UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE FÍSICA GLEB WATAGHIN

RODOLFO PAULA LEITE

Calculation of thermal properties using molecular simulation

Cálculo de propriedades térmicas utilizando simulação molecular

Campinas

2020

Rodolfo Paula Leite

Calculation of thermal properties using molecular simulation

Cálculo de propriedades térmicas utilizando simulação molecular

Tese apresentada ao Instituto de Física Gleb Wataghin da Universidade Estadual de Campinas para obtenção do título de Doutor em Ciências.

Thesis presented to the Gleb Wataghin Institute of Physics of the University of Campinas to obtain the Doctor degree in Science.

Supervisor/Orientador: Prof. Dr. Maurice de Koning

Este trabalho corresponde à versão final da Tese defendida pelo aluno Rodolfo Paula Leite e orientada pelo Prof. Dr. Maurice de Koning.

Campinas

2020

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Física Gleb Wataghin
Lucimeire de Oliveira Silva da Rocha - CRB 8/9174

L536c
Leite, Rodolfo Paula, 1991-
Calculation of thermal properties using molecular simulation / Rodolfo Paula Leite. – Campinas, SP : [s.n.], 2020.

Orientador: Maurice de Koning.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Física Gleb Wataghin.

1. Cálculos de energia livre. 2. Simulação molecular. 3. Fluídos. 4. Diagramas de fase. I. Koning, Maurice de, 1969-. II. Universidade Estadual de Campinas. Instituto de Física Gleb Wataghin. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Cálculo de propriedades térmicas utilizando simulação molecular
Palavras-chave em inglês:
Free energy calculations
Molecular simulation
Fluids
Phase diagrams
Área de concentração: Física
Titulação: Doutor em Ciências
Banca examinadora:
Maurice de Koning [Orientador]
Alex Antonelli
Ricardo Luís Doretto
Caetano Rodrigues Miranda
José Pedro Rino
Data de defesa: 08-05-2020
Programa de Pós-Graduação: Física

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0001-8479-348X
- Currículo Lattes do autor: http://lattes.cnpq.br/7450721054523054

MEMBROS DA COMISSÃO JULGADORA DA TESE DE DOUTORADO DE **RODOLFO PAULA LEITE – RA 095060** APRESENTADA E APROVADA AO INSTITUTO DE FÍSICA "GLEB WATAGHIN", DA UNIVERSIDADE ESTADUAL DE CAMPINAS, EM 08 / 05 / 2020.

COMISSÃO JULGADORA:

- Prof. Dr. Maurice de Koning – Orientador – DFMCIFGW/UNICAMP

- Prof. Dr. Alex Antonelli – DFMC/IFGW/UNICAMP

- Prof. Dr. Ricardo Luís Doretto - DFMC/IFGW/UNICAMP

- Prof. Dr. Caetano Rodrigues Miranda – IF/USP

- Prof. Dr. José Pedro Rino - DF/UFSC

OBS.: Ata da defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

CAMPINAS

2020

Este trabalho é dedicado a meus pais, Marcos e Edna, minha querida Stephanie e o amigo Marcelo Previatelli (In memoriam).

Acknowledgements

Em primeiro lugar, gostaria de agradecer a todas as pessoas que, de alguma forma, participaram da minha trajetória dentro da Unicamp, local no qual passei um terço de toda a minha vida. Tentar citar o nome de todos poderá resultar em grandes injustiças, então, se seu nome não for mencionado na sequência, meu muito obrigado.

Aos meus pais, Edna e Marcos, agradeço todo o amor, carinho e apoio incondicional dados a mim durante toda a minha vida. Sem vocês eu não seria absolutamente nada. Minha eterna gratidão.

À minha namorada Stephanie Manchado pelo amor, carinho e companheirismo de todos esses anos. Sem o seu apoio, esse trabalho não teria sido realizado.

Agradeço também de forma muito especial o Prof. Dr. Maurice de Koning pela orientação nesse trabalho. Me considero um privilegiado por poder conviver e aprender com um excelente físico, professor e pai. Obrigado pela paciência nos meus momentos de dúvidas, confusão, insegurança e, principalmente, na reta final de confecção dessa tese. Independentemente do caminho em que irei seguir, o considero como exemplo de profissional no qual me espelho e pretendo me tornar um dia. *Hartelijk dank voor uw inzet, geduld en vriendschap in dit werk.*

Aos professores Alex Antonelli, Eduardo Miranda e José Antonio Brum pelos excelentes e inspiradores cursos de pós-graduação ministrados durante esse período. Não posso deixar de citar o curso de "Teoria Quântica de Sistemas de Muitos Corpos" por todas as noites mal dormidas e pesadelos com os belíssimos diagramas de Feynman.

Agradeço a oportunidade de participar do programa de aperfeiçoamento didático (PED) ministrando as disciplinas F-328 e F-329, sob a supervisão dos professores José Antônio Roversi e Ricardo Urbano.

Aos meus amigos "Unicampers", companheiros do grupo de pesquisa e, principalmente, da Sala de Estudos por todos os momentos de convivência, correrias para terminar as listas de exercícios e estudar para as provas, todas as reuniões do café pós-almoço e momentos inesquecíveis no restaurante universitário (Bandeco). Em especial os amigos: Thiago, Helder, Lisan, Luiz Filipe, William, Thaís, Rene, Elohim, Lucas Madeira, Marina Lima (IMECC), Helena (IB), Luiz Guilherme (FEQ), Carol, Vitão, Murilo, Arthur, Alê, Adriana, Ingrid, Pedro, Jéssica, Danilo, Vitor Cândido e Rodrigo Freitas (Stanford).

Em nome do Sr. Oswaldo e da Sra. Emilia, meu muito obrigado a todos os funcionários do Departamento de Física da Matéria Condensada.

Gostaria de agradecer os professores Marcus Aguiar e Amir Caldeira pela presença no meu Exame de Qualificação de Doutorado. Agradeço também os professores Alex Antonelli e Francisco Rouxinol por comporem a banca do meu Seminário de Pré-Requisito de Defesa de Doutorado, principalmente pelas críticas e sugestões endereçadas a esse trabalho. Por fim, agradecer a disponibilidade dos professores Alex Antonelli, Ricardo Luís Doretto, Caetano Rodrigues Miranda e José Pedro Rino por comporem a banca julgadora desse trabalho em meio à pandemia de Covid-19 (Corona vírus).

À empresa Vitória locadora de veículos (Thiago, Maria e Oswaldo) responsáveis pela meu deslocamento entre minha casa (Indaiatuba) e a Unicamp (Campinas) durante todos esses anos.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, sob o número de processo: 1504593/2015.

Universidade Estadual de Campinas - UNICAMP
Campinas, SP - Brasil
Maio de 2020
Rodolfo Paula Leite.

"One of the principal objects of theoretical research is to find the point of view from which subject appears in its greatest simplicity."
Josiah Willard Gibbs

Resumo

A partir da utilização de técnicas de simulação de dinâmica molecular e integração termodinâmica de sistemas fora do equilíbrio, é possível calcular a energia livre de Helmholtz das fases cristalinas cúbica de corpo centrado (bcc), cúbida de face centrada (fcc), rede hexagonal compacta (hcp) e da fase fluida do modelo de Uhlenbeck-Ford (UF) e usar esses resultados para construir seu diagrama de fases. Foi encontrado que as estruturas bcc e fcc são as únicas fases cristalinas termodinamicamente estáveis no diagrama de fases. Além disso, é reportada a existência de duas sequências de transições reentrantes em função da densidade de número de partículas, sendo uma sucessão fluída-bcc-fluida e outra bcc-fcc-bcc próxima do ponto triplo. Para mais, é possível observar fortes semelhanças com o comportamento de fase de outros sistemas suaves e puramente repulsivos, como os potenciais de núcleo gaussiano (GCM), lei de potência inversa e de Yukawa. Em particular, observamos que o pontro triplo entre as fases fcc-bcc-fluído e os limites de fase em sua vizinhança estão de acordo com a previsão fornecida pelo recentemente proposto princípio de estados correspondentes. A particularmente forte semelhança entre o comportamento dos modelos UF e GCM também é discutida neste trabalho.

Na sequência, é apresentado um guia para calcular energias livres absolutas de fluídos clássicos utilizando técnicas de energia-livre fora do equilíbrio dentro do código `LAMMPS`. A principal abordagem adotada é baseada na construção de um caminho termodinâmico conectando o fluído de interesse para um dos variantes atômico ou molecular do modelo de UF como sistema de referência. Esses sistemas de referência são descritos matematicamente em detalhes além de sua implementação e disponibilização de códigos-fonte e arquivos auxiliares no pacote `LAMMPS`. Além disso, diversas aplicações envolvendo sistemas caracterizados por interações fundamentalmente distintas são ilustrados: dois modelos atômicos diferentes (água mono-atômica e liga binária líquida de cobre-zircônico) e três modelos moleculares para água, dois rígidos (TIP4P e SPC/E) e um flexível (q-SPC/E). Para os sistemas moleculares foi desenvolvido uma referência baseada no modelo UF na qual sua energia livre é dada pela soma de duas contribuições: uma parte intermolecular descrita pela energia livre já conhecida do modelo UF e uma parte intramolecular que é determinada de forma analítica. Portanto, as ferramentas descritas nesse trabalho fornecem uma plataforma na qual as energias livres de sistemas na fase fluida possam ser calculadas de maneira fácil e eficiente usando o código `LAMMPS`, permitindo obtenção de qualquer outra quantidade térmica de interesse.

Palavras-chave: Cálculos de energia livre. Simulação molecular. Fluidos. Diagrama de fase. LAMMPS.

Abstract

Using molecular dynamics simulations and nonequilibrium thermodynamic-integration techniques we compute the Helmholtz free energies of the body-centered-cubic (bcc), face-centered-cubic (fcc), hexagonal close-packed (hcp) and fluid phases of the Uhlenbeck-Ford model (UF) and use the results to construct its phase diagram. We find that the bcc and fcc are the only thermodynamically stable crystalline phases in the phase diagram. Furthermore, we report the existence of two reentrant transition sequences as a function of the number density, one featuring a fluid-bcc-fluid succession and another displaying a bcc-fcc-bcc sequence near the triple point. Moreover, we find strong resemblances to the phase behavior of other soft, purely repulsive systems such as the Gaussian-core model (GCM), inverse-power-law and Yukawa potentials. In particular, we find that the fcc-bcc-fluid triple point and the phase boundaries in its vicinity are in good agreement with the prediction supplied by a recently proposed corresponding-states principle. The particularly strong resemblance between the behavior of the UFM and GCM models are also discussed.

Subsequently, we present a guide to compute the absolute free energies of classical fluids using nonequilibrium free-energy techniques within the `LAMMPS` code. The main approach is based on the construction of a thermodynamic path connecting the fluid of interest to either atomic or molecular variants of the UF model as reference systems. We describe these reference systems in detail, discuss their implementation in the `LAMMPS` package and make available source code, scripts as well as auxiliary files. As an illustration we detail a number of distinct applications, involving systems characterized by fundamentally different interactions. In addition to two different atomic models (mW water and the MEAM-2NN CuZr liquid binary alloy), we consider three molecular models for water, two of them rigid (TIP4P and SPC/E) and one flexible (q-SPC/Fw). For the molecular models we develop UF-based reference systems for which the free energies are given by a sum of two contributions: an intermolecular part described by the known UF free energy and an intramolecular contribution that can be determined analytically. The tools described in this paper provide a platform on which fluid-phase free energies can be easily and efficiently computed using the `LAMMPS` code, allowing to obtain any other thermal quantity of interest.

Keywords: Free-energy calculation. Molecular simulation. Fluids. Phase diagram. LAMMPS.

List of Figures

List of Tables

List of abbreviations and acronyms

2NN	Second Nearest-Neighbor
AS	Adiabatic Switching
bcc	Body-Centered-Cubic
cd	Cubic Diamond
DFT	Density-Functional Theory
fcc	Face-Centered-Cubic
FL	Frenkel-Ladd
Fw	Flexible-water
GCM	Gaussian-Core Model
hcp	Hexagonal Close-Packed
hd	Hexagonal Diamond
HI	Hamiltonian Interpolation
IPL	Inverse Power Law
LAMMPS	Large-scale Atomic/Molecular Massively Parallel Simulator
LJ	Lennard-Jones
mW	Monoatomic Water
MEAM	Modified Embedded-Atom Model
MD	Molecular Dynamics
NEHI	Nonequilibrium Hamiltonian Interpolation
NERS	Nonequilibrium Reversible Scaling
NPT	Isothermal-Isobaric ensemble
NVT	Canonical ensemble
PPPM	Particle-Particle/Particle-Mesh

PBC	Periodic Boundary Condition
PEF	Potential-Energy Function
RS	Reversible Scaling
Rw	Rigid-water
SPC	Simple Point-Charge
TI	Thermodynamic Integration
TIP4P	Transferable Intermolecular Potential with 4 Points
UF	Uhlenbeck-Ford
UFM	Uhlenbeck-Ford Model
YK	Yoshida and Kamakura

Contents

1 Introduction

With the spectacular development of computer hardware over the past 50 years, computational tools have become an integral part of scientific investigation in fields as diverse as physics, chemistry, materials science, engineering and biology. A particularly important subset of these tools concerns computational simulations of systems at the atomistic scale. The purpose of such simulations is to, based on a specified interaction model between the atomic species, gain an understanding of the fundamental unit processes that are at the heart of the macroscopic properties of a system of interest. Given that such techniques allow one to observe the detailed motion of atoms and molecules, they are also referred to as tools for computational microscopy.

One of the main areas of application of such tools is the description and prediction the phase behavior of condensed matter systems. Specifically, this task amounts to determine the phase diagram of the substance of interest, which, in a plot in the pressure-temperature plane, consists of a set of lines that delimit the boundaries of thermodynamic stability between different phases. The melting curve, describing the boundary between liquid and solid phases, is an example of such a line and it describes the locus of temperature-pressure conditions in which the liquid and solid phases can coexist. The condition that must be satisfied for the coexistence is that the molar Gibbs free energy of the two phases are equal. Consequently, the determination of phase diagrams using atomistic simulation techniques requires methods to compute Gibbs free energies. The present Thesis is focused on this important task, developing methods and open source computational codes for this purpose.

All results described in this Thesis have been obtained using the molecular dynamics (MD) method, which is a simulation approach that implements the tools of classical statistical mechanics. Before presenting the results in the following chapters, we first describe the main ingredients of the MD technique, followed by a summary of the scope of the Thesis.

1.1 Molecular dynamics

Widely used in the research areas of physics, chemistry, engineering, biology and pharmacy industry [6, 7, 8, 9, 10, 11, 12], molecular dynamics (MD) is an atomistic simulation technique that analyzes the physical movement of atoms and/or molecules during specified intervals of time. In general, this dynamic and evolutionary analysis of the system is obtained by the resolution of Newtonian equations of motion. Once the particle trajectories have been determined, it is possible to obtain, on average, several physical quantities, both dynamic and static in nature [6, 13]. Due to the typically large number of atoms or molecules and the complex interatomic and molecular force fields used in describing the interactions between them, these equations cannot be solved analytically, leading to the need to use numerical methods and large-scale computing [14]. In the following, we briefly discuss the principal components of an MD simulation.

1.1.1 Force Field

The main ingredient in an MD simulation consists of the description of the interactions between the atoms and/or molecules that constitute the system of interest. Generally, these interactions are described in terms of a conservative force field represented by a potential-energy function (PEF)

$$V = V(\{\mathbf{r}_i\}), \tag{1.1}$$

that describes the potential energy V of the system as a function of the set of particle positions $\{\mathbf{r}_i\}$, where $i = 1 \cdots N$, with N the total number of particles.

There are basically two approaches toward constructing force field. The first is the so-called semi- empirical approach, which involves defining a specific explicit functional form containing a set of adjustable parameters. Their values are then chosen so as to best reproduce a set of reference values that can include experimental values and/or results from first-principles quantum-mechanical calculations. The chosen mathematical functional form is often inspired by physical insight. For instance, in the development of PEFs for silicon, the functional form explicitly incorporates the typical tetrahedral bonding structures that are characteristic in both its crystalline and liquid forms [15, 16]. One of the main issues involving semi-empirical force fields concerns their degree of transferability. Given that the model parameters are adjusted according to a limited set of reference values, there are no *a priori* guarantees that it can accurately describe properties that have not been included in this set.

The second approach provides a systematic improvement in this regard. Instead of relying on an explicit mathematical functional form, it is based on quantum-mechanical electronic-structure calculations such as density-functional theory (DFT) and Hartree-Fock [17], that numerically compute the potential energy and corresponding conservative forces for a given particle configuration. The advantage of this approach is that it is based on a quantum-mechanical description of the electronic system, in principle providing accurate energetics regardless of the type of configuration under consideration (e.g., crystalline, liquid-like, etc.). The biggest disadvantage of such first-principles approaches is their computational cost, which is often 2 to 3 orders of magnitude larger than those for semi-empirical models.

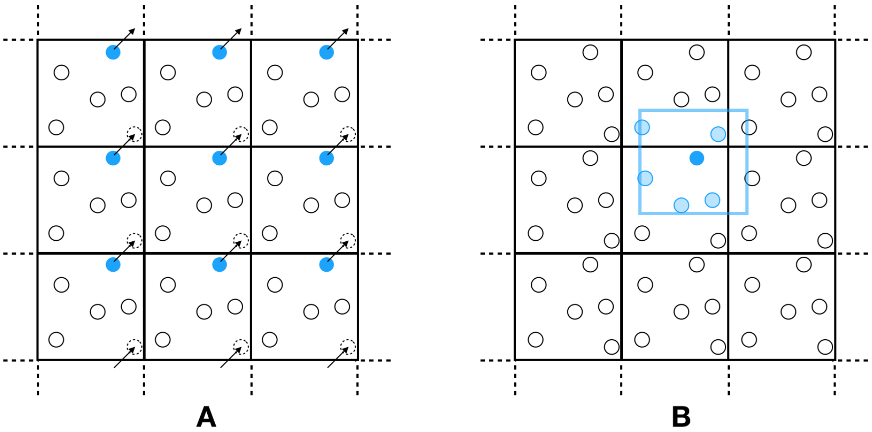

Figure 1 – (Color) Illustration of (A) PBC procedure. (B) Minimum image convention.

1.1.2 Boundary Conditions

When we imagine a “real” system, we might be talking about a system consisting of, at least, 6×10^{23} atoms or molecules. However, the length-scale of MD simulations is limited and usually consider at most $\sim 10^6$ atoms or molecules. Moreover, when putting these particles in a simulation box, a large fraction of them “feels” the presence of or are located at the surface. However, in many cases such surface effects are undesirable, for instance when one intends to study the bulk properties of a given substance. In this case, it is required to impose boundary conditions so as to eliminate such undesirable surface effects. The standard way to do so is to impose the so-called periodic boundary conditions (PBC) [6]. PBC entails a periodic repetition of the simulation box in all three directions, eliminating free surfaces and producing an effectively "infinite" system, as illustrated in Fig 1. When a particle leaves the box on one side, its re-enters on the other side and replace it. In addition, the particles do not only interact with the particles in the "primary" simulation box, but also with all their images. Usually, the interaction range between two particles is limited to a cut-off radius R_{cut} and, in this case, with the linear dimensions of the primary box being larger than $2\,R_{\text{cut}}$, particles interact only with their closest periodic image, which is referred to as the minimum image convention.

1.1.3 Time Evolution

After the force field and boundary conditions have been specified, one needs to define the equations of motion that describe the dynamical evolution of the system and choose the numerical algorithm for their integration.

Consider a system in which atoms interact through an interatomic potential $V(\{\mathbf{r}_i\})$. The simplest set of equations of motion is that given by Newton’s second law,

$$\frac{d^2\mathbf{r}_i(t)}{dt^2} = \frac{\mathbf{F}_i}{m_i} = -\frac{1}{m_i}\frac{\partial V(\{\mathbf{r}_i\})}{\partial \mathbf{r}_i} \quad ; \quad i = 1, 2, ..., N, \tag{1.2}$$

or, equivalently, its Hamiltonian formulation

$$\begin{aligned}\frac{d\mathbf{r}_i}{dt} &= \frac{\partial H(\{\mathbf{r}_i\},\{\mathbf{p}_i\})}{\partial \mathbf{p}_i}\\ \frac{d\mathbf{p}_i}{dt} &= -\frac{\partial H(\{\mathbf{r}_i\},\{\mathbf{p}_i\})}{\partial \mathbf{r}_i} \qquad ; \qquad i = 1,2,...,N,\end{aligned} \tag{1.3}$$

where $\mathbf{r}_i$ and $\mathbf{p}_i$ are the position and momentum of particle i and $\mathbf{F}_i$ is the force acting on it due to interactions with the other particles as derived from the interaction potential-energy function. Once the forces on each particle are known, the integration of the equations of motion will allow us to determine the microstates of the system in successive instant from an initial microstate $(\{\mathbf{r}_i(0)\},\{\mathbf{p}_i(0)\})$.

According to the Hamiltonian formulation of classical mechanics [18], these equations of motion have two properties: they respect the Louville theorem and are temporally reversible. Therefore, numerical integrators should incorporate these properties so as to give trajectories that are as close as possible to their "exact" counterparts. Over the years, several classes of such so-called symplectic integration algorithms [13] have been developed. One of the most frequently used integrators, also in this Thesis, is the velocity-Verlet algorithm [19], which is defined by the following update sequence:

1. $\mathbf{p}_i(t + \frac{\Delta t}{2}) = \mathbf{p}_i(t) + \frac{\mathbf{F}_i(t)}{m}\frac{\Delta t}{2}$
2. $\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \frac{\mathbf{p}_i(t + \frac{\Delta t}{2})}{m}\Delta t$
3. Compute $\mathbf{F}_i(t + \Delta t)$ using $\mathbf{r}_i(t + \Delta t)$
4. $\mathbf{p}_i(t + \Delta t) = \mathbf{p}_i(t + \frac{\Delta t}{2}) + \frac{\mathbf{F}_i(t + \Delta t)}{m}\frac{\Delta t}{2}$
5. Return to 1 with $t \equiv t + \Delta t$

The trajectories produced by the equations of motion in the Eqs. (1.2) and (1.3) sample configurations from the microcanonical ensemble [6], i.e., those with a constant total energy, volume and number of particles. However, if we want to consider the system under conditions of fixed temperature (canonical ensemble) or temperature and pressure (isobaric-isothermic ensemble) the equations of motion need to be adapted. While there is a large variety of such dynamical equations of motion [13], a particularly useful dynamics, widely used in this Thesis, is that provided by the Langevin thermostat [20] which, under conditions of fixed volume, guaranteed to the generate the canonical ensemble. The corresponding equations of motion are given by

$$\begin{aligned}\frac{d\mathbf{r}_i}{dt} &= \frac{\partial H(\{\mathbf{r}_i\},\{\mathbf{p}_i\})}{\partial \mathbf{p}_i}\\ \frac{d\mathbf{p}_i}{dt} &= -\frac{\partial H(\{\mathbf{r}_i\},\{\mathbf{p}_i\})}{\partial \mathbf{r}_i} + \mathbf{f}_i(\mathbf{p}_i, t) \qquad ; \qquad i = 1,2,...,N,\end{aligned} \tag{1.4}$$

where $\mathbf{f}_i(\mathbf{p}_i, t)$ is an external force that simulates an interaction with a heat bath at temperature T. Explicitly, this external force is given by

$$\mathbf{f}_i(\mathbf{p}_i, t) = -\gamma \mathbf{p}_i(t) + \sqrt{2m\gamma k_B T}\mathbf{R}_i(t), \tag{1.5}$$

where γ is a friction coefficient, m is the particle mass and $\mathbf{R}_i(t)$ is a Gaussian white noise that obeys the properties

$$\langle \mathbf{R}_i(t) \rangle = 0 \qquad \text{and} \qquad \langle \mathbf{R}_i(t)\mathbf{R}_j(t') \rangle = \delta(t - t')\delta_{ij},$$

which are dictated by the fluctuation-dissipation theorem [13].

1.1.4 Ensemble Averages

As discussed above, a classical state is completely described by the set of positions and velocities of all the atoms in the system. The theory responsible for making the connection between this set of microstates obtained from molecular dynamics and the calculation of physical quantities is statistical mechanics [6]. Suppose that we are interested in measuring some thermodynamic quantity $\mathcal{A}$. The latter specifies that the expectation value for $\mathcal{A}$ is given by an ensemble average, i.e.,

$$\langle \mathcal{A} \rangle = \int \rho(\Gamma)\mathcal{A}(\Gamma)d\Gamma, \tag{1.6}$$

where $\mathcal{A}(\Gamma)$ is the value of observable $\mathcal{A}$ calculated in the microscopic states $\Gamma \equiv (\{\mathbf{r}_i\}, \{\mathbf{p}_i\})$, $\rho(\Gamma)$ is the probability density of a given statistical ensemble and $d\Gamma = d\mathbf{r}^N d\mathbf{p}^N$ is the volume element of the phase space.

