A Feynman Path Centroid Effective Potential Approach for the Study of Low Temperature Parahydrogen Clusters and Droplets

by

Jing Yang

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

Waterloo, Ontario, Canada, 2012

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

The quantum simulation of large molecular systems is a formidable task. We explore the use of effective potentials based on the Feynman path centroid variable in order to simulate large quantum clusters at a reduced computational cost. This centroid can be viewed as the "most" classical variable of a quantum system. Earlier work has shown that one can use a pairwise centroid pseudo-potential to simulate the quantum dynamics of hydrogen in the bulk phase at 25 K and 14 K [Chem. Phys. Lett. 249, 231, (1996)]. Bulk hydrogen, however, freezes below 14 K, so we focus on hydrogen clusters and nanodroplets in the very low temperature regime in order to study their structural behaviours. The calculation of the effective centroid potential is addressed along with its use in the context of molecular dynamics simulations. The effective pseudo-potential of a cluster is temperature dependent and shares similar behaviour as that in the bulk phase. Centroid structural properties in three dimensional space are presented and compared to the results of reference path-integral Monte Carlo simulations. The centroid pseudo-potential approach yields a great reduction in computation cost. With large cluster sizes, the approximate pseudo-potential results are in agreement with the exact reference calculations. An approach to deconvolute centroid structural properties in order to obtain real space results for hydrogen clusters of a wide range of sizes is also presented. The extension of the approach to the treatment of confined hydrogen is discussed, and concluding remarks are presented.

Acknowledgements

First of all, I would like to thank my master's supervisor Dr. Pierre-Nicholas Roy who guided my step into the science research world. His encouragement and enthusiasm for the research have always been the impetus of the whole research group, including me. I started research with him in my last undergraduate year. I was always wondering what would be the pragmatic significance of the theoretical research since everything in the research has been made ideally and approximately. Pierre told me that not everything is born with its current use in the real life. In the most of the cases, the real use happens after people making the discovery. His words are imprinted in my mind. Since then, I decided to be the one to discover. Within the two-year research, trying to be a real discoverer, I learned a lot from Pierre, including things as simple as plotting a curve and as complicated as developing a computational method. I am glad I could have him as my torchbearer of the science world!

In addition, my degree would not be completed without the help from Dr. Robert J. Le Roy and Dr. Marcel Nooijen. They are both the members of my committee. Dr. Robert J. Le Roy, also known as Bob, was the undergraduate advisor of my academic program. I thank him for bringing me into the computational chemistry world. Also, during my potential calculation process, the BetaFit code from Bob has been a most useful potential-fitting program for my project. Without such a handy fitting program, the

further simulations could not be performed easily. Dr. Marcel Nooijen has always been the most careful and the most thorough thinking person. He can point out detailed problems that we missed, which are usually the essential points. Together with the suggestions from Dr. Robert J. Le Roy and Dr. Marcel Nooijen, I could finally finish my thesis.

Finally, I would like to thank my colleagues, Chris Ing, Steve Constable, Hui Li, and Tao Zeng, who have been helping me on code debugging and making enlightened suggestions on my project. In addition, I would like to thank two of my thesis proofreaders, Dr. Nike Dattani and, my brother in law, Dr. Stefan Wippermann, who spent days and nights reading my thesis and making corrections and comments. At the end, I would like to thank the whole theoretical chemistry group that provides me with a platform for science research and creates such a friendly research environment.

Dedication

To my parents who always support me whenever I need!

Table of Contents

List of Tables

List of Figures

List of Abbreviations

^{4}He	Helium-4
2D	Two Dimensions or two-dimensional
3D	Three Dimensions or three-dimensional
CMD	Centroid Molecular Dynamics
DMC	Diffusion Monte Carlo
DVR	Discrete Value Representation
EMO	Expanded Morse Oscillator
LJ	Lennard-Jones
MLR	Morse Long-Range
MMTK	Molecular Modelling ToolKit
NMM	Numerical Matrix Multiplication
p-H_2	Parahydrogen
PDF	Pair Distribution Function
PIGS	Path Integral Ground State

PIMC	Path Integral Monte Carlo
PMF	Potential of Mean Force
QDO	Quasi-Density Operator

Chapter 1

Introduction

The study of clusters has been an active field of research over the past half century. The discovery of clusters has lead to a new research world due to the unique property of these systems, which are neither "large" molecules nor a partial portion of the bulk [1]. Both theoretical and experimental research endeavours related to molecular clusters have uncovered many features that are not present in the bulk phase. The special properties of clusters is of importance in various research areas, such as surface science and catalysis[2]. Understanding the structural properties of clusters is a basic and essential starting point to unlock the mysteries of other features. This fact is true not only for studies of clusters, but also for the investigation of most atomic and molecular systems [3]. Earlier work has used ab initio calculations to define the general structures of noble gas clusters and

metal clusters. The rare-gas clusters show an icosahedral geometry [4], while metal clusters display more complicated morphologies such as icosahedral, decahedral and closed-packed structures [3]. In addition, clusters with some specific sizes tend to occur more frequently than to other sizes [5]. Those specific sizes are so-called "magic numbers". The occurrence of "magic numbers" is due to the fact that clusters with these sizes have lower building energies and are hence more stable and more compact in structure. Moreover, the "magic number species" adopt structures similar to those of nuclear shells [1]. Different types of clusters possess different "magic number species". For example, sodium clusters have magic numbers of 2, 8, 20, and 40 [6], whereas noble gas clusters such as Xenon (Xe) have magic numbers of 13, 19, 55, and 147 [7]. In the area of chemical physics, noble gas clusters are of special interest. In addition, the presence of dopants can be another factor that can influence the structural stability of clusters [1]. Studies of noble gas clusters with metal ion dopants in the form MX_n (M is metal ion, X=noble gas atom) have shown that the dopant size can affect the cluster structure by shifting the magic numbers [8].

After a series of sophisticated studies on noble gas systems, scientists turned their attention to lighter systems such as helium-4 (^{4}He) and parahydrogen (p-H_2). The first attention was drawn to the ^{4}He clusters due to their unique features at low temperature. Previous work had found that ^{4}He clusters are liquid at any temperature except when pressure is large enough, such as greater than 25 atm at 0 K where ^{4}He clusters become

solid [9]. Additionally, ^{4}He has simple electronic structure and its intermolecular potential is well known [10]. These features facilitate accurate theoretical calculations [11]. However, such high accuracy calculations are computationally demanding, because it is necessary to include nuclear quantum effects due to the low mass of the constituent atoms. Quantum calculation methods are required to include these effects and tackle associated phenomena such as superfluidity [12]. Superfluidity is a phenomenon where a substance, brought to a sufficiently low temperature, loses its viscosity. The substance can undergo "fountain" effects and tends to escape from a narrow capillary [13]. Furthermore, ^{4}He was found to be the first chemical substance to exhibit such a behaviour when the temperature goes down to 2.12 K [14]. In addition to ^{4}He, clusters composed of p-H_2 are also of great interest since they are potential candidates for the observation of superfluid effects [15].

There are various ways to calculate the properties of quantum clusters. One is to directly solve the Schrödinger equation by diagonalization to obtain energy states for both the bound states and the continuum. Those can then be used to directly construct the partition function. Once the partition function is known, any thermodynamic property can in principle be calculated. However, this exact approach only allows one to treat systems with 5 atoms or less. The computational cost of the exact diagonalization approach scales exponentially with the size of the system. For large many-body systems, the path integral formalism of Feynman [16] is an attractive alternative for calculating equilibrium properties

of systems such as p-H_2 clusters of various sizes, together with the inclusion of Bose-Einstein exchange effects [17]. The path integral Monte Carlo (PIMC) method is the tool of choice to obtain structural, energetic, and superfluidic properties of quantum systems such as ^{4}He clusters and p-H_2 clusters [11, 18, 12]. In addition, the path integral ground state (PIGS) method allows one to calculate the ground-state energy for many quantum systems [19]. The method was tested on finite-sized p-H_2 clusters with sizes in the range of $N = 2 - 20$ [20]. Path integral based simulation techniques are very well suited for large many-body quantum systems, but their use is still restricted to some finite system sizes due to the limitations of contemporary computer hardware.

In the thesis, a so-called centroid effective potential approach, based on a path integral formalism, will be introduced as a solution to circumvent the limitation of PIMC-like methods. The goal of the present work is to develop a method that allows one to obtain equilibrium properties of weakly bound many-body systems such as p-H_2 clusters at low temperatures for a broad range of sizes from clusters ($2 - 100$) to nanoscale droplets (up to 10^5 particles). The principal aim of this research is to establish the applicability of the effective centroid potential method for quantum p-H_2 clusters at a low temperature interval, 3 K<T<10 K. This achievement would be an important contribution since such effective potential approaches have been limited to bulk systems such as liquid p-H_2 at higher temperatures (14 K and 25 K) [21], where quantum effects are of less importance. Whether

such centroid pseudo-potential approach is valid for even low temperature with strong quantum effects remains unknown. We choose 10 K as the upper boundary is because the clusters may evaporate at the temperature higher than 10 K. Also, the permutation of the paths is suppressed when temperature is higher than 2.5 K [17], so that, the lower bound is larger than 2.5 K to exclude the exchange effects. The system is under Boltzmann statistics. To test the quality of the centroid effective potential approach, structural properties of p-H_2 clusters will be compared to their, in principle, exact counterparts. The essential elements of the theoretical formalism are presented in the forthcoming sections.

1.1 N-body problem

The study of the N-body problem, also referred to as the many-body problem, in quantum mechanics has some of its roots in the study of celestial interactions among planets in Newton's time. The gravitational field was the main focus in those years. Kepler's three laws can easily explain the motions of a system with two objects, including the description of the trajectory of each object and the interaction between the two [22]. In classical mechanics, the centre of mass motion was introduced when dealing with the two-body problem. The basic approach was to find a way to reduce the two-body problem to an effective one-body problem. Also, rather than absolute positions ($\vec{r}_1$, $\vec{r}_2$ corresponding to

the origin), the relative position, $\vec{r} = \vec{r}_2 - \vec{r}_1$, is more convenient when dealing with the two-body problem. Together with the centre of mass position $\vec{R}_{cm}$, the coordinates of the system can be properly represented as follows,

$$\vec{r}_1 = \vec{R}_{cm} - \left(\frac{m_2}{m_1 + m_2}\right)\vec{r}, \tag{1.1}$$

$$\vec{r}_2 = \vec{R}_{cm} + \left(\frac{m_1}{m_1 + m_2}\right)\vec{r}, \tag{1.2}$$

where m_1 and m_2 are the masses of the two particles. Under classical conditions, the total kinetic energy $\mathrm{KE}_{\mathrm{tot}}$ can be represented as,

$$\mathrm{KE}_{\mathrm{tot}} = \frac{1}{2}m_1|\vec{v}_1|^2 + \frac{1}{2}m_2|\vec{v}_2|^2 = \frac{1}{2}m_1|\dot{\vec{r}}_1|^2 + \frac{1}{2}m_2|\dot{\vec{r}}_2|^2, \tag{1.3}$$

or in relative representation,

$$\mathrm{KE}_{\mathrm{tot}} = \frac{1}{2}\left(m_1 + m_2\right)|\dot{\vec{R}}_{cm}|^2 + \frac{1}{2}\left(\frac{m_1 m_2}{m_1 + m_2}\right)|\dot{\vec{r}}|^2. \tag{1.4}$$

From the above equation, the second term corresponding to $|\dot{\vec{r}}|^2$ contains a special mass expression $\frac{m_1 m_2}{m_1+m_2}$. This mass is called reduced mass and is represented by the Greek symbol

μ,

$$\mu = \frac{m_1 m_2}{m_1 + m_2}. \tag{1.5}$$

The extension of the above ideas to the N-body problem is a very difficult task.

In quantum mechanics, the N-body problem is also referred to as the many-body problem. A two-body system such as, for instance, a diatomic molecule has already been solved. Even with diatomic systems, we have to consider vibrational and rotational motions, binding energies, and quantization features. With a system size greater than three, not only does the interaction between each pair of particles need to be addressed, but the exchange among all the particles and collisions should also be considered. Furthermore, those features are the limited issues we can encounter right now. To understand the whole many-body systems remains a challenge.

1.2 Approaches to solve the many-body problem

In the past century, theories and methods have been developed to tackle the quantum many-body problem. Those approaches include the Hartree-Fock equations [23] to calculate ground-state properties, perturbation theory based approaches, and Greens Function techniques to study excited states [24]. Due to computational limitations, directly solving the Schrödinger equation has always been a dilemma for many-body systems. The path

integral method from Feynman [16] opened the door to contemporary research on many-body bosonic systems. In our research, we combine approaches based on the direct solution of the Schrödinger equation with the concept of the centroid of a Feynman path in order to reduce computational cost while maintaining a high degree of accuracy.

In general, assuming the potential is only distance dependent, which can be represented by $U(x)$, the generic form of the Schrödinger equation for each particle can be written as,

$$\hat{H}\psi(x) = E\psi(x), \tag{1.6}$$

where E is the total energy for a single particle, $\hat{H}$ is the Hamiltonian of the system including both the kinetic and potential energy components. The Hamiltonian is,

$$\hat{H} = \hat{T} + \hat{V}. \tag{1.7}$$

The operator for the kinetic part $\hat{T}$ is,

$$\hat{T} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}, \tag{1.8}$$

for a single particle, and

$$\hat{T} = -\frac{\hbar^2}{2M}\sum_{i=1}^{N}\frac{\partial^2}{\partial x_i^2}, \tag{1.9}$$

for a system with N particles, where $\hbar$ is the reduced Planck's constant, M is the total mass of the system ($M = \sum_{i=1}^{N} m_i$, m_i is the mass for each particle), ψ is the wavefunction, and x is the distance between the particle and the origin. The potential part is just

$$\hat{V} = U(x). \tag{1.10}$$

Solving the Schrödinger equation is equivalent to deal with an eigenvalue problem, which requires diagonalizing the $\hat{H}$ matrix on the left. It is known that the diagonalization process can cost $O(n^3)$ of the CPU time. Even worse, as mentioned before, the calculation cost of finding all the wave functions will increase exponentially with respect to the number of basis sets in each wave function. A new method is required.

1.2.1 Feynman path integral in statistical mechanics

Feynman's path integral formalism provides an alternative way to treat quantum systems without requiring the solution of Schrödinger's equation [16]. When dealing with quantum systems, the uncertainty principle cannot be ignored. We tend to make our statement more accurate by studying the probability of occurrence of any motions related to quantum particles rather than their actual motions. In the case of quantum statistical mechanics, the partition function is the central quantity. It is the sum over all the possible occupations

corresponding to all the energy states. The form of the partition function for a canonical system is

$$Z = Tr\{e^{-\beta\hat{H}}\} = \sum_n e^{-\beta E_n}, \tag{1.11}$$

where $\beta = \frac{1}{k_B T}$, and k_B is the Boltzmann constant, and T is the temperature. The energy values, E_n are the eigenvalues of $\hat{H}$.

Rather than solving the eigenproblem, in the path integral approach, the partition function for a single quantum particle can be viewed as a summation of many small pieces. Each piece shares the same action. Before getting into any path integral formulation, it is necessary to make a proper representation of the system regarding path integrals. The so-called "path" can be viewed as one of many, given a number P, pieces which comprise the whole system. Taking a one-particle system as an example, the momentum p_i, and position, q_i, for each piece within the system can be represented as,

$$\begin{aligned} \mathbf{p} &= (p_1, \ldots, p_P), \\ \mathbf{q} &= (q_1, \ldots, q_P). \end{aligned} \tag{1.12}$$

Now, we write the partition function as shown above into a discretized path integral form [25],

$$Z = \lim_{P\to\infty} \left(\frac{mP}{2\pi\hbar^2\beta}\right)^{P/2} \int \mathrm{d}q_1 \cdots \int \mathrm{d}q_p \ \exp[-\beta V_p(\mathbf{q})], \tag{1.13}$$

where V_p is called the isomorphic quasiclassical polymer potential, while P is the number of quasiparticles (or beads) in the path. The exact form of V_p includes the spring term from the neighbouring particle and the interaction potential,

$$V_p(\mathbf{q}) = \sum_{i=1}^{P} \left[\frac{mP}{2\hbar^2\beta}(q_i - q_{i+1})^2 + \frac{V(q_i)}{P} \right]. \tag{1.14}$$

The interaction potential here is approximated as a pair-wise additive potential, analogous to that in classical systems. One important feature that should be mentioned is that the "path" is always a closed path, or a ring polymer since a trace is evaluated. The number P is also related to the thermal time [26] interval $\epsilon = \beta\hbar/P$. The fact that it is a closed path can be written as

$$q_{i+P} = q_i. \tag{1.15}$$

Furthermore, the parameter P should be large enough in order to achieve convergence [25, 27]. The finite P approximation comes from the Trotter factorization [28].

The path integral formulation, indeed, provides a more practical method for calculating equilibrium properties of many-body quantum systems than the direct solution of the Schrödinger's equation. However, difficulties still persist for the simulation of large systems [25]. Different physical properties reach their own convergence differently. We must find convergence separately when studying various features. In general, a large number of beads

(large P) and long computer simulations are required to achieve convergence. For larger systems, achieving convergence requires a considerable computational effort.

1.2.2 Reformulation of statistical mechanics in terms of Feynman's path centroid variables

To extend calculations to larger systems, Feynman also proposed another theory in which one defines the mean components of the ring polymer termed the "centroid variable" [16]. When going from path integral variables to centroid variables, the many-beads representing the motion of a quantum particle are averaged into a unique pseudoparticle, the centroid. The path centroid variable, q_0 (Figure (1.1), is defined as,

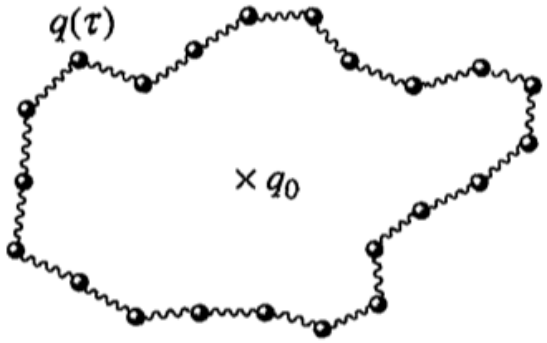

Figure 1.1: A ring polymer and its centroid value [25].

$$q_0 = \frac{1}{\beta\hbar}\int_0^{\beta\hbar} q(\tau)\,\mathrm{d}\tau, \tag{1.16}$$

where τ is the *imaginary time*, and $q(\tau)$ is the position at *imaginary time* τ. A more visual representation of the above equation is that the q_0 is equal to the average of positions over a circular trajectory of the particle travelling through a thermal time interval, $\beta\hbar$. For a finite number of quasiparticles, the path centroid variable can also be written as,

$$q_0 = \frac{1}{P}\sum_{i=1}^{P} q_i. \tag{1.17}$$

A functional path integral representation of the partition function as opposed to Eq.(1.13) is,

$$Z = \int \cdots \int \mathscr{D}q(\tau) \exp\left\{-S[q(\tau)]/\hbar\right\}, \tag{1.18}$$

in which $S[q(\tau)]$ is the *imaginary time action*, and $\mathscr{D}q(\tau)$ represents $dq_1 dq_2 \cdots dq_{P-1}$ [25]. According to Feynman, the centroid density is an important quantity in the study of finite temperature many-body quantum systems, and it allows a connection between quantum and classical mechanics [16]. A specific type of connection for structural properties between quantum mechanics and classical mechanics has been discovered already. Detailed information can be found in Chapter 3. Using the centroid variable, the partition function can be obtained by integrating the centroid density. The centroid density, $\rho_c(q_c)$, functional

can be defined as,

$$\rho_c(q_c) = \int \cdots \int \mathscr{D}q(\tau)\delta(q_c - q_0)\exp(-S[q(\tau)]/\hbar), \tag{1.19}$$

and the partition function in terms of the centroid variable

$$Z = \int \mathrm{d}q_c \ \rho_c(q_c). \tag{1.20}$$

In this research, we will discuss the centroid pseudo-potential approach related to centroid variables, especially in the study of p-H_2 clusters. The centroid variables are considered to be the most classical variables in the quantum system [16]. We tend to treat the quantum system in a classical way with the effective pseudo-potential to take care of the quantum effects. Since the calculation related to centroid variables is related to averages of many thermal properties such as kinetic energy and potential energy, the process may be a lot simpler than directly solving the Schrödinger equation or a full path integral simulation. It is expected that the centroid based method will provide better convergence and quicker computation time compared to path integral methods. Therefore, larger systems can be calculated within a shorter time limit without compromising the accuracy of the calculation.

1.3 Overview of Thesis

By thoroughly studying the centroid approach, centroid pairwise potentials for different temperatures are generated from the centroid pair density. Then, the potentials are implemented in an open-source molecular dynamics simulation tool, the Molecular Modelling ToolKit (MMTK) [29]. The centroid pair distribution functions of p-H_2 clusters are obtained by carrying out molecular dynamics simulations using MMTK and implemented centroid potentials. Prior to this research, low temperature structural properties of p-H_2 clusters had not been studied using centroid approaches. Only a few studies had been performed in this field, and applications were limited to bulk liquid p-H_2 [25, 30, 31, 32, 21]. The present research provides structural information for p-H_2 clusters at temperatures ranging from 3 K to 6 K with sizes from N= 4 up to N= 400. The simulation time is less than a day for the largest size of the cluster. In comparison, path integral methods may take several weeks, or even months.

In Chapter 2, the way of obtaining the centroid pair-potential for p-H_2 is presented, and the centroid density for the p-H_2 dimer is calculated. Different methods for density calculations, such as the Numerical Matrix Multiplication (NMM) [33] and PIMC methods are used. The NMM method is only practical for one-dimensional Hamiltonians, whereas the PIMC method works well for one, two and three dimensional systems. In Chapter 3,

the details of the molecular dynamics simulation with a centroid potential are illustrated. The centroid pair distribution functions (so-called $g(r)$) are extracted from the simulation. We compare the centroid $g(r)$ from centroid molecular dynamics (CMD) simulations and the centroid $g(r)$ from exact PIMC simulations. Furthermore, the connection between quantum mechanics and classical mechanics related to structural properties is explained in terms of a deconvolution kernel. It contains the information of both the centroid density and the real space quantum density. In that way, we can transform real-space densities to centroid densities and vice versa. Finally, in Chapter 4, conclusions and proposed future directions are presented.

Chapter 2

Centroid Pair-potential of the p-H_2 dimer

Feynman first suggested the use of centroid variables [16]. The concept was further developed in the study of p-H_2 with a lithium impurity [34], and the explicit formulations involving the centroid density soon followed [25, 30, 31, 32]. The corresponding equilibrium and dynamical properties were studied. The main idea of making use of centroid variables for the ensemble average of a quantum system [21] is to reduce the cost of quantum dynamics simulations. The centroids can form the classical variables of the quantum system. Therefore, the quantum simulation can be computed in a classical way. Similar to the classical dynamics, centroid molecular dynamics contains the free particle action and po-

tential action, but this potential is a unique temperature-related effective one. The unique potential is named centroid pseudo-potential. The exact pseudo-potential corresponds to the centroid potential of mean force (PMF) and is a many-body quantity that requires a full path integral treatment. To greatly reduce computational cost, it was suggested that a pairwise additive approximation could be made [21]. This approximation was found to be excellent for bulk p-H_2 at 14 K and 25 K [21]. However, bulk p-H_2 solidifies below 14 K. Since then, the pseudo-potential approach for p-H_2 has not been tested at lower temperatures where quantum effects are of greater importance. As opposed to the bulk form, p-H_2 clusters are believed to remain liquid-like down to very low temperatures [17]. In this thesis, the validity of the pseudo-potential in terms of centroids is tested in the low temperature regime for various sizes of p-H_2 clusters. As the pseudo-potential is related to the centroid density, we present multiple ways of computing the centroid pair-density in order to get the centroid potential efficiently. Once the centroid potential is obtained, the potential fitting to a proper analytical form is performed. The temperature related feature of the pseudo-potential will be studied.

