Coarse-graining and quantum-classical adaptive coupling in soft matter

Dissertation
zur Erlangung des Grades
“Doktor der Naturwissenschaften”

am Fachbereich Physik
der Johannes Gutenberg-Universität Mainz

Adolfo Poma

geboren in Peru (Lima)

Max-Planck-Institut für Polymerforschung
Mainz, February 2011

Tag der mündlichen Prüfung: 09. Mai 2011

Dedicated to my family

Zusammenfassung

Die vorliegende Arbeit untersucht den Zusammenhang zwischen Skalen in Systemen weicher Materie, der für Multiskalen-Simulationen eine wichtige Rolle spielt. Zu diesem Zweck wurde eine Methode entwickelt, die die Approximation der Separierbarkeit von Variablen für die Molekulardynamik und ähnliche Anwendungen bewertet. Der zweite und größere Teil dieser Arbeit beschäftigt sich mit der konzeptionellen und technischen Erweiterung des "Adaptive Resolution Scheme" (AdResS), einer Methode zur gleichzeitigen Simulation von Systemen mit mehreren Auflösungsebenen. Diese Methode wurde auf Systeme erweitert, in denen klassische und quantenmechanische Effekte eine Rolle spielen.

Die oben genannte erste Methode benötigt nur die analytische Form der Potentiale, wie sie die meisten Molekulardynamik-Programme zur Verfügung stellen. Die Anwendung der Methode auf ein spezielles Problem gibt bei erfolgreichem Ausgang einen numerischen Hinweis auf die Gültigkeit der Variablenseparation. Bei nicht erfolgreichem Ausgang garantiert sie, dass keine Separation der Variablen möglich ist. Die Methode wird exemplarisch auf ein zweiatomiges Molekül auf einer Oberfläche und für die zweidimensionale Version des Rotational Isomer State (RIS) Modells einer Polymerkette angewandt.

Der zweite Teil der Arbeit behandelt die Entwicklung eines Algorithmus zur adaptiven Simulation von Systemen, in denen Quanteneffekte berücksichtigt werden. Die Quantennatur von Atomen wird dabei in der Pfadintegral-Methode durch einen klassischen Polymerring repräsentiert. Die adaptive Pfadintegral-Methode wird zunächst für einatomige Flüssigkeiten und tetraedrische Moleküle unter normalen thermodynamischen Bedingungen getestet. Schließlich wird die Stabilität der Methode durch ihre Anwendung auf flüssigen para-Wasserstoff bei niedrigen Temperaturen geprüft.

Summary

This thesis investigates the connection between the length scales in soft matter systems, which is very important in the field of multiscale modeling. For this purpose a method was developed to evaluate the approximation of separation of variables in molecular dynamics and related fields. A second issue, and the main part of this thesis, concerns the conceptual and technical extension of the "Adaptive Resolution Scheme" (AdResS), a method that allows the simulation of a system with concurrent scales, to situations where quantum effects play a role.

The first method mentioned above requires only the analytical form of the potential as provided in most of the molecular dynamics programs. The outcome of the application to a particular problem gives, in the case of a positive assessment, a numerical indication about the validity of the separation of variables and in the negative case the evaluation guarantees strictly that no separation will be possible. This method is then applied to a diatomic molecule on a flat surface and the 2D version of the Rotational Isomer State (RIS) model of a polymer chain.

The second part of this thesis is about the development of an algorithm to perform an adaptive resolution simulation where quantum effects can be included, by mapping the quantum nature of an atom onto a classical polymer ring representation within the path integral formalism. The path integral/adaptive method is tested in a model liquid of monoatomic and tetrahedral molecules at standard (ambient) thermodynamic condition. Finally, the robustness of the method is assessed by using it to study liquid *para*-hydrogen at low temperatures.

Contents

Introduction **1**

1 Systematic Coarse Graining **3**
- 1.1 Effective potential for coarse graining 4
- 1.2 Structure-based coarse-graining 6
- 1.3 Nonbonded interaction potentials 6
  - 1.3.1 The Iterative Boltzmann Inversion 7
  - 1.3.2 Pressure correction 7
- 1.4 Limitation of structure-based coarse-graining 8

2 Approximation of Separation of Variables **9**
- 2.1 The ASV: a first criterion 10
- 2.2 Dependence of Q on the fixed point: A complementary criterion 11
- 2.3 A guiding example: A rigid diatomic molecule interacting with a surface 12
  - 2.3.1 The factor of quality Q 13
  - 2.3.2 The complementary criterion 14
- 2.4 Extension to higher dimensions: Parametric study 15
- 2.5 A second example: 2D version of the RIS model of a polymer chain 16
  - 2.5.1 A general problem in coarse grained modeling of bonded interactions 17
  - 2.5.2 Mapping schemes 17
- 2.6 Results 18
  - 2.6.1 The 1:2 MS 19
  - 2.6.2 The 1:3 MS 23
  - 2.6.3 Average of Q in fixed point and CG variable parametric space 27
- 2.7 Preliminary conclusions 29

3 The Adaptive Resolution Scheme **31**
- 3.1 The equation of motion 32
- 3.2 Thermodynamic equilibrium 33
- 3.3 Theoretical principles of thermodynamic equilibrium in AdResS . 34
  - 3.3.1 The geometrically induced phase transition 34

3.3.2 Thermodynamic quantities in AdResS 35
3.4 Application to simple liquid of tetrahedral molecules 36
3.5 Further development 38

4 Path Integral formulation 41
4.1 Derivation of the formalism 41
4.2 The Path Integral Molecular Dynamics (PIMD) 45
4.3 Static properties from PIMD 47
4.4 Limitations of the PIMD and possible solutions 48
4.5 The Staging transformation 49
4.6 The Normal Mode transformation 51
4.7 Similarities between the path integral formalism and the statistics of ring polymers 54
4.7.1 The free ring polymer 54
4.7.2 Rouse theory for ring polymers 57
4.7.3 Application of the Rouse theory for the *para*-hydrogen liquid 58

5 An adaptive classical-path integral scheme in Molecular Simulation 61
5.1 Quantum description in soft matter 61
5.2 The basic idea of the quantum-classical adaptivity 62
5.3 PIMD/CG approach in AdResS 64
5.3.1 Effective potential derived from a path integral representation 64
5.4 Results and Discussion 65
5.4.1 Case I: The monoatomic liquid 66
5.4.2 Case II: The molecular liquid 73

6 The *para*-hydrogen liquid in AdResS 79
6.1 Basic physics of hydrogen 80
6.2 The path integral description of *para*-hydrogen 80
6.3 *Para*-hydrogen in AdResS 81
6.4 Technical details 82
6.4.1 Effective coarse grained potential 84
6.5 Results and Discussion 84
6.5.1 For T=25 K 84
6.5.2 For T=14 K 88

7 Conclusions 91

Appendix A: The potential for the polymer chain. 93

Appendix B: CG variable as function of the atomistic variables. 95

Appendix C: Thermostat. 99

Appendix D: Forcefield of the tetrahedral molecule. 101

Appendix E: Implementation of the normal modes in AdResS. 103

Bibliography 106

Introduction

During the last few decades computer simulations have opened up the possibility to increase our understanding of nature in several fields of the science. Such versatility is due to the presence of a growing community of scientists and the large amount of computational facilities. Computer simulation has become extremely useful for scientists because it allows the study of complex systems. It can provide information that is sometime inaccessible by experiments due to the extreme conditions, impossible to reproduce in the laboratory. Furthermore, computer simulations are used to test various theoretical approaches.

However, there still remains a class of problems where brute force simulations are prohibitive due to the complexity of the system. Typically, the complexity involves a large amount of degrees of freedom and the interplay between different length and time scales. Many of these problems occur in the field of soft matter. A common strategy to overcome these problems is to create simplified models on each scale and then pass the information to next level of complexity in a hierarchical way. Naturally, this give rise to the idea of coarse graining procedures to obtain relevant information from each level of description and thus several methods have been developed for this purpose. So far, the validity of such procedures is based on the reproducibility of the properties of interest and a true control of the approximation can not be done in general a priori. In this context is necessary to develop numerical tools to control the underlying coarse grained procedures.

In the last few years several approaches have reached a level of technical sophistication to study multiscale problems and among them the Adaptive Resolution Scheme (AdResS) has maintained also a conceptual development as one can see in this thesis and references therein. The concept of adaptability helps to couple several levels of description on the fly, allowing the flux of information from one description to the other. However, the quantum-classical coupling is considered to be a hard conceptual problem and the adaptability adds a major difficulty. The present work extends the adaptability for the quantum description in the framework of the AdResS scheme.

This thesis consists of seven chapters:

- Chapter 1 introduces the basic concept of coarse graining in soft matter systems and discusses some numerical techniques, that are used later to obtain effective coarse grained potentials. We also discuss the possible limitations of this approach.

- Chapter 2 is dedicated to our method called the Approximation of Separation of variables (ASV). This method was developed to quantify the error introduced by techniques used to calculate the effective coarse grained (CG) potential under the approximation of separation of CG variables.

- Chapter 3 is entirely dedicated to the classical Adaptive Resolution Scheme (AdResS). We describe the concept of the AdResS equation of motion and the thermodynamic equilibrium is defined in systems which change their number of degrees of freedom on the fly. Then we use fractional calculus to generalize the concept of the equipartition theorem and formally define temperature in such conditions. For practical purposes, we comment on the results of AdResS for the tetrahedral molecules where each molecule is mapped onto one coarse-grained site. Finally, we dedicate a section to comment on further theoretical developments.

- Chapter 4 is about the path integral approach, which is considered to be an alternative formalism of quantum mechanics. This chapter starts by showing how to apply the path integral approach for a free particle and then generalizes the results for a quantum many-body system. Then the molecular dynamics implementation of path integrals, known as PIMD and the calculation of average of observables within the PIMD formalism is described. Last section discusses the limitations of such implementation and possible solutions.

- Chapter 5 is fully dedicated to our contribution which extends the concept of AdResS for certain problems where the quantum character of particles (e.g. delocalization) plays a central role and the adaptability of classical and quantum particles takes place. For the quantum description we use the path integral approach (see chapter 4). We describe how to obtain an effective coarse grained potential from the path integral representation. Finally, we applied this concept to two model systems, the monoatomic liquid and molecular liquid, for several degrees of "quantumness".

- Chapter 6 consists of the first real application reported with our adaptive/path integral method to the parahydrogen case.

- Finally, the conclusion and perspectives are presented in Chapter 7.

Chapter 1

Systematic Coarse Graining

Statistical thermodynamics describes the macroscopic state of N particles (e.g. $N_A = 6.0221 \times 10^{23}\text{mol}^{-1}$), in terms of a small set of variables, the so-called "thermodynamic variables" (e.g. pressure, entropy, etc), which depend on the microscopic states (e.g. positions and velocities of particles). One could interpret these macroscopic variables as an effective or a coarser description of the microscopic states. At equilibrium, statistical thermodynamics provides universally accepted recipes for such coarse graining. Thermodynamic potentials contain all the relevant information about a thermodynamic system in a compact format, and these potentials can be calculated via partition functions in terms of statistical weighting.

Nowadays, due to the rapid enhancement of computational resources [1], one may think that the computer simulation of soft condensed matter represents merely a technical task of running larger systems and longer times with standard simulation schemes and that there is no a conceptual challenge. Atomistic simulations based on molecular dynamics (MD) or monte carlo (MC) have shown to be successful in order to explore the potential energy surface (PES) [2] of systems at the molecular scale (i.e. few ps and nm), but there are still serious limitations for a detailed molecular description of mesoscopic scales. The wide time and length scales of such systems combined with the interplay between different scales makes the theoretical description harder and unfeasible to make meaningful comparison with experiment.One can have an idea of this issue, for instance, in the field of bio-molecular simulation, which requires one to access several scales with a large variety of complexity. The topic of protein-generated (or mediated) membrane curvature which is known as membrane "remodeling" for example: protein modules (e.g. Bin/amphiphysin/Rvs domain) can remodel liposomes having initial diameters of 20 nm into thin tubulated structures with diameters on the order of 20 to 50 nm over time-scales longer than microseconds [3]. Another common example is the folding of proteins of more than fifty residues using all-atom force fields. A computer simulation covering a time scale of 5×10^{-8}s for a heptapeptide in methanol succeeds in capturing the folding process [4, 5]. For larger systems direct atomistic simulation has been less successful [6,7]. Traditionally, computer simulations are restricted to short length scales or processes which relax in short time scales.

In order to bridge the atomistic and the mesoscopic scales, one has to develop novel approaches that can access longer times and larger length scales. One such approach is to coarse-grain a system, so that a group of atoms is clustered into a structureless CG bead (i.e. a superatom), which retains the essential physics of interest (see Figure 1.1). These CG beads interact through more efficient potentials where the "fast" variables responsible for the time limitation have been integrated out during its derivation. One expects to get softer potentials which allow the use of larger time steps. To sum up, the system is replaced by fewer particles (CG beads) which mimic the phenomena of interest as accurately as possible and it is able to reach longer time scales.

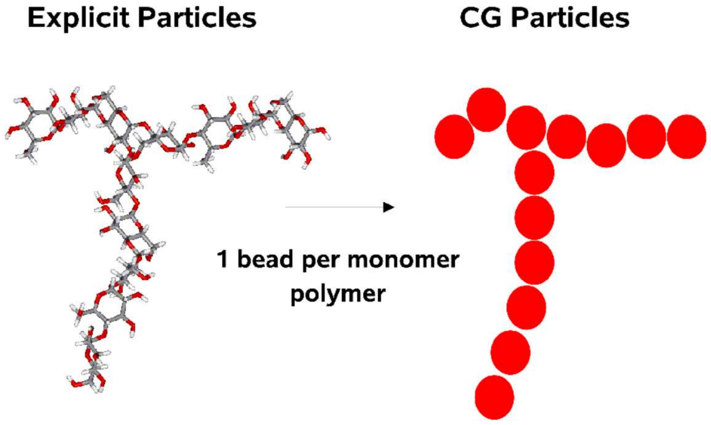

Figure 1.1: A pictorial coarse grained representation of the a-glucose molecule where one monomer has been mapped onto one bead [8].

The next section is about generalities of the effective coarse-grained potential.

1.1 Effective potential for coarse graining

As mentioned before, it is essential to preserve equilibrium thermodynamic properties of the original system in coarse-graining methodologies. This guarantees the predictability of the coarse grained models in a certain range of interest. Consider an explicit system of N particles with coordinates $\{\mathbf{r}\} = r_1 \ldots r_N$. Let us assume for simplicity that one is able to distinguish the "important" coordinates $\mathbf{r}'$ from the full set of variables. In general the partition function, Z_N, for a system of N distinguishable particles is given by

$$Z_N = \int d\mathbf{r}d\mathbf{p}\; e^{-\mathcal{H}/k_B T}, \tag{1.1}$$

where $\mathcal{H}$ is the total Hamiltonian (i.e kinetic and potential energy) and k_B is Boltzmann‘s constant. From this expression, one can obtain the Helmholtz free energy A for a given volume and temperature T as follows,

$$A = -k_B T \ln Z_N. \tag{1.2}$$

From the Eq. 1.1 one notices that the partition function contains two contributions

$$Z_N = Z_{\text{trans}} \times Z_{\text{conf}} \tag{1.3}$$

where Z_{trans} and Z_{conf} is the corresponding translational and configurational contributions. While Z_{tras} depends on the temperature and masses of the particles, its configurational part contains the conservative potential of the complete system, $V(r)$, as follows,

$$Z_{\text{conf}} = \int d\mathbf{r} e^{-V(\mathbf{r})/k_B T}. \tag{1.4}$$

In coarse graining one wishes to obtain an effective potential $V^{\text{eff}}(\mathbf{r'})$ such that the Boltzmann distribution in the reduced space of important variables remains the same as the equilibrium distribution of such varibles in the atomistic model. As a consequence the effective potential can be defined as

$$e^{-U^{\text{eff}}(\mathbf{r}')/k_B T} = \frac{Z'_{\text{conf}}}{Z_{\text{conf}}} \int dr\; e^{-V(\mathbf{r})/k_B T} \delta(\mathbf{r} - \mathbf{r}'). \tag{1.5}$$

From Eq. 1.5 it is clear that the effective coarse grained potential is not a conventional potential, but rather a many-body configuration free energy, the so called potential of mean force (PMF). Such an approach is considered as a bottom-up reconstruction of the effective potential because of the use of atomistic simulations (fine-grained) to derive such a potential in the CG scale. In practice it is very inefficient when one has to deal with multidimensional potentials (e.g. large molecules or membranes solvated in water). It is, however, convenient for low dimensional problems (e.g. a one dimensional reaction coordinate that describes the essential event).

In almost all cases the bottom-up reconstruction of the effective potential will not get the precision required to accurately predict thermodynamic quantities at the CG level. Alternatively one can take a top-down approach, which basically parameterizes the CG models from a macroscopic experimental data. For instance, according to important thermodynamic information [9] or important macroscopic structure of the reference system [10] in order to retain as much as possible the essential physics of interest. In the following section we will briefly introduce the structural coarse graining which has been used in this thesis.

1.2 Structure-based coarse-graining

The basic idea of structure-based coarse-graining is to guarantee a consistency between the structure of high resolution models (atomistic scale) and the low resolution ones (coarse grained scale). Ideally, the structure agreement should hold down to the smallest length scale, which is in the order of the CG unit. One has to be aware that in general all-atom configurations correspond to a single CG conformation [11]. Although there is not a one-to-one correspondence between the CG and full-atom configurations, it is very important that the ensemble of conformations of the CG model corresponds to the atomistic ones.In practice, the structural coarse-graining is done through the mapping of certain distribution functions between the mapping points (center of masses of CG beads) in the reference system (full atomistic). Although there is no unique set of mapping points, their choice relies on numerical convergence of the structural property [12] to be calculated.

At this point, one should comment of the methodologies used to generate the CG forced field. There are several ways to do this. A popular one is to fit a pair potential, so that it reproduces the structural quantity desired, such as the radial distribution function, $g(r)$. Ideally, the form of the potential should be independent of the procedure used to calculate it, as proved by Henderson [13], where two “pair potentials” that reproduce the same radial distribution function are equal up to a constant. A more sophisticated mathematical proof can be found in reference [14]. Conventionally, one derives this effective potential at a given state-point. Thus, one should expect “transferability problems” when that potential is used at different state points.

Technically, there is a set of very efficient methods used to obtain such effective potentials. It is important to note that, in these methods small numerical errors can lead to different effective potentials, during the fitting of the macroscopic property such as radial distribution function. Among the most popular ones are the iterative procedures, such as the Iterative Boltzmann Inversion (IBI) or Reverse Monte Carlo (RMC). They try to reproduce structural information (e.g. radial distribution functions) which can be taken from experiments or all-atom simulations for a given thermodynamic state. In 2009 a software named “VOTCA” [15] designed specifically for structural coarse graining was released and the current version contains several well-known traditional methods for constructing the effective potential. In this thesis work we used the Iterative Boltzmann Inversion for calculating the effective nonbonded interactions, and the next section is dedicated to it.

1.3 Nonbonded interaction potentials

The main aim of deriving an effective nonbonded potential is to reproduce structural properties. These are mainly contained in the radial distribution functions of some soft matter systems (e.g. liquids or polymer melts). This information is commonly obtained from experiment or atomistic simulations. The basic idea is to obtain numerical ("tabulated") potentials, which act between the coarse grained units. In a similar way the same coarse grained procedure can be made for bonded interactions (see ref. [16]).

1.3.1 The Iterative Boltzmann Inversion

The implementation for non-bonded interactions starts with an initial guess for the nonbonded potential, usually the Boltzmann inverse of the target $g_{\text{target}}(r)$ is chosen as a first guess,

$$V^{CG}_{NB,0}(r) = -k_B T \log g_{\text{target}}(r) \tag{1.6}$$

where T is the temperature and k_B is Boltzmann's constant. Last expression is known as the potential of mean force. One then runs a CG simulation and obtains a new $g(r)$ which usually does not match the target structure due to the multibody interaction. This is because the potential of mean force (Eq. 1.6) is a good estimate for the pair interaction of highly dilute systems. To achieve the desired convergence an additional correction has to be introduced through the following iterative scheme

$$V^{CG}_{NB,i+1} = V^{CG}_{NB,i} + k_B T \log \frac{g_i(r)}{g_{target}(r)} \tag{1.7}$$

Basically, the initial guess for the potential can be iteratively refined until the desired structure is obtained. For small molecules and simple liquids such as water and benzene at normal conditions, this process is powerful and straightforward to implement. On the other hand, for multicomponent system (several types of CG beads) e.g. liquid crystals or polymer melts, the process of determining the nonbonded potential is more complicated [17]. The iterative procedure guarantees a nice agreement with the target distribution but not with the pressure. Next section explains how to deal with this problem.

1.3.2 Pressure correction

During the structure-based coarse-graining, it is also important to fit the pressure at the density of the target system in order to retain as much as possible the state point of the all atom model. It is well-known that one can fit either the pressure or the compressibility, but not both simultaneously [18]. Typically, a linear term is added to the nonbonded potential interatively in order to fit the pressure:

$$\Delta V(r) = A(1 - \frac{r}{r_{cut}}) \tag{1.8}$$

for $r < r_{cut}$, where r_{cut} is the cutoff radius of the pair nonbonded potential and A is an arbitrary constant which can be estimated from the virial expansion [16]. Basically, the correction of Eq. 1.8 yields a constant force that makes the interaction repulsive if A is positive, and more attractive in the opposite case.

One runs into trouble if A is not small enough, it can affect the overall structure and, thus it will have to be readjusted until a good balance between pressure and the radial distribution can be obtained. Such a process could delay the convergence of the target radial distribution function of interest.

1.4 Limitation of structure-based coarse-graining

It is worthwhile to emphasize that **identifying the proper set of relevant variables** (CG ones) is the key to success in any systematic coarse-graining procedure. In other words, one has to first identify a suitable set of relevant variables for a coarser target level and then express them in terms of the variables of the finer resolution. The latter is taken as the reference in the simulation. In many cases, this is given by the atomistic level with particle positions and momenta as a variables.

Note that the assumption of a pairwise potential is suitable in many cases where the three-body or higher order forces are not so relevant for the level of description. However, one can not expect the same representation of "all properties of the system" between the coarser description and the atomistic one. The **representability problems** are widely spread in coarse-grained descriptions of soft matter [9, 18].

Another essential problem is the degree of **transferability** of the CG model from one thermodynamic condition to another. In principle, as stated before, CG potentials cannot be fully transferable due to the reduction of degrees of freedom (some information has been averaged out), which simplifies the complexity of the system.

Finally, the **dynamics** of coarse grained models in many cases does not correspond to the real dynamics. Generally, CG dynamics is faster than the underlying atomistic one. Thus the CG dynamics must be properly interpreted; in any case the fact that one can run much longer simulations implies a much more efficient statistical sampling for the calculation of static equilibrium properties.

Chapter 2

Approximation of Separation of Variables

A relevant problem in molecular-dynamics (MD) simulations is the determination of the minimal set of degrees of freedom (DOFs) to be employed in the simulation study by a systematic procedure. Particularly, in the field of coarse grained simulations, one would like to determine the set of relevant variables which are sufficient to properly characterize the phenomena under investigation. Also in complementary fields such as the study of rare events (e.g protein folding or crystal nucleation) the choice of the “collective variables (CVs)”, typically used in transition path sampling [19] or metadynamics [20] is characterized by the same problem. The choice of a small set of variables is generally guided by chemical or physical intuition and does not always allow for a systematic control of the underlying approximations. In this context it would be optimal to provide some criteria to control, in a systematic way, the choice of the set of DOFs (eg. reaction coordinate or CVs).

In this chapter we present the basis of a criterion to evaluate how separable two DOFs are and its extension to study the interdependency of several DOFs. Our algorithm [21] mainly requires as input the basic information contained in the potential energy surface (PES) [2]. The basic idea behind the separability of some DOFs in the PES typically leads us to propose a reduction in the dimensionality of the problem. In complex systems, the PES contains crucial information; in an extreme case, if two DOFs are independent their evolution occurs in two orthogonal spaces. This means that one could neglect one of these variables without altering the dynamic evolution of the system in the space of interest.

The present chapter begins by exposing the basic idea of how to evaluate the separability of two DOFs through the method developed and called the approximation of separation of variables (ASV). A methodological example of how to apply the ASV will illustrate the idea (the case of the diatomic molecules with a flat surface). And finally we apply the ASV method to a more realistic system, namely the RIS model of polymer chain.

2.1 The ASV: a first criterion

For practical purposes let us start considering a two dimensional potential of the form $V = V(x, y)$: the extension to more variables will be discussed later on. The analytical form of V is given and only in the case that the two variables are fully separable one has that:

$$V(x, y) = V_1(x) + V_2(y) \tag{2.1}$$

Or equivalently one could write V for any couple of fixed points x_0, y_0 as:

$$V(x, y) = V_1(x) + V_2(y) = V(x, y_0) + V(x_0, y) - V(x_0, y_0). \tag{2.2}$$

In other fields of science the fixed point has many meanings, for example, in mathematics is defined as a point that is mapped to itself by the function, in physics is commomly used in the renormalization group theory languagee, in chemistry is used as the point under which rotation of the molecule occurs, to name a few. In this thesis the fixed points are defined as the set of points that decouple the potential for each DOF (see Eq. 2.2).

Up to now this is exact and represents the ideal case of complete separability of two DOFs. A reasonable criterion to indicate how good the approximation of separation of variables is, is to calculate its deviation from the ideal case of Eq. 2.2. This can be done with the definition of an estimator Δ of the difference between the true potential where the variables are still coupled and the potential where one introduces by hand the separability in the fashion of Eq. 2.2. For instance, given a potential $V(x, y)$ where x and y are not decoupled one defines Δ as,

$$\Delta_{x_0,y_0}(x, y) = V(x, y) - [V(x, y_0) + V(x_0, y) - V(x_0, y_0)], \tag{2.3}$$

$\forall x, y \neq x_0, y_0$. Basically this is a point-by-point evaluation in order to compare the potential V and its respective version where the ASV was introduced.

To know how meaningful the energetic discrepancy between the coupled potential and the ASV for a specific problem is, one needs to define a scale of energy. Since we deal with energy scales of the order of thermal fluctuations ($\sim k_B T$ in classical simulation of soft matter systems). One could define the quality factor for the ASV as follows,

$$Q(x, y) = \frac{|\Delta_{x_0,y_0}|}{k_B T} \tag{2.4}$$

where k_B is Boltzmann's constant and T is the temperature. The simple form of Eq. 2.4 defines the "first criterion" of our algorithm; this form is easy to treat numerically. From Eq. 2.4 one sees that if Q is much larger than one, then the assumption of the approximation of separation of variables will lead to false dynamics and thus its exploration by MD can be questionable. The choice of the reference is system dependent, for instance, if one is studying the conformational space of a molecule which is characterized by an energy barrier ϵ_b that separates two important and well-defined states, in this case Δ can be compared with ϵ_b.

So far, we have not commented about the dependence of Q with respect to the fixed points (x_0, y_0). The treatment of this problem will strengthen the criterion designed previously and this is the focus of the next section.

2.2 Dependence of Q on the fixed point: A complementary criterion

A formal way to determine the dependence of Q on the fixed points (x_0, y_0) is to monitor the variation of Q upon the variation of x_0 or y_0. We define:

$$\delta_{x_0} = \frac{\partial Q(x_0, y_0, x, y)}{\partial x_0} \tag{2.5}$$

$$\delta_{y_0} = \frac{\partial Q(x_0, y_0, x, y)}{\partial y_0}. \tag{2.6}$$

These variations are calculated over a certain range of fixed points (x_0, y_0) and on a certain (x, y) domain. In general, if the ASV is reasonable, the dependence of Q, given by Eq. 2.5 and 2.6, on the fixed point is indeed negligible by construction. In such a case, one would have for Eq. 2.5 that,

$$\delta_{x_0} = \frac{1}{k_B T}[\frac{\partial V(x_0, y_0)}{\partial x_0} - \frac{\partial V(x_0, y)}{\partial x_0}] \ll 1, \tag{2.7}$$

if x and y are not highly correlated, and one immediately sees that $\partial V(x_0, y_0)/\partial x_0 \approx \partial V(x_0)/\partial x_0$ and similarly for $\partial V(x_0, y)/\partial x_0 \approx \partial V(x_0)/\partial x_0$. This demonstrates that $\delta_{x_0} \approx 0$ for the case of weak coupling between x and y, the same holds for δ_{y_0}. When the ASV is questionable, the quality factor Q will be dependent on the values of the fixed points. This seems to be a negative aspect of the algorithm proposed; but eventually here we propose to use it as a "complementary criterion" to identify the regions of (x, y) where the validity of the ASV is very critical. The recipe is the following: first one studies δ_{x_0} and δ_{y_0} as a function of x_0 and y_0 using x and y as parameters to vary, which helps us to identify the critical regions. Second, one chooses x_0 and y_0 outside the critical region. Finally, one applies the ASV over all relevant xy space to quantify the degree of separability (Δ). One could summarize these ideas as follows,

- It defines regions where the choice of the fixed points for Q is delicate and those where it is not.

- It also defines a region of the (x, y) space where the ASV is likely not to hold, compared with other (x, y) regions.

This part of the procedure is rather important. Instead of only minimizing the dependence of fixed points on Q, it also takes care of the fact that critical region may be too small and a fixed point taken from such a region may be too close to some of the (x, y) points, so that when one evaluates Q one gets small values because of the close values of $V(x, y)$, $V(x, y_0)$, $V(x_0, y)$ and $V(x_0, y_0)$. Basically, the analysis is prevented from being very local.

In summary, the simple recipe for the ASV criteria can be given as

1. Choose an arbitrary fixed point.
2. Calculate the quality factor Q.
3. Study the dependence on the fixed points by calculating the corresponding δ.
4. Identify the critical region.
5. Choose a fixed point outside the critical region (the optimal would be where $\delta = 0$).
6. Calculate Q once again using the fixed points of step 5 and analyze Q in the critical region.

In the next sections, we show first a simple example of how to apply the ASV. The second example is more complex and will help us to generalize our ideas for a multidimensional system.

2.3 A guiding example: A rigid diatomic molecule interacting with a surface

This is a simple example of how to apply the ASV criteria. The system consists of a diatomic molecule of equivalent atoms which interact with a uniform rigid surface via a potential given by

$$U(z_a, z_b) = \epsilon\{\frac{2}{5}[(\frac{\sigma}{z_a})^{10} + (\frac{\sigma}{z_b})^{10}] - [(\frac{\sigma}{z_a})^4 + (\frac{\sigma}{z_b})^4]\} \tag{2.8}$$

The question we want to address is whether there is a region z_a, z_b where the molecule can be treated as an effective "point-like" particle (see Figure 2.1) and whose interaction point is located at the center of mass.

An equivalent way to describe the same system is to transform the set of variables z_a, z_b to another set of degrees of freedom. One variable is the distance r from the surface to the center of mass, while the rotation around it by the variable θ. Now, the previous question becomes: is it possible to separate r and θ?

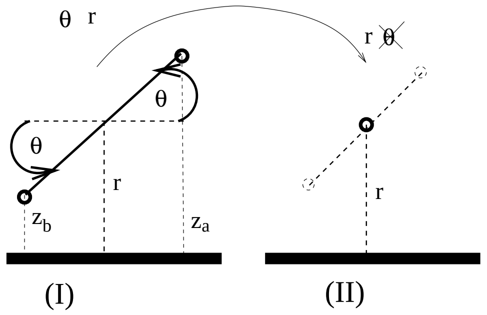

Figure 2.1: Schematic representation of the system considered. Part (I) shows pictorially the mapping from (z_a, z_b) to (r, θ) and part (II) describes the case when θ is completely independent from r.

