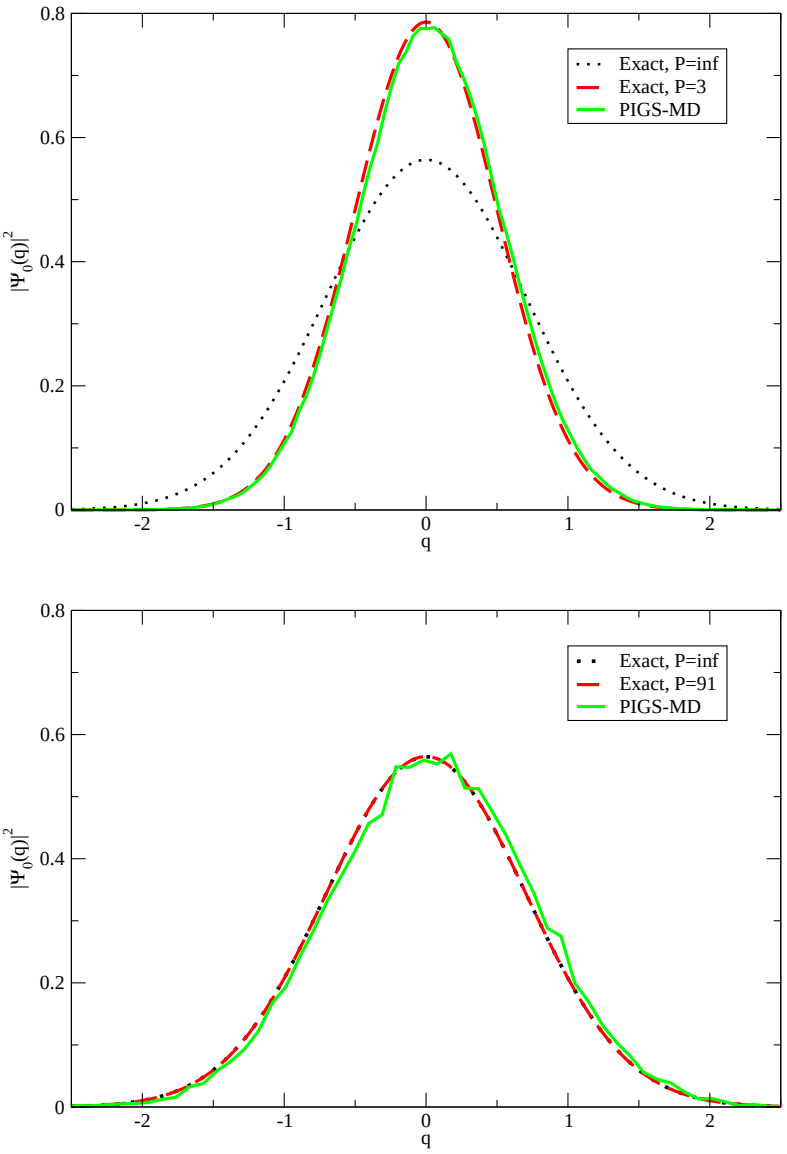


Figure 4.2: (Top panel) A $P = 3$ PIGS-MD simulation of a quantum harmonic oscillator with comparison to exact calculations for $\beta = 10$. (Bottom Panel) A $P = 91$ PIGS-MD simulation of the same type.

may be averaged at each sampling step to improve efficiency by Eq. 4.7.

The square ground-state wavefunction is sampled from the center of the path using Eq. 4.6. As the quantum harmonic oscillator can be solved analytically, both the exact ground-state energy and wavefunction can be calculated. The square of the ground-state wavefunction is

$$|\psi_0(q)|^2 = \left(\frac{m\omega}{\pi\hbar}\right)^{1/2} e^{-(m\omega/\hbar)q^2}. \tag{4.8}$$

More interestingly, an analytic expression was derived for the finite-bead path integral approximation to the ground state wavefunction [92, 86]. This quantity is given by

$$\begin{aligned} |\psi_0(q)|^2 &= \lim_{P\to\infty} \sqrt{\frac{mP}{2\pi\beta\hbar^2}} \times \sqrt{\frac{f^{P-1}(f^2-1)}{(f^{2P}-1)}} \\ &\times \exp\left[-(m\omega/2\hbar)q^2 \frac{(f+1)(f^P-1)}{f^{1/2}(f^P+1)}\right], \end{aligned} \tag{4.9}$$

which utilizes the following relations

$$\begin{aligned} R &= \frac{\omega\tau}{P} \\ f &= 1 + \frac{1}{2}R^2 + \frac{1}{2}(4+R^2)^{1/2} \end{aligned} \tag{4.10}$$

This truncated series expansion for a discretized ground-state wavefunction can thus be used to assess the sampling convergence even in the limit of a small number of path discretizations (small $P$), as will be depicted in the next section. Note that in practice both wavefunctions are renormalized to ensure the correct prefactors are used when comparing Eq. 4.8 to Eq. 4.9.

### 4.1.3 Results and Discussion

The square wavefunction of the harmonic oscillator is presented in Fig. 4.2 for large ($P = 91$) and small ($P = 3$) value for the number of discretizations at $\beta = 10$. Exact square ground-state wavefunctions are calculated using Eq. 4.8 to Eq. 4.9. It is determined from this structural analysis that $\beta = 10$ is sufficient imaginary time propagation to reproduce the quantum harmonic oscillator density using a large number of path discretizations ($P = 1000$).