2.1 Theory

2.1.1 Formal definition of the centroid potential

The approximate centroid potential, or the pseudo-potential, is defined as the PMF of the quantum system. We obtain the pseudo-potential by first calculating the centroid pair-density. As shown in the introduction, the partition function for the quantum system can be represented as the integral of the centroid pair-density,

$$Z = \int \mathrm{d}r_c \; \rho_c(r_c), \tag{2.1}$$

where r_c is the relative distance between any two centroid pseudoparticles in the system, and $\rho_c(r_c)$ is the centroid density. Instead of the expression shown in the introduction, a new representation of the centroid density is formed by using the Fourier integral,

$$\rho_c(r_c) = \int \mathrm{d}r \langle r| \int \frac{\mathrm{d}\vec{k}}{2\pi} \exp(i\vec{k}\cdot\vec{r_c}) \exp(-\beta\hat{H} - i\vec{k}\cdot\hat{r})|r\rangle, \tag{2.2}$$

where k is the Fourier basis vector, $\hat{H}$ is the Hamiltonian operator which contains kinetic energy and potential energy operators. The centroid potential is related to the logarithm

of the centroid density, as shown in the following equation,

$$V_c(r_c) = -k_B T \ln \left[\rho_c(r_c)/\rho_c^0\right], \tag{2.3}$$

where ρ_c^0 is the free particle centroid density [32, 21] given in form of

$$\rho_c^0 = \left[\prod_{i=1}^{N} \frac{m_i}{2\pi\hbar^2\beta}\right]^{3/2}, \tag{2.4}$$

where $\hbar$ is the reduced Plank's constant and β is $\frac{1}{k_B T}$.

The centroid potential problem boils down to finding $\rho_c(r_c)$. The centroid density given in Eq. (2.2) can be rewritten as,

$$\rho_c(r_c) = \int \frac{\mathrm{d}\vec{k}}{2\pi} \exp(i\vec{k}\cdot\vec{r_c}) \int \mathrm{d}r \langle r| \exp(-\beta\hat{H} - i\vec{k}\cdot\hat{r})|r\rangle. \tag{2.5}$$

Since the term $-i\vec{k}\cdot\hat{r}$ is only distance dependent, it can be recognized as a part of the

effective potential. The $\hat{H}$ can be rewritten in a complex form,

$$\begin{aligned}\hat{H}'(\hat{r},\vec{k}) &= \hat{H} + \frac{i\vec{k}\cdot\hat{r}}{\beta} \\ &= \hat{T} + \hat{V}(\hat{r}) + \frac{i\vec{k}\cdot\hat{r}}{\beta} \\ &= \hat{T} + \hat{V}'(\hat{r}). \end{aligned} \tag{2.6}$$

Now, the centroid density in the form of a complex $\hat{H}'$ is,

$$\rho_c(r_c) = \int \frac{\mathrm{d}\vec{k}}{2\pi} \exp(i\vec{k}\cdot\vec{r_c}) \int \mathrm{d}r \big\langle r| \exp\big(-\beta\hat{H}'(\hat{r},\vec{k})\big)|r\big\rangle. \tag{2.7}$$

2.1.2 Various forms of of the centroid potentials

Previous work has shown that linear algebra schemes can be used to calculate the density if we treat the density operator as a matrix [33]. The NMM method allows us to obtain the 1D centroid density as well as the corresponding 1D effective potential [35]. However, the many-body quantum system includes the actions of not only the linear motions which depend on distances between any two particles, but also the rotational or vibrational motions. Therefore, a 1D description is not sufficient to represent the system completely. The potential with two dimensions (2D) and three dimensions (3D) coordinates should be taken into account. The NMM could still be one of the ways to compute potential with 2D

or 3D coordinates. However, the complexity of the calculation becomes a big challenge, since the coordinates themselves are vectors that contain multi-dimensional elements.

2.1.3 NMM-1D potential related density calculation

The 1D centroid potential can be obtained by solving the 1D centroid density matrix related to the relative distances r_c between two centroid positions q_{c_i} and q_{c_j}. By taking the trace of the operator in 1D, $e^{\beta\hat{H}'(\hat{r},k)}$, the centroid density can be easily extracted. In a mathematical form, it is written as,

$$\rho_c(r_c) = Tr\left[e^{-\beta\hat{H}'(\hat{r},\vec{k})}\right] = \langle r|e^{-\beta\hat{H}'(\hat{r},\vec{k})}|r\rangle. \tag{2.8}$$

The idea of NMM is still a part of the application of path integral techniques. Based on Trotter's factorization [28], the density matrix can be obtained by the iterative matrix-multiplication of many discretized sub-density-matrices in the path integral representation. We write the complex centroid density in terms of sub-density matrix, $\rho(r, r'; \beta)$,

$$\rho(r, r'; \beta) = \langle r|e^{-\beta\hat{H}'(\hat{r},\vec{k})}|r'\rangle, \tag{2.9}$$

where the initial condition applied for the sub-density matrix is,

$$\lim_{\beta \to 0} \rho(r, r'; \beta) = \delta(r - r'). \tag{2.10}$$

We divide the β in the density matrix into many small pieces with each piece as ϵ. According to the Trotter's factorization, the simple separation is,

$$e^{-2\epsilon H} = e^{-\epsilon H} \cdot e^{-\epsilon H}, \tag{2.11}$$

and the full multiplication formula is,

$$e^{-\beta H} = \underbrace{e^{-\epsilon H} \cdots e^{-\epsilon H}}_{2^n \text{ terms}}. \tag{2.12}$$

Therefore, the sub-density in terms of ϵ is,

$$\rho(r, r'; 2\epsilon) = \int_{-\infty}^{+\infty} \mathrm{d}r'' \rho(r, r''; \epsilon) \rho(r'', r'; \epsilon). \tag{2.13}$$

With $\beta = 2^n \epsilon$, it is analogous to solving the path integral problem with 2^n identical pieces. The full density can be worked out by iteratively multiplying the sub-density pieces.

In the Hamiltonian, the free particle motion or kinetic motion can be calculated directly.

We shall leave the free particle motion as a pre-calculated parameter in the core. The Trotter's factorization is applied only to the potential component (V', which contains the imaginary Fourier components). The equation for the first iteration is provided in following,

$$\begin{aligned}\rho^{(0)}(r,r';\epsilon) &= \exp[-\epsilon(T+V')] \\ &= \exp[-(\epsilon/2)V'(r)]\exp(-\epsilon T)\exp[-(\epsilon/2)V'(r'))] \\ &= \exp[-(\epsilon/2)V'(r)]\rho_0(r,r',;\epsilon)\exp[-(\epsilon/2)V'(r'))],\end{aligned} \tag{2.14}$$

where ρ_0 is the density matrix for the exact free particle motion [16],

$$\rho_0(r,r';\epsilon) \equiv \left[\frac{\mu}{2\pi\hbar^2\epsilon}\right]^{1/2}\left\{\exp\left[-\frac{\mu}{2\hbar^2\epsilon}(r-r')^2\right] - \exp\left[-\frac{\mu}{2\hbar^2\epsilon}(r+r')^2\right]\right\}, \tag{2.15}$$

and

$$V'(r) = V(r) + \frac{i\vec{k}\cdot\vec{r}}{\beta} \tag{2.16}$$

with $V(r)$ being any kind of distance dependent 1D potential. Previous work has used the Lennard-Jones (LJ) potential, the Silver-Goldman potential [36] and the Buck potential [37] to represent the p-H_2 system. In this thesis, the Buck potential is used as a reference. In Eq. (2.16), the second term involves the dot product of the wave vector **k** and the

relative distance vector **r**. Expanding the product, this formula is written as,

$$\frac{i\vec{k}\cdot\vec{r}}{\beta} = \frac{i|\vec{k}||\vec{r}|\cos\theta}{\beta}. \tag{2.17}$$

For simplification, we choose $\vec{r}$ along the direction of $\vec{k}$. Now, the equation above becomes,

$$\frac{i\vec{k}\cdot\vec{r}}{\beta} = \frac{i|\vec{k}||\vec{r}|}{\beta}. \tag{2.18}$$

The NMM method works really well on 1D centroid density matrix calculations. Our earlier project was to apply the NMM method to the 1D centroid density to obtain the corresponding 1D centroid potential.

2.1.4 NMM-3D potential

In order to make better comparison with PIMC calculation, which obtained properties in 3D, the pseudo-potential in our research should also be generated in 3D. A modified NMM method with a basis vector as a propagator is implemented. Unlike the NMM in 1D, the 3D case involves the rotational motion of the p-H_2 dimer. Treating the p-H_2 dimer as a single diatomic molecule, the modified potential $V'(r)$ taken account the rotational motion

can be written as,

$$V_{\text{eff}}(r) = \frac{l(l+1)\hbar^2}{2\mu r^2} + V'(r), \tag{2.19}$$

where l is the angular momentum and μ is the reduced mass of the p-H_2 dimer. It is necessary to mention that the imaginary part of $V'(r)$ could make the calculation rather complicated since the angle between **k** and **r** cannot be ignored. Therefore, spherical harmonic terms should be introduced into the system. Spherical coordinates may be used in the calculation in order to easily interact with angles. The spherical coordinate representation of each Cartesian component is shown below,

$$\begin{cases} x = r\sin\theta\cos\phi \\ y = r\sin\theta\sin\phi \\ z = r\cos\theta. \end{cases} \tag{2.20}$$

Referring to Eq. (2.7), we set the second integral as the density core, $F(k)$. The centroid density in spherical coordinates can be easily written as,

$$\rho_c(r_c, \theta_c, \phi_c) = \int k^2 \mathrm{d}k \mathrm{d}\cos\theta_\mathrm{k} \mathrm{d}\phi_k \exp(i\vec{k}\cdot\vec{r_c}) F(k, \theta_k, \phi_k), \tag{2.21}$$

where θ_c, ϕ_c and r_c are regarding to $\vec{r_c}$, and θ_k, ϕ_k and k a set of components for $\vec{k}$. Here, we arbitrarily choose the axes. The θ's and ϕ's are the angles between those main axes. Since quantum numbers are also involved in the potential form (Eq. (2.19)), it is necessary to convert the equation into a Discrete Value Representation (DVR) [38] which consists of a basis of $|r, l, m\rangle$ functions, where r is the relative distance between the two particles in the dimer, and l is the angular momentum with magnetic quantum number m. Therefore, $F(k)$ in terms of $|r, l, m\rangle$ with $\tau = \frac{\beta}{/} P$ should be,

$$F(k,\tau) = \int d\theta_r d\phi_r \langle r,l,m|Y^*_{lm}(\theta_r,\phi_r) \exp(-\frac{\tau}{2}\frac{i\vec{k}\vec{r}}{\beta}) \exp(-\frac{V_{\text{eff}}}{2}) \tag{2.22}$$

$$\rho_0$$

$$\exp(-\frac{V_{\text{eff}}}{2}) \exp(-\frac{\tau}{2}\frac{i\vec{k}\vec{r}}{\beta}) Y_{lm}(\theta_r,\phi_r)|m,l,r\rangle,$$

and

$$F(k) = \langle r,l,m| \underbrace{\exp(-\frac{\tau}{2}\frac{i\vec{k}\vec{r}}{\beta}) \exp(-\frac{V_{\text{eff}}}{2})|r',l',m'\rangle\langle m',l',r'| \exp(\cdots)}_{2^{P/2}}$$

$$\rho_0(r,r';\beta)$$

$$\underbrace{\exp(\cdots)|r',l',m'\rangle\langle m',l',r'| \exp(-\frac{V_{\text{eff}}}{2}) \exp(-\frac{\tau}{2}\frac{i\vec{k}\vec{r}}{\beta})}_{2^{P/2}} |m,l,r\rangle, \tag{2.23}$$

where θ_r and ϕ_r are the corresponding spherical coordinates for vector $\mathbf{r}$. Here, we set the basis vector as,

$$\mathbf{v} = |mlr\rangle. \tag{2.24}$$

and it is simply a set of unit vectors. The $|r', l', m'\rangle$ and $\langle r', l', m'|$ in the middle are the general representation for the identities with different indices inserted to separate the multiplication pieces. Instead of taking the multiplication among all the matrices in the middle, we calculate the vector-matrix multiplication from the right to the left. In that way, for every multiplication, we only need to store one vector in the memory. We are still using the NMM formulation, but with a different way of computational implementation.

However, the actual calculation is cumbersome and outside the scope of this work, as described below. The final equation for the density core in 3D is

$$\begin{aligned} F(k) = & \langle r, l, m|\rho_0|m', l', r'\rangle_{\delta_{rr'}\delta_{ll'}\delta_{mm'}} \\ & + \langle r, l, m| \exp(-\frac{l(l+1)\hbar^2}{2\mu r^2} - V(r))|m', l', r'\rangle_{\delta_{rr'}\delta_{ll'}\delta_{mm'}} \\ & + \langle r, l, m| \exp(-\frac{i\vec{k}\cdot\vec{r}}{2\beta})|m', l', r'\rangle_{\delta_{rr'}}. \end{aligned} \tag{2.25}$$

The $|mlr\rangle$ based transformation is entangled with angular momenta and spherical coordinates. The method turns out to be hard to implement.

2.1.5 PIMC calculated potential

The actual finite temperature centroid potential calculation is needed only for the p-H_2 dimer. Later, the dynamics can be performed by using the pre-calculated potentials for each temperature. The PIMC method [39, 40] is still time consuming, but works well for calculations using Cartesian coordinates. Furthermore, we only need to determine the pair-centroid density once for calculating the potential. It is not harmful to perform a complicated calculation on such a simple system. The method utilizes some random distributions to sample all the possible paths that the molecule could go through. Then we collect potential energy for each possible path to construct the centroid density matrix. The path integral form of the partition function is written as,

$$Z = \int \mathrm{d}q_0 \cdots \mathrm{d}q_{P-1} \exp\big[-\frac{\beta}{P} H'(\hat{r}, \vec{k})\big], \tag{2.26}$$

where q_0 to q_{P-1} describe the path of the molecule, and $H'(\hat{r}, \vec{k})$ is the complex Hamiltonian defined in Eq. (2.6). The centroid density can be obtained from fixing an arbitrary configuration such q_0. Based on Eq. (1.17), the centroid density is,

$$\rho_c(q_0) = \int \mathrm{d}q_1 \cdots \mathrm{d}q_{P-1} \exp(-\frac{\beta}{P} H'). \tag{2.27}$$

The potential is still in form of the potential of mean force. The final potential form, V_{eff}, can be written as,

$$V_{\text{eff}} = -\frac{1}{\beta} \ln \big[\frac{1}{N} \sum_{n=1}^{N} \exp(-\frac{\beta}{P} \sum_{i=1}^{P} \frac{V'(r_i)}{P}) \big], \tag{2.28}$$

where N is the number of Monte Carlo steps, P is the number of beads, $V'(r_i)$ is the Fourier form potential with respect to a specific relative distance between two centroid positions of p-H_2 molecules, r_i.

2.1.6 Path integral molecular dynamics calculated potential

The centroid potential can also be obtained from path integral molecular dynamics. The preceding work has already shown that the integration of the centroid force can provide the centroid potential for the many-body system [21]. The equation is shown in the following,

$$V_c(r_c) = \int^{r_c} \mathrm{d}r f_c(r), \tag{2.29}$$

where $f_c(r)$ is the average of the centroid force between two particles of distance of r. The path integral formulation with centroid constraints has already been successfully implemented in the MMTK [41]. Therefore, the centroid forces of the p-H_2 clusters can be obtained directly from MMTK. The exact many-body centroid potential can be recovered by applying Eq. (2.29).

2.2 Choice for constructing potentials

In our preliminary work, we were testing the centroid potentials in 1D. The NMM method worked well for calculating the centroid density and potential. Later, we extended our interest to 3D many-body systems. We attempt to use NMM in 3D is because the matrix multiplication method can provide a potential that is smooth and continuous. It can make the subsequent potential fitting work much simpler. However, the 3D NMM requires large memory storage, long computation time and big modifications of the original NMM method. Eventually, we adopted the PIMC method to obtain the potential due to its flexible computation of the pair-potential in any number of dimensions. The PIMC method provides the same potential as the one in NMM. Furthermore, the PIMC calculated potential shows better behaviour of the repulsive wall, whereas the one from NMM appears to be noisy in the short-range between 0 to $\sim$3.5 Å (Figure (2.1)).

2.3 Effective potential representation

The potentials from PIMC or NMM cannot be directly put into the dynamics simulation package because those potentials are limited to a specific range and have boundary problems. As shown in Figure (2.1), the short-range of the potential from NMM contains lots of noise. The potential with noise may cause the energy calculation to blow up

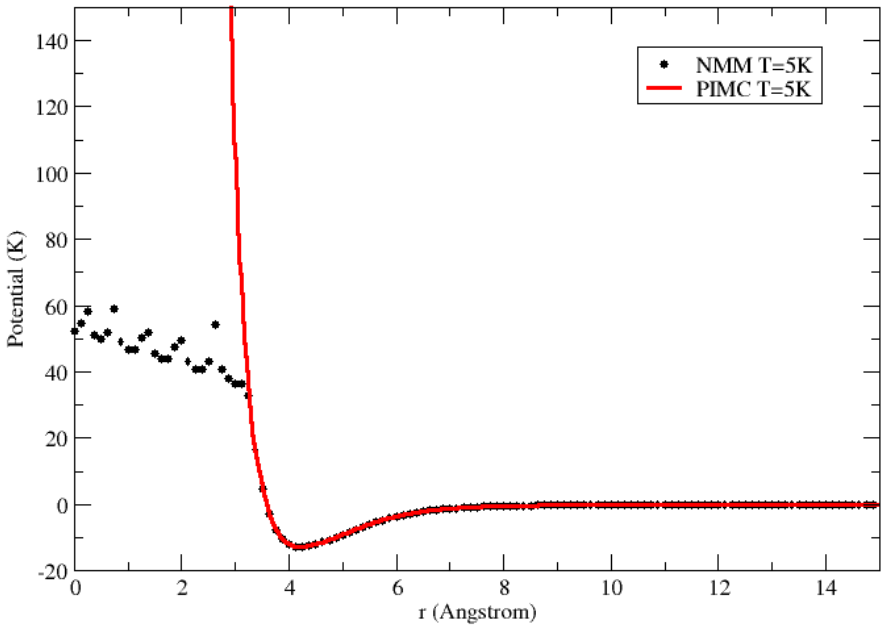

Figure 2.1: The 1D centroid pseudo-potential computed by NMM and PIMC at $T = 5$ K. The NMM and PIMC constructed potentials agree closely with each other except at short-range ($r < 3.5$ Å). The short-range of NMM becomes physically absurd. Note: both methods are obtained with the same number of beads.

when performing the molecular dynamics simulations later. Therefore, we need a smoother potential format with good behaviour in both the short-range and the long-range to represent the centroid potential. There are two types of potential forms we have been using, the Expanded Morse Oscillator (EMO) [42] and the Morse Long-Range (MLR) [43, 44].

2.3.1 EMO (previous work)

The EMO potential form is,

$$V_{\mathrm{EMO}}(r) = D_e \left[1 - e^{\beta(r)(r-r_e)}\right]^2 \tag{2.30}$$

where D_e is the potential well depth, r_e is the equilibrium distance, and $\beta(r)$ is called exponent coefficient function defined as,

$$\beta(r) = \sum_{i=0}^{N} \beta_i y_p^{\mathrm{ref}}(r)^i, \tag{2.31}$$

where β_i's are the β coefficients, and

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

with r_{ref} being the reference radius of r. The EMO form does not require any long-range terms in the potential, but it usually give a good overall shape of D_e, but incorrect long-range tail of the potential [45].

2.3.2 MLR (currently)

The MLR form contains the information of the long-range tail and can provide a better representation of our pseudo-potential than the EMO form. The form of the MLR is,

$$V_{\mathrm{MLR}}(r) = D_e \left[1 - \frac{u_{\mathrm{LR}(r)}}{u_{\mathrm{LR}}(r_e)} \exp\left[-\beta(r) \cdot y_p^{\mathrm{eq}}(r)\right]\right]^2, \tag{2.33}$$

here D_e and r_e have the same definition as in the EMO form. The radial variable is written as,

$$y_p^{\mathrm{eq}}(r) = \frac{r^p - r_e^p}{r^p + r_e^p}, \tag{2.34}$$

where p is some small integer. The long-range terms are contained inside of the u_{LR} term,

$$u_{\mathrm{LR}}(r) = \sum_{i=1}^{last} \frac{C_{m_i}}{r_{m_i}}, \tag{2.35}$$

in which C_m's are the long-range coefficients and i represents the index of the coefficient. The MLR form also contains the exponent coefficient function $\beta(r)$, but this $\beta(r)$ has a more complicated form than that in the EMO. This coefficient function is [43],

$$\beta(r) = y_p^{\mathrm{ref}}(r)\beta_\infty + \left[1 - y_p^{\mathrm{ref}}(r)\right] \sum_{i=0}^{N} \beta_i \left[y_q^{\mathrm{ref}}(r)\right]^i, \tag{2.36}$$

where β_∞ is the value of $\beta(r)$ as r approaches infinity,

$$\lim_{r\to\infty} \beta(r) \equiv \beta_\infty = \ln \frac{2D_e}{u_{\mathrm{LR}}(r_e)}, \tag{2.37}$$

and y_q^{ref} has the same form as y_p^{ref} but with the replacement of p by q. The formulas are similar,

$$y_p^{\mathrm{ref}}(r) = \frac{r^p - r_{\mathrm{ref}}^p}{r^p + r_{\mathrm{ref}}^p} \text{ and } y_q^{\mathrm{ref}}(r) = \frac{r^q - r_{\mathrm{ref}}^q}{r^q + r_{\mathrm{ref}}^q} \tag{2.38}$$

2.4 Computational details

The centroid potential is based on the centroid density of the p-H_2 dimer. The centroid potentials in 3D are calculated by PIMC. Those potentials are fitted to the MLR form using the BetaFit program [46]. In addition, partial potential data fits are performed by gnuplot [47].

2.4.1 PIMC potential construction

As mentioned before, the PIMC potential method is flexible with any dimension of the motions. The 3D potential can be computed easily. The method makes use of the Gaussian random distribution as a sampling tool to generate the paths for the p-H_2 dimer.

The centroid position of p-H_2 dimer is fixed during the simulation via the use of Fourier representation of path integral formulation. The centroid density is obtained via direct Monte Carlo integration and the centroid potential is obtained using Eq. (2.3).

2.4.2 Potential fitting

The 3D centroid potential behaves not as well as the 1D potential when using PIMC method. Noise can still be found in the 3D potential (Figure (2.2)) at short range, but is almost absence in 1D potential (Figure (2.1)). A fit to a smooth continuous potential form is highly recommended. We use the potential-fitting tool BetaFit to fit the potential data into the MLR form.

One dilemma of fitting the potential is to determine the long-range coefficients. The MLR form is an extended EMO form. It is meant to be a more accurate representation of the potential by taking into account the long-range information. Since the potential we are about to construct is an effective potential, it may not quite agree with the physical potential in real space. Hence, we cannot directly use the coefficients that have already been developed for p-H_2 in the previous work [48]. In order to obtain the rough shape of the potential long-range region, we perform the least-square fitting on the long-range in Gnuplot. By trial and error, we fitted the long range from 5.026 Å to 15.000 Åand obtained

the form,

$$u_{LR}(r) = \frac{C_8}{r^8} + \frac{C_{10}}{r^{10}}. \tag{2.39}$$

Another problem is the short-range noisy data fitting. As the 3D PIMC potential always presents a noisy short-range behaviour, the fitting program has difficulty to reaching convergence. From the fitting of inner wall turning points near to the potential minimum in the RKR1 fitting program [49], the potential may be represented by an exponential form for the repulsive wall, and the reciprocal distance terms are used to describe the long-range dispersion. Therefore, it is necessary to pre-fit the noisy data to an exponential form first. This exponential form is,

$$V_{\text{inner}}(r) = A + Be^{-Cr}. \tag{2.40}$$

After the pre-fitting of the short-range data, we replace the original data and use the fitted data as part of the input potential for BetaFit. The pre-fitting is also via Gnuplot.

Once the long-range coefficients are secured, and the short-range data have been taken care of, the potential data are ready to be put into BetaFit. BetaFit also requires the manual input of the estimated D_e, R_e and V_{min}. They can be easily found when plotting the potential data. V_{min} is the minimum energy of the potential.

2.5 Results and discussion

Through the PIMC method, we obtain the 3D centroid potentials at finite temperature from 3 K to 6 K. With BetaFit, we are able to obtain continuous forms for those centroid potentials.

