UiO : Department of Physics
University of Oslo

Evaluating the Phase Behaviour of Silica Modeled by the Vashishta Potential Using Free Energy Methods

Nicholas Karlsen

Master's Thesis, Autumn 2022

Abstract

In this thesis we study the phase behaviour of crystalline polymorphs of SiO_2 modeled by the Vashishta three-body interaction potential in classical molecular dynamics simulations. First the crystal free energies are computed using the Einstein crystal method, of which three variations was tested: (i) thermodynamic integration, (ii) Non-equilibrium thermodynamic integration and (iii) Free energy perturbation followed by thermodynamic integration. We found that out of these three methods, only the former two was able to accurately calculate the free energy of Vashishta SiO_2. Subsequently an efficient and automated computational workflow was developed and we searched for coexistence points the thermodynamically stable phases of SiO_2 in the (P, T) plane for three separate parameterizations of the Vashishta potential. We located the α-quartz/coesite coexistence points in all three parameterizations along the isotherm $T = 300\,\mathrm{K}$ at pressures $4.5, 48.6, 112.6\,\mathrm{kbar}$. Additionally we located the α-quartz/β-tridymite coexistence point in two of the parameterizations along the isobar $1\,\mathrm{bar}$ at temperatures $543.1, 818.9\,\mathrm{K}$.

We then use the Gibbs-Duhem integration method to draw out the remainder of the coexistence lines for the two parameterizations in which we obtained two coexistence points. Finding the qualitative reproduction of the α-quartz/β-tridymite line in both cases, barring the apparent phase-transition from α-quartz to β-quartz along the line. The α-quartz/coesite line is however not qualitatively reproduced in either case, as its gradient in $P(T)$ is opposite to that of the experimental phase diagram.

Acknowledgements

First and foremost, i would like to thank my main advisor, Henrik Andersen Sveinsson. Your helpful advice has been invaluable in the work on this thesis, as has your positive attitude. Further, i would like to thank my second advisor, Anders Malthe-Sørenssen, as well as the rest of the CCSE-MD group. It has been enjoyable to hear about the interesting things you have all been working on in our bi-weekly meetings! I would also like to give my thanks to Hilbert space and its inhabitants for providing an excellent environment in which to write a thesis.

Lastly, i would like to give thanks to Nora, for your patience and support while i have been working on this thesis.

Contents

Abstract i

Acknowledgements ii

Contents iii

1 Introduction 1
- 1.1 Goals . . . 2
- 1.2 My contributions . . . 3
- 1.3 Outline of this thesis . . . 3

I Theory & Background 5

2 Statistical Mechanics & Thermodynamics 6
- 2.1 The Canonical Ensemble . . . 6
- 2.2 Isothermal–isobaric ensemble . . . 7
- 2.3 The Ideal Einstein Crystal . . . 8
- 2.4 Phase Coexistence . . . 8

3 Molecular Dynamics 10
- 3.1 Verlet Integration . . . 10
- 3.2 Interaction Potentials . . . 11
- 3.3 Thermostats . . . 13

4 Free Energy Methods 15
- 4.1 Free Energy Perturbation . . . 15
- 4.2 Thermodynamic Integration . . . 17
- 4.3 Non-Equilibrium Thermodynamic Integration . . . 18
- 4.4 The Einstein Crystal Method . . . 19

II Implementation, Results & Discussion 22

5 Computing the Free Energy of Vashishta SiO_2 in LAMMPS 23
- 5.1 Computational Pipeline . . . 23
- 5.2 Crystal Initialization . . . 24

5.3 Thermodynamic Integration 26
5.4 Non-Equilibrium Thermodynamic Integration 34
5.5 Testing thermodynamic integration and non-equilibrium thermodynamic integration 37
5.6 Free Energy Perturbation Followed By Thermodynamic Integration 38
5.7 Summary 42

6 Finding points of coexistence between Silica polymorphs 43
6.1 Coexistence points along an isobar 44
6.2 Coexistence points along an isotherm 52

7 Computing the phase diagram of Silica 56
7.1 Gibbs-Duhem Integration in LAMMPS 56
7.2 Phase diagram of Vashishta Silica 59

III Conclusion 62

8 Summary and Outlook 63
8.1 Summary 63

Appendices 66

A Mass independence of coexistence points 67

B Vashishta parameters 68

C LAMMPS Scripts 70
C.1 Crystal Initialization 70
C.2 Thermodynamic Integration 72
C.3 Non-Equilibrium Thermodynamic Integration 73
C.4 Free Energy Perturbation (Vega Path) 74
C.5 Thermodynamic Integration (Vega Path) 75

Bibliography 77

CHAPTER 1

Introduction

Silicon dioxide, or Silica, is a naturally occurring chemical of great geological [1] as well as industrial [2] importance. Quartz, a crystal structure of silica, is of particular importance which by virtue of being stable in standard conditions is estimated to make up 12% of the earths crust by volume [3]. In addition to quartz, crystalline silica exists in many different polymorphic structures such as tridymite, cristobalite, coesite and stishovite. Each of which may be further subdivided into groups of closely related structures [4, 5] separated by displacive phase transitions [1].

The mechanisms driving geological processes is an inherently multi-scale problem, and thus understanding of the nanoscale behaviour is a crucial component in explaining and understanding the driving mechanisms [6]. Ideally we would study these processes at the nanoscale by so-called ab initio, quantum mechanical, methods where the electronic structure is determined by solving the many-body Schrodinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{R},t) = \hat{H}\Psi(\boldsymbol{R},t). \tag{1.1}$$

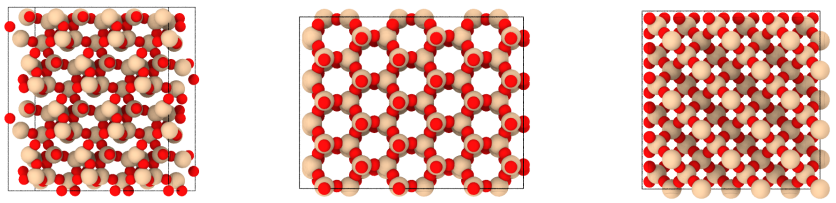
Unfortunately even approximate methods in this regime, like density functional theory, is limited both in spatial and temporal resolution, even on large super computers [7, 8]. Classical molecular dynamics (MD) offers a good approximation, even down to the nanoscale, whilst being much more computationally efficient allowing the study of systems resolved in much greater detail, by several orders of magnitude both temporally and spatially. There exists a wide variety of classical interaction potentials aimed at modeling different phenomena, and developing an interaction potential that accurately depicts the behaviour of Silica has been a long-standing issue [1, 9, 10]. A few notable examples include the BKS potential [11], Tsuneyuki potential [12], and ReaxFF [13].