The cumulative average ground state is presented in Fig. 4.3 as a function of PIGS-MD sampling steps. This plot shows fast convergence over sampling time. When $\beta = 10$ is fixed, the convergence of the ground-state energy is studied as a function of path discretization error $\tau$. A quadratic fit indicates a slight deviation from the expected infinite bead limit of $< E_0 >= 0.5$ which may be due to some correlation in sampling, or perhaps insufficient lengths of imaginary time $\beta$.

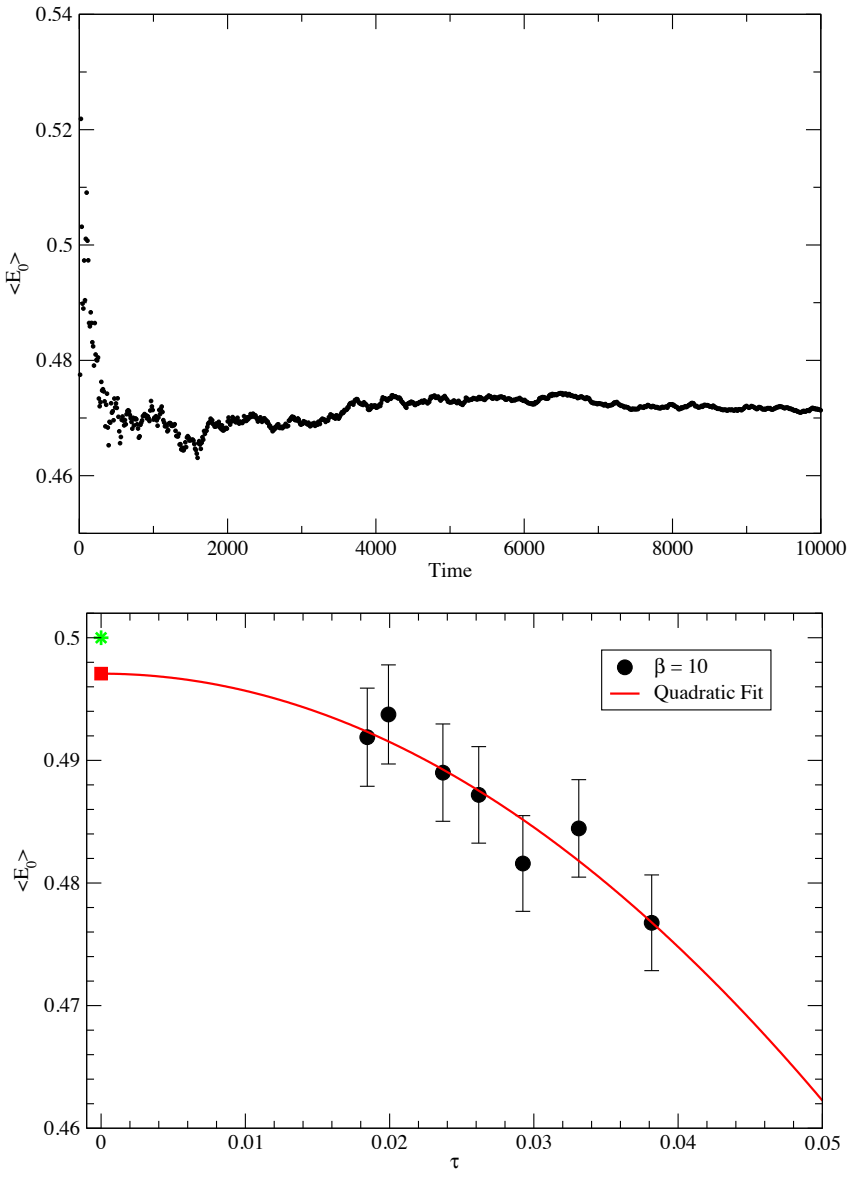


Figure 4.3: (Top panel) Cumulative average ground-state energy of the quantum harmonic oscillator for a PIGS-MD simulation as a function of time for $P = 91$, $\beta = 10$. (Bottom panel) Ground-state energy convergence of the quantum harmonic oscillator as a function of $\tau = \beta/P$ and a quadratic fit (red line), with exact result (green star). Error bars are calculated using the standard error of mean.

## 4.2 Biomolecular Simulation

Drug-resistant strains of the bacteria that cause tuberculosis, such as *Mycobacterium tuberculosis*, have lead scientists to probe the molecular activity that occurs on the mycobacterial cell wall [93, 94]. The treatment of bacterial diseases often begins by disrupting the formation of the wall itself [95]. A first step to understanding the interactions that occur on the mycobacterial cell wall is to investigate the dynamics of its components in isolation, one of such components being Furanose (or five-membered ring) carbohydrates. Furanosides are believed to play a role in the protection of cells against its hostile environment promoting the formation of a lipophilic barrier at the outer part of the cell wall preventing antibiotics and hindering immune response[93].