2.5.1 Potential energy convergence

Getting convergence in path-integral calculations means minimizing the sampling error. Two aspects are usually addressed when dealing with the convergence of the PIMC method; one is the number of time slices or number of beads, the other is the number of Monte Carlo steps. The path integral calculated potential can converge really fast. The number of beads does not make significant changes to either the shape and the well-depth of the potential when going over 128 beads (Figure (2.2)). We show the potential convergence at $T = 5$ K as an example. In the 1D case, a large number of beads, such as $P = 640$, in the calculation tends to give no noise in the short-range, but the same result has already be found with 128 beads (Figure (2.3)). In 3D potentials, the number of beads does not significantly affect the region where noise occurs. As shown in Figure (2.4), the noise starts around 3.0 Å for the potential with 640 beads. Even with $P = 1024$, the noisy short-range still exists and starts almost at the same position. Therefore, we keep using the potential

calculated with a small number of beads.

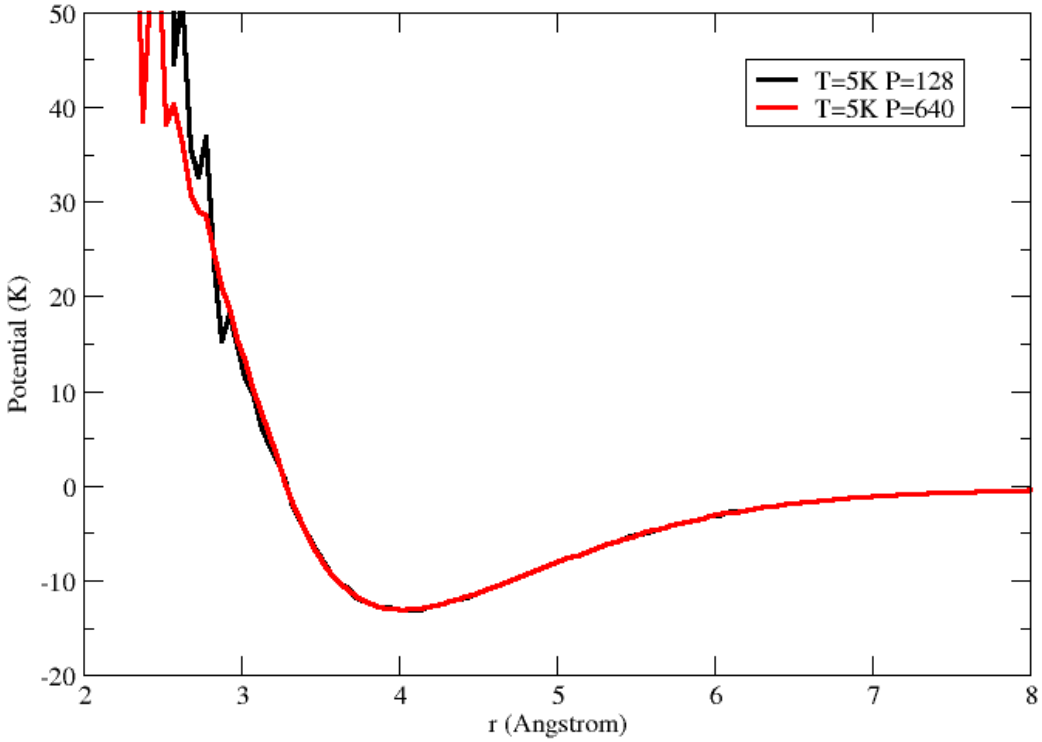

Figure 2.2: The 3D centroid potential convergence with $P = 128$. Even if the number of beads increases to 640, the potential well depth and long-range do not change much.

The noise behaviour of potentials depends on the number of Monte Carlo steps. With the same number of beads, the potentials become smoother when increasing the number of steps. As with the convergence with respect to the number of beads, it is a rapid convergence. The noise can be effectively minimized by using 10^4 Monte Carlo steps for the

1D potential. In addition, the 1D potential converges quicker than the 3D one, which needs at least 100000 steps. In the 1D case, the short-range noise can be virtually eliminated when the number of steps reaches 10^5 (Figure (2.5)). As opposed to the behaviour in 1D, the 3D potential still shows pretty noisy short-range behaviour. The number of Monte Carlo steps tends to affect the noise of the short-range part of the potential more than of the long-range part. As the number of steps goes larger, the shape of the potential is more clearly defined and less noisy. With 10^5 Monte Carlo steps, the long-range noise is essentially vanished (Figure (2.6)). To ensure the convergence is reached, we use 10^6 steps for all the calculations of the 3D potential.

2.5.2 Centroid pair-potential with finite temperature

The preliminary results for the 1D potentials show the significant variation on cutoff distance compared with that in the 3D potential. The cutoff distance is the radius that the corresponding potential function goes to zero. The 1D potential has a cutoff distance of ~ 3.7 Å which is 0.6 Å larger than that in the 3D potential (Figure (2.7)). The significant shift on the cutoff distance of 1D and 3D pseudo-potentials explains why the centroid radial distribution function with a 1D potential refer previously has a shift when comparing to that from the PIMC calculation (Figure (2.8)).

In this thesis, we focus on the study of the 3D centroid potential. With the PIMC

method, we are able to obtain the centroid potential for the p-H_2 dimer at temperatures starting from 3 K. Still, results for lower temperatures can be obtained; however, they will not be discussed here. The p-H_2 clusters start to show some signs of superfluidity at temperatures around or lower than 3 K [50, 51]. Rather than discussing the lower temperature behaviour, we would like to focus on the potential behaviour at 3 K and above (Figure (2.9)). As shown in Figure (2.9), all the effective potentials tend to keep the dissociation asympototte the same, only varying D_e and R_e. As the temperature increases, D_e becomes deeper, and R_e tends to approach values of the reference Buck potential. For the short-range behaviour, the noise of potential virtually disappears for the temperatures greater than 10 K. Also, the repulsive wall of the effective potentials before the cutoff distance becomes almost parallel to that of the Buck potential after 10 K. The pair-potentials at $T < 10$ K always show a wiggling repulsive wall and have a larger cutoff distance than that of the Buck potential. This is due to the quantum effects that make the Monte Carlo sampling really hard to converge, and quantum melting allows p-H_2 clusters stay liquid-like below 4 K. Furthermore, the interaction among molecules is weak at low temperatures, which allows molecules stay even closer compared to high temperature conditions.

The changing of R_e contains a sensitivity interval of at least 10 K when at low temperatures. In the thesis, we are only going to focus on the physical behaviour of p-H_2 clusters at finite low temperatures such as at 3 K, 4 K, 5 K and 6 K. Within the intervals of the

temperature changing 1 K, the equilibrium position holds the same value up to the third decimal place as shown in Table (2.1). This unchanging of the R_e does not violate the general trend which is stated that the temperature-dependent effective potential approaches to the Buck potential as the temperature reaches infinity. As for the temperature spacing bigger than 10 K, the changes on R_e can be observed (Table (2.2)).

Table 2.1: The R_e and corresponding D_e with respect to different temperatures that are smaller than 10 K

Temperature (K)	R_e (Angstrom)	D_e (Hartree$\times 10^{-5}$)
3	4.025	3.41059
4	4.025	3.79630
5	4.025	4.12738
6	4.025	4.43724

Table 2.2: The equilibrium position changes when the temperature is beyond the 10 K interval.

Temperature (K)	R_e (Angstrom)	D_e (Hartree$\times 10^{-5}$)
6	4.025	4.43724
10	3.925	5.44899
20	3.775	7.01757
30	3.625	7.89795
200	3.475	10.19320
∞ (Buck Potential)	3.425	10.71170

2.5.3 MLR fitting

As mentioned before, the MLR potential can provide a better shape on fitting the shallow effective potential than other potential forms. We monitor the centroid potential behaviours at T =3, 4, 5 and 6 K (Figure (2.10)). For the fitting of the long-range, the effective potential may not follow the regular physical rules. However, we should still start from basic physical parameter guessing. Since the potential for p-H_2 molecules has already been studied for many years, potential forms such as the Buck [37, 48] and the Silver-Goldman [36] are all well-developed. The Buck form containing the dispersion coefficients C_6, C_8 and C_{10} has been our first guess. For the 3D effective potentials, we applied the simple form of the Buck potential long-range which is

$$V_{tail}(r) = -\frac{C_6}{r^6} - \frac{C_8}{r^8} - \frac{C_{10}}{r^{10}}. \tag{2.41}$$

Here, $V_{tail}(r)$ is a set of truncated data with long-range features starting from $r = 5.025$ Å. The corresponding long-range coefficients for the p-H_2 dimer at different temperatures are presented in Table (2.3). All the data are magnified by 10^5 times since the effective potential is too shallow to catch the behaviour. Taking the potential at $T = 5$ K as an example, we get the fitted long-range coefficients $C_6 = 1492.72 \pm 674.7$ with a percentage error of 45.2%, $C_8 = (3.24495 \pm 0.04467) \times 10^6$ with percentage uncertainty of 1.376% and

Table 2.3: C_6, C_8 and C_{10} coefficients at various temperatures with the reference Buck potential [37].

	$C_6(10^{-5}Hartree$ Å$^6)$	$C_8(10^{-5}Hartree$ Å$^8)$	$C_{10}(10^{-5}Hartree$ Å$^{10})$
T=3 K	7347.27	3.78×10^6	7.67×10^7
T=4 K	2828.01	3.79×10^6	6.89×10^7
T=5 K	1492.72	3.24×10^6	5.52×10^7
T=6 K	2990.37	2.64×10^6	4.24×10^7
T$\rightarrow\infty(Buck)$	26657.7	1.32×10^5	8.29×10^5

$C_{10} = (5.51724 \pm 0.07253) \times 10^7$ with percentage uncertainty of 1.315%. The long-range fitting looks pretty accurate. However, it is noticeable that the C_6 term is not as important as C_8 and the C_{10} terms. Also, the percentage uncertainty is bigger with the C_6 coefficient than with C_8 and C_{10}. In other words, the C_6 term only contributes less than 10% of the whole long-range behaviour (Table (2.4)). For the potential at 5 K, $\frac{C_6}{r^6} = 0.0927176$ compared to other terms $\frac{C_8}{r^8} = 7.98$ and $\frac{C_{10}}{r^{10}} = 5.37$.

Table 2.4: C_6/r^6, C_8/r^8 and C_{10}/r^{10} coefficients at various temperatures with the Buck potential coefficients

	$C_6/r^6(10^{-5}Hartree)$	$C_8/r^8(10^{-5}Hartree)$	$C_{10}/r^{10}(10^{-5}Hartree)$
T=3 K	0.45636213	9.31	7.47
T=4 K	0.1756566	9.33	6.71
T=5 K	0.0927176	7.98	5.37
T=6 K	0.18574132	6.49	4.13
T$\rightarrow\infty(Buck)$	1.65589431	8.22	51.5

It is obvious that C_6 is not a dominant coefficient here. Omitting C_6 terms, we found the form of the long-range consisting of only C_8 and C_{10} terms provides almost the same accuracy as was obtained on including C_6 (Figure (2.11)). The fitted curve matches perfectly to the long-range of the potential and also to the curve obtained with the C_6 parameter included. The new coefficients are listed in Table (2.5). The insensitivity to the long-range

Table 2.5: C_6, C_8 and C_{10} coefficients at various temperatures

	$C_8(10^{-5} Hartree$ Å$^8)$	$C_{10}(10^{-5} Hartree$ Å$^{10})$
T=3 K	4.26×10^6	8.43×10^7
T=4 K	3.61×10^6	6.60×10^7
T=5 K	3.15×10^6	5.36×10^7
T=6 K	2.83×10^6	4.55×10^7

C_6 term may be due to the low temperature centroid potential for the p-H_2 dimer being very shallow compared to the classical Buck potential. The interaction between two p-H_2 molecules of the dimer is too weak to "see" the dispersion behaviour represented by the C_6 coefficient.

The noisy wall of the PIMC potential is due to the sampling error, which cannot be avoided. The noise starts 1.0 to 2.0 Å away from the potential minimum. The exponential form showed in Eq. (2.40) is used to represent the short-range. The fitting of the inner wall of the potential is tough. The choice of parameters and the selection of data points are the key to appropriate fitting of the potential. An example of the final fitting for the

potential at $T = 5$ K is presented in Figure (2.12). The short-range noise ends around 3.0 Å. Instead of fully fitting the noise, the exponential form shown in Eq. (2.40) is meant to fit the noise as close as possible. Finally, a smooth curve can be obtained and is able to represent the potential well and long-range information almost perfectly. Since the later molecular dynamics will only make use of the potential outside the cutoff distance, the inaccuracy at short range can be ignored. The summary of potential fitting error for different temperatures is listed in Table (2.6). Two types of errors are used to describe how well the fitting is: the one is the dimensionless root mean square deviation, $\overline{dd}$; and another is the dimensionless standard error, DSE [46]. The potential fitting errors for the potential at 6 K is one magnitude higher than the other errors for potentials at lower temperatures. Large errors may be due to the difficulty of fitting the steepest short-range of the potential at 6 K, but these errors are still acceptable. The summary of all the

Table 2.6: Potential fitting error with BetaFit

	$\overline{dd}$	DSE
T=3 K	4.37×10^{-2}	4.42×10^{-2}
T=4 K	8.30×10^{-2}	8.39×10^{-2}
T=5 K	3.06×10^{-2}	3.10×10^{-2}
T=6 K	3.19×10^{-1}	3.23×10^{-1}

potential fitting parameters is presented in Table 2.7. The fitted effective potential with temperature varied from 3 K to 6 K is given in Figure (2.13).

2.5.4 Singularity

As shown in Section 2.5.2, the PIMC calculated potentials do not have a normal short-range behaviour. Except for the short-range noise, the potentials also present a constant value from the origin to the beginning of the noise (Figure (2.14)). This constant value may not correspond to the systematic error as it can also be found in the NMM method. It may be a unique feature of the centroid potential. It is as if the centroid potential avoids the singularity of the classical potential at zero distance. Paths with a zero relative centroid distance, but where the actual beads are away from the origin, will still contribute to the centroid density leading to a finite centroid potential.

2.6 Conclusion

The use of a centroid pseudo-potential provides us with a way to reduce the complexity of the quantum many-body problem and allows us to extend our research on the physical properties of larger N-body systems. The centroid pseudo-potential for the N-body system is considered to be a pairwise additive potential, similar to that in the classical system. Using p-H_2 clusters as an example, we computed the centroid potential for the p-H_2 dimer. The potential was studied for both 1D and 3D systems. For the 1D potential, the NMM method can provide a well converged potential for the dimer. For 3D systems, the PIMC

method is feasible for calculating the potential, however, noisy short-range behaviour. In order to make a smooth potential for later use, we applied the potential fitting program BetaFit to cast the potentials into the MLR form. The well established Buck potential was used as a basis to start the fitting. The centroid potential is a temperature dependent effective potential and cannot be fully represented by the Buck potential form with the absence of the long-range C_6 term. The constant short-range behaviour may be a unique feature for the centroid potential. This singularity of the centroid potential will be studied in future research. If the NMM 3D method can finally be available, it may provide more details on the angular momentum dependence and the interaction between radial and angular properties, and may explain the singularity problem in a better way. In order to reduce the noise of the short range, the umbrella sampling is a possible method to used [52].

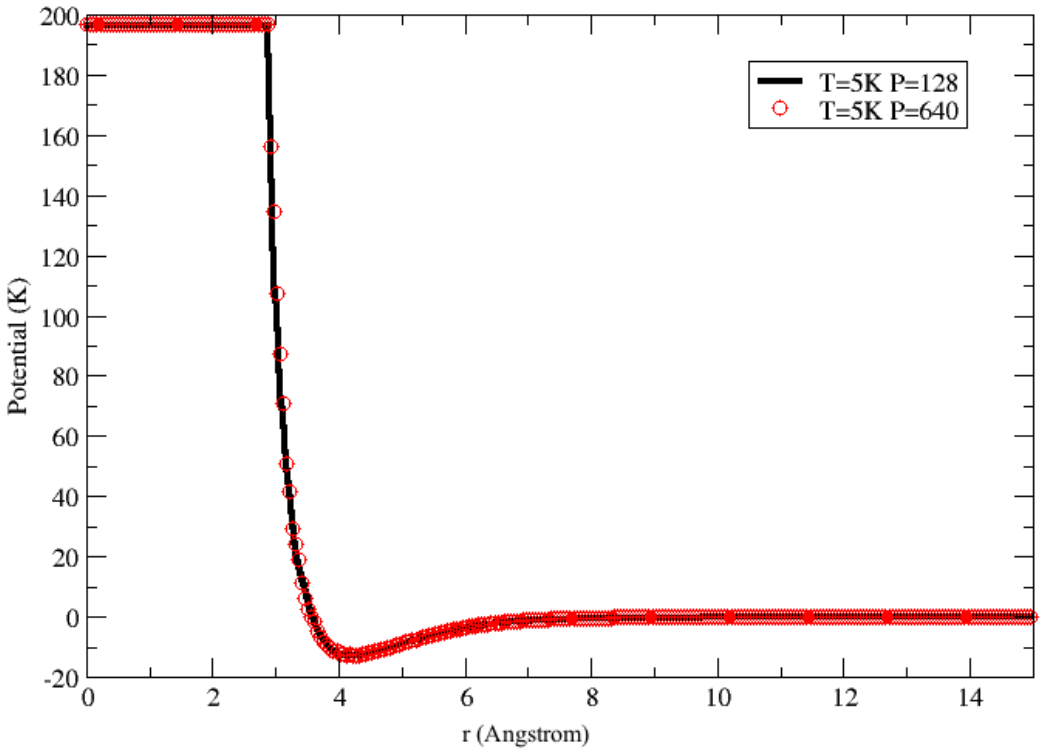

Figure 2.3: Centroid potential of p-H_2 dimer in 1D converges with P=128 at a temperature of 5 K. The short-range behaviour shows no noise.

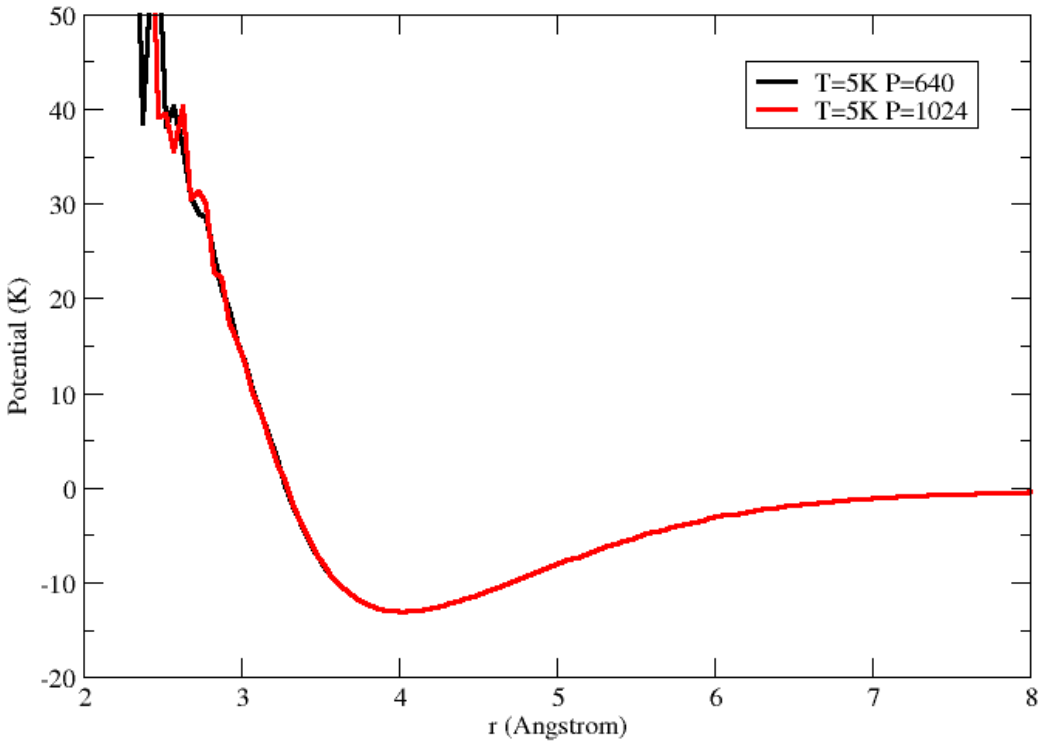

Figure 2.4: The 3D Centroid potential with an even larger number of beads in the calculation. The final picture does not change too much; here, P=640 and P=1024.

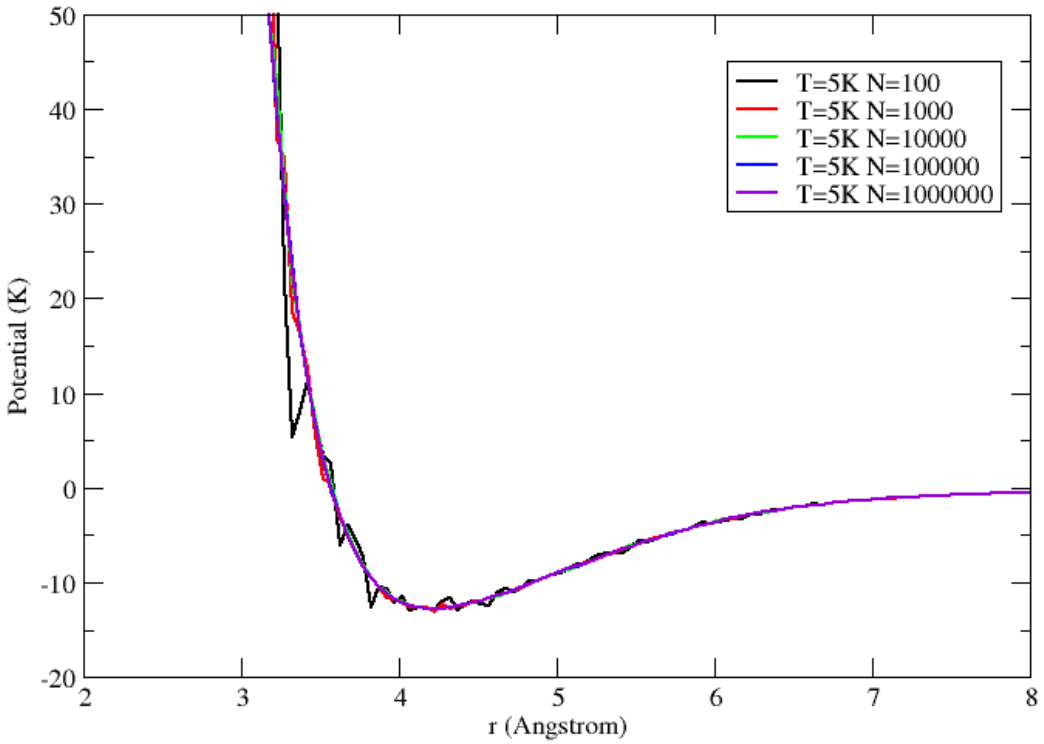

Figure 2.5: The 1D centroid potential with number of Monte Carlo steps varying from 100 to 1000000 at $T = 5$ K and $P = 128$. The noise can be essentially eliminated when the number of steps goes to 10000.