For this purpose, one has to transform the old set of variables to the new ones:

$$z_a = r + d \cdot sin(\theta); z_b = r - d \cdot sin(\theta) \tag{2.9}$$

and thus, substituting the last expression into Eq. 2.8 one writes the potential in terms of the new variables r, θ as

$$\begin{aligned} U(r,\theta) &= \epsilon\left(\frac{2}{5}\left[\left(\frac{\sigma}{r + d\cdot sin(\theta)}\right)^{10} + \left(\frac{\sigma}{r - d\cdot sin(\theta)}\right)^{10}\right]\right) - \\ &- \epsilon\left(\left[\left(\frac{\sigma}{r + d\cdot sin(\theta)}\right)^{4} + \left(\frac{\sigma}{r - d\cdot sin(\theta)}\right)^{4}\right]\right) \end{aligned} \tag{2.10}$$

where $\theta \in [0, \pi/2]$ due to the symmetry of the system and d is the fixed distance from the center of mass to each single atom. The values used here for various parameters were taken from an atomistic model which was employed to study the absorption of a molecule on a surface [22]. Firstly, by using $Q(r, \theta)$ one could determine the minimum distance r from the surface for which the separation is still reasonable, and thus for distances larger than this one is able to neglect the molecular rotation and represent the molecule (with respect to the surface) as one effective interaction site located at the center of mass. Now we show the factor of quality for this problem and its study with respect to the fixed points.

2.3.1 The factor of quality Q

One should start to define the threshold for the potential discrepancy. Let us use for this $2k_BT$ in Eq. 2.4. Energy errors that exceed this number may alter the dynamics of the process of absorption. With the help of Eq. 2.4 one computes the quality factor of the problem. The result of this calculation is shown in

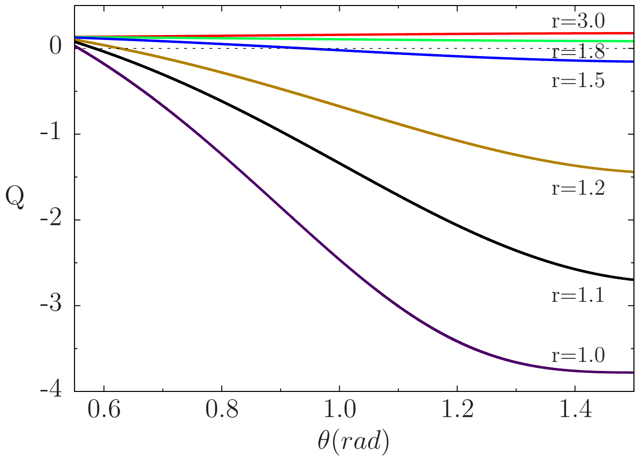

Figure 2.2: The quality factor Q studied as a function of the angle θ varying parametrically r from 3.0 nm to 1.0 nm. The fixed point used in this example is: $\theta_0 = \frac{\pi}{6}$ and $r_0 = 2.0$ nm. Here $\sigma = 0.5$ nm; $l = 0.5$ nm, $\epsilon = 10k_BT$. As the distance of the center of mass from the surface decreases the dependence on θ becomes stronger and for approximately $r \leq 1.2$ nm the ASV starts to be questionable since the error induced can be larger than 2 k_BT.

Figure 2.2. As long as r takes values in the range of $1.0 - 1.2$ nm the potential where the separation is applied overestimates the true potential of the relevant quantity by around $2 - 3k_BT$, as expected, and this overestimation increases as θ increases.

From this part one sees the following for the separability: as the molecule is close, i.e. $r \leq 1.2$ nm, the assumption of separation of r and θ not longer holds. For this example we use the values of the fixed points which at least affect the quality factor Q in the range of interest. This will be explained in detail in the next section.

2.3.2 The complementary criterion

Methodologically, as we pointed out before, one can start to look at the small variation of Q as a function of the fixed points. For this specific example, we show the variation of Q with respect to r_0 and θ_0. The results are shown in Figure 2.3 and 2.4 respectively. For δ_{r_0} the critical region corresponds to the region where $r < 1.4$ nm (by construction r cannot be less than 0.5 nm). For δ_{θ_0} the critical region varies from $\pi/4 \leq \theta \leq \pi/2$. In fact, Figure 2.3 shows, for different values of θ and θ_0, a trend according to which for $r_0 \leq 1.4$ nm the dependence of Q on r_0 becomes relevant. Similarly, one sees in Figure 2.4 that the critical region is located in $\pi/4 \leq \theta \leq \pi/2$.

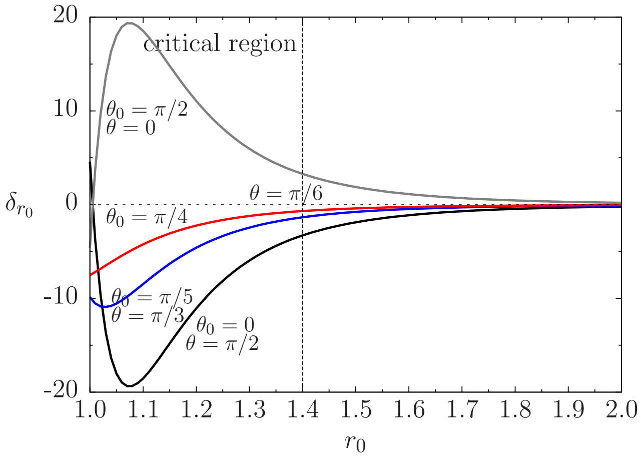

Figure 2.3: The variation of Q w.r.t. r_0, (δ_{r_0}), for some example values of θ_0 and θ. We plot also the extreme values taken by θ_0 and θ (symmetric curves) allowed by the atomistic model and, in between, two more examples. The message of this plot is that there is a general trend according to which the dependence of Q on r_0 becomes crucial for, approximatively, $r_0 \leq 1.4$ nm. The vertical dashed line indicates that the region $r_0 \leq 1.4$ nm should be considered as the critical region for the variable r.

From these studies one can optimize the choice of fixed points, r_0 and θ_0, for Q. Thus, if one performs the analysis of Q in the region of $r \leq 1.4$ nm and $\pi/4 \leq \theta \leq \pi/2$, then the optimal choice of the fixed points would be $r_0 = 1.2$ nm and $\theta_0 = \pi/6$ as given in Figure 2.2. From this it emerges easily that for $r \leq 1.2$ nm the ASV starts to be questionable. Basically, these studies of δr_0 and $\delta\theta_0$ show how to refine the analysis of Q for the region of large discrepancy. The studies of Q and $\delta r_0, \delta\theta_0$ complement each other and shows how one can determine the validity of the ASV.

2.4 Extension to higher dimensions: Parametric study

The criteria presented before can be in principle extended to more dimensions although the computational demands would inevitably increase. The most simple way to proceed is by looking at two variables at a time with the rest frozen. Let us suppose the potential in a generic form $U(x, y, z, \ldots)$. Then, one may focus on two variables per time and see how the separation can be carry out. For simplicity, we consider $U = U(x, y, z)$ as function of three variables x, y and

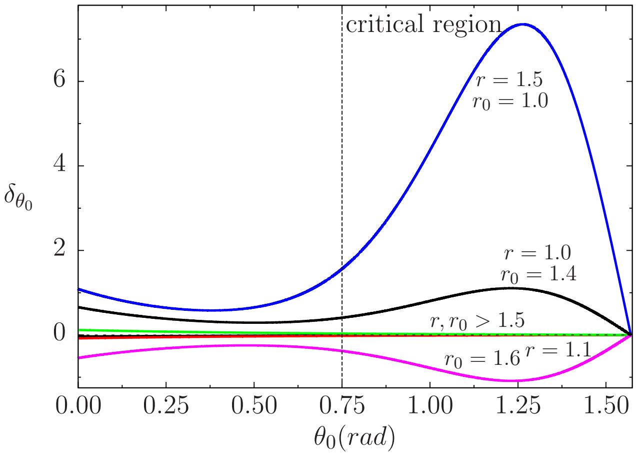

Figure 2.4: The variation of Q w.r.t. θ_0, (δ_{θ_0}), for some examples values of r_0 and r. For values of r_0 and r outside the critical region defined by Figure 2.3, the dependence of Q from θ_0 is negligible, however for values within the critical region of $r, r_0 \leq 1.4$ nm the dependence becomes stronger and identifies the critical region of the variable θ. The vertical dashed line indicates that the region $\theta_0 \geq \frac{\pi}{4}$ is the critical one.

z. In this case Δ is written as

$$\Delta_{x_0,y_0}(x, y, \bar{z}) = U(x, y, \bar{z}) - [U(x, y_0, \bar{z}) + U(x_0, y, \bar{z}) - U(x_0, y_0, \bar{z})] \quad (2.11)$$

if one is interested in how separable x and y are, under the assumption that z can be factorized from the others. In Eq. 2.11, $\bar{z}$ means all the possible values of z that could decouple it from x and y. In the same way one could analyze $\Delta_{y_0,z_0}(\bar{x}, y, z)$ if one were interested in how separable are y and z, under the implicit hypothesis that x can be separated from the others or similarly for $\Delta_{x_0,z_0}(x, \bar{y}, z)$, if the interest is in the separation of x and z.

2.5 A second example: 2D version of the RIS model of a polymer chain

The Rotational Isomeric State (RIS) theory is considerd to be the standard method [23] to study the conformational properties of macromolecules. Here we define the RIS model of our polymer chain, where all torsions are set to zero (i.e. 2D case). such a condition will be fulfilled for all the dihedral angles around the bonds. In our model system we apply the ASV criteria for two different mapping schemes (MSs) (see section 2.5.2). Let us explain in the next section the problem of choosing the mapping scheme for a polymer system.

2.5.1 A general problem in coarse grained modeling of bonded interactions

In the previous chapter we comment about the mapping scheme (MS) as the common CG strategy to reduce the large number DOFs in soft matter simulations. After one defines a MS (which also defines implicitly the CG variables of interest) one performs simulations to obtain the effective potential (bonded and nonbonded) among the set of CG variables. Here, let us focus on the case of bonded potential for a polymer. For example one replaces each chemical group in a polymer by spheres that become the new effective particles of such system. The new CG variables are the following: bond lengths (r), bond angles (θ) and torsion (ϕ) associated with two, three or four chemical groups respectively. Next, one has to think about the effective potential for this minimal set of variables. Some methods are based on all-atom simulations of a single chain in vacuum, the corresponding distribution of the CG variables is obtained in these MD runs and under **the approximation that they are decoupled**. Then one can write,

$$P\left(r,\theta,\phi,T\right)=P\left(r,T\right)P\left(\theta,T\right)P\left(\phi,T\right) \tag{2.12}$$

and by Boltzmann inversion [16] at the given temperature T one has:

$$U\left(r,\theta,\phi\right)=U\left(r\right)+U\left(\theta\right)+U\left(\phi\right) \tag{2.13}$$

This new potential mainly reproduces some features of the full atomistic model but at much lower computational demands (less variables to integrate in MD). However, *a priori* one does not know how far the hypothesis of separability could hold for a given MS. This means, so far, that there is not a systematic approach to control the underlying ASV implied by the potential derived in Eq. 2.13. Here we can apply straightforwardly the criteria developed previously.

2.5.2 Mapping schemes

In order to make a comparison between our model system and a real polymer chain we use the parameters that characterize the energy and length scales corresponding to the n-alkane chain (see Appendix A). The structure of an alkane chain (AC) is illustrated in Figure 2.5. This simple polymer is composed of several repeat units with each one corresponding to either methyl groups (CH_3) in both ends or ethyl groups (CH_2) along the backbone. We study two different MSs for our model system which are depicted in Figure 2.6. They are indicated as the 1:2 and 1:3 mapping schemes (MSs). By convention the right number denotes the number of carbons which will be replace by a sphere (left number), this means that the 1:3 MS is more coarser than 1:2 MS. Now we define the variables as follows:

- the distance between close beads $R^{(1)}$ and $R^{(2)}$
- the angle Ω between the vectors $\vec{R}^{(1)}$ and $\vec{R}^{(2)}$.

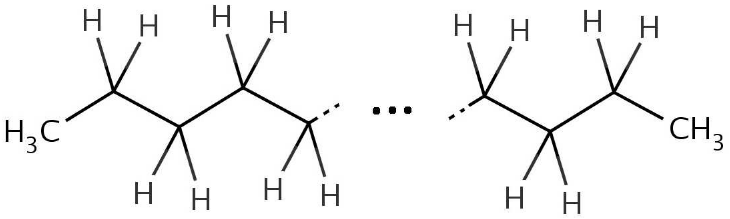

Figure 2.5: Chemical structure of the n-alkane chain in which each vertex of the backbone represents a carbon.

Even though in Figure 2.6 both MSs result in the same set of CG variables, they depend on different atomistic variables. The point we want to address is how separable are these CG variables, that is how independent are $R^{(1)}$ and $R^{(2)}$, $R^{(1)}$ and Ω, and $R^{(2)}$ and Ω.

2.6 Results

There are some basic steps that one has to follow in order to evaluate the quality of each MSs. First, one has to express the atomistic potential as a function of the CG variables $R^{(1)}$, $R^{(2)}$ and Ω (see Appendix B). After that we follow the steps of the ASV discussed at the begin of this chapter. We discuss here some technical details for this system which can be useful for more complex systems.

First, we calculate Q in regions of the CG variables allowed by the atomistic conformations. This means that the regions of Q where bond breaking or overlap of two atoms occur, are not taken in account during the analysis.

The choice of fixed points is not so trivial ($\delta = 0$) as in the first example. This time δ is not likely to be zero given the complexity of the molecule. Thus, we extend the previously used criteria as follows:

a) If δ has a region where it varies slowly and then a region where its variation increases rapidly, we define the first region as non critical and choose the final fixed point from there (possibly the point corresponding to the minimum value of δ, that is, the minimum dependence on the fixed point) or we choose several fixed points and sample Q over all of them.

b) If δ is constant, but it is characterized by a high value or it increases rapidly over the whole domain, then the whole domain is defined as critical, which means that the ASV does not hold. Since we need a fixed point to quantify the error introduced by the ASV. This time one could choose several fixed points all over the whole domain of fixed points and for each choice calculate the Q separately. For the final Q one takes the average (plus the fluctuations) of all the values obtained for each study. What changes from

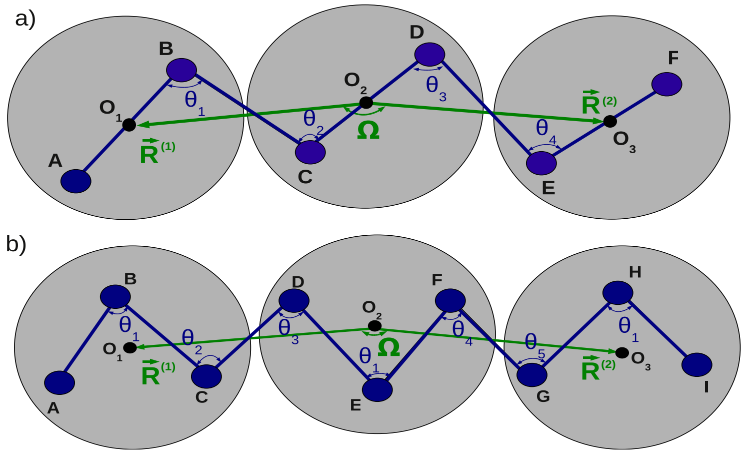

Figure 2.6: The figure shows the typical CG structure of the polymeric chain studied. Each chemical group is represented by blue circles close to a letter. Part (a) shows the MS where there are two carbon groups per bead (1:2), while part (b) shows three per bead (1:3). The “atomistic” variables $\theta_i \, \forall i = 1, 2, 3, 4$ and the CG variables $R^{(1)}$, $R^{(2)}$and Ω are also shown.

> case (a) is that one needs a large sample of fixed points. Typically in this case one can expect that the approximation leads to very large errors.

With these remarks in mind, we proceed to show the results for each mapping scheme.

2.6.1 The 1:2 MS

As we stated before one performs a systematic study of Q for each parametric CG variable. Since we have three CG variables one defines also three different Q for each parametric value. For example Q_1 indicates the case for $\bar{R}^{(1)}$ as a parametric, Q_2 for $\bar{R}^{(2)}$ and Q_3 for $\bar{\Omega}$.

To obtain Q_1, see Figure 2.7(a), we first choose an arbitrary set of fixed points $(R_0^{(2)}, \Omega_0)$ and then analyze Q and its dependence on the fixed points as shown in part (b). Next, similar to the procedure shown for the diatomic molecule, we can determine if there are critical regions, which in this case are given by $\forall R_0^{(2)} \in [2.425, 2.525]$ and $\forall \Omega_0^{(2)} \in [145°, 180°]$. Accordingly, we have chosen the fixed points outside these critical regions to be $R_o^{(2)} = 2.4$Å and $\Omega_0 = 140°$. With this optimized set of fixed points one estimates the error in the potential that one makes under the hypothesis of separability of Ω and $R^{(2)}$ which turns out to be $\sim 9k_BT$; this is larger than the expected thermal fluctuations.

Similarly, in Figure 2.8 and 2.9 we show the results for Q_2 and Q_3 respectively. We note that in the case of Figure 2.8 the analysis of δ does not lead to identify proper critical regions. This time we use the extension (b) of the ASV given in the previous section to quantify the quality factor (Q_2) of the MS. The general message is that in each plot for the 1:2 MS there are extended regions where the error varies between $6k_BT$ and $9k_BT$, which is much larger than the reference energy.

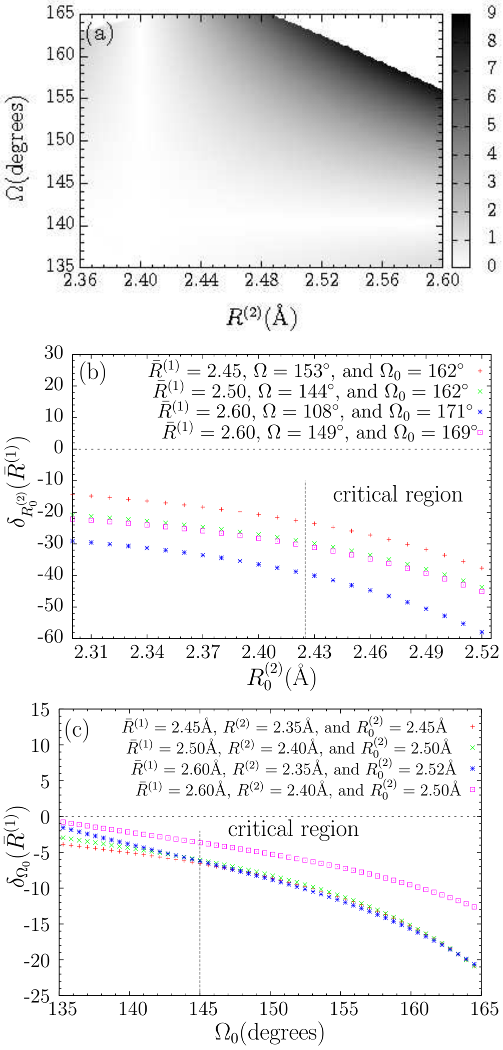

Figure 2.7: The figure (a) shows the quality factor $Q_1 = Q(\bar{R}^{(1)}) = \Delta_{R_0^{(2)},\Omega_0}(\bar{R}^{(1)})/k_BT$ for the mapping 1:2 with $\bar{R}^{(1)} = 2.4$Å as a parametric value and the fixed points equal to $R_o^{(2)} = 2.4$Å and $\Omega_0 = 140°$. In (b) and (c) are shown critical regions $\forall R_o^{(2)} \in [2.425, 2.525]$ and $\forall \Omega_o \in [145°, 180°]$. In (a), the light gray region represents the vicinity of the fixed point $(R_0^{(2)}, \Omega_0)$ where CG potential is decoupled (by definition) and thus of no interest in this context. The real interest is in the critical regions determined by the plot shown in (b) and (c). There the potential is no more decoupled and the variables are highly correlated. Note also that the white regions are those were the CG variables are not defined and thus Q is not calculated. The same applies for all the next figures.

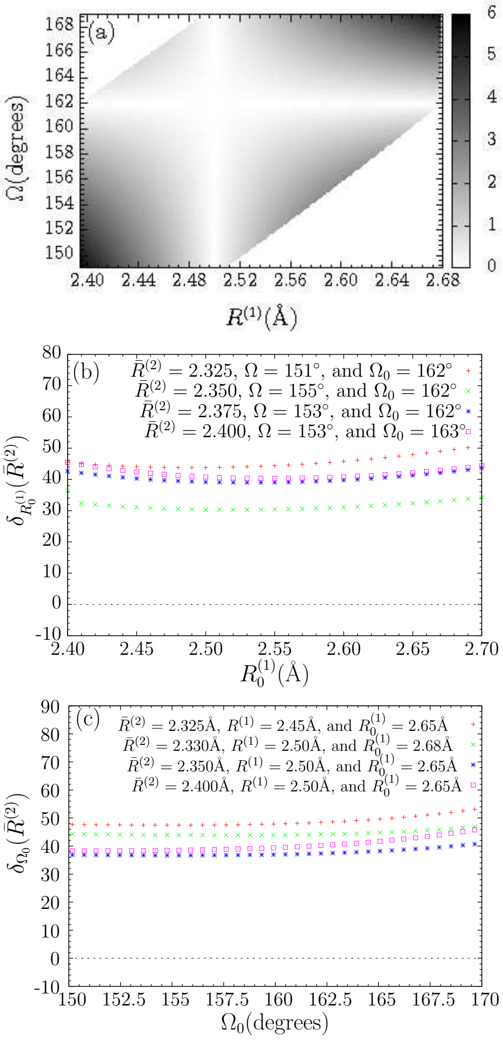

Figure 2.8: The plot (a) shows the quality $Q_2 = Q(\bar{R}^{(2)}) = \Delta_{R_0^{(1)},\Omega_0}(\bar{R}^{(2)})/k_BT$ for the mapping 1:2 with $\bar{R}^{(2)} = 2.35$Å as a parametric value and the fixed points equal to $R_o^{(1)} = 2.5$Å and $\Omega_0 = 162°$. To be noticed that the value of the derivatives shown in (b) and (c) are almost constant for each parametric curve. This means that there exists no real definition of critical region and in this case one can proceed by using several fixed points in the domain and averaging the quality factor resulting from each study. In (a) the light gray colour indicates the region close to the fixed points $(R_o^{(1)}, \Omega_0)$ where the potential is decouple by definition. Totally white regions indicate configuration space of $R^{(1)}$ and Ω not allowed by the atomistic model.

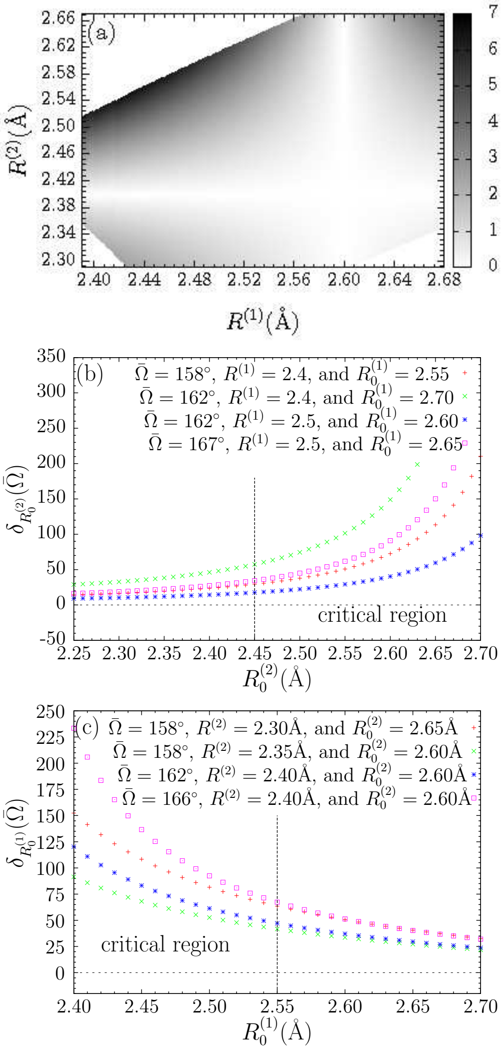

Figure 2.9: The plot (a) shows the quality factor $Q_3 = Q(\bar{\Omega}) = \Delta_{R_0^{(1)},R_0^{(2)}}(\bar{\Omega})/k_BT$ for the mapping 1:2 with $\bar{\Omega} = 162^\circ$ as a parametric value and the fixed points equal to $R_o^{(1)} = 2.6$Å and $R_o^{(2)} = 2.4$Å. The procedure for analyzing (b) and (c) is the same as described by the previous figures.

2.6.2 The 1:3 MS

Similarly as in the previous case we perform a study of Q for each CG variables. The results for each Q and the analysis of δ with respect to the critical points are shown from Figure 2.10 to 2.12. A common aspect for Q_1 and Q_2 is that one cannot find a critical region within the domain of the fixed points, $(R_0^{(2)}, \Omega_0)$

and $(R_0^{(1)}, \Omega_0)$, respectively. Thus, one has to use the extension of our criteria given in section 2.6.

In comparison to the 1:2 MS the maximum value of Q is between $6k_BT$ and $7k_BT$. This means that the ASV is a better approximation for the 1:3 MS than for the 1:2 MS. This is true in polymer theory [24] because the coarser the system becomes the closer it is to a freely jointed chain and thus more separable. Implicitly the 1:3 MS becomes more efficient than the 1:2 MS (see also the analysis of the average values of Q reported in the next section), this means that a CG model using the former MS will better resemble the underlying atomistic model than the latter.

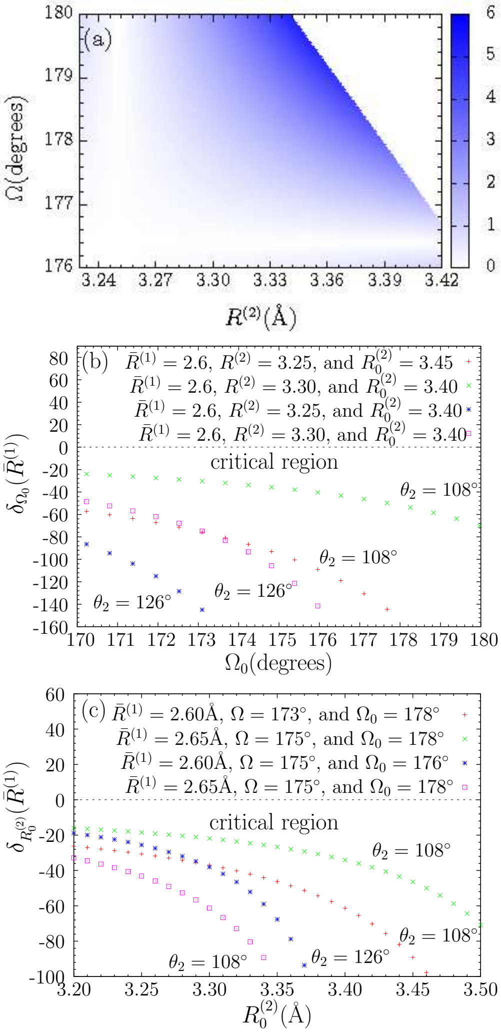

Figure 2.10: Plot (a) shows the quality factor $Q_1 = Q(\bar{R}^{(1)}) = \Delta_{R_0^{(2)},\Omega_0}(\bar{R}^{(1)})/k_BT$ of the mapping 1:3 with $\bar{R}^{(1)} = 2.7$Å and $\theta_2 = 126°$ as parametric values. The corresponding fixed points are $R_o^{(2)} = 3.25$Å and $\Omega_0 = 176.5°$. As for Figure 2.7, but in this case due to the rapidly varying dependence on the fixed point (see (b) and (c)) one has to use several fixed points and consider the Q resulting from the average of each study, in order to have a more valid quantitative information. The light blue colour indicates the region close to the fixed points $(R_0^{(2)}, \Omega_0)$. The totally white colour is not allowed to be sample by the underlying atomistic model (e.g. a chemical bond is broken).

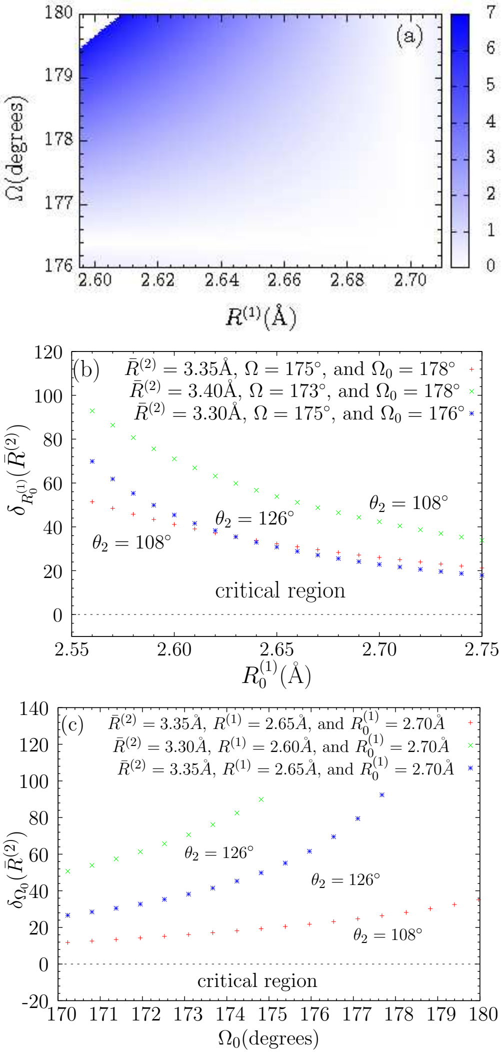

Figure 2.11: Plot (a) shows the quality factor $Q_2 = Q(\bar{R}^{(2)}) = \Delta_{R_0^{(1)},\Omega_0}(\bar{R}^{(2)})/k_B T$ for the mapping 1:3 with $\bar{R}^{(2)} = 3.25$Å and $\theta_2 = 126°$ as parametric values. The corresponding fixed points are $R_o^{(1)} = 2.7$Å and $\Omega_0 = 176.5°$. As the previous figure, does not exist a non critical region, see (b) and (c). This means that one must explore critical points over the whole domain in order to estimate quantitatively Q.