Understanding the conformational preferences of these biomolecules is an area of active interest [96, 97, 98, 99]. These studies have focussed on developing models for furanosides and subsequently obtaining equilibrium properties that match NMR acquired J-couplings measurements. Previous theoretical studies have made use of high-level ab initio simulations but no study on these biomolecules has been published including nuclear quantum effects with a method like PIMD. It is the aim of this work to present simulation results on five-membered rings to determine if quantum effects have any significant role.

The application of PIMD to the simulation of such biomolecules is often restricted due

to complexity of both the forcefields that must be used and the large number of atoms present in these systems. Though the current PIMD implementation deals with forcefields in a robust manner, simulations are still hindered by long computing times. As such, only preliminary findings will be presented, and more thorough convergence analysis will be performed in subsequent research.

### 4.2.1 Computational Methods

Constant temperature simulation is performed with normal-mode path integral Langevin equation simulations at T=300K with a trajectory snapshot shown in Fig. 4.4. Qualitatively, the lighter hydrogen atoms appear delocalized compared to heavier carbon and oxygen atoms, and could be considered an approximate measure of quantum effects. Remaining consistent with previous molecular dynamics studies of methyl $\beta$-D-arabinofuranoside, the AMBER forcefield [87] augmented with the carbohydrate specialized GLYCAM parameter set is used[100].

For solution-based simulations, a bounding box of the sugar plus 10 Å for each coordinate axis is used as a periodic universe and filled with density of $1g/cm^3$ of water placed stochastically and minimized using a universe scaling and shrinking solvation technique. This resulted in the sugar being solvated with 110 water molecules. An Ewald summation cut-off of 9Å is set to handle electrostatic interactions. The water molecules are simulated

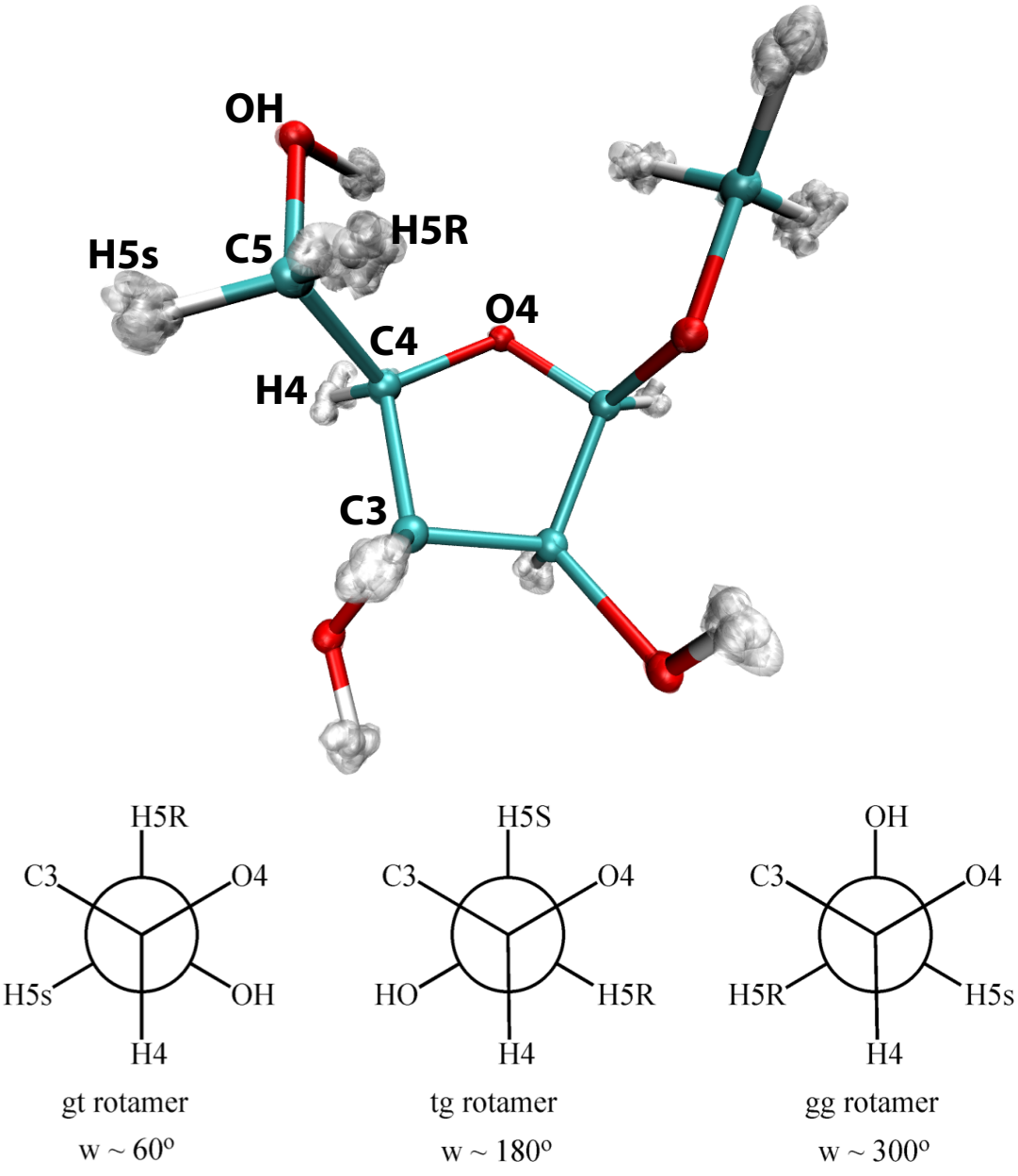


Figure 4.4: A snapshot from a gas-phase molecular dynamics trajectory of methyl $\beta$-D-arabinofuranoside with $P = 32$ beads for each atom simulated at $T = 300K$. Atoms labels correspond to the definitions of the three rotamers about the C4-C5 bond. Sugar is depicted in the *gt* rotamer.

with the SPC-Fw-q potential[101], a flexible water model which is parameterized in its development to reproduce the correct radial distributions of water when used in path integral molecular dynamics simulations. A preliminary simulation of 10 ns of simulation time is used for all solution results presented.