The ensemble average $\mathcal{A}(\Gamma)$ cannot be determined directly by MD simulations, however, given that they provide a temporal sequence of microstates. In this fashion, MD simulations can estimate the temporal averages $\overline{\mathcal{A}(\Gamma)}$ defined as

$$\overline{\mathcal{A}(\Gamma)} = \frac{1}{\tau}\int_0^\tau \mathcal{A}(\Gamma(t))dt \approx \frac{1}{M}\sum_i \mathcal{A}(\Gamma(t_i)), \tag{1.7}$$

where the average is calculated over M states along the MD trajectory, usually sampled at time intervals longer than the system correlation time. When comparing equations (1.6) and (1.7), it is evident that they are different. However, it is possible to make a connection between these two averages quantities if one assumes the validity of the ergodic hypothesis which states that, for sufficiently long times, the system trajectory will visit all accessible states of the phase space, such that the ensemble average is equal to the temporal average:

$$\langle \mathcal{A} \rangle = \overline{\mathcal{A}(\Gamma)}. \tag{1.8}$$

Accordingly, MD simulations estimate desired ensemble averages as temporal averages of thermodynamic observables.

1.2 Scope of Thesis

The main objective of this Thesis involves the calculation of a particular class of thermodynamic observables, namely free energies. The main difficulty in this task is that they cannot be expressed in terms of ensemble averages and, accordingly, cannot be directly measured in an MD simulation. As such, their calculation using MD simulations requires specialized methodologies and our main purpose here is to provide the scientific community with systematic and accurate method for free-energy calculations. We focus on classical systems in the fluid phase, for which such computations are most difficult. All the results from molecular simulations that are presented in this Thesis were obtained using the large-scale atomic/molecular massively parallel simulator `LAMMPS` package. The developed computational codes are also available.

In chapter 2, we describe all the methodology and simulation techniques used to compute the Helmholtz free energies of crystalline and fluid phases of the Uhlenbeck-Ford model (UF) and then construct its phase diagram applying the common-tangent procedure. Discussions about the comparison of its phase behaviour with other soft, purely repulsive systems within the context of the corresponding-states-principle [2, 4] are also presented.

In chapter 3, we present our systematic and accurate method for free-energy calculation of fluid-phase systems characterized by fundamentally different interactions: two atomic liquids (single and binary compounds) and three molecular liquids (rigid and flexible) that mimics water, focusing on the implementation in the widely used `LAMMPS` package. For each considered system we describe in details all the methodology and reference systems based on UF model that was used during the molecular simulations. All the results for each kind of system were compared with other calculations reported in the Literature. In order to obtain these results, many different new functionalities were programmed in such a way that they can be used within the `LAMMPS` package.

At the end, in chapter 4, we summarize all the main results and conclusions obtained in these research and discuss directions and suggestions for future work.

2 Uhlenbeck-Ford Model: Phase diagram and corresponding-states analysis

This chapter was reproduced from the published paper with permission from R. Paula Leite, P. A. Santos-Flórez, and M. de Koning, Phys. Rev. E 96, 032115 (2017) [21]. © 2017 American Physical Society. All rights reserved.

2.1 Introduction

Originally named Gaussian gas, the model devised by Uhlenbeck and Ford [22] was proposed to provide a system for which the virial equation of state could be handled analytically [23]. The model, here referred to as the Uhlenbeck-Ford model (UFM) to avoid confusion with the well-known Gaussian-core model (GCM), is characterized by an ultrasoft, purely repulsive pairwise interaction potential that diverges logarithmically at the origin and features an energy scale that coincides with the thermal energy unit k_BT, with k_B Boltzmann's constant and T the absolute temperature. The particular functional form of the potential permits, in principle, that the virial coefficients and, therefore, the equation of state and excess free energies for the fluid phase be evaluated analytically. Moreover, Baram and Rowlinson [24], based on the convergence properties of the virial expansion, argued that the fluid is the only thermodynamically stable phase for the UFM, regardless of the number density.

These properties motivated a recent investigation into the suitability of the UFM to serve as a reference system for fluid-phase free-energy computations [25]. The results demonstrate that, while the original UFM is too soft to be useful, the scaling of the energy unit k_BT by a factor p provides a convenient reference system for values of $p \sim 50 - 100$. However, although the original model characterized by $p = 1$ exists only in the fluid phase, the phase behavior for $p \neq 1$ is not known. Therefore, to assess the regions of applicability of the UFM as a reference system in free-energy computations for fluid-phase systems, knowledge of its phase diagram is required.

Moreover, such knowledge is also of interest from a more general point of view, for instance in the context of understanding the generic phase behavior of systems characterized by soft purely repulsive pairwise interactions. This class of systems, aside from the UFM, includes the GCM [26, 27, 28, 1, 29], the inverse-power-law (IPL) interaction [30, 31, 32, 33, 34, 35, 36, 37, 28, 38] and the Yukawa-type potentials [39, 40, 41, 42, 43]. Indeed, invoking a corresponding-states principle, Khrapak *et al.* [2] have recently shown that the melting curves of various different pairwise interaction potentials display universal behavior, collapsing to the near vicinity of a master curve when using appropriately selected scaled variables. In particular, based on the same principle, similar universal behavior has been observed for the location of fcc-bcc-fluid triple point for purely repulsive potentials [4]. Interestingly, even other kinds of pairwise interactions such as the Buckingham or exp-6 model [44], the Yoshida and Kamakura (YK) potentials [45, 46, 47] and the modified inverse-power potentials [47] show very similar phase behavior for low densities.

In this setting, the purpose of the present paper is the construction of the UFM's phase diagram. To this end we perform extensive molecular dynamics (MD) simulations using the nonequilibrium free-energy techniques to determine its phase behavior. Computing the Helmholtz free energies of the bcc, fcc, hcp and fluid phases and applying the common-tangent procedure, we construct the phase diagram of the UFM under controlled conditions of the pressure and scaling factor p and locate the fcc-bcc-fluid triple point. Subsequently, the position of the latter is analyzed in further detail by using a finite-size scaling extrapolation procedure. Finally, the results are compared to those obtained for other purely-repulsive pair potentials within the context of the corresponding-states principle.

The remainder of the paper has been organized as follows. In Section 2.2, we describe the UFM and summarize its principal characteristics, followed by a brief outline of the used nonequilibrium free-energy techniques and a description of the MD simulation details. In Section 2.3, we present and discuss the results of the free-energy calculations for the solid and fluid phases as well as the corresponding phase behavior characteristics. Based on the results, we compare the phase diagram topology and location of the triple point for the UFM to those of the GCM, IPL and Yukawa systems within the context of the corresponding-states principle proposed by Khrapak *et al.* [2, 4]. We conclude with a summary in Section 2.4.

2.2 Model and Simulation Methods

2.2.1 The Uhlenbeck-Ford Model

The UFM is defined by the interatomic pair potential

$$U_{\mathrm{UF}}(r) = -\frac{p}{\beta} \ln\left(1 - e^{-(r/\sigma)^2}\right), \tag{2.1}$$

where $\beta \equiv (k_B T)^{-1}$, σ is a length-scale parameter and $p > 0$ is a scaling factor. Fig. 2 displays the UFM potential for a number of different scaling factors p. It is characterized by a smooth and purely repulsive soft-sphere interaction that diverges logarithmically as $r \to 0$ and decays rapidly for increasing distances. Increasing the value of p gives rise to a stronger repulsion.

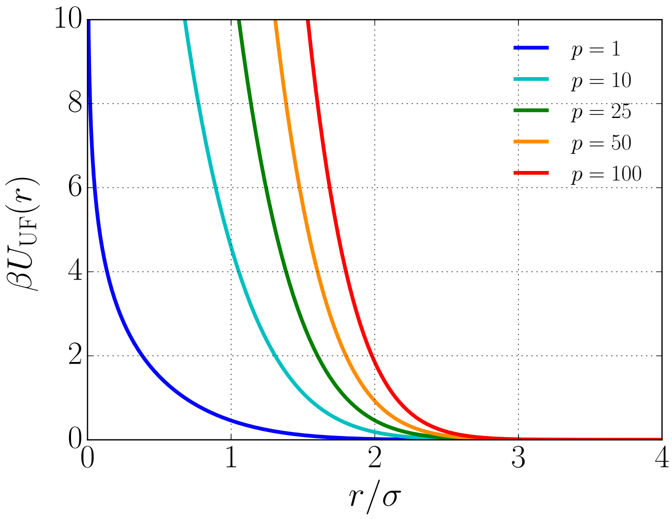

Figure 2 – (Color) Interatomic potential associated with the UFM for different values of p.

Because of the temperature-dependent functional form of the interaction potential in Eq. (2.1), the partition function is independent of temperature and eliminates it as a relevant thermodynamic parameter in the model. On the other hand, given that the parameter p is a scaling factor of the potential energy, it plays the role of an effective inverse temperature T^*, with $T^* \equiv 1/p$. Furthermore, given that the pair potential is of the generic form

$$U(r) = \Phi(r/\sigma), \tag{2.2}$$

it can be shown [23] that the Helmholtz free energy per particle of a collection of N UFM particles confined to a volume V is a function of only adimensional variables and can be expressed in the form

$$\frac{\beta F_{\mathrm{UF}}}{N} = f(\rho\sigma^3, p), \tag{2.3}$$

with $\rho \equiv N/V$ the number density. In particular, as discussed recently [25], the virial-expansion expression for the excess free energy per particle of the fluid phase can be written as

$$\frac{\beta F_{\mathrm{UF}}^{\mathrm{exc}}(x,p)}{N} = \sum_{n=1}^{\infty} \frac{\tilde{B}_{n+1}(p)}{n} x^n, \tag{2.4}$$

where the $\tilde{B}_{n+1}(p)$ are reduced virial coefficients and x is the adimensional variable

$$x \equiv b\rho, \tag{2.5}$$

with $b \equiv \frac{1}{2}(\pi\sigma^2)^{3/2}$. Numerical values of the virial coefficients for several p's can be found in Ref. [25].

2.2.2 Free-energy calculations

For each phase we compute the Helmholtz free energy as a function of p on a predefined grid of x values, the set $\{x_i\}$, with $0.195 \leqslant x_i \leqslant 2.784$. In other words, for each phase we compute a collection of Helmholtz free-energy functions $F_{\mathrm{UF}}(p; x_i)$ of the scaling variable p at fixed x, one for each of the values in the set $\{x_i\}$.

Each free-energy function $F_{\mathrm{UF}}(p; x_i)$ is computed on an interval of p-values $[p_l, p_u]$, where the lower and upper limits, p_l and p_u are chosen so as to guarantee that the phase under consideration is at least metastable. In particular, for the considered crystalline phases, the upper limit is always chosen to be $p_u = 1000$. The values of the lower limit, on the other hand, are adjusted for each value of x so as to remain in the metastable regime. For the fluid phase, the lower limit is always set at $p_l = 50$ while the upper limit was chosen as large as possible while avoiding freezing transitions.

The free-energy functions are computed combining the reversible-scaling (RS) path [48, 49] and the adiabatic switching (AS) approach [50, 51, 52]. In this scheme, the Helmholtz free-energy as a function of p is computed as

$$F_{\mathrm{UF}}(p; x_i) = F_{\mathrm{UF}}(p_0; x_i) + W_{\mathrm{rev}}(p; x_i), \tag{2.6}$$

where $F_{\mathrm{UF}}(p_0; x_i)$ is the Helmholtz free-energy at scaling factor p_0, and W_{rev} is the work done on the system during a reversible process in which the scaling factor in the potential varies from p_0 to p. This process is described by the switching Hamiltonian

$$H(\lambda) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m} - \frac{\lambda}{\beta} \sum_{i>j} \ln[1 - e^{-(r_{ij}/\sigma)^2}], \tag{2.7}$$

where N is the number of particles and the $\mathbf{p}_i$ denote the particle momenta. The reversible work $W_{\mathrm{rev}}(p; x_i)$ is given by the well known relation

$$W_{\mathrm{rev}}(p; x_i) = \int_{p_0}^{p} \left\langle \frac{\partial H}{\partial \lambda} \right\rangle d\lambda, \tag{2.8}$$

where the angular brackets denote canonical equilibrium ensemble averages at fixed values of β, x_i and λ. In the AS approach, the reversible-work integral is estimated along an explicitly time dependent process in which $\lambda = \lambda(t)$ according to

$$W_{\mathrm{dyn}}(p; x_i) = \int_{0}^{t_p} \left(\frac{\partial H}{\partial \lambda} \right) \frac{d\lambda}{dt} dt, \tag{2.9}$$

where t_p is the duration of the dynamical process in which λ varies between p_0 and p and the integral is over instantaneous values of the thermodynamic driving force $\partial H / \partial \lambda$ along the process.

Due to the intrinsic nonequilibrium nature of these processes, W_{dyn} is a stochastic variable whose mean value $\overline{W}_{\mathrm{dyn}}$, determined by averaging over a set of different realizations of the process, overestimates the reversible work, i.e.,

$$\overline{W}_{\mathrm{dyn}} \geqslant W_{\mathrm{rev}}, \tag{2.10}$$

with the equality being valid only in the limit of an infinitely slow, quasistatic process. However, if the process is sufficiently slow for linear-response theory to be valid, the systematic error can be eliminated by combining the results of forward and backward processes [53], i.e.,

$$W_{\mathrm{rev}}(p) = \tfrac{1}{2}\left[\overline{W}_{\mathrm{dyn}}^{p_0\to p} - \overline{W}_{\mathrm{dyn}}^{p\to p_0}\right]. \tag{2.11}$$

In this manner, the reversible work function $W_{\mathrm{rev}}(p;x_i)$ in Eq. (2.8) is determined by carrying out a number of replicas of the nonequilibrium process in which λ varies in the interval $[p_l, p_u]$, in both directions. $W_{\mathrm{rev}}(p;x_i)$ is estimated by combining the forward and backward results as prescribed by Eq. (2.11).

Finally, to find the final Helmholtz free energy $F_{\mathrm{UF}}(p;x_i)$ using Eq. (2.8), one needs to determine the reference value $F_{\mathrm{UF}}(p_0;x_i)$. For the crystalline phases, it is determined for $p_0 = p_u = 1000$ using the standard Frenkel-Ladd (FL) [54, 49] switching path to the Einstein crystal in combination with the AS approach in forward and backward process directions. For the fluid phase, on the other hand, we use the known Helmholtz free-energy values for $p_0 = p_l = 50$ as computed in Ref. [25].

2.2.3 Common-tangent construction

To construct the phase diagram in terms of pressure and scaling factor p, we adopt the common-tangent procedure to determine the values of p at which the pressure and chemical potentials of two phases are equal. This is achieved as follows. Using the Helmholtz free-energies as a function of p for the set of densities $\{x_i\}$, i.e. the functions $F_{\mathrm{UF}}(p;x_i)$ described in Section 2.2.2, we construct a set of Helmholtz free energies $F_{\mathrm{UF}}(v;p_i)$ as a function of the volume per particle $v = b^3/x$ for a particular set $\{p_i\}$ of scaling factor values. Subsequently, each of these curves is adjusted to a third-degree polynomial by means of a least-squares regression analysis, followed by the determination of the common tangent and corresponding per-particle volumes for pairs of phases. This procedure is carried out for a predefined grid of p-values in the set $\{p_i\}$, with $50 \leqslant p_i \leqslant 1000$.

2.2.4 Simulation details

All MD simulations have been carried out using the `LAMMPS` code [55]. The phase diagram reported in the next section was obtained using cubic computational cells containing $\sim 10^4$ particles, subject to standard periodic boundary conditions. A Langevin thermostat with a damping time scale of 100 time steps was used to control the temperature at 1000 K and, choosing a particle mass of 100 grams/mol, the equations of motion were integrated using the velocity-Verlet algorithm with a timestep $\Delta t = 1$ fs . The UF length-scale was set at $\sigma = 1.0$ Å, and a cutoff radius of $r_c = 4.0\,\sigma$ was adopted.

To compute the absolute Helmholtz free-energy for the bcc, fcc and hcp phases for $p_0 = 1000$ we perform 10 independent forward and backward FL switching processes at fixed center of mass, using the polynomial $\lambda(t)$ protocol given in Ref. [49]. To reduce dissipation [49], the force constants of the Einstein crystal are chosen such that the mean-square displacement of the Einstein oscillators closely matches those of the particles in the crystalline UFM phase. Before each FL process, the system is first equilibrated during a time interval $t_{\text{eq}} = 10^5 \Delta t$, followed by the switching procedure carried out in the switching time $t_{\text{sw}} = 10^6 \Delta t$. Following the procedures detailed in Ref. [49] we verified that this value is sufficiently large for linear response theory to hold and the reversible-work estimator Eq. (2.11) to be valid.

A similar procedure was adopted for the RS simulations of the bcc, fcc, hcp and fluid phases. We carried out 10 independent forward and backward RS simulations for each process, using a linear $\lambda(t)$ protocol. The employed equilibration and switching times were the same as those used in the FL switching runs. As for the FL calculations, these choices are sufficient for the processes to be in the linear-response regime.

2.3 Results and Discussion

2.3.1 Phase diagram

Fig. 3 depicts typical results for the Helmholtz free energies $F_{\text{UF}}(p; x_i)$ as a function of the scaling parameter p for given x, displaying the bcc and fcc Helmholtz free energies per particle relative to that of the fluid for $x = 0.2673$. There are three p-values at which the Helmholtz free energies of two phases are equal, one for fluid-bcc pair near $p = 309.9$, another for the fluid-fcc pair near $p = 310.9$ and a third for the bcc and fcc phases close to $p = 316.5$. Furthermore, the areas numbered I through V represent intervals of p-values that describe different regions of thermodynamic phase stability for this particular value of x, as determined using the common-tangent construction further discussed below. In addition to the bcc and fcc phases we also considered other crystalline structures such as the hcp, diamond-cubic and simple-cubic crystals. While the hcp phase is found to be only metastable, the other two were found to be not even mechanically stable under any conditions of x and p.

Subsequently, to determine the coexistence lines of the UFM phase diagram, we apply the common-tangent construction described in Sec. 2.2.3 to the Helmholtz free-energy data for the entire range of p-values between $50 \leqslant p_i \leqslant 1000$. The resulting diagram for controlled conditions of the scaling factor p and the pressure is shown in Fig. 4A. It is characterized by the existence of a single fluid phase and two crystalline forms, i.e., the bcc and fcc phases. Each symbol type represents coexistence points of equal pressure and chemical potentials for specific pairs of phases, with the triangles,

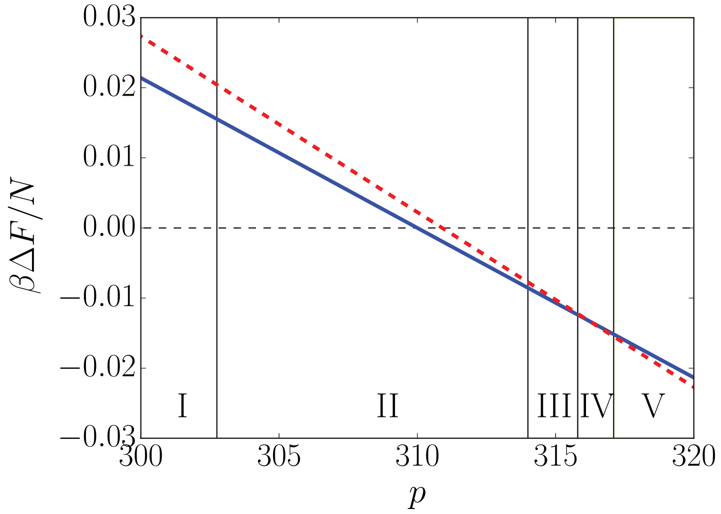

Figure 3 – (Color) Helmholtz free energies per particle of fcc (red dashed line) and bcc phase (blue line) relative to that of the fluid as a function of p for $x = 0.2673$. The areas I, III and V represent the intervals of p-values where, respectively, the fluid, bcc and fcc forms are the thermodynamically stable phases as determined using the common-tangent construction. Regions II and IV depict the intervals of fluid-bcc and bcc-fcc coexistence, respectively (see text). The values of the excess Helmholtz free energies of fluid, bcc and fcc phases at the three crossing points are $\beta F^{\text{exc}}_{\text{fluid}}/N = \beta F^{\text{exc}}_{\text{bcc}}/N = 7.42531 \pm 0.00006$ and $\beta F^{\text{exc}}_{\text{fcc}}/N = 7.42587 \pm 0.00003$ for $p = 309.9$, $\beta F^{\text{exc}}_{\text{fluid}}/N = \beta F^{\text{exc}}_{\text{fcc}}/N = 7.44544 \pm 0.00006$ and $\beta F^{\text{exc}}_{\text{bcc}}/N = 7.44352 \pm 0.00008$ for $p = 310.9$, and $\beta F^{\text{exc}}_{\text{fluid}}/N = 7.56563 \pm 0.00006$ and $\beta F^{\text{exc}}_{\text{bcc}}/N = \beta F^{\text{exc}}_{\text{fcc}}/N = 7.55066 \pm 0.00008$ for $p = 316.2$.

circles and squares representing the fluid-bcc, bcc-fcc and fluid-fcc coexistence, respectively. The three coexistence lines meet at a triple point and its coordinates are estimated by determining the crossing points of cubic spline fits to the fluid-bcc, bcc-fcc and fluid-fcc phase boundaries, locating it at a reduced pressure $P^*_t \equiv \beta P_t \sigma^3/p = (7.0 \pm 0.2) \times 10^{-3}$ and a p-value $p_t = 320 \pm 2$. The error bars depict typical variations in the results when using different sets of data points in the spline fits for the phase boundaries. Fig. 4B displays the phase diagram of the UFM in (x, p) representation, displaying the widths of the coexistence regions.

To assess the influence of the finite size of the simulation cells we have repeated the calculations for three cell sizes containing particle numbers $N = 1024$, 5488 and 11664, respectively, for all three phases. The corresponding results for the triple point are shown in Fig. 5, plotting the triple-point coordinates as a function of $1/N$. The scaling is approximately linear for both P^*_t and p_t, allowing an extrapolation of the triple point position to the $N \to \infty$, giving $P^*_t = (7.2 \pm 0.1) \times 10^{-3}$ and $p_t = 316 \pm 2$. This extrapolated result is shown as the light-blue hexagon in Fig. 4A.

The choice of a finite interaction cut-off at $r_c = 4.0\,\sigma$ is also found to have a

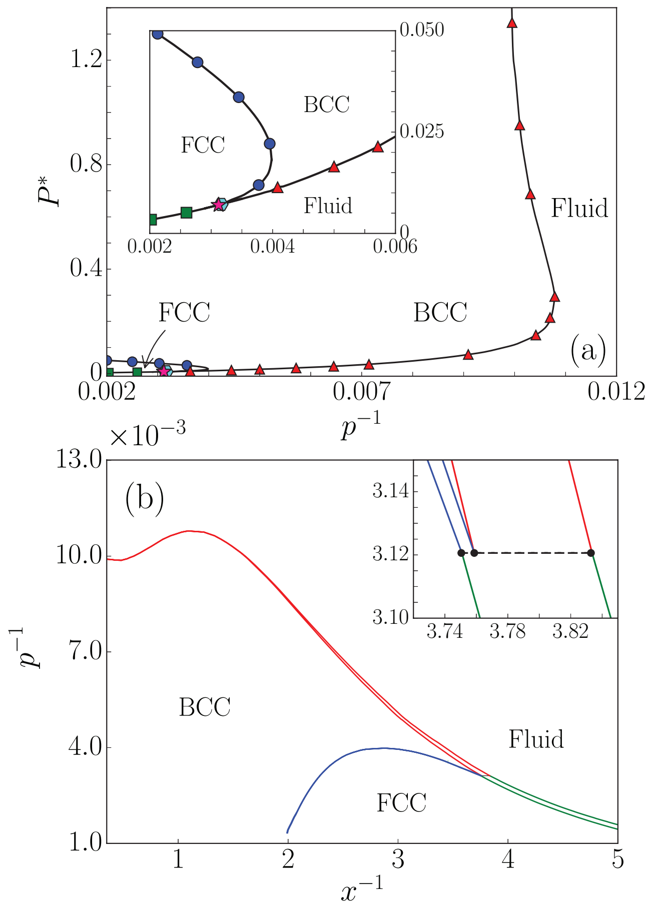

Figure 4 – (Color) A) Phase diagram of the UFM in the (P^*, p) plane, with P^* the reduced pressure defined as $P^* \equiv \beta P \sigma^3 / p$. Symbols represent points of equal chemical potential for specific pairs of phases: bcc-fluid, bcc-fcc and fcc-fluid pairs are depicted by red triangles, blue circles and green squares, respectively. Error bars are smaller than symbol size. Black lines are cubic-spline curves that serve as guides to the eye. Pink star represents the location of the triple point. Light blue hexagon depicts triple-point position obtained after finite-size scaling analysis (see text). Inset provides a zoom of the triple-point region. B) Phase diagram of the UFM in the (x, p) plane. Green, blue and red lines delimit fcc-fluid, fcc-bcc and bcc-fluid coexistence regions, respectively. Inset provides a zoom of the triple-point region, where the dashed line represents the width of the triple point.

negligible influence on the results. Using the same cells as those used to construct the phase diagram we recomputed the Helmholtz free energies for fcc, bcc and fluid phases in the vicinity of the triple point for an increased cut-off radius of $5.0\,\sigma$. The relative differences between the free-energies values for both cut-offs are found to be no larger than 2.0×10^{-4}.