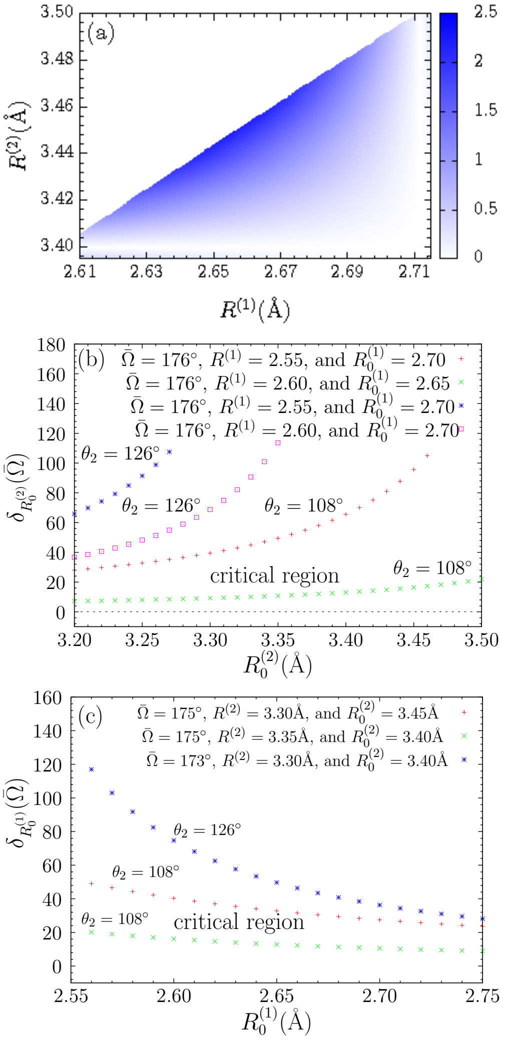

Figure 2.12: Plot (a) shows the quality factor $Q_3 = Q(\bar{\Omega}) = \Delta_{R_0^{(1)},R_0^{(2)}}(\bar{\Omega})/k_BT$ for the mapping 1:3 with $\bar{\Omega} = 174°$ and $\theta_2 = 126°$ as parametric values. The corresponding fixed points are $R_o^{(1)} = 2.71$Å and $R_o^{(2)} = 3.4$Å. As in the other cases for Q_1, Q_2 and now for Q_3 non critical regions within the domains of the CG variables cannot be found (see (b) and (c)) and thus it applies the same considerations of the previous two figures.

2.6.3 Average of Q in fixed point and CG variable parametric space

In this section we explain in detail how to deal numerically with systems where a critical region cannot be defined due to the strong dependency of quality factor

(Q) with respect to the fixed points. Using the extensions of our criteria given in the section 2.6 we can still get a reasonable estimate of the quality factor. We provide as an example the case of the RIS model for a polymer chain. Due to its complexity we cannot expect to always find critical regions. The same may occur in more complex systems. In the case of the 1:2 MS, one notes that for Q_2 (with $\bar{R}^{(2)}$ as a parametric value) in Figure 2.8 that the study of $\delta_{R_0^{(1)}}$ (part (a)) and δ_{Ω_0} (part (b)) do not provide a critical region, this means that $R^{(1)}$ and Ω are not likely to be separable under the assumption that both CG variables are independent from $R^{(2)}$ in the PES. This case represents the negative assessment of the ASV. With this in mind we proceed to quantify how much the total energy deviates, when one assumes separability of these two CG variables. In order to do that one samples a few different fixed points in both regions and for each pair $(R_0^{(1)}, \Omega_0)$ we quantify the corresponding value of Q_2. Next one takes the average over all of them. This procedure optimizes Q_2 making it less dependent on its fixed points.

As we report in Table 2.1 for the 1:2 MS the mean value of Q_2 is characterized by a large value compared to Q_1 and Q_3 in the same table. For example, the value 16.0 ± 4.6 tells us that on average the error introduced is $16 k_B T$ with a corresponding maximum of $20.6 k_B T$ and a minimum of $11.4 k_B T$. In the case of the 1:3 MS reported in Table 2.2 it was also not possible to find any critical region. Thus, we employed a similar procedure as for Q_2 in the 1:2 MS case. Comparing both tables we observe that the errors introduced by the ASV in the 1:3 MS are much smaller than in the 1:2 MS.

Finally, in Table 2.3 we show the average of Q over the parametric values of the CG variables. This evaluation of Q is an indirect indication of the three-variable dependence. According to Table 2.3 the correlations due to the third variables do not alter our conclusions, that is the two variables correlation are more representative for this system than the three-variables correlations.

Table 2.1: Quality Factor for 1:2 MS (average in fixed points space)

$< Q\left(\bar{R}^{(1)} = 2.44\text{Å}\right) >$	$< Q\left(\bar{R}^{(2)} = 2.47\text{Å}\right) >$	$< Q\left(\bar{\Omega} = 162°\right) >$
9.10 ± 0.40	16.00 ± 4.60	6.20 ± 1.60

Table 2.2: Quality Factor for 1:3 MS (average in fixed points space)

$< Q\left(\bar{R}^{(1)} = 2.7\text{Å}\right) >$	$< Q\left(\bar{R}^{(2)} = 3.29\text{Å}\right) >$	$< Q\left(\bar{\Omega} = 170°\right) >$
3.50 ± 1.20	7.20 ± 2.70	2.10 ± 0.60

Table 2.3: Quality Factor (average in parametric space)

MS	$< Q\left(\bar{R}^{(1)}\right) >$	$< Q\left(\bar{R}^{(2)}\right) >$	$< Q(\bar{\Omega}) >$
1:2	8.20 ± 1.50	15.0 ± 6.00	5.40 ± 2.50
1:3	4.90 ± 2.90	6.50 ± 3.70	3.00 ± 1.60

2.7 Preliminary conclusions

The aim of this part of the thesis was to develop a formal procedure for analyzing the approximation of separation of variables in certain problems where it is required. First we introduced the mathematical basis of the procedure and summarized in section 2.2. The extension to higher dimensions (e.g. complex systems) was given in section 2.4.

This procedure was applied first to a simple system, namely the diatomic molecule on a flat surface. Our physical intuition tells us that a clear separation of variables can be obtained among the distance r from the center of mass to the surface and the orientation angle θ of molecule, as long as the molecule explores configurations far away from the surface. In fact, this is corroborated from the application of our procedure for this system. The advantage is that now we are able to know in all the configurational space of the molecule the regions where the separation of variables is still questionable up to a known error in energy.

Finally, the second system represented a non-trivial case due the larger number of DOFs compared to the previous system. For this case, we generalized the criteria used in the previous example in order to estimate the error of the separation of variables among the collective variables (e.g. $R^{(1)}, R^{(2)}$ and Ω). Methodologically we tested our procedure on two different mapping schemes, which maintain the same set of CVs, but not the same dependence on the atomistic variables. In summary, our method showed that the 1:3 MS, which results in a coarser model, introduces a smaller error in the energy than the 1:2 MS.

Chapter 3

The Adaptive Resolution Scheme

In recent years there has been a continuing growth of interest in multiscale modeling due to the applications in many fields such as chemistry, biology and material science. Most problems in such areas are related to the interplay between different time and length scales; this means that relevant properties of many systems are typically determined by the interplay of these various temporal and spatial scales. Generally, it is useful to divide the multiscale problem into several scales making a simpler description of the system possible.

However, there are certain categories of problems where it is not possible to perform such separation within reasonable numerical accuracy. Typical examples of such situation is found in the description of edge dislocation in metal [25], cracks propagation in solid materials [26] or large molecules on metal surfaces [27] where the local chemistry affects the large scale properties and vice versa. As a consequence, the system of interest has to be described in a detailed manner, which turns out impossible to be handled computationally due to the large number of degrees of freedom.

In order to overcome this bottleneck, and to study such systems was first incorporated several levels of descriptions based on a hierarchy of theories, for instance, whether the problem requires, from a quantum until a mesoscopic description, which will take account of both the small and large scale phenomena. A systematic coarse graining (CG) procedure may help to build each level of description based on information accessible on previous scales. So far, there are diverse hybrid multiscale techniques aiming to bridge the gap between closer scales, for instance, the atomistic and mesoscopic scales [27–31] or the quantum and classical scales [32,33]. However, in all those methods the regions or parts of the system treated at different level of resolution are fixed and thus the exchange of particles among these regions is not allowed. This approximation turns out to be not very relevant in hard condensed matter, but it becomes crucial in soft matter systems. In the former bulk properties are determined by the strength of the intermolecular forces, which restrict the motion of particles, and in the latter thermal fluctuations (e.g. density fluctuations) contribute to the overall behavior of the system.

In this context, the adaptive resolution methods which not only couple diverse scales (or resolutions) and also allow for particle fluctuations represent the most natural way to overcome such problems. Recently, some methods based on this idea have captured the attention of many researchers and several schemes have been developed for coupling the atomistic and coarse grained level of description in classical MD. Typically, the coupling can be performed through the smooth interpolation of forces (AdResS) [34], potentials [35] and Lagrangians [36] by using a switching or interpolating function. Although the equation of motion in the AdResS method cannot be derived within a Hamiltonian formalism, it has been shown to preserve the thermodynamic equilibrium of the whole system and the result is independent of the switching function. The second method integrates in principle the same set of equation as the AdResS method and it claims to conserve the total energy of the system. This has been shown to be flawed, because the total Hamiltonian cannot be defined in both hybrid schemes [37]. Finally, the last method presents an energy conserving protocol whose dynamics depends explictly on the switching function and its derivative. Moreover, its implementation is too complex for large applications.

In any case, all of them must not change the physics of the system, since the change of resolution does not affect the physical nature of particles. Thus, adaptive schemes should preserve the thermodynamic equilibrium during the simulation, this implies that thermal, mechanical and chemical equilibrium should not be modified by the scheme of interpolation used.

As a part of this thesis work, we have studied the quantum/classical adaptivity in the framework of force interpolation and this topic will be discussed in the following chapters. In the present chapter, we will give an overview of the Adaptive Resolution Scheme (AdResS); we start describing the equation of motion of AdResS, then a description of the thermodynamic equilibrium of a system follows, where the change of degrees of freedom is allowed. Then we comment about the theoretical foundation of the AdResS scheme. Finally, the method is applied to a liquid system of tetrahedral molecules as proof of validity of the scheme.

3.1 The equation of motion

The first step is to derive the effective Coarse-Grained potential between the interacting sites, mapped at the center of mass of each molecule, by the iterative Boltzmann inversion method, presented in chapter 1. This CG potential has to be obtained at the same thermodynamic state point, thus a pressure correction must be done to retain the same pressure. Once this is done, one proceeds to couple the forces derived from the atomistic (AT) and coarse-grained (CG) potentials by the following expression,

$$F_{\alpha\beta} = w(R_\alpha)w(R_\beta)F^{\mathrm{AT}}_{\alpha\beta} + (1 - w(R_\alpha)w(R_\beta))F^{\mathrm{CG}}_{\alpha\beta} \tag{3.1}$$

where α an β are the labels for two different molecules. $F_{\alpha\beta}^{\mathrm{AT}}$ is the corresponding force derived from the atomistic force field where each atom of molecule α interacts with each atom of β, and $F_{\alpha\beta}^{\mathrm{CG}}$ is the pairwise force obtained from the CG potential between the centers of mass of the coarse grained molecules. One important element of this equation is given by the "weighting function", $w(R)$, whose functional form is shown in Figure 3.1, varying from 0 to 1. This function depends on the position (R) of the center of mass of the molecules α and β. A simple way to interpret this function is by switching of "degrees of freedom". From the Eq. 3.1, it is evident that $w = 0$ represents the case of pure coarse-grained force field, while $w = 1$ keeps the system fully atomistic. The region of non-integer values of w is called the "hybrid region" and there the particles maintain at the same time a double resolution or representation (eg. AT/CG).

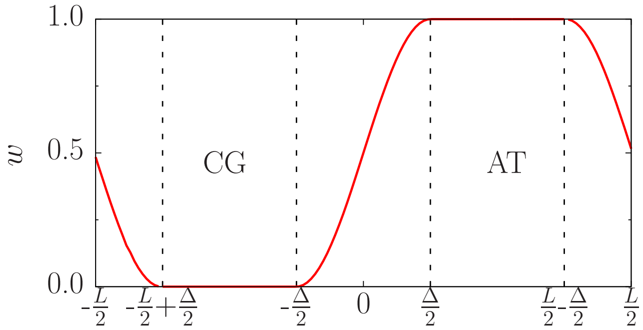

Figure 3.1: Here is depicted the functional form of the weighting function $w(x) \in [0, 1]$. The values 1 and 0 correspond to the regions where the molecules are fully atomistic (AT) or fully coarse grained (CG) respectively. While values in between, $0 < w(x) < 1$, correspond to the hybrid region. In this figure, the total box length is equal to L, atomistic and CG region have the same length and Δ is the length of hybrid region.

An important consequence of the analytical form of Eq. 3.1 is that, by construction, it preserves Newton's third law (conservation of linear momentum), despite the fact that a Hamiltonian in the transition region cannot be defined. This guarantees that the diffusion of particles between regions is not affected by the change of resolution.

3.2 Thermodynamic equilibrium

The Eq. 3.1 cannot be obtained from a potential and thus there would not have an energy to conserve in such circumstances. This scheme resembles a non-Hamiltonian equation of motion, where new DOFs are couple to the system, in order to design new MD schemes for different equilibrium ensembles [38]. Since we want to study systems in equilibrium, a natural question arises immediately, how to control the thermodynamic equilibrium. Conceptually, in the adaptive scheme the number of degrees of freedom is not homogeneous in the space and therefore the free energy density is not uniform. This situation creates non-physical flux of particles in the direction of less DOFs in order to lower the free energy of the whole system. In spite of that, one expects the same physical scenario everywhere by construction (i.e. same state point), this means that all molecules must maintain the same underlying physical nature in all the space and later one must learn how to deal with the artifact of the formalism used. An illustrative way to understand this process is the following: when a molecule goes from an atomistic to a coarse grained region, it experiences a transition where it loses vibrational and rotational DOFs and when it arrives to the CG region a natural process of accommodating its exclude volume may take place. The inverse process is more complicate, a molecule in this case acquires rotational and vibration DOFs and tries to enter in a region where the other molecules are locally in equilibrium. In such circumstances, a way back to the coarse grained region is more preferable than remaining in the atomistic one. In thermodynamic terms, as an artifact of the method, the different regions are characterized by a different chemical potential. Since, this is a consequence of the formalism and it is not generated by the physics of the system, thus one has to correct this thermodynamic unbalance. Based on these arguments, one sees that Eq. 3.1 alone cannot maintain the thermodynamic equilibrium and further considerations concerning the variables of the problem, should be used to guarantee the equilibrium. This is the aim of the next sections, by analyzing the process of changing degrees of freedom from a thermodynamic and statistical framework.

3.3 Theoretical principles of thermodynamic equilibrium in AdResS

We present the theoretical basis for the thermodynamic equilibrium of a system where the number of DOFs are, by construction, space dependent and yet molecular properties are maintained as close as possible to the reference system in all the space.

3.3.1 The geometrically induced phase transition

We provide a parallel description between the space dependent change of resolution and the physical phase transition, in this context we will identify the former as a ficticious geometrically induced phase transition. To describe a phase transition, one uses the concept of the latent heat to associate the energy required by the system to account for a transition. For example, typically the transition from a liquid to a gas phase requires energy (latent heat) to activate those vibrational modes that make the molecules free from the tight bonding of the liquid state. In the adaptive scheme a similar process occurs to a molecules which passes from a coarse grained to an atomistic resolution, in this case such molecule needs latent heat to reactivate the vibrational and rotational DOFs in order to reach the equilibrium with the atomistic surrounding. In the reverse process, the molecule releases latent heat, when a transition from a gas to a liquid phase occurs, during such transition a molecule increases the bond to the other molecules at each time, in the same way in the adaptive scheme, the passage from the atomistic to the coarse grained description happens, formally losing DOFs and therefore the associate heat. All this is synthesized in the following relation:

$$\mu^{AT} = \mu^{CG} + \phi, \tag{3.2}$$

where μ^{AT} is the chemical potential calculated with the atomistic representation, μ^{CG} that of the coarse grained one, and ϕ is the latent heat associated to the process. To satisfy Eq. 3.2 a simple solution is devised, basically one has to couple the system to a local thermostat (see Appendix C), whose main function is to provide (or removes) the required latent heat. Such thermostat ensures the equilibrium and the stability of the algorithm. Naturally, such a coupling raises serious questions on how to define thermodynamic quantities in a region where the number of DOFs is space-dependent and Hamiltonian is not defined. This question is answered in the next section.

3.3.2 Thermodynamic quantities in AdResS

In this section we describe all the thermodynamic quantities used typically in the terminology of AdResS. These quantities are relevant for a cross check of the thermodynamic equilibrium during an AdResS simulation. As we stated before the thermodynamic equilibrium is maintained in each region provided that the average of temperature, pressure and chemical potential do not change in the MD simulation.

In order to define the pressure in a system where atomistic and coarse-grained particles coexist, one proceeds to use the concept of molecular pressure instead of the atomistic one. The equivalence between these two expression has been proved by Ciccotti et al. (1986) [39] and discussed recently [40]. The molecular pressure is given by,

$$p = \frac{1}{V}[Nk_BT + \frac{1}{3}\sum_{\alpha}\sum_{\beta>\alpha} R_{\alpha\beta}\cdot F_{\alpha\beta}] \tag{3.3}$$

where N is the number of molecules, V is the volume of system, T is the temperature and $R_{\alpha\beta}$ and $F_{\alpha\beta}$ correspond to distance and total force (see Eq. 3.1) between the molecules α and β.
The temperature is provided by the equipartition theorem [41].

$$T^{\mathrm{AT/CG}} = 2\frac{< K^{\mathrm{AT/CG}} >}{N^{\mathrm{AT/CG}}}, \tag{3.4}$$

where $< K^{AT/CG} >$ represent the average kinetic energy of the atomistic/coarse grained region and $N^{AT/CG}$ is the total average number of degrees of freedom

(DOFs) in the respective representation. This principle is well-defined in each region (AT/CG), but the same can not be applied in the transition region where $\mathcal{N} = \mathcal{N}(x)$. Therefore, in the hybrid region one has to extend the principle of equipartition for a switchable DOF q to properly define its kinetic contribution to the temperature. To account for that, we observe the space dependency of such DOF in each region, being fully represented in the atomistic region and vanishing in the coarse grained region. This behavior should be taken into account, when calculating the average of statistical quantities. In a formal mathematical language, this is a common problem in fractional calculus [42]. Using this mathematical tool to describe the change of dimensionality of the phase space of q (between one and zero), one has that

$$dV_w = \frac{\Gamma(\frac{w}{2})}{2\pi^{w/2}\Gamma(w)} d^w q = \frac{|q|^{w-1}}{\Gamma(w)} dq = \frac{1}{w\Gamma(w)} dq^w \tag{3.5}$$

where Γ is well-known Γ-function [43]. The kinetic contribution of a quadratic DOFs is given by

$$< K_q >_w = \frac{\int_0^\infty e^{-\beta q^2} q^{w+1} dq}{\int_0^\infty e^{-\beta q^2} q^{w-1} dq} \tag{3.6}$$

and the solution of Eq. 3.6 was demonstrated to be [44]:

$$< K_q >_w = \frac{w}{2\beta}. \tag{3.7}$$

The last result generalizes the equipartition theorem for a non integer DOFs whose functional form is quadratic. This states that the kinetic energy is proportional to its dimensionality (w).

So far, all the concepts presented in this section help to control thermodynamic quantities (Eq. 3.3) and establish a thermodynamic consistency (Eq. 3.6) in our studies within the AdResS framework.

3.4 Application to simple liquid of tetrahedral molecules

The generality of the results shown in this section are independent of the system under study. More complex systems as the solvation of an ideal bead-spring model for a polymer in a tetrahedral liquid and liquid water were carried out in refs. [45, 46]. The results for the tetrahedral molecules in AdResS were compared to the atomistic ones. All the result were obtained in a cubic symmetry, however in Figure 3.2 we depict the same model system for another symmetry (i.e. spherical). The following functional form was proposed for the weighting

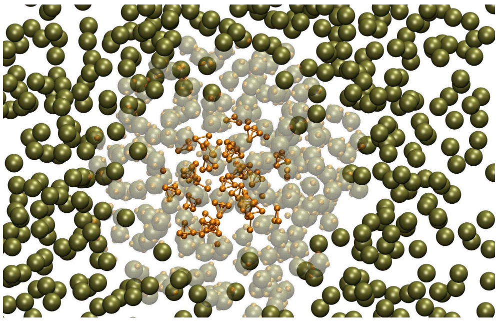

Figure 3.2: Snapshot of tetrahedral molecules in a spherical symmetric in AdResS. Atomistic molecules are represented in the inner shell, subsequently follows the shell that contains hybrid molecules which is surrounded by CG particles.

function $w(x)$:

$$w(x) = \begin{cases} 1; & d < x \leq \frac{a}{2} - d \\ 0; & -\frac{a}{2} + d \leq x < -d \\ \sin^2[\frac{\pi}{4d}(x+d)]; & -d \leq x \leq d \\ \cos^2[\frac{\pi}{4d}(x - \frac{a}{2} + d)]; & \frac{a}{2} - d < x \leq \frac{a}{2} \\ \cos^2[\frac{\pi}{4d}(x + \frac{a}{2} + d)]; & -\frac{a}{2} \leq x < -\frac{a}{2} + d, \end{cases} \tag{3.8}$$

where a is the box length and d is the half-width of the hybrid region. In general this function has to be monotonic, continue, differentiable and with zero slope at boundaries of the atomistic and coarse grained regions. These mathematical assumptions guarantee a smooth transition of one molecule from the CG region to the atomistic and vice versa [47].

Once the atomistic force field (see Appendix D) is well-defined we proceed to derive from it the effective coarse grained potential. In Figure 3.3 the CG potential, between the center of mass of the molecules U^{cm}, obtained for a liquid of tetrahedral molecules at $\rho^* = 0.1$ and $T^* = 1.0$ in the reduced Lennard-Jones units ($\epsilon = 1$ and $\sigma = 1$ as the unit of energy and length respectively).

The first way to compare the reliability of the AdResS method is by comparison of the the global structure of the liquid in AdResS with the result of the full atomistic simulation. This is done by computing the center of mass radial distribution function (RDF). The result is shown on the left side of the Figure 3.4. One can see a satisfactory agreement between the AdResS method and the full atomistic reference calculations. Similarly, the density profile of the molecules in AdResS (See right side of Figure 3.4) is presented . From this figure one can see how the density is maintained homogeneous in the atomistic and coarse grained regions, however, in the hybrid region the density drops about 5% compared to to the average value of $\rho^* = 0.175$ of the reference all atom case. However, this small perturbation neither affects the structure nor the thermodynamics of the system.

As a final test we show the study of the diffusion of molecules close to the

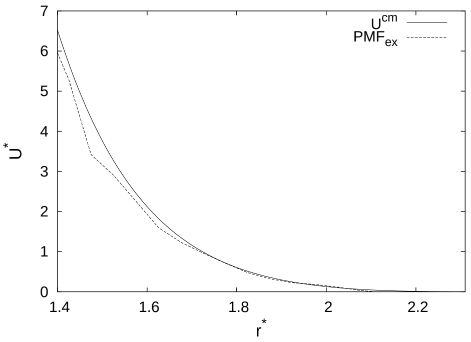

Figure 3.3: The effective CG potential obtained by the iterative Boltzmann method [16] and the potential of mean force (PMF) for the highly diluted explicit system at $\rho^* = 0.0025$ used as the first guest in the iterative procedure are shown. (Figure was taken from Ref. [34]).

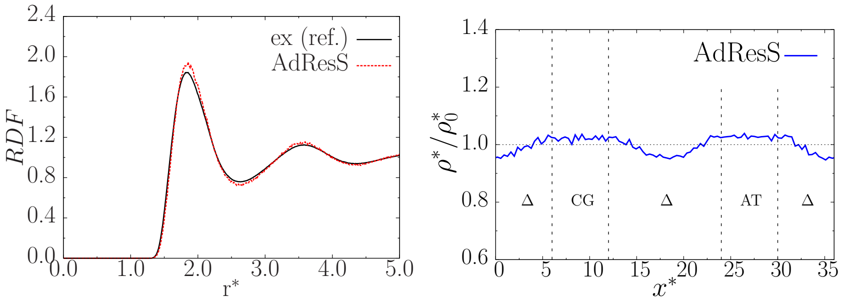

Figure 3.4: Left part shows the radial distribution function of the center of masses at $\rho^* = 0.175$ and $T^* = 1$. Right part shows the typical normalized density profile in the x direction in AdResS with 12.0σ interface layer width. Vertical lines denote the boundary between atomistic, hybrid and coarse grained region.

hybrid region in Figure 3.5. This result tells us that no artificial effect like spurious reflexion of molecules, occurs while they move from one region to the other one.

3.5 Further development

There are mainly two new contributions to the AdResS method. One has to do with the generalization of the scheme based on thermodynamic arguments, by means of a thermodynamic force to ensure the equilibrium from first principles of thermodynamics. Concerning this first part one must say that equilibrium in the traditional AdResS is ensured by the use of a stochastic thermostat which removes or adds latent heat to the system in order to maintain the thermal

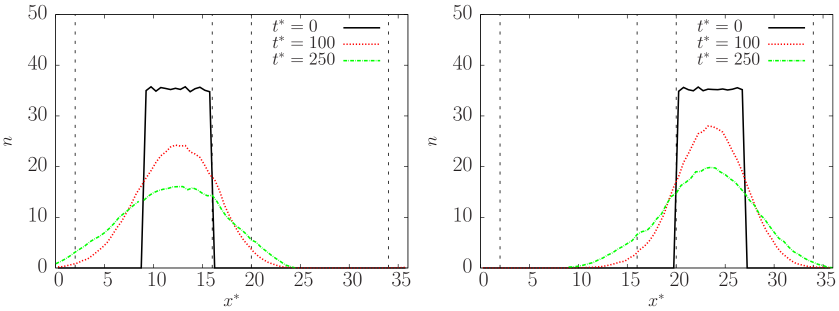

Figure 3.5: Time evolution of diffusion profiles of molecules for two different initial condition at two neighboring slabs of the hybrid region. Left side corresponds to the case of molecules localized on the CG side and in right side to molecules localized in the atomistic region.

equilibrium. The formal demonstration of this generalized scheme can be found in [48].

The other important contribution is a conceptual improvement to the method. It concerns the extension of the scheme for a wide variety of problems where quantum mechanical description may take place. This is related to the current thesis work and is primarily devoted to the understanding of the quantum-classical adaptivity. The next chapter provides a general overview of the path integral formalism. Such a method was used in our studies of quantum-classical coupling due to a key feature of its formalism. This is given by the fact that path integral method translates the representation of a quantum particle into the evaluation of certain classical objects (i.e. ring polymers).

Chapter 4

Path Integral formulation

The basic idea of the path integral formulation of quantum mechanics can be traced back to P. A. M. Dirac in his original book [49] of 1930 and later on his paper [50] of 1933. Later, the completed method was developed by Richard Feynman [51] in 1948 as an alternative formulation of the non-relativistic quantum mechanics and since then it is widely used in several fields of science, for example, many-body theoretical quantum physics [52], superfludity [53], polymer science [54, 55], financial markets [56], to name a few.

The path integral formalism is convenient not only for its mathematical elegance, but also for its treatable numerical form with a structure that is suitable for an implementation in parallel computing. Furthermore, path integrals have been succeeded in calculating several equilibrium properties like the free energy and structural quantities of quantum systems [57–59] in comparison with other quantum methods. Finally, path integral can be used in several thermodynamic ensembles, for instance, the microcanonical (NVE) [60], the canonical (NVT) [61], isothermal-isobaric (NPT) [62] and gran-canonical (μVT) [63]. This versatility allows us to study systems under several experimental conditions. The path integral formulation translates the quantum description of a many body problem into the classical representation due to the so-called isomorphism of the quantum partition function.

The next section describes in detail the aforementioned isomorphism, then we introduce the idea of how to combine path integral formulation with molecular dynamics (PIMD) and the calculation of statistical properties. In the end we comment about the limitation of the PIMD scheme and possible numerical solutions.

4.1 Derivation of the formalism

One quantum particle

The definition of the partition function for a system of a single quantum particle is

$$\mathcal{Z} \equiv \mathcal{Z}(N, V, T) = \mathrm{Tr}[\mathrm{e}^{-\beta\mathcal{H}}] \tag{4.1}$$

where "Tr" denotes a trace and $\beta = 1/k_BT$. T is the temperature and k_B is the Boltzmann constant. This trace can be evaluated in the position eigenstates, $|R\rangle$, as follows

$$\mathcal{Z} = \int dR\ \langle R|e^{-\beta\mathcal{H}}|R\rangle, \tag{4.2}$$

Given that the kinetic and the potential energy operator do not commute,

$$[\hat{T}, \hat{V}] \neq 0, \tag{4.3}$$

one can use the Trotter theorem [64], which states that for any two operator, A and B, which in general do not commute

$$e^{\lambda(A+B)} = \lim_{n\to\infty}[e^{\frac{\lambda}{2n}B}e^{\frac{\lambda}{n}A}e^{\frac{\lambda}{2n}B}]^n, \tag{4.4}$$

where n is an integer and known as the Trotter number. Now substituting the Trotter theorem into Eq. 4.2 yields,

$$\mathcal{Z} = \lim_{n\to\infty}\int dR\ \langle R|\Omega^n|R\rangle = \lim_{n\to\infty}\int dR\ \langle R|\Omega\cdot\Omega\cdots\Omega|R\rangle \tag{4.5}$$

For simplicity, we define

$$\Omega = e^{-\beta\hat{V}/(2n)}e^{-\beta\hat{T}/n}e^{-\beta\hat{V}/(2n)}. \tag{4.6}$$

Introducing the identity operator

$$I = \int dR\ |R\rangle\langle R|, \tag{4.7}$$

$n-1$ times in the Eq. 4.5 in the following way

$$\Omega^n = \Omega I_2 \Omega I_3 \cdots \Omega I_n \Omega \tag{4.8}$$

and using the definition of the identity, the integration term in Eq. 4.5 results

$$\begin{aligned} \langle R|\Omega^n|R\rangle &= \int dR_2\ldots dR_n\langle R|\Omega|R_2\rangle\langle R_2|\Omega|R_3\rangle\times \\ &\quad \times\langle R_3|\ldots|R_n\rangle\langle R_n|\Omega|R\rangle \\ &= \int dR_2\ldots dR_n[\prod_{i=1}^{n}\langle R_i|\Omega|R_{i+1}\rangle|_{R_1=R_{n+1}} \end{aligned} \tag{4.9}$$

where the condition $R_1 = R_{n+1}$ is the result of the trace. One can evaluate each matrix element of Ω,

$$\langle R_i|\Omega|R_{i+1}\rangle = \langle R_i|e^{-\beta\hat{V}/(2n)}e^{-\beta\hat{T}/n}e^{-\beta\hat{V}/(2n)}|R_{i+1}\rangle. \tag{4.10}$$

One knows that the potential operators are space dependent and they are acting on the coordinate eigenstates in the last equation. Thus, the following expression is derived

$$\langle R_i|\Omega|R_{i+1}\rangle = e^{-\beta V(R_i)/(2n)}\langle R_i|e^{-\beta\hat{T}/n}|R_{i+1}\rangle e^{-\beta V(R_{i+1})/(2n)}. \tag{4.11}$$

Now, we use the identity operator in the representation of momentum eigenstates,

$$I = \int dp\ |p\rangle\langle p|, \tag{4.12}$$

then the remaining matrix elements can be written as

$$\langle R_i|e^{-\beta\hat{T}/n}|R_{i+1}\rangle = \int dp\ \langle R_i|p\rangle\langle p|e^{-\beta\hat{T}/n}|R_{i+1}\rangle. \tag{4.13}$$

In the last expression, $\hat{T} = p^2/2m$ acts on one of its eigenstates from the left, yielding:

$$\langle R_i|e^{-\beta\hat{T}/n}|R_{i+1}\rangle = \int dp\ \langle R_i|p\rangle\langle p|R_{i+1}\rangle\ e^{-\beta p^2/(2mn)}. \tag{4.14}$$

The projection of a momentum eigenstate on a position eigenstate is given by

$$\langle R|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{ipR/\hbar}. \tag{4.15}$$

Then we find that

$$\langle R_i|e^{-\beta\hat{T}/n}|R_{i+1}\rangle = \frac{1}{2\pi\hbar}\int dp\ e^{ip(R_i-R_{i+1})/\hbar}e^{-\beta p^2/(2mn)}. \tag{4.16}$$

To solve the integral, we complete the square in the exponential and then integrate as follows:

$$\begin{aligned}\langle R_i|e^{-\beta\hat{T}/n}|R_{i+1}\rangle &= \frac{1}{2\pi\hbar}\int dp\ e^{-\frac{\beta}{2mn}[p-(\frac{mni}{\hbar\beta}(R_i-R_{i+1}))]^2} \times \\ &\quad \times e^{-\frac{mn}{2\beta\hbar^2}(R_i-R_{i+1})^2}.\end{aligned} \tag{4.17}$$

The first exponential in the integral is a Gaussian-like and the second is a constant, thus one obtains,

$$\langle R_i|e^{-\beta\hat{T}/n}|R_{i+1}\rangle = (\frac{mn}{2\pi\beta\hbar^2})^{1/2}\ e^{-\frac{mn}{2\beta\hbar^2}(R_i-R_{i+1})^2} \tag{4.18}$$

Substituting our last result into the expression for the whole partition function

$$\begin{aligned}\mathcal{Z} &= \lim_{n\to\infty}(\frac{mn}{2\pi\beta\hbar^2})^{n/2}\int dR_1\ldots dR_n \times \\ &\quad exp\left(-\beta\sum_{i=1}^{n}[\frac{1}{2}m\omega_p^2(R_i-R_{i+1})^2+\frac{1}{n}V(R_i)]\right)_{R_{n+1}=R_1},\end{aligned} \tag{4.19}$$

and,

$$\omega_n = \frac{\sqrt{n}}{\beta\hbar}, \tag{4.20}$$

is the "frequency" of the ring polymer.