The three-body interaction potential by Vashishta et al. [14] is an appealing candidate in the study of Silica, and has been shown to accurately reproduce many of the properties of its crystal structures [14, 15]. A particularly appealing feature of the Vashishta potential is that its bonded interactions are implicit, rather than explicitly defined in the molecular geometry as if often the case for interaction potentials, the latter of which is more computationally demanding. This makes the potential particularly suited for the study of friction and fracture [16, 17], during which bonds are continuously broken and formed.

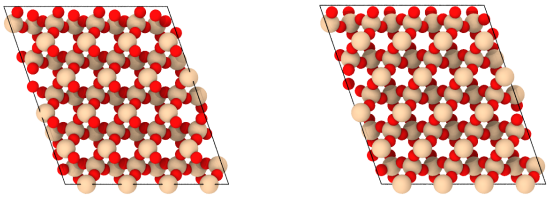
When subject to external load, silica may undergo phase transitions. For example, the transition between α and β-quartz, which may be a driving

factor in the formation of cracks in the continental crust [18]. For this reason, knowledge of the phase diagram is an important component in studying these phenomenon.

The Vashishta potential has been fit to reproduce many of the physical properties of silica [14, 15, 19]. Yet, this gives no indication that its phase behaviour should be accurately reproduced [20], and as of now, no comprehensive study of the Vashishta silica phase diagram has been done. We will therefore tackle three different parameterizations of Vashishta silica and evaluate their phase behaviour using free energy methods. In particular, we will utilize the Einstein crystal method by Frenkel and Ladd [21] to compute the free energy of Vashishta silica at different temperatures and pressures in order to locate coexistence points between the different polymorphs. Following this, we draw out the coexistence lines using the Gibbs-Duhem integration method by Kofke [22], producing the phase diagram.



(a) Coesite (b) α-Tridymite (c) β-Cristobalite



(d) α-Quartz (e) β-Quartz

Figure 1.1: Various Silica polymorphs

1.1 Goals

The overarching goal in this thesis is to gain an increased understanding of the phase behaviour of crystalline polymorphs of silica modelled by the Vashishta potential. To that end, we have set the following goals:

1. Implement the Einstein crystal method in the LAMMPS molecular dynamics software to compute free energies of crystalline systems.
2. Develop a computational workflow that enables efficient preparation and handling of a large number of molecular dynamics simulations, as well as the subsequent analysis.

3. Locate points of coexistence in (P, T) between crystalline polymorphs of SiO_2.

4. Compute the phase diagram of Vashishta SiO_2 using Gibbs-Duhem integration.

1.2 My contributions

- A Python package which enables SiO_2 free energy calculations as well as the subsequent analysis en masse with the Vashishta potential in LAMMPS.
- A Python package which performs Gibbs-Duhem Integration with LAMMPS.
- An increased understanding of how to properly impose center of mass constraints in the Einstein crystal methods, and its effects on the computed free energies.
- A fork of LAMMPS which introduces a "scale" parameter in the implementation of the Vashishta potential in the MANYBODY package which enables the user to scale the strength of the potential with the adapt command, a crucial component of thermodynamic integration computations in the Einstein crystal method.
- Knowledge of the silica phase diagram for the Vashishta et al. [14], Broughton et al. [15] and Wang et al. [19] parameterisations of the Vashishta potential.

1.3 Outline of this thesis

This thesis is structured in three main parts. In the first part, theory and background, we will give a brief reminder of some of the most central elements of statistical mechanics and thermodynamics that are of particular relevance to the work presented in this thesis. We then move on to introduce molecular dynamics in some detail, assuming little existing familiarity with the method. Lastly, we give an introduction to the central free energy methods that make up the so-called Einstein crystal method, as well as the Einstein crystal method itself.

We then move on to the second, and most central part of the thesis. Here we start by giving detailed descriptions in both the implementation and other practical details concerning the Einstein crystal method. In particular, we explain in detail how the method may be implemented in LAMMPS, with emphasis on the considerations that had to be made when applying the method to Vashishta SiO_2. The method is implemented in 3 variations, which are all tested and compared against each other. We also outline an efficient computational workflow to perform the Einstein crystal en masse, with statistically robust estimates of the standard error. Following this, we move on to the next chapter in which the Einstein crystal method is used to compute the free energy of several SiO_2 polymorphs in (P, T) to determine

their region of stability, and to locate points of coexistence between phases, this done concurrently for the 1990, [14] 1997 [15] and Wang [19] parameterizations of the Vashishta potential. We then move on to discuss the implementation of the Gibbs-Duhem integration method in LAMMPS, and use the method to draw out the remainder of the coexistence lines in the 1997 [15] and Wang [19] parameterizations of the Vashishta potential, producing their phase diagrams.

Finally, we conclude with a summary where we reflect on what has been done, as well discussing some future prospects on how this work may be extended further.

PART I

Theory & Background

CHAPTER 2

Statistical Mechanics & Thermodynamics

In this chapter we provide a brief reminder of some elements in statistical mechanics and thermodynamics that are of particular relevance to the work presented in this thesis. This summary is not intended to provide a complete background, and a general understanding of statistical mechanics and thermodynamics is assumed throughout this thesis. For a more complete overview of the subject we refer the reader to the textbook by Swendsen [23] or a similar texts.