The time-step for these simulations is selected as 10 times smaller than the fastest vibrational mode including the full path-integral potential. Although a rigorous bead convergence study is not performed as is done for He-$CO_2$ system, trial calculations for $P = 4, 8, 16$ are performed. Autocorrelation times for observables such as the dihedral angle and dipole are obtained in short simulations in order to determine an efficient rate at which to sample data points for correct equilibrium averages.

RPMD computational methods are similar to those presented in Chapter 3. 1000 statistically independent constant temperature configurations and momenta at $T = 300K$ are sampled and used as initial conditions for constant energy simulations of 50ps in length.

### 4.2.2 Results and Discussion

NMR J couplings can be compared to a calculated value using a Karplus relation and C4-C5 dihedral angle densities [96, 97, 98, 99]. The three primary rotamer populations are defined in the lower panel of Fig. 4.4, where each diagram depicts a rotamer conformation

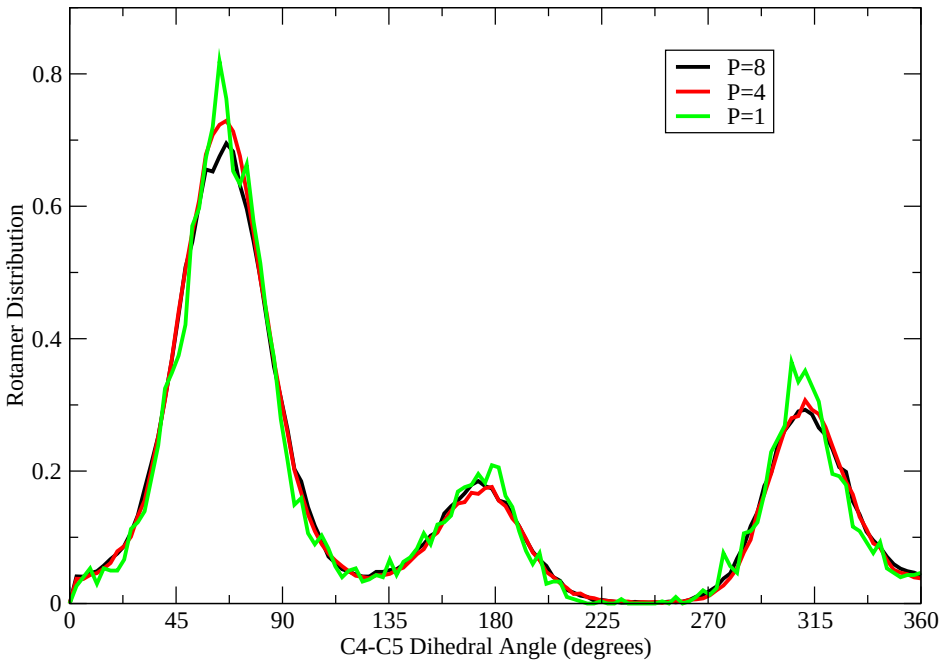


Figure 4.5: $\beta$-D-arabinofuranoside dihedral distribution using SPC-Fw water model for P=1 and SPC-Fw-$q$ for P=4,8 at 300K.

from a point of view parallel with the C4-C5 bond. Due to the limited extent of this study, the calculated J coupling values with comparison to experiment are not presented. The rotamer populations about the C4-C5 bond using classical molecular dynamics and PIMD for several numbers of beads is shown in Fig. 4.5. Integration of these peaks for both PIMD using SPC-Fw-$q$ and normal SPC-Fw water molecular dynamics simulations yields the results shown in Table 4.1. Previous research has indicated that the correct dihedral distribution converges in the limit of near 200ns of simulation [97]. It's expected

Table 4.1: Rotamer population distribution about the C4-C5 bond (*gt tg gg*)

| Method | Population | | |
|---|---|---|---|
| PIMD w/ SPC-Fw-$q$ | 62 | 15 | 23 |
| MD w/ SPC-Fw | 60 | 16 | 24 |
| Experiment [97] | 57 | 8 | 35 |
| TIP5P [97] | 46 | 20 | 34 |
| TIP4P [97] | 46 | 13 | 41 |
| TIP3P [97] | 40 | 12 | 48 |

that the short simulations presented will overestimate the gt rotamer and underestimate the gg rotamer. It is certain that these population results are not converged to the correct equilibrium values but what it is interesting to note that quantum effects do not play a significant role in these short simulations, despite the number of light hydrogen atoms in the simulation. These results offer a promising outlook on the use of PIMD for similar biomolecules like $\beta$-D-arabinofuranoside.