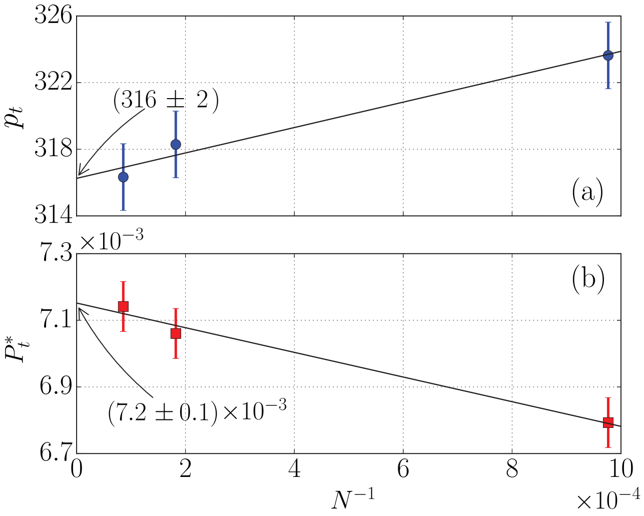

Figure 5 – (Color) Finite-size scaling of the fcc-bcc-fluid Helmholtz free-energy triple point. (A) Linear scaling of p_t as a function of $1/N$. Blue circles represent data for p_t obtained from simulations and line is the best linear fit. (B) Linear scaling of P_t^* as a function of $1/N$. Red squares represent data obtained from simulations and line is the best linear least-squares fit.

2.3.2 Discussion

The phase diagrams of the UFM in Fig. 4 are very similar to those seen in other systems characterized by purely-repulsive interactions, such as the GCM, Yukawa and IPL models. All of these feature phase diagrams contain a single fluid phase and the fcc and bcc crystalline forms. The resemblance with the GCM is particularly striking [1], displaying the same two reentrant transition sequences as can be seen in Fig. 6. Specifically, there is a fluid-bcc-fluid reentrant melting transition as the density is increased for values of p close to 100, i.e., $T^* \equiv 1/p$ close to 0.01, and a bcc-fcc-bcc-fluid sequence in the vicinity of the triple-point region. On the other hand, in contrast to the case of the GCM, for increasing densities the bcc structure remains the stable thermodynamic phase and does not remelt into the fluid.

Prestipino and co-workers [1] discussed the similarities between the phase behaviors of various model systems characterized by soft repulsive interactions. Specifically, they established a criterion to relate the phase behaviors of the IPL and Yukawa models to that of the GCM by requiring that the logarithmic derivatives of the corresponding potentials be equal to that of the GCM for interparticle separations close to the mean distance $\Delta = \rho^{-1/3}$, providing mapping values of potential parameters that play the role that of an effective temperature.

More recently, Khrapak and co-workers [2, 4] proposed a corresponding-states principle by which the melting curves and fcc-bcc-fluid triple points of various different pairwise interaction potentials are shown to display universal behavior when using ap-

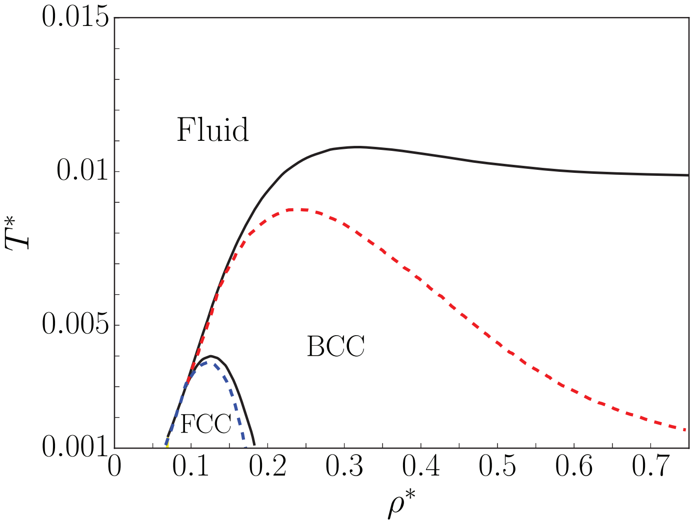

Figure 6 – (Color) Comparison between the phase diagrams of the UFM and GCM in the (ρ^*, T^*) representation. Black lines are cubic-spline curves representing the coexistence conditions for the UFM. On the displayed scale the thickness of the black lines is larger than widths of the coexistence regions. Red and blue dashed lines represent the coexistence conditions for the GCM as obtained from Ref. [1].

propriately selected scaled variables. For a pair potential $U(r)$ these variables are the generalized *softness* parameter

$$s = [-1 - U''(\Delta)\Delta/U'(\Delta)]^{-1} \tag{2.12}$$

and the generalized *interaction* (or reduced force) parameter

$$\mathcal{F} = -\beta\, U'(\Delta)\Delta, \tag{2.13}$$

where $\Delta \equiv \rho^{-1/3}$ is the mean interparticle distance. Using the functional form of Eq. (2.1), the corresponding expressions for the UFM are

$$s = -\frac{1}{2}\left[1 - \left(\frac{\Delta}{\sigma}\right)^2 \frac{e^{(\Delta/\sigma)^2}}{(e^{(\Delta/\sigma)^2} - 1)}\right]^{-1}, \tag{2.14}$$

and

$$\mathcal{F} = \left(\frac{2p}{e^{(\Delta/\sigma)^2} - 1}\right)\left(\frac{\Delta}{\sigma}\right)^2. \tag{2.15}$$

Using these definitions, one can determine the values of the variables s_t and $\mathcal{F}_t$ that correspond to the triple point. Table 1 presents these values for the GCM, IPL, Yukawa and UFM systems, where the data for the former three have been reproduced from Table 1 in Khrapak and Morfill [4]. To compute these parameters for the UFM we have used the triple-point result from Fig. 4, using $p = 320$ and a mean inter-particle distances Δ that corresponds to the average density across the coexistence interval, i.e. $x = 0.26377$.

Table 1 – Scaled variables s and $\mathcal{F}$ at the triple points for the GCM, IPL, Yukawa and UFM models. The numerical values for the first three have been reproduced from Table 1 in Ref. [4]. The UFM results were obtained using the triple-point location present in Fig. 4.

Model	s_t	$\mathcal{F}_t$
GCM	0.129	24.08
IPL	0.14	28.46
IPL	0.16	31.57
Yukawa	0.145	31.26
Yukawa	0.142	27.65
Yukawa	0.128	27.09
UFM	0.13	25.29

They observed that the scaled variables s_t and $\mathcal{F}_t$ display only a relatively narrow range of values between these models, with $s_t \simeq 0.14 \pm 0.02$ and $\mathcal{F}_t \simeq 28 \pm 4$. We find that the corresponding values for the UFM also fall within these intervals. Specifically, when using the above mentioned values of x and p for the UFM, one finds $s_t = 0.13$ and $\mathcal{F}_t = 25.29$. As argued in Ref. [4] this proximity of values for different potentials suggests the existence of a corresponding-states principle that allows the triple-point regions of different model systems to be *approximately* located by finding the thermodynamic conditions that satisfy $s_t^c \simeq 0.14$ and $\mathcal{F}_t^c \simeq 28$. Doing so for the UFM one finds the corresponding-states triple point to be at $p_t^c = 281.7$ and $x_t^c = 0.289574$, which is fairly close to the triple point obtained in the simulations.

In addition to the position of the fcc-bcc-fluid triple point itself, the universal-like behavior also appears for the phase boundary lines near the triple point. Again following Khrapak *et al.* [2, 4], this can be seen if one plots the fluid-solid and bcc-fcc boundaries in terms of the scaled variables s/s_t and $\mathcal{F}/\mathcal{F}_t$, as has been done in Fig. 7 for the GCM, IPL, Yukawa and UFM systems, respectively. The numerical results for the GCM have been taken from Refs. [28], those for the IPL data from Refs. [28, 31], and the Yukawa data from Refs. [40, 41]. Representing the phase diagrams in this manner, it becomes clear that the GCM, Yukawa and IPL have topologically equivalent phase diagrams near the triple point, where it separates a region in which the fluid freezes directly into the fcc phase to another in which the bcc structure is intermediate. Moreover, the rescaled melting lines of all three previously considered models collapse essentially on a single master melting curve, described by $\mathcal{F}(s) \simeq 106\, s^{2/3}$ [2, 4]. Our results here show that the phase behavior of the UFM is equivalent to the other three and is consistent with the corresponding-states principle.

In particular, comparing the rescaled phase behavior to that of the other three models it is evident that the UFM is most similar to the GCM. Indeed, both the rescaled

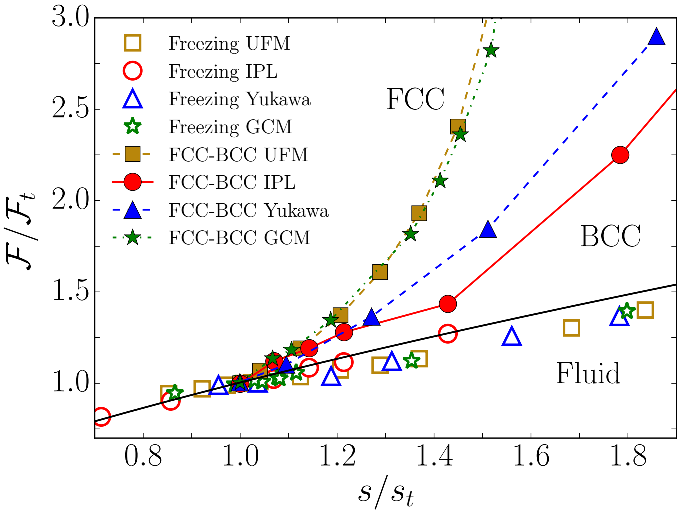

Figure 7 – (Color) Phase boundaries in the vicinity of the fcc-bcc-fluid triple point of the UFM, GCM, IPL and Yukawa systems on the plane of reduced parameters s/s_t and $\mathcal{F}/\mathcal{F}_t$. The black solid curve corresponds to the universal melting curve $\mathcal{F}(s) \simeq 106\, s^{2/3}$ [2]. Open symbols correspond to the fluid-solid phase transitions and the solid symbols represent the fcc-bcc phase transition. (For data references, see the text).

melting curves as well as the bcc-fcc boundaries lines essentially fall on top of each other. Following Ref. [4], this similarity between the UFM and GCM also manifests itself when plotting the potential-energy functions of the models in the vicinity of the triple point in terms of the rescaled form

$$u(r/\Delta) \equiv \beta U(r/\Delta), \tag{2.16}$$

and where the potential parameters have been set such that $s = 0.14$ and $\mathcal{F} = 28$. The results are shown in Fig. 8. While the GCM, IPL and Yukawa potentials touch only at a single point around $x \simeq 0.85$, the GCM and UFM essentially overlap across the entire region $x \gtrsim 0.7$. Only for shorter distances, for which the GCM tends to a constant whereas the UFM diverges, do the potential-energy expressions deviate substantially. This resemblance becomes more explicit when writing the potential-energy function in Eq. (2.1) in terms of its Taylor series representation

$$U_{\mathrm{UF}}(r) = \frac{p}{\beta} \sum_{k=1}^{\infty} \frac{e^{-k(r/\sigma)^2}}{k} = \frac{p}{\beta} e^{-(r/\sigma)^2} \left[1 + \frac{1}{2} e^{-(r/\sigma)^2} + \cdots \right], \tag{2.17}$$

which converges for any $r \geqslant 0$. Except for small values of r/σ, the predominant term in the series corresponds to the GCM. This implies that, as the density of the system is reduced, the behavior of the UFM and the GCM should become progressively more alike. For high densities, on the other hand, it is expected that the two models display different phase behavior. This indeed appears to be the case. While the bcc phase in the GCM melts for any energy scale as long as the system is sufficiently dense, our results indicate that the bcc phase in the UFM is stable for arbitrarily high densities for $p \gtrsim 100$.

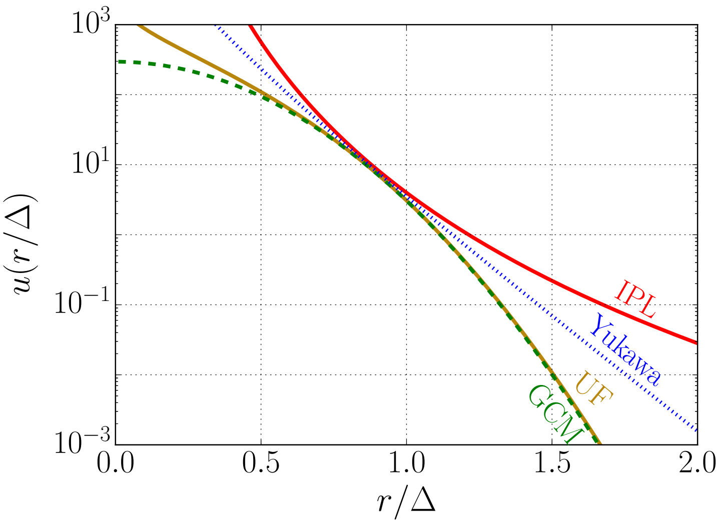

Figure 8 – (Color) Rescaled potential energy functions in the vicinity of the triple point for the IPL, Yukawa, GCM and UFM models, with model parameters corresponding to $s = 0.14$ and $\mathcal{F} = 28$.

2.4 Summary

In summary, we have determined the phase diagram of the UFM using state-of-the-art nonequilibrium free-energy calculation techniques. Similar to other soft purely repulsive systems, we find the phase diagram to contain a single fluid phase and the crystalline bcc and fcc structures, the three of which can coexist at a triple point. The hcp phase is only metastable with respect to the fcc phase and other crystalline phases are found to not even be mechanically stable.

Using a finite-size scaling procedure we determine the location of the fcc-bcc-fluid triple point and analyze its position as well as the phase boundaries in its vicinity in terms of the corresponding-states principle proposed by Khrapak *et al.* [2, 4]. Applying its scaling approach, it is found that the UFM phase behavior is very similar to that of the GCM, IPL and Yukawa systems. The UFM is particularly resemblant of the GCM, with their melting curves and bcc-fcc phase boundaries effectively overlapping. This similarity can be traced back to the particular functional form of the UFM, which can be written in terms of an infinite series of Gaussians. Its first term corresponds to the GCM and, except for large densities, dominates the value of the UFM potential-energy function. For large densities, on the other hand, all terms contribute and give rise to the logarithmic divergence of the UFM. It is suggested that this is related to the fact that, while the GCM is expected to melt as the density is increased, the UFM displays a stable bcc phase for arbitrarily high densities.

3 Nonequilibrium free-energy calculations of fluids using LAMMPS

This chapter was reproduced from the published paper with permission from R. P. Leite, and M. de Koning, Comp. Mat. Sci. 159, 316 (2019) [56]. © 2018 Elsevier B.V. All rights reserved.

3.1 Introduction

Free energies are among the most important quantities in the description of condensed-phase systems and atomistic simulation techniques are very frequently used for their calculation. Nonetheless, neither free energies nor entropies can be expressed in terms of ensemble averages and therefore cannot be determined directly from simulations. By virtue of this difficulty, countless indirect methods to compute free energies have been developed [57]. One of the most applied approaches is the Hamiltonian interpolation (HI) method (also referred to as thermodynamic integration (TI) and λ-integration), [58, 59, 5, 60] which computes the free-energy difference between the system of interest and a reference, for which the free energy is known, by constructing a sequence of equilibrium states along a thermodynamic path between them. Specifically, the free-energy difference is determined by computing ensemble averages of the thermodynamic driving force for a number of states on this path through a set of independent equilibrium simulations, followed by numerical integration. In recent years, nonequilibrium (NE) versions of the HI approach have become popular [50, 53, 49], in particular due to the rigorous connection between NE processes and equilibrium free-energy differences as encoded in Jarzynski's equality [61, 62, 63]. Contrary to equilibrium HI methods, NE approaches estimate the desired free-energy difference by traversing the thermodynamic path between the system of interest and the reference in an explicitly time-dependent process and have shown to give accurate results using only a few relatively short non-equilibrium simulations

For the particular case of solids, either crystalline or amorphous, the choice of reference system is straightforward, with the Einstein crystal [54, 51, 64] providing a system with analytically known free energy. For fluid-phase systems, on the other hand, the choice is less obvious. At first sight, the ideal gas seems a natural pick as a reference since its free energy is known analytically. However, a direct switching path connecting an interacting fluid of interest to the ideal gas may cross a liquid-vapor coexistence line, hampering reversibility due to the presence of appreciable hysteresis [65]. Several strategies have been devised to avoid these difficulties. One of these is to divide the switching process into two stages [66, 67], first introducing an intermediate, purely repulsive reference system, followed by a second stage transforming the repulsive model into the ideal gas. Another frequently adopted approach is to use an interacting fluid with a known free energy as a reference. One example is the Lennard-Jones (LJ) fluid [68, 3], for which extensive numerical data is available [69].

Recently, we discussed the Uhlenbeck-Ford (UF) model and its generalizations as an alternative interacting reference system for fluid-phase free-energy calculations [22, 25, 21]. They are characterized by ultrasoft, purely repulsive pairwise interactions for which the Helmholtz free energies are fully characterized by functions of a single adimensional density variable and temperature appears only as a scaling factor. Moreover, the interactions decay as quickly and smoothly as a Gaussian, dispensing the need for choosing a particular truncation and/or shifting scheme with accompanying long-range corrections, as usually required for the LJ system [6].

Here we discuss the application of the UF-based reference system in the calculation of fluid-phase free energies, focusing on the implementation in the widely used `LAMMPS` molecular dynamics (MD) package [55]. We show how, together with NE free-energy techniques, the UF models can be used to accurately and efficiently compute free energies of both atomic as well as molecular liquids. We provide excerpts from the used `LAMMPS` scripts to exemplify the practical details of the calculations. Complete `LAMMPS` scripts, source codes and postprocessing tools are also made available. In this sense, the present paper is similar to Ref. [49] in which the calculation of solid-phase free energies using LAMMPS were discussed in detail. As illustrations we perform free-energy computations for two atomic liquids and three molecular models for water. Specifically, for the former we consider the coarse-grained atomic water (mW) model [70] and a Cu-Zr liquid alloy described in terms of an MEAM-2NN potential [71, 72, 73, 74, 75]. The molecular water systems concern the rigid TIP4P [76] and SPC/E [77] potentials as well as the flexible q-SPC/Fw model [78]. For these water models we construct rigid and flexible interacting molecular reference systems based on the UF models such that their free energies are given by a sum of two contributions: an intermolecular part described by the known UF free energy and an intramolecular contribution that can be determined analytically.

The remainder of the paper is organized as follows. In Section 3.2 we provide a brief review of the NE methods used to compute free-energy differences between two equilibrium states. In Section 3.3, we describe in detail the UF-based reference systems for atomic and water-like molecular fluids. Next, in Section 3.4, we illustrate the implementation and application of these methods within the `LAMMPS` code, reporting the results for the above-mentioned systems and providing script excerpts. Finally, we end with a summary in Section 3.5.

3.2 Nonequilibrium free-energy methods

One of the widely-used methods to estimate free energy differences is the equilibrium TI technique introduced by Kirkwood in 1935 [58]. It consists of the construction of a path connecting two generic well-defined thermodynamic states by defining a parametrized Hamiltonian $H(\lambda)$, where λ is a coupling parameter describing the interpolation between the two ends of the thermodynamic path.

According to the second law of thermodynamics, the reversible work W_{rev} done along such a Hamiltonian-interpolation (HI) path is equal to the free-energy difference between the two equilibrium systems corresponding to $H(\lambda_i)$ and $H(\lambda_f)$, i.e.

$$\Delta F \equiv F(\lambda_f) - F(\lambda_i) = W_{\text{rev}} = \int_{\lambda_i}^{\lambda_f} \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_{\lambda} d\lambda, \tag{3.1}$$

where $\langle \cdots \rangle_\lambda$ is the canonical ensemble average at a particular value of λ parameter and $\partial H/\partial\lambda$ is the so-called driving-force. In this approach, an equilibrium simulation is performed for a set of λ-values in the interval between λ_i and λ_f, followed by numerical integration of the corresponding driving-force values.

As an alternative, intrinsically NE processes can be used to estimate the free-energy difference ΔF. In such a NE process, the coupling parameter $\lambda = \lambda(t)$ changes continuously throughout the simulation, and instead of an integration over equilibrium ensemble averages, Eq.(3.1), is replaced by an integration over instantaneous values of the driving-force,

$$W_{\text{dyn}} = \int_0^{t_s} \frac{d\lambda}{dt} \left. \frac{\partial H(\lambda)}{\partial \lambda} \right|_{\lambda(t)} dt, \tag{3.2}$$

where t_s is the switching time and W_{dyn} is the dynamical work done. Due to the irreversible nature of the process, dissipative entropy is produced, causing W_{dyn} to be a stochastic variable whose mean value, by the second law of thermodynamics, differs from W_{rev} by the relation

$$\Delta F = W_{\text{rev}} = \overline{W_{\text{dyn}}} - \overline{Q_{\text{diss}}}, \tag{3.3}$$

where the overbar means an average over an ensemble of realizations of the NE process and $\overline{Q_{\text{diss}}} \geqslant 0$ is the average dissipated heat. The latter is zero only in the quasistatic limit ($t_s \to \infty$). However, it can be shown [53] that the systematic error can be eliminated by combining the results of the processes realized in both directions (forward and backward) as long as the process is executed sufficiently slowly for linear-response theory to be valid. Accordingly, one has

$$\begin{aligned}
\Delta F &\equiv \frac{1}{2}\left[W_{\text{rev}}^{i\to f} - W_{\text{rev}}^{f\to i}\right] \\
&= \frac{1}{2}\left\{\left[\overline{W_{\text{dyn}}^{i\to f}} - \overline{Q_{\text{diss}}^{i\to f}}\right] - \left[\overline{W_{\text{dyn}}^{f\to i}} - \overline{Q_{\text{diss}}^{f\to i}}\right]\right\} \\
&= \frac{1}{2}\left[\overline{W_{\text{dyn}}^{i\to f}} - \overline{W_{\text{dyn}}^{f\to i}}\right],
\end{aligned} \tag{3.4}$$

where we have used the fact that $\overline{Q_{\text{diss}}^{i\to f}} = \overline{Q_{\text{diss}}^{f\to i}}$ under these conditions. Similarly, the systematic error can be estimated by

$$\overline{Q_{\text{diss}}^{i\to f}} = \overline{Q_{\text{diss}}^{f\to i}} = \frac{1}{2}\left[\overline{W_{\text{dyn}}^{i\to f}} + \overline{W_{\text{dyn}}^{f\to i}}\right]. \tag{3.5}$$

Eqs. (3.4) and (3.5) allow a systematic monitoring of the convergence of the results with the process rate, running a number of forward and backward realizations for a set of t_s values and plotting ΔF and the dissipation as a function of t_s.

While the comparison between the computational efficiencies of the equilibrium and NE methods is an active topic of investigation [79, 80, 81, 82], it has been shown that the latter allows one to obtain accurate estimates of ΔF using only a few relatively short non-equilibrium simulations [53, 49]. Indeed, due to the availability of straightforward error analysis and convergence protocols, the NE techniques serve as an attractive alternative to the standard equilibrium methodology.

In the following we present two specific thermodynamic paths $H(\lambda)$ that, respectively, are used to compute (i), the free-energy difference between Hamiltonians describing two different systems and, (ii) the temperature-dependence of the free energy for a given system Hamiltonian.

3.2.1 Free-energy difference between two systems: Hamiltonian interpolation method

Suppose we wish to compute the absolute free-energy of some system of interest described by a Hamiltonian H_{int} of the form

$$H_{\text{int}} = K + U_{\text{int}}(\{\mathbf{r}_i\}), \tag{3.6}$$

where K is the kinetic energy and U_{int} is the system of interest's interaction potential, which is a function of the set of particle coordinates and/or molecular degrees of freedom $\{\mathbf{r}_i\}$. Suppose further that there is a second system, in the same thermodynamic phase as that of the previous one, having a known free energy (analytical or numerical) and described by the Hamiltonian H_{ref}, given by

$$H_{\text{ref}} = K + U_{\text{ref}}(\{\mathbf{r}_i\}), \tag{3.7}$$

with $U_{\text{ref}}(\{\mathbf{r}_i\})$ its interaction potential.

The HI method consists of defining a parameterized Hamiltonian $H(\lambda)$ as a linear interpolation between the system of interest and reference Hamiltonians:

$$H(\lambda) = \lambda H_{\text{int}} + (1-\lambda)H_{\text{ref}}. \tag{3.8}$$

We define the forward process according to $\lambda_i = 1$ and $\lambda_f = 0$, transforming H_{int} into H_{ref}.

The dynamical work for a given realization of the HI process is given by

$$W_{\mathrm{dyn}}^{i\to f} = \int_0^{t_s} \frac{d\lambda}{dt} \left(U_{\mathrm{int}} - U_{\mathrm{ref}}\right) dt, \tag{3.9}$$

which, in practice, is estimated numerically.