2.1 The Canonical Ensemble

The canonical ensemble describes a system in which the number of particles (N), volume (V) and temperature (T) are conserved quantities. The ensemble may also interchangeably be referred to as the NVT ensemble. Physically, the ensemble can be thought to describe a fixed-size box with volume V containing N particles submerged in a much larger heat bath at temperature T, with which the box is at thermal equilibrium. In statistical mechanics each micro-state of the system is assigned a probability density

$$\mathcal{P}(\boldsymbol{p}^N, \boldsymbol{r}^N) = \frac{1}{Z} e^{-\beta \mathcal{H}(\boldsymbol{p}^N, \boldsymbol{r}^N)} \tag{2.1}$$

with Hamiltonian $\mathcal{H}$, and the notation

$$\begin{aligned} \boldsymbol{p}^N &= \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_N\} \\ \boldsymbol{r}^N &= \{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N\} \end{aligned} \tag{2.2}$$

which will be used frequently throughout the text. We further have the normalization factor

$$Z(N, V, T) = \int \frac{d\boldsymbol{p}^N d\boldsymbol{r}^N}{h^{3N}} \, e^{-\beta \mathcal{H}(\boldsymbol{p}^N, \boldsymbol{r}^N)}, \tag{2.3}$$

the canonical partition function, comprised of a $6N$ dimensional integral over all possible configurations of $(\boldsymbol{p}^N, \boldsymbol{r}^N)$ spanning the accessible region of phase-space for the system. This is visualized in Fig. 2.1, in which each point denotes one configuration. Whilst Z essentially acts as a normalization, it can be shown

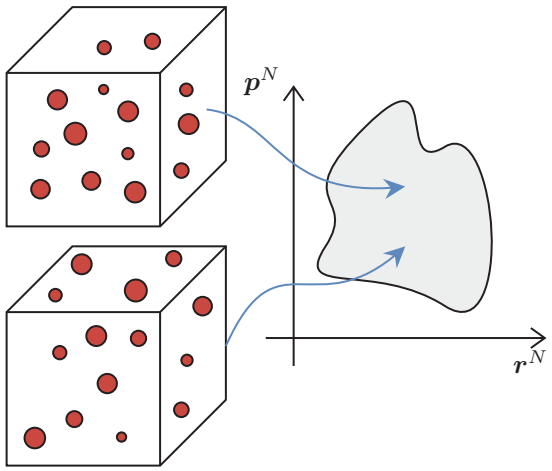


Figure 2.1: Two distinct configurations of a system mapping to different points in the accessible region of phase space.

that is is in fact related to Helmholtz free energy by the relation

$$F(N,V,T) = -\frac{1}{\beta}\ln Z(N,V,T), \tag{2.4}$$

the thermodynamic equation of state in the canonical ensemble. Hence, may in principle compute the free energy by solving the integral in Eqn. 2.3. Unfortunately this is only possible to do analytically for a few select systems.

2.2 Isothermal–isobaric ensemble

Closely related to the canonical ensemble we have the so-called isothermal–isobaric ensemble that describes a system in which N, P and T are conserved quantities, hence, we may interchangeably refer to it as the NPT ensemble. Physically, the NPT ensemble can be thought to represent the conditions of an open container in a lab. That is, one that is under pressure as well as tempered by the atmosphere around it. The probability density function in the NPT ensemble is

$$\mathcal{P}(\boldsymbol{p}^N, \boldsymbol{r}^N, V) = \frac{1}{Z(N,P,T)} e^{-\beta\mathcal{H}(\boldsymbol{p}^N,\boldsymbol{r}^N)+PV} \tag{2.5}$$

with the corresponding partition function

$$Z(N,P,T) = \int dV\, Z(N,V,T) e^{-\beta PV}. \tag{2.6}$$

In a similar fashion to the canonical ensemble, $Z(N,P,T)$ may be related to its corresponding thermodynamic equation of state, the Gibbs free energy, by

$$G(N,P,T) = -\frac{1}{\beta}\ln Z(N,P,T). \tag{2.7}$$

2.3 The Ideal Einstein Crystal

The Einstein crystal[1] is a system composed of N independently oscillating harmonic oscillators which may be described by the Hamiltonian

$$\mathcal{H} = \sum_{i}^{N} \left[\frac{\boldsymbol{p}_i^2}{2m_i} + \frac{1}{2} k_i (\Delta \boldsymbol{r}_i)^2 \right] \tag{2.8}$$

where $\Delta \boldsymbol{r}_i = \boldsymbol{r}_i - \boldsymbol{r}_{i,0}$ denotes the atoms displacement from its rest position and k_i the spring constant of the i-th oscillator. The canonical partition function for this system may be written as

$$Z_E(N,V,T) = \int \frac{d\boldsymbol{p}^N d\boldsymbol{r}^N}{h^{3N}} \, e^{-\beta \mathcal{H}_E} \tag{2.9}$$

where h denotes Plancks constant. This integral is solved analytically by considering the momenta (P_E) and configurational (Q_E) contributions individually, both of which are standard gaussian integrals with solutions

$$P_E(N,V,T) = \int d\boldsymbol{p}^N \prod_{i=1}^{N} \exp\left[-\beta \frac{\boldsymbol{p}_i^2}{2m_i} \right] = \prod_{i=1}^{N} \left(\frac{2\pi m_i}{\beta} \right)^{3/2} \tag{2.10}$$

and

$$Q_E(N,V,T) = \int d\boldsymbol{r}^N \prod_{i=1}^{N} \exp\left[-\beta \frac{1}{2} k_i (r_i - r_{i,0})^2 \right] = \prod_{i=1}^{N} \left(\frac{2\pi}{\beta k_i} \right)^{3/2} \tag{2.11}$$

which combined yields the full partition function

$$Z_E(N,V,T) = \frac{P_E Q_E}{h^{3N}}. \tag{2.12}$$

With this, one may compute an analytic expression for the free energy

$$F_E = \sum_{i=1}^{N} \ln \left(\frac{\beta k_i \Lambda_i^2}{2\pi} \right)^{3/2} \tag{2.13}$$

where $\Lambda_i = \left(\beta h^2 / 2\pi m_i \right)^{1/2}$ is the thermal de Broglie wavelength.