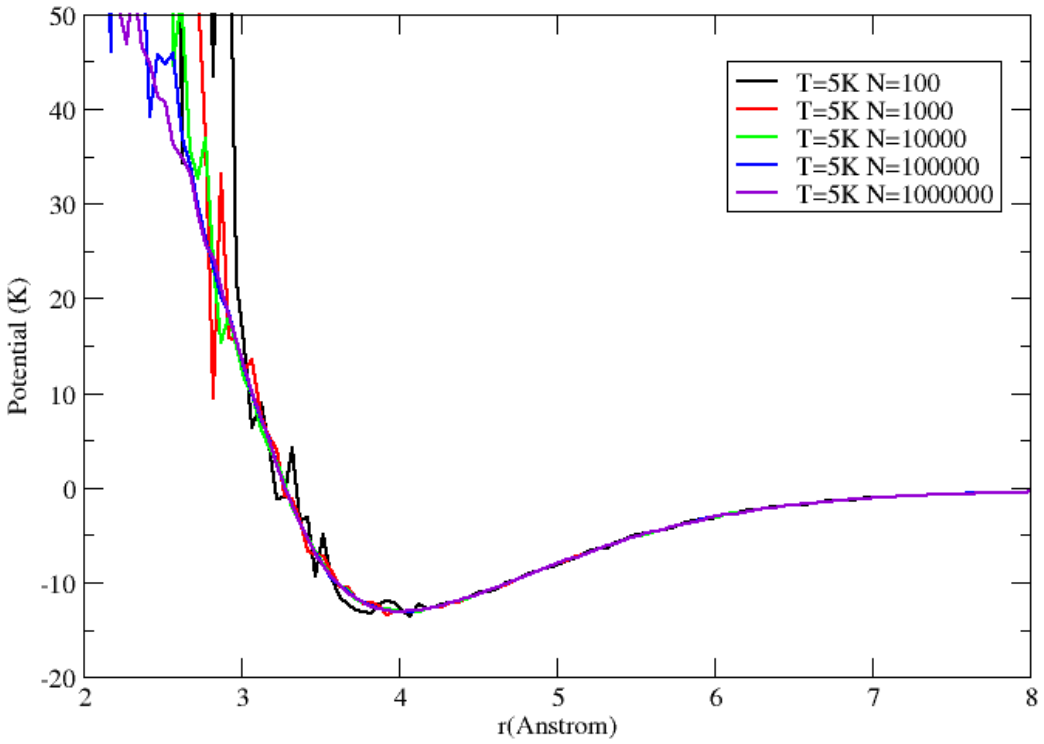

Figure 2.6: The 3D centroid potential obtained with the number of Monte Carlo steps varying from 100 to 10^6 at $T = 5$ K and $P = 128$. The potential curve reaches convergence at 10^6 steps. The short-range noise does not go away even with large numbers of steps.

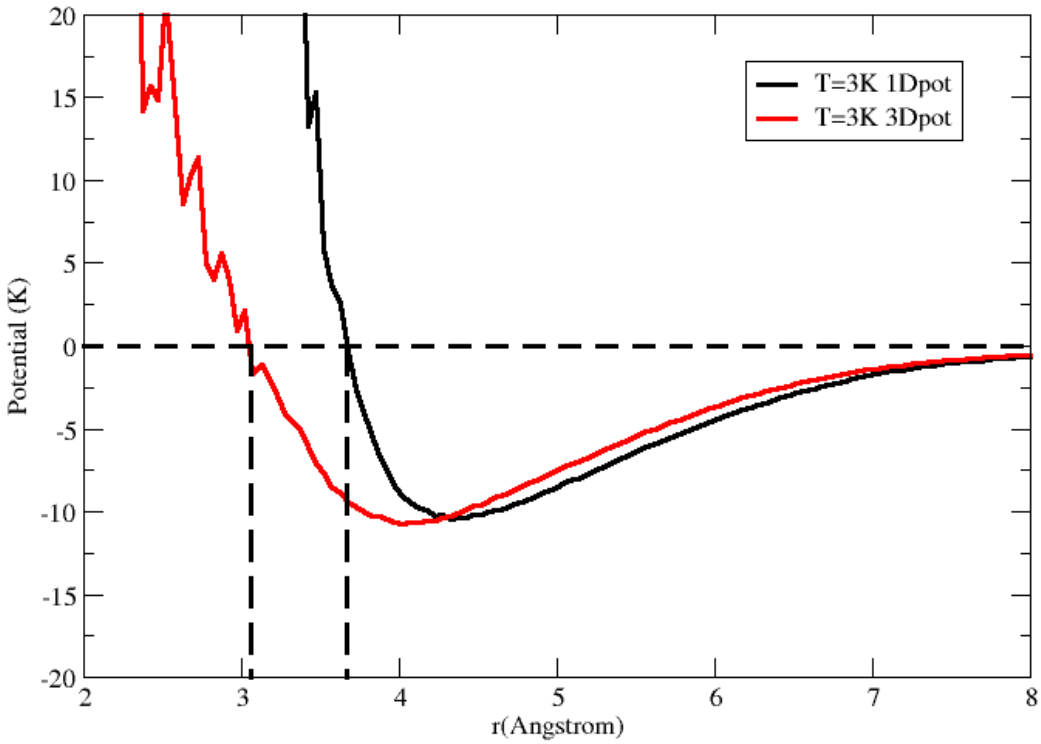

Figure 2.7: The centroid pair-potential under 1D coordinates versus the centroid pair-potential under 3D coordinates. The 3D potential is wider than the 1D potential because of taking account of the angular momenta in 3D space.

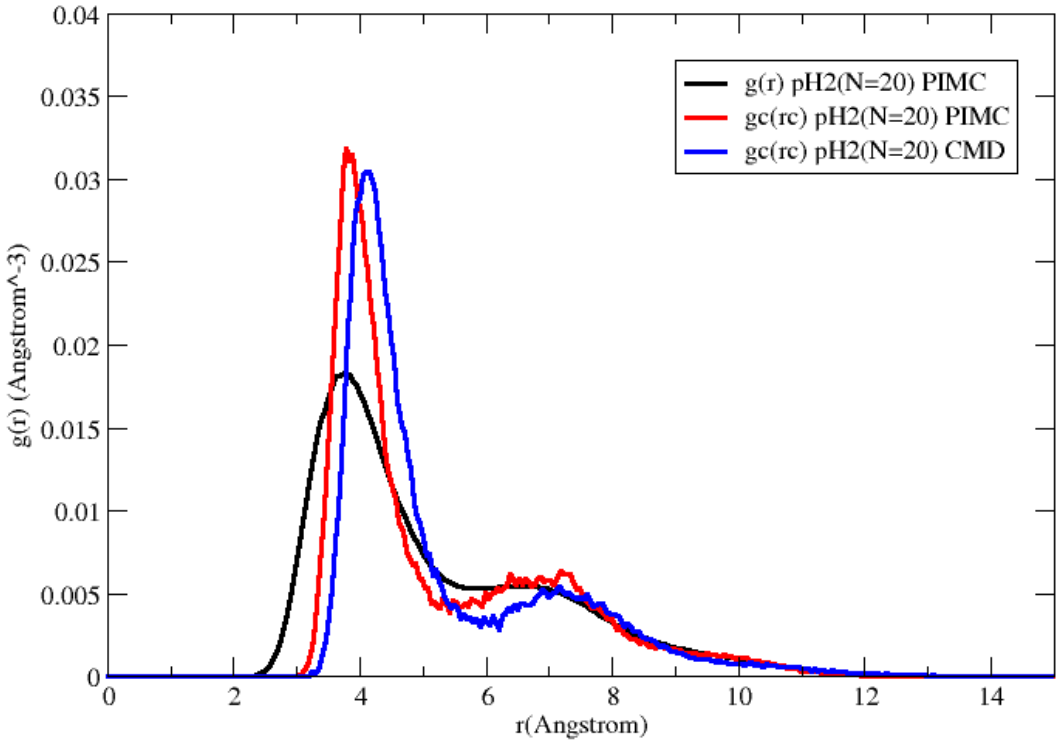

Figure 2.8: The centroid and the regular radial distribution functions of $(\text{p-H}_2)_{N=20}$ at $T = 5$ K. The centroid radial distribution function obtained from CMD is with the 1D NMM pseudo-potential. The shift can be found when comparing to the exact centroid radial distribution from PIMC.

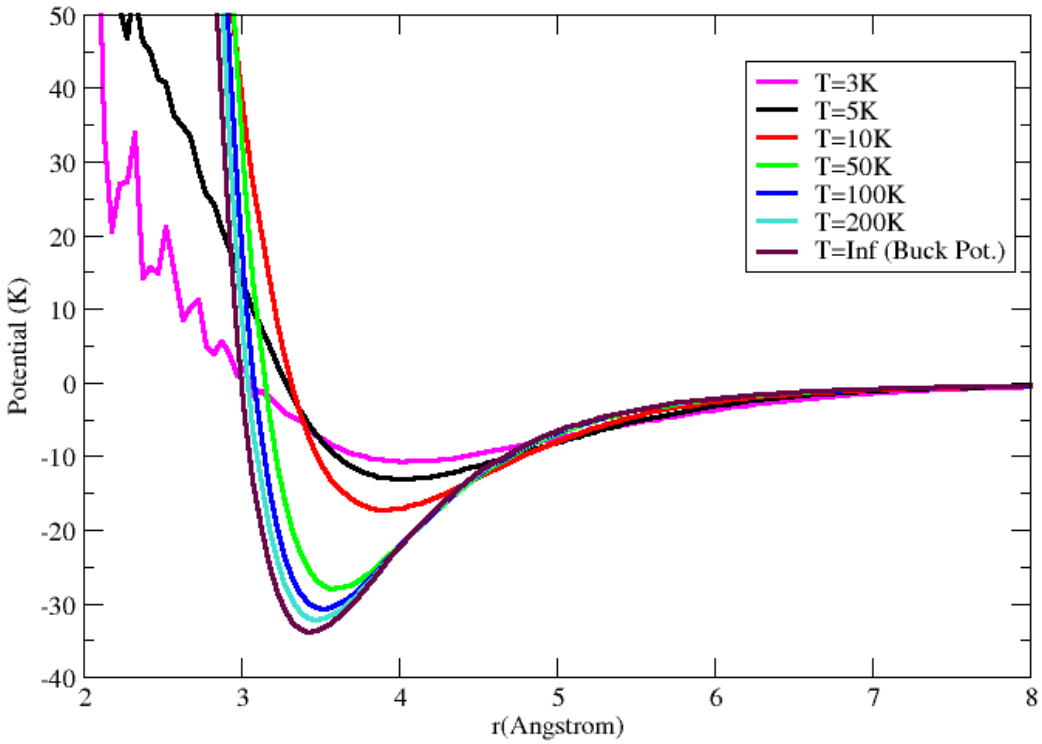

Figure 2.9: The centroid potential for p-H_2 at various finite temperatures. The equilibrium position becomes smaller as the temperature goes higher. Both the potential well-depth and equilibrium position approach the Buck potential at high temperature.

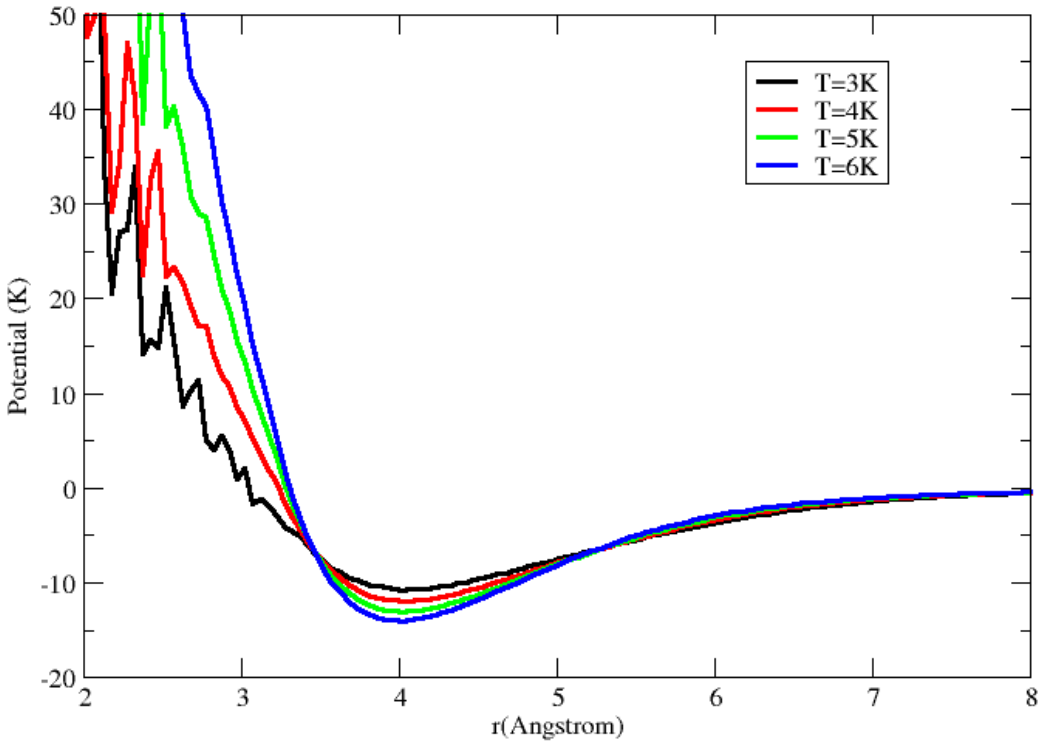

Figure 2.10: The centroid pseudo-potentials directly from the PIMC calculation at $T = 3 - 6$ K.

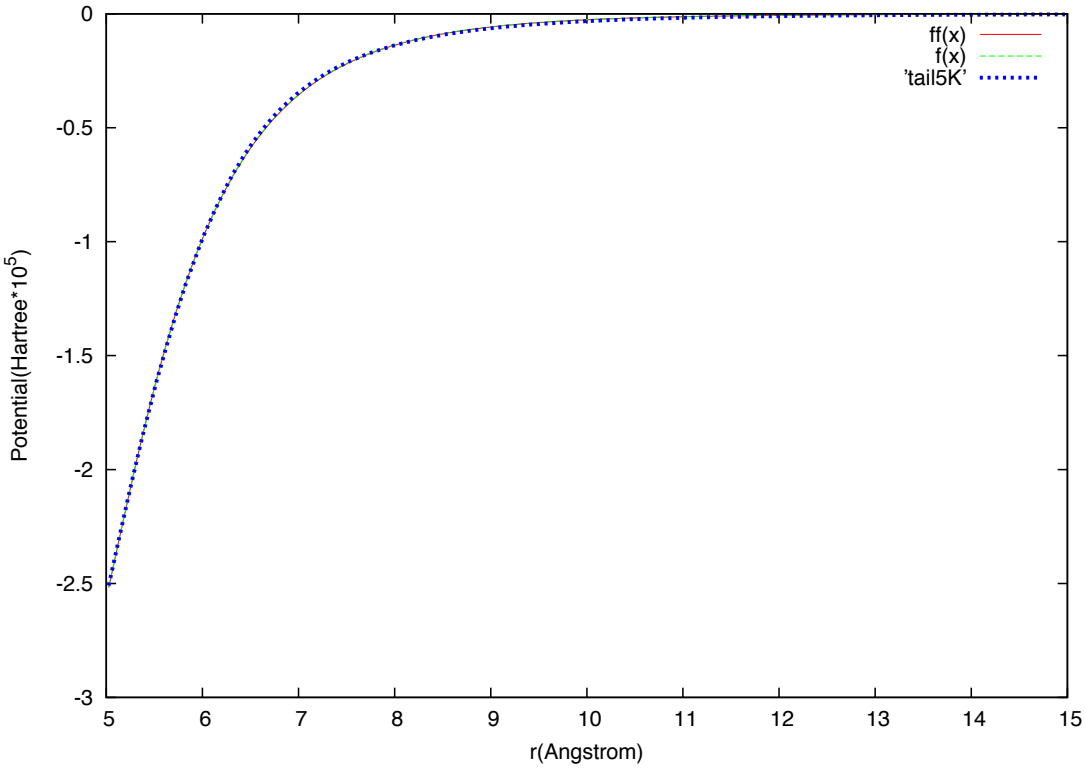

Figure 2.11: The long-range fitting with $-\frac{C_8}{r^8} - \frac{C_{10}}{r^{10}}$ and $-\frac{C_6}{r^6} - \frac{C_8}{r^8} - \frac{C_{10}}{r^{10}}$. Both with the C_6 coefficient and without the C_6 coefficient provides almost the same curvature that matches to the original data pretty well. In the plot, $\mathrm{ff}(x) = -\frac{C_8}{r^8} - \frac{C_{10}}{r^{10}}$ and $\mathrm{f}(x) = -\frac{C_6}{r^6} - \frac{C_8}{r^8} - \frac{C_{10}}{r^{10}}$.

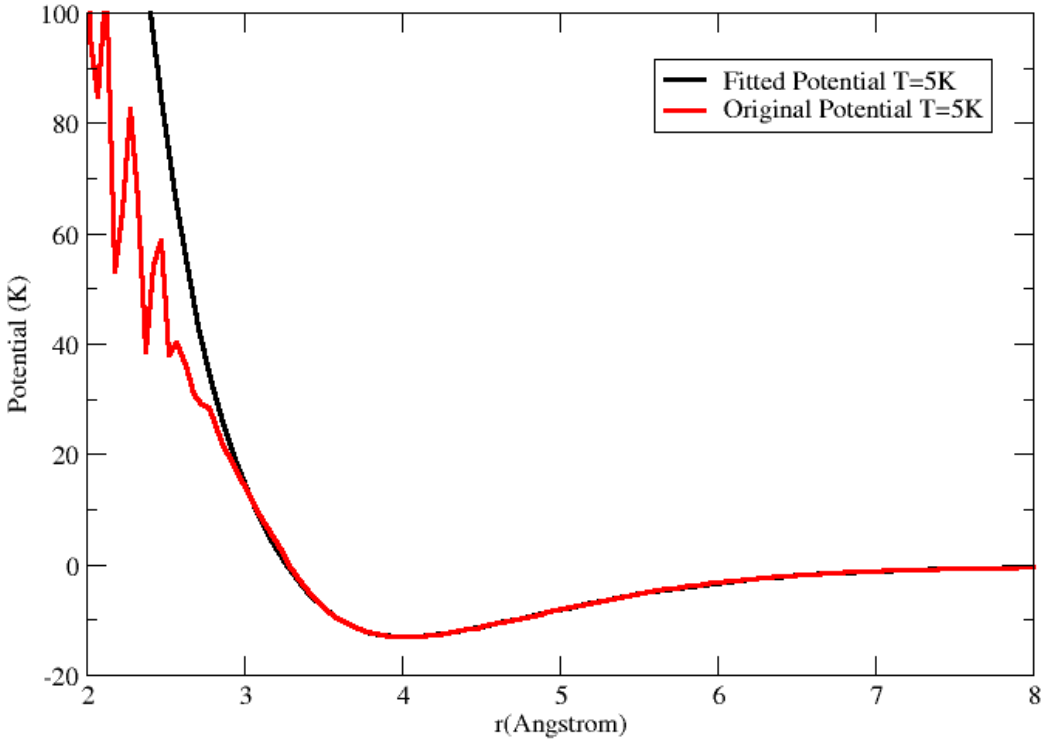

Figure 2.12: The centroid potential fitted to the MLR form at $T = 5$ K with a fitting error $\overline{dd} = 3.06 \times 10^{-2}$

Table 2.7: Parameters of the MLR fit of the helium dimer potential.

T=3 K	β_0	-2.8104	β_3	-3.5542
	β_1	0.2675	β_4	6.3496
	β_2	-3.7116	β_5	7.0412
	β_∞	-2.7379		
	C_8/(Hartree $\times 10^{-5}$Å^8)	4.2673×10^6		
	C_{10}/(Hartree $\times 10^{-5}$Å^{10})	8.4378×10^7		
	R_e/ Å	4.025	D_e / Hartree $\times 10^{-5}$	3.40349
T=4 K	β_0	-2.5166	β_3	-1.8808
	β_1	0.2763	β_4	4.0267
	β_2	-2.7881	β_5	4.0994
	β_∞	-2.4232		
	C_8/(Hartree $\times 10^{-5}$Å^8)	3.6064×10^6		
	C_{10}/(Hartree $\times 10^{-5}$Å^{10})	6.5976×10^7		
	R_e/ Å	4.025	D_e / Hartree $\times 10^{-5}$	3.7775
T=5 K	β_0	-2.4821	β_3	-0.2979
	β_1	0.0361	β_4	3.6399
	β_2	-1.5235	β_5	3.2247
	β_∞	-2.1681		
	C_8/(Hartree $\times 10^{-5}$Å^8)	3.1409×10^6		
	C_{10}/(Hartree $\times 10^{-5}$Å^{10})	5.3449×10^7		
	R_e/ Å	4.025	D_e / Hartree $\times 10^{-5}$	4.1274
T=6 K	β_0	-2.3751	β_3	1.3819
	β_1	-0.1212	β_4	5.3722
	β_2	-1.4598	β_5	3.5801
	β_∞	-1.967189		
	C_8/(Hartree $\times 10^{-5}$Å^8)	2.8333×10^6		
	C_{10}/(Hartree $\times 10^{-5}$Å^{10})	4.5529×10^7		
	R_e/ Å	4.025	D_e / Hartree $\times 10^{-5}$	4.4372

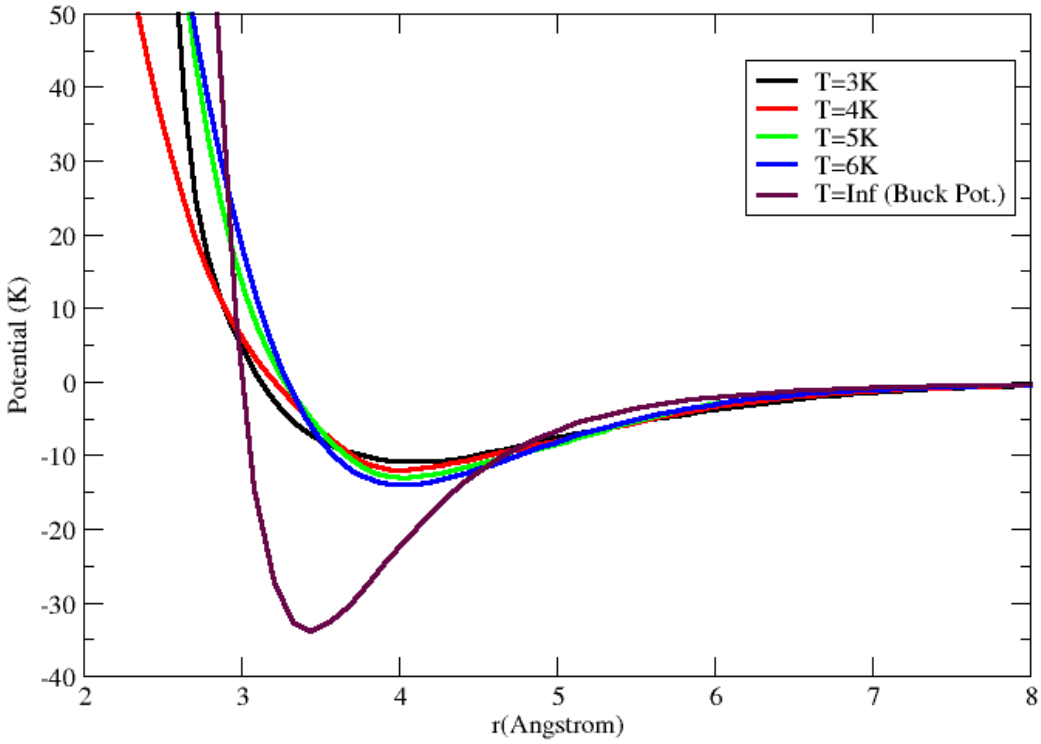

Figure 2.13: The final fitted effective pseudo-potentials at T =3, 4, 5 and 6 K.