The outcome indicates the isomorphism between a static quantum mechanical problem and the classical problem of a replicate classical particle which interacts with a potential $V(R)/n$ and two of its own images through a quadratic potential. We illustrate the idea of such a isomorphism in Figure 4.1. This result is well-known as the discrete path integral for the quantum partition function, which is exact for a large value of Trotter number (i.e, $n \to \infty$).

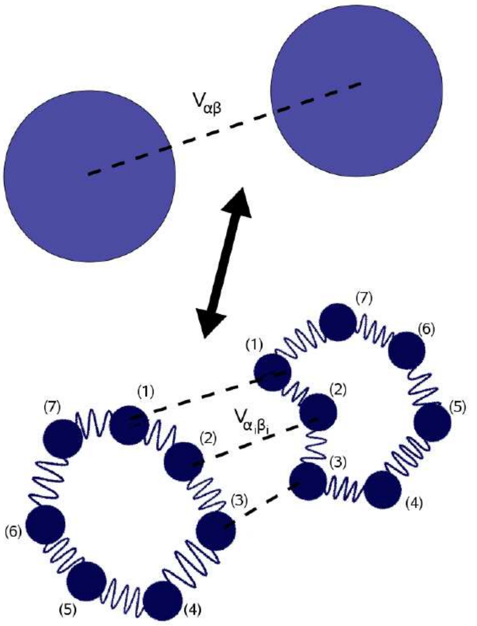

Figure 4.1: Path integral representation of 2 quantum particles with $n = 7$ beads each. In the path integral formalism the interaction between the beads with the same label is given by the classical potential $V_{\alpha\beta}$.

Many quantum particle

The same result obtained for an isolated particle can be extended for a system of N-particles interacting by the potential $V(\{R_I\})$. This time each particles is represented by a ring polymer and so that the interaction is delocalized among its conforming beads. In other words, beads with the same label "s" will interact by the potential $V(\{R_I^{(s)}\})$ and no cross interactions are allowed as it is depicted in Fig. 4.1. For example, the bead 1 of the molecule α will only interact with the bead 1 of the molecule β and so on. Such generalization is not trivial and can be found in traditional textbooks of path integral methods as [65, 66]. In summary, the quantum partition function of N particles is given by,

$$\begin{aligned} \mathcal{Z} &= \lim_{n\to\infty} \mathcal{Z}_n(\beta) = \lim_{n\to\infty} [\prod_{I=1}^{N} (\frac{m_I n}{2\pi\beta\hbar^2})^{n/2} \int dR_I^{(1)} \ldots dR_I^{(n)}] \times \\ & exp\left(-\beta \sum_{s=1}^{n} \{ \sum_{I=1}^{N} \frac{1}{2} m_I \omega_n^2 (R_I^{(s)} - R_I^{(s+1)})^2 + \frac{1}{n} V(\{R_I^{(s)}\}) \}\right) \end{aligned} \quad (4.21)$$

where $R_I^{(n+1)} = R_I^{(1)}$ and ω_n is given by Eq. 4.20. Finally, the above expression can be sampled by Monte Carlo (MC) methods or by molecular dynamics (MD) adding some effective Gaussian distribution in the momentum space, $P_I^{(s)}$. The next section will show how to perform a MD implementation of the path integral formalism.

4.2 The Path Integral Molecular Dynamics (PIMD)

The path integral expression of Eq. 4.21 of the quantum N-particles system can be written in the following form

$$\begin{aligned} \mathcal{Z}_n(\beta) &= [\prod_{I=1}^{N} (\frac{m_I n}{2\pi\beta\hbar^2})^{n/2} \int dR_I^{(1)} \ldots dR_I^{(n)}] \times \\ & \times exp\left(-\beta \sum_{s=1}^{n} \left\{ \sum_{I=1}^{N} \frac{1}{2} m_I \omega_n^2 (R_I^{(s)} - R_I^{(s+1)})^2 + \frac{1}{n} V(\{R_I^{(s)}\}) \right\}\right) \\ &= \prod_{I=1}^{N} \mathcal{N} \int dR_I^{(1)} \ldots dR_I^{(n)} e^{-\beta \mathcal{U}_{\text{eff}}(R_I^{(1)} \ldots R_I^{(n)})} \end{aligned} \quad (4.22)$$

where $\mathcal{N} = (\frac{m_I n}{2\pi\beta\hbar^2})^{n/2}$ is a constant and

$$\mathcal{U}_{\text{eff}} = \sum_{s=1}^{n} \{ \sum_{I=1}^{N} \frac{1}{2} m_I \omega_n^2 (R_I^{(s)} - R_I^{(s+1)})^2 + \frac{1}{n} V(\{R_I^{(s)}\}) \} \quad (4.23)$$

is considered as the effective potential.

The Eq. 4.22 looks like the classical configuration partition function for a system of N ring polymers with n particles (or beads) each. Thus, it is the so-called isomorphism of the quantum partition function due to the path integral formalism.

To perform a molecular dynamics implementation [67] one needs to use conservative forces derived from the Eq. 4.23, which also obey the equipartition theorem. From Eq. 4.22 one can see the connection with Molecular Dynamics (MD) by adding n-Gaussian integrals under the condition that,

$$\left(\frac{\beta}{2\pi m_I}\right)^{n/2}\int dP_I^{(1)}\cdots dP_I^{(n)} exp(-\beta\sum_{s=1}^{n}\sum_{I=1}^{N}\frac{P_I^{(s)2}}{2m_I'})=1 \tag{4.24}$$

where m_I' is an arbitrary mass parameter. The value of this parameter affects the rate at which the molecular dynamics trajectories move and thus, the efficiency of the sampling. Inserting last equation into Eq. 4.22 gives,

$$\begin{aligned}\mathcal{Z}_n(\beta) &= \prod_{I=1}^{N}\mathcal{N}'\int dR_I^{(1)}\ldots dR_I^{(n)}\int dP_I^{(1)}\ldots dP_I^{(n)}\times\\ &\times exp(-\beta\sum_{s=1}^{n}\{\sum_{I=1}^{N}\frac{P_I^{(s)2}}{2m_I^{'}}+\frac{1}{2}m_I\omega_n^2(R_I^{(s)}-R_I^{(s+1)})^2+\\ &+\frac{1}{n}V(\{R_I^{(s)}\})\}),\end{aligned} \tag{4.25}$$

where $\mathcal{N}'$ is a new normalization factor. The Gaussian variables are uncoupled and can be integrated analytically to obtain the prefactor $\mathcal{N}$ from Eq. 4.22. The derivation presented here only involves the partition function, thus only statistical properties of the quantum system can be calculated. This means, that although one can derive a corresponding Hamiltonian for the whole partition function as

$$\mathcal{H}=\sum_{s=1}^{n}\{\sum_{I=1}^{N}\frac{P_I^{(s)2}}{2m_I^{'}}+\frac{1}{2}m_I\omega_n^2(R_I^{(s)}-R_I^{(s+1)})^2+\frac{1}{n}V(\{R_I^{(s)}\})\}, \tag{4.26}$$

the dynamics of this system cannot be directly related to the quantum system, but rather only the statistical properties of the ensemble which are the results of the sampling of this pseudo-dynamics. Formally, now we can use the term PIMD to denominate the Hamiltonian obtained above. The PIMD technique uses the classical dynamics of the ring-polymer Hamiltonian, i.e.,

$$\begin{aligned}\dot{R}_I^{(s)} &= \frac{\partial\mathcal{H}}{\partial P_I^{(s)}}=\frac{P_I^{(s)}}{m_I^{'}},\\ \dot{P}_I^{(s)} &= -\frac{\partial\mathcal{H}}{\partial R_I^{(s)}}=-m_I\omega_n^2[2R_I^{(s)}-R_I^{(s+1)}-R_I^{(s-1)}]\\ &\quad -\frac{dV(\{R_I^{(s)}\})}{dR_I^{(s)}}\end{aligned} \tag{4.27}$$

to propagate the trajectories and thus sample the phase space of the ring polymers.

The final result of the path integral in a MD scheme is considered ideal for the design of parallel implementation, turning such representation in a powerful tool to calculate statistical properties of materials in the presence of quantum effects.

4.3 Static properties from PIMD

The path integral formulation of the quantum statistical mechanics provides an efficient method to evaluate the quantum static equilibrium properties of a system of N particles. Suppose we want to calculate the expectation value of the position-dependent operator $\hat{A}$. By definition, the expectation value of $\hat{A}$ is:

$$\langle \hat{A} \rangle = \frac{\mathrm{Tr}[\hat{\mathrm{A}}\mathrm{e}^{-\beta \mathcal{H}}]}{\mathrm{Tr}[\mathrm{e}^{-\beta \mathcal{H}}]} = \frac{\mathrm{Tr}[\hat{\mathrm{A}}\mathrm{e}^{-\beta \mathcal{H}}]}{\mathcal{Z}} \tag{4.28}$$

This evaluation can be performed using the path integral approach described above. The result is

$$\begin{aligned} \langle \hat{A} \rangle \quad = \quad & \frac{1}{\mathcal{Z}(\beta)} \lim_{n\to\infty} \prod_{I=1}^{N} (\frac{m_I n}{2\pi\beta\hbar^2})^{n/2} \int dR_I^{(1)} \ldots dR_I^{(n)} A(R_I^{(1)}) \times \\ & exp(-\beta \sum_{s=1}^{n} \{ \sum_{I=1}^{N} \frac{1}{2} m_I \omega_n^2 (R_I^{(s)} - R_I^{(s+1)})^2 + \frac{1}{n} V(\{R_I^{(s)}\}) \}) \end{aligned} \tag{4.29}$$

with $R_I^{(n+1)} = R_I^{(1)}$. The above integral is invariant under a cyclic relabeling of all the path integration variables, $R_I^{(1)} \to R_I^{(2)}, R_I^{(2)} \to R_I^{(3)}$, and so forth. Such a relabeling can be carried out n times, the resulting expression added together and divided by n equals:

$$\begin{aligned} \langle \hat{A} \rangle \quad = \quad & \frac{1}{\mathcal{Z}} \lim_{n\to\infty} \prod_{I=1}^{N} (\frac{m_I n}{2\pi\beta\hbar^2})^{n/2} \times \\ & \int dR_I^{(1)} \ldots dR_I^{(n)} \frac{1}{n} \sum_{s=1}^{n} A(R_I^{(s)}) e^{-\beta \mathcal{U}_{\mathrm{eff}}} \end{aligned} \tag{4.30}$$

where $\mathcal{U}_{\mathrm{eff}}$ is given in Eq. 4.23. Now a finite expression for $\langle \hat{A} \rangle$ can be obtained by substituting $\mathcal{Z}$ from Eq. 4.21. This yields,

$$\begin{aligned} \langle \hat{A} \rangle_n \quad = \quad & \frac{1}{\mathcal{Z}_n(\beta)} \lim_{n\to\infty} \prod_{I=1}^{N} (\frac{m_I n}{2\pi\beta\hbar^2})^{n/2} \times \\ & \int dR_I^{(1)} \ldots dR_I^{(n)} \frac{1}{n} \sum_{s=1}^{n} A(R_I^{(s)}) e^{-\beta \mathcal{U}_{\mathrm{eff}}} \end{aligned} \tag{4.31}$$

from which one obtains the true value of the expectation of $\hat{A}$ in the limit of $n \to \infty$:

$$\langle \hat{A} \rangle = \lim_{n\to\infty} \langle \hat{A} \rangle_n \tag{4.32}$$

Eq. 4.31 can be evaluated using a Monte Carlo (MC) scheme, since such average is computed in the configuration space. A trick can be done in order to compute the same average in a Molecular Dynamics (MD) scheme as we did in Eq. 4.25. Substituting the identity 4.24 into Eq. 4.31 we obtain that,

$$
\begin{aligned}
\langle \hat{A} \rangle_n \quad = \quad & \prod_{I=1}^{N} \mathcal{N} \int dR_I^{(1)} \cdots dR_I^{(n)} \int dP_I^{(1)} \cdots dP_I^{(n)} \times \\
& \frac{1}{n} \sum_{s=1}^{n} A(R_I^{(s)}, P_I^{(s)}) e^{-\beta \mathcal{H}}, \qquad (4.33)
\end{aligned}
$$

where $\mathcal{H}$ is the Hamiltonian defined in Eq. 4.26. Since the momenta are only a trick to compute the averages of A in a different scheme, no meaningful quantum dynamics can be extracted from the new set of trajectories. The next section shows the limitations of the MD implementation and a posteriori solutions.

4.4 Limitations of the PIMD and possible solutions

The PIMD equations of motion derived in Eq. 4.27 explore the phase space that is consistent with their energy (i.e. microcanonical ensemble). Thus, they fail to sample a canonical distribution. These fixed-energy trajectories must be modified if they pretend to produce a fixed temperature distribution (canonical distribution). A natural solution for achieving this involves the coupling to a thermostat. Several approaches have been used so far to maintain the temperature, a very popular one involves linking additional vibrational modes onto the physical degrees of freedom of the system [68, 69]. The ficticious thermostat modes are coupled to the momenta of the physical DOFs and regulate the kinetic fluctuations to produce constant temperature trajectories. It turns out that DOFs which are dominated by harmonic motion require the use of not one but a chain of thermostats [70]. Such systems plus the thermostat will require a massive solution of a set of equation of motion. Although the canonical distribution is rigorously reproduced with a thermostat chains of sufficient length [71], the dynamics of the thermostated and the unthermostated system is not clear. This last point is not of relevance but will be crucial in algorithms based on path integral formulation where quantum dynamics can be extracted from [72, 73]. An alternative technique for generating a canonical distribution from microcanonical trajectories involves the use of thermostats whose physical principle involves a periodically resampling of their momenta from a Maxwell-Boltzmann distribution [74]. Physically this thermostat mimics inelastic collision with a thermal bath at fixed temperature. Another type of thermostat commonly used in the molecular dynamics community are the stochastic thermostats due to their local behavior and easy implementation. One can use a Langevin thermostat [75] to compute static properties, but it is well-known that such a thermostat does not preserve the true dynamics of the system. Recent development of more sofisticated stochastic thermostats such as the Dissipative Particle Dynamics (DPD) has shown to preserve the hydrodynamics [76] or the possibility to control transport coeficients [77]. Since hydrodynamics is not important in the scale of observation, therefore a simple langevin thermostat will be sufficient to properly thermostat the ring polymers.

A second limitation of the direct implementation of the Eq. 4.27 is due to the large number of beads required to approach the true value of quantum partition function. As one increases this number, the harmonic spring terms become stiffer and start to dominate the dynamics. Thus, the sampling of the entire spectrum of frequencies becomes poorly evaluated. It is known that microcanonical trajectories in such system may not event follow a microcanonical distribution and ergodicity problems may arise [78], i.e. on the time scale accessible to a computer simulation. This can be synthesized as follows:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} A(R^{(i)}, t_i) \neq \frac{1}{\Omega(E)} \int dP \int dR\ A(R)\delta(E-\mathcal{H}) \tag{4.34}$$

where $t_i = i\Delta t$ for some time interval Δ along the microcanonical trajectory with fixed energy E and $\Omega(E)$ as the microcanonical partition function at that energy.

The last problem is due to the choice of the mass parameter, m'_I, in the PIMD scheme. To illustrate this problem we can see in Figure 4.2 how the size of the ring polymer decreases for heavier particles. Thus, one has to decrease the integration time to sample high frequencies. However, the choice of the mass parameter will affect the efficiency of the molecular dynamics sampling since it governs the rate at which the trajectories moves through the phase space. A large-mass trajectory will move relatively slow and will take a long time to sample the whole phase space. During the last decades some solutions to tackle this particular problem of the molecular implementation of path integral have been developed. For instance, one is the change of variable (Staging transformation) or the use of the normal modes (NM algorithm) to evolve the equation of motion of the ring polymer. Only the NM algorithm was implemented and used in the last chapter for numerical accuracy. These two solution will be discussed in detail in the following section.

4.5 The Staging transformation

If n is large, the springs are n-times stiffer and the potential n-times weaker. Then the spring dominates the dynamics and the system does not explore the rest of the phase space required to compute properly the static properties. Therefore, for a large n the harmonic modes have to be decoupled so that one could move all the modes in the same time scale. This can be done by the staging transformation [79], which was developed in analogy to the staging Monte Carlo method [80]. Each degree of freedom is transformed as

$$\begin{aligned} u_I^{(1)} &= R_I^{(1)}, \\ u_I^{(s)} &= R_I^{(s)} - R_I^{(s)*},\ s = 2, \ldots, n, \end{aligned} \tag{4.35}$$

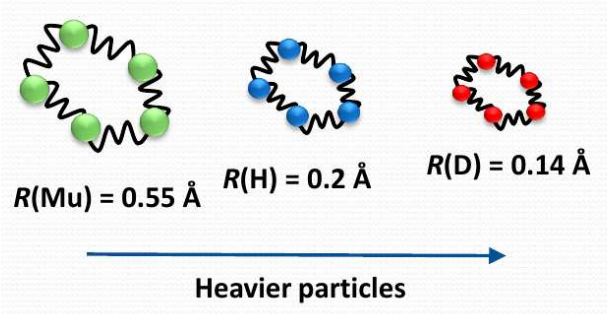

Figure 4.2: Path integral collapse due to the effect of heavier particles (with Muon(Mu), Hydrogen(H) and Deuterium(D)).

with

$$R_I^{(s)*} = \frac{(s-1)R_I^{(s+1)} + R_I^{(1)}}{s} \tag{4.36}$$

Such transformation can be inverted as

$$\begin{aligned} R_I^{(1)} &= u_I^{(1)}, \\ R_I^{(s)} &= u_I^{(s)} + \sum_{t=s}^{n} \frac{s-1}{t-1} u_I^{(t)}, \; s = 2, \ldots, n, \end{aligned} \tag{4.37}$$

Substituting into Eq. 4.25, the partition function results

$$\begin{aligned} \mathcal{Z}_n(\beta) &= \prod_{I=1}^{N} \mathcal{N} \int du_I^{(1)} \ldots du_I^{(n)} \int dP_I^{(1)} \ldots dP_I^{(n)} \times \\ & \times exp(-\beta \sum_{s=1}^{n} \{ \sum_{I=1}^{N} \frac{P_I^{(s)2}}{2m_I^{\prime(s)}} + \frac{1}{2} m_I^{(s)} \omega_n^2 (u_I^{(s)})^2 + \\ & + \frac{1}{n} V(R_I^{(s)}(\{u_I^{(s)}\}))\}), \end{aligned} \tag{4.38}$$

where the staging masses $m_I^{(s)}$ are defined as

$$\begin{aligned} m_I^{(1)} &= m_I \\ m_I^{(s)} &= \frac{s}{s-1} m_I, \; s = 2, \ldots, n. \end{aligned} \tag{4.39}$$

The harmonic oscillators are decoupled from each other in the staging variables. From Eq. 4.38 one can notice that the Hamiltonian in the staging variables is given by

$$\mathcal{H}_{stage} = \sum_{s=1}^{n} \{ \sum_{I=1}^{N} \frac{P_I^{(s)2}}{2m_I^{\prime(s)}} + \frac{1}{2} m_I^{(s)} \omega_n^2 (u_I^{(s)})^2 + \frac{1}{n} V(R_I^{(s)}(\{u_I^{(s)}\}))\}, \tag{4.40}$$

using this Hamiltonian, the path integral calculation is expected to sample its corresponding phase space much faster than the Hamiltonian (see Eq. 4.26) in the primitive variables. The Hamiltonian in Eq. 4.40 suggests that an optimal choice of the mass parameters is given by

$$\begin{aligned} m_I'^{(1)} &= m_I^{(1)} \\ \vdots &= \vdots \\ m_I'^{(s)} &= m_I^{(s)} \end{aligned} \tag{4.41}$$

Such choice will move the staging modes, $u_I^{(1)}, \ldots, u_I^{(n)}$ on the same time scale facilitating the sampling of all the modes during the MD run. Alternatively to the staging transformation and with the same efficiency was developed the normal modes algorithm.

4.6 The Normal Mode transformation

NM transformation has been extensively used in the field of polymer science to study the Rouse dynamics of unentangled short chains [81]. In our system, for $n \to \infty$, the simulation of ring polymer will tend to suffer of convergence problems as we described before. This is partly due to the time scale separation between the intermolecular and intramolecular potentials. While the former scales as $1/n$, n being the number of beads, the latter follows n. However, in the path integral approach, the internal modes of the ring polymer are mixed due to the interaction between beads of different polymer rings and this is the physical scenario that one wants to preserve in the simulation. The numerical implementation of the normal modes (see Appendix E) takes into account the mixing of the modes in a good approximation. This transformation has been used in this thesis to express the harmonic spring potential $V_{harm}(\{R_I^{(s)}\})$ from Eq. 4.26 as a sum of n uncoupled harmonic oscillators. The harmonic potential for a system of N particles can be written as

$$V_{harm}(\{R_I^{(s)}\}) = \sum_{I=1}^{N} V_I(\{R_I^{(s)}\}) \tag{4.42}$$

where the potential due to the Ith ring-polymer is

$$\begin{aligned} V_I(\{R_I^{(s)}\}) &= \sum_{s=1}^{n} \frac{1}{2} m_I w_n^2 (R_I^{(s)} - R_I^{(s+1)})^2 \\ &= \sum_{s=1}^{n} \frac{1}{2} m_I w_n^2 (R_I^{(s)2} - R_I^{(s)} R_I^{(s+1)} - \\ &\quad R_I^{(s+1)} R_I^{(s)} + R_I^{(s+1)2}). \end{aligned} \tag{4.43}$$

After rearranging the terms in the sum, it becomes

$$V_I(\{R_I^{(s)}\}) \;=\; \sum_{s=1}^{n} \frac{1}{2} m_I w_n^2 (2R_I^{(s)2} - R_I^{(s)} R_I^{(s+1)} - R_I^{(s)} R_I^{(s-1)}) \tag{4.44}$$

which in vector notation is

$$V_I(\{R_I^{(s)}\}) \;=\; \frac{1}{2} m_I w_n^2 \mathbf{R}_I^T \cdot \mathbf{A} \cdot \mathbf{R}_I \tag{4.45}$$

where

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & & & -1 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ -1 & & & -1 & 2 \end{pmatrix}$$

is independent of the particular value of I.

It can be noted that for the Ith ring polymer, the coordinates of different beads are coupled by the off-diagonal therm of the matrix $\mathbf{A}$. Thus, the uncoupling of the coordinates can be done by diagonalizing $\mathbf{A}$, i.e., this means, finding the matrices $\mathbf{a}$ and $\mathbf{C}$ such that

$$\mathbf{A} = \mathbf{C} \cdot \mathbf{a} \cdot \mathbf{C}^T, \tag{4.46}$$

where $\mathbf{a}$ is a diagonal matrix and $\mathbf{C}$ is an orthogonal matrix such that

$$\mathbf{C}^T \cdot \mathbf{C} = \mathbf{I}, \tag{4.47}$$

due to the symmetry of $\mathbf{A}$. Theses matrices can be found by exploiting the analytical form of $\mathbf{A}$ and writing the Eq. 4.46 in the following form:

$$-c_{k-a,s} + 2c_{k,s} - c_{k+1,s} = c_{k,s} a_s, \ \forall k, s = 1, \ldots, n. \tag{4.48}$$

The boundary condition $c_{n+1,s} = c_{1,s}$ and $c_{0,s} = c_{n,s}$ can be satisfied by the following independent solutions

$$c_{k,s} = N \cos(2\pi ks/n) \tag{4.49}$$

for $s = 1, \ldots, n/2$ and

$$c_{k,s} = N \sin(2\pi ks/n), \tag{4.50}$$

for $s = n/2 + 1, \ldots, n$ where N is a normalization constant. Combining Eq. 4.46 and Eq. 4.49 gives

$$0 = -\cos(2\pi(k-1)/n) + (2 - a_s)\cos(2\pi ks/n) - \cos(2\pi(k+1)s/n). \tag{4.51}$$

and using the identity:

$$\cos(x \pm y) = \cos(x)\cos(y) \pm \sin(x)\sin(y), \tag{4.52}$$

the expression above can be arranged and yields

$$a_s 4\sin^2(s\pi/n),\ \forall s = 1, \ldots, n. \tag{4.53}$$

This implies that the eigenvalue $a_{n/2+s'}$ is degenerate with $a_{n/2-s'}$ for $s' = 1, \ldots, (n/2-1)$. Thus, the normalization constants are

$$N \;=\; \begin{cases} \sqrt{1/n}, & \text{for } s = n/2 \text{ or } n \text{ (distinct eigenvalues)} \\ \sqrt{2/n}, & \text{for all other } s \text{ (degenerate eigenvalues).} \end{cases} \tag{4.54}$$

With the diagonalization of $\mathbf{A}$, one can insert Eq. 4.46 into Eq. 4.45 and this yields

$$V_I(\{Q_I^{(s)}\}) = \frac{1}{2} m_I w_n^2 \mathbf{Q}_I^T \cdot \mathbf{a} \cdot \mathbf{Q}_I, \tag{4.55}$$

where

$$\mathbf{Q} = \mathbf{C}^T \cdot \mathbf{R}, \tag{4.56}$$

and it defines the transformation to the ring polymer in the Normal Modes coordinates. And now using the property of $\mathbf{a}$ being diagonal, the Eq. 4.44 becomes

$$V_I(\{Q_I^{(s)}\}) = \frac{1}{2} m_I \omega_n^2 \sum_{s=1}^{n} a_s Q_s^2 = \sum_{s=1}^{n} \frac{1}{2} m_I \Omega_s^2 Q_s^2 \tag{4.57}$$

which is the potential for a collection of n uncoupled harmonic oscillators with normal-mode frequencies equal to

$$\Omega_s = \omega_n \sqrt{a_s} = 2\omega_n \sin(s\pi/n). \tag{4.58}$$

Similarly as in Eq. 4.38 one can notice that the Hamiltonian in the normal modes (NM) is given by

$$\mathcal{H}_{NM} = \sum_{s=1}^{n} \{ \sum_{I=1}^{N} \frac{P_I^{(s)2}}{2m_I^{'(s)}} + \frac{1}{2} m_I \Omega_s^2 (Q_I^{(s)})^2 + \frac{1}{n} V(R_I^{(s)}(\{Q_I^{(s)}\})) \}. \tag{4.59}$$

For $s = n$ in Eq. 4.53, one obtains $a_n = 0$. Thus, the Eq. 4.57 defines a potential for a zero-frequency. We also define the ring polymer centroid as

$$R_c = \frac{1}{n} \sum_{s=1}^{n} R_I^{(s)}. \tag{4.60}$$

So that, one defines the nth normal mode with elements $c_{k,n} = \frac{1}{\sqrt{n}}$ independent of the value of k and therefore Eq. 4.56 yields

$$Q_n = \frac{1}{\sqrt{n}} \sum_{s=1}^{n} R_I^{(s)} = \sqrt{n} R_c \tag{4.61}$$

Through this expression, the normal-modes transformation introduces the centroid variable. In the Appendix E is described in detail the numerical implementation of the normal mode transformation .

4.7 Similarities between the path integral formalism and the statistics of ring polymers

4.7.1 The free ring polymer

Let us start the comparison between both approaches by studying the behaviour of a single ring polymer in the absence of an intermolecular potential between beads of different ring polymers (see Eq. 4.21). This case has an analytical solution and will illustrate the similarities. Due to the simplicity the distribution of beads is given by,

$$e^{-\beta m \omega_n^2 (R_i - R_{i+1})^2/2}, \tag{4.62}$$

and from which the average distance between bead is,

$$\langle (R_i - R_{i+1})^2 \rangle^{1/2} = \sqrt{\frac{1}{\beta m \omega_n^2}} = \frac{\beta \hbar^2}{mn}. \tag{4.63}$$

Thus, as the number of beads, mass of the particle or the temperature is increased the average distance between the bead decreases (see Fig. 4.3).

Another important property which can be quantified is the radius of gyration, r_G, which describes the spread of an individual bead from the centroid (center of mass) of the ring polymer. In one dimension can be defined as

$$x_G^2 = \frac{1}{n} \sum_{k=1}^{n} |x_k - x_c|^2 \tag{4.64}$$

where x_c is the position of the ring centroid,

$$x_c = \frac{1}{n} \sum_{k=1}^{n} x_k \tag{4.65}$$

The thermal average of the radius of gyration can be calculated exactly for a free ring polymer as follows

$$\langle x_G^2 \rangle = \frac{1}{n} \sum_{k=1}^{n} \langle (x_k - x_c)^2 \rangle = \frac{1}{n} \sum_{k=1}^{n} (\langle x_k^2 \rangle - \langle x_c^2 \rangle) \tag{4.66}$$

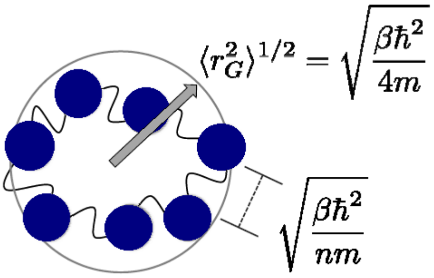
$\langle r_G^2\rangle^{1/2} = \sqrt{\frac{\beta\hbar^2}{4m}}$
$\sqrt{\frac{\beta\hbar^2}{nm}}$

Figure 4.3: Length scale of the free ring polymer in three dimensions. It is shown the thermal average of the root mean square radius of gyration and root mean square bond length of the ring polymer.