2.4 Phase Coexistence

In general, the thermodynamically stable phase of a system corresponding to a given macroscopic state, (P,T), is the one which minimizes the free energy. This is the phase preferred by nature, and given time a system will tend towards this phase. However, there may also exist additional, so called meta-stable phases, that whilst not minimizing the global free energy, a corresponds to a local minima in the free energy landscape.

For some (P,T), there may be two phases that have the same free energy. These are said to be in phase coexistence with one another, and the macroscopic

[1] Alternatively referred to as an Einstein solid

state at which this occurs, the so-called coexistence point. A coexistence point, is part of a coexistence line, along which the two phases always has the same free energy, and the gradient of this line is given by the so-called Clauseus-Clapeyron equation

$$\frac{\mathrm{d}P}{\mathrm{d}T} = \frac{\Delta h}{T\Delta v} \tag{2.14}$$

where Δh is the difference in enthalphy per particle between the two phases, and Δv the difference in volume per particle.

CHAPTER 3

Molecular Dynamics

Molecular Dynamics (MD) is a computational method in which the dynamics of atomistic and molecular systems are simulated in the classical regime. The simulations are performed by solving Newtons equations of motion

$$\boldsymbol{f}_i = -\nabla_i U(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \tag{3.1}$$

numerically, where $\boldsymbol{f}_i$ denotes the net force acting on the $i-th$ atom and U the interaction potential. The method allows one to observe the dynamics of atoms at full resolution and to make direct measurements based on their trajectories. By solving these equations directly, one naturally samples the microcanonical ensemble where the number of particles (N), volume (V) and total energy (E) are conserved. One may also sample from different ensembles by the use of so-called thermostats and barostats which constrains the temperature and pressure of the system respectively.

MD commonly employs periodic boundary conditions, wherein particles are allowed to interact across the boundaries and pair-wise interactions will always prefer the minimal distance between pairs of atoms. This is done to minimize surface effects due to a finite system size and to better mimic the bulk properties of a system.

3.1 Verlet Integration

When integrating the equations of motion in molecular dynamics, it is of particular importance to use an integration scheme that preserves the energy of the system over time. For example, if one was to perform MD simulations using the Forward Euler method to integrate the equations of motion, one would expect the energy to drift over time. This is especially problematic if one wishes to sample from the NVE ensemble, in which the energy is an explicitly conserved quantity.

A popular, and commonly used integrator in MD simulations is the Velocity Verlet method, which integrates the equations of motion according to

$$\begin{aligned} \boldsymbol{r}(t+\Delta t) &= \boldsymbol{r}(t) + \boldsymbol{v}(t)\Delta t + \frac{1}{2}\boldsymbol{a}(t)\Delta t^2 \\ \boldsymbol{v}(t+\Delta t) &= \boldsymbol{v}(t) + \frac{1}{2}(\boldsymbol{a}(t) + \boldsymbol{a}(t+\Delta t))\Delta t. \end{aligned} \tag{3.2}$$

Velocity Verlet is shown to be good in preserving the energy of MD systems over time, even with relatively long timesteps [24]. For this reason, it is a is a standard method in MD simulations.

3.2 Interaction Potentials

3.2.1 The Leonard Jones Potential

Perhaps the most widely studied interaction potential is that of Jones and Chapman [25], the so-called Leonard-Jones potential. The potential is simple in its form and models the pairwise interaction between atoms according to

$$U(r_{ij}) = 4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right]. \tag{3.3}$$

where $r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$ and ε, σ are free parameters to tune the depth of the potential well and zero of the potential respectively, see Fig. 3.1. The $1/r^6$ term in Leonard-Jones aims to model the Van der Waals interaction, and the $1/r^{12}$ term steric repulsion. It should be noted that the 12 in the exponent is chosen for efficiency reasons as a direct multiple of 6, whilst steric repulsion should reality go as an exponential, as is the case in the closely related Buckingham potential [26].

Note in Fig. 3.1 that $U \to 0$ as $r \to 3\sigma$, meaning that particles separated by distances greater that 3σ will hardly interact at all. Summing over all the pairwise interactions in the system goes like $\mathcal{O}(N^3)$, hence in situations like these it is common to set an explicit cut-off distance. In combination with so-called neighbour lists, that tracks the particles that are within the cut-off distance of each other, one may significantly reduce the number of pairs for which one has to evaluate the pairwise interaction at the cost of some memory.

For mono-atomic substances, Leonard-Jones is capable of modeling gas, liquid and solid states, being stable in both FCC and HCP crystal lattices. For systems composed of different atomic types there exists several so-called mixing rules, for example the Goode-Hope rules

$$\begin{aligned} \varepsilon_{\mathrm{A,B}} &= \sqrt{\varepsilon_{\mathrm{A}}\varepsilon_{\mathrm{B}}} \\ \sigma_{\mathrm{A,B}} &= \sqrt{\sigma_{\mathrm{A}}\sigma_{\mathrm{B}}} \end{aligned} \tag{3.4}$$

which yields the parameters for the interaction between atom types A and B. Chemical bonds may also be added as either holonomic constrains enforced efficiently in with the SHAKE [27] algorithm or related methods, or simply as an harmonic interaction between pairs of atoms. In either case, the bonds need be explicitly defined and kept track of during simulation at some computational cost.