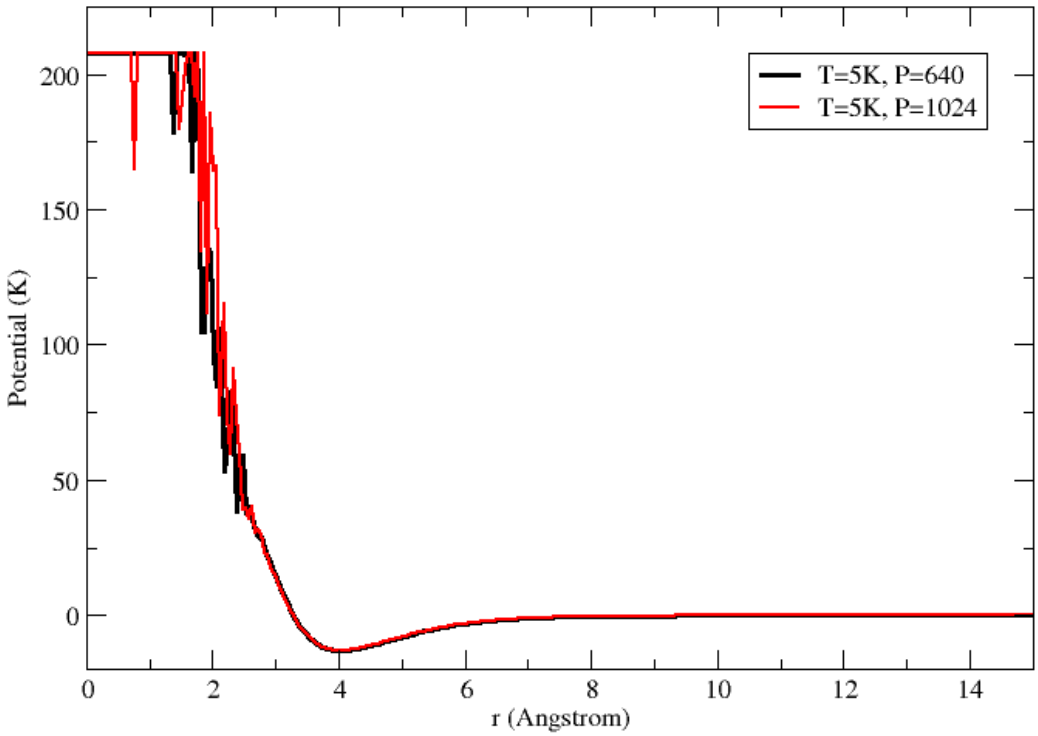

Figure 2.14: The full view of the centroid potential at $T = 5$ K with number of beads, P=640 and P=1024. It is the same with Figure (2.4) but in different scale of y-axis.

Chapter 3

CMD of parahydrogen clusters

The study of CMD dates back to the early 1990s and is based on the concept of a centroid variable introduced by Feynman [16]. The idea of centroid density, and its associated centroid effective potential, is very useful since it implicitly accounts for zero-point motion and thermal quantum delocalization effects [34]. The development of CMD is based on the Feynman path centroid density [53]. Both equilibrium and dynamical properties have been addressed within the centroid formulation of quantum statistical mechanics [25, 30, 31, 54]. CMD has been applied to low temperature quantum systems, namely liquid parahydrogen [21]. The self-diffusion constants (D) of bulk liquid parahydrogen obtained from CMD has been shown to be very close to that from experiments. At 14 K, the experimental value of D is 0.4 $\text{Å}^2\text{ps}^{-1}$, while the value from CMD is 0.48 $\text{Å}^2\text{ps}^{-1}$ [21].

Care must be taken when analyzing structural properties based on centroid variables since they do not correspond directly to their real space counterparts [55, 56]. One can use a deconvolution kernel to connect centroid and real space structural properties [57].

In this thesis, we use CMD to obtain equilibrium properties of parahydrogen clusters. We focus especially on the pair distribution function, $g(r)$, for various low temperatures and cluster sizes.

3.1 Theory

3.1.1 CMD formulation

CMD is analogous to classical molecular dynamics although the potential is replaced by an effective potential energy surface containing quantum information [31]. The equations of motion can be written as,

$$\begin{aligned} \dot{q_c}(t) &= p_c(t)/m, \\ \dot{p_c}(t) &= F_c(q_c), \end{aligned} \tag{3.1}$$

where the centroid force F_c is written as,

$$F_c(q_c) = -\frac{dV_c(q_c)}{dq_c}, \tag{3.2}$$

where $V_c(q_c)$ is the centroid potential. The centroid potential has been defined in the previous chapter and has the form,

$$V_c(q_c) = -k_B T \ln[\rho_c(q_c)/\rho_c^0], \tag{3.3}$$

relative to the centroid density ρ_c of the system. Based on the equations of motion, we can perform CMD simulations for our N-body system of interest, parahydrogen clusters, to obtain trajectories for different cluster sizes at low temperatures. The trajectories allow us to calculate many properties of these clusters.

3.1.2 Quasi-density operator

It was shown that the introduction of a centroid quasi-density operator (QDO), analogous to the usual thermal density operator, is useful to make further developments [58, 35, 59,

57]. This QDO is defined as,

$$\begin{aligned}\hat{\delta}_c(r_c, p_c) &\equiv \hat{\phi}(r_c, p_c)/\rho_c(r_c, p_c) \\ &= \int \mathrm{d}r' \int \mathrm{d}r''|r'\rangle\Big\{\frac{e^{ip_c(r'-r'')/\hbar}}{e^{-m(r'-r'')^2/2\beta\hbar^2}}\frac{\langle r'|\hat{\phi}(r_c)|r''\rangle}{\rho_c(r_c)}\Big\}\langle r''|. \end{aligned} \tag{3.4}$$

The operator $\hat{\phi}_c$ is defined as,

$$\hat{\phi}_c(r_c, p_c) = \int \frac{d\vec{k}}{2\pi}\exp(-\beta\hat{H}'(\vec{k}, \hat{r})), \tag{3.5}$$

where

$$\hat{H}' = \hat{T} + \hat{V} + \frac{i\vec{k}\cdot\hat{r}}{\beta}. \tag{3.6}$$

All the arguments are illustrated in detail in Chapter 2. The centroid distribution function as $\rho_c(r_c, p_c)$ in Eq. (3.4) is just the centroid density combined with a centroid momentum distribution,

$$\rho_c(r_c, p_c) = e^{-\beta p_c^2/2m}\rho_c(r_c). \tag{3.7}$$

The concept of QDO allows one to introduce the centroid representation of several physical observables.

3.1.3 Centroid equilibrium properties

Any centroid variable, A_c, can be represented as the trace of the product of the quasi-density operator and the operator of interest itself [35]. It can be written as,

$$A_c(x_c, p_c) = Tr\{\hat{\delta}_c(x_c, p_c)\hat{A}\}, \tag{3.8}$$

where $\hat{A}$ is the operator associated with A_c. In order to simplify the mathematical illustration, one-dimensional notations are used in this section. Therefore, x_c represents the centroid distance in one-dimension. A_c can be described as the average of a physical variable mapped by the QDO. It is also possible to obtain the average over the centroid distribution,

$$\begin{aligned}\langle A_c\rangle_c &\equiv \frac{1}{Z}\int\int \frac{\mathrm{d}x_c\,\mathrm{d}p_c}{2\pi\hbar}\rho_c(x_c, p_c)A_c \\ &= \frac{1}{Z}Tr\{\int\int \frac{\mathrm{d}x_c\,\mathrm{d}p_c}{2\pi\hbar}\rho_c(x_c, p_c)\hat{\delta}_c(x_c, p_c)\hat{A} \\ &= \frac{1}{Z}Tr\{e^{-\beta\hat{H}}\hat{A}\} \equiv \langle\hat{A}\rangle. \end{aligned} \tag{3.9}$$

For the centroid position and momentum, based on Eq. (3.8),

$$x_c = Tr\{\hat{\delta}_c(x_c, p_c)\hat{r}\}, \tag{3.10}$$

$$p_c = Tr\{\hat{\delta}_c(x_c, p_c)\hat{p}\}. \tag{3.11}$$

For the centroid force, F_c, we have

$$\begin{aligned} F_c &= Tr\{\hat{\delta}_c(x_c, p_c)\hat{F}(\hat{x})\} \\ &= Tr\{\frac{\hat{\varphi}(x_c)}{\rho_c(x_c)}F(\hat{x})\} = \frac{1}{\beta}\frac{d}{d\ x_c}\ln\{\rho_c(x_c)\}. \end{aligned} \tag{3.12}$$

The centroid potential of mean force (PMF), V_{cm}, can then be defined as,

$$V_{cm}(x_c) = -\frac{1}{\beta}\ln\{\rho_c(x_c)\}. \tag{3.13}$$

Since the potential does not relate directly to the real physical properties, it is also called the centroid effective pseudopotential. A detailed proof was shown in Blinov and Roy's work in 2001 [59]. The centroid force Eq. (3.12) can be obtained from the centroid effective potential,

$$F_c = -\frac{d}{d\ x_c}V_{cm}(x_c). \tag{3.14}$$

Note that the centroid effective potential is a temperature dependent curve, and a free particle factor, ρ_c^0, can be used such that the potential goes to zero in the long range and

Eq. (3.13) can be written as,

$$V_{cm}(x_c) = -\frac{1}{\beta} \ln\{\rho_c(x_c)/\rho_c^0\}. \tag{3.15}$$

3.1.4 Structural properties study of parahydrogen clusters

The feature that can directly describe the structural factor for the system is the radial distribution function (or pair correlation function), $g(r)$ [60]. It fixes one of the particles in the system as a reference and measures the density within an area covered by the radius r, when the system has spherical symmetry. Specifically, the pair distribution function (PDF) is the radial distribution function that considers the probability of finding only one atom within distance of r to the reference particle.

For the parahydrogen clusters, researchers have been putting a great deal of effort into determining the proper structures for different clusters. They have found that the structure of parahydrogen clusters depends on both the temperature and the size of the cluster. The structure determination has been debated over the past few years. Earlier work has shown that finite temperature parahydrogen clusters most likely adopt icosahedral-based structures [17]. These adopted structures have been confirmed in subsequent works [61, 62, 63]. However, in the ground state ($T = 0$ K), small parahydrogen clusters ($N \leqslant 20$)

tend to remain liquid-like [20]. Later, the concept of quantum melting [50] explained that both liquid-like and solid-like features are in coexistence in parahydrogen clusters at low temperature [51, 64, 65]. Recent studies have successfully predicted the exact phases and structures of parahydrogen clusters [66, 67].

These findings about the properties of parahydrogen clusters are based primarily on one of the leading methods of many-body computation, PIMC. The path-integral ground state method [20] and the diffusion Monte Carlo (DMC) [68] method have also been used extensively. Regarding our CMD, so far, work related to the structural properties was performed only for parahydrogen bulk liquids. We are unaware of any further studies for parahydrogen clusters employing the CMD method. It is necessary to mention that our CMD is a little bit different from the work carried out in previous studies. In this thesis, we make use of pseudo-potential implemented CMD to determine the structural properties, whereas previous work focused on the dynamical properties [21, 69].

The centroid structural properties were found to be different from those obtained from previous calculation with path-integral methods [55, 56]. The centroid radial distribution function, $g_c(r_c)$, was more sharply peaked than the one from PIMC (Figure (3.1)). This difference was explained by the existence of a transition function connecting the centroid and regular structural properties. This transition function is termed the "deconvolution kernel" [57]. The deconvolution kernel is written as $f(r, r_c)$.

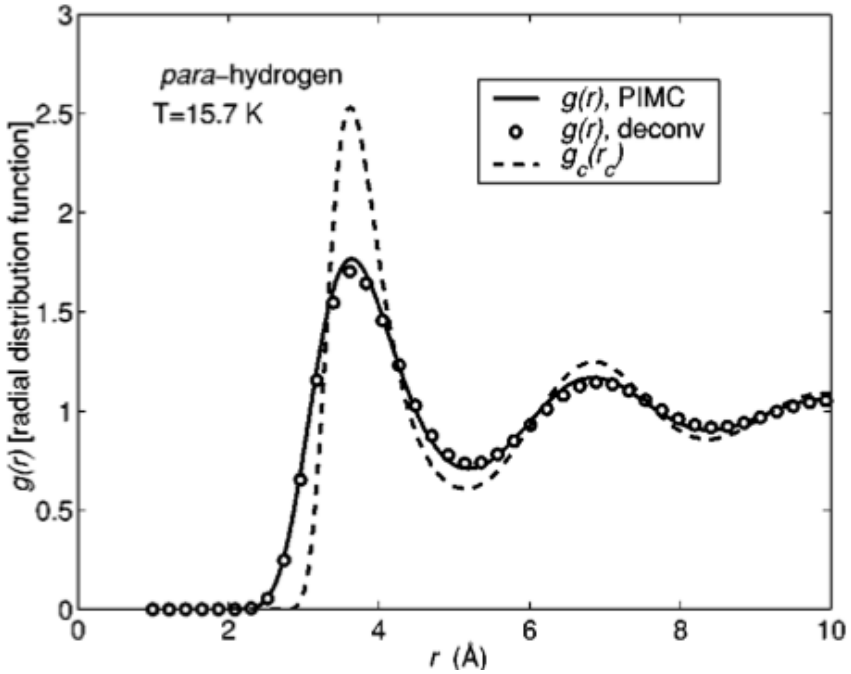

Figure 3.1: Ccomparison of centroid versus regular $g(r)$. For bulk parahydrogen liquid, the first $g_c(r_c)$ peak is one-and-half times higher than that from PIMC [57].

Deconvolution of centroid properties

To illustrate the deconvolution idea, it is appropriate to define the $g(r)$ in terms of CMD.

The actual $g(r)$ can be written as

$$g(\mathbf{r}) = \frac{1}{Z} \int g_c(q_c; \mathbf{r}) \rho_c(r_c) \, \mathrm{d}r_c, \tag{3.16}$$

where $q_c = (\mathbf{r}_{c1}, \ldots, \mathbf{r}_{cN})$ is the phase space vector of the centroid coordinates in 3N dimensions, $\mathbf{r}$ is the relative distance between two particles, Z is the partition function of the system, and g_c is a centroid variable defined as,

$$g_c(r_c; \mathbf{r}) = \langle \mathbf{r} | \hat{\delta}_c(r_c) | \mathbf{r} \rangle. \tag{3.17}$$

The centroid variable $g_c(r_c; \mathbf{r})$ is the partial trace of the QDO. Using the functional derivative of the centroid density with respect to the classical potential, $g_c(q_c; \mathbf{r})$ can be written as,

$$\langle \mathbf{r} | \hat{\delta}_c(r_c) | \mathbf{r} \rangle = \frac{2V}{N(N-1)} \frac{1}{(-\beta)} \frac{\delta \ln \rho_c(r_c)}{\delta v(\mathbf{r})}, \tag{3.18}$$

where V is the volume, N is number of particles in the system, and v is the classical potential. Taking advantage of the functional derivative of Eq. (3.18) and with some algebra, the connection between the $g_c(r_c)$ and the physical one can be written as,

$$g(\mathbf{r}) = \frac{1}{Z} \int g_c(\mathbf{r}_c) f(\mathbf{r}_c; \mathbf{r}) \, \mathrm{d}\mathbf{r}_c, \tag{3.19}$$

where $\mathbf{r}_c$ is the relative distance between two centroid positions, and $\mathbf{r}$ is the physical relative distance between particles. For this pairwise radial distribution, the deconvolution

kernel is the partial trace of the QDO of the two-particle system,

$$f(\mathbf{r}_c;\mathbf{r}) = \langle \mathbf{r}|\hat{\delta}_c^{(2)}(\mathbf{r}_c)|\mathbf{r}\rangle \equiv \frac{\tilde{\rho}_{\mathrm{rel}}^{(2)}(\mathbf{r}_c;\mathbf{r})}{\tilde{\rho}_{\mathrm{rel}}^{(2)}(\mathbf{r}_c)}, \tag{3.20}$$

where $\tilde{\rho}_{\mathrm{rel}}^{(2)}(\mathbf{r}_c;\mathbf{r})$ is the quasi-density for the two-particle system, which is the trace of Eq. (3.4) and $\tilde{\rho}_{\mathrm{rel}}^{(2)}(\mathbf{r}_c)$ is the corresponding centroid density equivalent to Eq. (3.7) when considering the interaction only between two particles in the system.

3.1.5 Energy estimator and chemical potential

The energy estimator has long been studied as one of the important energetic properties of many-body systems [40]. For a canonical system, the average internal energy can be calculated by

$$\langle E\rangle = -\frac{\partial \ln Z(\beta)}{\partial \beta}. \tag{3.21}$$

In the case of the centroid partition function,

$$Z = \int \mathrm{d}r_c \rho_c(r_c), \tag{3.22}$$

where $\rho_c(r_c)$ holds a Boltzmann factor form only related to the effective pair potential (Chapter 2),

$$\rho_c(r_c) = \exp(-\beta V_{\text{eff}}) = \exp(-\beta \sum_{i<j} v_{ij}), \tag{3.23}$$

with v_{ij} as an additive pair potential, similar to the method used in the classical system. The additive pair potential has a Fourier integral form related to the pair density,

$$\begin{aligned} v_{ij} &= -k_B T \ln \rho_c^{(2)} = -\frac{1}{\beta} \ln \rho_c^{(2)} \\ &= -\frac{1}{\beta} \ln \big(\int \frac{\mathrm{d}k}{2\pi} \exp(i\vec{k} \cdot \vec{r_c}) \mathrm{Tr}(\exp(-\beta H - i\vec{k} \cdot \hat{r})) \big). \end{aligned} \tag{3.24}$$

The average internal energy can be derived according to,

$$\begin{aligned} \langle E \rangle &= -\frac{1}{Z} \frac{\partial Z}{\partial \beta} \\ &= -\frac{1}{Z} \big[- \int \mathrm{d}r_c \exp(-\beta V_{\text{eff}}) V_{\text{eff}} - \beta \int \mathrm{d}r_c \exp(-\beta V_{\text{eff}}) \frac{\partial V_{\text{eff}}}{\partial \beta} \big] \\ &= \langle V_{\text{eff}} \rangle + \langle E^{(2)} \rangle. \end{aligned} \tag{3.25}$$

The above equation consists of two parts: the average potential energy and the average internal energy for a parahydrogen dimer. Note that the superscript "(2)" in this equation represents that the variables are in pairs.

The chemical potential, the measure of Gibbs free energy changes for adding one more parahydrogen molecule in the system, is also presented in this thesis. The proper equation for the chemical potential is,

$$\mu(N) = E(N-1) - E(N), \tag{3.26}$$

where $E(N-1)$ and $E(N)$ are the total energy for $N-1$ and N particles in the system.

3.2 Computational details

From Sindzingre and coworkers [70], p-H_2 clusters start to show superfluidity at $T \sim$2.5 K and below. The clusters at higher temperature with permutation being supressed can be considered to be following the Boltzmann distribution. For the CMD simulations, we first work out the forcefield from the centroid potential dicussed in Chapter 2. The MMTK [29] is used to perform the simulations. We obtain the $g_c(r_c)$ from two different methods: the PIMC method with exact $g_c(r_c)$ calculation, and CMD with the centroid pseudo-potential. The exact centroid calculation is performed for the purpose of testing the accuracy of our pseudo-potential approximation. The PIMC usually takes very long time when taking account into the quantum effects, whereas the pseudo-potential approach, classical-like simulation, is shorter than PIMC.

3.2.1 Forcefield construction

The forcefield consisting of all the gradient components of each particle in the system is manually implemented in MMTK. In order to get rid of the repetitive calculation for the gradients, we pre-calculate the force gradient and save it in a lookup table. In that way, for a specific position of the particle, the program will just look in the table and perform a simple interpolation to obtain the gradient value. The force can be calculated by

$$f = -\frac{dV_c}{dr_c}, \tag{3.27}$$

where V_c is the centroid potential for the relative distance r_c between two p-H_2 molecules. In 3D , the gradient should be representated on a $grid(x, y, z)$ according to the formula,

$$grid(x, y, z) = \nabla V_c(x, y, z) = (\frac{dV_c}{dx}, \frac{dV_c}{dy}, \frac{dV_c}{dz}). \tag{3.28}$$

However, the potential is given in terms of distances. The chain rule is used to connect Eq. (3.27) and Eq. (3.28). The gradient along each axis for a specific molecule can be written

as,

$$
\begin{aligned}
grid_{x_i} &= -\frac{dV_c}{dx_i} = \frac{dV_c}{dr_c}\frac{dr_c}{dx_i}, \\
grid_{y_i} &= -\frac{dV_c}{dy_i} = \frac{dV_c}{dr_c}\frac{dr_c}{dy_i}, \\
grid_{z_i} &= -\frac{dV_c}{dz_i} = \frac{dV_c}{dr_c}\frac{dr_c}{dz_i},
\end{aligned} \tag{3.29}
$$

where r_c is the length of a relative position vector $\vec{r_c}$

$$
r_c = |\vec{r_c}| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}. \tag{3.30}
$$

For different temperatures of the system, different forcefields should be computed. In this thesis, the centroid potentials are obtained at 3 K, 4 K, 5 K and 6 K. We also create the forcefields for all those temperatures.

3.2.2 Pair-distribution function calculation

Once the forcefield is ready to be put into MMTK, the centroid molecular dynamics simulation can be performed to generate the trajectory for each cluster at different temperatures. The $g(r)$ function is calculated by extracting the position information from the trajectory data file and plotted it as a histogram.

3.3 Results and discussions

3.3.1 Structural properties from CMD

The $g(r)$ function contains information regarding the structures of the clusters and can be used to estimate the number of nearest neighbours and the shell structures. The $g_c(r_c)$ function is calculated in the CMD simulation. By utilizing the centroid pseudo-potential, the calculation on the clusters can be performed for relatively large cluster sizes, even up to $N = 400$ for various temperatures. The discussion of $g_c(r_c)$ is separated into three parts: we categorize those clusters with $N \leqslant 10$ as small clusters, those with $N \leqslant 100$ as medium clusters, and those with $N > 100$ as large clusters.

The $g_c(r_c)$ calculated from CMD is time saving and easy to converge. Unlike PIMC simulations, which usually take over a week to get equilibrium results and over a month to achieve convergence, CMD simulations take less than two hours for small and medium sized clusters with $N \leqslant 60$ holding a time step (dt) of 0.01fs and 10^8 steps in total. For large clusters, such as $N = 100$, the simulations need only a day or less to finish. The reducing in the simulation time is huge compared to PIMC. The excessive computational time problem and limitation to small cluster sizes can be solved by the centroid method.

Superfluid behaviour is suppressed at temperatures above 2 K. Both the CMD and PIMC simulations for small clusters difficulty reaching equilibrium there. The PIMC

method cannot provide stable results when the cluster size is less than 10 particles. Therefore, we do not provide any results from PIMC, as the acceptance ratio is too high to capture the correct behaviour. However, the CMD method is able to achieve stable and converged results for small clusters, even the dimer and trimer (Figure (3.2)).

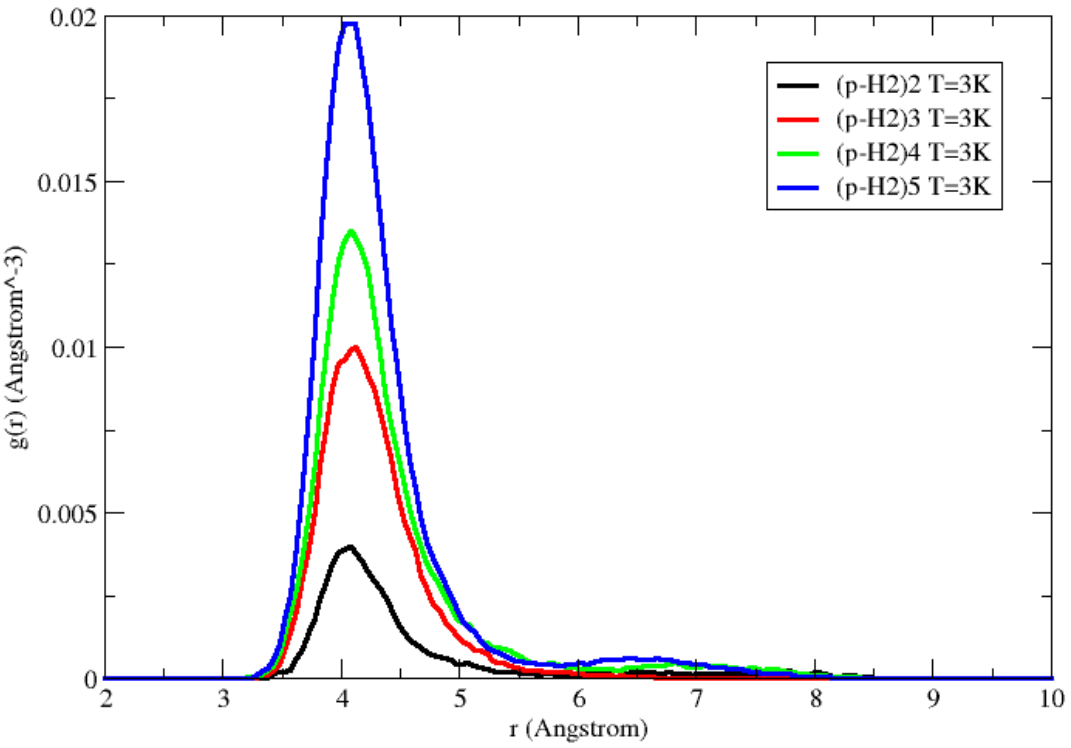

Figure 3.2: $g_c(r_c)$ of p-H_2 at 3 K for $N = 2 - 5$.