Finally, combining the average results obtained from a number of independent realizations in the forward and backward switching processes, the desired free-energy is estimated as

$$F_{\mathrm{int}} = F(\lambda_i) = F_{\mathrm{ref}} + \frac{1}{2}\left[\overline{W_{\mathrm{dyn}}^{i\to f}} - \overline{W_{\mathrm{dyn}}^{f\to i}}\right]. \tag{3.10}$$

3.2.2 Free energy as a function of the temperature: Reversible Scaling method

Suppose the Helmholtz free energy $F_{\mathrm{int}}(T_0)$ of a system of interest is known at some temperature T_0 (e.g., computed using the method of previous section), and that we now wish to determine $F_{\mathrm{int}}(T)$ for other temperatures T. One way of doing this is repeating the free-energy calculation for each temperature of interest. Alternatively, this can be achieved using a single, constant-temperature MD simulation using the reversible-scaling (RS) technique [48, 83]. The RS method is based on the parametric Hamiltonian

$$H_{\mathrm{RS}}(\lambda) = K + \lambda U_{\mathrm{int}}(\{\mathbf{r}_i\}), \tag{3.11}$$

where the potential energy is scaled by the coupling parameter λ. The configurational part of the classical partition function $Z_{\mathrm{RS}}(\lambda)$ of the RS Hamiltonian at temperature T_0 is given by

$$Z_{\mathrm{RS}}(\lambda) = \int d^{3N}\mathbf{r}\,\exp\left[-\lambda U(\{\mathbf{r}\})/k_B T_0\right] = Z_{\mathrm{int}}(T_0/\lambda), \tag{3.12}$$

which is equal to that of the partition function of the system of interest at temperature $T = T_0/\lambda$. In view of this relationship it can be shown [48] that the free energies of the scaled and physical systems are related according to

$$F_{\mathrm{int}}(T) = \frac{F_{\mathrm{RS}}(T_0;\lambda)}{\lambda} + \frac{f}{2} N k_B T_0 \frac{\ln\lambda}{\lambda}, \tag{3.13}$$

where N is the number of entities (i.e, atoms or molecules) in the system and f is the number degrees of freedom for each of them, e.g., $f = 3$ for atomic systems and $f = 6$ for a fluid composed of rigid molecules. Equation (3.13), implies that each value of λ in the scaled Hamiltonian H_{RS} at a *fixed* temperature T_0 corresponds to the system of interest described by H_{int} at a temperature $T = T_0/\lambda$. In other words, the free energy of physical system as a function of temperature T can be obtained from $H(\lambda)$ by varying the scaling parameter λ at fixed temperature T_0. To this end the AS procedure can be applied, with $\lambda(t)$ varying from $\lambda(0) = 1$ to $\lambda(t_s) = \lambda_f$ to estimate the forward dynamical work along the isothermal scaling process,

$$W_{\mathrm{dyn}}^{1\to f} = \int_0^{t_s} \frac{d\lambda}{dt} U_{\mathrm{int}}\, dt. \tag{3.14}$$

Also carrying out the scaling process in the opposite direction and using Eq.(3.10), $F_{\text{int}}(T)$ on the temperature interval between T_0 and T_0/λ_f is given by

$$\begin{aligned} F_{\text{int}}(T) &= \frac{F_{\text{int}}(T_0)}{\lambda} + \frac{f}{2} N k_B T_0 \frac{\ln \lambda}{\lambda} \\ &+ \frac{1}{2\lambda}\left[\overline{W_{\text{dyn}}^{1\to\lambda}} - \overline{W_{\text{dyn}}^{\lambda\to 1}}\right], \end{aligned} \tag{3.15}$$

where λ varies between 1 and λ_f

It is important to note that the application of this approach requires knowledge of the absolute free energy of the system of interest at a temperature T_0, which can be computed using the HI method detailed in Sec. 3.2.1.

Finally, the RS method summarized above permits one to compute the *Helmholtz* free energy as a function of temperature for fixed volume. However, RS can also been generalized to compute the *Gibbs* free energy as a function of temperature for any desired pressure P [83]. In particular, for zero pressure, the scaling simulation detailed above, when applied under conditions of constant zero pressure $P = 0$ and temperature T_0, transfers directly to the temperature-dependence of the Gibbs free energy, replacing the Helmholtz free energy F_{int} by the Gibbs free energy G_{int} in Eq. (3.15), i.e.,

$$\begin{aligned} G_{\text{int}}(P=0,T) &= \frac{G_{\text{int}}(P=0,T_0)}{\lambda} + \frac{f}{2} N k_B T_0 \frac{\ln \lambda}{\lambda} \\ &+ \frac{1}{2\lambda}\left[\overline{W_{\text{dyn}}^{1\to\lambda}} - \overline{W_{\text{dyn}}^{\lambda\to 1}}\right]. \end{aligned} \tag{3.16}$$

3.3 Uhlenbeck-Ford reference systems

In this section we describe the UF-based fluid-phase reference systems that we use to compute the absolute free energies of atomic and molecular fluids. In view of the particular applications that will be presented, we describe references for atomic fluids, as well as for rigid and flexible models for liquid water. All the reference systems are interacting fluids composed of atoms or molecules that repel each other according to the UF model, defined as [22, 25, 21]

$$U_{\text{UF}}(\mathbf{r}) = -\frac{p}{\beta} \ln\left(1 - e^{-(r/\sigma)^2}\right), \tag{3.17}$$

where $\beta \equiv (k_B T)^{-1}$, σ is a length-scale parameter, r is an inter-particle distance and $p > 0$ is a scaling factor that controls the softness of the interactions. The UF model is a purely repulsive, smooth soft-sphere pairwise interaction that decays rapidly for increasing distances, diverges logarithmically at the origin, and is characterized by an energy scale controlled by the absolute temperature T.

The UF models feature a number of properties that render it a suitable choice to serve as a reference system for fluid-phase free-energy computations:

i) The excess free energy of the UF fluid can be represented as a function of a single adimensional density parameter for any desired temperature,

$$F_{\mathrm{UF}}^{(\mathrm{exc})}(x,T) = k_B T \sum_{n=1}^{\infty} \frac{\tilde{B}_{n+1}(p)}{n} x^n, \tag{3.18}$$

where

$$x \equiv b\rho, \tag{3.19}$$

with ρ the system's number density, $b \equiv (\pi\sigma^2)^{3/2}/2$ and the $\tilde{B}_{n+1}(p)$ are reduced virial coefficients that depend only on the scaling factor p and, in principle, can be computed exactly [25]. However, the effort associated with their calculation increases extremely rapidly with the order n such that only a limited number of them can actually be evaluated [25]. Nevertheless, a set of very accurate numerical representations of the free-energy functions in Eq. (3.18) is available for $p = 1, 25, 50, 75, 100$, and can be found in the Supplementary Material of Ref. [25].

ii) There is only a single fluid phase, i.e., there is no liquid-gas transition, and this fluid phase is the only thermodynamically stable phase for $p \lesssim 100$ [21].

In the following we provide details concerning the UF-based reference systems and discuss how to apply them using the NE Hamiltonian-interpolation (NEHI) free-energy technique discussed in the previous Section.

3.3.1 Atomic models

As described previously [25], the application as a reference system for the calculation of free energies of atomic fluids is straightforward. In this case, the thermodynamic path is defined as

$$H(\lambda) = \lambda U_{\mathrm{int}} + (1-\lambda) U_{\mathrm{UF}}, \tag{3.20}$$

transforming the interactions in the physical system into those of the UF model as λ varies from 1 to 0. Even when dealing with atomic fluids containing different species, such as in the case of liquid alloys, all particles interact according to the same UF model.

Determining the free-energy difference based on the forward and backward dynamical work estimators according to Eq. (3.4), the absolute Helmholtz free energy of the fluid of interest is given by

$$F_{\mathrm{int}} = F_{\mathrm{ig}} + F_{\mathrm{UF}}^{(\mathrm{exc})} + \frac{1}{2}\left[\overline{W_{\mathrm{dyn}}^{1\to 0}} - \overline{W_{\mathrm{dyn}}^{0\to 1}}\right], \tag{3.21}$$

where $F_{\mathrm{UF}}^{(\mathrm{exc})}$ is the known UF excess Helmholtz free energy and F_{ig} is the kinetic ideal gas contribution. In the case of a monoatomic system of N atoms with mass m, it is given by

$$F_{\mathrm{ig}} = \frac{N}{\beta}\left[3\ln(\Lambda) + \ln(\rho) - 1 + \frac{1}{2N}\ln(2\pi N)\right], \tag{3.22}$$

where Λ is the de Broglie's thermal wavelength

$$\Lambda = \sqrt{\frac{\beta h^2}{2\pi m}}, \tag{3.23}$$

and last term on the right side is a logarithmic correction to Stirling's approximation.

For a binary mixture containing a total of N particles with concentrations X_A and X_B and masses m_A and m_B for species A and B, respectively, the ideal-gas part is given by

$$\begin{aligned} F_{\text{ig}} &= \frac{N}{\beta}\Big\{ X_A \left[3\ln\left(\Lambda_A\right) + \ln\left(\rho\right) - 1 + \ln\left(X_A\right)\right] \\ &+ X_B \left[3\ln\left(\Lambda_B\right) + \ln\left(\rho\right) - 1 + \ln\left(X_B\right)\right] \\ &+ \frac{1}{2N}\ln\left(2\pi N\right)\Big\}, \end{aligned} \tag{3.24}$$

with Λ_A and Λ_B the de Broglie wavelengths associated with the masses m_A and m_B.

3.3.2 Rigid-body water models: UF/Rw

In the last 50 years, a plethora of empirical water models have been developed. A substantial fraction of these describe the water molecule as a rigid body in which the intramolecular degrees of freedom, such as OH bond length r_{eq} and HOH bond angle θ_{eq}, are fixed. Often, the interactions between the water molecules involve a LJ-type interaction between the oxygen atoms, supplied with Coulomb interactions between point charges located on a variety of positions, giving a potential of the form

$$U_{\text{int}} = \sum_{i,j>i} \left\{ 4\varepsilon \left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6} \right] + \sum_{m,n} \frac{q_{im}q_{jn}}{r_{im,jn}} \right\} \tag{3.25}$$

where the indices i and j label molecules, m and n label the point charges within the molecules, r_{ij} is the distance between the oxygens of molecules i and j and the $r_{im,jn}$ are the distances between their respective charge centers m and n.

Because of the rigidity of the water molecules in these models, a natural choice for a reference system is one consisting of molecules with the same intramolecular structure. In addition, to decouple translational degrees of freedom from the rotational ones, it is convenient to design the interaction between molecules of the reference system to act on their respective centers of masses [68]. In this manner, the reference system we employ for rigid water models is characterized by the following elements:

- The reference system consists of rigid molecules that have the same geometry as that of the water model under consideration.
- The interaction between two molecules is given by the UF model, with the interparticle distance between them defined as the distance between their centers of masses.

- The reference molecules do not possess any charges.

The Hamiltonian of this reference system, which we will refer to as UF/Rw, is given by

$$H_{\mathrm{UF/Rw}} = K_{\mathrm{rot}}(\boldsymbol{\varphi}_i, \mathbf{p}_{\varphi_i}) + K_{\mathrm{trans}}(\{\mathbf{P}_i\}) + U_{UF}(\{\mathbf{R}_i\}), \tag{3.26}$$

where K_{rot} describes the rotational kinetic energy in terms of the molecules' Euler angles and associated canonical momenta, K_{trans} is the translational kinetic energy, which is a function of the total linear momenta $\{\mathbf{P}_i\}$ of the molecules of total mass M, and $U_{UF}(\{\mathbf{R}_i\})$ is the UF interaction between the molecular centers of mass $\{\mathbf{R}_i\}$.

The classical canonical partition function for a system of N such molecules can be written as

$$Z = \frac{1}{N!} Z_{\mathrm{rot}}^N Z_{\mathrm{trans}}^N, \tag{3.27}$$

where the rotational part is given by (see Appendix A.1)

$$Z_{\mathrm{rot}} = \sqrt{\frac{2\pi I_1\, I_2\, I_3}{\beta^3 \hbar^6}}, \tag{3.28}$$

where I_1, I_2 and I_3 are the principal moments of inertia of the molecule, and the translational part is identical to that of an atomic system of N UF particles with mass M [25]. The corresponding absolute Helmholtz free energy of the UF/Rw system is then given by

$$\begin{aligned} F_{\mathrm{UF/Rw}}(T) &= -\frac{N}{\beta}\left[\frac{3}{2}\ln\left(\frac{2\pi M}{\beta h^2}\right) - \ln(\rho) + 1\right. \\ &- \left.\frac{1}{2N}\ln(2\pi N) + \frac{1}{2}\ln\left(\frac{2\pi I_1 I_2 I_3}{\beta^3\hbar^6}\right)\right] \\ &+ F_{\mathrm{UF}}^{(\mathrm{exc})}(x,T), \end{aligned} \tag{3.29}$$

where x is the adimensional scaled density variable defined in Eq. (3.19). In this expression, the intermolecular contribution to the free energy is described entirely by $F_{\mathrm{UF}}^{(\mathrm{exc})}(x,T)$, with the other terms representing the translational and rotational ideal-gas contributions.

With the available numerical expressions for $F_{\mathrm{UF}}^{(\mathrm{exc})}(x,T)$, we use the thermodynamic path

$$H(\lambda) = \lambda H_{\mathrm{int,rigid}} + (1-\lambda)H_{\mathrm{UF/Rw}}, \tag{3.30}$$

connecting the fully-interacting rigid water model to the rigid UF/Rw reference. Then, performing forward and backward processes to estimate the work done along this thermodynamic path, the absolute Helmholtz free-energy of the fully-interacting water system is computed as

$$F_{\mathrm{int,rigid}} = F_{\mathrm{UF/Rw}} + \frac{1}{2}\left[\overline{W_{\mathrm{dyn}}^{1\to 0}} - \overline{W_{\mathrm{dyn}}^{0\to 1}}\right] \tag{3.31}$$

3.3.3 Flexible water models: UF/Fw

Another class of water models incorporates explicit flexibility into the molecule structure, allowing both stretching of the OH bonds as well as variation of the ∠HOH angle. To compute the free energy of fluids composed of such molecules we employ a flexible reference system in which the intermolecular interactions acting between the oxygen atoms are described by the UF model, whereas the intramolecular bond and angular dynamics are characterized by purely harmonic potentials. Specifically, the Hamiltonian of the flexible UF water (UF/Fw) model is given by

$$H_{\text{UF/Fw}} = K + U_{\text{intra}} + U_{\text{inter}} \tag{3.32}$$

where K is the sum of the kinetic energies of all atoms, U_{intra} is the sum of the intramolecular bond energies,

$$U_{\text{intra}} = \sum_{i=1}^{N} \left[U_{\text{OH}}(r_{i,1}) + U_{\text{OH}}(r_{i,2}) + U_{\text{HOH}}(\theta_i) \right], \tag{3.33}$$

in which $r_{i,1}$ and $r_{i,2}$ are the two OH bond lengths of molecule i and θ_i is its ∠HOH bond angle. The bond and angle potential energies are given by

$$U_{\text{OH}}(r) = \frac{1}{2} k_r (r - r_{\text{eq}})^2, \tag{3.34}$$

and

$$U_{\text{HOH}}(\theta) = \frac{1}{2} k_\theta (\theta - \theta_{\text{eq}})^2, \tag{3.35}$$

in which k_r and k_θ are the bond and angle spring constants, respectively, and r_{eq} and θ_{eq} are the corresponding equilibrium bond length and angle.

The intermolecular potential energy is given by

$$U_{\text{inter}} = \sum_{i>j} U_{\text{UF}}(r_{ij}), \tag{3.36}$$

with r_{ij} the distance between the oxygen atoms of molecules i and j.

Computing the canonical partition function (see Appendix A.2) one can show that the Helmholtz free energy of the flexible reference fluid is given by a sum of the UF free energy and a contribution due to the intramolecular interactions that can be expressed in analytical form,

$$\begin{aligned} F_{\text{UF/Fw}}(T) &= -\frac{N}{\beta} \left[\frac{3}{2} \ln \left(\frac{8\pi^3 m_{\text{O}} m_{\text{H}}^2}{\beta^3 h^6} \right) - \ln \left(\frac{\rho}{8\pi^2} \right) + 1 \right. \\ &- \left. \frac{1}{2N} \ln (2\pi N) + 2 \ln (I_r) + \ln (I_\theta) \right] \\ &+ F_{\text{UF}}^{(\text{exc})}(x, T), \end{aligned} \tag{3.37}$$

in which m_{O} and m_{H} are the masses of the oxygen atom and the proton, respectively, and where

$$\begin{aligned} I_r(k_r,\beta) &= \frac{\sqrt{\frac{\pi}{2}}(\beta r_{\mathrm{eq}}^2 k_r + 1)\left[\mathrm{erf}\left(\mathrm{r}_{\mathrm{eq}}\sqrt{\frac{\beta \mathrm{k}_{\mathrm{r}}}{2}}\right) + 1\right]}{(\beta k_r)^{3/2}} \\ &+ \frac{r_{\mathrm{eq}}\exp\left(-\frac{1}{2}\beta r_{\mathrm{eq}}^2 k_r\right)}{\beta k_r}, \end{aligned} \tag{3.38}$$

and

$$\begin{aligned} I_\theta(k_\theta,\beta) &= \sqrt{\frac{\pi}{2\beta k_\theta}}\exp\left(-\frac{1}{2\beta k_\theta}\right) \\ &\times \ \mathrm{Re}\left[\mathrm{e}^{\mathrm{i}\theta_{\mathrm{eq}}}\mathrm{erfi}\left(\frac{1-\mathrm{i}\theta_{\mathrm{eq}}\beta \mathrm{k}_\theta}{\sqrt{2\beta \mathrm{k}_\theta}}\right)\right. \\ &- \left. e^{i\theta_{\mathrm{eq}}}\mathrm{erfi}\left(\frac{1+\mathrm{i}(\pi-\theta_{\mathrm{eq}})\beta \mathrm{k}_\theta}{\sqrt{2\beta \mathrm{k}_\theta}}\right)\right], \end{aligned} \tag{3.39}$$

involve the imaginary error function $\mathrm{erfi}(z)$ [84]. These results are consistent with those reported recently [85]. As in Eq. (3.29), the intermolecular contribution to the free energy is described entirely by $F_{\mathrm{UF}}^{(\mathrm{exc})}(x,T)$, with the remaining terms representing the intramolecular vibrational and ideal-gas contributions.

With the availability of $F_{\mathrm{UF/Fw}}$ we use the thermodynamic path

$$H(\lambda) = \lambda H_{\mathrm{int,flexible}} + (1-\lambda)H_{\mathrm{UF/Fw}}, \tag{3.40}$$

connecting the flexible water model to the UF/Fw reference. Performing forward and backward processes, the absolute Helmholtz free-energy of the water system is calculated as

$$F_{\mathrm{int,flexible}} = F_{\mathrm{UF/Fw}} + \frac{1}{2}\left[\overline{W_{\mathrm{dyn}}^{1\to 0}} - \overline{W_{\mathrm{dyn}}^{0\to 1}}\right]. \tag{3.41}$$

3.4 Applications: Results and Discussion

In this section we describe how to implement and apply the described reference systems in the `LAMMPS` code to compute fluid free energies. Complete `LAMMPS` scripts, source code and auxiliary files are available [86].

As a first application, we determine the melting temperature of the hexagonal and cubic diamond phases of mW model. Next, we compute the free energies of the liquid Cu-Zr alloy as a function of composition. Finally, we apply the rigid and flexible molecular UF references to compute the Helmholtz free energies of the rigid TIP4P, SPC/E models as well as the flexible q-SPC/Fw potential.

3.4.1 Mono-atomic water (mW) model

To determine the melting temperature of the hexagonal and cubic-diamond structures of the mW model at zero pressure we compute the Gibbs free-energy curves

$G(P,T)$ for the solid and liquid phases and find the temperature T_m at which they cross. To achieve this we follow Ref. [87], carrying out the following sequence of steps:

For the liquid:

i) Perform an NPT MD simulation at zero pressure to determine the equilibrium density ρ_l of the liquid at a temperature $T_2 > T_m$.

ii) Perform NVT MD simulations to compute the Helmholtz free energy F_l of the liquid at a density ρ_l and temperature T_2 using the AS procedure transforming the liquid of interest into the UF reference. For this case, $F_l(V_l,T_2) = G_l(P=0,T_2)$.

iii) Apply the RS method in the zero-pressure NPT ensemble to construct the Gibbs free energy curve $G_l(P=0,T)$ of the liquid between T_1 and T_2.

For the solid:

i) Perform a NPT MD simulation to determine the equilibrium density ρ_s of the crystalline solid at a temperature $T_1 < T_m$.

ii) Perform NVT MD simulations to compute the Helmholtz free energy F_s of the crystalline solid at density ρ_s and temperature T_1. This is done applying the AS procedure connecting the interacting crystalline solid to an Einstein crystal. Since we consider the system under zero pressure, $F_s(\rho_s,T_1) = G_s(P=0,T_1)$.

iii) Apply the RS method under NPT ensemble at zero pressure $P=0$ to construct the Gibbs free energy curve $G_s(P=0,T)$ of the solid between T_1 and T_2.

For all phases we use computational cells containing $8,000$ mW atoms subject to periodic boundary conditions. Pressure and temperature control was obtained using a Parrinello-Rahman-type barostat [88] and a Langevin thermostat [20]. The corresponding equations of motion were integrated using velocity-Verlet algorithm with a time step of $\Delta t = 2$ fs.

For the liquid phase, step i) entails determining the zero-pressure equilibrium volume at $T_2 = 290K$. This is achieved using the following fix commands in the `LAMMPS` script:

```
fix f1 all nph iso 0.0 0.0 3.0
fix f2 all langevin 290 290 0.2 666
```

After an initial equilibration of 0.2 ns, we determine the average value of the volume over a time interval of 3.0 ns, giving an equilibrium number density of $\rho_l = 0.0334045(1)$ Å^{-3}, where the number in parentheses denotes the uncertainty in the final digit.

Next, in step ii), we compute the Helmholtz free energy of liquid at this density, applying the NEHI procedure to compute the Helmholtz free-energy difference between the mW and UF fluids at $T = 290$ K and $\rho = \rho_l$. The switching process is implemented in the `LAMMPS` script by invoking the `pair_style hybrid/overlay` command, followed by the `pair_coeff` instructions for the mW and UF models:

```
pair_style hybrid/overlay sw ufm 10.0
pair_coeff 1 1 sw mW.sw H2O
pair_coeff 1 1 ufm 1.2495 2.0
```

The value 10.0 in the `hybrid/overlay` command corresponds to a cut-off radius $r_c = 10$ Å. When involving the UF model, it should always be chosen such that $r_c = 5\sigma$, with σ the length scale parameter in the UF model. This requirement is due to the fact that the available excess free energies of the UF models [25] are given for this specific cut-off distance.

The parameter values of a reference system in the NEHI approach should be chosen such that the driving force $\partial H/\partial\lambda = U_{\text{int}} - U_{\text{ref}}$ is as smooth as possible to minimize the dissipated heat for a given switching rate. To this end it is useful to carry out a few short NEHI simulations to assess the behavior of the driving force and systematic error for different choices of the reference-system parameters. Fig 9 shows the results of such an analysis obtained from simulations with $t_s = 10$ ps. In Fig. 9a) we consider the value of the scaling parameter p during a single forward process from the mW liquid to the UF reference, monitoring the driving-force behavior for $p = 1$, 10 and 50 at $\sigma = 2.0$ Å. It can be seen that for small values of p the UF reference is too soft, allowing particles to become too close to each other and causing a near divergence [25] in the driving force close to $\lambda_f = 0$. Increasing p to 50 eliminates this problem, giving a slowly varying driving force over the entire λ-interval. Indeed, this choice for p has shown to be adequate for all NEHI calculations using the UF model, providing smooth driving-force curves characterized by small fluctuations. Similarly, if the length scale σ of the particle size is chosen too small, an analogous near-divergence issue occurs, as can be seen in Fig. 9b). This can also be resolved by increasing its value. For this particular case a choice of $\sigma = 2.0$ Å gives a smooth slowly varying driving force. This choice can be further validated by measuring the mean dissipated heat $\overline{Q_{\text{diss}}}$ during the switching process as a function of σ, as displayed in Fig. 9c). While Q_{diss} is large for smaller values, it decreases by an order of magnitude upon

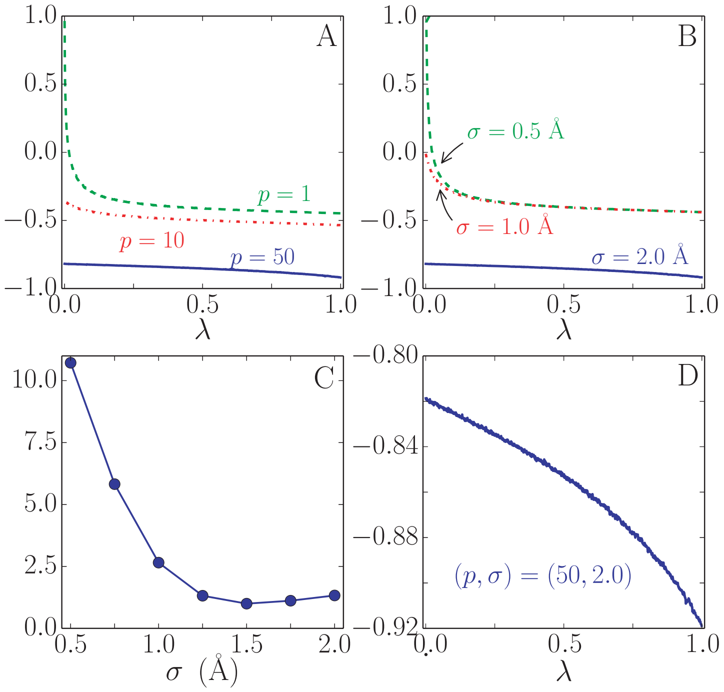

Figure 9 – (Color) Choice of parameters for UF reference to compute the free energy of the mW model. (A) Driving force $\partial H/\partial\lambda$ in units of eV per atom as a function of λ for various values p with $\sigma = 2.0$ Å. (B) Driving force $\partial H/\partial\lambda$ in units of eV per atom as a function of λ for various values of σ with $p = 50$. (C) Average dissipated heat $\overline{Q_{\text{diss}}}$ in units of meV per atom as a function of σ for $p = 50$. (D) Driving force as a function of λ for chosen parameter set $p = 50$ and $\sigma = 2.0$ Å.

increasing σ, reaching a shallow minimum close to $\sigma = 1.5$ Å. The chosen parameter set with $p = 50$ and $\sigma = 2.0$ Å gives a slowly varying driving force with very small fluctuations, as can be seen in Fig. 9d).