3.2.2 The Vashishta Potential

We then have the potential developed by Vashishta et al. [14], a three-body potential with the functional form

$$U = \sum_i \sum_{j>i} U^{(2)}_{ij}(r_{ij}) + \sum_i \sum_{j\neq i} \sum_{\substack{k>j \\ k\neq i}} U^{(3)}_{ijk}(r_{ij}, r_{ik}, \theta_{ijk}). \tag{3.5}$$

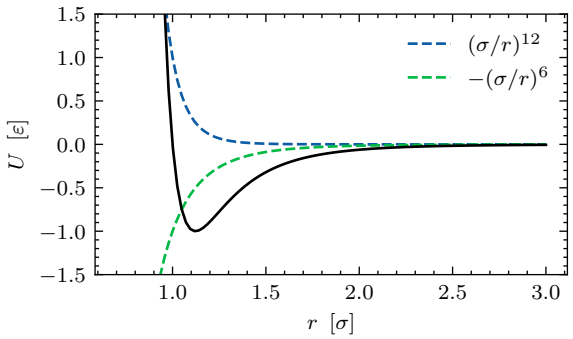


Figure 3.1: Functional form of the Leonard-Jones potential

The two body term of the potential which models pair-wise interactions between atoms has the form

$$U_{ij}^{(2)}(r) = \frac{H_{ij}}{r^{\eta_{ij}}} + \frac{Z_i Z_j}{r} e^{-r/r1s} - \frac{D_{ij}}{r^4} e^{-r/r4sj} - \frac{W_{ij}}{r^6}. \tag{3.6}$$

where the term $H_{ij}/r^{\eta_{ij}}$ models steric repulsion. The second term, $Z_i Z_j/r$, models the coulomb interaction with a screening factor $\exp(-r/r_{1s})$. The third term, D_{ij}/r^4, models charge-dipole interactions, again with a screening factor. Lastly, we have the term W_{ij}/r^6, which models the Van der Waals interaction. Then we also have the three-body interaction

$$U_{ijk}^{(3)}(r_{ij}, r_{ik}, \theta_{ijk}) = B_{ijk} \frac{(\cos\theta_{ijk} - \cos\theta_0)^2}{1 + C_{ijk}(\cos\theta_{ijk} - \cos\theta_0)^2} e^{\frac{\xi}{r_{ij} - r_0}} e^{\frac{\xi}{r_{ik} - r_0}}, \tag{3.7}$$

where the B_{ijk} denotes the strength of the interaction, cosines model the effects of bond bending and the exponential bond stretching and r_0 the cut-off. To illustrate, the geometry of a three-body system is illustrated in Fig. 3.2. Note in particular that the bonds in the Vashishta potential are not explicitly defined, but will dynamically break and form based on the cut-off distance r_0. In a computational perspective, this means this makes the potential highly efficient as you forego much of the much of the computational work associated with maintaining the bond topology of the system.

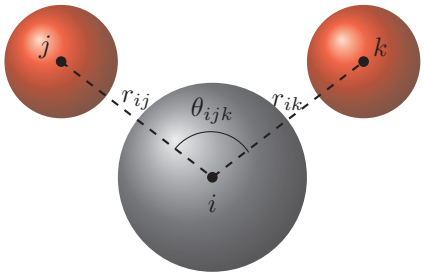


Figure 3.2: Sketch depicting the bond angles and distances between atoms labeled i, j, k in a three-body interaction

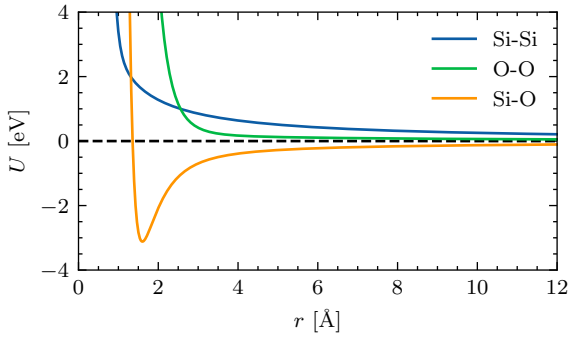


Figure 3.3: The Two-body interactions, $U^{(2)}$, in the 1990 parameterization of the Vashishta potential

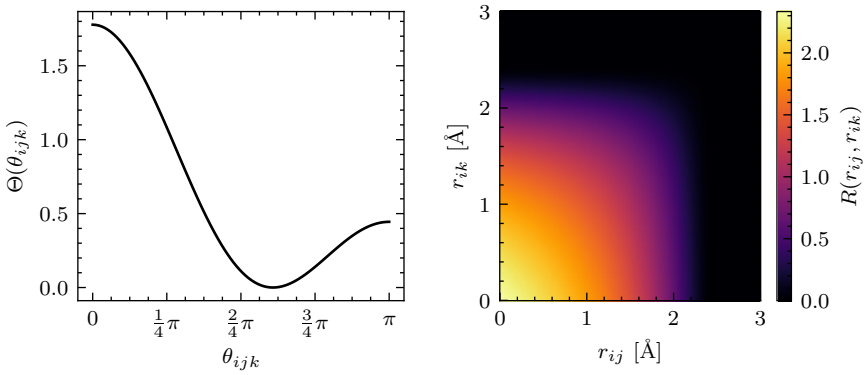


Figure 3.4: The Three-body interactions in the 1990 parameterization of the Vashishta potential for Si-O-O decomposed to its radial and angular contributions to $U^{(3)}(r_{ij}, r_{ik}, \theta_{ijk}) = B_{ijk} R(r_{ij}, r_{ik}) \Theta(\theta_{ijk})$

3.3 Thermostats

The thermostat by Berendsen et al. [28] was one of the earliest, and perhaps simplest examples of a thermostat in MD. The Berendsen thermostat works by imposing that the temperature of the system should evolve according to the differential equation

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{T_0 - T(t)}{\tau} \tag{3.8}$$

where T_0 is the temperature of the heat bath, $T(t)$ the temperature of the system at time t and τ a parameter indicating the strength of the coupling. This is them explicitly enforced by re-scaling the velocities of each atom in the system after each time-step according to

$$v_i(t) \to \lambda v_i(t) \tag{3.9}$$

where the scaling parameter λ is given by

$$\lambda = \sqrt{1 + \frac{\Delta t}{\tau}\left(\frac{T_0}{T} - 1\right)}. \tag{3.10}$$

Unfortunately, the dynamics of a system that under the Berendsen thermostat does not sample from the canonical ensemble. Furthermore, the thermostat actively dampens internal degrees of freedom in systems favouring instead translational degrees of freedom, known as the "Flying ice-cube" effect [29]. These problems are remedied in the closely related thermostat by Bussi, Donadio and Parrinello [30], which in essence may be thought of us Berendsen with added random noise.