For (p-H_2)$_N$ with N =2−5, the $g_c(r_c)$'s show a similar position for the major peak,

which is located around 4.0 Å. Note that the horizontal axis in Figure (3.2) represents the distance between any two p-H_2 molecules within the cluster. The peak height increases with cluster size. Starting with (p-H_2)$_{N=4}$, the second shell appears to be filled. A small bump can be observed between 6.0 and 8.0 Å. The actual structure of the cluster is a slightly distorted tetrahedral shape see Figure (3.3) in which light blue spherical particles.

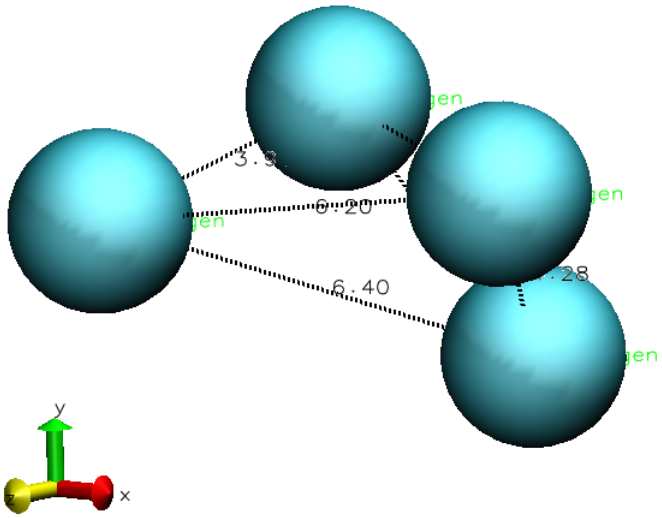

Figure 3.3: The distorted tetrahedral structure of (p-H_2)$_{N=4}$

Unfortunately, the $g_c(r_c)$'s of small clusters at higher temperatures fail to provide a smooth structure, even for very long simulations. For example, the $g_c(r_c)$'s at $T = 4$ K still show the expected general shape, but are very noisy(Figure (3.4)). The difficulty in finding convergence may be due to the tendency for cluster evaporation above the critical temperature for superfluidity (~ 2.5 K). Additionally, the interaction in a small cluster is weak and may not hold the cluster in a fixed structure.

The $g_c(r_c)$ starts to show normal behaviour when the size increases to 10 where two clear peaks can be found on the $g_c(r_c)$ plot (Figure (3.5)). The major peak lies around 4.0 Å, and a minor peak is centred around 7.0 Å. It is interesting to note that the minor peaks for each $g_c(r_c)$ shares almost the same peak position and height. However, although the peak positions are the same, the first peak heights decrease as the temperature increases. This fact is raised because the potential well depth is deeper at higher temperature. The movement among molecules may more active than at lower temperature. It suggests the possibility of the p-H_2 molecules contributing less to the first shell and a tendency to diffuse into the next shell.

The $g_c(r_c)$'s of medium clusters, $10 < N \leqslant 100$, can be obtained from both the PIMC and CMD methods. Taking $(\text{p-H}_2)_{N=30}$ as an example, the $g_c(r_c)$ from CMD contains three distinct peaks representing a three-shell structure (Figure (3.6)). At T =3 K and 4 K, the behaviours of the $g_c(r_c)$'s are similar except for a slightly higher first peak at 4 K. Both 3

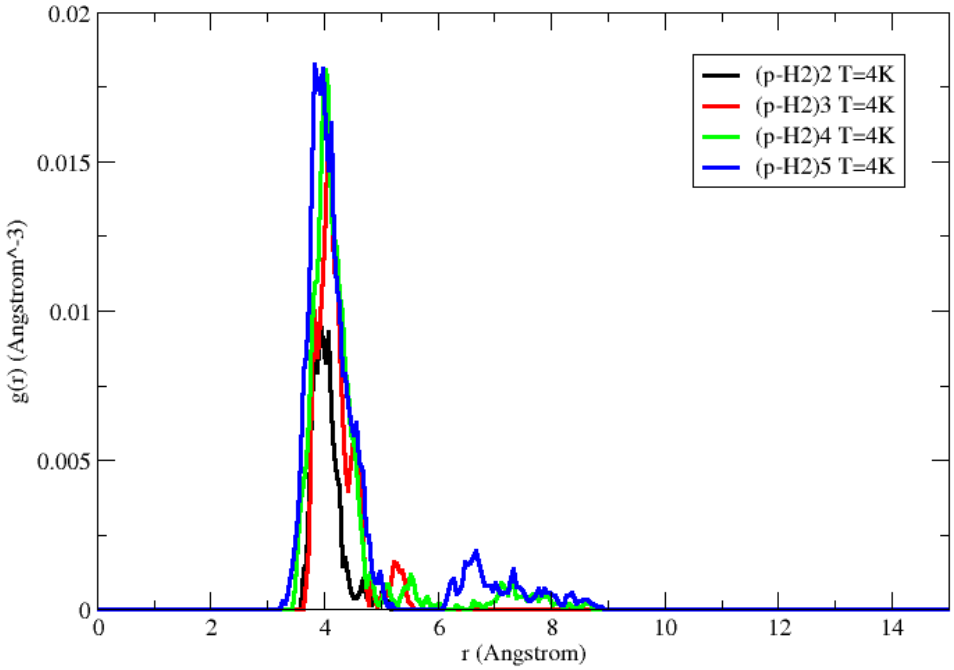

Figure 3.4: $g_c(r_c)$ of p-H_2 at 4 K for $N = 2 - 5$. The plot remains noisy even for long simulation and small time step. The cluster tends to evaporate at temperature higher than 2.5 K.

K and 4 K $g_c(r_c)$'s show a shoulder behind the second peak indicating that the molecules start entering the next shell. This feature is less visible on the $g_c(r_c)$ at 5 K and 6 K for which a broad peak can be found for the second shell, which provides evidence for melting and hence a reduction in the solid-like structural behaviour.

It is noticeable that the $g_c(r_c)$ of medium clusters at 4 K shows a higher peak in the

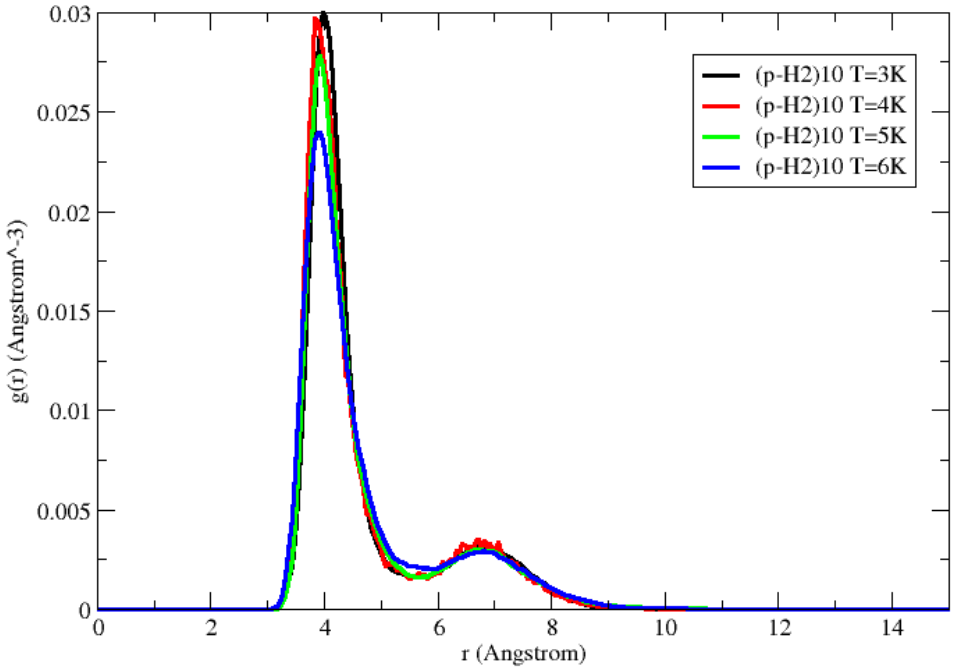

Figure 3.5: $g_c(r_c)$ of $(\text{p-H}_2)_{N=10}$ at $T = 3 - 6$ K.

first shell compares to the ones at other temperatures. This feature is universal for all the medium clusters. The heights and corresponding shell radii of the first peaks are listed in Table (3.1). We collect the data from $g_c(r_c)$'s with $N = 20-60$ (spacing of 10) at $T = 3, 4, 5$ and 6 K. For all the $g_c(r_c)$'s at $T = 4$ K, there exists a turnaround on the height of the first peak. It obeys our expectation that the clusters tend to be less structured with increasing temperature. The turnaround reminds us of quantum melting [50]. Quantum melting of a cluster refers to a cluster appearing to be solid-like at a certain temperature, whereas the

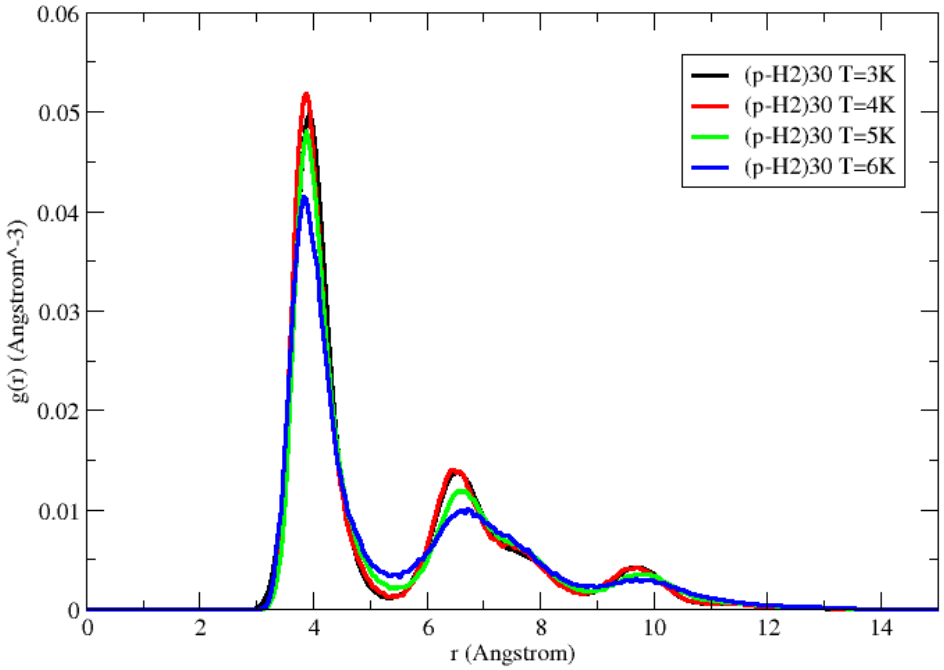

Figure 3.6: $g_c(r_c)$ of $(\text{p-H}_2)_{N=30}$ at $T = 3-6$ K. The $g(r)$ at 3 K is shown in black matches with the one at 4 K shown in red.

liquid-like features are retained at lower temperature. Our research may benchmark the preceding work related to the quantum melting of $(\text{p-H}_2)_N$. Previously, researchers found that quantum melting can be observed within $T \leqslant 4$ K and for a size range of $22 \leqslant N \leqslant 30$ [50, 20, 51]. With CMD method, we are able to observe this quantum melting feature in even larger clusters such as $(\text{p-H}_2)_{N=50}$ and $(\text{p-H}_2)_{N=60}$. As shown in Figure (3.7), for all the $g_c(r_c)$'s of $(\text{p-H}_2)_N$, the turnaround stands out at $T = 4$ K. After the turnaround, the

Table 3.1: Peak height and corresponding peak position for the first shell for different cluster sizes within a temperature range from 3 K to 6 K.

Cluster size	Temperature (K)	Shell Radius	Peak Height
20	3	3.98591	0.0454152
	4	3.93527	0.049265
	5	3.87264	0.0419065
	6	3.88399	0.0358134
30	3	3.93522	0.0496897
	4	3.88145	0.0517313
	5	3.8707	0.0479828
	6	3.838844	0.0414228
40	3	3.88725	0.0505882
	4	3.86275	0.0537395
	5	3.83824	0.0486975
	6	3.83824	0.0421849
50	3	3.89672	0.049609
	4	3.83485	0.0568619
	5	3.84722	0.0512623
	6	3.82247	0.0454493
60	3	3.83524	0.0516667
	4	3.83824	0.0601225
	5	3.83824	0.0544853
	6	3.78922	0.048848

first peak height decreases linearly with on temperature, which reveals quantum melting. In addition, the pseudo-potential obtained at $T = 4$ K shows larger can cutoff radius than the one at $T = 5$ K (Figure (2.10)). This feature indicates that larger cutoff radius allows more particles enter to the first shell, and hence higher first peak of the $g_c(r_c)$ can be observed.

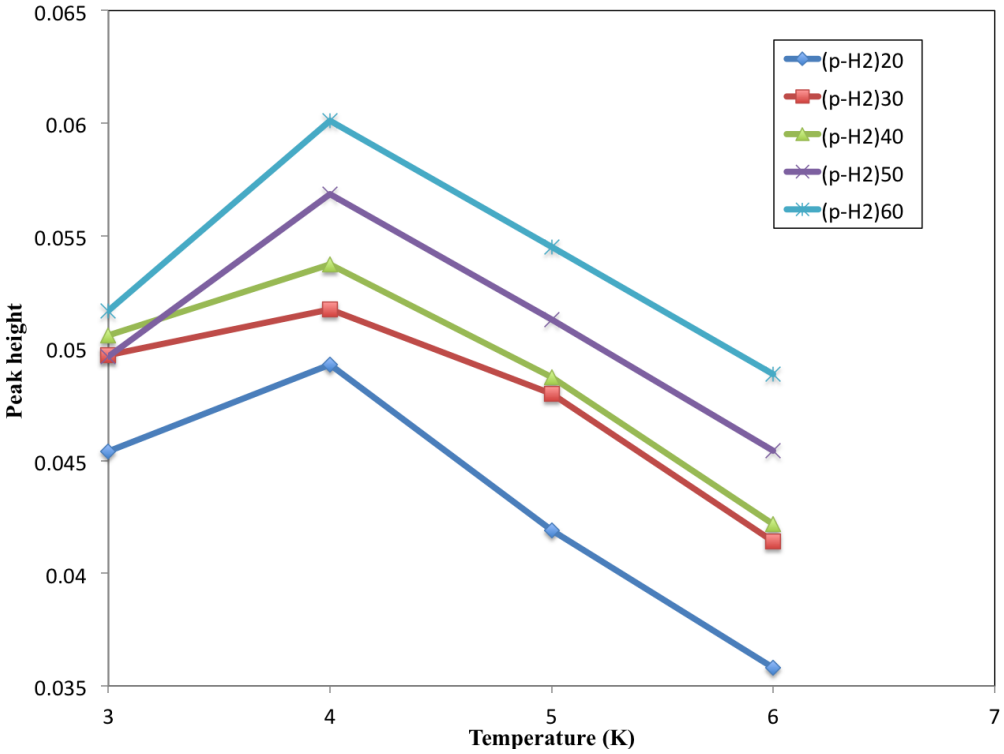

Figure 3.7: The height of the first peak in $g(r)$ of $(\text{p-H}_2)_N$ $N = 20, 30, 40, 50$ and 60 at $T = 3 - 6$ K.

CMD is able to perform calculations for the large clusters with sizes up $N = 400$ (Figure (3.8)). The $(\text{p-H}_2)_{N=400}$ distribution displays a liquid-like feature with the shell radius extending to 30.0 Å. Another example for large clusters is $(\text{p-H}_2)_{N=100}$, for which we present $g_c(r_c)$ at $T = 3, 4, 5$ and 6 K (Figure (3.9)). It is evident that $(\text{p-H}_2)_{N=100}$ contains three distinct peaks corresponding to three distinct shells. The fourth and fifth shell are also visible but do not make much contribution to the $g_c(r_c)$. Across this temperature range, the cluster starts melting and tends to become structureless. Still, a point of turnaround

can be found at $T = 4$ K, for which an extra shoulder occurs on the second peak.

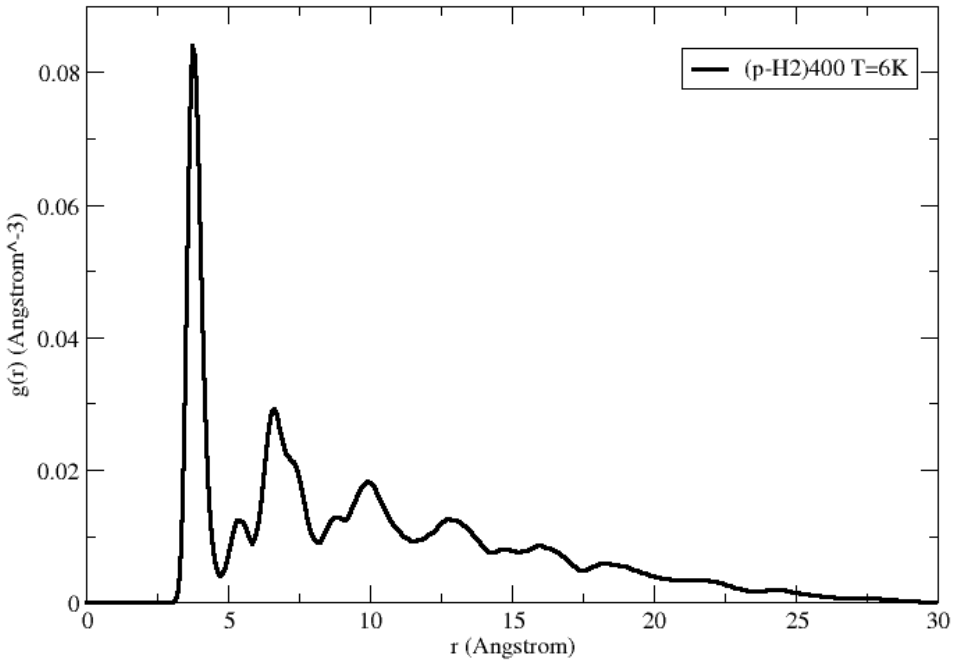

Figure 3.8: $g_c(r_c)$ of $(\text{p-H}_2)_{N=400}$ at 6 K. The cluster contains multiple peaks indicating the liquid-like feature. The shell radius extends up to 30.0 Å.

From Wales and his coworkers [71], we know that a LJ cluster containing 100 atoms displays a C_s symmetry and looks like an icosahedron. For our calculated $(\text{p-H}_2)_{N=100}$, we obtain a structure similar to that of LJ cluster (Figure (3.10)). By taking look at the trajectory after the CMD simulation, the movements of the molecules in the cluster are trapped in an icosahedral-like framework and tremble about.

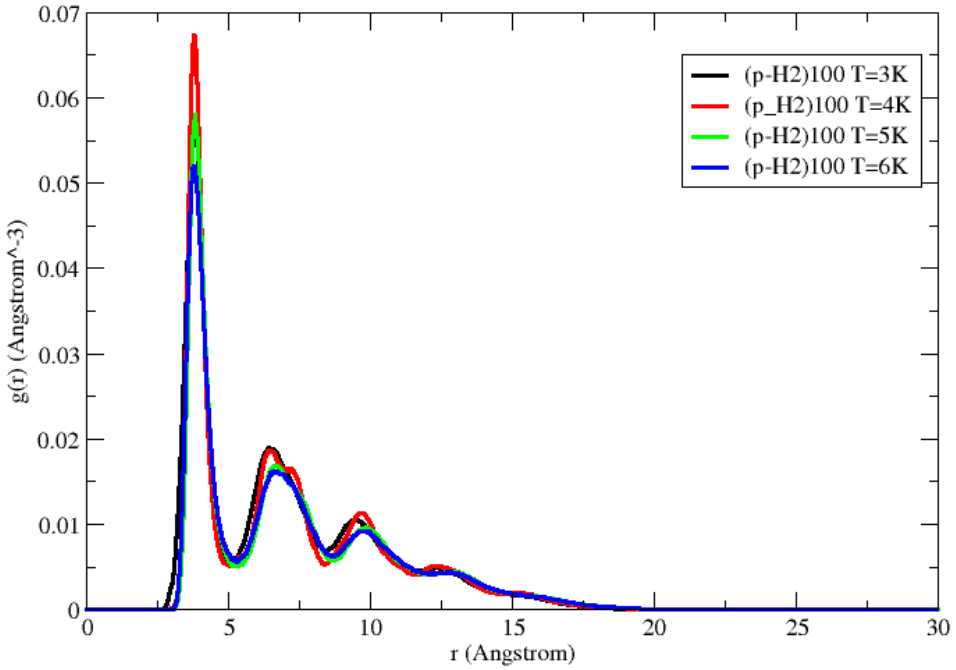

Figure 3.9: $g_c(r_c)$ of $(\text{p-H}_2)_{N=100}$ at $T = 3 - 6$ K.

To assess the accuracy of this CMD approach, we compare our $g_c(r_c)$ from CMD with that calculated by PIMC, since PIMC is able to provide the exact centroid results. A variation can be found in small clusters (Figure (3.11)). However, the difference there is mainly in the first peak. The $g_c(r_c)$ from CMD contains a higher first peak compared to that from PIMC. Other than that, a close match can be seen for the second peak at longer range. The similarity between the $g_c(r_c)$'s from CMD and PIMC increases with increasing cluster size (Figure (3.12)). As the size increase, the radial distribution functions tend

to approach to those calculated from PIMC (Figure (3.12)). The difference for the small clusters may be due to the convergence problem of PIMC. Since Monte Carlo simulations require large number of steps to ensure convergence, we may need even longer simulations on small clusters because of quantum effects. As for larger clusters, quantum effects are less important, and the system behave more classically. Therefore, better match is observed for $g_c(r_c)$'s of larger clusters. This demonstrates that the CMD method is a better solution than path integral methods of quantum many-body system for calculating structural properties.

3.3.2 Average centroid potential energy

The average potential energy, $\langle V_c \rangle$, is one of the major contributions to the energy estimator (Eq. **??**). We obtain the average potential energy by integrating the product of $g_c(r_c)$ and the centroid pair potential, $V_c(r_c)$, as shown in the equation below,

$$\langle V \rangle = \int \mathrm{d}r_c \; g_c(r_c) \; V_c(r_c). \tag{3.31}$$

Since the centroid pair potential for each temperature is given, $\langle V \rangle$ can be calculated from both the PIMC and CMD methods. The value of $\langle V \rangle$ obtained from CMD is in very good agreement with the one from PIMC. This agreement applies to all the temperatures and cluster sizes we have been studying (Figure (3.13), Figure (3.14) and Figure (3.15)).

Overall, the $\langle V \rangle$ decays slowly when the size of cluster is small and more rapid as the cluster size increases. The value of $\langle V \rangle$ from CMD decays faster than those from PIMC. In addition, the splitting between $\langle V \rangle$ from PIMC and CMD becomes more noticeable when the size gets bigger.

3.3.3 Deconvolution test

As mentioned earlier in this chapter, differences can always be found between the centroid method and the regular path-integral method in the description of $g(r)$ at low temperatures. It has been proved that the deconvolution kernel can be used to correct the centroid property and convert it back to the regular property [57].

Based on CMD and PIMC calculations, the $g_c(r_c)$'s have twice the height on the first peak compared to the regular $g(r)$. The trend of decreasing peak height is observed as the temperature increases (Figure (3.16) − Figure (3.17)).

The deconvolution is performed on both $g_c(r_c)$'s from CMD and PIMC methods. Specifically, we focus on the deconvolution of the medium clusters to figure out the reliability of the $g_c(r_c)$ on transforming it back to the regular $g(r)$. In this research, the 3D QDO's are obtained together with the centroid pseudo-potential. A close match can be found between the $g_c(r_c)$'s from both methods and the regular $g(r)$'s for the long range, starting from the

second peak of the $g(r)$'s (Figure (3.18) and Figure (3.19)). Both the deconvoluted $g(r)$'s from CMD and PIMC method show a higher first peak than do the regular $g(r)$'s. The results from the CMD method usually provide the highest peaks among the comparisons of $g_c(r_c)$'s and $g(r)$'s. In general, the structure of the regular $g(r)$'s can be revealed via application of the deconvolution kernel.