For a gaussian chain the radius of gyration scales as $r_G^2 \sim n$ [81], however, for a ring polymer, the path intgral approach predicts a constant value in the limit of $n \to \infty$ (see below). This is due to the non-trivial n-dependence of the spring constant in the path integral approach Eq. 4.21. In the case of a free ring polymer, the thermal average is given by

$$\langle \ldots \rangle = \frac{\int d^n \mathbf{x} e^{-\beta \sum_{k=1}^{n}[m\omega_n^2 (x_k - x_{k+1})^2/2]}(\ldots)}{\int d^n \mathbf{x} e^{-\beta \sum_{k=1}^{n}[m\omega_n^2 (x_k - x_{k+1})^2/2]}} \tag{4.67}$$

where $\omega_n = \sqrt{n}/\beta\hbar$. The above expression can be analytically evaluated by transforming the coordinate system to the normal modes coordinates of the ring polymer (i.e. $\mathbf{Q}_l = (\tilde{x}_l, \tilde{y}_l, \tilde{z}_l), \forall l = 1, \ldots, n$). The main result that we used is that the free ring polymer potential energy expressed in terms of bead coordinates,

$$V(\mathbf{x}) = \sum_{k=1}^{n} \frac{m\omega_n^2}{2}(x_k - x_{k+1})^2, \tag{4.68}$$

can be transformed using the orthogonal transformation $\mathbf{C}$ as follows:

$$\tilde{x}_l = \sum_{k=1}^{n} C_{l,k} x_k \tag{4.69}$$

to a set of normal modes coordinates $\tilde{x}_l$.

In this new set of coordinates the ring polymer potential energy can be expressed as a set of uncoupled harmonic oscillators

$$V(\tilde{x}) = \sum_{l=1}^{n} \frac{1}{2} m\Omega_l^2 \tilde{x}_l^2 \tag{4.70}$$

where Ω_l are the frequencies of the free ring polymer,

$$\Omega_l = 2\omega_n \sin(l\pi/n) \tag{4.71}$$

and n-th normal modes frequency is given by $\Omega_n = 0$ and is related to the centroid of the ring polymer by

$$\tilde{x}_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} x_k = \sqrt{n} x_c. \tag{4.72}$$

Transforming to the normal modes coordinates and using

$$\sum_{k=1}^{n} x_k^2 = \sum_{l=1}^{n} \tilde{x}_l^2, \tag{4.73}$$

which follows directly from the orthogonality of the transformation, the average of radius of gyration can be expressed as

$$\langle x_G^2 \rangle = \frac{1}{n} \sum_{l=1}^{n} (\langle \tilde{x}_l^2 \rangle - \langle (\tilde{x}_n/\sqrt{n})^2 \rangle) = \frac{1}{n} \sum_{l=1}^{n-1} \langle \tilde{x}_l^2 \rangle. \tag{4.74}$$

The expectation value of $\langle \tilde{x}_l^2 \rangle$ is therefore

$$\langle \tilde{x}_l^2 \rangle = \frac{\int d^n \tilde{x}\ e^{-\beta \sum_{l=1}^{n} [m \Omega_l^2 \tilde{x}_l^2/2]} (\tilde{x}_l^2)}{\int d^n \tilde{x}\ e^{-\beta \sum_{l=1}^{n} [m \Omega_l^2 \tilde{x}_l^2/2]}}, \tag{4.75}$$

which is a typical Gaussian integral, that can be easily evaluated and gives

$$\langle \tilde{x}_l^2 \rangle = \frac{1}{\beta m \Omega_l^2} = \frac{\beta \hbar^2}{4nm \sin^2(l\pi/n)}. \tag{4.76}$$

Finally, the square radius of gyration is

$$\langle x_G^2 \rangle = \frac{\beta \hbar^2}{4mn} \sum_{l=1}^{n-1} \frac{1}{\sin(\pi/n)}. \tag{4.77}$$

Using the identity,

$$\sum_{l=1}^{n} \frac{1}{\sin^2(l\pi/n)} = \frac{1}{3}(n^2 - 1), \tag{4.78}$$

the square radius of gyration in one dimension is,

$$\langle x_G^2 \rangle = \frac{\beta \hbar^2}{12m} \left(1 - \frac{1}{n^2}\right). \tag{4.79}$$

So in three dimension we have,

$$\langle r_G^2 \rangle = \langle x_G^2 + y_G^2 + z_G^2 \rangle = 3 \langle x_G^2 \rangle = \frac{\beta \hbar^2}{4m} \left(1 - \frac{1}{n^2}\right), \tag{4.80}$$

which is consequence of the isotropy of the space. As previously mentioned, the radius of gyration approaches a constant value in the limit of $n \to \infty$. This is an important consequence of the path integral formalism with respect to the classical ring polymers. Fig. 4.3 depicts the radius of gyration in the path integral approach.

4.7.2 Rouse theory for ring polymers

In this section we reproduce first some important results from the well-known Rouse theory for a "classical" ring polymer and later we provide a comparison with the path integral approach. Such a theory became extremly useful for the early understanding of the dynamics of short (non-entangled) polymer chains. Here we describe a bead-spring model of a classical ring polymer under the Rouse theory. The potential between beads is given by

$$U = \frac{1}{2} k \sum_{i=1}^{n} (r_i - r_{i-1})^2 \tag{4.81}$$

where $k = 3k_B T/b^2$ is the spring constant. It is important to emphasize the differences between the potential energy for a ring polymer under the classical and the path integral approaches. In the former, the spring constant depends on the temperature the effective bond length. In the latter, it is also proportional to the temperature and additionally to the number of beads (Trotter number). In the Rouse model each monomer is subjected to a Brownian motion. Thus, one could assume that each bead will experience a drag force proportional to the velocitiy and random kicks which rapidly decorrelates in time.

The position of a single bead is given by,

$$\xi \frac{dr_i}{dt} = -\frac{\partial U}{\partial r_i} + \Gamma_i, \text{ for } i = 1, \ldots, n \tag{4.82}$$

here ξ is the friction coeffient of a bead and Γ_i is the noise acting on the bead with the following properties:

$$\begin{aligned} \langle \Gamma_i \rangle &= 0 \\ \langle \Gamma_i^\alpha(t) \Gamma_{i'}^\beta(t') \rangle &= 2k_B T \xi \delta(t - t') \delta_{ii'} \delta_{\alpha\beta}, \end{aligned} \tag{4.83}$$

where $\alpha, \beta = 1, 2, 3$ are cartesian indexes. One could see from Eq. 4.81 and 4.82 that the equations of motion (EOM) will be coupled due to the form of the potential. A simple idea to overcome this difficulty is to introduce normalized coordinates which decompose the motion into independent modes (i.e. "Rouse modes"). Following the treatment of the Rouse model for ring polymers [82], to solve Eq. 4.82 in the continuos limit (i.e $n \to \infty$), we introduce the normal coordinates as follows:

$$Q_l(t) = \frac{1}{n} \int_0^n dn \; r_i(t) \cos\left(\frac{l\pi i}{n}\right) \text{ for } l = 0, 1, 2, 3, \ldots \tag{4.84}$$

and the inverse transform is

$$r_i(t) = Q_0 + 2 \sum_{l=1}^{\infty} Q_l \cos\left(\frac{l\pi i}{n}\right) \text{ for } l = 0, 1, 2, 3, \ldots. \tag{4.85}$$

Now, applying the boundary condition for the case of a ring polymer, $r_0 = r_n$ and $\frac{\partial r_i}{\partial i}|_{i=0} = \frac{\partial r_i}{\partial i}|_{i=n}$, we find that all odd modes vanish. This means that

$$\begin{aligned} r_i(t) &= Q_0 + 2\sum_{l,\text{even}}^{\infty} Q_l \cos\left(\frac{l\pi i}{n}\right) \\ Q_l(t) &= \frac{1}{n}\int_0^n dn\ r_i(t)\cos\left(\frac{l\pi i}{n}\right) \quad \text{for } l = 0, 2, 4, \ldots \end{aligned} \tag{4.86}$$

where the summation is evaluated for even modes (i.e. $l, \text{even} = 2, 4, 6, \ldots$).

The Eq. 4.82 in the continuous limit can be written as:

$$\xi_l \frac{\partial Q_l}{\partial t} = -k_l Q_l + f_l \tag{4.87}$$

and using Eq. 4.83, we also find that

$$\langle Q_l^{\alpha}(t) Q_{l'}^{\beta}(0)\rangle = \delta_{\alpha\beta}\delta_{ll'}\frac{k_B T}{\kappa_l}e^{-t/\tau_l}, \text{ for } l = 2, 4, 6, \ldots \tag{4.88}$$

where

$$\begin{aligned} \kappa_l &= \frac{2\pi^2 l^2}{n}k = \frac{6\pi^2 k_B T l^2}{b^2 n} \\ \tau_l &= \frac{2n\xi}{\kappa_l} = \frac{\xi n^2 b^2}{3\pi^2 k_B T l^2}, \text{ for } l = 2, 4, 6, \ldots \end{aligned} \tag{4.89}$$

and τ_l is known as the relaxation time of the Rouse modes. The derivation of Eq. 4.88 and Eq. 4.89 is fully transferable to the case of the ring polymer in the path integral (PI) approach and when the proper spring constant, $k = mn/(\beta\hbar)^2$, is used we have that,

$$\begin{aligned} \kappa_l^{\text{PI}} &= = \frac{2\pi^2 (k_B T)^2 m l^2}{\hbar^2} \\ \tau_l^{\text{PI}} &= \frac{\xi n \hbar^2}{\pi^2 (k_B T)^2 m l^2}, \text{ for } l = 2, 4, 6, \ldots. \end{aligned} \tag{4.90}$$

Comparing the relaxation time predicted by the Rouse theory between Eq. 4.89 and Eq. 4.90, one could see that modes of the polymer rings in the path integral approach relax faster than classical ring polymers.

4.7.3 Application of the Rouse theory for the *para*-hydrogen liquid

A test case is studied under the Rouse theory for polymer rings. More details about the forced field between *para*-hydrogen molecules and the application of our novel approach will be presented later in the chapter 6. Here we report briefly the analysis of the Rouse modes using path integral molecular dynamic

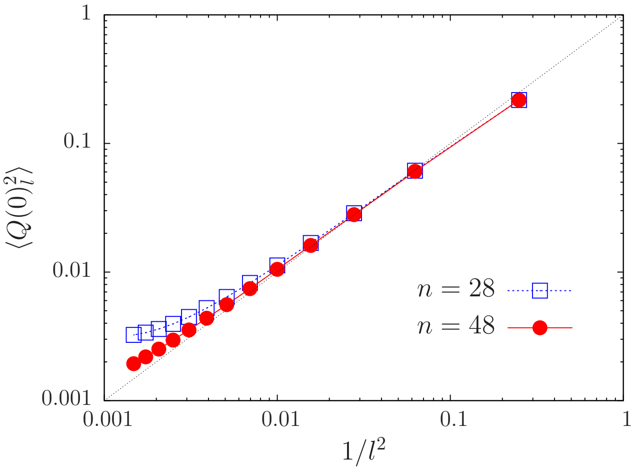

Figure 4.4: The square amplitude of the Rouse normal modes $\langle Q_l^2(0) \rangle$ for the path integral representation of para-hydrogen molecules as a function of $1/l^2$. Here is depicted for the ring polymer with lentgh n=28 and n=48 beads at T=25 K.

(PIMD) simulations with normal modes (see section 4.6 and Appendix E). We perform simulations for two different Trotter number, namely $n = 28$ and $n = 48$ at T=25 K. The trajectories accumulated in the simulation serve as a starting point for testing the new insight about the Rouse theory for the ring polymer representation of the *para*-hydrogen molecules.

There are several test that one can think of: the first is to test the Eq. 4.88 that the mean square amplitude $\langle Q_l^2(0) \rangle$ of the l-th Rouse modes scales with l as $1/l^2$. Results obtained for the two ring polymers are shown in the Fig. 4.4. For a given number of beads n, the Rouse scaling is seen to be followed only for the lower l modes, for higher l modes there are significant deviation from the Rouse theory, specially the shorter one.

A second test of the Rouse mode for the ring polymer refers to the time decay of the normalized autocorrelation function $\langle Q_l(t) Q_l(0) \rangle / \langle Q_l^2(0) \rangle$. According to Eq. 4.88, a log-linear plot of $\langle Q_l(t) Q_l(0) \rangle / \langle Q_l^2(0) \rangle$ versus $t/(1/l^2)$ should yield a straight line. Numerical results for the relation times are shown in Fig. 4.5. If the ring polymer behaves identically as a Rouse chain, all the curve should have collapsed on a straight line. This is not the case in the data reported in the figure and significant deviation from the ideal Rouse behaviour are observed, in special for higher modes ($l > 2$). The same behaviour dominates the larger ring polymer (not reported here).

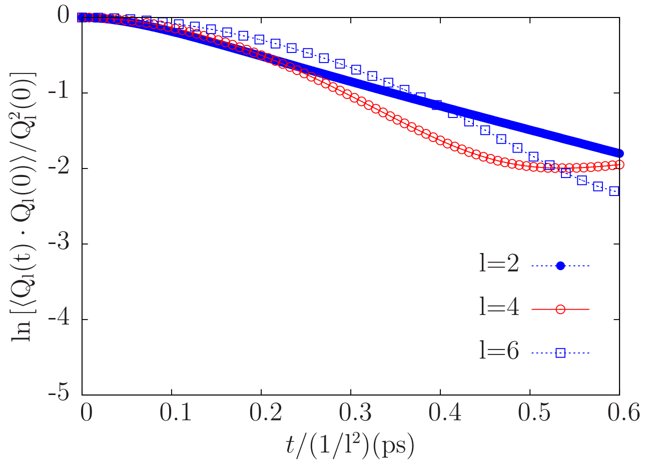

Figure 4.5: Time autocorrelation function of the first normal modes for ring polymer representation of para-hydrogen molecules with n=28 (beads) plotted against $t/(1/l^2)$ in a log-linear plot.

Chapter 5

An adaptive classical-path integral scheme in Molecular Simulation

In Chapter III, we described the classical adaptive resolution simulations where the simulation domain is subdivided in regions of different molecular resolution and particles can easily diffuse between these two resolutions. However, while the adaptive process of changing resolution on the fly can be described within a reasonable simulation accuracy according to the basic principles of classical dynamics and thermodynamics, the same cannot be said when quantum mechanics becomes relevant. The proper coupling of quantum and classical mechanics is known to be a non trivial (and open) problem [83] and hence the adaptive character adds up as a further difficulty [84]. Practical methods [85–87] that couple the two regimes, in general, do not take into account the "conceptual" discontinuity of going from a probabilistic (quantum) to a deterministic (classical) approach (and vice versa), and usually base their validity on empirical and numerical criterion. In this chapter, we present a new scheme [88] where the coupling between classical and quantum regime can be achieved in a smooth and consistent way. This chapter is organized as follows: First a section about the importance of the quantum effects in the matter, then a section is dedicated to the basic idea of quantum-classical coupling and finally a test of our approach PIMD/CG within AdResS framework in two model systems, namely the monoatomic liquid and molecular liquid.

5.1 Quantum description in soft matter

Generally, the extent to which the quantum nature of the systems matters is a crucial aspect of modeling any soft matter system (e.g. proteins, liquid, polymers, etc). This is partly due to the variety of processes whose quantum character influence several other scales. It is undeniable that a full quantum description of a system (by Density Functional Theory (DFT) methods) is often computationally too expensive. Therefore, typically the modeling of many soft matter systems starts to incorporate partial quantum effects into the so-called classical force fields (e.g. CHARMM [89], GROMOS [90] and several others). However, there are many interesting problems where such approximation is not accurate and detailed quantum description is needed in certain regions of the space, for example, the inherently quantum nature of a chemical reaction in biological processes [91] and the quantum nature of the nuclei of light particles as hydrogens in important molecules as water [92] at room temperature require a proper quantum description, to name a couple.

The quantum character of particles becomes particularly relevant for the regime of low temperatures and light particles. In such conditions, the quantum nature of matter play important roles in modifying the structure and dynamics of the entire system. Soft matter systems are typically around room temperature and their behavior is generally determined by the thermal fluctuations, which are of the order of a few k_BT. However, as we illustrate in Figure 5.1, some biological systems may need to incorporate quantum details to completely describe the structure and dynamics. In this figure we illustrate a large molecule solvated in a model system of tetrahedral molecules, in part (a) of Figure 5.1 the full quantum mechanical description of the system with path integral description is shown. This makes the molecule more flexible and new conformations can be explored in presence of the solvent. For example, the red circle indicates a typical region where the atoms become delocalized and thus they induce conformational changes in the molecule. Let us assume that such conformations are only observed in presence of the quantum character of the particles and therefore, a conventional simulation with the path integral approach will be computationally expensive, due to the system size. In Figure 5.1 (b) we depict a solution which combines the advantage of a systematic structural coarse graining (see chapter I) to reduce the number of degrees of freedom in a region of no interest and the adaptive resolution scheme (see chapter III) to allow the change of degrees of freedom on the fly. The preceding description maintains particle fluctuations between all the regions. The basic idea of how one can implement such approach is presented in the next section.

5.2 The basic idea of the quantum-classical adaptivity

As discussed in the chapter III, the AdResS method is numerically robust and its theoretical background has been well-established. In our case, the AdResS method becomes an important tool for the adaptivity of the quantum-classical system. More specifically, for the adaptive process the path integral approach of atoms has far-reaching consequences, because it translates the quantum-classical coupling into the coupling of two effective classical regions characterized by a

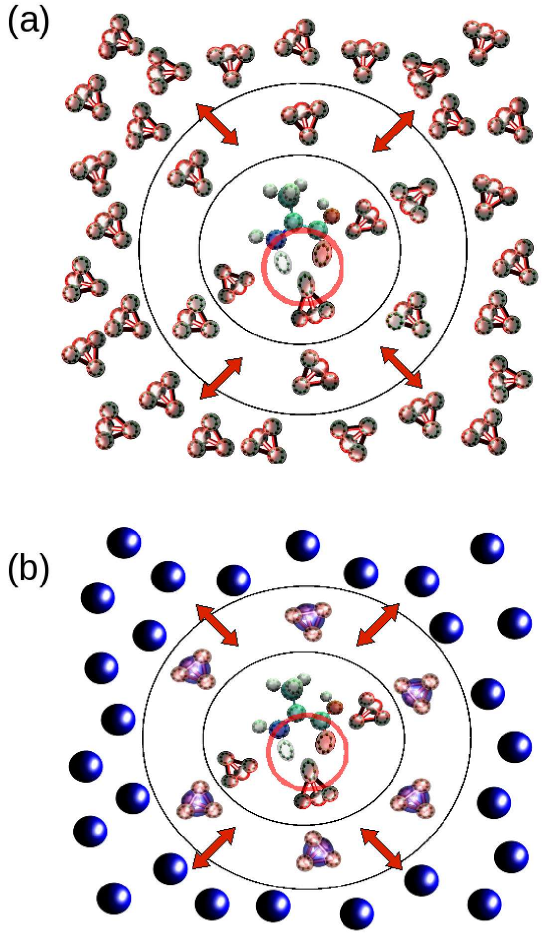

Figure 5.1: A pictorial representation of a large molecule solvated in a liquid of tetrahedral molecules. In part (a) the quantum mechanics description of the whole system generates conformational changes that are not observed in classical MD simulations. In (b) the PIMD/CG approach is depicted, the space is partitioned in a central region described with quantum mechanics, molecules that are far of the central region can be replaced by CG spheres and in between a transition region with hybrid particles.

different number of (as a matter of fact) "classical" degrees of freedom; thus the whole machinery of classical adaptive methods would apply straightforwardly.

In this chapter, we show that this indeed is the case. It must be noticed that our PIMD/CG approach in AdResS (see next section) aims to calculate static equilibrium properties within a quantum mechanical description and not for the true quantum dynamics. Practical adaptive methods where electrons are considered are already available [85–87], however the nuclei are in these cases classical. In the next section, we test our approach systematically in two systems, namely the monoatomic liquid and molecular (tetrahedral) liquid. As in the classical AdResS a coarse graining procedure is used to obtain an effective one site potential which is coupled to the explicit force field using the equation of motion of AdResS. Similarly here, one obtains an effective potential from the path integral representation and then performs PIMD/CG simulation within the AdResS framework.

5.3 PIMD/CG approach in AdResS

The purpose of our hybrid scheme is to use a path integral (PI) description of the molecules in a certain region of interest, instead of using PI representation (i.e. ring polymers) throughout the simulation domain. We use a reduced number of path integral molecules in equilibrium with an effective coarse grained (CG) region (classical particles), where the effective potential is derived from a full path integral simulation. The main advantage of the hybrid scheme is that we do not need to represent all the particles by the path integral formalism, which is computationally expensive for large systems. Details of the procedure used to characterize the CG particles are given in the next sections. Let us now consider the total number of particles, N, in the simulation box to be fixed, which represents the total number of atoms or molecules. Our results in the next sections show that a fraction of the total number of particles can be represented by ring polymers, which remain in thermodynamic equilibrium with the classical CG representation and at the same time preserve quantum properties, for example the distributions of delocalized particles in the quantum region.

5.3.1 Effective potential derived from a path integral representation

The force field needed for the path integral simulation has to be parametrized without the quantum effects, which are important at a given thermodynamic condition. Otherwise, the system under study will overcount the quantum effects due to the additional path integral approach. Such a problem has been addressed recently for the case of path integral simulation of a flexible water model [92]. In our study, we use a classical force field that does not include quantum effects. Then the effective coarse-grained model is derived from a full path integral representation. To obtain an effective coarse grained potential from a path integral representation of atoms/molecules we have used the iterative Boltzmann inversion procedure available in the literature [16]. The main idea is to calculate the non-bonded potential $V^{\text{eff}}(R)$ between the center of masses ($R = R_{\text{CM}}$) of the particles from their path integral representation at given temperature. For example the CG procedure for the monoatomic liquid consists of replacing the path integral representation of an atoms (i.e. ring polymers) by an effective particle (CG bead), which interacts through the potential $V^{\text{eff}}(R)$. Thus, such CG potential takes into account (in average) the quantum corrections to the classical CG potential as shown below,

$$V^{\text{eff}}(R) = V^{\text{cl}}(R) + \Delta U^{\text{QM}}(R), \tag{5.1}$$

where V^{cl} is the classical CG potential and ΔU^{QM} is the corresponding correction due to quantum effects. Therefore, as the temperature T or the mass of the particle decreases the quantum corrections to the V^{eff} become more important. To guarantee that our hybrid scheme reproduces the same thermodynamic state point one has to correct the shift in the pressure produced by the artifact of the iterative procedure (see section 1.3.2).

5.4 Results and Discussion

In this section, we present the results of our PIMD/CG approach in AdResS for the two model systems studied in a regime of low temperatures, which corresponds to the extreme thermodynamic condition where the quantum description provided by path integral becomes relevant. In general, for soft matter applications the temperature is usually high and thus our approach, if it works at low temperature, should then work even better. As we have seen before the quantum character of the path integral representation (see chapter IV) is characterized by the number “n” of beads (Trotter number) at given temperature. In our simulation, we have fixed the Trotter number and change the thermal energy from a high temperature (classical) to a lower temperature regime where quantum effects become important. A procedure to achieve such conditions is to decrease the temperature T. However, in the path integral formalism one can also modify the spring constant (due to its temperature dependence, $k = \frac{mn(k_B T)^2}{\hbar^2}$) to resemble such conditions. Thus, one can mimic temperature effects by tuning the spring constant. This plays the same role as the temperature and helps us to explore a broad picture of delocalization. This procedure is used here to test our PIMD/CG approach in AdResS from a moderate to a strong quantum regime. Note that in real physical systems each temperature defines one fixed value of k. In the first part of this section we show the results for a simple model of a repulsive monoatomic liquid and later we test the method in a rather more complicated molecular system.

5.4.1 Case I: The monoatomic liquid

To begin with our analysis, we start to simulate a generic simple model of a monoatomic system within its quantum limit. The idea here is to test the switching of degrees of freedom from the path integral representation in the "primitive coordinates" to an effective CG one site representation. The Figure 5.2 illustrates the process of changing on the fly the representation of a single quantum atom in its path integral representation to the effective classical model. The process that we address is the free passage of atoms from a path integral region to a coarser one and vice versa.

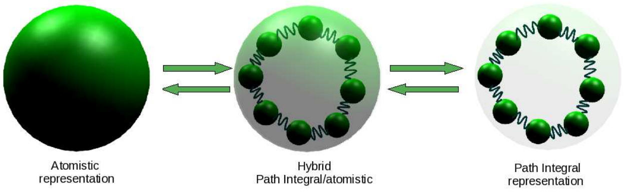

Figure 5.2: The on-the-fly interchange among the path integral (PI) and coarse-grained (CG) representation. Here we depict the case of a quantum particle in the high resolution (Right side) given by the PI representation which due to the isomorphism of the quantum partition function becomes a polymer ring.

To account for several degrees of "quantumness" (i.e. delocalization) we perform a systematic study decreasing the temperature associated with each quantum regime as we disscused at the beginning of this section. For the purpose of testing, we define the reference temperature, $T^* = 1$ in the Lennard-Jones units (with $\epsilon = 1$ and $\sigma = 1$ as the unit of energy and length respectively), to be the temperature associated with classical regime. We perform simulations at different temperatures $T^*/\sqrt{10}$, $T^*/5$, $T^*/\sqrt{50}$ and $T^*/10$. We show the effect of decreasing the temperature in Table 5.1, where we compare the radius of gyration for the free ring polymer $r_g(\text{free})$ and the one obtained in our simulations of full path integral for each respective temperature. The contraction observed in the r_g is due to the presence of the intermolecular interactions in the simulation.

The force field used to describe the classical condition is a repulsive Morse potential parameterized as follows:

$$V(r^*) = \gamma^*\{1 - exp(-\alpha^*(r^* - r_0^*)\}^2 \tag{5.2}$$

where $\gamma^* = \gamma/\epsilon = 0.105$, $\alpha^* = \alpha\sigma = 2.4$ and $r_0^* = r_0/\sigma = 2.31$, and cutoff at r_0^*. In this case σ and ϵ represent the length and energy units.
The next section focus on the technical details about the calculation of the effective potential using the iterative Boltzmann inversion method.

Table 5.1: Radius of gyration obtained from PIMD simulation and its theoretical value in the approach of free ring polymer, i.e. $r_g^2(\text{free}) = \frac{\beta\hbar^2}{4m}(1 - \frac{1}{n^2})$.

$Temperature$	r_g	r_g(free)
$T^*/10$	0.59 ± 0.01	0.70
$T^*/\sqrt{50}$	0.42 ± 0.01	0.49
$T^*/5$	0.29 ± 0.02	0.32
$T^*/\sqrt{10}$	0.20 ± 0.01	0.22
T^*	0.05 ± 0.01	0.07

Determination of the effective potential

We calculate the effective non-bonded pair potential $V^{\text{eff}}(R)$ between the CM of the polymer rings using the Iterative Boltzmann Inversion (IBI) method [16] and subsequent pressure correction for a given degree of delocalization. As mentioned earlier the "quantumness" becomes more evident at lower temperatures. For each temperature a numerical set of effective potentials is obtained, which reproduces the center-of-mass radial distribution function (RDF) and total pressure for a given thermodynamic state point. This process was repeated for two number densities which correspond to the medium density liquid with $\rho^* = 0.1$ and the high density liquid with $\rho^* = 0.175$.

Figure 5.3 (a) shows the effective potentials for the very quantum case equal to $T^*/10$ and for $T^*/5$ where the "quantumness" is negligible at density $\rho^* = 0.1$. Part (b) shows the distribution from the beads to the center-of-mass of each ring polymer in the path integral representation which provides a signal of the delocalization of particles for lower temperatures. We infer from both pictures that for this particular system as the temperature becomes lower, the effective coarse grained potential is more softer. Thus, atoms can occupy more space in the statistical sense (i.e quantum delocalization) due to quantum aspects of the matter present at given temperature.

Statistical properties

In this part, we present the numerical results of our adaptive simulation of PIMD/CG in AdResS for the monoatomic liquid. It is important to emphasize that the atoms diffuse freely in time traveling from the path integral region to the coarser region and vice versa. Thus, one must check the thermodynamic consistency of our results based on the criteria introduced in the chapter of the AdResS (see Chapter III). As we described for the classical system the path integral particles interact with CG ones by the interpolation force scheme [93]. Our simulations are tested by comparing the computed statistical properties of the PIMD/CG in AdResS with the corresponding properties in the full path integral (reference system). As we show below, the results are in good agreement up to approximately the same error known from the classical adaptive simulation.

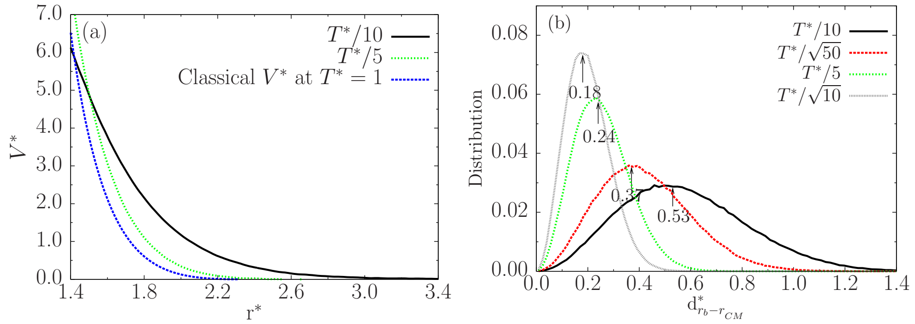

Figure 5.3: (a) The effective potential V^* between center of masses (CMs) of the polymer rings for $T^*/10$ (more quantum regime) and for $T^*/5$ (classical regime) with $\rho^* = 0.1$. As the temperature decreases the effective CG potential, V^*, approaches to the classical force field (in blue). (b) Shows the distribution distance ($d^*_{r_b - r_{\mathrm{CM}}}$) from each single beads to its corresponding CM of the ring polymers, for lower temperatures the distribution spreads more in the space.

We start to compare compare the CG level of description in our adaptive simulation. In part (a) of Figures 5.4, 5.5 and 5.6 is compared the center-of-mass radial distribution function (RDF) calculated in all the space in the adaptive scheme with the full explicit path integral simulation. We note that for $T > T^*/5$ the effect of the delocalization implicitly determined by the temperature ceases as soon as the temperature approaches to the reference (classical) value, $T^* = 1$.

Figure 5.7 shows the normalized density profile along the x direction for the same set of temperatures. Once again, for $T \geq T^*/5$ the drop of density disappears indicating that both representations: path integral (PI) and coarse grained (CG) are thermodynamically equivalent. To guarantee the free diffusion of particles and their subsequent transformation from one representation to the other one, we plot the center-of-mass diffusion profile for two slab of particles, one is localized on the quantum region and the other in CG one in Figure 5.8. This tells us that the atoms can freely diffusive and we can see how the initial set of particles spread as the time goes.

The second level of comparison is due to the local information accessible only in the path integral region. To show the consistency for this level, we plot the partial bead-bead RDF which is calculated only in path integral region in our adaptive simulation. This is compared with RDFs calculated in the same subregion but from a full explicit path integral simulation (see part (b) of Figures 5.4, 5.5 and 5.6). For the set of temperatures the agreement is highly satisfactory. The coexistence of path integral, hybrid and CG particles can be observed in the Table 5.2. From this table, we note that approximately 1/3 (due to the size of the hybrid zone) of the total number of atoms are represented as ring polymers and they are in good thermodynamic equilibrium with 1/3 of coarse grained particles. The remaining particles are represented by hybrids.

It is important to note that the thermodynamic equilibrium is maintained for each region. This means the pressure P, temperature T and average density are kept in equilibrium. Same results are obtained for the density $\rho^* = 0.175$, not reported here.