In addition to these, there exists a wide range of different thermostats fit for different scenarios. However, the one which we will use in this thesis is the so-called Langevin thermostat.

The Langevin thermostat takes an entirely different approach, where instead of re-scaling the velocities, the equations of motion of each atom is changed by introducing a new set of forces. These forces are aimed to model the presence of a solvent that exchanges heat with the system. This is done by introducing adding the two forces

$$\boldsymbol{f}_i = -\nabla_i U(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) + \boldsymbol{f}_f + \boldsymbol{f}_r \tag{3.11}$$

where the term, $\boldsymbol{f}_f$, models frictional drag between the atoms and the solvent and $\boldsymbol{f}_r$ the force due to the atoms bumping into the solvent.

CHAPTER 4

Free Energy Methods

The Free energy of a system is linked to the partition function via the relation

$$F(N,V,T) = -\frac{1}{\beta}\ln Z(N,V,T) \tag{4.1}$$

and thus you could in principle determine the free energy by computing the partition function. Recall however that Z is composed of an integral over the $6N$ dimensional phase-space. Hence, it is only analytically solvable for a select few simple systems. Numerical integration is also unfeasible given the curse of dimensionality, whereby the required computational time scales far beyond what is practical even on the largest of supercomputers. The usual solution to this problem is to employ monte-carlo methods, or alternatively for molecular systems, molecular dynamics. In the case of computing free energies, this is however not sufficient. The probability that we sample a configuration in both MC and MD is

$$\mathcal{P}(\boldsymbol{p}^N, \boldsymbol{r}^N) \propto e^{-\beta\mathcal{H}(\boldsymbol{p}^N,\boldsymbol{r}^N)} \tag{4.2}$$

with $\beta = 1/k_BT$. Hence the probability of sampling configurations with energies greater than the thermal energy k_BT is exponentially unlikely. This is the fundamental problem that free energy methods aims to solve, and there exists a plethora of methods aimed study the free energy of a wide range of different systems Chipot and Pohorille [31] and Lelievre, Stoltz and Rousset [32].

In this chapter we start by giving a brief overview of the Free energy methods that have been employed in this thesis from a general perspective, where Chipot and Pohorille [31] has been used extensively as a reference in addition to the textbooks Frenkel and Smit [33], Allen and Tildesley [24]. Following this, we move on to look at the Einstein crystal method introduced by Frenkel and Ladd [21], as a method for computing the Free energy of solids for which Frenkel and Smit [33] and Vega et al. [34] provide the main reference.

4.1 Free Energy Perturbation

Consider a Hamiltonian in the form

$$\mathcal{H}_1(\boldsymbol{p}^N, \boldsymbol{q}^N) = \mathcal{H}_0(\boldsymbol{p}^N, \boldsymbol{r}^N) + \Delta\mathcal{H}(\boldsymbol{p}^N, \boldsymbol{r}^N) \tag{4.3}$$

where $\Delta\mathcal{H}$ may interpreted as a perturbation to $\mathcal{H}_0$. The Helmholtz free energy is related to the canonical partition function by

$$F(N,V,T) = -\frac{1}{\beta}\ln Z(N,V,T) \tag{4.4}$$

thus the difference in free energies between the perturbed and unperturbed systems may be written as

$$\Delta F = F_1 - F_0 = -\frac{1}{\beta}\ln\frac{Z_1}{Z_0} = -\frac{1}{\beta}\frac{\int d\boldsymbol{p}^N d\boldsymbol{r}^N\, e^{-\beta\mathcal{H}_1(\boldsymbol{p}^N,\boldsymbol{r}^N)}}{\int d\boldsymbol{p}^N d\boldsymbol{r}^N\, e^{-\beta\mathcal{H}_0(\boldsymbol{p}^N,\boldsymbol{r}^N)}} \tag{4.5}$$

where the integrals are to be understood as $6N$ dimensional integrals over phase-space, with N denoting the number of particles in the system. Substituting in the RHS of Eqn. 4.3 to the integral in the numerator we get

$$\Delta F = -\frac{1}{\beta}\ln\left[\frac{\int d\boldsymbol{p}^N d\boldsymbol{r}^N\, e^{-\beta\Delta\mathcal{H}(\boldsymbol{p}^N,\boldsymbol{r}^N)}e^{-\beta\mathcal{H}_0(\boldsymbol{p}^N,\boldsymbol{r}^N)}}{\int d\boldsymbol{p}^N d\boldsymbol{r}^N\, e^{-\beta\mathcal{H}_0(\boldsymbol{p}^N,\boldsymbol{r}^N)}}\right] \tag{4.6}$$

and because the Boltzmann factors $e^{-\beta\mathcal{H}_0(\boldsymbol{p}^N,\boldsymbol{r}^N)}$ constitutes a probability density function, the integral in the above expression can be interpreted as taking the ensemble average of the observable $e^{-\beta\Delta\mathcal{H}(\boldsymbol{p}^N,\boldsymbol{r}^N)}$ measured in the unperturbed system, which dynamics is governed solely by $\mathcal{H}_0$. We may thus compute the difference in free energies between the perturbed and unperturbed systems as

$$\Delta F = -\frac{1}{\beta}\ln\left\langle e^{-\beta\Delta\mathcal{H}(\boldsymbol{p}^N,\boldsymbol{r}^N)}\right\rangle_0 \tag{4.7}$$

where $\langle\ldots\rangle_0$ indicates an ensemble average in the $\mathcal{H}_0$-system. That is, $\Delta\mathcal{H}$ is measured in a system governed by the dynamics of $\mathcal{H}_0$. Eqn. 4.7 defines the technique of free energy perturbation [35], a powerful and broadly applicable technique that allows the computation of changes in free energies. If both systems are governed by conservative force fields and the mass of the constituent particles are kept constant, the momenta contribution to the partition functions in Eqn. 4.6 will be equal and thus cancel out, reducing Eqn. 4.7 to just