In particular, at the same temperature, the deconvoluted $g(r)$'s of small-size clusters closely match the corresponding regular $g(r)$'s in contrast with those of larger clusters. For example, Figure (3.18) shows that at $T = 3$ K, the deconvoluted $g(r)$'s are better matched to the regular $g(r)$ for $(\text{p-H}_2)_{N=20}$ rather than for $(\text{p-H}_2)_{N=40}$. For the same cluster size, we observe the regular $g(r)$'s and the deconvoluted $g(r)$'s from both methods to coincide increasingly with increasing temperature. For $(\text{p-H}_2)_{N=60}$, the close match can be found on the second peak of the $g(r)$'s (Figure (3.19)). At $T = 5$ K, the deconvoluted $g(r)$'s show a slightly better matching behaviour than that at $T = 4$ K.

Both $g_c(r_c)$'s from PIMC and CMD method are applicable for deconvolution and can obtain the $g(r)$'s similar to the regular ones. It turns out the $g_c(r_c)$'s from PIMC tend to provide better match at the higher temperature (Figure (3.19) lower panel), whereas the $g_c(r_c)$'s from CMD show a closer fit for small clusters (Figure (3.18) top panel).

3.3.4 Mass effects

The question of mass dependence was raised when testing the deconvolution function. In the PIMC simulation, the full mass of the p-H_2 dimer is used as one of the initial parameters to calculate $g_c(r_c)$. In contrast, the CMD uses the reduced mass to obtain the pseudo-potential and the deconvolution kernel. Since the centroid effective potential contains kinetic free particle motion in the Hamiltonian, the centroid potential should not depend only on temperature but also on mass. Therefore, the properties based on the pseudo-potential should also be mass dependent. However, the CMD generated $g_c(r_c)$'s that are related to full-mass and reduced-mass potentials do not differ. The only difference is found when performing the deconvolution test. The deconvoluted $g(r)$'s obtained by applying the reduced-mass kernel lose the long-range information (Figure (3.20)). Note that the results shown in the previous section are obtained using a full-mass kernel. Furthermore, since the pseudo-potential is obtained from the centroid density which contains the free particle motion involving the mass, the mass effects should be taken into account for the centroid molecular dynamics research.

3.4 Conclusion

CMD with a pairwise pseudo-potential opens the door to a wealth of applications to many-body clusters. Compared to traditional path-integral methods, CMD presents tremendous computational advantages. It also allows us to calculate the properties of much larger clusters, or even of droplets. By performing CMD simulation, we are able to obtain $g_c(r_c)$ and $\langle V_c \rangle$ for structural and energetic studies. We presented the $g_c(r_c)$'s for cluster sizes up to 400 at a series of temperatures from 3 K to 6 K. Based on the analysis of these $g_c(r_c)$'s, we are able to delineate the structural behaviour for large clusters. Also, we found a signature for quantum melting at 4 K. The $\langle V_c \rangle$'s obtained from CMD and PIMC are in agreement and decrease as the cluster size grows. The 3D deconvolution kernel was computed and used to convert the $g_c(r_c)$'s to real space. In general, the regular $g(r)$ can be determined. As the temperature and the cluster size increase, the deconvolution function appears to work increasingly well. The investigation of the deconvolution approach shows that the centroid pseudo-potential method is applicable for large quantum clusters. Issues such as the "mass effect" remain to be further investigated.

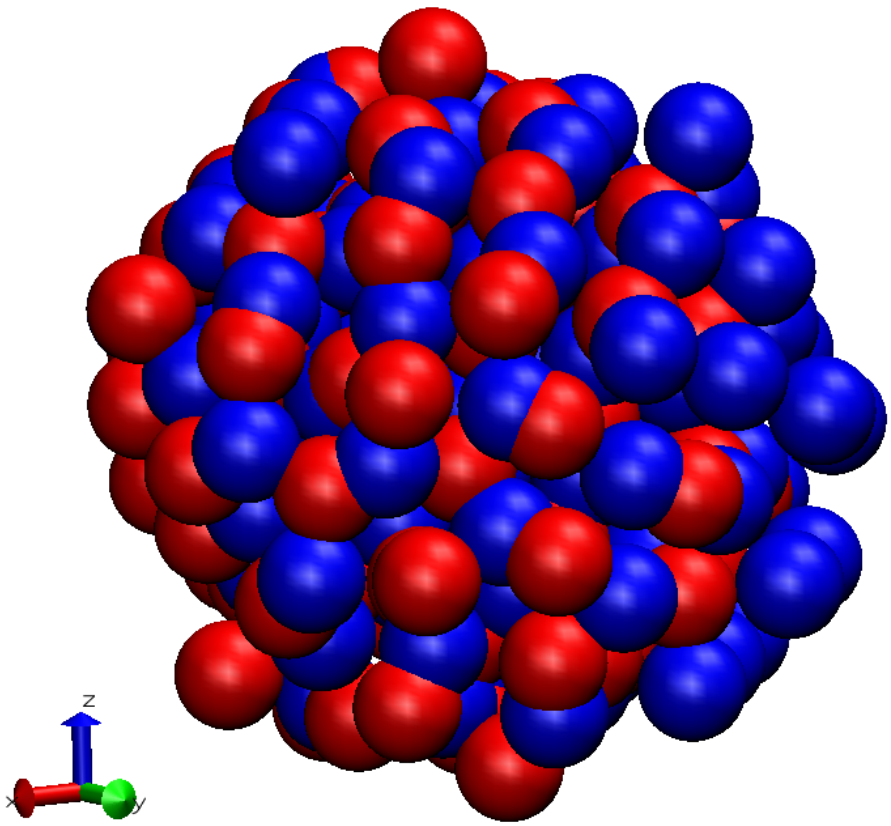

Figure 3.10: Comparison of the structure of $(\text{p-H}_2)_{N=100}$ obtained from CMD and LJ clusters [71]. Both structures show a rough icosahedral-like shape. Red spheres represent the parahydrogen molecules in the CMD simulation, whereas blue spheres are atoms in the LJ cluster.

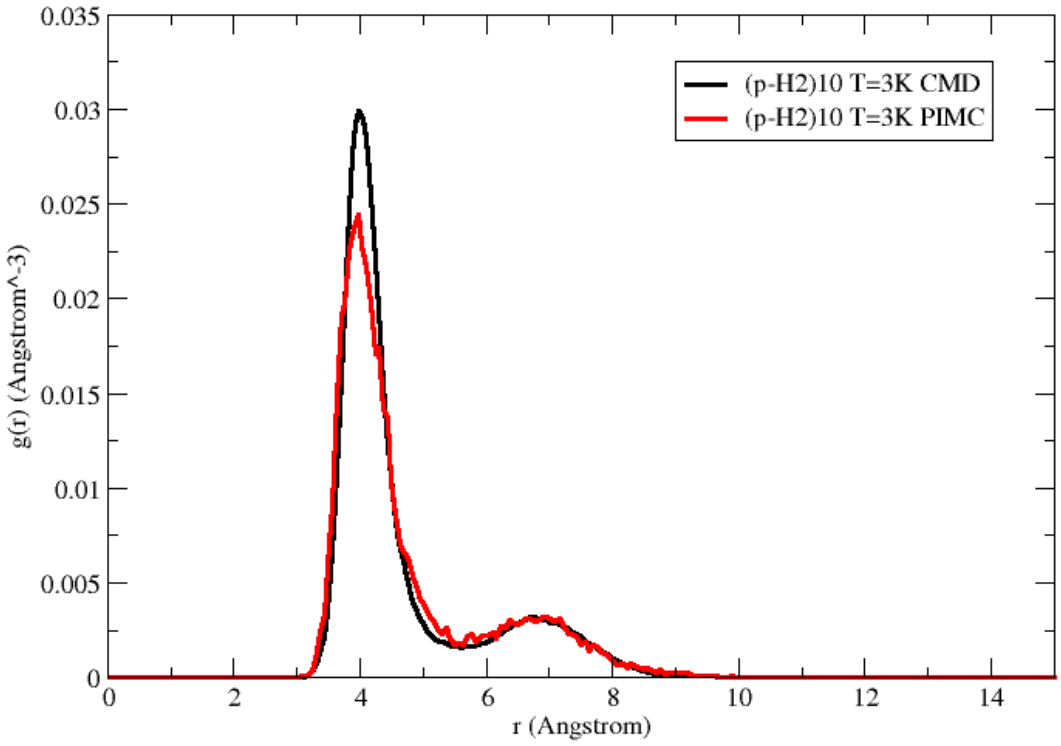

Figure 3.11: Comparison of the $g_c(r_c)$ of $(\text{p-H}_2)_{N=10}$ at $T = 3$ K obtained from CMD and PIMC.

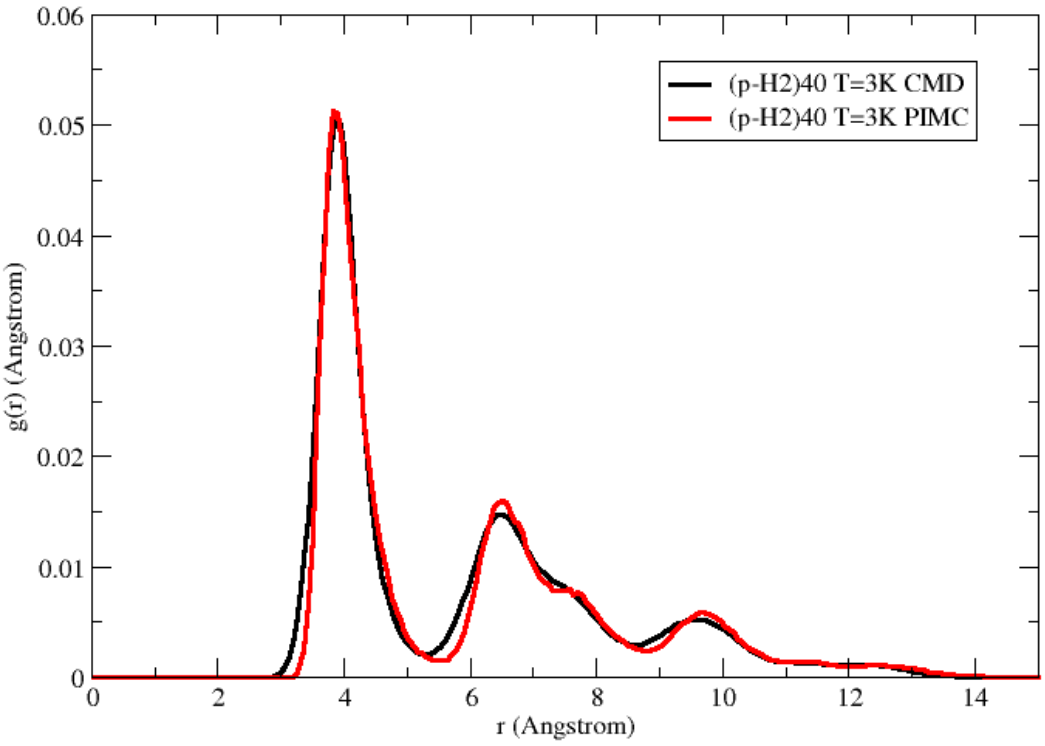

Figure 3.12: Comparison of the $g_c(r_c)$ of $(\text{p-H}_2)_{N=40}$ at $T = 3$ K obtained from CMD and PIMC. With large cluster size, the $g_c(r_c)$ from both methods are in closer agreement.

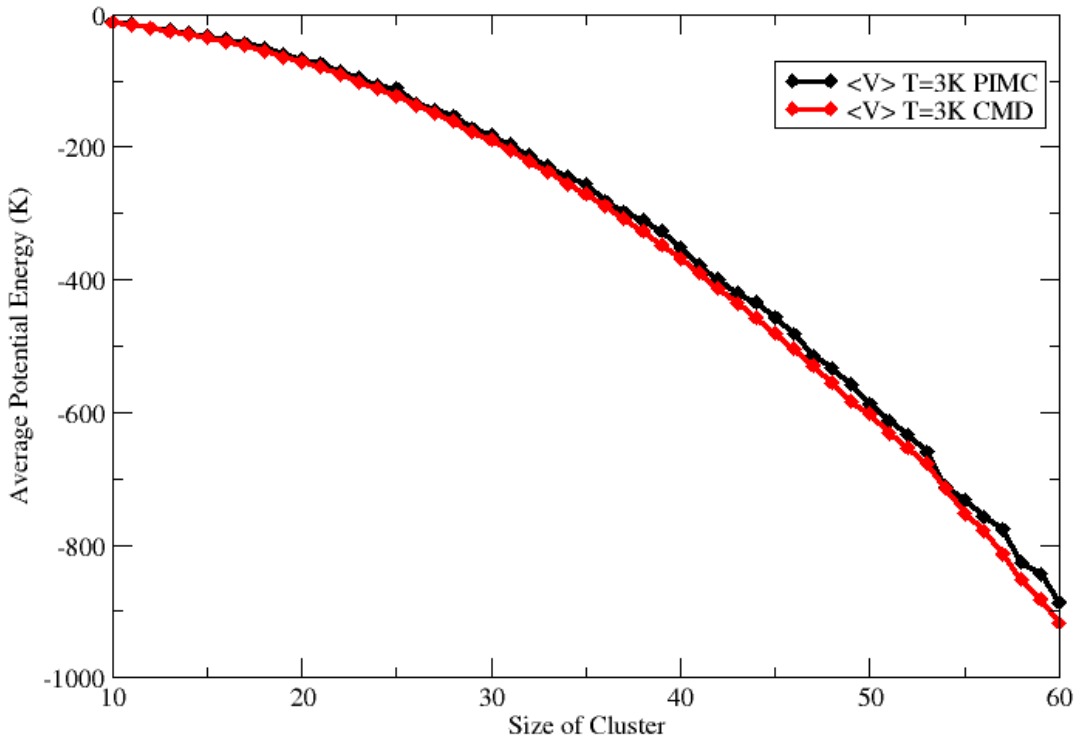

Figure 3.13: Comparison of $\langle V_c \rangle$ of (p-H_2)$_N$ $N = 10 - 60$ at $T = 3$ K from CMD and PIMC.

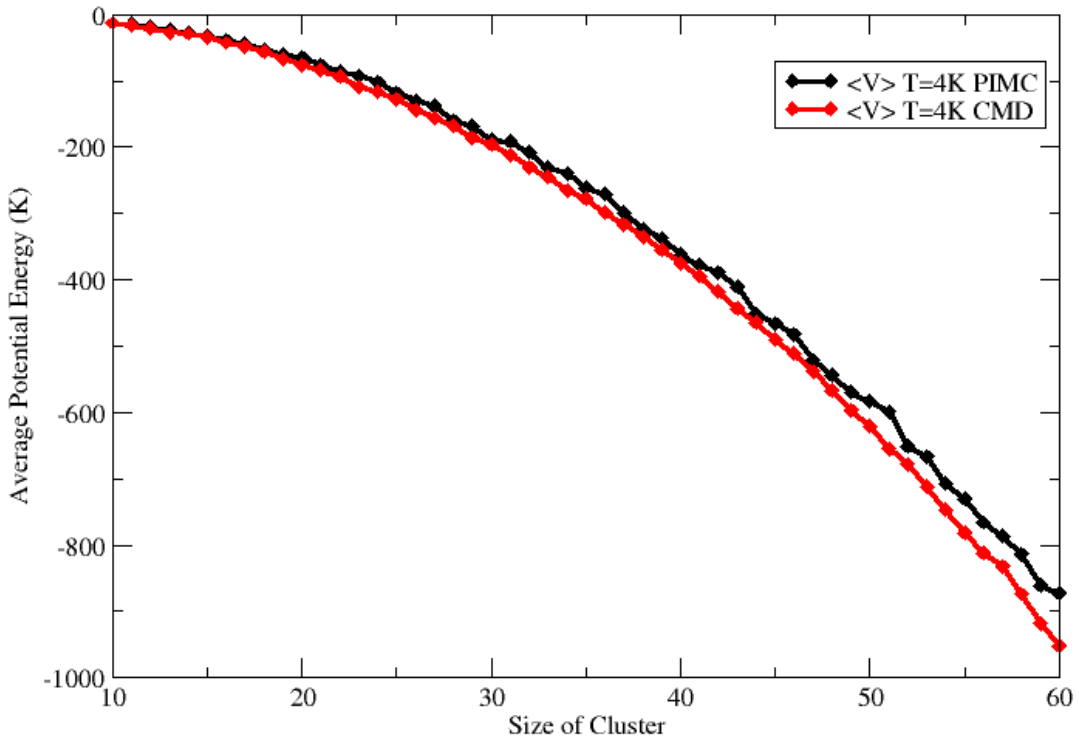

Figure 3.14: Comparison of $\langle V_c \rangle$ of $(\text{p-H}_2)_N$ $N = 10 - 60$ at $T = 4$ K from CMD and PIMC.

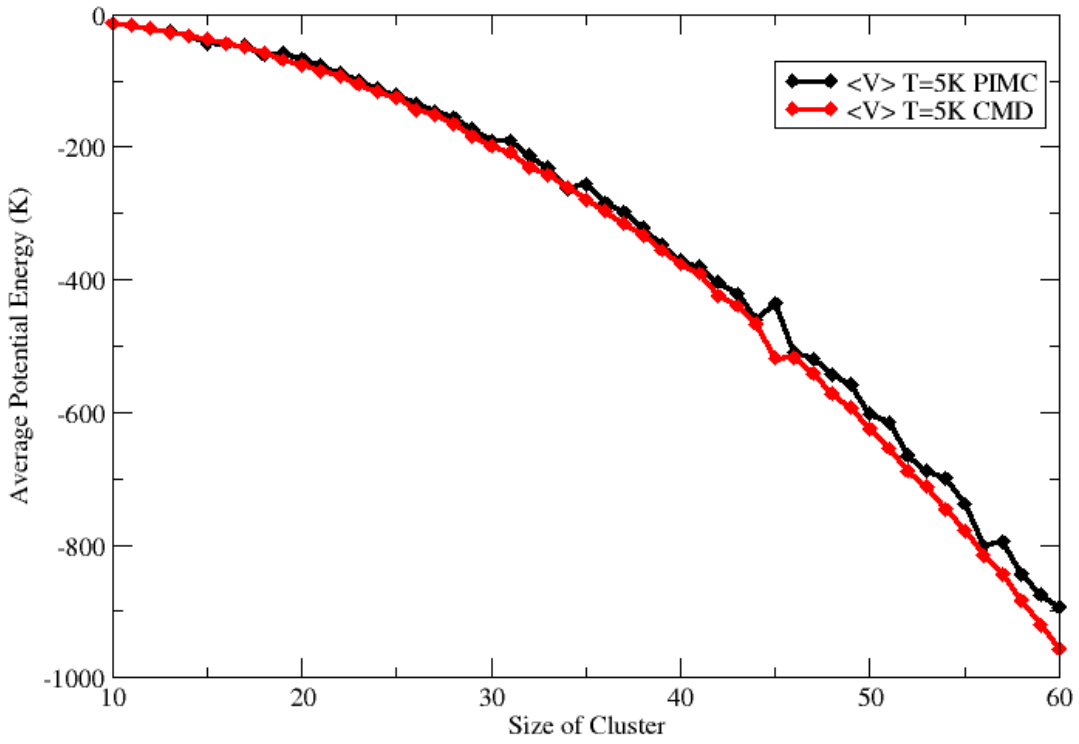

Figure 3.15: Comparison of $\langle V_c \rangle$ of (p-H_2)$_N$ $N = 10 - 60$ at $T = 5$ K from CMD and PIMC.

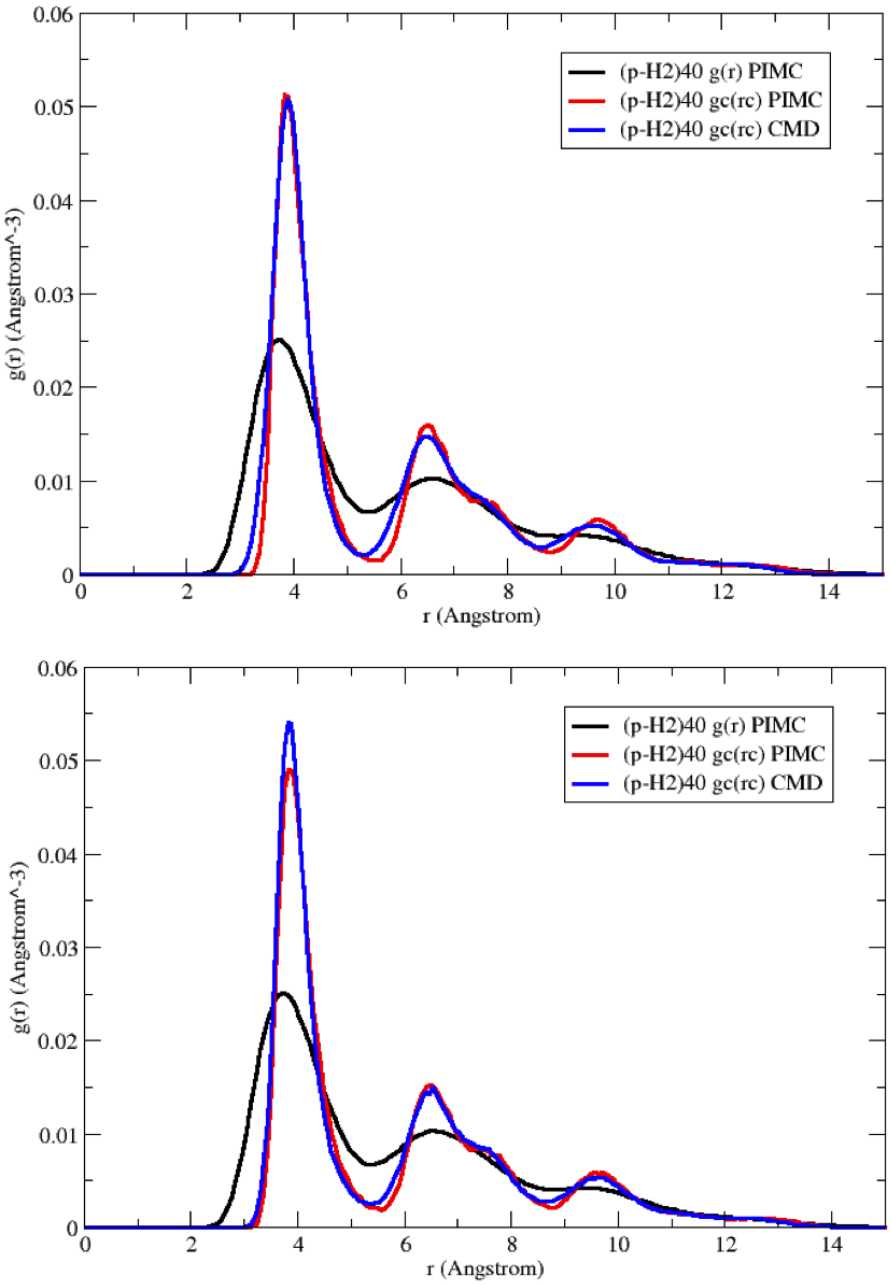

Figure 3.16: Comparison of centroid and regular $g(r)$ of $(\text{p-H}_2)_{N=40}$ at $T = 3$ K (top panel) and $T = 4$ K (lower panel). The regular $g(r)$ is obtained from PIMC.