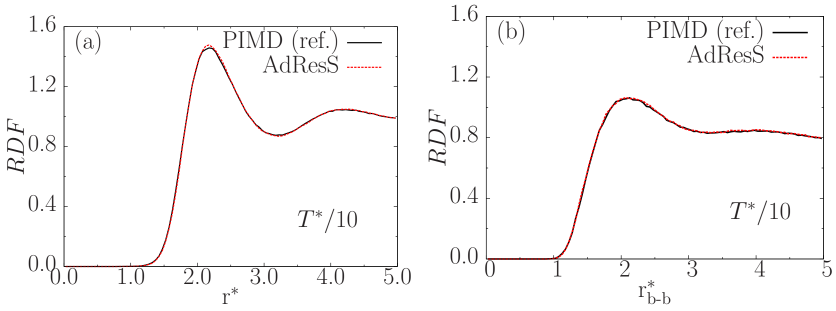

Figure 5.4: (a)Center-of-mass radial distribution function (RDF) evaluated in the whole box. The dashed lines (in red) indicates the result in AdResS which is compared with the RDF obtained from a full path integral simulation (reference system) for $T^*/10$, $\rho^* = 0.1$. (b) Shows the bead-bead radial distribution function calculate in the path integral region. This is compared with the RDF from the reference system.

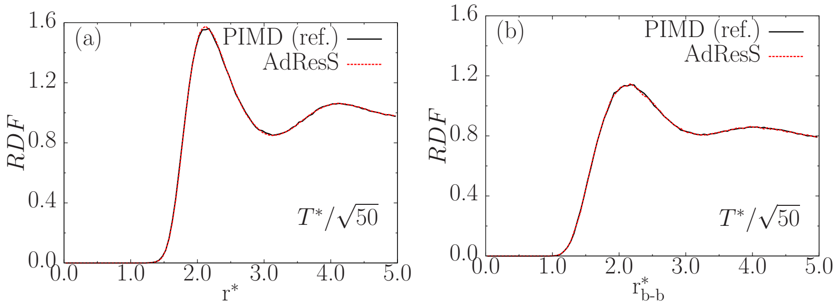

Figure 5.5: As for Figure 5.4, but in this case for $T^*/\sqrt{50}$. The same agreement for the global center of mass and local bead-bead radial distribution functions.

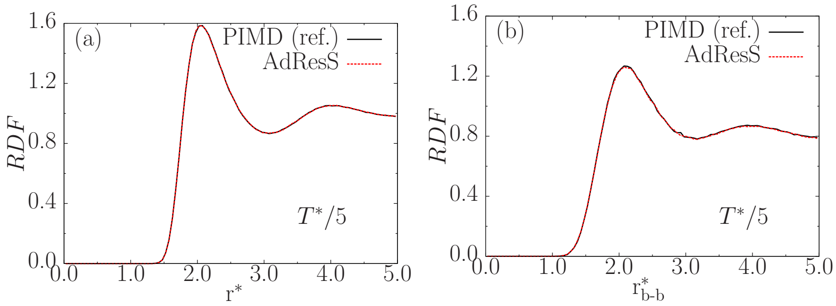

Figure 5.6: As for Figure 5.4, but this time for $T^*/5$. The same agreement as before is reported in this figure.

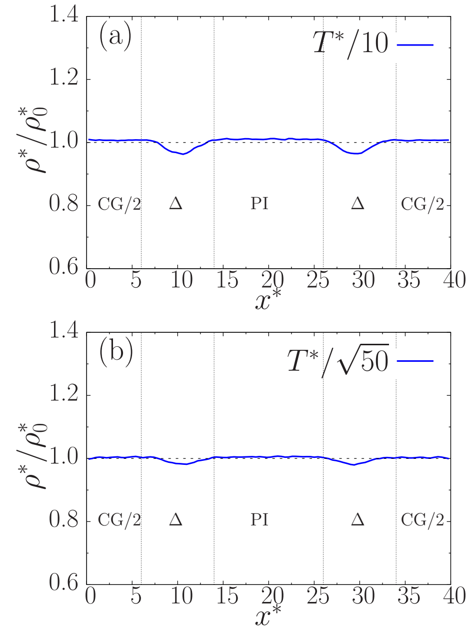

Figure 5.7: Normalized density profile in the x direction for the monoatomic liquid in AdResS. In all the graphics the interface region layer is given by 8σ. The vertical lines denote the boundaries between the path integral (PI), coarse grained (CG) and hybrid (Δ) regions of the system. In figure (a) is showed the case of lowest temperature studied $T^*/10$, in (b) for $T^*/\sqrt{50}$ and (c) for $T^*/5$. From all of them, one observes how the drop in density for high temperatures is suppressed by the thermodynamic similarity between the PI and CG model.

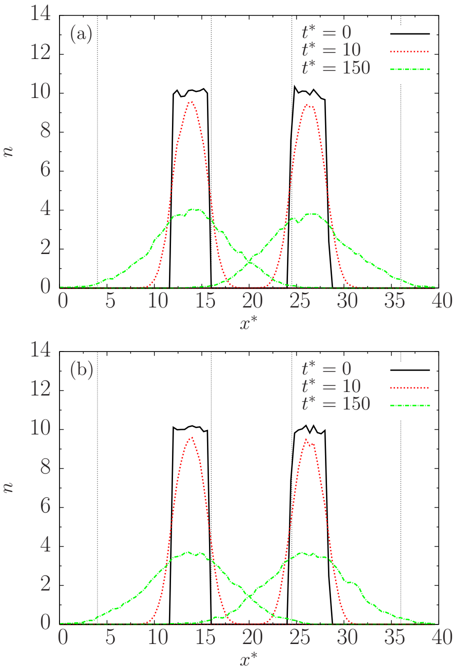

Figure 5.8: Time evolution of diffusion profile for particles that are initially, at time $t^* = 0$, localized at two neighboring slabs of the midinterface layer with 8σ interface layer width (n is the number of atoms with the CM position at a given coordinate x^*). The width of the slab is $L_x/10$. Vertical lines denote the boundary of the hybrid regions. (a) The diffusion profile for the monoatomic liquid at $T^*/10$, averaged over 1000 different time origins, at $t^* = 0$, $t^* = 10$ and $t^* = 150$ for atoms that are initially localized at the slab on the CG region (left part) and also on the path integral region (right part) for the temperature $T^*/10$. In b) and c) for the temperature $T^*/\sqrt{50}$ and $T^*/5$ respectively.

5.4.2 Case II: The molecular liquid

Here we use a tetrahedral model [93] for the molecular liquid which can be used to test the methodology introduced at the begin of this chapter. The quantum representation for each atom of the tetrahedral molecule is given by a ring polymer whose Trotter number or number of beads "n" is set to 30, which is usually used in the path integral simulation of water molecules in the liquid phase [92]. The corresponding coarse-grained model for the tetrahedral molecule is seen as an effective one-site classical force field. For this system the idea of changing the molecular representation on the fly is "more challenging" due to the very extended reduction of degrees of freedom in comparison to the monatomic case. The Figure 5.9 illustrates the process of changing representation on the fly starting on the right side with the path integral representation of tetrahedral molecule and ending in the opposite side as one coarse-grained sphere, in between the system becomes a hybrid particle. Note: the center of mass (CM) of a path integral tetrahedral molecule is obtained as the average position of all the beads per molecule (i.e $R_{\rm CM} = \sum_{\alpha=1}^{4} \sum_{i=1}^{30} \frac{m_i^{(\alpha)} r_i^{(\alpha)}}{m_i^{(\alpha)}}$). We have studied a system of thousand molecules at temperature $T_1 = \frac{T^*}{\sqrt{2}}$ and $T_2 = \frac{T^*}{\sqrt{10}}$ where $T^* = 1$ (in the reduced Lennard-Jones units, $\epsilon = 1$ and $\sigma = 1$) corresponds to the classical limit. For the purpose of testing we have chosen the temperatures to be lower than that employed in the classical simulation T^*, such condition mimics the thermodynamic conditions of a "more quantum" system.

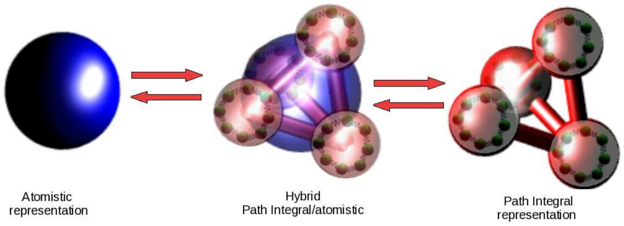

Figure 5.9: The on-the-fly interchange among the path integral (PI) and coarse-grained (CG) representation. Here we depict the case of a quantum molecule in the high resolution (right side) given by the PI representation which due to the isomorphism of the quantum partition function, each atom becomes a ring polymer.

Determination of the effective potential

Here we perform also the Iterative Boltzmann Inversion [16] method and pressure correction to obtain the effective potential, $V^{\rm eff}(R)$ for the CM between tetrahedral molecules. We emphasize, that each single atom in the tetrahedral molecule is represented by 30 beads due to the path integral approach. Thus, the effective potential takes into account the spatial reduction of degrees of freedom (DOF) from 360 to 3 per molecule. As in the monoatomic case, we obtained a set of potentials corresponding to different temperatures and for density $\rho^* = 0.1$. We depict in Figure 5.10 (a) the CG potential obtained by IBI with pressure correction. One can see that as the temperature decreases, the CG potential becomes more repulsive (less softer), contradicting our previous results (monoatomic liquid). This effect can be explained if we see the part (b) of this figure, where the bond length distribution, $P(d^*)$, between the atoms in the tetrahedral molecules is plotted. Therefore, one can infer how the molecule is delocalized as the temperature decreases in its path integral representation, generating the net effect of moving away the center of masses.

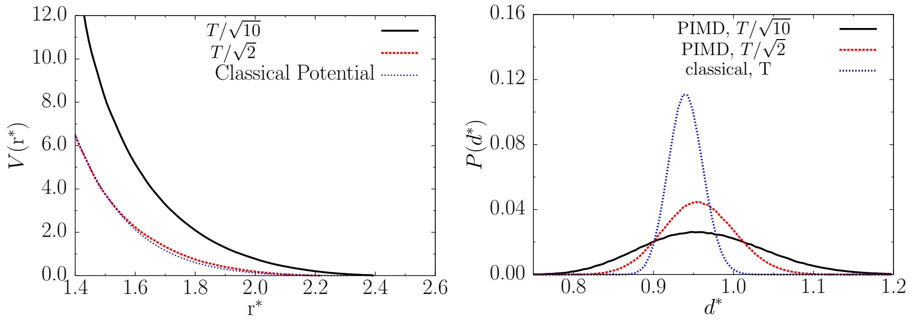

Figure 5.10: (a) The effective potential V^* for the center of masses (CMs) of the tetrahedral molecules in their path integral representation for $\rho^* = 0.1$ at $T^* = 1$. Similarly as in Figure 5.3 when the temperature is high the effective CG potential approaches to the classical CM-CM effective potential obtained from a classical simulation. (b) Shows the bond length distribution for the tetrahedral molecule in the path integral approach.

Statistical properties

Here we report two important aspects of our PIMD/CG approach with AdResS. The first has to do with the macroscopic information such as the center of masses radial distribution function (RDF) and the diffusion of center of masses along the x direction (see part (a) in Figures 5.11, 5.12 and 5.13). The second aspect relates the microscopic information (quantum delocalization) of the molecules. To guarantee that we have sampled appropriately the bead-bead configurations in our hybrid algorithm within AdResS framework, we calculate the partial bead-bead RDF in the quantum region of the adaptive simulation (see part (b) in Figures 5.11 and 5.12). This is compared with the RDF calculated in the same subregion, but from a full explicit path integral simulation. The local and global information is preserved in our hybrid scheme in all over the simulation box. This means that a quantum molecule is able to change representation from a detailed path integral description to a single classical CG site and vice versa and hence such a process maintains the thermodynamic equilibrium. Table 5.3 shows the number of particles for each description at given temperature, the path integral (PI) and coarse grained (CG), which maintain almost the same value (not spurious flux of particles) and same pressure, P, is ensured for each resolution, maintaining the mechanical equilibrium.

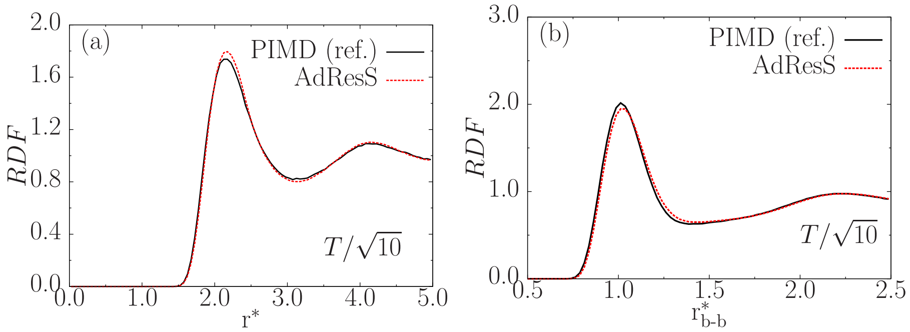

Figure 5.11: Center-of-mass radial distribution function (RDF) for all the molecules in the box. We compare the RDF of CMs in AdResS and the corresponding full explicit PI simulation (reference system), the agreement is acceptable in all box. (b) Corresponds to bead-bead radial distribution function obtained with AdResS in the quantum region compared with that of the full path integral reference system for the temperature, $T^*/\sqrt{10}$.

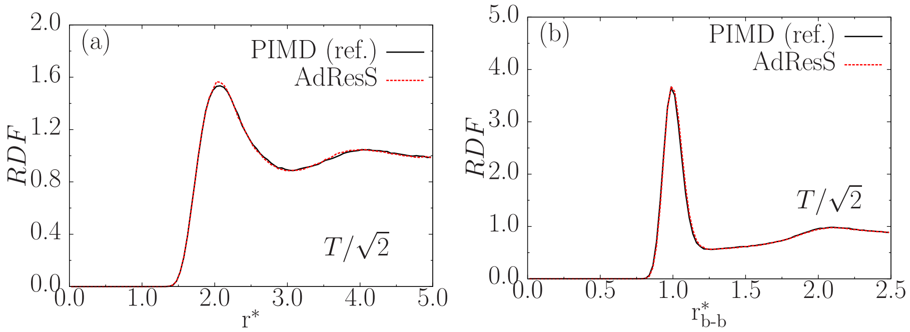

Figure 5.12: As for Figure 5.11, but in this case for $T^*/\sqrt{2}$. The same agreement between the global center-of-mass and local bead-bead radial distribution function with the full path integral reference system.

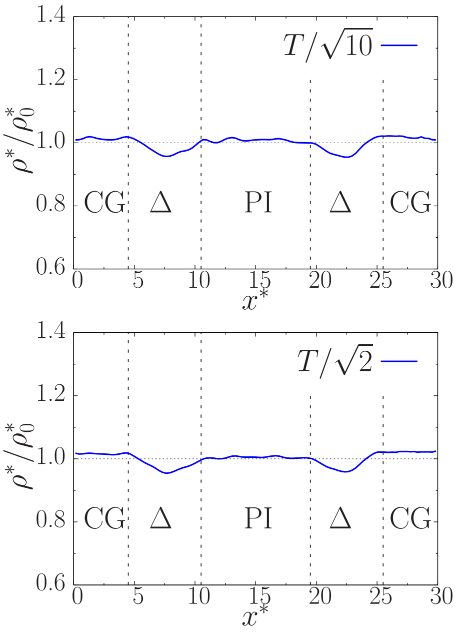

Figure 5.13: Normalized density profile of the center of masses in the x direction for tetrahedral molecules in AdResS. The width of the interface regions layer is given by 6σ. The vertical lines denote the boundaries between explicit path integral (PI), coarse grained (CG) and hybrid (Δ) regions of the system. Top (a) corresponds to the temperature, $T^*/\sqrt{10}$ (more quantum case). Bottom (b) for $T^*/\sqrt{2}$.

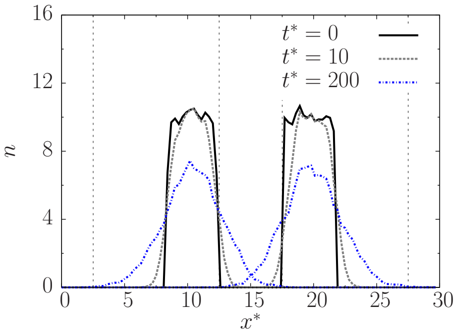

Figure 5.14: Time evolution of diffusion profile for the tetrahedral molecules that are initially, at time $t^* = 0$, localized at two neighboring slabs of the midinterface layer with 6σ interface layer width (n is the number of atoms/molecules with the CM position at a given coordinate x^*). The width of the slab is 4σ. Vertical lines denote the boundary of the hybrid regions. The diffusion profile is averaged over 1000 different time origins, at $t^* = 0$, $t^* = 10$ and $t^* = 200$ for molecules that are initially localized at the slab on the CG region (left part) and also on the path integral region (right part) for the temperature $T^*/\sqrt{10}$ (more quantum regime).

Table 5.2: **The monoatomic liquid**: Number of explicit (PI representation), hybrid and CG particles as a function of the different temperatures studied within AdResS method. The interface layer width is given by $d^*_{hyb} = 8$. Column 5^{th}, 6^{th} and 7^{th} show the pressure calculated in the path integral zone (P_{ex}), CG zone (P_{cg}) and full hybrid system with $w = 0.5$.

Temperature	n_{ex}	n_{cg}	n_{hyb}	P_{ex}	P_{cg}	$P_{w=1/2}$
$T/10$	310 ± 10	307 ± 14	382 ± 16	0.797 ± 0.013	0.797 ± 0.008	0.874 ± 0.013
$T/\sqrt{50}$	306 ± 17	304 ± 12	389 ± 21	0.638 ± 0.013	0.639 ± 0.008	0.668 ± 0.012
$T/5$	303 ± 13	297 ± 18	400 ± 20	0.481 ± 0.013	0.481 ± 0.015	0.485 ± 0.014
$T/\sqrt{10}$	295 ± 18	300 ± 14	405 ± 21	0.443 ± 0.015	0.444 ± 0.015	0.450 ± 0.014
T	301 ± 12	299 ± 11	406 ± 15	0.382 ± 0.012	0.382 ± 0.012	0.381 ± 0.013

Table 5.3: **The molecular liquid**: Number of explicit (PI representation), hybrid and CG particles as a function of different temperatures studied with AdResS. The interface layer width is given by $d^*_{hyb} = 6$. Column 5^{th} and 6^{th} show the pressure of the path integral (P_{ex}) and cg systems (P_{cg}).

Temperature	n_{ex}	n_{cg}	n_{hyb}	P_{ex}	P_{cg}
$T/\sqrt{10}$	305 ± 15	304 ± 16	382 ± 14	0.582 ± 0.08	0.581 ± 0.021
$T/\sqrt{2}$	303 ± 14	302 ± 15	387 ± 16	0.431 ± 0.08	0.430 ± 0.021
T	300 ± 15	300 ± 15	400 ± 15	0.382 ± 0.08	0.381 ± 0.021

Chapter 6

The *para*-hydrogen liquid in AdResS

In the previous chapter, we developed the basic idea of a quantum-classical adaptive coupling by mapping the quantum nature of an atom onto a classical polymer ring representation. To test the applicability of the method we have used a toy model of tetrahedral molecules with limited physical meaning, however it was appropriate to test our approach for a broad range of thermodynamic situations. The tetrahedral molecule possesses enough structural complexity, typical of small multiatomic molecules, and we have proven that a large reduction of the number of degrees of freedom in the adaptive process could be described by the AdResS method. Typically, there are critical situations (e.g. low temperatures or light particles) where the quantum description of the matter is mandatory in order to describe the relevant properties of the system. In such cases, a good starting point to include quantum effects, e.g. "delocalization or tunneling" of particles, is given by path integral methods. Here we study the *para*-hydrogen molecules at low temperatures as the first real application of our method because of: (a) the extreme thermodynamic condition at low temperature and pressure and (b) the strong quantum nature of the hydrogen molecule under these conditions. We do not aim to explore the low temperature physics within the AdResS framework, since our interest is primarily focused on the soft matter scale at ambient condition, where quantum effects are not very dominat. If our approach works technically and conceptually, one should expect even a better agreement at ambient conditions.

This chapter is organized as follows. First we introduce a brief description of the physics of the hydrogen, then we present the path integral representation of one of its molecular states, the *para*-hydrogen at low temperature. Next section presents the basis of the adaptive resolution scheme used to study the *para*-hydrogen molecules. The last two sections show the technical details of our simulation, the results and discussion.

6.1 Basic physics of hydrogen

During the early development of the quantum mechanics the significance of understanding a single hydrogen atom/molecule played an important role in the theoretical and experimental foundation of this theory. For instance, the atomic and molecular form of hydrogen were used to study the quantum structure of matter by N. Bohr and A. Sommerfeld [94, 95] (e.g. atomic shell model), and later it was also used for the development of the chemical binding by L. Pauling and several others [96, 97].

The liquid and solid phases of molecular hydrogen and deuterium have been studied experimentally [98, 99] and theoretically [100, 101]. These many body systems continue to attract great deal of interest due to its fundamental nature. It is known that hydrogen molecules are the principal constituents of distant planets [102] and in the field of high-pressure physics the hydrogen exhibits the fluid metal-insulator transition [103, 104]. In soft matter systems, the hydrogen becomes important as a constituent of important molecules such as water, DNA and proteins. The presence of hydrogen in such systems play an important role in the structure of the molecules by forming covalent or hydrogen bonds.

It is known that the spectrum of molecular hydrogen in the liquid phase exhibits the effect of the internal nuclear degrees of freedom which lead to the characterization of two spin isomers of hydrogen diatomic molecules and different levels of excitations. One of the isomeric forms is the *ortho*-hydrogen where the two proton spins are aligned parallel and form a triplet with a total spin quantum number of $J = 1$; in the *para*-hydrogen form the proton spins are aligned antiparallel and form a singlet with a total spin quantum number of $J = 0$. At standard temperature and pressure (STP) of 0 °C and 1 atm respectively, the hydrogen gas consists of about 25% *para*-hydrogen and 75% *ortho*-hydrogen which is a consequence of the spin degeneracy ratio. This scenario changes significantly if thermal equilibrium is established at low temperatures between the two forms of hydrogen. At 20 K, for example, natural hydrogen consists of 99.8% of parahydrogen. In the following section the quantum description of *para*-hydrogen is introduced by the path integral method.

6.2 The path integral description of *para*-hydrogen

As we described previously the *para*-hydrogen represents an ideal system for testing of new methods which include quantum effects in molecular dynamics simulation. In general, liquid hydrogen does not exhibit the strong identical particle exchange effects typically observed in liquid helium, and thus the physics of the system is simplified. Even at very low temperature (e.g. around the triple point, 13.8 K) the thermal de Broglie wavelength $\lambda = h/(2\pi m k_B T)^{1/2} = 3.3$ Å is slightly larger than the mean distance between two hard spheres in the classical representation of hydrogen molecules ($\sigma = 3.0$ Å). This implies that the exchange of identical *para*-hydrogen molecules will not have a pronounced effect in the properties of the liquid phase [105] and therefore the *para*-hydrogen molecules follow Boltzmann statistics. Another characteristic of the liquid molecular hydrogen at low temperatures is the inversion of the predominant population, i.e.; *ortho*-hydrogen at room temperature to the *para*-hydrogen at lower temperatures. Since the majority of the *para*-hydrogen molecules are in the ground state ($J = 0$), the wave functions that characterize such a molecular state are spherically symmetric, and hence the interaction between molecules can be modeled by an effective isotropic pair potential [106] (see Figure 6.4).

Due to the inherent quantum behavior of the *para*-hydrogen molecules at low temperatures, several computational techniques based on the path integral formulation of quantum mechanics (see chapter IV) became routine methods to calculate static [107–109] and dynamic [110–112] properties of this quantum liquid. In Figure 6.1, we show the conceptual transition from the "effective spherical" classical model to the corresponding quantum "ring polymer" model by means of the path integral approach.

In the next section, we describe the coupling of the path integral representation with an effective classical representation, where only the classical spherical symmetry is preserved and still, locally, quantum information can be extracted in the spirit of AdResS.

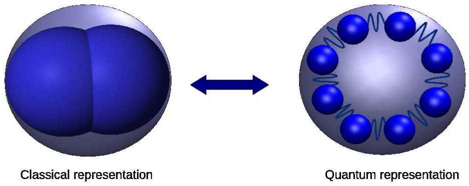

Figure 6.1: Quantum representation of the *para*-hydrogen by the path integral approach. On the left side the classical form for a diatomic hydrogen molecule is illustrated and in the opposite side the quantum description of the molecule by a ring-polymer as it is known from the path integral approach.

6.3 *Para*-hydrogen in AdResS

As stated previously the *para*-hydrogen can be accurately described by the path integral approach and represents a valid conceptual and technical test for the adaptive simulation approach in its extension to (some) quantum problems. For this reason we have performed simulations of the liquid *para*-hydrogen using AdResS. In Figure 6.2 we illustrate the particle transition from a classical, one-site, coarse grained representation to a quantum representation by means of the path integral approach. In between, we depict the hybrid resolution that characterizes the transition region. Our scheme allows the switching of representations according to the molecular position in the simulation box. This will be explained further in the next section.

An important key feature of the adaptive simulation scheme is the interpo-

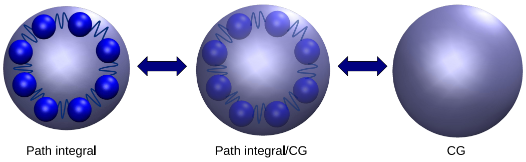

Figure 6.2: Adaptive resolution scheme for *para*-hydrogen: the high resolution on the left corresponds to the quantum path integral representation. The low resolution on right corresponds to the effective spherical classical model obtained by coarse-graining the high resolution.

lation function, $w(x)$, which weighs the force contributions coming from each representation (i.e. path integral, hybrid and coarse grained) and also makes a smooth transition from the classical to quantum regime and vice versa. The functional form of $w(x)$ is the same as the one used in classical adaptive simulation. In figure 6.3, we depict the weighting function and the *para*-hydrogen molecules in our adaptive simulation as a function of their position. The classical CG, path integral and hybrid description of the molecules are restricted in the space, but free diffusion of the molecules takes place and thus single molecules undergo an adaptive process from the path integral to classical representation.

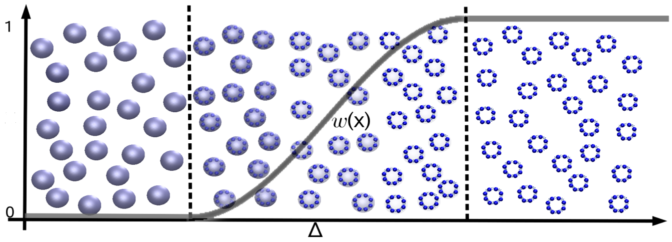

Figure 6.3: AdResS set up for the *para*-hydrogen liquid. In this figure the weighting function is shown and it varies from 0 in the coarse grained region to 1 in the path integral region, in between takes continuous values and molecules become hybrid in the sense of classical/quantum description.

6.4 Technical details

To perform an adaptive PIMD/CG simulation of liquid *para*-hydrogen we have used the Silvera-Goldman pair potential [106],

$$V(r) = e^{\alpha-\beta r-\gamma r^2} - \left(\frac{C_6}{r^6} + \frac{C_8}{r^8} + \frac{C_{10}}{r^{10}}\right) f_c(r) + \frac{C_9}{r^9} f_c(r), \tag{6.1}$$

where

$$f_c(r) = \left\{ \begin{array}{rl} e^{-(r_c/r-1)^2}, & \text{if } r \leq r_c \\ 1, & \text{otherwise,} \end{array} \right. \tag{6.2}$$

here the interacting diatomic molecules are treated as spheres. This is justified by the fact that molecules are in the rotational ground state (J-0) at the two temperatures of interest, i.e.; 25 K and 14 K. The first term on the Eq. 6.1 corresponds to the short-range repulsive interaction, while the second term describes the long-range attractive interaction. The last term, C_9/r^9, is an effective two body approximation to the three-body dispersion interaction. The $f_c(r)$ is used to screen the effect of the attractive interaction at short distance. The values of the parameter are listed in Table 6.1. V(r) is shown in Figure 6.4

Table 6.1: Parameters used in the Silvera-Goldman pair potential (in atomic units).

α 1.713	C_6 12.14
β 1.5671	C_8 215.2
γ 0.00993	C_9 143.1
r_c 8.32	C_{10} 4813.9

In our adaptive simulation of the *para*-hydrogen, we used the theoretical number density obtained from the earlier path integral Monte Carlo (PIMC) [107] ($\rho = 0.0035(\text{bohr})^{-3}$ at 14 K and $\rho = 0.0028(\text{bohr})^{-3}$ at 25 K), under the condition of almost zero pressure. Important quantities are expressed in atomic units (i.e. $e, \hbar, m = 1$). This gives the unit of energy in Hartree ($E_h = 4.3597 \times 10^{-18}$J) and the distance in Bohr radius ($a_0 = 0.529$Å) The bead-bead interaction of neighboring ring-polymers is truncated at 15 bohr.

From the earlier simulation with PIMC it is known that at the low temperature ($T = 14\ K$) the degree of quantumness for *para*-hydrogen become stronger and the Trotter number used in the path integral approach should be increased to 48 beads. For the high temperature ($T = 25\ K$) case, 28 beads are enough to guarantee the convergence of the static properties. For comparison we perform adaptive simulations of two kinds of path integrals; (a) path integral formulation in the primitive (real) space PIMD (see chapter IV), which is computationally less demanding, but shown to be not accurate enough for the low temperature limit and (b) the path integral formulation in terms of the normal modes (PIMD+NM), computationally more demanding but more accurate at the low temperature. In fact, at the low temperatures, the use of the large number of beads leads to the situation that the interaction between the corresponding beads of different ring polymers becomes less relevant because it follows $\frac{1}{n}V_{i\alpha,j\beta}$ ($i\alpha$ indicates the i-th bead of the molecules α, same for β, and n is the number of beads). However, the bead-bead interaction between bonded neighboring beads of the same ring polymer becomes dominant, because it scales as n. This leads to the problems of ergodicity in a real space approach, which can be overcome in the normal mode approach, by ensuring the proper sampling of all the frequency of the system, solves the problem [113] (see appendix E).

Each simulation was equilibrated for 100 ps and then the static properties were calculated by averaging over 1000 consecutive 10 ps path integral trajectories with a time step of 0.5 fs.

6.4.1 Effective coarse grained potential

To obtain an effective one-site coarse grained (CG) potential from the path integral representation of *para*-hydrogen, we have used the standard procedure known as the Iterative Boltzmann Inversion (IBI) [16]. Such procedure derives an effective potential by using the full explicit (path integral) center-of-mass radial distribution function (RDF) as a target. Also for this system a pressure correction has been employed to ensure as close as possible to the target pressure. In this aspect, the coarse-graining procedure at these thermodynamic conditions becomes more difficult since the large pressure fluctuations (of the order of 10^{-7} E_h/bohr^{-3}) and the inherent error of the iterative procedure lead to a large relative error on the pressure of the coarse-grained model. For each system the IBI was applied over 30 iterations until the convergency of the target radial distribution was reached. Each iteration consists of 50 ps and 500 ps of equilibration and production respectively. A smoothing procedure over the potential was applied 5 times per iteration.