$$\Delta F = -\frac{1}{\beta}\ln\left\langle e^{-\beta\Delta U(\boldsymbol{r}^N)}\right\rangle_0 \tag{4.8}$$

furthermore, by absorbing the factor -1 in Eqn. 4.6 into the logarithm we may flip the fraction reformulating the expression to

$$\Delta F = \frac{1}{\beta}\ln\frac{Z_0}{Z_1} = \frac{1}{\beta}\ln\left[\frac{\int d\boldsymbol{p}^N d\boldsymbol{r}^N\, e^{-\beta\mathcal{H}_0(\boldsymbol{p}^N,\boldsymbol{r}^N)}}{\int d\boldsymbol{p}^N d\boldsymbol{r}^N\, e^{-\beta\mathcal{H}_1(\boldsymbol{p}^N,\boldsymbol{r}^N)}}\right] \tag{4.9}$$

then by adding and subtracting a factor $\beta\Delta\mathcal{H}$ to the exponent of the partition function in the numerator we get

$$\Delta F = \frac{1}{\beta}\ln\frac{Z_0}{Z_1} = \frac{1}{\beta}\ln\left[\frac{\int d\boldsymbol{p}^N d\boldsymbol{r}^N\, e^{\beta\Delta\mathcal{H}(\boldsymbol{p}^N,\boldsymbol{r}^N)}e^{-\beta\mathcal{H}_1(\boldsymbol{p}^N,\boldsymbol{r}^N)}}{\int d\boldsymbol{p}^N d\boldsymbol{r}^N\, e^{-\beta\mathcal{H}_1(\boldsymbol{p}^N,\boldsymbol{r}^N)}}\right] \tag{4.10}$$

which leaves us with

$$\Delta F = \frac{1}{\beta}\ln\left\langle e^{\beta\Delta\mathcal{H}(\boldsymbol{p}^N,\boldsymbol{r}^N)}\right\rangle_1 \tag{4.11}$$

as an alternative and equivalent way of computing the change of free energy when going from system 0 to 1. If we again follow the same assumptions as before we may reduce this expression to

$$\Delta F = \frac{1}{\beta} \ln \left\langle e^{\beta \Delta U(\boldsymbol{r}^N)} \right\rangle_1 . \tag{4.12}$$

4.2 Thermodynamic Integration

Consider a Hamiltonian in the form

$$\mathcal{H}(\boldsymbol{p}^N, \boldsymbol{r}^N, \lambda) = \sum_i^N \frac{\boldsymbol{p}^2}{2m_i} + \lambda U(\boldsymbol{r}^N) \tag{4.13}$$

where λ is a tuning parameter, usually chosen to be in the interval $[0, 1]$, which meaning will be made clear later in this section. In to simplify the notation $(\boldsymbol{p}^N, \boldsymbol{r}^N)$ will be dropped from the Hamiltonian and related quantities when in the context of thermodynamic integration in this and subsequent sections.

In order to derive the expression used for thermodynamic integration we may start by taking the derivative of Helmholtz free energy with respect to the hyper-parameter λ as

$$\frac{\partial F(\lambda)}{\partial \lambda} = -\frac{1}{\beta} \frac{\partial}{\partial \lambda} \ln Z(\lambda) = -\frac{1}{\beta Z(\lambda)} \int d\boldsymbol{p}^N d\boldsymbol{r}^N \, \frac{\partial}{\partial \lambda} \Big(e^{-\beta U(\lambda)} \Big). \tag{4.14}$$

The partial derivative inside of the integral pulls out a factor $-\beta U(\lambda)$ from the exponential and leaves us with

$$\frac{1}{Z(\lambda)} \int d\boldsymbol{p}^N d\boldsymbol{r}^N \, \frac{\partial U(\lambda)}{\partial \lambda} e^{-\beta U(\lambda)} = \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda \tag{4.15}$$

where the λ suffix indicates that the average is taken for a given λ. We are then left with the differential equation

$$\frac{\partial F(\lambda)}{\partial \lambda} = \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda \tag{4.16}$$

which we integrate on both sides with respect to λ over its domain. The LHS then gives us

$$\int_0^1 d\lambda \frac{\partial F(\lambda)}{\partial \lambda} = F(\lambda = 1) - F(\lambda = 0) = \Delta F \tag{4.17}$$

which leaves us with the final expression

$$\Delta F = \int_0^1 d\lambda \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda \tag{4.18}$$

as a tool for finding the Free energy difference between the states corresponding to $\lambda = 1$ and $\lambda = 0$. This technique is often used in conjunction with a potential in the from

$$U(\lambda) = \lambda U_B + (1 - \lambda) U_A \tag{4.19}$$

a so-called linear homotopy. This potential continuously transforms your system from A to B, and Eqn.4.18 will yield the change in free energy provided there occurs no phase transitions during the transformation.

4.3 Non-Equilibrium Thermodynamic Integration

We know from thermodynamics that the work done, W, during some process is related to the free energy by the relation

$$\Delta F \leq W. \tag{4.20}$$

The equality is in this case only obtained for quasi-static processes, which is indeed the case during the thermodynamic integration method discussed in the previous section in which the system must be allowed to reach equilibrium between each step. We will in this section introduce a non-equilibrium variant of thermodynamic integration that allows us to perform the computation in a more efficient manner [36], as we forego the need for equilibration between each step.