The quantum dynamics of this sugar in gas-phase is also studied for short-times. By performing a point-charge dipole calculation and obtaining the autocorrelation function of this property, one can gain insight into the vibrational response of this molecule. The dipole autocorrelation is shown in Fig. 4.6 at T=300K for different numbers of path integral beads. This result indicates that the dipole relaxation time is less than 10ps. Despite PIMD being a seemingly useful tool for equilibrium properties, we have no evidence that the dynamics of this sugar are influenced by quantum effects at this temperature.

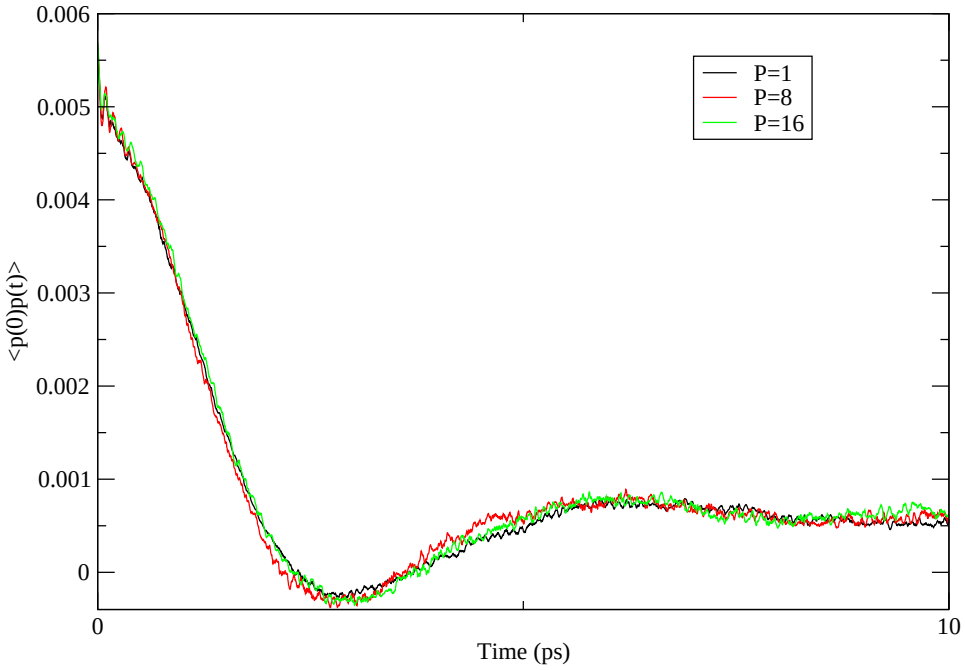


Figure 4.6: methyl $\beta$-D-arabinofuranoside point-charge dipole-dipole correlation function for varying numbers of beads. Calculated using RPMD at $T = 300K$.

## 4.3 Conclusion and Outlook

Both the biomolecular and ground-state simulations are useful studies where the contribution of quantum effects from path integrals range from negligible to completely essential. Both of these PIMD applications offer a stunning level of detail fit for thesis projects within themselves. Outside of cluster studies, the path integral ground state method is relatively unused. The application of these methods to biomolecules would be of great interest. Ad-

ditional studies of PIGS-MD should be performed in order to compare to the ground-state energy of formaldehyde [102].

Future research pertaining to the quantum dynamics of this biomolecule should investigate the correct Fourier transformation of this dipole autocorrelation to obtain infrared spectra and compare to spectroscopic data to understand the accuracy of RPMD for this purpose, as well, the determination of what temperature these quantum effects are significant.

# Chapter 5

# Conclusion

As diagonalizing the Hamiltonian to obtain exact quantum mechanical properties is a difficult problem, for all but model systems and few-body problems, approximate methods must be used. As such, there is no question that the path integral formalism is a powerful tool to circumvent this limitation, and will continue to serve as one of the gold standards for the calculation of equilibrium and dynamic properties including nuclear quantum effects.

Due to the substantial increase in computational overhead associated with path integrals, efficient theoretical foundations and software development are of great importance. This work demonstrates that the thermostatting technique of the path integral Langevin equation offers a robust method to obtain constant temperature simulation which was not benchmarked for low temperature weakly bound systems prior to this study.

By calculating equilibrium properties of doped helium clusters, we gained both a better understanding of the parameters and steps required to obtain efficiently calculated and accurate results and significant scientific insights. Benchmark convergence studies for doped helium clusters in terms of both structural and energetic properties confirmed the necessity for between 256-512 beads at our simulation temperature of 0.37K. Studying the effect of the Langevin thermostat equipartitioning to the center of mass indicated the importance of both the normal-mode integrator and correct thermostat parameters. Rotational and angular distributions confirmed previously calculated results, as well as vibrational shifts associated with these distributions. As the cluster size dependence of vibrational shifts showed, these structural distributions resulted in values that differed from experimental values. This indicates the importance of rotation for clusters, a subject of a more detailed research paper to follow.