In the corresponding `pair_coeff` command these values of p and σ are encoded in the first and second arguments after the `ufm` specification, respectively, with 1.2495 corresponding to the energy-scale $p\,k_BT = 1.2495$ eV, with $T = 290$ K.

In the `hybrid/overlay` pair style, the interactions between atoms are superpositions of two specified models and one can control the magnitude of each interaction type by changing the scale factor that multiplies the forces on the atoms using the fix `adapt` command. For example, to equilibrate a fluid in which the atoms interact only through the mW potential within the defined `hybrid/overlay` pair style given above, one can turn-off the UF interactions by scaling to zero the corresponding force contribution. This is accomplished by the commands

```
variable lambda equal 0
fix f3 all adapt 0 pair ufm fscale 1 1 v_lambda
run 0
```

Here, `fscale` is a force scale factor defined in the employed pair styles. By default it is initialized at a value 1.0, but in the command sequence above it is set equal to the value of the variable `lambda`. Note that only the force contributions are scaled, whereas the potential-energy values are not. This is required since the driving force $\partial H/\partial\lambda$ always involves unscaled values of the potential energies. The purpose of the `run 0` command is to immediately update the forces according to this definition, without performing an MD time step.

The definitions above can be used to carry out the AS procedure linking the mW and UF fluids and compute the associated dynamical work value. In practice, this is accomplished by carrying out the following sequence of steps:

1) First we set $\lambda = 1$ and equilibrate the system $H(\lambda = 1) = H_{\mathrm{mW}}$ during a time interval of t_{eq}.
2) We linearly change the λ parameter from $\lambda = 1$ to $\lambda = 0$ in a switching time t_{s}. This amounts to changing the particle interactions from mW at $t = 0$ to UF at $t = t_{\mathrm{s}}$. We define this procedure to be the forward process.
3) Next we set $\lambda = 0$ and equilibrate the system described by $H(\lambda = 0) = H_{\mathrm{UF}}$ during a time of t_{eq}.
4) We linearly change the λ parameter from $\lambda = 0$ to $\lambda = 1$ in a switching time t_{s}. This amounts to changing the particle interactions from UF at $t = 0$ to mW at $t = t_{\mathrm{s}}$. We define this procedure to be the backward process.

Translating into `LAMMPS` script language, the forward process (steps 1-2) can be formulated as:

```
run t_eq
variable lambda1 equal ramp(1,0)
fix f4 all adapt 1 pair sw fscale 1 1 v_lambda1
variable lambda2 equal ramp(0,1)
fix f5 all adapt 1 pair ufm fscale 1 1 v_lambda2
run t_s
```

where `ramp(x,y)` is a function that linearly interpolates between the initial value `x` and the final value `y` during the simulation. The backward process (steps 3-4) is executed by repeating the same block of code, exchanging the values 1 and 0 in the `ramp` function.

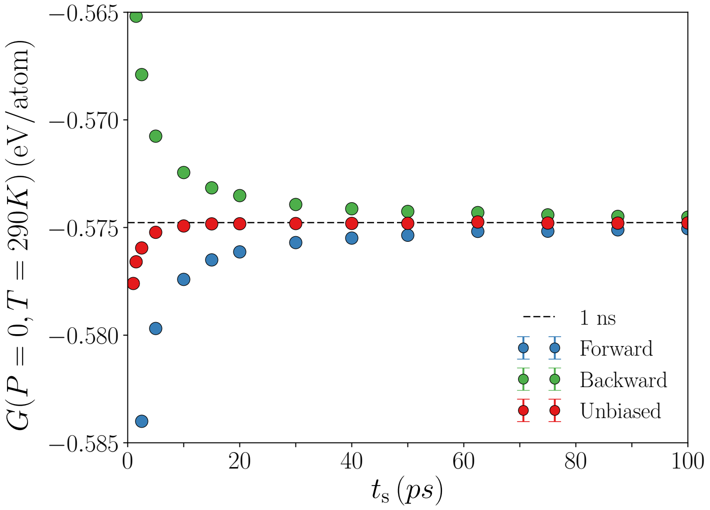

Figure 10 – (Color) Convergence of the absolute zero-pressure Gibbs free energy of the mW liquid at $T_0 = 290K$ obtained from NEHI simulations as function of the switching time t_s. Blue, green and red circles represent the results of forward, backward and the unbiased estimators, respectively. The error bars describing the statistical uncertainties are smaller than the symbol size for all cases. Dashed line represents unbiased estimator value for $t_\mathrm{s} = 1.0$ ns.

We perform ten independent AS realizations (forward and backward) for the fluid system composed of $8,000$ atoms at $T_0 = 290$ K and the corresponding mW equilibrium density determined from the NPT simulations mentioned above. Before both switching processes the fluid system is equilibrated during $t_\mathrm{eq} = 0.1$ ns. To verify the convergence of the free-energy in the AS processes, we execute simulations for different values of the switching time t_s. The results, depicted in Fig. 10, clearly show that the unbiased estimator, given by Eq. (3.10), converges very quickly. For instance, the relative difference between the results obtained for $t_\mathrm{s} = 30$ ps and 1.0 ns (shown as the dashed line) is less than 10^{-4}.

Finally, in step (iii), we use the RS method to construct the zero-pressure Gibbs free-energy curve as a function of temperature for the mW liquid. As detailed earlier, these simulations are carried out at constant temperature and zero pressure. In the `LAMMPS` script this is accomplished by repeating the same combination of `fix nph` and `langevin` that is used to determine the equilibrium density of the fluid for the AS calculations. The RS path involves scaling of the Hamiltonian of interest by a factor of λ according to Eq. (3.11). Here, for the mW liquid, the RS simulations are carried out at $T_0 = 290$K, for which the Gibbs free energy is known from the AS calculations described in the previous paragraph. The scaling parameter is varied between $\lambda = 1$ and $\lambda = \lambda_f = 290/260$, such that the covered temperature range is that between $T_0/\lambda = 290$ K and $T_0/\lambda_f = 260$ K. The temperature dependence of the Gibbs free energy is then computed using Eq. (3.16), where we use the reference value $G(T_0 = 290K) = -0.574774(4)$ eV/atom obtained from the NEHI simulations with a switching time of $t_\mathrm{s} = 1.0$ ns and determine the dynamical

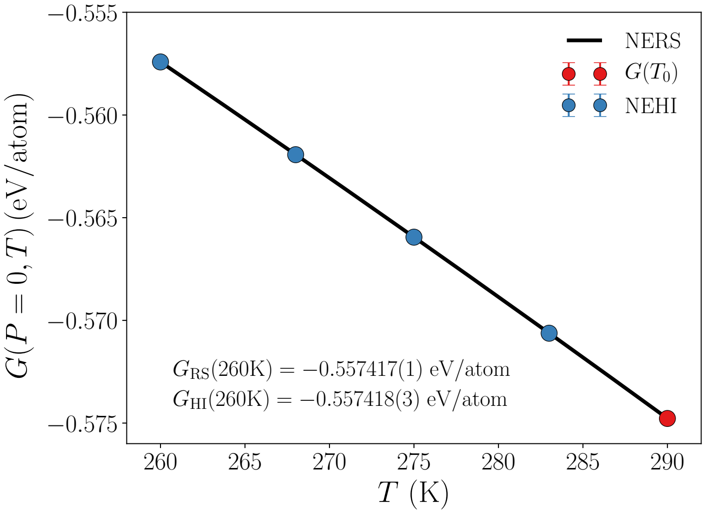

Figure 11 – (Color) Gibbs free-energy per atom of the mW liquid as a function of T at zero pressure. Black line depicts RS results. Red circle is the reference free-energy value $G(P=0,T_0)$ used in Eq. (3.16). Blue circles represent free-energy values obtained using independent HI calculations. In all cases, the error bars are smaller than the symbol size.

work values $W^{1\to\lambda}$ and $W^{\lambda\to 1}$ as the average over ten independent realizations of the RS process in forward and backward directions.

The implementation of the RS procedure is based on the same sequence of steps (1-4) used for the Hamiltonian interpolation simulations described above. The only difference is that the RS simulations require only a single `fix adapt` command to scale the Hamiltonian of interest. Before initiating the scaling, the mW liquid system is first equilibrated for $t_{\mathrm{eq}} = 0.1$ ns. Next, the scaling parameter $\lambda(t)$ is varied linearly between 1 and λ_f during the switching time $t_{\mathrm{s}} = 1.0$ ns. Fig. 11 shows the resulting zero-pressure Gibbs free energy of the mW liquid as a function of temperature. To verify the accuracy of the RS results, we have carried out a number of additional NEHI simulations (using $t_{\mathrm{s}} = 1.0$ ns and 10 independent forward and backward realizations) to compute the Gibbs free energy in an independent manner at a number of temperatures in the interval between 260 and 290 K. The agreement between the results obtained in both methods is excellent. For instance, at $T = 260$ K, the relative error in the free-energy results is very small, $\sim 10^{-4}$ %, and both methods are in agreement within these error bars.

Having computed the Gibbs free energy of the liquid phase, we now determine the zero-pressure absolute Gibbs free energies as a function of temperature of the hexagonal-diamond (*hd*) and cubic-diamond (*cd*) structures. This is accomplished using procedures that are essentially identical to those applied to the calculations for the liquid phase, using HI to compute the Gibbs free energy at a reference temperature of $T_0 = 260$ K, followed by RS simulations to determine its further temperature dependence. The main difference concerns the reference system in the HI calculations, which, for the crystalline phases, is chosen to be the Einstein crystal with a spring constant of 0.5 eV/Å^2. For both the HI and RS runs the systems were equilibrated during 0.2 ns prior to the switching runs, and the unbiased dynamical-work estimators were obtained from 10 independent forward and backward realizations using switching times of 0.4 and 0.8 ns, respectively. Full details are described in Ref. [49].

With the resulting RS free-energy curves for both crystalline phases we now determine their respective melting points by analyzing the free-energy differences $\Delta G(T) \equiv G_{\text{solid}}(T) - G_{\text{liquid}}(T)$ as a function of temperature. The results are shown in Fig. 12, which displays $\Delta G(T)$ for the hexagonal and diamond cubic phases on the temperature interval between 260 and 290 K. A first observation is that, as expected, the *cd* phase is only metastable with respect to the *hd* structure, with the free-energy difference between them being positive across the entire temperature interval. Next, to determine the melting temperatures of both phases, we locate the temperature values T_m for which $\Delta G(T_m) = 0$. To this end we generate cubic-spline representations of the RS data for $\Delta G_{cd}(T)$ and $\Delta G_{hd}(T)$ and locate their respective roots, giving $T_m^{hd} = 276.13(3)$ K and $T_m^{cd} = 275.83(3)$ K, respectively. Here, the uncertainties in the final digit of the T_m values have been computed using the approach outlined in Ref. [87].

These results are consistent with the findings of Molinero and Moore [70], who reported the thermodynamic stability of the *hd* phase with respect to the *cd* structure. Our results for the melting-temperature values, however, are somewhat higher than those reported in Ref. [70], in particular for the *cd* phase: $T_m^{hd} = 275(1)$ K and $T_m^{cd} = 272(1)$ K. It is possible that these differences are related to the different methodologies used to determine the melting temperatures. The values reported in Ref. [70] were obtained using the coexistence method [89, 90], and it has been shown [91] that this approach gives rise to melting temperatures that are systematically lower compared to those extracted from free-energy computations.

3.4.2 Cu-Zr liquid alloy

In the second application we apply the formalism to compute the free energy of a multicomponent fluid. In particular we consider the Cu-Zr liquid alloys systems, which have shown to be valuable in engineering applications due to their ability to produce bulk metallic glasses. We apply the nonequilibrium HI procedure to calculate the zero-pressure Gibbs free energy of Cu-Zr alloys as a function of composition at $T = 1800$ K, using the UF model as reference system.

All simulations were performed using computational cells containing $2,000$ atoms subject to periodic boundary conditions, with varying Cu and Zr fractions, X_{Cu} and X_{Zr}, respectively. The interparticle interactions are modeled by a second nearest-neighbor

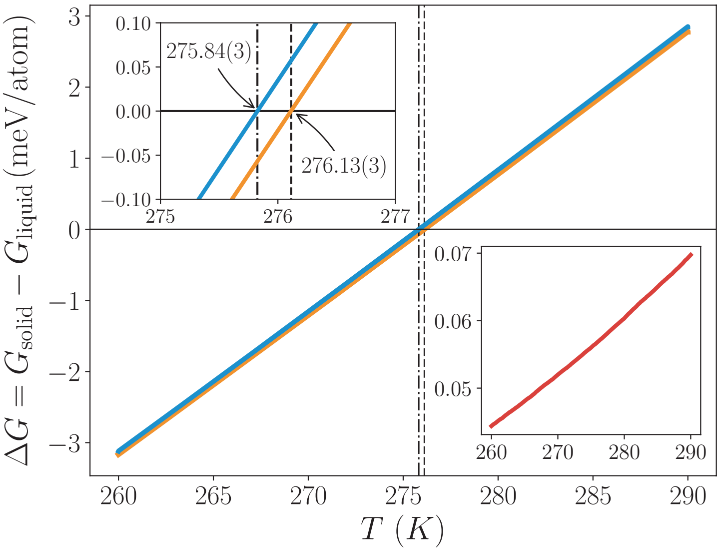

Figure 12 – (Color) Zero-pressure Gibbs free-energy difference between solid and liquid phases of mW model as a function of T. Orange and blue lines are RS curves for *hd* and *cd* structures, respectively. Vertical dashed lines describe the melting-temperature locations for both phases. Upper inset provides a zoom of melting point region. Red line in lower inset shows the free-energy difference $G_{cd}(T) - G_{hd}(T)$ between the *cd* and *hd* phases as a function of temperature.

modified embedded atom method (MEAM-2NN) potential [73], with the adjustable parameters for Cu-Cu, Zr-Zr and Cu-Zr interactions taken from Refs. [75], [92] and [93], respectively. The equations of motion are integrated using a time step of $\Delta t = 1$ fs.

As we are interested in determining the Gibbs free energy under zero pressure, we repeat the same procedure followed in Section 3.4.1, first performing equilibrium NPT MD simulations for a set of different Zr fractions to obtain the corresponding equilibrium number density, followed by NVT nonequilibrium HI calculations at the determined equilibrium density to compute the free-energy difference between the Cu-Zr liquid and the UF model. The HI procedure can be implemented in `LAMMPS` script following the same commands used for the mW model, with `fix adapt` scaling the respective intermolecular potentials. However, since we are dealing with more than one species, we must first define the reference interactions for all three types. Here, we choose all of them to be identical UF interactions, i.e., the reference UF models for the Cu-Cu, Zr-Zr and Cu-Zr interactions are specified by the same values of p, and σ. With this choice only the ideal gas part of the absolute reference free energy, Eq. (3.24), depends on the composition, whereas the interaction part does not. This can be encoded in the `fix adapt` command by introducing asterisks in its pair-type option:

```
fix f4 all adapt 1 pair meam/c fscale * * &
v_lambda
```

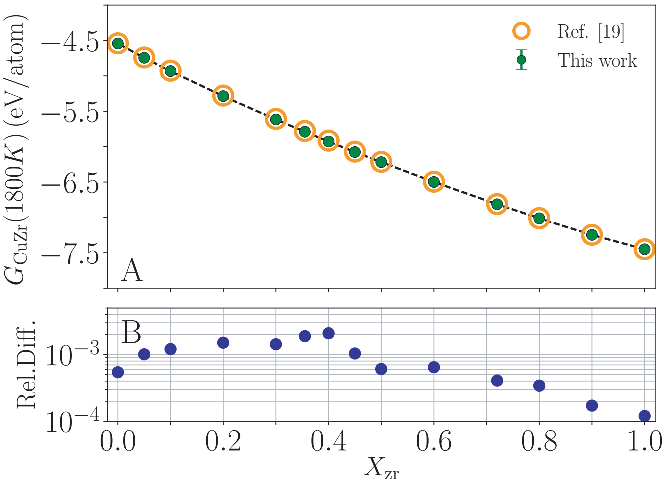

Figure 13 – (Color) Zero pressure Gibbs free energy as a function of Zr composition for the Cu-Zr liquid alloy at $T = 1800$ K. (A) Green circles depict present results, orange circles are data taken from Ref. [3]. The latter were computed using the equilibrium HI method with the LJ potential as a reference system. Dashed line serves as a guide to the eye. (B) Relative differences between present values and those of Ref. [3].

```
fix f5 all adapt 1 pair ufm fscale * * v_lambda
```

Using this approach we compute the free energy of the Cu-Zr liquid according to the unbiased estimator obtained from ten independent forward and backward realizations of the nonequilibrium HI process. We employ the parameter values $p = 50$ and $\sigma = 1.5$ Å for the UF reference. Before all HI realizations, the systems are first equilibrated for an interval $t_{\mathrm{eq}} = 0.1$ ns, followed by an HI process with a switching time $t_{\mathrm{s}} = 0.5$ ns.

The results are presented in Fig. 13 and Table 2. In the former, the present absolute free-energy values are compared to previously published data [3], which were obtained using equilibrium thermodynamic integration techniques using the LJ system as a reference system. Table 2 gives the corresponding numerical values. The agreement between the present calculations and the results reported in Ref. [3] is excellent, with relative discrepancies below ~ 0.1 % for all compositions.

3.4.3 Rigid models: TIP4P and SPC/E

After the applications concerning atomic fluids we now turn to the case of molecular liquids, focusing in particular on liquid water. First, we consider two popular rigid-molecule descriptions, namely the SPC/E [77] and TIP4P [76] models, and compute their Helmholtz free energies as a function of temperature at a fixed density.

The adopted computational approach is similar to the 2-step procedure em-

Table 2 – Numerical values as a function of Zr composition. The columns contain, respectively, the atomic equilibrium number density ρ, the per-particle Gibbs free-energy difference ΔG between the MEAM and UF fluids, the absolute per-particle free-energy values for the UF model and the absolute zero-pressure Gibbs free energy per particle for the Cu-Zr liquid.

	ρ	$\Delta G/N$	G_{UF}/N	G_{MEAM}/N
X_{Zr}	atom/ Å^3	eV/atom	eV/atom	eV/atom
1.000	0.03973	-6.28736(6)	-1.16271	-7.45007(6)
0.900	0.04221	-6.15403(2)	-1.09207	-7.24610(3)
0.800	0.04460	-6.01482(8)	-0.99818	-7.01300(8)
0.720	0.04665	-5.90906(9)	-0.90551	-6.81458(9)
0.600	0.05000	-5.76125(8)	-0.73821	-6.49947(8)
0.500	0.05306	-5.64985(9)	-0.57098	-6.22094(9)
0.450	0.05469	-5.59854(9)	-0.47768	-6.07622(9)
0.400	0.05635	-5.54842(2)	-0.37898	-5.92741(2)
0.355	0.05791	-5.50595(9)	-0.28401	-5.78996(9)
0.300	0.05987	-5.45625(3)	-0.16057	-5.61682(3)
0.200	0.06352	-5.36667(5)	0.08028	-5.28639(5)
0.100	0.06721	-5.27332(4)	0.34066	-4.93267(4)
0.050	0.06903	-5.22252(1)	0.47763	-4.74489(1)
0.000	0.07119	-5.19153(2)	0.64899	-4.54254(2)

Table 3 – Helmholtz free-energy per molecule of liquid water as described by the SPC/E and TIP4P models at density $d = 1.05$ g/cm^3, obtained using the NEHI method. The values originally reported in Ref. [5] does not include the rotational contribution to the free energy and used a generic value of 1 Å for the thermal de Broglie wavelength. To allow a comparison with the present results, we have added the kinetic contribution to the data from Ref. [5] and adjusted the thermal wave-length value.

	T	$\Delta F/N$	$F_{\mathrm{UF}}^{(\mathrm{exc})}/N$	$F_{\mathrm{UF/Rw}}/N$	F_{liquid}/N	Ref. [5]
Model	K	kcal/mol	kcal/mol	kcal/mol	kcal/mol	kcal/mol
SPC/E	225	-25.7214(2)	17.9129	-5.2191	-13.0276(2)	-13.0298
SPC/E	443	-39.8195(4)	35.2684	-12.0651	-16.6162(4)	-16.6276
TIP4P	225	-16.4540(3)	9.6944	-5.1712	-11.9308(3)	-11.9307
TIP4P	443	-22.8003(4)	19.0872	-11.9707	-15.6838(4)	-15.6961

ployed to compute the temperature dependence of the free energy for the mW model. First, we compute the Helmholtz free energy at a reference temperature T_0 using the NEHI method, transforming the water-model interactions into those of the rigid UF water reference model, UF/Rw, described in Sec. 3.3.2. Next, we apply the NERS technique to extrapolate the free energy for temperatures beyond the reference value T_0.

For all calculations we use computational cells containing 4,000 molecules at a density of 1.05 g/cm^3, subject to the usual periodic boundary conditions and the integration of the NVT MD equations of motion is carried out using a time step of $\Delta t = 2.0$ fs. The LJ parts for both rigid models are truncated, without shifting, at a cutoff distance of 8.5 Å and standard long-range corrections [6] are added to its energy. The long-range intermolecular electrostatic interactions are calculated using the particle-particle particle-mesh (PPPM) method [94] and intramolecular bond lengths and angles are held fixed using `SHAKE` [95] algorithm.

In `LAMMPS`, the SPC/E interactions with PPPM electrostatics are invoked by the generic `pair lj/cut/coul/long` and `kspace pppm` styles commands. For the TIP4P model, on the other hand, the model-specific `pair lj/cut/tip4p/long` and `kspace pppm/tip4p` styles are used. The UF/Rw reference interactions are activated by the `pair ufm/rw` command.

Although the implementation of the NEHI procedure for the water models into `LAMMPS` script is similar to that described in Sec. 3.4.1 for atomic fluids, additional care must be given to the PPPM force contributions. As the electrostatic interactions are switched off during the HI process, so must the long-range part evaluated using the PPPM method. This can be done using the same `fscale` variable introduced earlier in Sec. 3.4.1, invoking a `fix adapt` command in which the both the force contributions of the `pair` style as well as the long-range parts evaluated by the `kspace` command are scaled by the variable `lambda`:

```
fix f4 all adapt 1 pair lj/cut/coul/long &
fscale * * v_lambda kspace v_lambda fscale yes
```

For both rigid water models we compute the corresponding Helmholtz free-energy differences with respect to the UF/Rw system at $T = 225$ and 443 K, for which comparative free-energy information is available [5]. The values are computed using the unbiased estimator Eq. (3.10) obtained from 10 independent forward and backward NEHI realizations with a switching time of $t_s = 0.6$ ns. The parameters for the UF model are chosen to be $\sigma = 2.0$ Å and $p = 50$ for both cases. The corresponding absolute Helmholtz free-energy values for both rigid models are then computed using the absolute UF/Rw free energy given by Eq. (3.29), including the translational and rotational contributions to the ideal gas part. The resulting values for the absolute free-energy per molecule are given in Table 3 and they are in excellent agreement with those obtained by Vega *et al.* [5].

After the NEHI calculations for 225 and 443 K, we compute intermediate free-energy values form both models using the NEHI and NERS methods. For the former we consider the temperature $T = 250, 300, 350$ and 400 K and use the same simulation parameters as those used for 225 and 443 K. For the NERS computations we set $T_0 = 225$ K, obtaining the RS data from the unbiased estimator, Eq. (3.15), using 10 independent forward and backward simulations with $t_s = 1.0$ ns. The results, shown in Fig. 14, attest to the excellent agreement between both techniques.

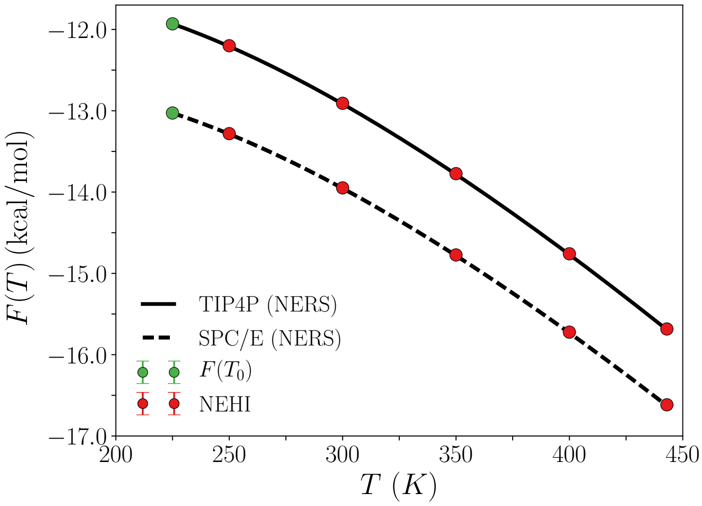

Figure 14 – (Color) Helmholtz free-energy per molecule as a function of temperature of TIP4P and SPC/E liquid water for a density of 1.05 g/cm^3. Circles represents results obtained using the NEHI approach. Error bars are smaller than symbol size. Full and dashed lines depict NERS results for the TIP4P and SPC/E models, respectively, using the green data points as reference values $F(T_0 = 225\text{ K})$ in Eq. (3.15).