Our results are shown in Figure 6.4 where we plot the effective CG potential obtained for each temperature studied and for comparison is depicted the classical Silvera and Goldman potential from Eq. 6.1. One can see the effective potential becomes less attractive and the minima of the potential is shifted as the temperature decreases from T=25 K to T=14 K.

6.5 Results and Discussion

In this section we present the results of our approach in AdResS for the two thermodynamic state points studied. The path integral implementation in the real space will be denoted by PIMD while the normal modes implementation as PIMD+NM.

6.5.1 For T=25 K

This thermodynamic state represents a "less quantum" system with Trotter number (number of beads) n=28 compared to the other system at T=14 K shown later. As one can see in the Figure 6.5 (a) for the bead-bead RDF for the full (explicit) path integral simulation the PIMD and PIMD+NM implementation perform quite well and no differences with the reference data is found compared to the reference data [107]. In part (b) and (c) we report the partial bead-bead RDF obtained within AdResS, with PIMD and the PIMD+NM, calculated in the quantum region of the adaptive resolution system. This is

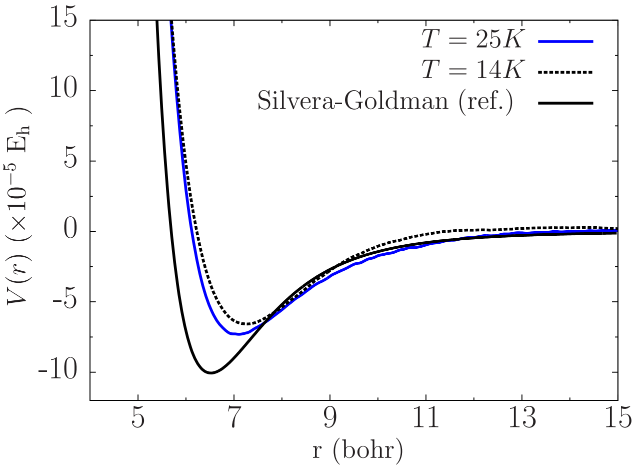

Figure 6.4: A plot of the potentials resulting from the Iterative Boltzmann Inversion (IBI) procedure. The solid (blue) line corresponds to T=25 K, $\rho = 0.0028(\text{bohr})^{-3}$ and the (black) dashed line for T=14 K, $\rho = 0.0035(\text{bohr})^{-3}$. The classical Silvera and Goldman potential [106] in solid (black) line is also depicted.

compared with the RDF calculated in the same subregion but from a full explicit path integral simulation. The agreement is highly satisfactory. In Figure 6.6 we compare the center-of-mass RDF over the entire simulation box for the PIMD and PIMD+NM in AdResS and compared it with the reference data; also in this case a full agreement is found at this temperature.

To further prove the validity of the adaptive simulation to achieve thermodynamic equilibrium, we have calculated the density profile in Figure 6.7. This shows the typical 5% drop of density in the hybrid region (Δ), which is compensated by the increment of the density in the path integral and coarse grained regions. The slightly large density in the coarse-grained region is mostly due to the abovementioned problem of the difficulties of targeting the coarse-grained model to the exact pressure of the high resolution system. However, the overall agreement is rather satisfactory.

Finally, in order to check the free diffusion of particles across the region we have calculated the diffusion profile in Figure 6.8; which proves that indeed the *para*-hydrogen molecules move through the hybrid region diffusing from the high resolution to the low resolution and vice versa.

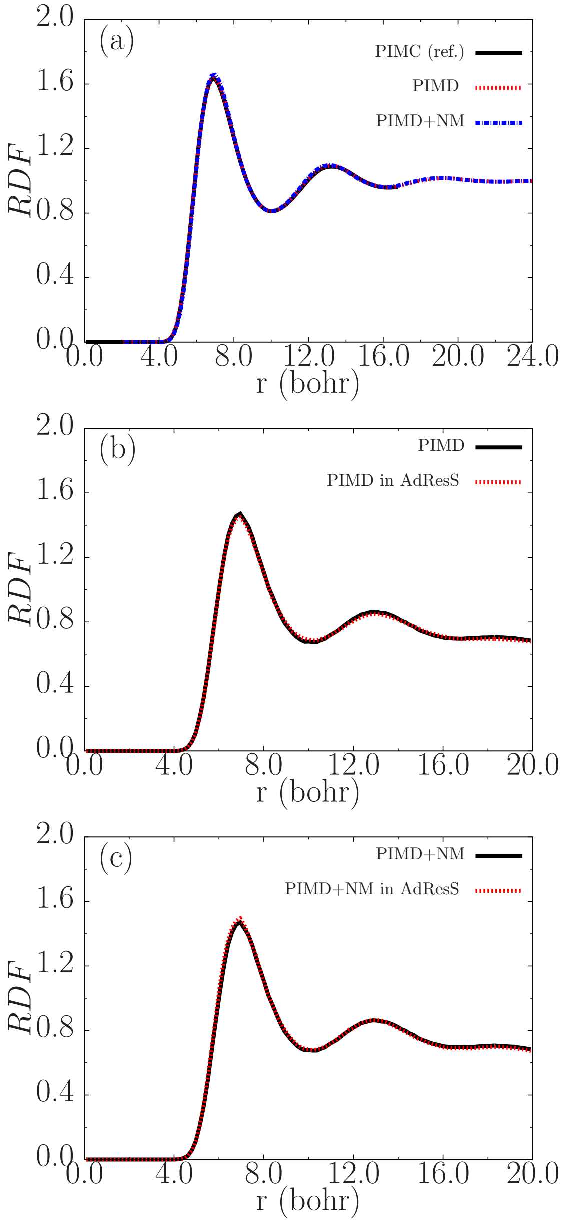

Figure 6.5: (a) Comparison of the bead-bead radial distribution functions (RDF) in a full path integral simulation at T=25 K. The primitive path integral (PIMD) and the path integral in the normal modes (PIMD+NM) are compared to earlier work [107] and full agreement is obtained. (b) The bead-bead partial RDF calculate only in the quantum region in AdResS for the PIMD compared with the same quantity calculated in the PIMD/AdResS approach. (c) Shows the same as in (b), but for PIMD+NM implementation.

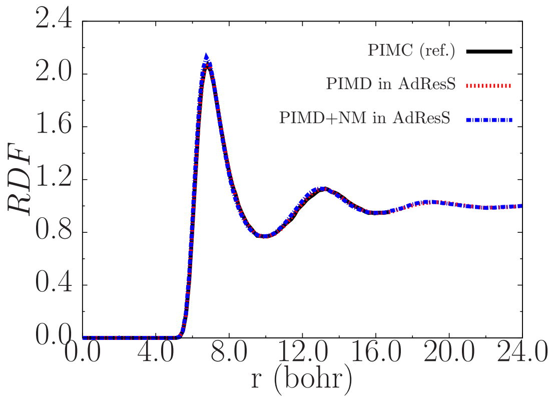

Figure 6.6: Center-of-mass radial distribution function (RDF) evaluated in the whole box at T=25 K. The agreement between the reference [107] calculation, the primitive path integral (PIMD) in AdResS and PIMD+NM in AdResS are highly satisfactory.

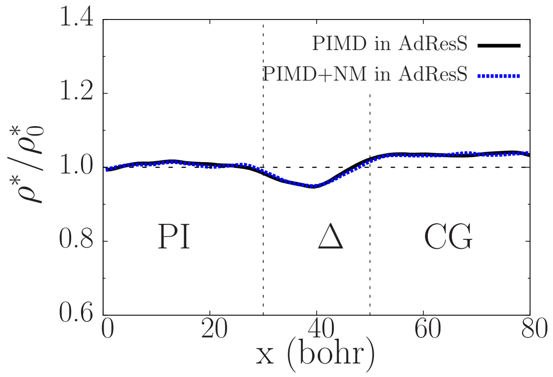

Figure 6.7: Normalized density profile along the x direction at T=25 K in AdResS. The interface region layer is given by 20 bohr. The vertical lines denote the boundaries between the path integral (PI), coarse grained (CG) and hybrid (Δ) regions of the system. In the figure is depicted the case for the primitive path integral in the real space (PIMD) and in the normal modes (PIMD+NM) in AdResS. Drop of density in the hybrid region is about 5% of the reference value (horizontal dashed line), while the overestimation in the CG region is below 3%.

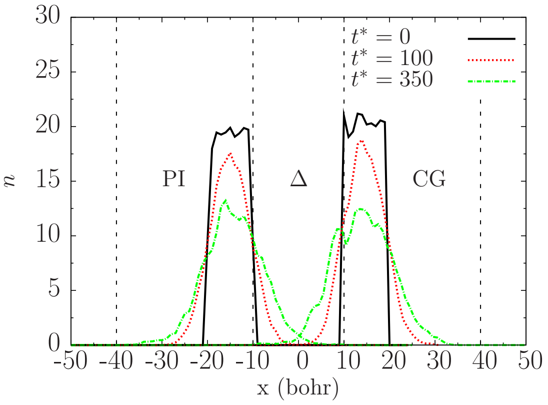

Figure 6.8: Time evolution of diffusion profile for *para*-hydrogen molecules that are initially, a time t=0, localized at two neighboring slabs of the midinterface layer with 20 bohr. interface layer width. The width of the slab is $L_x/10$, Vertical lines denote the boundary of the hybrid regions. The diffusion profile obtained by averaging over 1000 different time origins, at t=0, t=100 and t=350 for molecules that are initially localized as the slab on the CG region (right side) and also on the path integral (PI) region (left side) for the temperature T=25 k.

6.5.2 For T=14 K

This case represents the “more quantum” situation due to the larger Trotter number used, n=48, in which AdResS has been tested. One can appreciate in the Figure 6.9 (a) that the path integral implementation in real space (PIMD) does not converge to the reference data. However, the normal modes implementation reproduces the reference structure of the system [107]. This indicates only that PIMD in AdResS may not be suitable to reproduce the true physics of the system, and hence we will consider here the PIMD+NM in the adaptive scheme. Figure 6.9 (b) shows the comparison between the partial bead-bead RDF (calculated as in Figure 6.5) for the full explicit PIMD+NM and that of the PIMD+NM in AdResS and in part (c), the center-of-mass RDF for the PIMD+NM in AdResS with the reference data [107] is compared. A satisfactory agreement is found in both cases.

Similarly, to the previous case we show the density profile in Figure 6.10 for the PIMD+NM and once again we note a satisfactory agreement. The same can be said for the diffusion profile of *para*-hydrogen molecules shown in Figure 6.11.

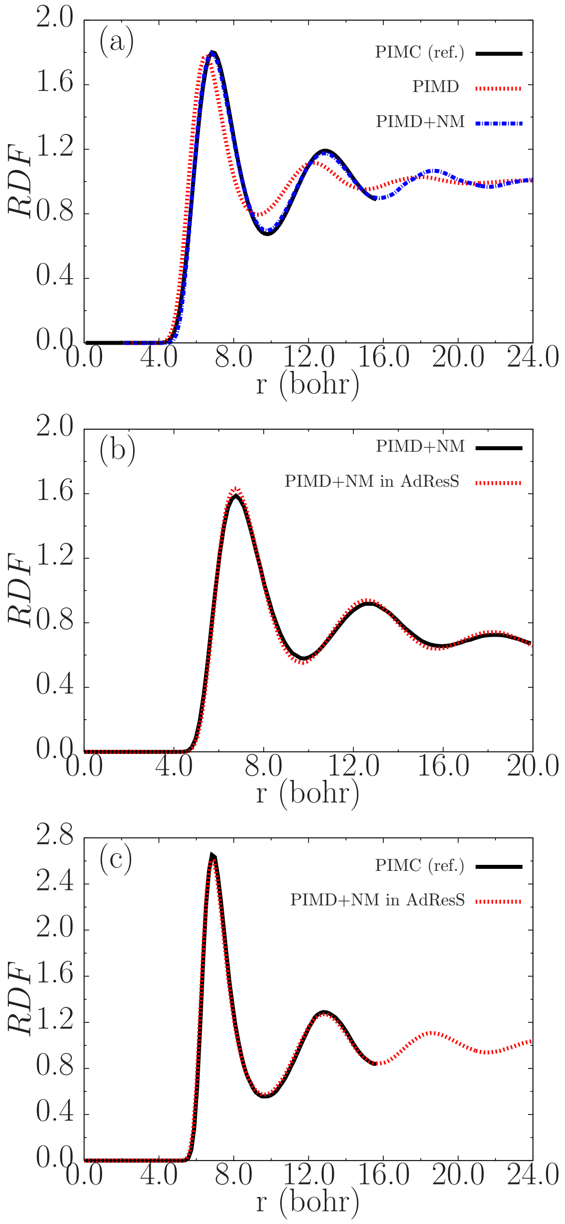

Figure 6.9: (a) Comparison of the bead-bead radial distribution functions (RDF) in a full path integral simulation at T=14 K. PIMD and PIMD+NM are compared to earlier work [107]. Agreement between the reference data and PIMD+NM is obtained while the PIMD shows to no be appropriate at this temperature. (b) Bead-bead partial RDF for a full explicit PIMD+NM and in AdResS. (c) Center-of-mass RDF for PIMD+NM in AdResS compared to the reference data (available up to 16 bohr).

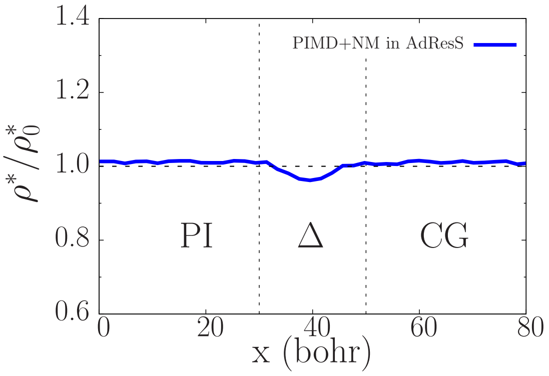

Figure 6.10: As in Figure 6.7 but for T=14K and only in the case of PIMD+NM.

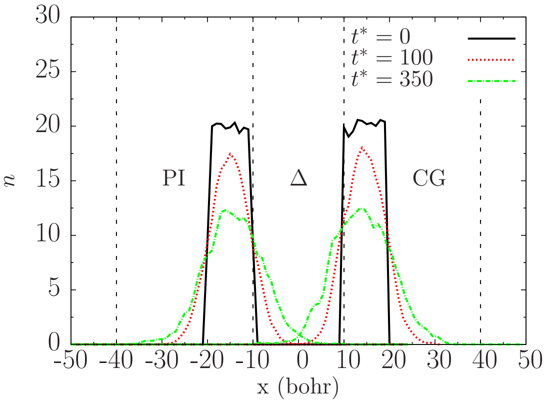

Figure 6.11: As in Figure 6.8 but at the temperature T=14K and only for the PIMD+NM case.

Chapter 7

Conclusions

In this thesis, we have dealt with two contributions in the field of method development for multiscale problems. Such methods are commonly used to overcome serious problems of time and length scales in soft matter systems. Our first contribution dealt with the numerical control of the coarse graining procedures, typically used in the reduction of degrees of freedom (DOFs) in complex systems. And the second contribution was a conceptual extension and technical development of the Adaptive Resolution Scheme (AdResS) for quantum systems. Let us explain in detail the summary of our findings and future perspectives:

In the first part, we proposed a systematic procedure to estimate the validity of the approximation of separation of variables (ASV). The advantage of our method is that only the analytical form of the potential is needed. This helps us to define regions where the ASV is reasonable a priori without the task of performing molecular dynamics simulation to calculate the correlation among DOFs. As a result, in the positive assessment our procedure indicates the error introduced by the assumption of ASV and in the negative evaluation guaratees that no separation will be possible. Our procedure can be seen as a numerical tool to systematically compare a set of coarse grained (CG) variables and choose the ones that will better mimic the dynamics of the atomistic system. For this purpose, we have studied a simple model, namely the diatomic molecule on a surface. For this system our procedure has shown the regions where the ASV becomes questionable. A second application of the procedure for two differente CG mappping schemes of a simple polymer chain has shown its practical use for polymeric systems.

Further studies can merge our procedure with several other methods, which implicitly make use of the ASV or require the a priori knowledge of the separability. For instance, the choice of collective variables (CVs) or order parameters as in the metadynamics [20] or transition path sampling [19], used commonly in the study of rare events. All these methods assume one or several CVs, which are independent from other variables whose contribution is not considered. In this context, our procedure may help to indentify a minimal set of CVs. Previously, we learnt about the problem of choosing the proper set of variables for a coarse grained description. Then one starts typically the simulation in the CG level of resolution. However, interesting multiscale problems involve several levels of description as we described in this thesis, and the need for coupling these levels and perform one single simulation is now technically possible between several scales (e.g. atomistic-mesoscopic and mesoscopic-continuum scales). Only the quantum-classical coupling presents conceptual and technical problems in adaptive simulations. In this context the Adaptive Resolution Scheme (AdResS) has shown to be a robust and successful scheme for the study of a system with concurrent resolutions.

In the second part of the thesis we extended its applicability to the quantum description based on the path integral approach of atoms/molecules. Concerning this part of the thesis, the conceptual/technical extension maintains the quantum character of the particles in the region of interest and away of it, an effective coarse grained classical description can be used to decrease the computational demands of performing full path integral simulations in the full space. We tested successfully our approach in monoatomic and molecular toy models at standart (ambient) thermodynamic conditions. To end this work we performed an application with our adaptive/path integral method in a system where quantum effects play a central role. The *para*-hydrogen molecules were studied at two different temperatures, namely 25 K and 14 K. We aimed to show the robustness of the adaptive/path integral within AdResS framework, even for such critical conditions. To account properly for the quantum effects we implemented the normal modes (NM), which was necessary for the lower temperature. Our method reproduces, in a rather satisfactory way, the structural properties when compared to the results of full explicit path integral simulation and to those available in literature. To summarize since, at ambient conditions, the quantum effects can be important, but not dominant, and the coarse-graining procedure is technically simpler, the adaptive/path integral method can be applied to standard systems in soft matter and condensed matter.

The adaptive/path integral method allows us to study the static properties of quantum systems in the region of interest while the rest of the system remains classical. Thus, real quantum “dynamics” cannot be obtained from our approach. In this respect, possible ways to overcome this problem will involve the use of more sophisticated techniques, which are still based on the path integral formalism, such as the Centroid Molecular Dynamics [73] or the Ring Polymer Molecular Dynamics [72]. Additionally, our adaptive/path integral method could be merged in the near future with practical adaptive methods for electrons [85–87], where the adaptivity of nuclei could be taken into account by our approach.

Appendix A: The RIS model for a polymer chain

The RIS model used in this study resembles the one of n-alkane chain in a plane. Here the interaction between sites (e.g. atomic or molecular) are described only by a bond-bending potential [114, 115] of the form

$$V_{\text{bending}}(\theta_i) = \frac{1}{2} K \left(\theta_i - \theta_i^0\right)^2 \tag{7.1}$$

with $K = 115.2\,\text{kcal/mol}$ and $\theta_i^0 = 112°$for T = 450K. θ_i is the angle formed by three consecutive particles, as shown in Figure 2.6. Some DOFs, such as the torsion angles are disregarded, as we focus on the separation of the variables $R^{(1)}, R^{(2)}$ and Ω which (for symmetry) are independent of the torsions in the RIS model. In both systems with 1:2 or 1:3, we have used a stiff bond length between carbon atoms equal to $l_1 = 1.54$. The difference between the two cases studied lies in the choice of the center for each bead and the number of particles per bead. For instance, in the first case we take the center of the distance between two particles as the center of the bead and in the last one we fix the internal angle formed by three particles and then use the barycenter of this triangle as the center of the bead. In both cases the angle θ_1 is kept fixed at its equilibrium value, that is $\theta_1 = 108°$, while the other angles are allowed to vary in such a way that $\theta_i = 115° \pm 10°, \forall i = 2, 3, 4$ as suggested by atomistic simulations [114, 115].

Appendix B: CG variable as function of the atomistic variables

Once the relevant CG variables are determined, one has to proceed to express the CG variables as a function of the explicit (atomistic) variables. After that one can apply systematically the criterion of the ASV and calculate the factor of quality (Q) for this particular mapping scheme . Here, for technical convenience, we proceed first by expressing the atomistic variables as a function of the new CG variables and then use the inverse function in order to get the CG variables as a function of the explicit dependencies. Due to the high complexity of the system we have fixed some atomistic variables and consider them as possible parametric variables. This is a useful procedure and a common way to treat a complex system with a high number of DOFs where some variables are more relevant than others. In order to obtain a set of equations, we analyze the geometrical properties of the MS and search for suitable relations between the explicit and CG variables. The geometrical conditions that we have used to obtain a system of equations involve the square of the absolute value of $\vec{R}^{(1)}$ and $\vec{R}^{(2)}$, which are expressed as the resultants of the sequence of collinear vectors for each case . For the first MS, now we define $\vec{r}_1 = \overrightarrow{O_1B}$, $\vec{r}_2 = \overrightarrow{BC}$, $\vec{r}_3 = \overrightarrow{CO_2}$, $\vec{r}_4 = \overrightarrow{DE}$ and $\vec{r}_5 = \overrightarrow{EO_3}$ (see Figure 2.6). So we have:

$$
\begin{array}{rcl}
\vec{R}^{(1)} & = & -(\vec{r}_1 + \vec{r}_2 + \vec{r}_3) \\
\vec{R}^{(2)} & = & \vec{r}_3 + \vec{r}_4 + \vec{r}_5
\end{array}
$$

then the square of the absolute value,

$$
\begin{array}{rcl}
(R^{(1)})^2 & = & (r_1)^2 + (r_2)^2 + (r_3)^2) \\
 & & +2\left(-r_1 r_2 \cos(\theta_1) + r_1 r_3 \cos(\theta_2 - \theta_1) - r_2 r_3 \cos(\theta_2)\right) \\
(R^{(2)})^2 & = & (r_3)^2 + (r_4)^2 + (r_5)^2 \\
 & & +2\left(-r_3 r_4 \cos(\theta_3) + r_3 r_5 \cos(\theta_4 - \theta_3) - r_4 r_5 \cos(\theta_4)\right)
\end{array}
\qquad (7.2)
$$

And similarly for the second MS having defined $\vec{r}_1 = \overrightarrow{O_1B}$, $\vec{r}_2 = \overrightarrow{BC}$, $\vec{r}_3 = \overrightarrow{CD}$,

$\vec{r}_4 = \overrightarrow{DE}$, $\vec{r}_5 = \overrightarrow{EO_2}$, $\vec{r}_6 = \overrightarrow{EF}$, $\vec{r}_7 = \overrightarrow{FG}$, $\vec{r}_8 = \overrightarrow{GH}$ and $\vec{r}_9 = \overrightarrow{HO_3}$.

$$\begin{aligned}
\vec{R}^{(1)} &= -(\vec{r}_1 + \vec{r}_2 + \vec{r}_3 + \vec{r}_4 + \vec{r}_5) \\
\vec{R}^{(2)} &= -\vec{r}_5 + \vec{r}_6 + \vec{r}_7 + \vec{r}_8 + \vec{r}_9
\end{aligned}$$

once again the square of the absolute value,

$$\begin{aligned}
(R^{(1)})^2 &= (r_1)^2 + (r_2)^2 + (r_3)^2 + (r_4)^2 + (r_5)^2 + 2(-r_1 r_2 \cos(\frac{\theta_1}{2}) \\
&\quad + r_1 r_3 cos(\theta_2 - \frac{\theta_1}{2}) - r_1 r_4 \cos(\theta_3 + \frac{\theta_1}{2} - \theta_2) + r_1 r_5 \cos(\theta_2 - \theta_3) \\
&\quad - r_2 r_3 \cos(\theta_2) + r_2 r_4 \cos(\theta_2 - \theta_3) - r_2 r_5 \cos(\frac{\theta_1}{2} + \theta_2 - \theta_3) \\
&\quad - r_3 r_4 \cos(\theta_3) + r_3 r_5 \cos(\theta_3 - \frac{\theta_1}{2}) - r_4 r_5 \cos(\frac{\theta_1}{2})) \\
(R^{(2)})^2 &= (r_5)^2 + (r_6)^2 + (r_7)^2 + (r_8)^2 + (r_9)^2 + 2(-r_5 r_6 \cos(\frac{\theta_1}{2}) \\
&\quad + r_5 r_7 cos(\theta_4 - \frac{\theta_1}{2}) - r_5 r_8 \cos(\theta_5 + \frac{\theta_1}{2} - \theta_4) + r_5 r_9 \cos(\theta_4 - \theta_5) \\
&\quad - r_6 r_7 \cos(\theta_4) + r_6 r_8 \cos(\theta_5 - \theta_4) - r_6 r_9 \cos(\frac{\theta_1}{2} + \theta_4 - \theta_5) \\
&\quad - r_7 r_8 \cos(\theta_5) + r_7 r_9 \cos(\theta_5 - \frac{\theta_1}{2}) - r_8 r_9 \cos(\frac{\theta_1}{2})) \qquad (7.3)
\end{aligned}$$

Despite the scalar product of $\vec{R}^{(1)} \cdot \vec{R}^{(2)}$ being a valid relation to obtain Ω, we notice that it is not well handled by the conventional inverse procedure because it involves no simple argument dependencies. Hence we use a particular relation for each MS as shown below.

1. Case Figure 2.6(a) : For Ω we have: $\Omega = \pi - D\widehat{O_2}O_3 + O_1\widehat{O_2}C$, then use the scalar product of $\vec{r}_3 \cdot \vec{r}_4 = -r_3 r_4 cos\,(\theta_3)$ and since $\vec{r}_4 = \vec{R}^{(2)} - \vec{r}_3 - \vec{r}_5$, we have that,

   $$\begin{aligned}
   -r_3 r_4 cos\,(\theta_3) &= \vec{r}_3 \cdot \left(\vec{R}^{(2)} - \vec{r}_3 - \vec{r}_5\right) \\
   &= \left(r_3 R^{(2)} cos\left(D\widehat{O_2}O_3\right) - r_3^2 - r_3 r_5 cos(\theta_4 - \theta_5)\right)
   \end{aligned}$$

   thus inverting this relation we obtain $D\widehat{O_2}O_3$. The remaining angle $O_1\widehat{O_2}C$ can be expressed as a function of the $R^{(1)}$ and is equal to:

   $$O_1\widehat{O_2}C = \arccos\left[\frac{(R^{(1)})^2 - 4(r_1)^2(1 - 2cos(\theta_1)}{2R^{(1)} r_1}\right]$$

2. Case Figure 2.6(b) : $\Omega = 2\pi - D\widehat{O_2}F - O_1\widehat{O_2}D - O_3\widehat{O_2}F$ and due to the symmetry of the system the last two angles on the r.h.s of the equation

are mathematically similar. We now show the geometrical procedure to determine the expression for one of them, namely $\widehat{O_1O_2D}$.

$$\begin{aligned} \triangle O_1DC: \quad \overline{O_1D}^{\,2} &= \overline{O_1C}^{\,2} + \overline{CD}^{\,2} - 2\overline{O_1C}\,\overline{CD}\,cos\left(\widehat{O_1CD}\right) \\ \triangle O_1DO_2: \quad \overline{O_1D}^{\,2} &= \overline{O_1O_2}^{\,2} + \overline{DO_2}^{\,2} - 2\overline{O_1O_2}\,\overline{DO_2}\,cos\left(\widehat{O_1O_2D}\right) \end{aligned}$$

Where the symbol $\triangle$ indicates the triangle under consideration defined by the letters of its vertices (e.g. $\triangle O_1DC$ is the triangle whose vertices are the points O_1, D and C. Next we solve for $\widehat{O_1O_2D}$ and considering that $\overline{O_1C} = \overline{DO_2}$ we have

$$\widehat{O_1O_2D} = \frac{\arccos\left(\overline{O_1O_2}^{\,2} - \overline{CD}^{\,2} + 2\overline{O_1C}\,\overline{CD}\,cos\left(\widehat{O_1CD}\right)\right)}{2\overline{O_1O_2}\,\overline{O_1C}},$$

for analogy, for the other angle $\widehat{O_3O_2F}$:

$$\widehat{O_3O_2F} = \frac{\arccos\left(\overline{O_2O_3}^{\,2} - \overline{FG}^{\,2} + 2\overline{GO_3}\,\overline{FG}\,cos\left(\widehat{FGO_3}\right)\right)}{2\overline{O_2O_3}\,\overline{GO_3}}.$$

where $\widehat{O_1CD} = \theta_2 + \widehat{O_1CB}$ and $\widehat{O_3O_2F} = \theta_5 + \widehat{HGO_3}$. Due to the symmetry we have that $\widehat{O_1CB} = \widehat{HGO_3}$, this angle can be calculated using the geometrical properties of the triangles and is equal to,

$$\widehat{O_1CB} = \frac{3}{8} + \frac{1}{2}\left(\frac{5}{4} - cos\left(\theta_1\right)\right)^{\frac{1}{2}}$$

From our notation for CG variables we have,

$$\overline{O_1O_2} = R^{(1)},$$

$$\overline{O_2O_3} = R^{(2)}$$

and,

$$\overline{CD} = \overline{FG} = l_1.$$

Despite the appearance of a complicated mathematical procedure, technically this relations are not difficult to obtain and by using standard computational tools as Mathematica [116], one can obtain "numerical" expressions of the transformation which can be directly plugged into a computer code. The explicit expressions are rather lengthy (but easy to use into the numerical procedure of the ASV) and would occupy too much space, thus they are not reported here. In any case the formal procedure reported above is sufficient to reproduce all the calculations we are performing. We have tested the correctness of the explicit expressions obtained by calculating several values of the potential using both the explicit atomistic coordinates and the corresponding CG value of $R^{(1)}, R^{(2)}$ and Ω. Below the formal relations corresponding to the two MS are reported;

for the 1:2 MS, we have the following dependencies with θ_1 and l_1 as parametric values:

$$\begin{aligned} \theta_1 &= \bar{\theta}_1 \\ R^{(1)} = R^{(1)}(\bar{l}_1, \bar{\theta}_1, \theta_2) \longrightarrow \theta_2 &= \theta_2(\bar{l}_1, \bar{\theta}_1, R^{(1)}) \\ R^{(2)} = R^{(2)}(\bar{l}_1, \bar{\theta}_1, \theta_2, \theta_3, \theta_4) \longrightarrow \theta_3 &= \theta_3(\bar{l}_1, \bar{\theta}_1, R^{(1)}, R^{(2)}, \Omega) \\ \Omega = \Omega(\bar{l}_1, \bar{\theta}_1, \theta_2, \theta_3, \theta_4) \longrightarrow \theta_4 &= \theta_4(\bar{l}_1, \bar{\theta}_1, R^{(1)}, R^{(2)}, \Omega) \qquad (7.4) \end{aligned}$$

For the 1:3 MS, we fixed θ_1,θ_2 and l_1 as a parametric values and obtain:

$$\begin{aligned} \theta_1 &= \bar{\theta}_1 \\ \theta_2 &= \bar{\theta}_2 \\ R^{(1)} = R^{(1)}\left(\bar{l}_1, \bar{\theta}_1, \bar{\theta}_2, \theta_3\right) \longrightarrow \theta_3 &= \theta_3\left(\bar{l}_1, \bar{\theta}_1, \bar{\theta}_2, R^{(1)}\right) \\ R^{(2)} = R^{(2)}\left(\bar{l}_1, \bar{\theta}_1, \theta_4, \theta_5\right) \longrightarrow \theta_4 &= \theta_4\left(\bar{l}_1, \bar{\theta}_1, \bar{\theta}_2, R^{(1)}, R^{(2)}, \Omega\right) \\ \Omega = \Omega\left(\bar{l}_1, \bar{\theta}_1, \bar{\theta}_2, \theta_3, \theta_4, \theta_5\right) \longrightarrow \theta_5 &= \theta_5\left(\bar{l}_1, \bar{\theta}_1, \bar{\theta}_2, R^{(1)}, R^{(2)}, \Omega\right) \qquad (7.5) \end{aligned}$$

Appendix C: Thermostat

In computer simulations a thermostat is needed to define properly the NVT ensemble, being it responsible for the control of the temperature. In the case of the adaptive resolution scheme the thermostat is used to compensate the switching of the interactions, which occurs during the transition from an atomistic to a coarse grained region or viceversa. Such thermostat must ensure that the atoms of a molecules have the correct velocity distribution when entering or leaving the switching region. To fulfill this task, we have used the concept of the Langevin idea or stochastic dynamics to guarantee the sampling of the appropriate distributions [117]. This is done by adding a random and damping force as shown below,

$$\dot{p}_i = \nabla_i U + F_i^D + F_i^R, \tag{7.6}$$

where the damping force F_i^D is a Stokes-like friction force which acts in the oposite direction of the velocity,

$$F_i^D = -\xi_i / m_i \; p_i \tag{7.7}$$

where ξ_i is the friction constant. To compensate for this friction one adds a random force F_i^R which acts in a random direction and is completely local,

$$F_i^R = \sigma_i \eta_i(t), \tag{7.8}$$

where σ_i is the noise amplitute and η_i is a noise with certain properties:

- Homegeneity: $< \eta_i(t) >= 0$
- Independecy of the time and space: $< \eta_i(t)\eta_j(t') >= \delta_{ij}\delta(t-t')$

The ratio between ξ_i and η_i can be obtained from the Fokker-Planck formalism [118]. Let us write the corresponding Fokker-Planck operator for the stochastic part of the langevin equation (Eq. 7.6) as follows:

$$\mathcal{L}_{SD} = \sum_i \frac{\partial}{\partial p_i}[\xi_i \frac{\partial \mathcal{H}}{\partial p_i} + \sigma_i^2 \frac{\partial}{\partial p_i}] \tag{7.9}$$

where this sum runs over all the particles. By assuming that the equilibrium distribution is a Boltzmann type, one has that:

$$\mathcal{L}_{SD} e^{-\mathcal{H}/k_B T} = 0 \tag{7.10}$$

and from the above equation one obtains that:

$$\sigma_i^2 = k_B T \xi_i \tag{7.11}$$

this relation is known as the Fluctuation-Dissipation theorem (FDT). The Langevin thermostat is one of the standard local stochastic thermostats, which generates a canonical ensemble. This local thermostat has shown to stabilize the systems, which a global thermostat cannot do. A drawback of the Langevin thermostat is its lack of Galilei invariance and the strong dependece of the dynamics on the friction strength.