The extension of these equilibrium techniques to the calculation of correlation functions resulted in accurate quantum dynamics of low-temperature clusters. The position autocorrelation results for the quartic oscillator matched exact values, confirming the validity of the RPMD method in our software package. The limitations of RPMD were tested for weakly bound clusters, indicating that RPMD reproduced the exact correlation for times of less than 5ps. Although an exact calculation is not presented, internal $CO_2$ motion also exhibited correlation decay on the order of 5ps ps and would be expected to

reproduce short-time exact results less than that. Both the normal mode integrator and cartesian integrators were used to present this correlation, indicating that both converge to a similar autocorrelation function for short times. Further dopant dynamics studies should be performed to understand the effect of bosonic exchange on dynamics. Additionally, biomolecular dipole autocorrelation is calculated using RPMD but further analysis and comparison to spectroscopic data is needed to confirm these calculations.

Finally, both the challenges of ground-state properties and large biomolecular systems were addressed with preliminary studies using minor modifications to the PIMD code developed for the previous two projects. The calculation of the ground-state energy of the quantum harmonic oscillator is a necessary step to assess the accuracy of this method. A convergence study is presented, indicating the correct parameters for efficient calculations of this nature. In solution phase, the dihedral angle about the C4-C5 bond of methyl $\beta$-D-arabinofuranoside shows significant deviation from classical results. For a quantum simulation at 300K, there is only slight deviation from the classical result and quantum effects showed negligible contribution to the dipole autocorrelation calculation. Additional studies should be performed to confirm that each of the main rotamers are effectively sampled and that the results shown in this paper are sufficiently converged. A bead convergence study should be performed at multiple temperatures to determine at temperature quantum effects become significant for this sugar.

## 5.1 Future Directions

Though the scope of chemical systems for applications for PIMD is vast, the essential direction of this research is to facilitate the simulation of larger and larger systems and processes probing where quantum effects may play a critical role. However, there are numerous aspects to this research that could benefit from a detailed study before advancing to new applications.

With regards to computational methods, the use of the PILE thermostat should be compared to alternative techniques such as Nose-Hoover chains in terms of efficiency and accuracy [103]. For systems with a large range of vibrational frequencies, the Generalized Langevin Equation thermostat should be considered [16]. This study would put the PILE method in better context as far as constant temperature simulation of highly quantum systems is concerned. Furthermore, the single free parameter of the PILE thermostat, the centroid friction, and the white-noise Langevin thermostat should be tested with a range of values for the systems presented. The quantum dynamics method of RPMD should also be compared to the more rigorously defined centroid molecular dynamics.

In a purely computational study, the benefit of parallelization by using multithreading or across multiple CPUs in a cluster would be a highly useful piece of information for all PIMD simulations using the code presented in this research. Efficiency benchmarks of

computer wall-time should be supplied and compared to alternative sampling methods like Monte Carlo to obtain similar error bars on ensemble averages.

Generally speaking, this thesis demonstrates the robustness of PIMD for a number of systems even those where quantum effects were not particularly significant. The detailed computational methodologies presented offer a viable template for future path integral studies. As such, several promising research projects are in progress which utilize the foundation presented in this research.

In a similar study to that of Chapter 2 of this thesis, a 3D parahydrogen-water potential is being rigorously tested using the methodology presented. Both energetic and structural convergence plots with comparison to exact basis set calculations will be presented. Normal-mode path integral Langevin equation calculations for this potential are being performed with intention to study large clathrate water cages for potential hydrogen storage applications.

Using the normal mode path integral integrator, one can selectively propagate certain modes. Conveniently, the zeroth mode from this normal mode transformation refers to the centroid motion of the atom. By removing the centroid propagation, one can calculate constrained centroid forces along a pair distance coordinate and calculate a potential of mean force for a diatomic system. This force can then be used as a course grained quantum potential with 1 bead instead of simulating a full path integral system of hundreds of beads

[104]. Using this theory, one may be able to simulate quantum clusters at low temperatures with sizes that approach the nanodroplet limit and hopefully reproduce spectroscopic measurements in that range.

As was stated in the biomolecule study of Chapter 4, there is a great depth of research left to be done to properly simulate a solvated sugar with path integrals. In previous molecular dynamics studies of this solvated sugar, 200ns of simulation in large boxes of water was necessary to converge structural properties. In order to run simulations of that length and size with PIMD, one needs to extend the capabilities of our software to either run in parallel or on high-performance GPU clusters. Current research is aimed to enable GPU computation through customized integrators and facilitate the simulation of any path integral system. Numerous publications praise the use of GPU computing for performing molecular dynamics [105, 106].

All of these applications build strongly on the results presented in this work and offer a positive outlook for numerous path integral simulations in the near future. As the code used in this thesis will be released as open-source software, PIMD will become a more accessible tool for researchers in the field of computational chemistry.

# References


[1] E. Schrodinger, Annalen der Physik **385**, 437 (1926).

[2] H. Schaefer, *Quantum Chemistry: The Development of Ab Initio Methods in Molecular Electronic Structure Theory*, Dover books on chemistry (Dover Publications, 2004).

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

[4] J. L. Klepeis, K. Lindorff-Larsen, R. O. Dror, and D. E. Shaw, Curr Opin Struc Biol **19**, 120 (2009).

[5] P. Sindzingre, M. Klein, and D. Ceperley, Phys. Rev. Lett. **63**, 1601 (1989).

[6] D. Frenkel and B. Smit, *Understanding Molecular Simulation, Second Edition* (Academic Press, San Diego, 2002).