3.4.4 Flexible model: q-SPC/Fw

Finally, we compute the free energy of the flexible q-SPC/Fw water model [78], using the flexible UF/Fw system as reference. The adopted simulation approach is the same as those used in the previous calculations, using the NEHI technique to compute the free-energy value at a specified temperature, followed by an NERS calculation to extrapolate its temperature dependence.

Compared to the calculations for the rigid water models the main difference is the computational cost. Due to the stiffness of the intramolecular harmonic interactions, the MD time step needs to be reduced substantially to preserve the numerical stability of the MD simulations. In particular, we use $\Delta t = 0.1$ fs, a factor 20 smaller than the value used for the rigid water molecules. Given the increased computational effort, we use a smaller simulation cell, containing 512 water molecules at a density of 1.0129 g/cm^3 subject to periodic boundary conditions.

We first compute the absolute Helmholtz free energy at $T = 255$ K applying the NEHI technique using the UF/Fw as the reference system, for which we set $p = 50$ and $\sigma = 1.8$ Å. The results are obtained from Eq. (3.41), using ten independent forward and backward realizations at a switching time of $t_s = 2.0$ ns. These results are shown in Table 4. Next, using the reference value at $T_0 = 255$ K, we extrapolate the temperature

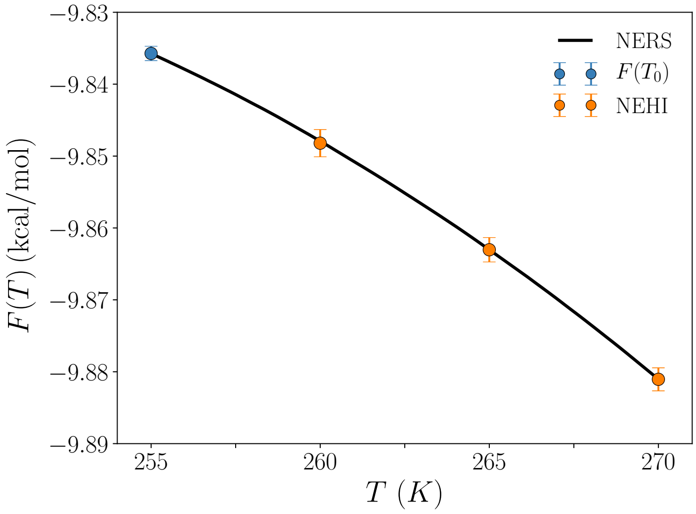

Figure 15 – (Color) Helmholtz free-energy per molecule as a function of temperature for q-SPC/Fw (full line) liquid water at a density of 1.0129 g/cm^3. Circles represent value computed using the NEHI technique. Solid line represents NERS results obtained using the reference free-energy value $F(T_0 = 255\text{ K})$ indicated by the green circle.

dependence up to $T = 270$ K using NERS with a switching time of $t_s = 1.0$ ns. The results are depicted in Fig. 15 and show excellent agreement between the NERS results and the NEHI data.

3.4.5 Efficiency of UF-based reference systems

The results presented in the previous sections attest to the applicability of UF-based models as reference systems in free-energy calculations for a variety of fluid-phase systems characterized by distinct interaction types.

Not only do they allow free-energy values to be computed with high-precision, they also do so in a computationally efficient manner. The rapid convergence of the unbiased estimator in Fig. 10 is illustrative for this, with trajectories as short as only 30 ps

Table 4 – Helmholtz free energy per molecule of the liquid q-SPC/Fw water model at density $d = 1.0129$ g/cm^3 for different temperatures as computed with the NEHI method using the UF/Rw system as reference.

T K	$\Delta F/N$ kcal/mol	$F_{\text{UF}}^{(\text{exc})}/N$ kcal/mol	$F_{\text{UF/Fw}}/N$ kcal/mol	F_{liquid}/N kcal/mol
255	-13.531(1)	6.061	-2.365	-9.835(1)
260	-13.556(2)	6.180	-2.472	-9.848(2)
265	-13.582(2)	6.298	-2.580	-9.864(2)
270	-13.610(2)	6.417	-2.688	-9.881(2)

giving results that are essentially indistinguishable from those obtained from simulations covering a time interval of 1.0 ns.

Quick convergence is a consequence of small dissipation along the switching simulations, characterizing a smooth, close-to-equilibrium process. For a given system of interest, the rate of dissipation in NEHI calculations is largely determined by the choice of reference system. While the above results indicate that the UF-based models are a competitive option, it is not clear whether or not they are superior to other common choices. To shed light on this issue we carry out a comparison with one of the most-frequently employed reference systems in fluid-phase free-energy calculations, namely those based on the LJ model.

For this purpose, we carry out additional NEHI simulations for the atomic $\mathrm{Cu}_{50}\mathrm{Zr}_{50}$ liquid alloy and molecular TIP4P fluid, respectively, using UF and LJ-based reference systems for both. In particular, as a quantitative measure of efficiency, we compute the ratio of the average dissipated heat $\overline{Q_{\mathrm{diss}}}$ and the unbiased estimator for the free-energy difference ΔF as a function of the switching time t_{s}. The values for $\overline{Q_{\mathrm{diss}}}$ and ΔF were computed using Eqs. (3.4) and (3.5), employing the results of ten independent forward and backward processes. The UF-model parameters are the same as those used in the previous sections. For the liquid alloy, the chosen parameters of the LJ reference coincide with those used in Ref. [3]. For the TIP4P liquid, the LJ parameters are the same as those in the proper model definition.

The resulting ratios are depicted in Fig. 16. They decay with increasing switching time according to $\sim 1/t_{\mathrm{s}}$, which is expected for nonequilibrium processes in the linear regime [53]. In addition, the UF-based references display superior efficiency for both cases, giving a substantially smaller dissipation for given switching time. Indeed, for the TIP4P case, the ratio obtained using the UF reference system is almost an order of magnitude smaller compared to the results for the LJ reference. These results clearly show that the UF model provides a general and efficient reference system for the calculation of fluid-phase free energies of systems characterized by different interactions.

3.5 Summary

In this paper we provide a guide for computing free energies of fluid-phase systems using nonequilibrium techniques within the `LAMMPS` MD simulation package. In addition to describing `LAMMPS` implementation details and making available the computational tools in the form of full source code, scripts and auxiliary files [86], we discuss in detail the purely repulsive, ultrasoft UF pair potential as a reference system. Not only is the UF model useful for atomic fluids, either monoatomic or mixtures, it can also be used to construct reference systems for molecular fluids.

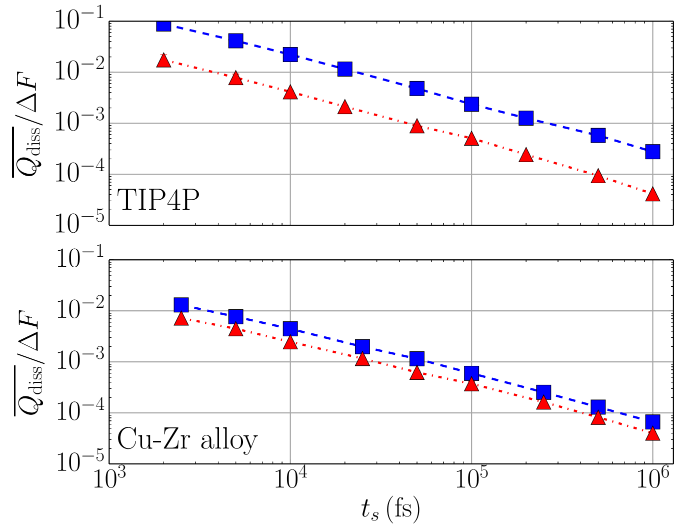

Figure 16 – (Color) Average ratios of dissipated heat and free-energy difference as a function of t_s in NEHI simulations using the UF (triangles) and LJ (squares) models as reference systems. Results are shown for the TIP4P water model and the $Cu_{50}Zr_{50}$ liquid alloy as described by the MEAM-2NN potential. Dashed lines serve as guides to the eye.

This is illustrated in a number of application in which free energies of fluids characterized by fundamentally different interaction types are computed. In particular, we consider atomic mW model for water, a binary liquid alloy described by an MEAM-2NN potential and three molecular models for water, two of them rigid and one flexible. In all applications, the corresponding UF-based reference systems provide smooth thermodynamic paths that allow accurate and efficient free-energy calculations. In particular, comparing the results to the LJ model, which has been frequently been used as a reference system for fluids, the UF is much smoother in that it leads to substantially lower dissipation in a given switching processes.

The techniques described in this paper, together with the supplied source code, scripts and post-processing files, provide a platform on which fluid-phase free energies can be easily and efficiently computed using the `LAMMPS` code. In addition to being useful for the development of new models for liquid phases, the tools may also find applications in the construction of community databases containing thermodynamic properties of existing models.

4 Conclusions and Outlook

In this chapter we summarize the main results and conclusions from this Thesis. We also present an Outlook that discusses possible future research directions based on the results obtained in this work.

4.1 UFM Phase Diagram

In chapter 2 state-of-the-art nonequilibrium free-energy calculation techniques, combined with a large-scale computing facilities, were employed to determine the phase diagram of the UFM. We find that a single fluid phase and the crystalline bcc and fcc structures are the only thermodynamically stable phases that can coexist at a unique triple point. The location of this triple-point was determined to high precision using a finite-size scaling procedure. In addition to these stable condensed forms, the hcp crystalline phase was found to be the only metastable crystalline phase of the UFM. Other structures, such as the diamond-cubic structure were observed to not even be mechanically stable.

Furthermore, the UFM equilibrium phase diagram is characterized by the existence of two reentrant transition sequences. The first concerns a fluid-bcc-fluid series as the density is increased for values of p close to 100. In addition, there is a bcc-fcc-bcc-fluid succession in the vicinity of the triple-point. These patterns are also found in other purely repulsive systems such as GCM, IPL and Yukawa model, as predicted by the corresponding-states principle. In particular, we find that the UFM and GCM have their melting curves and bcc-fcc phase boundaries effectively overlapping. Nevertheless, for increasing densities the bcc structures remains the stable thermodynamic phase on UFM's phase diagram and does not re-melt into the fluid as is the case for the GCM. This difference is due to the logarithmic divergence of the UFM for shorter distances as opposed to the finite value at zero separation for the GCM.

Finally, our results imply that the UFM is a very convenient reference system when performing fluid-phase free-energy calculations due to the fact that, for arbitrary densities, its single fluid phase is the only stable phase for $p \lesssim 100$. Accordingly, there are no liquid-gas or solid-fluid transitions under these conditions, eliminating the risk of encountering a first-order phase transition during the free-energy computations.

4.2 Fluid-phase Free-energy Calculations

In chapter 3 we successfully implemented a systematic and accurate method for computing free energies of fluid-phase systems using nonequilibrium techniques (Adiabatic Switching and Reversible Scaling) within the `LAMMPS` MD simulation package. To this end, we developed a family of reference systems based on UFM pairwise potential for atomic (monoatomic and mixtures) and molecular (rigid and flexible water models) liquids. For each kind of these reference systems, we provide a set of free-energy expressions which are given by a sum of two contributions: an intermolecular part described by an accurate numerical representation of UFM free energy that is already known (atomic fluids) and an intramolecular contribution that can be determined analytically (molecular fluids).

The application of these methods are illustrated in the calculation of free energies of fluids characterized by fundamentally different interaction types. As a first illustration, we have calculated the free energies as a function of temperature for the liquid phase and crystalline hexagonal and cubic diamond structures of the mW water in order to determine its melting temperature. We find that hexagonal diamond is more stable than cubic diamond and that the melting temperature is 276.13(3) K, which is somewhat higher than that reported by Molinero and Moore, 275(1) K, using the coexistence method. This finding is consistent with other reports that melting-temperature results obtained using the coexistence method are systematically lower compared to free-energy-based results.

Further applications involved a multicomponent atomic fluid (MEAM-2NN) and different molecular versions of liquid water (TIP4P, SPC/E, q-SPC/Fw). For all of these cases we obtained a set of absolute free-energy values for a variety of temperatures (molecular fluids) and compositions (CuZr alloy) which was compared to other results reported in literature using equilibrium techniques and using the LJ potential as a reference system. We found the relative discrepancies to be below $\sim 0.1\%$, attesting the applicability of UF-based models as reference systems in free-energy calculations of fluids. Furthermore, comparing our results to the frequently used LJ model as a reference system, we find that the UFM is a more efficient reference system, leading to a substantially lower dissipation in a given switching processes.

Recently published papers have successfully used our method to compute free energies of fluid-phase systems in order to find efficient and accurate schemes to calculate the melting temperature [96], configurational and vibrational entropy of metallic glasses [97] and study fundamental atomistic events that leading to silicon crystallization [98]. Furthermore, these applications confirm that the p parameter of ~ 50 is the best choice to perform nonequilibrium free-energy calculations, leading to a small driving-force fluctuations and systematic errors.

4.3 Development of Computational Codes

All the results from molecular simulations presented in this Thesis were obtained using the `LAMMPS` package. To this end, since the UFM pairwise potential and its variants had never been implemented in the original `LAMMPS` versions, we implemented all necessary source codes and scripts to perform the free-energy calculations. Some of these functionalities are already available in the main `LAMMPS` distribution. All other tools are available on GitHub [86].

4.4 Outlook

The results in this Thesis provide a general and practical framework to carry out free-energy calculations of fluids. In particular, the UFM system has shown to be a versatile and robust reference system that allowing application to both atomic as well as molecular fluids. All the applications presented here concern semi-empirical interaction potentials. However, in principle, the UFM should also be useful for fluid systems described by first-principles quantum approaches such as density-functional theory (DFT) [17].

A particularly interesting prospect in this context involves the calculation of the melting temperature of water ice for a diverse set of exchange-correlation functionals. A tremendous amount of work is being done toward the development of functionals that accurately describe a variety of properties of water's condensed phases. The search for appropriate functionals is a very active area of investigation and, due to the enormous difficulty in describing hydrogen bonds, water-based systems are often chosen to assess the quality of a given functional. In most cases, they are tested for a number of properties of the liquid phase, water clusters of various sizes as well as a number of crystalline phases [99]. For the latter, most efforts have focused on quantities such as the sublimation energy and equilibrium volume. Melting temperatures, however, have not been taken into account in such assessments and, given that melting provides key insight into the strength of the cohesion in the solid phases. As such, DFT melting-temperature calculations would provide a stringent test for exchange-correlation functionals and we believe the UFM model and its variations can contribute to such efforts in the form of an efficient reference system for *ab initio* liquid water free-energy computations.

Bibliography

[1] Prestipino, S., Saija, F. & Giaquinta, P. V. Phase diagram of the gaussian-core model. *Phys. Rev. E* **71**, 050102 (2005). URL `http://link.aps.org/doi/10.1103/PhysRevE.71.050102`. pages 12, 25, 33, 34

[2] Khrapak, S. A., Chaudhuri, M. & Morfill, G. E. Communication: Universality of the melting curves for a wide range of interaction potentials. *J. Chem. Phys.* **134**, 241101 (2011). URL `http://dx.doi.org/10.1063/1.3605659`. `http://dx.doi.org/10.1063/1.3605659`. pages 12, 24, 25, 26, 33, 35, 36, 37

[3] Harvey, J.-P., Gheribi, A. E. & Chartrand, P. Accurate determination of the gibbs energy of cu–zr melts using the thermodynamic integration method in monte carlo simulations. *J. Chem. Phys.* **135**, 084502 (2011). URL `https://doi.org/10.1063/1.3624530`. `https://doi.org/10.1063/1.3624530`. pages 13, 39, 57, 62

[4] Khrapak, S. A. & Morfill, G. E. fcc-bcc-fluid triple point for model pair interactions with variable softness. *E* **100** (2012). URL `http://iopscience.iop.org/article/10.1209/0295-5075/100/66004/pdf`. pages 14, 24, 26, 33, 34, 35, 36, 37

[5] Vega, C., Sanz, E., Abascal, J. L. F. & Noya, E. G. Determination of phase diagrams via computer simulation: methodology and applications to water, electrolytes and proteins. *J. Phys.: Condens. Matter* **20**, 153101 (2008). URL `http://stacks.iop.org/0953-8984/20/i=15/a=153101`. pages 14, 38, 58, 59

[6] Allen, M. P., Tildesley, D. J. & Tildesley, D. J. *Computer Simulation of Liquids* (Oxford University Press, 2017). URL `https://books.google.com.br/books?id=PA9rnQAACAAJ`. pages 19, 20, 21, 22, 23, 39, 59

[7] Karplus, M. & McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Mol. Biol.* **9**, 646–652 (2002). URL `https://doi.org/10.1038/nsb0902-646`. pages 19

[8] Hollingsworth, S. A. & Dror, R. O. Molecular dynamics simulation for all. *Neuron* **99**, 1129 – 1143 (2018). URL `http://www.sciencedirect.com/science/article/pii/S0896627318306846`. pages 19

[9] Vlachakis, D., Bencurova, E., Papangelopoulos, N. & Kossida, S. Chapter seven - current state-of-the-art molecular dynamics methods and applications. vol. 94 of *Advances in Protein Chemistry and Structural Biology*, 269 – 313 (Academic Press, 2014). URL `http://www.sciencedirect.com/science/article/pii/B978012800168400007X`. pages 19

[10] Ma, Z. & Ranjith, P. G. Review of application of molecular dynamics simulations in geological sequestration of carbon dioxide. *Fuel* **255**, 115644 (2019). URL `http://www.sciencedirect.com/science/article/pii/S0016236119309962`. pages 19

[11] Hospital, G. J. R. O. M. . G. J. L., A. Molecular dynamics simulations: advances and applications. advances and applications in bioinformatics and chemistry. *AABC* **8**, 37–47 (2015). URL `https://doi.org/10.2147/AABC.S70333`. pages 19

[12] Komanduri, R. & Raff, L. M. A review on the molecular dynamics simulation of machining at the atomic scale. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* **215**, 1639–1672 (2001). URL `https://doi.org/10.1177/095440540121501201`. `https://doi.org/10.1177/095440540121501201`. pages 19

[13] Tuckerman, M. *Statistical Mechanics: Theory and Molecular Simulation*. Oxford Graduate Texts (OUP Oxford, 2010). URL `https://books.google.com.br/books?id=Lo3Jqc0pgrcC`. pages 20, 22, 23

[14] Zepeda-Ruiz, L. A., Stukowski, A., Oppelstrup, T. & Bulatov, V. V. Probing the limits of metal plasticity with molecular dynamics simulations. *Nature* **550**, 492 (2017). URL `http://dx.doi.org/10.1038/nature23472`. pages 20

[15] Stillinger, F. H. & Weber, T. A. Computer simulation of local order in condensed phases of silicon. *Phys. Rev. B* **31**, 5262–5271 (1985). URL `https://link.aps.org/doi/10.1103/PhysRevB.31.5262`. pages 20

[16] Tersoff, J. New empirical approach for the structure and energy of covalent systems. *Phys. Rev. B* **37**, 6991–7000 (1988). URL `https://link.aps.org/doi/10.1103/PhysRevB.37.6991`. pages 20

[17] Martin, R. M. *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2004). pages 20, 66

[18] Goldstein, H., Poole, C. & Safko, J. *Classical Mechanics: Pearson New International Edition* (Pearson Education Limited, 2014). URL `https://books.google.com.br/books?id=Xr-pBwAAQBAJ`. pages 22

[19] Swope, W. C., Andersen, H. C., Berens, P. H. & Wilson, K. R. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics* **76**, 637–649 (1982). URL `https://doi.org/10.1063/1.442716`. `https://doi.org/10.1063/1.442716`. pages 22

[20] Schneider, T. & Stoll, E. Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. *Phys. Rev. B* **17**, 1302–1322 (1978). URL `https://link.aps.org/doi/10.1103/PhysRevB.17.1302`. pages 22, 49

[21] Paula Leite, R., Santos-Flórez, P. A. & de Koning, M. Uhlenbeck-ford model: Phase diagram and corresponding-states analysis. *Phys. Rev. E* **96**, 032115 (2017). URL `https://link.aps.org/doi/10.1103/PhysRevE.96.032115`. pages 25, 39, 43, 44

[22] Uhlenbeck, G. & Ford, G. *Studies in Statistical Mechanics - The Theory of Linear Graphs with Application to the Theory of the Virial Development of the Properties of Gases*, vol. Volume 2 (North-Holland, 1962). pages 25, 39, 43

[23] McQuarrie, D. *Statistical Mechanics* (University Science Books, 2000). URL `http://books.google.com.br/books?id=itcpPnDnJM0C`. pages 25, 27

[24] Baram, A. & Rowlinson, J. Studies of the gaussian model. *Mol. Phys.* **74**, 707–713 (1991). URL `http://dx.doi.org/10.1080/00268979100102521`. `http://dx.doi.org/10.1080/00268979100102521`. pages 25

[25] Leite, R. P., Freitas, R., Azevedo, R. & de Koning, M. The uhlenbeck-ford model: Exact virial coefficients and application as a reference system in fluid-phase free-energy calculations. *J. Chem. Phys.* **145**, 194101 (2016). URL `http://dx.doi.org/10.1063/1.4967775`. `http://dx.doi.org/10.1063/1.4967775`. pages 25, 27, 29, 39, 43, 44, 46, 50

[26] Stillinger, F. H. Phase transitions in the gaussian core system. *J. Chem. Phys.* **65**, 3968–3974 (1976). URL `http://dx.doi.org/10.1063/1.432891`. `http://dx.doi.org/10.1063/1.432891`. pages 25

[27] Lang, A., Likos, C. N., Watzlawek, M. & Löwen, H. Fluid and solid phases of the gaussian core model. *J. Phys.: Condens. Matter* **12**, 5087 (2000). URL `http://stacks.iop.org/0953-8984/12/i=24/a=302`. pages 25

[28] Prestipino, S., Saija, F. & Giaquinta, P. V. Phase diagram of softly repulsive systems: The gaussian and inverse-power-law potentials. *J. Chem. Phys.* **123**, 144110 (2005). URL `http://dx.doi.org/10.1063/1.2064639`. `http://dx.doi.org/10.1063/1.2064639`. pages 25, 35

[29] Zachary, C. E., Stillinger, F. H. & Torquato, S. Gaussian core model phase diagram and pair correlations in high euclidean dimensions. *J. Chem. Phys.* **128**, 224505 (2008). URL `http://dx.doi.org/10.1063/1.2928843`. `http://dx.doi.org/10.1063/1.2928843`. pages 25

[30] Bharadwaj, A. S. & Singh, Y. Density-functional theory for fluid-solid and solid-solid phase transitions. *Phys. Rev. E* **95**, 032120 (2017). URL `http://link.aps.org/doi/10.1103/PhysRevE.95.032120`. pages 25

[31] Dubin, D. H. E. & Dewitt, H. Polymorphic phase transition for inverse-power-potential crystals keeping the first-order anharmonic correction to the free energy. *Phys. Rev. B* **49**, 3043–3048 (1994). URL `http://link.aps.org/doi/10.1103/PhysRevB.49.3043`. pages 25, 35

[32] Hoover, W. G. *et al.* Soft sphere equation of state. *J. Chem. Phys.* **52**, 4931–4941 (1970). URL `http://dx.doi.org/10.1063/1.1672728`. `http://dx.doi.org/10.1063/1.1672728`. pages 25

[33] Hoover, W. G., Young, D. A. & Grover, R. Statistical mechanics of phase diagrams. i. inverse power potentials and the close-packed to body-centered cubic transition. *J. Chem. Phys.* **56**, 2207–2210 (1972). URL `http://dx.doi.org/10.1063/1.1677521`. `http://dx.doi.org/10.1063/1.1677521`. pages 25

[34] Hoover, W. G., Gray, S. G. & Johnson, K. W. Thermodynamic properties of the fluid and solid phases for inverse power potentials. *J. Chem. Phys.* **55**, 1128–1136 (1971). URL `http://dx.doi.org/10.1063/1.1676196`. `http://dx.doi.org/10.1063/1.1676196`. pages 25

[35] Laird, B. B. & Haymet, A. Phase diagram for the inverse sixth power potential system from molecular dynamics computer simulation. *Mol. Phys.* **75**, 71–80 (1992). URL `http://dx.doi.org/10.1080/00268979200100071`. `http://dx.doi.org/10.1080/00268979200100071`. pages 25

[36] Agrawal, R. & Kofke, D. A. Solid-fluid coexistence for inverse-power potentials. *Phys. Rev. Lett.* **74**, 122–125 (1995). URL `http://link.aps.org/doi/10.1103/PhysRevLett.74.122`. pages 25

[37] Agrawal, R. & Kofke, D. A. Thermodynamic and structural properties of model systems at solid-fluid coexistence. *Mol. Phys.* **85**, 23–42 (1995). URL `http://dx.doi.org/10.1080/00268979500100911`. `http://dx.doi.org/10.1080/00268979500100911`. pages 25

[38] Wang, D. C. & Gast, A. P. Crystallization of power-law fluids: A modified weighted density approximation model with a solid reference state. *J. Chem. Phys.* **110**, 2522–2528 (1999). URL `http://dx.doi.org/10.1063/1.477957`. `http://dx.doi.org/10.1063/1.477957`. pages 25

[39] Hamaguchi, S., Farouki, R. T. & Dubin, D. H. E. Phase diagram of yukawa systems near the one-component-plasma limit revisited. *J. Chem. Phys.* **105**, 7641–7647

(1996). URL http://dx.doi.org/10.1063/1.472802. http://dx.doi.org/10.1063/1.472802. pages 25

[40] Hamaguchi, S., Farouki, R. T. & Dubin, D. H. E. Triple point of yukawa systems. *Phys. Rev. E* **56**, 4671–4682 (1997). URL http://link.aps.org/doi/10.1103/PhysRevE.56.4671. pages 25, 35