Appendix D: Forcefield of the tetrahedral molecule

A tetrahedral molecules is characterized by N=4 atoms of the same mass m_0 connected by anharmonic bonds. All atoms in the system interact according to a purely repulsive shifted 12-6 Lennard-Jones potential with a cutoff at $2^{1/6}\sigma$ of the form

$$U_{LJ}^{atom}(r_{i\alpha j\beta}) = \begin{cases} 4\varepsilon\big[\big(\frac{\sigma}{r_{i\alpha j\beta}}\big)^{12} - \big(\frac{\sigma}{r_{i\alpha j\beta}}\big)^{6} + \frac{1}{4}\big]; & r_{i\alpha j\beta} \leq 2^{1/6}\sigma \\ 0; & r_{i\alpha j\beta} > 2^{1/6}\sigma \end{cases} \tag{7.12}$$

$r_{i\alpha j\beta}$ defines the distance between the atom $i\alpha$ of the molecule α and the atom $j\beta$ of the molecule β. We define ε and σ as a the reference length and energy units. Neighboring atoms of a molecule are linked via an attractive finite extensible nonlinear elastic (FENE) potential

$$U_{FENE}^{atom}(r_{i\alpha j\alpha}) = \begin{cases} -\frac{1}{2}kR_0^2\ln\big[1 - \big(\frac{r_{i\alpha j\alpha}}{R_0}\big)^2\big]; & r_{i\alpha j\alpha} \leq R_0 \\ \infty; & r_{i\alpha j\alpha} > R_0 \end{cases} \tag{7.13}$$

being $R_0 = 1.5\sigma$ the divergence length and $k = 30\varepsilon/\sigma^2$ the stiffness. The equilibrium bond length at $k_BT = \varepsilon$ is approximately 1.0σ.

Appendix E: Implementation of the normal modes in AdResS

We use the Hamiltonian from Eq. 4.59 in the normal modes variables without the external potential to derive the equation of motion (EOM) of a free ring polymer. In the new system of coordinates the EOM take the form:

$$\begin{aligned} P_I^{(s)} &= \frac{dQ_I^{(s)}}{dt} \\ \frac{d}{dt}P_I^{(s)} &= -m_I \Omega_s^2 Q_I^{(s)}, \ \forall s = 1, \ldots, n, \end{aligned} \tag{7.14}$$

where Ω_s are the normal mode frequencies and they are defined in Eq. 4.58. These are the EOM of a n-decoupled harmonic oscillators and the solution is known to be at time $t + \Delta t$

$$\begin{aligned} Q_I^{(s)}(t+\Delta t) &= Q_I^{(s)}(t)\cos(\Omega_s \Delta t) + \frac{P_I^{(s)}(t)}{m_I \Omega_s}\sin(\Omega_s \Delta t) \\ P_I^{(s)}(t+\Delta t) &= -Q_I^{(s)}(t)\Omega_s m_I \sin(\Omega_s \Delta t) + P_I^{(s)}(t)\cos(\Omega_s \Delta t) \end{aligned} \tag{7.15}$$

for all $s = 1, \ldots, n$ and where $Q_I^{(s)}(t)$ and $P_I^{(s)}(t)$ are the initial conditions at time t. In a matrix form,

$$\begin{bmatrix} P_I^{(s)}(t+\Delta t) \\ Q_I^{((s)}(t+\Delta t) \end{bmatrix} = \begin{pmatrix} \cos(\Omega_s \Delta t) & -\Omega_s m_I \sin(\Omega_s \Delta t) \\ \dfrac{\sin(\Omega_s \Delta t)}{\Omega_s m_I} & \cos(\Omega_s \Delta t) \end{pmatrix} \begin{bmatrix} P_I^{(s)}(t) \\ Q_I^{(s)}(t) \end{bmatrix}$$

For simplicity we show the component of the matrices for the zero-frequency (i.e. for the centroid), this is the case of $s = 0$ and from the previous matrix form we get

$$\begin{bmatrix} P_I^{(0)}(t+\Delta t) \\ Q_I^{((0)}(t+\Delta t) \end{bmatrix} = \begin{pmatrix} 1 & 0 \\ \Delta t / m_I & 1 \end{pmatrix} \begin{bmatrix} P_I^{(0)}(t) \\ Q_I^{(0)}(t) \end{bmatrix}$$

where we used the property of $\lim_{x\to 0}\frac{\sin(x)}{x} = 1$. This form is ideal for the numerical implementation of the algorithm. Now let us synthesize the changes in the velocity Verlet algorithm to perform the normal modes (NM) in few step:

STEPs of Velocity Verlet with NM in AdResS:

1. Calculate forces using AdResS. Evaluate $F_I^{(s)}(t)$ only from the external potential in the primitive space.
2. Update velocities, $v_I^{(s)}(t+\frac{\Delta t}{2}) = v_I^{(s)}(t)+\frac{\Delta t}{2}F_I^{(s)}(t)$, in the primitive space.
3. Evolve positions and velocities at $t+\Delta t$ with **NM algorithm** (See below).
4. Calculate forces, $F_I^{(s)}(t+\frac{\Delta t}{2})$ as in STEP 1 from positions of STEP 3.
5. Update velocities, $v_I^{(s)}(t+\Delta t) = v_I^{(s)}(t+\frac{\Delta t}{2}) + \frac{\Delta t}{2}F_I^{(s)}(t+\Delta t)$, in the primitive space.

The STEP 3 concerns the implementation of the normal modes and this is presented below

NM algorithm:

1. Apply the forward FFT to $\{r_I^{(s)}(t)\}, \{p_I^{(s)}(t)\} \longrightarrow \{Q_I^{(s)}(t)\}, \{P_I^{(s)}(t)\}$ (coordinates and momenta in NM).
2. Evolve $\{Q_I^{(s)}(t)\}, \{P_I^{(s)}(t)\}$ according EOM of free ring-polymer (Eq. 7.15) to $t+\Delta t$.
3. Apply the backward FFT to $\{Q_I^{(s)}(t+\Delta t)\}, \{P_I^{(s)}(t+\Delta t)\} \longrightarrow \{r_I^{(s)}(t+\Delta t)\}, \{p_I^{(s)}(t+\Delta t)\}$.

Bibliography

[1] Top 500 supercomputer, http://www.top500.org/.

[2] D. J. Wales. *Energy Landscape*. Cambridge University Press, 1 edition, 2003.

[3] G. S. Ayton, E. Lyman and G. A. Voth. **Hierarchical coarse-graining strategy for protein-membrane systems to access mesoscopic lengthscales**. *Faraday Discuss.*, **144**, 347 (2010).

[4] X. Daura et al. **Reversible peptide folding in solution by molecular dynamics simulation**. *J. Mol. Biol.*, **280**, 925 (1998).

[5] X. Daura, W. van Gunsteren and A. Mark. **Folding-unfolding thermodynamics of a ß-heptapeptide from equilibrium simulations**. *Proteins: Struct.: Gen.*, **34**, 269 (1999).

[6] Y. Duan and P. Kollman. **Pathways to a Protein Folding Intermediate Observed in a 1-Microsecond Simulation in Aqueous Solution**. *Science*, **282**, 740 (1999).

[7] J. Xanthipe et al. **Solvation Dynamics in Protein Environments Studied by Photon Echo Spectroscopy**. *J. Phys. Chem. B*, **103**, 7995 (1999).

[8] V. Molinero and W. A. Goddard. **M3B: A CG force field for molecular simulation of Malto-Oligosaccharides and their water mixtures**. *J. Phys. Chem. B*, **108**, 1414 (2004).

[9] A. A. Louis. **Beware of density dependent pair potentials**. *J. Phys.: Condens. Matter*, **14**, 9187 (2002).

[10] R. L. C. Akkermans and W. J. Briels. **A structure-base coarse-grained model for polymer melts**. *J. Chem. Phys.*, **114**, 1020 (2001).

[11] C. F. Abrams and K. Kremer. **Combined Coarse-Grained and Atomistic Simulation of Liquid Bisphenol A−Polycarbonate: Liquid Packing and Intramolecular Structure**. *Macromolecules*, **36**, 260 (2003).

[12] V. A. Harmandaris, D. Reith N.F.A., Van der Vegt and K. Kremer. **Comparison Between Coarse-Graining Models for Polymer Systems: Two Mapping Schemes for Polystyrene**. *Macromol. Chem. Phys.*, **208**, 2109 (2007).

[13] R. L. Henderson. **A uniqueness theorem for fluid pair correlation functions**. *Phys. Lett. A*, **49**, 197 (1974).

[14] J. T. Chayes, L. Chayes and E. H. Lieb. **The inverse problem in classical statistical mechanics**. *Commun. Math. Phys.*, **93**, 57 (1984).

[15] V. Ruehle, C. Junghans, A. Lukyanov, K. Kremer and D. Adrienkov. **Versatile Object-oriented Toolkit for Coarse-graining Applications**. *J. Chem. Theor. Comp.*, **5**, 3211 (2009).

[16] D. Reith, M. Puetz and F. Mueller-Plathe. **Deriving effective mesoscale potentials from atomistic simulations**. *J. Comput. Chem.*, **24**, 1624 (2003).

[17] D. Fritz, V. harmandaris, K. Kremer and N. van der Vegt. **Coarse-Grained polymer melts based on isolated atomistic chains: Simulation of polystyrene of different tacticities**. *Macromolecules*, **42**, 7579 (2009).

[18] H. Wang, C. Junghans and K. Kremer. **Comparative atomistic and coarse-grained study of water: What do we lose by coarse-graining?** *Euro. Phys. J. E*, **28**, 221 (2009).

[19] C. Dellago, P. G. Bolhuis, F. S. Csajka and D. Chandler. **Transition path sampling and the calculation of rate constants**. *J. Chem. Phys.*, **108**, 1964 (1998).

[20] A. Laio and M. Parrinello. **Escaping free-energy minima**. *Proc. Natl. Acad. Sci. USA*, **99**, 12562 (2002).

[21] A. B. Poma and L. Delle Site. **Separation of variables in molecular-dynamics simulations: A criterion to estimate the quality of the approximation**. *Phys. Rev. E*, **78**, 056703 (2008).

[22] P. Schravendijk, L. M. Ghiringhelli, L. Delle Site and N. F. A. van der Vegt. **Interaction of Hydrated Amino Acids with Metal Surfaces: A Multiscale Modeling Description**. *J. Phys. Chem. C*, **111**, 2631 (2007).

[23] P.J. Flory. **Foundations of Rotational Isomeric State Theory and General Methods for Generating Configurational Averages**. *Macromolecules*, **7**, 381 (1974).

[24] P. J. Flory. *Statistical Mechanics of Chain Molecules.* Interscience Publishers, 1 edition, (1969).

[25] V. Bulatov, F. F. Abraham, L. Kubin B. Devincre and S. Yip. **Connecting atomistic and mesoscale simulations of crystal plasticity**. *Nature*, **591**, 669 (1998).

[26] F. F. Abraham et al. **Simulating materials failure by using up to one billion atoms and the world's fastest computer: Work-hardening**. *Proc. Natl. Acad. Sci. USA*, **99**, 5783 (2002).

[27] L. Delle Site, C.F. Abrams, A. Alavi and K. Kremer . **Polymers near Metal Surfaces: Selective Adsorption and Global Conformations**. *Phys. Rev. Lett.*, **89**, 156103 (2002).

[28] S. T. O Connell and P. A. Thompson. **Molecular dynamics–continuum hybrid computations: A tool for studying complex fluid flows**. *Phys. Rev. E.*, **52**, R5792 (1995).

[29] N. G. Hadjicostantino. **Combining atomistic and continuum simulations of contact-line motion**. *Phys. Rev. E.*, **59**, 2475 (1999).

[30] J. A. Smirnova, L. V. Zhigilei and B. J. Garrison. **A combined molecular dynamics and finite element method technique applied to laser induced pressure wave propagation**. *Comp. Phys. Commun.*, **118**, 11 (1999).

[31] J. Q. Broughton, F. F. Abraham, N. B. Bernstein and E. Kaxiras. **Concurrent coupling of length scales: Methodology and application**. *Phys. Rev. B*, **60**, 2391 (1999).

[32] G. Csanyi, T. Albaret, M. C. Payne and A. D. Vita. **"Learn on the Fly": A Hybrid Classical and Quantum-Mechanical Molecular Dynamics Simulation**. *Phys. Rev. Lett.*, **93**, 175503 (2004).

[33] A. Laio, J. van de Vondele and U. Rothlisberger. **A Hamiltonian electrostatic coupling scheme for hybrid Car–Parrinello molecular dynamics simulations**. *J. Chem. Phys.*, **116**, 6941 (2002).

[34] M. Praprotnik, L. Delle Site and K. Kremer. **Adaptive resolution molecular-dynamics simulation: Changing the degrees of freedom on the fly**. *J. Chem. Phys.*, **123**, 224106 (2005).

[35] B. Ensing, S. O. Nielsen, P. B. Moore, M. L. Klein and M. Parrinello. **Energy conservation in adaptive hybrid atomistic/coarse-grain molecular dynamics**. *J. Chem. Theory Comput.*, **3**, 1100 (2007).

[36] A. Heyden and D. G. Truhlar. **Conservative algorithm for an adaptive change of resolution in mixed Atomistic/Coarse-Grained multiscale simulations**. *J. Chem. Theory Comput.*, **4**, 217 (2008).

[37] L. Delle. Site. **Some fundamental problems for an energy-conserving adaptive-resolution molecular dynamics scheme**. *Phys. Rev. E*, **76**, 047701 (2007).

[38] Mark E. Tuckerman and Glenn J. Martyna. **Understanding Modern Molecular Dynamics: Techniques and Applications**. *J. Phys. Chem. B*, **104**, 159 (2000).

[39] G. Ciccotti and J. P. Ryckaert. **Molecular dynamics simulation of rigid molecules**. *Comp. Phys. Rep.*, **4**, 345 (1986).

[40] R. L. C. Akkermans and G. Ciccotti. **On the Equivalence of Atomic and Molecular Pressure**. *J. Phys. Chem. B*, **21**, 6866 (2004).

[41] K. Huang. *Statistical Mechanics*. John Wiley and Sons, 2 edition, (1987).

[42] R. Hilfer. *Applications of Fractional Calculus in Physics*. World Scientific Publishing Company, 1 edition, (2000).

[43] F. W. J. Olver, D. W. Lozier, R. F. Boisvert and C. W. Clark. *NIST Handbook of Mathematical Functions*. Cambridge University Press, 1 edition, (2010).

[44] M. Praprotnik, L. Delle Site and K. Kremer. **Fractional dimensions of phase space variables: A tool for varying the degrees of freedom of a system in a multiscale treatment.** *J. Phys. A: Math. Theor.*, **40**, F281 (2007).

[45] M. Praprotnik, L. Delle Site and K. Kremer. **A macromolecule in a solvent: Adaptive resolution molecular dynamics simulation**. *J. Chem. Phys.*, **126**, 134902 (2007).

[46] M. Praprotnik, S. Matysiak, L. Delle Site, K. Kremer and C. Clementi. **Adaptive resolution simulation of liquid water**. *J. Phys. Cond. Matt.*, **19**, 292201 (2007).

[47] L. Delle Site. **Some fundamental problems for an energy-conserving adaptive-resolution molecular dynamics scheme**. *Phys. Rev. E*, **76**, 047701 (2007).

[48] S. Poblete, M. Praprotnik, K. Kremer and L. Delle Site. **Coupling different levels of resolution in molecular simulations**. *J. Chem. Phys*, **132**, 114101 (2010).

[49] P. A. M. Dirac. *The principles of Quantum Mechanics*. Oxford University Press, 1 edition, (1930).

[50] P A. M. Dirac. **The Lagrangian in Quantum Mechanics**. *Physikalische Zeitschrift der Sowjetunion*, **3**, 64 (1933).

[51] R. P. Feynman. **Space-Time Approach to Non-Relativistic Quantum Mechanics**. *Rev. Mod. Phys.*, **20**, 367 (1948).

[52] R. J. Rivers. *Path integral methods in quantum field theory.* Cambridge University Press, 1 edition, (1987).

[53] D. M. Ceperley. **Path integrals in the theory of condensed helium**. *Rev. Mod. Phys.*, **67**, 279 (1995).

[54] G. J. Papadopoulos and J. Thomchick. **On a path integral having application in polymer physics**. *J. Phys. A: Math. Gen.*, **10**, 1115 (1997).

[55] R. Martonak, W. Paul, K. Binder. **Orthorhombic Phase of Crystalline Polyethylene: A Constant Pressure Path integral Monte Carlo Study**. *Phys. Rev. E*, **57**, 2425 (1998).

[56] V. Linetsky. **The Path Integral Approach to Financial Modeling and Options Pricing**. *Computational Economics*, **11**, 129 (1998).

[57] H. Gai and B. C. Garrett. **Path Integral Calculations of the Free Energies of Hydration of Hydrogen Isotopes (H, D, and Mu)**. *J. Phys. Chem.*, **98**, 9642 (1994).

[58] L. M. Sesé. **Thermodynamic and structural properties of the path-integral quantum hard-sphere fluid**. *J. Chem. Phys.*, **108**, 9086 (1998).

[59] C. P. Herrero and R. Ramirez. **Structural and thermodynamic properties of diamond: A path-integral Monte Carlo study**. *Phys. Rev. B*, **63**, 024103 (2000).

[60] J. W. Lawson. **Path integrals for the quantum microcanonical ensemble**. *Phys. Rev. E*, **61**, 61 (2000).

[61] M. E. Tuckerman, B. J. Berne, G. J. Martyna, and M. L. Klein. **Efficient molecular dynamics and hybrid Monte Carlo algorithms for path integrals**. *J. Chem. Phys.*, **99**, 2796 (1993).

[62] G. J. Martyna, A. Hughes, and M. E. Tuckerman. **Molecular dynamics algorithms for path integrals at constant pressure**. *J. Chem. Phys.*, **110**, 3275 (1999).

[63] Q. Wang, J. K. Johnson and J. Q. Broughton. **Path integral grand canonical Monte Carlo**. *J. Chem. Phys.*, **107**, 5108 (1997).

[64] L. Schulman. *Techniques and Application of Path Integration.* John WIley & Sons, New York, 1 edition, (1981).

[65] R. P. Feynman and A. R. Hibbs. *Quantum Mechanics and Path Integrals.* New York: McGraw-Hill, 1 edition, (1965).

[66] H. Kleinert. *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets.* World Scientific Publishing Company, 4 edition, (2004).

[67] M. Parrinello, A. Rahman. **Study of an F center in molten KCl**. *J. Chem. Phys.*, **80**, 860 (1984).

[68] S. Nose. **A unified formulation of the constant temperature molecular dynamics methods**. *J. Chem. Phys.*, **81**, 511 (1984).

[69] W. G. Hoover. **Canonical dynamics: Equilibrium phase-space distributions**. *Phys. Rev. A*, **31**, 1695 (1985).

[70] G. J. Martyna, M. L. Klein and M. Tuckerman. **Nosé–Hoover chains: The canonical ensemble via continuous dynamics**. *J. Chem. Phys.*, **97**, 2635 (1992).

[71] M. E. Tuckerman, Y. Liu, G. Ciccotti and G. J. Martyna. **Non-Hamiltonian molecular dynamics: Generalizing Hamiltonian phase space principles to non-Hamiltonian systems**. *J. Chem. Phys.*, **115**, 1678 (2001).

[72] I. R. Craig and D. E. Manolopoulos. **Quantum statistics and classical mechanics: Real time correlation functions from ring polymer molecular dynamics**. *J. Chem. Phys.*, **121**, 3368 (2004).

[73] J. Cao and G. A. Voth. **A new perspective on quantum time correlation functions**. *J. Chem. Phys.*, **99**, 10070 (1993).

[74] H. C. Andersen. **Molecular dynamics simulations at constant pressure and/or temperature**. *J. Chem. Phys.*, **72**, 2384 (1980).

[75] G. S. Grest and K. Kremer. **Molecular-dynamics simulation for polymers in the presence of a heat bath**. *Phys. Rev. A*, **33**, 3628 (1986).

[76] T. Soddemann, B. Dünweg, and K. Kremer. **Dissipative particle dynamics: A useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations**. *Phys. Rev. E*, **4**, 046702 (2003).

[77] C. Junghans, M. Praprotnik and K. Kremer. **Transport properties controlled by a thermostat: An extended dissipative particle dynamics thermostat**. *Soft Matter*, **68**, 156 (2008).

[78] R. W. Hall and B. J. Berne. **Nonergodicity in path integral molecular dynamics**. *J. Chem. Phys.*, **81**, 3641 (1984).

[79] M. E. Tuckerman et al. *J. Chem. Phys.*, **99**, 2796 (1993).

[80] E. L. Pollock and D. M. Ceperley. **Simulation of quantum many-body systems by path-integral methods**. *Phys. Rev. B*, **30**, 2555 (1984).

[81] M. Doi and S.F. Edwards. **The theory of polymer dynamics**. Oxford University press, 2 edition, (1998).

[82] G. Tsolou, N. Stratikis, C. Baiget, P.S. Stephanou, and V.G. Mavrantzas. **Melt Structure and Dynamics of Unentangled Polyethylene Rings: Rouse Theory, Atomistic Molecular Dynamics Simulation, and Comparison with the Linear Analogues**. *Macromolecules*, **43**, 10692 (2010).

[83] D. M. Kernan, G. Ciccotti and R. Kapral. **Trotter-Based Simulation of Quantum-Classical Dynamics**. *J. Phys. Chem. B*, **112**, 424 (2008).

[84] C. Junghans, M. Praprotnik and L. Delle Site. **Adaptive Simulation Method**. In S. Bluegel J. Grotendorst, N. Attig and D. Marx, editors, *Multiscale Simulation Methods in Molecular Sciences*, page 77. NIC Series, 1 edition, February (2009).

[85] M. G. Guthrie, A. D. Daigle and M. R. Salazar. **Properties of a Method for Performing Adaptive, Multilevel QM Simulations of Complex Chemical Reactions in the Gas-Phase**. *J.Chem.Theor.Comp.*, **6**, (2010).

[86] A. Heyden, H. Lin and D. G. Truhlar. **Adaptive partitioning in combined quantum mechanical and molecular mechanical calculations of potential energy functions for Multiscale simulations**. *J.Phys.Chem. B*, **111**, 2231 (2007).

[87] R. Bulo, B. Ensing, J. Sikkema and L. Visscher. **Toward a Practical Method for Adaptive QM/MM Simulations**. *J. Chem. Theory Comput.*, **5**, 2212 (2009).

[88] A. B. Poma and L. Delle Site. **Classical to path-integral sdaptive resolution in molecular simulation: Towards a smooth quantum-classical coupling**. *Phys. Rev. Lett.*, **104**, 250201 (2010).

[89] B. R. Brooks et al. **CHARMM: The Biomolecular Simulation Program**. *J. Comp. Chem.*, **30**, 1545 (2009).

[90] W. F. van Gunsteren et al. **The GROMOS software for biomolecular simulation: GROMOS05**. *J. Comput. Chem.*, **26**, 1719 (2005).

[91] A. de la Lande, N. S. Babcock, J. Rezac, B. C. Sanders, D. R. Salahub **Surface residues dynamically organize water bridges to enhance electron transfer between proteins**. *Natl. Acad. Sci. USA*, **107**, 11799 (2010).

[92] S. Habershon, T. E. Markland and D. E. Manolopoulos. **Competing quantum effects in the dynamics of a flexible water model**. *J. Chem. Phys.*, **131**, 024501 (2009).

[93] M. Praprotnik, L. Delle Site and K. Kremer. **Adaptive resolution scheme for efficient hybrid atomistic-mesoscale molecular dynamics simulations of dense liquids**. *Phys. Rev. E*, **73**, 066701 (2006).

[94] N. Bohr. **Atomic Structure**. *Nature*, **107**, 104 (1921).

[95] H. E. White. *Introduction to atomic Spectra*. McGraw-Hill, New York, 1 edition, (1934).

[96] L. Pauling. **The nature of the chemical bond. Application of results obtained from the quantum mechanics and from the theory of paramagnetic suceptibility to the structure of molecules**. *J. Am. Chem. Soc.*, **53**, 1367 (1931).

[97] H. M. James and A. S. Coolidge. **The Ground State of the Hydrogen Molecule**. *J. Chem. Phys.*, **1**, 825 (1933).

[98] H. K. Mao and R. J. Hemley. **Ultrahigh-pressure transitions in solid hydrogen**. *Rev. Mod. Phys.*, **66**, 671 (1994).

[99] I. I. Mazin, R. J. Hemley, A. F. Goncharov, M. Hanfland and H. K. Mao. **Quantum and Classical Orientational Ordering in Solid Hydrogen**. *Phys. Rev. Lett.*, **78**, 1066 (1997).

[100] D. M. Ceperley and B. J. Alder. **Ground state of solid hydrogen at high pressures**. *Phys. Rev. B*, **36**, 2092 (1987).

[101] T. Cui, Y. Ma and G. Zou. **Crystal structure of solid molecular hydrogen under high pressures**. *J. Phys.: Condens. Matter*, **14**, 10901 (2002).

[102] G. E. Ballester, D. K. Sing and F. Herbert. **The signature of hot hydrogen in the atmosphere of the extrasolar planet HD 209458b**. *Nature*, **445**, 511 (2007).

[103] K. T. Delaney, C. Pierleoni and D. M. Ceperley. **Quantum Monte Carlo simulation of the high-pressure molecular-atomic crossover in fluid hydrogen**. *Phys. Rev. Lett.*, **97**, 235702 (2006).

[104] M. A. Morales, C. Pierleoni, E. Schwegler and D. M. Ceperley. **Evidence for a first-order liquid-liquid transition in high-pressure hydrogen from ab initio simulations**. *Natl. Acad. Sci. USA*, **107**, 12799 (2010).

[105] K. A. Gernoth, T. Lindenau and M. L. Ristig. **Screening of particle exchange in quantum boltzmann liquids**. *Phys. Rev. B*, **75**, 174204 (2007).

[106] I. F. Silvera and V. V. Goldman. **The isotropic intermolecular potential for H2 and D2 in the solid and gas phases**. *J. Chem. Phys.*, **69**, 4209 (1978).

[107] D. Scharf, G. J. Martyna and M. L. Klein. **Structure and energy of fluid para-hydrogen**. *Low. Temp. Phys.*, **19**, 364 (1993).

[108] M. Zoppi, M. Neumann and M. Celli. **Microscopic structure factor of liquid para-hydrogen**. *Phys. Rev. B*, **65**, 092204 (2002).

[109] Q. Wang and J. K. Johnson. **Hydrogen adsorption on graphite and in carbon slit pores from path integral simulations**. *Mol. Phys.*, **95**, 299 (1998).

[110] T. F. Miller and D. E. Manolopoulos. **Quantum diffusion in liquid para-hydrogen from ring-polymer molecular dynamics**. *J. Chem. Phys.*, **122**, 184503 (2005).

[111] S. Izvekov and G. A. Voth. **Fast centroid molecular dynamics: A force-matching approach for the predetermination of the effective centroid forces**. *J. Chem. Phys.*, **122**, 054105 (2005).

[112] J. A. Poulsen, G. Nyman and P. J. Rossky. **Quantum Diffusion in Liquid Para-hydrogen: An Application of the Feynman−Kleinert Linearized Path Integral Approximation**. *J. Phys. Chem. B*, **108**, 19799 (2004).

[113] B. J. Berne and D. Thirumalai. **On the Simulation of Quantum Systems: Path Integral Methods**. *Ann. Rev. Phys. Chem.*, **37**, 401 (1986).

[114] V. A. Harmandaris, V. G. Mavrantzas and D. N. Theodorou. **Atomistic Molecular Dynamics Simulation of Polydisperse Linear Polyethylene Melts**. *Macromolecules*, **31**, 7934 (1998).

[115] V. A. Harmandaris, N. P. Adhikari, N.van der Vegt and K. Kremer. **Hierarchical Modeling of Polystyrene: From Atomistic to Coarse-Grained Simulations**. *Macromolecules*, **39**, 6708 (2006).

[116] Wolfram Research, Inc., Mathematica, Version 7, http://www.wolfram.com/.

[117] B. Dünweg. **Langevin Methods**. In B. Dünweg, A. I. Milchev, and D. P. Landau, editors, *Computer Simulations of Surfaces and Interfaces: Proceedings of the NATO Advanced Study Institute, Albena, Bulgaria, from 9 to 20 September 2002*, page 77. Springer Netherlands, 1 edition, February (2004).

[118] H. Risken. *The Fokker-Planck Equation.* Berlin: Springer Verlag, 2 edition, (1989).