[7] M. E. Tuckerman and G. J. Martyna, J. Phys. Chem. B **104**, 159 (2000).

[8] R. Feynman and A. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).

[9] H. Trotter, Proceedings of the American Mathematical Society **10**, 545 (1959).

[10] D. Chandler and P. Wolynes, J. Chem. Phys. **74**, 4078 (1981).

[11] R. Hall and B. Berne, J. Chem. Phys. **81**, 3641 (1984).

[12] M. E. Tuckerman, B. J. Berne, G. J. Martyna, and M. L. Klein, **99**, 2796 (1993).

[13] G. J. Martyna, A. Hughes, and M. E. Tuckerman, **110**, 3275 (1999).

[14] R. D. Coalson, D.L.Freeman, and J. Doll, J. Chem. Phys. **85**, 4567 (1986).

[15] C. Chakravarty, M. Gordillo, and D. Ceperley, J. Chem. Phys. **109**, 2123 (1998).

[16] M. Ceriotti, G. Bussi, and M. Parrinello, J. Chem. Theory Comput. **6**, 1170 (2010).

[17] M. Ceriotti, G. Bussi, and M. Parrinello, J. Chem. Theory Comput. **6**, 1170 (2010).

[18] G. Bussi and M. Parrinello, Phys Rev E Stat Nonlin Soft Matter Phys **75** (2007).

[19] L. Verlet, Phys. Rev. **159**, 98 (1967).

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

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

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

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

[24] A. R. W. McKellar, J. Chem. Phys. **128**, 044308 (2008).

[25] S. Miura and J. Tanaka, J. Chem. Phys. **120**, 2160 (2004).

[26] S. Miura, J. Chem. Phys. **126**, (2007).

[27] S. Miura, J. Chem. Phys. **126**, (2007).

[28] S. Miura, J. Phys-Condens. Mat. **20**, (2008).

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

[30] Y. Mizumoto and Y. Ohtsuki, Chem. Phys. Lett. **501**, 304 (2011).

[31] M. Herman, E. Bruskin, and B. Berne, J. Chem. Phys. **76**, 5150 (1982).

[32] M. Allen and D. Tildesley, *Computer simulation of liquids* (Clarendon Press, 1999).

[33] T. Schneider and E. Stoll, Phys. Rev. B **17**, 1302 (1978).

[34] T. Rog, K. Murzyn, K. Hinsen, and G. Kneller, J. Comp. Chem. **24**, 657 (2003).

[35] W. Humphrey, A. Dalke, and K. Schulten, Journal of Molecular Graphics **14**, 33 (1996).

[36] H. Li and R. J. Le Roy, Phys Chem Chem Phys **10**, 4128 (2008).

[37] R. J. Le Roy, *betaFIT 2.0 A Computer Program to Fit Pointwise Potentials to Selected Analytic Functions* (University of Waterloo, 2009).

[38] R. A. Aziz, A. R. Janzen, and M. R. Moldover, Phys. Rev. Let. **74**, 1586 (1995).

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

[40] N. Blinov, X. Song, and P.-N. Roy, J. Chem. Phys. **120**, 5916 (2004).

[41] S. Moroni, N. Blinov, and P. N. Roy, J. Chem. Phys. **121**, 3577 (2004).

[42] N. Blinov and P.-N. Roy, J. Low Temp. Phys. **140**, 235 (2005).

[43] W. Topic et al., J. Chem. Phys. **125**, 144310 (2006).

[44] Y. J. Xu, N. Blinov, W. Jäger, and P. N. Roy, J. Chem. Phys. **124**, 081101 (2006).

[45] N. Blinov and P.-N. Roy, ACS Symposium Series (Oxford University Press) **953**, 165 (2007).

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

[47] L. Wang et al., J. Mol. Spec. **267**, 136 (2011).

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

[49] X. Li and J. Broughton, J. Chem. Phys. **86**, 5094 (1987).

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

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

[52] D. Thirumalai and B. Berne, Comput Phys Commun **63**, 415 (1991).

[53] J. Cao and G. Voth, J. Chem. Phys. **100**, 5093 (1994).

[54] J. Cao and G. Voth, J. Chem. Phys. **100**, 5106 (1994).

[55] J. Cao and G. Voth, J. Chem. Phys. **101**, 6157 (1994).

[56] J. Cao and G. Voth, J. Chem. Phys. **101**, 6168 (1994).

[57] S. Jang and G. Voth, J. Chem. Phys. **111**, 2371 (1999).

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

[59] B. J. Braams and D. E. Manolopoulos, J. Chem. Phys. **125**, 124105 (2006).

[60] T. Hone, P. Rossky, and G. Voth, J. Chem. Phys. **124**, (2006).

[61] A. Yoshimori, J. Chem. Phys. **128**, (2008).

[62] T. D. Hone, J. A. Poulsen, P. J. Rossky, and D. E. Manolopoulos, J. Phys. Chem. B **112**, 294 (2008).

[63] A. Pérez, M. E. Tuckerman, and M. H. Müser, J. Chem. Phys. **130**, 184105 (2009).

[64] A. Witt, S. D. Ivanov, M. Shiga, H. Forbert, and D. Marx, J. Chem. Phys. **130**, (2009).