[41] Hoy, R. S. & Robbins, M. O. Fcc-bcc transition for yukawa interactions determined by applied strain deformation. *Phys. Rev. E* **69**, 056103 (2004). URL http://link.aps.org/doi/10.1103/PhysRevE.69.056103. pages 25, 35

[42] Hynninen, A.-P. & Dijkstra, M. Phase diagrams of hard-core repulsive yukawa particles. *Phys. Rev. E* **68**, 021407 (2003). URL http://link.aps.org/doi/10.1103/PhysRevE.68.021407. pages 25

[43] Meijer, E. J. & Frenkel, D. Melting line of yukawa system by computer simulation. *J. Chem. Phys.* **94**, 2269–2271 (1991). URL http://dx.doi.org/10.1063/1.459898. http://dx.doi.org/10.1063/1.459898. pages 25

[44] Malescio, G., Saija, F. & Prestipino, S. Anomalous melting behavior under extreme conditions: Hard matter turning "soft". *The Journal of Chemical Physics* **129**, 241101 (2008). URL http://dx.doi.org/10.1063/1.3050315. http://dx.doi.org/10.1063/1.3050315. pages 26

[45] Saija, F., Prestipino, S. & Malescio, G. Anomalous phase behavior of a soft-repulsive potential with a strictly monotonic force. *Phys. Rev. E* **80**, 031502 (2009). URL https://link.aps.org/doi/10.1103/PhysRevE.80.031502. pages 26

[46] Prestipino, S., Saija, F. & Malescio, G. Anomalous phase behavior in a model fluid with only one type of local structure. *The Journal of Chemical Physics* **133**, 144504– (2010). URL http://dx.doi.org/10.1063/1.3499830. pages 26

[47] Malescio, G. & Saija, F. A criterion for anomalous melting in systems with isotropic interactions. *The Journal of Physical Chemistry B* **115**, 14091–14098 (2011). URL http://dx.doi.org/10.1021/jp203939c. PMID: 21899259, http://dx.doi.org/10.1021/jp203939c. pages 26

[48] de Koning, M., Antonelli, A. & Yip, S. Optimized free-energy evaluation using a single reversible-scaling simulation. *Phys. Rev. Lett.* **83**, 3973–3977 (1999). URL https://link.aps.org/doi/10.1103/PhysRevLett.83.3973. pages 28, 42

[49] Freitas, R., Asta, M. & de Koning, M. Nonequilibrium free-energy calculation of solids using {LAMMPS}. *Comput. Mater. Sci.* **112, Part A**, 333 – 341 (2016). URL http://www.sciencedirect.com/science/article/pii/S0927025615007089. pages 28, 29, 30, 38, 39, 41, 55

[50] Watanabe, M. & Reinhardt, W. P. Direct dynamical calculation of entropy and free energy by adiabatic switching. *Phys. Rev. Lett.* **65**, 3301–3304 (1990). URL `http://link.aps.org/doi/10.1103/PhysRevLett.65.3301`. pages 28, 38

[51] de Koning, M. & Antonelli, A. Einstein crystal as a reference system in free energy estimation using adiabatic switching. *Phys. Rev. E* **53**, 465– (1996). URL `http://link.aps.org/doi/10.1103/PhysRevE.53.465`. pages 28, 38

[52] de Koning, M. & Antonelli, A. Adiabatic switching applied to realistic crystalline solids: Vacancy-formation free energy in copper. *Phys. Rev. B* **55**, 735–744 (1997). URL `http://link.aps.org/doi/10.1103/PhysRevB.55.735`. pages 28

[53] de Koning, M. Optimizing the driving function for nonequilibrium free-energy calculations in the linear regime: A variational approach. *J. Chem. Phys.* **122**, 104106 (2005). URL `https://doi.org/10.1063/1.1860556`. `https://doi.org/10.1063/1.1860556`. pages 29, 38, 40, 41, 62

[54] Frenkel, D. & Ladd, A. J. C. New monte carlo method to compute the free energy of arbitrary solids. application to the fcc and hcp phases of hard spheres. *J. Chem. Phys.* **81**, 3188–3193 (1984). URL `http://dx.doi.org/10.1063/1.448024`. `http://dx.doi.org/10.1063/1.448024`. pages 29, 38

[55] Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1 – 19 (1995). URL `http://www.sciencedirect.com/science/article/pii/S002199918571039X`. pages 29, 39

[56] Leite, R. P. & de Koning, M. Nonequilibrium free-energy calculations of fluids using lammps. *Computational Materials Science* **159**, 316 – 326 (2019). URL `http://www.sciencedirect.com/science/article/pii/S0927025618308085`. pages 38

[57] Kofke, D. A. Free energy methods in molecular simulation. *Fluid Phase Equilib.* **228-229**, 41 – 48 (2005). URL `http://www.sciencedirect.com/science/article/pii/S0378381204004431`. PPEPPD 2004 Proceedings. pages 38

[58] Kirkwood, J. G. Statistical mechanics of fluid mixtures. *J. Chem. Phys.* **3**, 300–313 (1935). URL `http://dx.doi.org/10.1063/1.1749657`. `http://dx.doi.org/10.1063/1.1749657`. pages 38, 40

[59] Frenkel, D. & Smit, B. *Understanding Molecular Simulation* (Academic Press, San Diego, San Diego, 2002). pages 38

[60] Li, L., Totton, T. & Frenkel, D. Computational methodology for solubility prediction: Application to the sparingly soluble solutes. *J. Chem. Phys.* **146**, 214110 (2017). URL `https://doi.org/10.1063/1.4983754`. pages 38

[61] Jarzynski, C. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **78**, 2690– (1997). URL `http://link.aps.org/doi/10.1103/PhysRevLett.78.2690`. pages 38

[62] Jarzynski, C. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Phys. Rev. E* **56**, 5018– (1997). URL `http://link.aps.org/doi/10.1103/PhysRevE.56.5018`. pages 38

[63] Jarzynski, C. Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale. *Annu. Rev. Condens. Matter Phys.* **2**, 329–351 (2011). URL `http://dx.doi.org/10.1146/annurev-conmatphys-062910-140506`. pages 38

[64] Vega, C. & Noya, E. G. Revisiting the frenkel-ladd method to compute the free energy of solids: The einstein molecule approach. *J. Chem. Phys.* **127**, 154113 (2007). URL `http://scitation.aip.org/content/aip/journal/jcp/127/15/10.1063/1.2790426`. pages 38

[65] Abramo, M. C. *et al.* On the determination of phase boundaries via thermodynamic integration across coexistence regions. *J. Chem. Phys.* **142**, 214502 (2015). URL `https://doi.org/10.1063/1.4921884`. `https://doi.org/10.1063/1.4921884`. pages 38

[66] Broughton, J. Q. & Gilmer, G. H. Molecular dynamics investigation of the crystal–fluid interface. i. bulk properties. *J. Chem. Phys.* **79**, 5095–5104 (1983). URL `https://doi.org/10.1063/1.445633`. `https://doi.org/10.1063/1.445633`. pages 39

[67] Broughton, J. Q. & Li, X. P. Phase diagram of silicon by molecular dynamics. *Phys. Rev. B* **35**, 9120–9127 (1987). URL `https://link.aps.org/doi/10.1103/PhysRevB.35.9120`. pages 39

[68] Habershon, S. & Manolopoulos, D. E. Free energy calculations for a flexible water model. *Phys. Chem. Chem. Phys.* **13**, 19714–19727 (2011). URL `http://dx.doi.org/10.1039/C1CP21520E`. pages 39, 45

[69] Johnson, J. K., Zollweg, J. A. & Gubbins, K. E. The lennard-jones equation of state revisited. *Mol. Phys.* **78**, 591–618 (1993). URL `http://dx.doi.org/10.1080/00268979300100411`. pages 39

[70] Molinero, V. & Moore, E. B. Water modeled as an intermediate element between carbon and silicon. *J. Phys. Chem. B* **113**, 4008–4016 (2009). URL `https://doi.org/10.1021/jp805227c`. PMID: 18956896, `https://doi.org/10.1021/jp805227c`. pages 39, 55

[71] Baskes, M. I. Modified embedded-atom potentials for cubic materials and impurities. *Phys. Rev. B* **46**, 2727–2742 (1992). URL https://link.aps.org/doi/10.1103/PhysRevB.46.2727. pages 39

[72] Baskes, M. Atomistic potentials for the molybdenum–silicon system. *Materials Science and Engineering: A* **261**, 165 – 168 (1999). URL http://www.sciencedirect.com/science/article/pii/S0921509398010624. pages 39

[73] Lee, B.-J. & Baskes, M. I. Second nearest-neighbor modified embedded-atom-method potential. *Phys. Rev. B* **62**, 8564–8567 (2000). URL https://link.aps.org/doi/10.1103/PhysRevB.62.8564. pages 39, 56

[74] Lee, B.-J., Baskes, M., Kim, H. & Koo Cho, Y. Second nearest-neighbor modified embedded atom method potentials for bcc transition metals. *Phys. Rev. B* **64**, 184102 (2001). URL https://link.aps.org/doi/10.1103/PhysRevB.64.184102. pages 39

[75] Lee, B.-J., Shim, J.-H. & Baskes, M. I. Semiempirical atomic potentials for the fcc metals cu, ag, au, ni, pd, pt, al, and pb based on first and second nearest-neighbor modified embedded atom method. *Phys. Rev. B* **68**, 144112 (2003). URL https://link.aps.org/doi/10.1103/PhysRevB.68.144112. pages 39, 56

[76] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983). URL https://doi.org/10.1063/1.445869. https://doi.org/10.1063/1.445869. pages 39, 57

[77] Berendsen, H. J. C., Grigera, J. R. & Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **91**, 6269–6271 (1987). URL https://doi.org/10.1021/j100308a038. pages 39, 57

[78] Paesani, F., Zhang, W., Case, D. A., Cheatham, T. E. & Voth, G. A. An accurate and simple quantum model for liquid water. *J. Chem. Phys.* **125**, 184507 (2006). URL https://doi.org/10.1063/1.2386157. https://doi.org/10.1063/1.2386157. pages 39, 60

[79] Hummer, G. Fast-growth thermodynamic integration: Error and efficiency analysis. *J. Chem. Phys.* **114**, 7330–7337 (2001). URL https://doi.org/10.1063/1.1363668. pages 41

[80] Oberhofer, H., Dellago, C. & Geissler, P. L. Biased sampling of nonequilibrium trajectories: Can fast switching simulations outperform conventional free energy calculation methods? *J. Phys. Chem. B* **109**, 6902–6915 (2005). URL http://dx.doi.org/10.1021/jp044556a. pages 41

[81] Daura, X., Affentranger, R. & Mark, A. E. On the relative merits of equilibrium and non-equilibrium simulations for the estimation of free-energy differences. *ChemPhysChem* **11**, 3734–3743 (2010). URL https://doi.org/10.1002/cphc.201000562. pages 41

[82] Dellago, C. & Hummer, G. Computing equilibrium free energies using non-equilibrium molecular dynamics **16** (2014). pages 41

[83] de Koning, M., Antonelli, A. & Yip, S. Single-simulation determination of phase boundaries: A dynamic clausius–clapeyron integration method. *J. Chem. Phys.* **115**, 11025–11035 (2001). URL https://doi.org/10.1063/1.1420486. https://doi.org/10.1063/1.1420486. pages 42, 43

[84] Powers, J. & Sen, M. *Mathematical Methods in Engineering* (Cambridge University Press, 2015). pages 48, 87

[85] Farhi, A. & Singh, B. Calculation of molecular free energies in classical potentials. *New J. Phys.* **18**, 023039– (2016). URL http://stacks.iop.org/1367-2630/18/i=2/a=023039. pages 48

[86] GitHub. For full source code, scripts and auxiliary files, see http://www.github.com/plrodolfo/FluidFreeEnergyforLAMMPS. pages 48, 62, 66

[87] Ryu, S. & Cai, W. Comparison of thermal properties predicted by interatomic potential models. *Modell. Simul. Mater. Sci. Eng.* **16**, 085005 (2008). URL http://stacks.iop.org/0965-0393/16/i=8/a=085005. pages 49, 55

[88] Shinoda, W., Shiga, M. & Mikami, M. Rapid estimation of elastic constants by molecular dynamics simulation under constant stress. *Phys. Rev. B* **69**, 134103– (2004). URL https://link.aps.org/doi/10.1103/PhysRevB.69.134103. pages 49

[89] Morris, J. R., Wang, C. Z., Ho, K. M. & Chan, C. T. Melting line of aluminum from simulations of coexisting phases. *Phys. Rev. B* **49**, 3109–3115 (1994). URL https://link.aps.org/doi/10.1103/PhysRevB.49.3109. pages 55

[90] Morris, J. R. & Song, X. The melting lines of model systems calculated from coexistence simulations. *The Journal of Chemical Physics* **116**, 9352–9358 (2002). URL https://doi.org/10.1063/1.1474581. pages 55

[91] García Fernández, R., Abascal, J. L. F. & Vega, C. The melting point of ice ih for common water models calculated from direct coexistence of the solid-liquid interface. *J. Chem. Phys.* **124**, 144506 (2006). URL http://scitation.aip.org/content/aip/journal/jcp/124/14/10.1063/1.2183308. pages 55

[92] Kim, Y.-M., Lee, B.-J. & Baskes, M. I. Modified embedded-atom method interatomic potentials for ti and zr. *Phys. Rev. B* **74**, 014101 (2006). URL `https://link.aps.org/doi/10.1103/PhysRevB.74.014101`. pages 56

[93] Kim, Y.-M. & Lee, B.-J. A modified embedded-atom method interatomic potential for the cu–zr system. *J. Mater. Res.* **23**, 1095–1104 (2008). pages 56

[94] Hockney, R. W. *Computer Simulation Using Particles* (Taylor & Francis Group, 2017). URL `https://books.google.com.br/books?id=dt9-tAEACAAJ`. pages 59

[95] Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977). URL `http://www.sciencedirect.com/science/article/pii/0021999177900985`. pages 59

[96] Zou, Y., Xiang, S. & Dai, C. Investigation on the efficiency and accuracy of methods for calculating melting temperature by molecular dynamics simulation. *Computational Materials Science* **171**, 109156 (2020). URL `http://www.sciencedirect.com/science/article/pii/S0927025619304550`. pages 65

[97] Alvarez-Donado, R. & Antonelli, A. Splitting up entropy into vibrational and configurational contributions in bulk metallic glasses: A thermodynamic approach. *Phys. Rev. Research* **2**, 013202 (2020). URL `https://link.aps.org/doi/10.1103/PhysRevResearch.2.013202`. pages 65

[98] Freitas, R. & Reed, E. J. Uncovering atomistic mechanisms of crystallization using machine learning (2019). `1909.05915`. pages 65

[99] Gillan, M. J., AlfÃÂ, D. & Michaelides, A. Perspective: How good is dft for water? *J. Chem. Phys.* **144**, 130901 (2016). URL `http://scitation.aip.org/content/aip/journal/jcp/144/13/10.1063/1.4944633`. pages 66

[100] Greiner, R. D. N. L., W. & Stocker, H. *Thermodynamics and Statistical Mechanics* (Springer New York, 2012). pages 80

[101] Greiner, W. & Bromley, D. A. *Classical Mechanics: Systems of Particles and Hamiltonian Dynamics* (Springer, 2003). pages 80

[102] Honnell, K. G., Hall, C. K. & Dickman, R. On the pressure equation for chain molecules. *J. Chem. Phys.* **87**, 664–674 (1987). URL `https://doi.org/10.1063/1.453562`. `https://doi.org/10.1063/1.453562`. pages 84

Appendix

APPENDIX A – Statistical Mechanics

In this appendix we describe in full details the calculation of classical canonical partition function $Z(N, V, T)$ and Helmholtz free energy $F(N, V, T)$ of two interacting molecular systems: The rigid (UF/Rw) and flexible (UF/Fw) Uhlenbeck-Ford models for water.

A.1 Rigid-water/Uhlenbeck-Ford Model (UF/Rw)

A.1.1 Center of mass and inertia tensor

Firstly, we going to determinate the coordinates of the center of mass (COM) of the water molecule, composed of two protons with masses m_{H} each one rigidly connected to an oxygen atom with mass m_{O}, giving rise to a rigid body. To this, we chose the origin of our coordinates system to be centered at the oxigen atom with the x-axis parallel to the bissector line, as illustrate in Fig. 17. Thus, its COM is given by

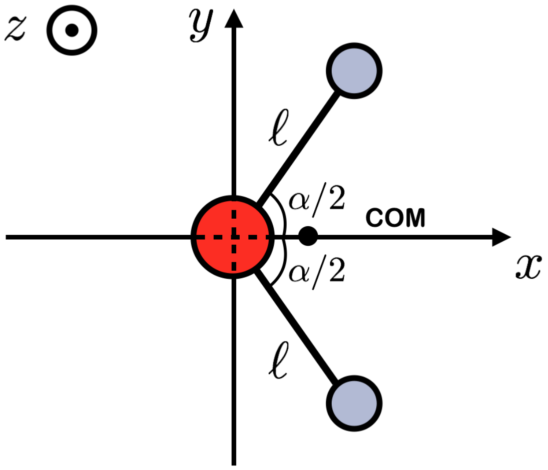

Figure 17 – Adopted coordinates to describe the rigid-body water model.

The center of mass is given by

$$\mathbf{X}_{\mathrm{COM}} = \frac{1}{M}\sum_{i=1}^{3} m_i\,\mathbf{r}_i, \tag{A.1}$$

where $M = m_O + 2m_H$ is the total mass of the water molecule. For each coordinate, it gives:

$$x_{\mathrm{COM}} = \frac{m_O(0) + m_H\ell\cos(\alpha/2) + m_H\ell\cos(-\alpha/2)}{M} = \frac{2\ell m_H\cos(\alpha/2)}{M}, \tag{A.2}$$

$$y_{\text{COM}} = \frac{m_O(0) + m_H\ell\text{sen}(\alpha/2) + m_H\ell\text{sen}(-\alpha/2)}{M} = 0, \tag{A.3}$$

$$z_{\text{COM}} = 0. \tag{A.4}$$

Furthermore, we want to calculate the Inertia tensor $\mathcal{I}$ for rotations over the axes passing thru the COM, which is given by

$$\mathcal{I} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} = \begin{pmatrix} \sum\limits_i [m_i(y_i^2 + z_i^2)] & \sum\limits_i (m_i x_i y_i) & \sum\limits_i (m_i x_i z_i) \\ \sum\limits_i (m_i y_i x_i) & \sum\limits_i [m_i(x_i^2 + z_i^2)] & \sum\limits_i (m_i y_i z_i) \\ \sum\limits_i (m_i z_i x_i) & \sum\limits_i (m_i z_i y_i) & \sum\limits_i [m_i(x_i^2 + y_i^2)] \end{pmatrix}.$$

In the COM frame, the atoms coordinates are: $O[-x_{CM}, 0, 0], H_1[\ell\cos{(\alpha/2)} - x_{CM}, \ell\text{sen}(\alpha/2), 0]$, and $H_2[\ell\cos{(\alpha/2)} - x_{CM}, -\ell\text{sen}(\alpha/2), 0]$. Using these values to obtain $\mathcal{I}$, for each element, we have:

$$\begin{aligned} I_{xx} &= m_O(0^2 + 0^2) + m_H(\ell^2\text{sen}^2(\alpha/2) + 0^2) + m_H(\ell^2\text{sen}^2(\alpha/2) + 0^2) \\ &= 2\ell^2 m_H\text{sen}^2(\alpha/2), \end{aligned} \tag{A.5}$$

$$\begin{aligned} I_{yy} &= m_O(x_{CM}^2 + 0^2) + m_H[(\ell\cos{(\alpha/2)} - x_{CM})^2 + 0^2] + \\ &+ m_H[(\ell\cos{(\alpha/2)} - x_{CM})^2 + 0^2] \\ &= 2\ell^2 m_H \cos^2{(\alpha/2)} \left(\frac{m_O}{M}\right), \end{aligned} \tag{A.6}$$

$$\begin{aligned} I_{zz} &= m_O(x_{CM}^2 + 0^2) + m_H[(\ell\cos{(\alpha/2)} - x_{CM})^2 + \ell^2\text{sen}^2(\alpha/2)] + \\ &+ m_H[(\ell\cos{(\alpha/2)} - x_{CM})^2 + \ell^2\text{sen}^2(\alpha/2)] \\ &= 2\ell^2 m_H[\text{sen}^2(\alpha/2) + \left(\frac{m_O}{M}\right)\cos^2{(\alpha/2)}] \\ &= I_{xx} + I_{yy}, \end{aligned} \tag{A.7}$$

$$\begin{aligned} I_{xy} &= m_O(x_{CM}.0) + m_H[(\ell\text{sen}(\alpha/2)).(\ell\cos{(\alpha/2)} - x_{CM})] + \\ &+ m_H[(-\ell\text{sen}(\alpha/2)).(\ell\cos{(\alpha/2)} - x_{CM})] = 0 \\ &= I_{yx} = I_{xz} = I_{zx} = I_{yz} = I_{zy}. \end{aligned} \tag{A.8}$$

Therefore,

$$\mathcal{I} = 2\ell^2 m_H \begin{pmatrix} \text{sen}^2(\alpha/2) & 0 & 0 \\ 0 & \left(\frac{m_O}{M}\right)\cos^2{(\alpha/2)} & 0 \\ 0 & 0 & \text{sen}^2(\alpha/2) + \left(\frac{m_O}{M}\right)\cos^2{(\alpha/2)} \end{pmatrix}.$$

A.1.2 Rotational Lagrangian and Hamiltonian

Afterwards, to determine the rotational Lagrangian of this rigid-body water molecule we need to invoke the Euler angles, which are defined as follows: The first, second and third rotations are performed about the z-axis, the new X-axis and Z'-axis, which one by ϕ, θ and ψ angles, respectively. Following reference [100, 101], if $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ define the laboratory system and $(\mathbf{i}', \mathbf{j}', \mathbf{k}')$ a body-fixed system of principal axes, after some algebra we get:

$$\begin{aligned} \mathbf{i} &= (\cos\phi\cos\psi - \mathrm{sen}\phi\cos\theta\mathrm{sen}\psi)\mathbf{i}' + (-\cos\phi\mathrm{sen}\psi - \mathrm{sen}\phi\cos\theta\cos\psi)\mathbf{j}' \\ &+ (\mathrm{sen}\phi\mathrm{sen}\theta)\mathbf{k}', \end{aligned} \tag{A.9}$$

$$\begin{aligned} \mathbf{j} &= (\mathrm{sen}\phi\cos\psi + \cos\phi\cos\theta\mathrm{sen}\psi)\mathbf{i}' + (-\mathrm{sen}\phi\mathrm{sen}\psi + \cos\phi\cos\theta\cos\psi)\mathbf{j}' \\ &- (\cos\phi\mathrm{sen}\theta)\mathbf{k}', \end{aligned} \tag{A.10}$$

$$\mathbf{k} = (\mathrm{sen}\theta\mathrm{sen}\psi)\mathbf{i}' + (\mathrm{sen}\theta\cos\psi)\mathbf{j}' + (\cos\theta)\mathbf{k}'. \tag{A.11}$$

In terms of matrices, the angular velocity ω in the body-fixed frame is given by

$$\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \begin{pmatrix} \mathrm{sen}\theta\mathrm{sen}\psi & \cos\psi & 0 \\ \mathrm{sen}\theta\cos\psi & -\mathrm{sen}\psi & 0 \\ \cos\theta & 0 & 1 \end{pmatrix} \begin{pmatrix} \dot\phi \\ \dot\theta \\ \dot\psi \end{pmatrix}.$$

Then, the rotational Lagrangian expression $\mathcal{L}_{\mathrm{rot}}$ of the rigid-body is then:

$$\begin{aligned} \mathcal{L}_{\mathrm{rot}} &= \frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2) \\ &= \frac{I_1}{2}(\dot\phi\mathrm{sen}\theta\mathrm{sen}\psi + \dot\theta\cos\psi)^2 + \frac{I_2}{2}(\dot\phi\mathrm{sen}\theta\cos\psi - \dot\theta\mathrm{sen}\psi)^2 \\ &+ \frac{I_3}{2}(\dot\phi\cos\theta + \dot\psi)^2 \end{aligned} \tag{A.12}$$

The rotational Hamiltonian is expressed by the following Legendre transform:

$$\mathcal{H}_{\mathrm{rot}} = \sum_i (\dot\varphi_i p_{\varphi_i}) - \mathcal{L}_{rot} = \dot\theta p_\theta + \dot\phi p_\phi + \dot\psi p_\psi - \mathcal{L}_{rot}, \tag{A.13}$$

where the canonical momenta is expressed as the derivative of rotational Lagrangian with respect to the time derivative of the Euler angles, $p_{\phi_i} = \partial\mathcal{L}_{rot}/\partial\dot\phi_i$. Then, using these results into Eq. (A.13) and writing this expression in a more fashion way, we obtain

$$\begin{aligned} \mathcal{H}_{\mathrm{rot}} &= \frac{\mathrm{sen}^2\psi}{2I_1}\left[p_\theta - \frac{\cos\psi}{\mathrm{sen}\theta\mathrm{sen}\psi}(p_\phi - p_\psi\cos\theta)\right]^2 + \\ &+ \frac{\cos^2\psi}{2I_2}\left[p_\theta + \frac{\mathrm{sen}\psi}{\mathrm{sen}\theta\cos\psi}(p_\phi - p_\psi\cos\theta)\right]^2 + \frac{1}{2I_3}p_\psi^2, \end{aligned} \tag{A.14}$$

where the Euler angles θ, ϕ, and ψ are in the ranges $\theta \in [0, \pi]$, $\phi \in [0, 2\pi]$, and $\psi \in [0, 2\pi]$. The next step is to rewrite the Eq. (A.15) in a Gaussian form for each canonical momenta p_θ, p_ϕ, and p_ψ. Therefore, after some mathematical manipulation, we finally have

$$\begin{aligned} \mathcal{H}_{\text{rot}} &= \Psi \left[p_\theta + \left(\frac{1}{I_2} - \frac{1}{I_1} \right) \frac{\text{sen}\psi \cos\psi}{2\Psi \text{sen}\theta} (p_\phi - p_\psi \cos\theta) \right]^2 \\ &+ \frac{1}{4 I_1 I_2 \Psi \text{sen}^2\theta} (p_\phi - p_\psi \cos\theta)^2 + \frac{1}{2I_3} p_\psi^2, \end{aligned} \tag{A.15}$$

with

$$\Psi = (\text{sen}^2\psi / 2I_1 + \cos^2\psi / 2I_2). \tag{A.16}$$

A.1.3 Classical Partition Function

Consider a fluid of N water-like rigid-bodies molecules, as described in the last section, in which each molecule interacts through the UF potential, given by:

$$U_{\text{UF}}(\mathbf{R}_{ij}) = -\frac{p}{\beta} \ln \left(1 - e^{-(\mathbf{R}_{ij}/\sigma)^2} \right) \tag{A.17}$$

where $\mathbf{R}_{ij}$ is a distance between the COM of two molecules, σ is a distance-scale and p is a positive parameter which is responsible to control the softness of interactions. As the forces are acting on the COM, we can assume that the single degrees of fredom of a rigid molecule are independent of each other. Therefore, the total Hamiltonian $\mathcal{H}$ of the system can be written as:

$$\mathcal{H} = \mathcal{H}_{\text{trans}}(\mathbf{R}, \mathbf{P}) + \mathcal{H}_{\text{rot}}(\varphi_i, p_{\varphi_i}), \tag{A.18}$$

with

$$\mathcal{H}_{\text{trans}} = \sum_{i=1}^{N} \frac{\mathbf{P}_i^2}{2M} - \sum_{i,j>i}^{N} \frac{p}{\beta} \ln \left(1 - e^{-(\mathbf{R}_{ij}/\sigma)^2} \right), \tag{A.19}$$

where, $M = m_O + 2m_H$ is the sum of particles mass, $\mathcal{H}_{\text{trans}}$ describes the translation and interactions on $\mathbf{R}$ of the molecule and $\mathcal{H}_{\text{rot}}$ is the rotational energy, Eq. (A.15), which depends on the Euler angles $\varphi_i = (\theta, \phi, \psi)$ and the corresponding angular momenta.