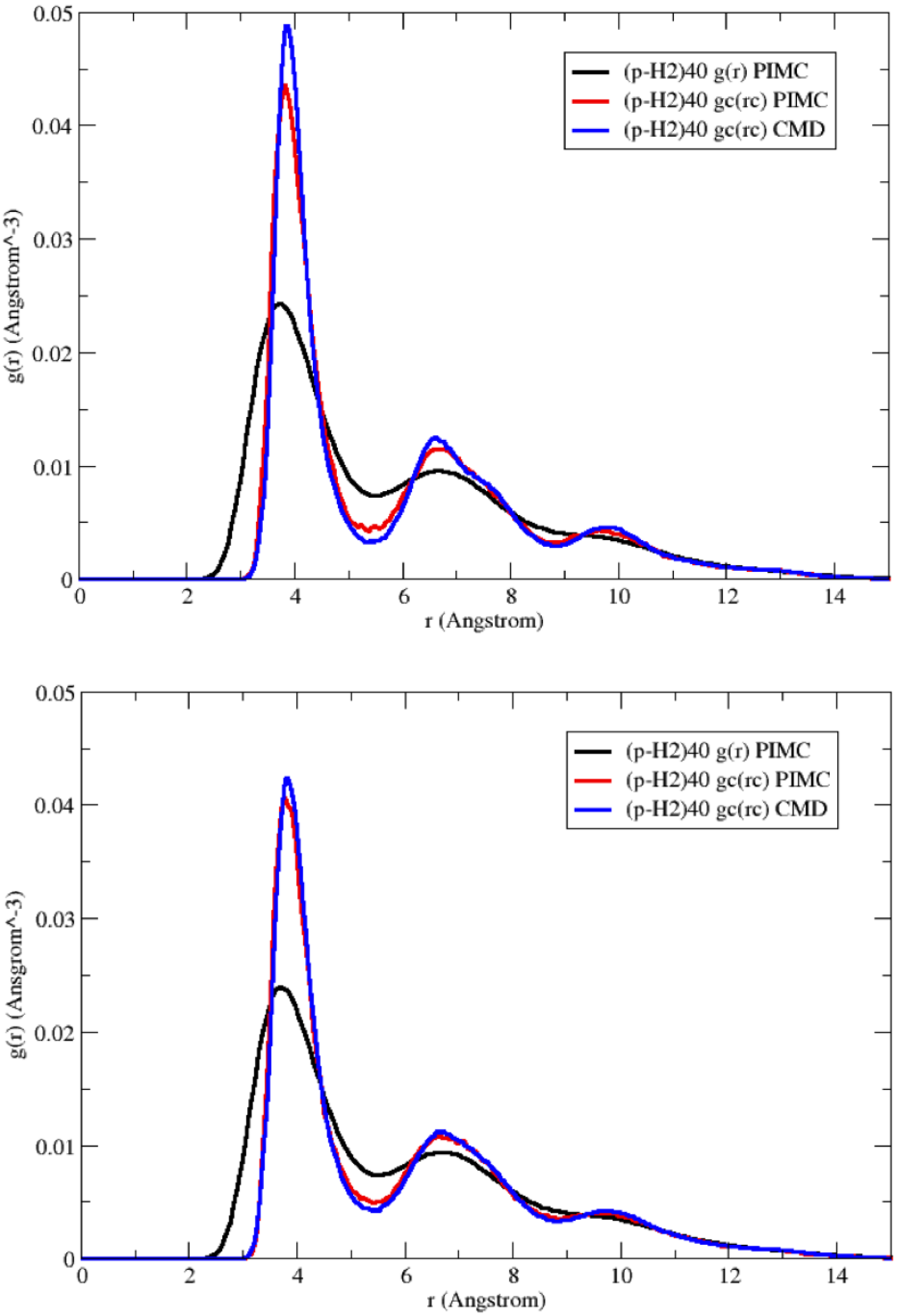

Figure 3.17: Comparison of centroid and regular $g(r)$ of $(\text{p-H}_2)_{N=40}$ at $T = 5$ K (top panel) and $T = 6$ K (lower panel).

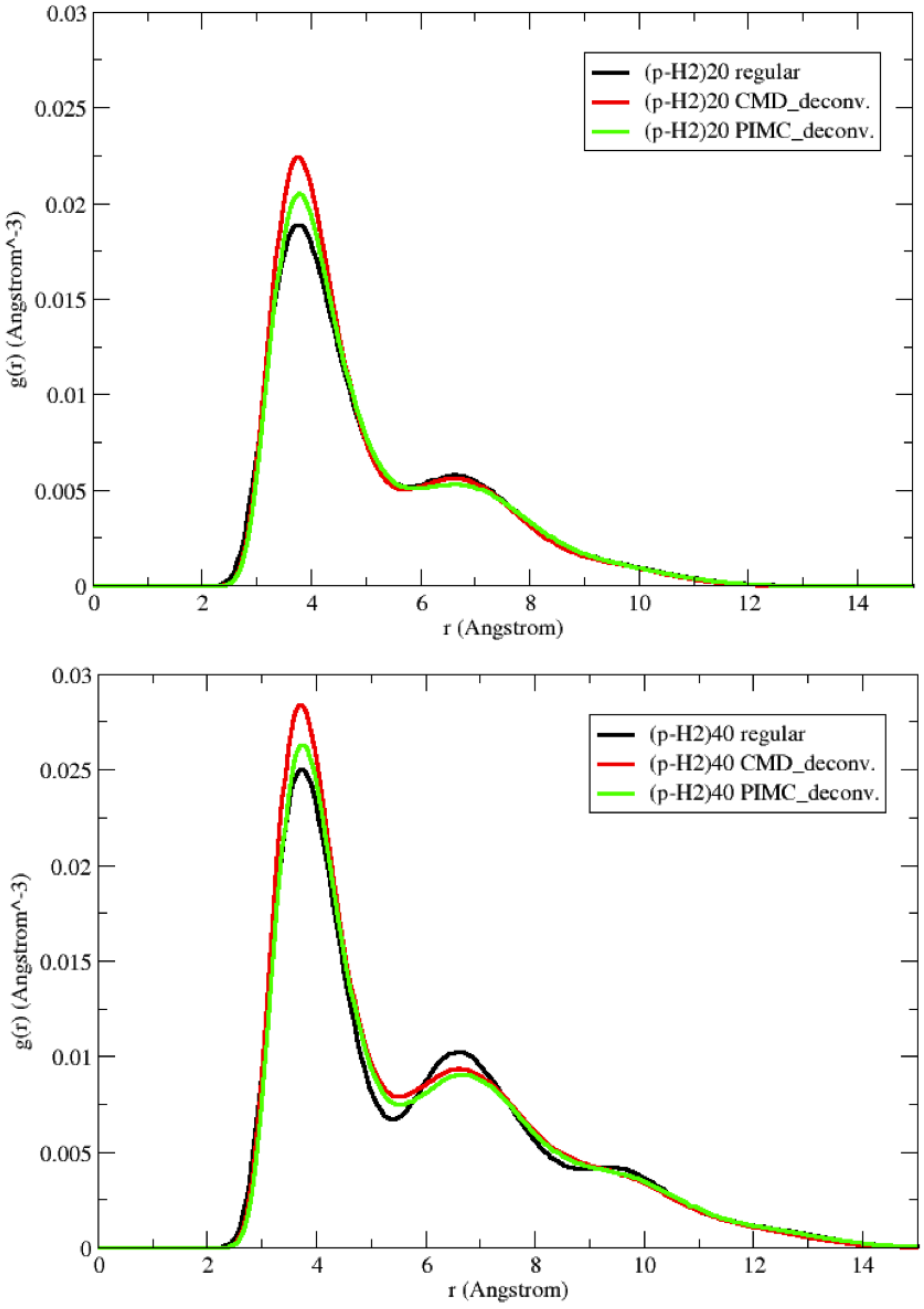

Figure 3.18: Comparison of the deconvoluted and regular $g(r)$ of $(\text{p-H}_2)_{N=20}$ (top panel) and $(\text{p-H}_2)_{N=40}$ (lower panel) at $T = 3$ K. The deconvolution is performed on both CMD and PIMC calculated $g_c(r_c)$.

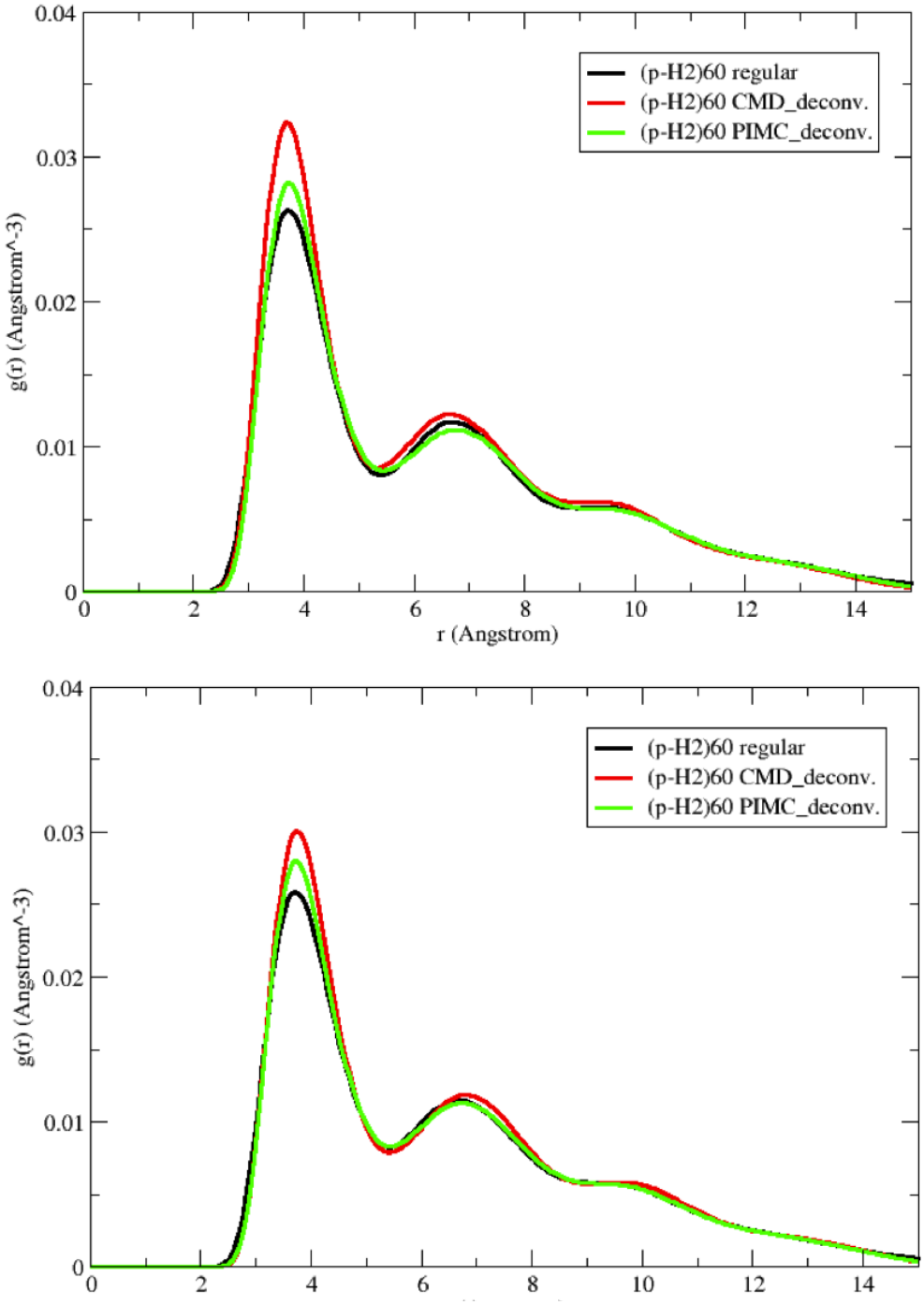

Figure 3.19: Comparison of deconvoluted and regular $g(r)$ of $(\text{p-H}_2)_{N=60}$ at $T = 4$ K (top panel) and $T = 5$ K (lower panel).

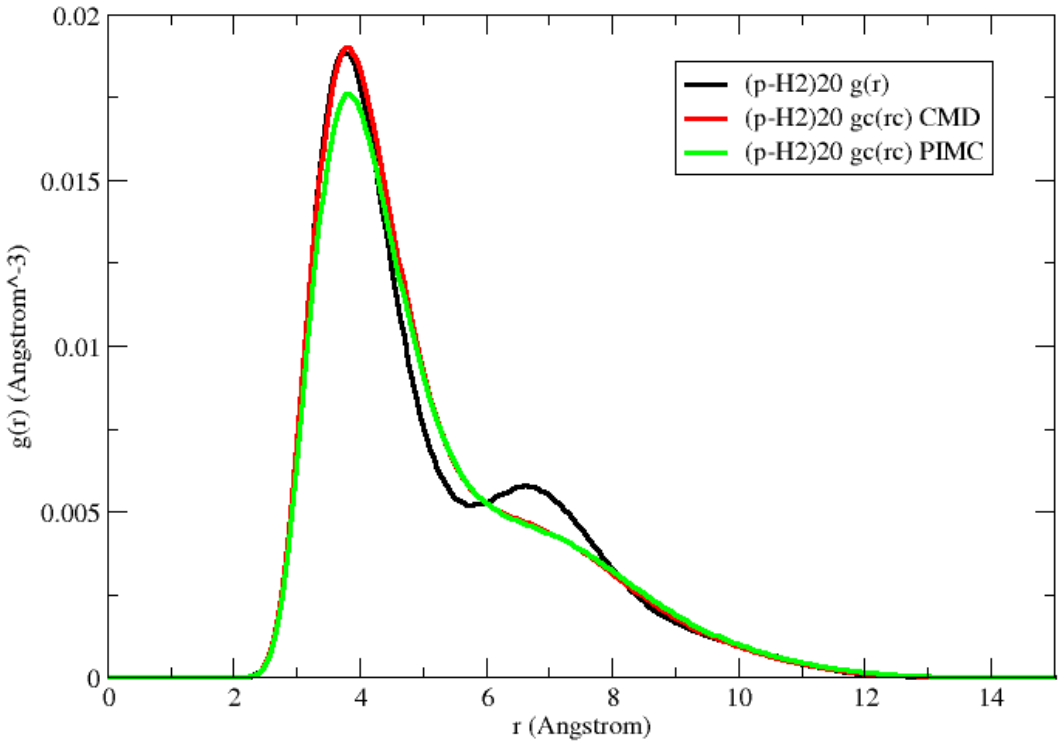

Figure 3.20: Comparison of the regular and deconvoluted $g(r)$ with the deconvolution kernel involving the reduced mass of the p-H_2 dimer. The example has been given with (p-H_2)$_{N=20}$ at $T = 3$ K.

Chapter 4

Conclusion

This thesis focused on the applicability of the centroid pseudo-potential approach for solving many-body quantum system problems at low temperature. Parahydrogen clusters were used as a test case, and structural and energetic properties were studied and compared to their exact counterparts obtained from the PIMC method. The pseudo-potential for the parahydrogen dimer was obtained from the centroid potential of mean force of the dimer. An analytical potential based on the MLR form with long-range coefficients C_8 and C_{10} was used to represent the pair centroid pseudo-potential. These analytic pseudo-potentials were then used to perform centroid molecular dynamics simulations and obtain cluster properties for various sizes and temperatures. The results were in quite good agreement with the exact PIMC simulations. This work has shown that the use of pseudo-potentials

is a powerful tool for solving large many-body problems.

For the construction of the centroid pseudo-potential, the centroid density for the parahydrogen dimer was the first quantity to compute. Initial efforts were limited to 1D since the NMM method is costly in higher dimensions. Instead, the PIMC method was rather used to the obtain a 3D-based centroid density. This method is flexible for calculating any potential for any number of dimension. The centroid potential obtained from the density formula contains noisy data, and hence needed to be fitted to a smooth analytical form. The MLR form was chosen as it provide a close fit to the centroid potential. The centroid pseudo-potential is temperature dependent. The potential becomes deeper and its equilibrium distance approaches that of the classical potential when the temperature is increased. Unlike the classical potential which has a continuous repulsive wall at the short-range, the pseudo-potential has a finite value at short range.

Once the centroid potential was computed, the centroid molecular dynamics can be carried out. The calculation is fast and can easily be used to treat large clusters. We were able to obtain the centroid radial distribution function for parahydrogen clusters with size up to $N = 400$ within the temperature range of 3 K to 6 K. It is found that the centroid molecular dynamics reached convergence quicker than does the PIMC method. Based on the study of structural properties of parahydrogen clusters in terms of the centroid radial distribution function, quantum melting can be found to occur at 4 K. Compared to the

centroid radial distribution function calculated from PIMC, a close match can be observed at large cluster sizes. Through the observation of the average potential energy, both the CMD and PIMC method provided similar results showing decreasing values with increasing cluster size. Furthermore, a deconvolution approach was used to recover to real space radial distribution from the centroid disbution. The deconvolution function is found to be mass dependent. Future research should examine this mass effect.

4.1 Ongoing efforts

Many questions regarding parahydrogen clusters still remain unanswered. During the period of potential fitting, we were trying to find the connection between temperature and all the parameters entering the potential form. The first problem is the missing C_6 term in the long-range. The basis for the effective potential is the Buck potential [48] which contains dispersion coefficients C_6, C_8 and C_{10}. However, the effective potential was found to be better behaved with only C_8 and C_{10} in the form. In the future, an even wider range of temperature for the potential should be studied, and the general trend for the long-range coefficients could be found.

For the centroid molecular dynamics, the simulation can be performed successfully for cluster sizes less than 500. However, if the cluster size goes even larger, the energy

calculation blows up right after the energy minimization steps, and the dynamics could not be performed. Further analysis of simulation parameters such as the friction is required for larger clusters. Use of the newly implemented MMTK GPU acceleration should allow us to tackle much larger problems, all the way to droplets.

Finally, the mass dependent deconvolution function needs to be studied more carefully. There are two problems we need to solve. First, as mentioned in Chapter 3, the deconvolution function generated with the full mass of the parahydrogen dimer can recover regular radial distribution function but not the one with the reduced mass. We need to find out what makes the deconvolution function with the reduced mass miss the description of the long-range. Secondly, even if the deconvolution recovers the regular function, a variation on the first peak is still observed. This may be due to the importance of the low temperature quantum effects that lie beyond our simple pair approximation. As the temperature goes higher, the variation of the first peak of the radial distribution function may disappear. Therefore, a larger range of temperatures must be investigated.

4.2 Future directions

Several areas still deserve future investigation. According to the recent paper published by my colleague Chris Ing [41], the path integral molecular dynamics approach has been

implemented in MMTK and applied to He-CO_2 clusters. A portion of the code allows the calculation of centroid constrained trajectories. We can make use this approach to calculate the centroid potential by integrating the centroid force. This calculated potential will be the true many-body centroid potential. We can compare this potential with the pseudo-potential we obtained from a pairwise-additive approximation and to get a better idea of the nature of our approximation. Furthermore, the short-range behaviour can be sampled with the umbrella sampling tool. In that way, we can figure out whether the noise is physically existed or due to sampling error.

Much work remains to be performed on parahydrogen clusters. The dynamical properties and the confinement at low temperatures come to mind. With a deep study of those properties, we may be able to provide further insight and fundamental knowledge that could be used in the area of hydrogen storage. Large scale simulations of low temperature molecular hydrogen adsorption on organometallic surfaces [72] could be achieved using the pseudo-potential approach and possibly contribute to the field of hydrogen storage.

References

[1] P. Jena, B. K. Rao, and S. N. Hibbs, *Physics and Chemistry of Small Clusters* (New York, 1987 Plenum Press, 1987).

[2] A. Hira, A. Ray, and J. Fry, *An Ab-Initial Study of the Interaction of Atomic Hydrogen with Lithium Clusters*, chapter Electronic Structure and Properties, pages 413–419, New York, Plenum Press, 1986.

[3] J. Doye and D. Wales, New J. Chem. **22**, 733 (1998).

[4] J. Doye, D. Wales, and R. Berry, J. Chem. Phys. **103**, 4234 (1995).

[5] B. Teo and N. Sloane, Inorg. Chem. **24**, 4545 (1985).

[6] W. Knight et al., Phys. Rev. Lett. **52**, 2141 (1984).

[7] O. Echt, K. Sattler, and E. Recknagel, Phys. Rev. Lett. **47**, 1121 (1981).

[8] D. Prekas, C. Luder, and M. Velegrakis, J. Chem. Phys. **108**, 4450 (1998).

[9] H. R. Glyde, *Rare Gas Solids* (Academic, New York, 1976, 1976).

[10] R. Aziz, V. Nain, J. Carley, W. Taylor, and G. McConville, J. Chem. Phys. **70**, 4330 (1979).

[11] C. Cleveland, U. Landman, and R. Barnett, Phys. Rev. B **39**, 117 (1989).

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

[13] P. Kapitza, Nature **141**, 74 (1938).

[14] J. Allen and A. Misener, Nature **142**, 643 (1938).

[15] V. Ginzburg and A. Sobyanin, Jetp Lett-Ussr **15**, 242 (1972).

[16] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, chapter Statistical Mechanics, pages 267–320, New York, Mcgraw-Hill C1965, 1965.

[17] D. Scharf, M. Klein, and G. Martyna, J. Chem. Phys. **97**, 3590 (1992).

[18] T. Beck, J. Doll, and D. Freeman, J. Chem. Phys. **90**, 5651 (1989).

[19] J. Cuervo, P. Roy, and M. Boninsegni, J. Chem. Phys. **122** (2005).

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

[21] M. Pavese and G. Voth, Chem. Phys. Lett. **249**, 231 (1996).

[22] G. Herbert, *Classical Mechanics* (Addison-Wesley Press 1957, c1950, 1922).

[23] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems*, chapter Fermi Systems, pages 120–170, McGraw-Hill C1971, 1937.

[24] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems*, chapter Bose Systems, pages 198–223, McGraw-Hill C1971, 1937.

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

[26] J. Doll, T. Beck, and D. Freeman, J. Chem. Phys. **89**, 5753 (1988).

[27] B. Berne and D. Thirumalai, Annu. Rev. Phys. Chem. **37**, 401 (1986).

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

[29] K. Hinsen, J. Comput. Chem. **21**, 79 (2000).

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

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

[32] J. Cao and G. Voth, J. Chem. Phys. **101**, 6184 (1994).

[33] D. Thirumalai, E. Bruskin, and B. Berne, J. Chem. Phys. **79**, 5063 (1983).

[34] D. Scharf, G. Martyna, and M. Klein, J. Chem. Phys. **99**, 8997 (1993).

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

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

[37] U. Buck, F. Huisken, J. Schleusener, and J. Schaefer, J. Chem. Phys. **74**, 535 (1981).

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

[39] J. Barker, J. Chem. Phys. **70**, 2914 (1979).

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

[41] C. Ing et al., J. Chem. Phys. **136** (2012).

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

[43] R. J. Le Roy et al., J. Chem. Phys. **131** (2009).

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

[45] R. J. Le Roy et al., J. Chem. Phys. **123**, 204304 (2005).

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

[47] T. Williams and C. Kelley, *gnuplot4.7 An Interactive Plotting Program*, 2011.

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

[49] R. J. Le Roy, *RKR1 2.0 A Computer Program Implementing the First-Order RKR Method for Determining Diatomic Molecule Potential Energy Functions* (University of Waterloo, 2004).

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

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

[52] G. M. Torrie and J. P. Valleau, J. Comput. Phys. **23**, 187 (1977).

[53] J. Cao and G. Voth, J. Chem. Phys. **99**, 10070 (1993).

[54] J. Cao and G. Voth, journal of chemical physics **101**, 6168 (1994).

[55] F. Bermejo et al., Phys. Rev. Lett. **84**, 5359 (2000).

[56] M. Zoppi, M. Neumann, and M. Celli, Phys. Rev. B **65** (2002).

[57] N. Blinov and P. Roy, J. Chem. Phys. **120**, 3759 (2004).

[58] S. Jang and G. Voth, J. Chem. Phys. **111**, 2357 (1999).

[59] N. Blinov and P. Roy, J. Chem. Phys. **115**, 7822 (2001).

[60] C. David, *Introduction to Modern Statistical Mechanics* (New York : Oxford University Press, 1987).

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

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

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

[64] R. Guardiola and J. Navarro, J. Chem. Phys. **128** (2008).

[65] F. Mezzacapo and M. Boninsegni, Phys. Rev. Lett. **100** (2008).

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

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

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

[69] T. Hone and G. Voth, J. Chem. Phys. **121**, 6412 (2004).

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

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

[72] S. J. Kolmann, B. Chan, and M. J. T. Jordan, Chem. Phys. Lett. **467**, 126 (2008).

[73] P. Roy, S. Jang, and G. Voth, J. Chem. Phys. **111**, 5303 (1999).

[74] N. Blinov and P.-N. Roy, A simple model for superfluid environment based on an effective centroid potential, Unpublished.

[75] D. A. McQuarrie, *Quantum Mechanics* (University Science Books, 1983).