[65] L. de la Pena, M. Razul, and P. Kusalik, J. Chem. Phys. **123**, (2005).

[66] F. Paesani and G. A. Voth, J. Chem. Phys. **132**, (2010).

[67] D. T. Major and J. Gao, J. Chem. Theory Comput. **3**, 949 (2007).

[68] E. A. Polyakov, A. P. Lyubartsev, and P. N. Vorontsov-Velyaminov, J. Chem. Phys. **133**, (2010).

[69] K. Kinugawa, P. Moore, and M. Klein, J. Chem. Phys. **106**, 1154 (1997).

[70] Q. Shi and E. Geva, J. Chem. Phys. **116**, 3223 (2002).

[71] T. F. Miller and D. E. Manolopoulos, J. Chem. Phys. **122**, 184503 (2005).

[72] T. Miller and D. Manolopoulos, J. Chem. Phys. **123**, (2005).

[73] T. E. Markland, S. Habershon, and D. E. Manolopoulos, J. Chem. Phys. **128**, (2008).

[74] I. Craig and D. Manolopoulos, Chem Phys **322**, 236 (2006).

[75] I. Craig and D. Manolopoulos, J. Chem. Phys. **122**, 084106 (2005).

[76] I. Craig and D. Manolopoulos, J. Chem. Phys. **123**, (2005).

[77] A. Kaczmarek, M. Shiga, and D. Marx, J Phys Chem a **113**, 1985 (2009).

[78] T. F. I. Miller, J. Chem. Phys. **129**, (2008).

[79] A. R. Menzeleev and T. F. I. Miller, J. Chem. Phys. **132**, (2010).

[80] S. Habershon, B. J. Braams, and D. E. Manolopoulos, J. Chem. Phys. **127**, (2007).

[81] S. Habershon and D. E. Manolopoulos, J. Chem. Phys. **131**, (2009).

[82] R. Collepardo-Guevara, Y. V. Suleimanov, and D. E. Manolopoulos, J. Chem. Phys. **130**, (2009).

[83] Y. V. Suleimanov, R. Collepardo-Guevara, and D. E. Manolopoulos, J. Chem. Phys. **134**, (2011).

[84] T. E. Markland et al., Nat Phys **7**, 134 (2011).

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

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

[87] D. A. Case et al., J. Comput. Chem. **26**, 1668 (2005).

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

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

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

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

[92] K. Schweizer, R. Stratt, D. Chandler, and P. Wolynes, J. Chem. Phys. **75**, 1347 (1981).

[93] P. Brennan, Tuberculosis **83**, 91 (2003).

[94] P. Brennan and H. Nikaido, Annu Rev Biochem **64**, 29 (1995).

[95] C. Barry, Biochem Pharmacol **54**, 1165 (1997).

[96] M. Seo et al., J. Chem. Theory Comput. **4**, 184 (2008).

[97] H. A. Taha, N. Castillo, P.-N. Roy, and T. L. Lowary, J. Chem. Theory Comput. **5**, 430 (2009).

[98] H. A. Taha et al., J. Chem. Theory Comput. **6**, 212 (2010).

[99] H. A. Taha, P.-N. Roy, and T. L. Lowary, J. Chem. Theory Comput. **7**, 420 (2011).

[100] K. N. Kirschner et al., J. Comput. Chem. **29**, 622 (2008).

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

[102] S. Y. Y. Wong, D. M. Benoit, M. Lewerenz, A. Brown, and P.-N. Roy, J. Chem. Phys. **134**, 094110 (2011).

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

[104] M. Pavese and G. A. Voth, Chemical Physics Letters **249**, 231 (1996).

[105] M. S. Friedrichs et al., J. Comp. Chem. **30**, 864 (2009).

[106] J. A. Anderson, C. D. Lorenz, and A. Travesset, J. Comp. Phys. **227**, 5342 (2008).

[107] Z. Li et al., J. Chem. Phys. **128**, 224513 (2008).

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

[109] M. E. Tuckerman, Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms, Lecture Notes **10**, 269 (2002).

[110] A. M. Reilly, S. Habershon, C. A. Morrison, and D. W. H. Rankin, J. Chem. Phys. **132**, (2010).

[111] M. Shiga, M. Tachikawa, and S. Miura, J. Chem. Phys. **115**, 9149 (2001).

[112] S. Miura, M. Tuckerman, and M. Klein, J. Chem. Phys. **109**, 5290 (1998).

[113] S. Miura, S. Okazaki, and K. Kinugawa, J. Chem. Phys. **110**, 4523 (1999).

[114] M. Shiga and A. Nakayama, Chem Phys Lett **451**, 175 (2008).

[115] R. Collepardo-Guevara, I. R. Craig, and D. E. Manolopoulos, J. Chem. Phys. **128**, (2008).

[116] E. Geva, Q. Shi, and G. Voth, J. Chem. Phys. **115**, 9209 (2001).

[117] Y. Kwon, D. Ceperley, and K. Whaley, J. Chem. Phys. **104**, 2341 (1996).

[118] S. Baroni and S. Moroni, Phys. Rev. Lett. **82**, 4745 (1999).

[119] W. Foulkes, L. Mitas, R. Needs, and G. Rajagopal, Rev Mod Phys **73**, 33 (2001).