The classical partition function $Z_N(N, V, T)$ of this system is given by

$$Z_N(N, V, T) = \frac{1}{N!} Z_{\text{trans}} Z_{\text{rot}} = \frac{1}{N!} Z_{\text{trans}} \left[Z_{\text{rot}}^{(1)} \right]^N, \tag{A.20}$$

where we are using the fact that the rotational contribution factor just as in the case of noninteracting particles, it means that we can write $Z_{\text{rot}} = \left[Z_{\text{rot}}^{(1)} \right]^N$, with $Z_{\text{rot}}^{(1)}$ being the rotational partition function of a single molecule.

We start calculating the translational partition function Z_{trans}, which is

$$\begin{aligned}
Z_{\text{trans}}(N,V,T) &= \frac{1}{h^{3N}}\int d^3\mathbf{R}_1\cdots d^3\mathbf{R}_N \int d^3\mathbf{P}_1\cdots d^3\mathbf{P}_N \,\exp\left(-\beta\mathcal{H}_{\text{trans}}\right) \\
&= \frac{1}{h^{3N}}\int d^3\mathbf{P}_1\cdots d^3\mathbf{P}_N \,\exp\left(-\beta\sum_{i=1}^{N}\frac{\mathbf{P}_i^2}{2M}\right)\times \\
&\times \int d^3\mathbf{R}_1\cdots d^3\mathbf{R}_N \,\exp\left[\beta\sum_{i>j}^{N}\frac{p}{\beta}\ln\left(1-e^{-(\mathbf{R}_{ij}/\sigma)^2}\right)\right]. \qquad (A.21)
\end{aligned}$$

Transforming the exponential of the sum into a product of exponentials, we have

$$\begin{aligned}
Z_{\text{trans}}(N,V,T) &= \frac{1}{h^{3N}}\int d^3\mathbf{P}_1\cdots d^3\mathbf{P}_N \prod_{i=1}^{N}\left[\exp\left(-\beta\frac{\mathbf{P}_i^2}{2M}\right)\right]\times \\
&\times \int d^3\mathbf{R}_1\cdots d^3\mathbf{R}_N \underbrace{\left[\prod_{i,j>i}^{N}\left(1-e^{-(\mathbf{R}_{ij}/\sigma)^2}\right)\right]^p}_{[g(\mathbf{R})]^p} \\
&= \left(\frac{2\pi M}{\beta h^2}\right)^{3N/2}\times\left[V^N\underbrace{\left(\frac{\int d^3\mathbf{R}_1\cdots d^3\mathbf{R}_N\,[g(\mathbf{R})]^p}{\int d^3\mathbf{R}_1\cdots d^3\mathbf{R}_N}\right)}_{\langle[g(\mathbf{R})]^p\rangle}\right]. \qquad (A.22)
\end{aligned}$$

Therefore,

$$Z_{\text{trans}}(N,V,T) = \left(\frac{2\pi M}{\beta h^2}\right)^{3N/2} V^N\langle[g(\mathbf{R})]^p\rangle. \qquad (A.23)$$

On the other hand, the rotational partition function of a single molecule $Z_{\text{rot}}^{(1)}$ is given by:

$$Z_{\text{rot}}^{(1)} = \frac{1}{h^3}\int d\theta\, d\phi\, d\psi \int_{-\infty}^{+\infty} dp_\theta\, dp_\phi\, dp_\psi \,\exp\left[-\beta\mathcal{H}_{\text{rot}}\right]. \qquad (A.24)$$

First, we perform the integration over p_θ, which is a Gaussian function with its integral having a closed form

$$\int_{-\infty}^{+\infty} e^{-a(x+b)^2}\,dx = \int_{-\infty}^{+\infty} e^{-ax^2}\,dx = \sqrt{\frac{\pi}{a}}, \qquad (A.25)$$

Therefore, we obtain as a result for p_θ integration

$$\int_{-\infty}^{+\infty}\exp\left\{-\beta\Psi\left[p_\theta+\left(\frac{1}{I_2}-\frac{1}{I_1}\right)\frac{\text{sen}\psi\cos\psi}{2\Psi\text{sen}\theta}(p_\phi-p_\psi\cos\theta)\right]^2\right\}dp_\theta = \sqrt{\frac{\pi}{\beta}}\,\Psi^{-1/2}. \qquad (A.26)$$

The next integration is over p_ϕ. One more time using Eq. (A.25), we get

$$\int_{-\infty}^{+\infty}\exp\left\{\frac{-\beta}{4I_1I_2\Psi\text{sen}^2\theta}(p_\phi-p_\psi\cos\theta)^2\right\}dp_\phi = \sqrt{\frac{4\pi I_1I_2\text{sen}^2\theta}{\beta}}\,\Psi^{1/2}. \qquad (A.27)$$

The last canonical momenta integration is over p_ψ,

$$\int_{-\infty}^{+\infty} \exp\left\{\frac{-\beta}{2I_3}p_\psi^2\right\} dp_\psi = \sqrt{\frac{2\pi I_3}{\beta}}. \tag{A.28}$$

Returning to Eq. (A.24) with the previous results and performing the last integrals over the 3 Euler angles

$$Z_{\text{rot}}^{(1)} = \frac{1}{h^3}\sqrt{\frac{\pi}{\beta}}\sqrt{\frac{4\pi I_1 I_2}{\beta}}\sqrt{\frac{2\pi I_3}{\beta}}\underbrace{\left(\int_0^\pi \text{sen}\theta\, d\theta\right)}_{2}\underbrace{\left(\int_0^{2\pi} d\phi\right)}_{2\pi}\underbrace{\left(\int_0^{2\pi} \Psi^{-1/2}\Psi^{1/2}\, d\psi\right)}_{2\pi}. \tag{A.29}$$

Therefore

$$Z_{\text{rot}}^{(1)} = \frac{1}{\nu\pi}\sqrt{\frac{2\pi I_1}{\beta\hbar^2}}\sqrt{\frac{2\pi I_2}{\beta\hbar^2}}\sqrt{\frac{2\pi I_3}{\beta\hbar^2}}, \tag{A.30}$$

where ν is a molecule symmetry factor, which takes the value of 2 for H_2O molecule. The rotational partition function is quite an analogous form to that of the translational, however, we must replace $V^{1/3}\sqrt{2\pi M k_B T/h^2}$ by $\pi^{-1/3}\sqrt{2\pi I k_B T/\hbar^2}$.

A.1.4 Helmholtz free energy

The Helmholtz free energy becomes

$$\begin{aligned} F_{\text{Rw}}(N,V,T) &= -\frac{1}{\beta}\ln Z(N,V,T) \\ &= -\frac{1}{\beta}\left[\ln Z_{\text{trans}} + N\ln Z_{\text{rot}}^{(1)} - \ln N!\right], \end{aligned} \tag{A.31}$$

and, using Eqs. (A.23) and (A.30), is given by

$$\begin{aligned} F_{\text{UF/Rw}}(T) &= -\frac{N}{\beta}\left[\frac{3}{2}\ln\left(\frac{2\pi M}{\beta h^2}\right) - \ln(\rho) + 1\right. \\ &- \left.\frac{1}{2N}\ln(2\pi N) + \frac{1}{2}\ln\left(\frac{2\pi I_1 I_2 I_3}{\beta^3\hbar^6}\right)\right] \\ &+ F_{\text{UF}}^{(\text{exc})}, \end{aligned} \tag{A.32}$$

where

$$F_{\text{UF}}^{(\text{exc})} = -\frac{1}{\beta}\ln\langle[g(\mathbf{R})]^p\rangle, \tag{A.33}$$

is the excess Helmholtz free-energy of the UF.

A.2 Flexible-water/Uhlenbeck-Ford Model (UF/Fw)

A.2.1 Classical Partition Function

We consider a fluid of N water-like molecules in a volume V and in thermal equilibrium with a heat bath at a temperature T. The oxygen atom (atom 1) has mass m_{O} and is connected to the two protons (atoms 2 and 3) with masses m_{H} by springs with constant k_r and equilibrium length d. The energetics of the bond angle $\alpha \equiv \angle HOH$ is also harmonic, with spring constant k_α and equilibrium angle α_0. The Hamiltonian describing such a molecule is

$$\begin{aligned} H &= \sum_{i=1}^{N}\left[\frac{\mathbf{p}_{\mathrm{O},i}^2}{2m_{\mathrm{O}}} + \frac{\mathbf{p}_{\mathrm{H1},i}^2 + \mathbf{p}_{\mathrm{H2},i}^2}{2m_{\mathrm{H}}} + \frac{k_r}{2}\left[(|\mathbf{r}_{\mathrm{H1},i} - \mathbf{r}_{\mathrm{O},i}| - d)^2 + (|\mathbf{r}_{\mathrm{H2},i} - \mathbf{r}_{\mathrm{O},i}| - d)^2\right]\right. \\ &+ \left.\frac{k_\alpha}{2}(\alpha_i - \alpha_0)^2\right] - \sum_{i,j>i}^{N}\frac{p}{\beta}\ln\left\{1 - \exp\left[-\frac{1}{\sigma^2}\left(|\mathbf{r}_{\mathrm{O},j} - \mathbf{r}_{\mathrm{O},i}|^2\right)\right]\right\}, \end{aligned} \tag{A.34}$$

where $\mathbf{p}_{\mathrm{O},i}$ is the momentum of the oxygen atom and $\mathbf{p}_{\mathrm{H1},i}$ and $\mathbf{p}_{\mathrm{H2,i}}$ are the momenta of the two protons in a molecule i. The same notation applies for the position vectors $\mathbf{r}_{\mathrm{O},i}$ and $\mathbf{r}_{\mathrm{H1},i}$ and $\mathbf{r}_{\mathrm{H2},i}$.

The classical canonical partition function for this system is given by

$$Z(N,V,T) = \frac{1}{N!}\left(\frac{2\pi m_{\mathrm{O}}}{\beta h^2}\right)^{3N/2}\left(\frac{2\pi m_{\mathrm{H}}}{\beta h^2}\right)^{6N/2} Z_N, \tag{A.35}$$

where $\beta = 1/k_B T$, h Planck's constant and Z_N is configurational integral

$$\begin{aligned} Z_N(N,V,T) &= \int_V d^3\mathbf{r}_{\mathrm{O},1}\cdots d^3\mathbf{r}_{\mathrm{O},N}\int_V d^3\mathbf{r}_{\mathrm{H1},1}\cdots d^3\mathbf{r}_{\mathrm{H1},N}\int_V d^3\mathbf{r}_{\mathrm{H2},1}\cdots d^3\mathbf{r}_{\mathrm{H2},N} \times \\ &\times \prod_{i=1}^{N}\left[\exp\left(-\frac{k_r\beta}{2}\left[(|\mathbf{r}_{\mathrm{H1},i} - \mathbf{r}_{\mathrm{O},i}| - d)^2 + (|\mathbf{r}_{\mathrm{H2},i} - \mathbf{r}_{\mathrm{O},i}| - d)^2\right]\right) \times\right. \\ &\times \left.\exp\left(-\frac{k_\alpha\beta}{2}(\alpha_i - \alpha_0)^2\right) \times \prod_{j>i}^{N}\left\{1 - \exp\left[-\frac{1}{\sigma^2}|\mathbf{r}_{\mathrm{O},j} - \mathbf{r}_{\mathrm{O},i}|^2\right]\right\}^p\right] \end{aligned} \tag{A.36}$$

To evaluate the above integrals it is useful to define the relative coordinate vectors [102]

$$\mathbf{r}_i \equiv \mathbf{r}_{\mathrm{O},i} - \mathbf{r}_{\mathrm{H1},i} \tag{A.37}$$

$$\mathbf{r}_i' \equiv \mathbf{r}_{\mathrm{H2},i} - \mathbf{r}_{\mathrm{O},i}, \tag{A.38}$$

where the absolute position vectors of the 3 atoms are represented using a Cartesian laboratory frame (X,Y,Z). The two relative coordinate vectors, on the other hand, are represented by the Cartesian coordinate systems (x,y,z) and (x',y',z'), respectively, that are fixed with respect to the molecule. The first has its origin located at the position of

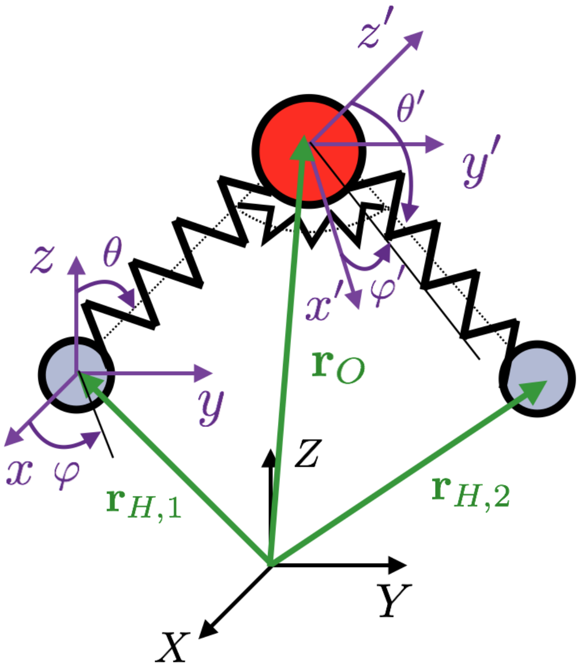

Figure 18 – Adopted coordinates to describe the flexible water molecule.

$\mathbf{r}_{\mathrm{H},1}$ and the directions of its coordinate axes coincide with those of the laboratory frame. The origin of the second coincides with $\mathbf{r}_{\mathrm{O}}$ and the direction of its z'-axis is such that it is along $\mathbf{r}$, as depicted in Fig. 18. The directions of the x' and y' axes can be chosen arbitrarily, provided it leads to an orthogonal right-handed axis system (x', y', z').

The convenience of these choices becomes apparent when we adopt spherical coordinates to describe both relative coordinate vectors. Specifically, the vectors $\mathbf{r}$ and $\mathbf{r}'$ are specified by the coordinates (r, θ, φ) and (r', θ', φ'), respectively, as shown in Fig. 18.

Writing $\tilde{\mathbf{r}}_i \equiv \mathbf{r}_{\mathrm{O},i}$, the integral in Eq. (A.36) can be transformed into

$$\begin{aligned} Z_N(N,V,T) \;=\; & \int_V d^3\tilde{\mathbf{r}}_1 \cdots d^3\tilde{\mathbf{r}}_N \int_0^L d^3\mathbf{r}_1 \cdots d^3\mathbf{r}_N \int_0^L d^3\mathbf{r}'_1 \cdots d^3\mathbf{r}'_N \times \\ & \times \prod_{i=1}^{N} \left[\exp\left(-\frac{k_r\beta}{2} \left[(r_i - d)^2 + (r'_i - d)^2 \right] \right) \times \exp\left(-\frac{k_\alpha\beta}{2} (\alpha_i - \alpha_0)^2 \right) \times \right. \\ & \left. \times \prod_{j>i}^{N} \left\{ 1 - \exp\left[-\frac{1}{\sigma^2} |\tilde{\mathbf{r}}_j - \tilde{\mathbf{r}}_i|^2 \right] \right\}^p \right] \end{aligned} \tag{A.39}$$

where L is the linear size of the containing volume V

Performing the integrations over the relative coordinates in spherical coordinates (assuming that V is so large that the shape of the boundary of volume V becomes irrelevant), the classical partition function can be written as

$$\begin{aligned}
Z_N(N,V,T) &= \int\limits_V d^3\tilde{\mathbf{r}}_1\cdots d^3\tilde{\mathbf{r}}_N \underbrace{\prod_{i,j>i}^{N}\left\{1-\exp\left[-\frac{1}{\sigma^2}|\tilde{\mathbf{r}}_j-\tilde{\mathbf{r}}_i|^2\right]\right\}^p}_{[g(\tilde{\mathbf{r}})]^p} \times \\
&\times \int_0^R dr_1 r_1^2\cdots dr_N r_N^2 \prod_{i=1}^{N}\exp\left(-\frac{k_r\beta}{2}(r_i-d)^2\right)\int_0^{2\pi} d\varphi_1\cdots d\varphi_N \times \\
&\times \int_0^{\pi} d\theta_1 \mathrm{sen}\theta_1\cdots d\theta_N \mathrm{sen}\theta_N \times \\
&\times \int_0^R dr'_1 r'^2_1\cdots dr'_N r'^2_N \prod_{i=1}^{N}\exp\left[-\frac{k_r\beta}{2}(r'_i-d)^2\right]\int_0^{2\pi} d\varphi'_1\cdots d\varphi'_N \times \\
&\times \int_0^{\pi} d\theta'_1 \mathrm{sen}\theta'_1\cdots d\theta'_N \mathrm{sen}\theta'_N \prod_{i=1}^{N}\exp\left[-\frac{k_\alpha\beta}{2}\left(\pi-\theta'_i-\alpha_0\right)^2\right]
\end{aligned} \tag{A.40}$$

where $R=(3V/4\pi)^{1/3}$ is the radius of a sphere of volume V. This reduces to

$$Z_N(N,V,T)=\left[8\pi^2 V\, I_r^2(R,\beta)\, I_\alpha(\beta)\right]^N \langle[g(\tilde{\mathbf{r}})]^p\rangle, \tag{A.41}$$

with

$$I_r(R,\beta)=\int_0^R dr\; r^2\exp\left(-\frac{k_r\beta}{2}(r-d)^2\right), \tag{A.42}$$

$$\begin{aligned}
I_\alpha(\beta) &= \int_0^{\pi} d\theta \mathrm{sen}\theta\, \exp\left(-\frac{k_\alpha\beta}{2}(\pi-\theta-\alpha_0)^2\right) \\
&= \int_0^{\pi} d\theta \mathrm{sen}\theta\, \exp\left(-\frac{k_\alpha\beta}{2}(\theta-\alpha_0)^2\right),
\end{aligned} \tag{A.43}$$

and

$$\langle[g(\tilde{\mathbf{r}})]^p\rangle = \frac{\int\limits_V d^3\tilde{\mathbf{r}}_1\cdots d^3\tilde{\mathbf{r}}_N[g(\tilde{\mathbf{r}})]^p}{\int\limits_V d^3\tilde{\mathbf{r}}_1\cdots d^3\tilde{\mathbf{r}}_N} \tag{A.44}$$

$$\tag{A.45}$$

The integrations in Eqs. (A.42) and (A.43) can be performed explicitly, giving

$$\begin{aligned}
I_r(R,\beta) &= \frac{\sqrt{\frac{\pi}{2}}\,(\beta\, d^2\, k_r+1)\left[\mathrm{erf}\left(d\sqrt{\frac{1}{2}\,\beta k_r}\right)-\mathrm{erf}\left((d-R)\sqrt{\frac{1}{2}\,\beta k_r}\right)\right]}{(\beta k_r)^{3/2}} \\
&+ \frac{d\exp\left(-\frac{1}{2}\beta\, d^2\, k_r\right)-(d+R)\exp\left(-\frac{1}{2}\beta(d-R)^2 k_r\right)}{\beta k_r},
\end{aligned} \tag{A.46}$$

which, in the limit $R\to\infty$, gives

$$\lim_{R\to\infty} I_r(R,\beta) = \frac{\sqrt{\frac{\pi}{2}}\,(\beta d^2 k_r+1)\left(\mathrm{erf}\left(\frac{d\sqrt{\beta k_r}}{\sqrt{2}}\right)+1\right)}{(\beta k_r)^{3/2}} + \frac{d\,\exp\left(-\frac{1}{2}\,\beta d^2 k_r\right)}{\beta k_r}, \tag{A.47}$$

and

$$
\begin{aligned}
I_\alpha(\beta) &= \sqrt{\frac{\pi}{8\,\beta\,k_\alpha}}\exp\left(-\frac{1}{2\beta k_\alpha}\right)\left[e^{i\alpha_0}\mathrm{erfi}\left(\frac{1-i\,\alpha_0\,\beta\,k_\alpha}{\sqrt{2\,\beta\,k_\alpha}}\right)+e^{-i\alpha_0}\mathrm{erfi}\left(\frac{1+i\,\alpha_0\,\beta\,k_\alpha}{\sqrt{2\,\beta\,k_\alpha}}\right)\right.\\
&- \left. e^{i\alpha_0}\mathrm{erfi}\left(\frac{1+i\,\beta\,k_\alpha(\pi-\alpha_0)}{\sqrt{2\,\beta\,k_\alpha}}\right)-e^{-i\alpha_0}\mathrm{erfi}\left(\frac{1-i\,\beta\,k_\alpha(\pi-\alpha_0)}{\sqrt{2\,\beta\,k_\alpha}}\right)\right]\\
&= \sqrt{\frac{\pi}{2\,\beta\,k_\alpha}}\exp\left(-\frac{1}{2\beta k_\alpha}\right)\times \qquad (A.48)\\
&\quad \mathrm{Re}\left[e^{i\alpha_0}\mathrm{erfi}\left(\frac{1-i\,\alpha_0\,\beta\,k_\alpha}{\sqrt{2\,\beta\,k_\alpha}}\right)-e^{i\alpha_0}\mathrm{erfi}\left(\frac{1+i\,\beta\,k_\alpha(\pi-\alpha_0)}{\sqrt{2\,\beta\,k_\alpha}}\right)\right]. \qquad (A.49)
\end{aligned}
$$

Here, $\mathrm{erfi}(z)$ is the imaginary error function defined as [84]

$$\mathrm{erfi}(z) \equiv -i\,\mathrm{erf}(i\,z), \tag{A.50}$$

$\mathrm{Re}[\cdots]$ indicates the real part of the expression in brackets, and we have used the following properties of the error function,

$$\mathrm{erf}(z^*) = (\mathrm{erf}(z))^*, \tag{A.51}$$

with the asterisk indicating complex conjugation.

A.2.2 Helmholtz free energy

The Helmholtz free energy of N molecules is defined as

$$F_{\mathrm{Fw}}(N,V,T) = -\frac{1}{\beta}\ln Z(N,V,T), \tag{A.52}$$

and, using Eqs. (A.35), (A.39), (A.41), is given by

$$
\begin{aligned}
F_{\mathrm{Fw}}(T) &= -\frac{N}{\beta}\left[\frac{3}{2}\ln\left(\frac{8\pi^3 m_{\mathrm{O}} m_{\mathrm{H}}^2}{\beta^3 h^6}\right)-\ln\left(\frac{\rho}{8\pi^2}\right)+1\right.\\
&- \left.\frac{1}{2N}\ln(2\pi N)+2\ln(I_r)+\ln(I_\alpha)\right]\\
&+ F_{\mathrm{UF}}^{(\mathrm{exc})}, \qquad (A.53)
\end{aligned}
$$