In order to compute the change in free energy during a non-equilibrium process we start by considering the reversible work done during a quasi-static process from state A to B. By virtue of the process being reversible, it follows that

$$\Delta F = W^{\text{rev}}_{A\to B} = \frac{1}{2}[W^{\text{rev}}_{A\to B} - W^{\text{rev}}_{B\to A}] \tag{4.21}$$

where $W^{\text{rev}}_{A\to B}$ denotes the reversible work from A to B. In several realizations of a finite time, non-equilibrium process, we may decompose the work into two parts: The amount of reversible work and the dissipated energy [36]. This leads to the relationship

$$W^{\text{rev}}_{A\to B} = \left\langle W^{\text{irr}}_{A\to B} \right\rangle - \left\langle E^{\text{diss}}_{A\to B} \right\rangle. \tag{4.22}$$

In addition, it may be shown that for if the non-equilibrium process is sufficiently close to the quasi-static one [36], we have that

$$\left\langle E^{\text{diss}}_{A\to B} \right\rangle = \left\langle E^{\text{diss}}_{B\to A} \right\rangle. \tag{4.23}$$

By inserting Eqn. 4.22 and Eqn. 4.23 into Eqn. 4.24 we get

$$\Delta F = \frac{1}{2}\left[\left\langle W^{\text{irr}}_{A\to B} \right\rangle + \left\langle W^{\text{irr}}_{B\to A} \right\rangle\right] \tag{4.24}$$

as an expression for the change in free energy as a function of the average amount of irreversible work done during a forward and backward non-equilibrium process.

As in equilibrium thermodynamic integration, the transformation between systems A and B is done along a linear homotopy Eqn. 4.19, where in this case λ is varied in a time dependent switching function that goes from 0 to 1 as $t \to t_s$. We may for example employ a switching function in the simple form,

$$\lambda(t) = t. \tag{4.25}$$

It has however been shown [37] that a switching function in the form

$$\lambda(t) = t^5\left(70t^4 - 315t^3 + 540t^2 - 420t + 126\right) \tag{4.26}$$

is particularly suitable for non-equilibrium thermodynamic integration. Both switching functions are shown in Fig. 4.1 for illustration.

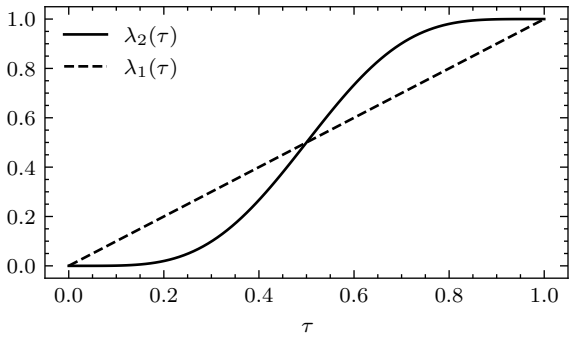


Figure 4.1: The switching functions defined in Eqn. 4.25 and Eqn. 4.26

The irreversible work done is then computed by the integral

$$W^{\mathrm{irr}} = \int_0^{t_s} dt \, \frac{\partial U(\lambda_t)}{\partial \lambda_t} \frac{\partial \lambda_t}{\partial t}, \tag{4.27}$$

which related to Eqn. 4.18, but with a change in variables from λ to t. Eqn. 4.27 is thus repeated several times in both directions, that is $A \to B$ and $B \to A$, the work done is averaged and inserted in Eqn. 4.24, producing the change in free energy, equivalent to that which is computed in equilibrium thermodynamic integration.

4.4 The Einstein Crystal Method

The Einstein Crystal method proposed by Frenkel and Ladd [21] introduced a method for computing the free energy of crystalline solids by defining a thermodynamic pathway between an ideal Einstein crystal,

$$U_E(\boldsymbol{r}^N) = \sum_i^N \frac{1}{2} k_i (\boldsymbol{r}_i - \boldsymbol{r}_{i,0})^2, \tag{4.28}$$

and solid governed by an interaction potential $U_C(\boldsymbol{r}^N)$. The idea is that the free energy of the solid may be decomposed into two parts

$$F_C(N, V, T) = F_E + \Delta F \tag{4.29}$$

where F_E denotes the Helmholtz free energy of the Einstein crystal, determined analytically, and the difference in free energy between the two systems, ΔF, which may be computed by the use of free energy methods. In their original paper Frenkel and Ladd [21] suggested a pathway in which the harmonic potential governing the ideal Einstein crystal is gradually turned off whilst the interaction potential of the target system is simultaneously turned on. In the formalism of thermodynamic integration this is performed by defining a potential

$$U(\lambda) = \lambda U_C(\boldsymbol{r}^N) + (1 - \lambda) U_E \tag{4.30}$$

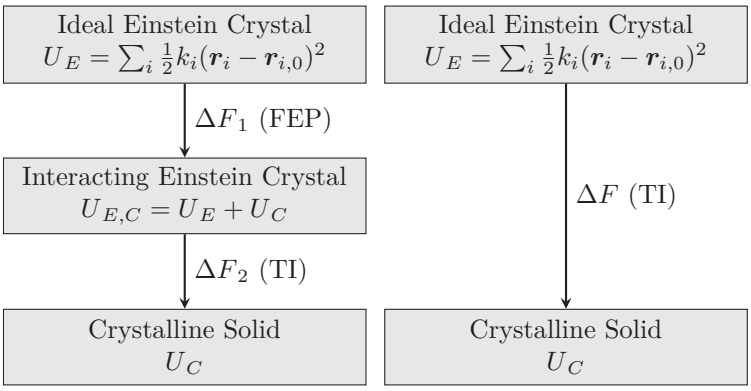
and computing the free energy difference either by thermodynamic integration, or non-equilibrium thermodynamic integration. However, there is no strict

requirement that this is the thermodynamic pathway that must be taken, and different pathways between the two systems may be used provided they are free of any phase-transitions. One such alternate pathway is the one described in Vega et al. [34]. Here, the authors instead compute the free energy difference between the ideal Einstein crystal and a system that is governed by both the harmonic potential and the interaction potential with free energy perturbation. Following this, the harmonic potential is gradually removed during thermodynamic integration along the linear homotopy

$$U(\lambda) = (1-\lambda)U_E + U_C \tag{4.31}$$

and the free energy of the crystal is given by

$$F_C(N,V,T) = F_E + \left\langle e^{-\beta p e_C} \right\rangle_E + \int_0^1 d\lambda \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda . \tag{4.32}$$



(a) The Vega et. al Path (b) The Frenkel and Ladd Path

Figure 4.2: Qualitative overview of the difference in the thermodynamic pathways between the ideal Einstein crystal and a solid (i.e Vashishta SiO_2) taken in the Vega et. al as well as the Frenkel and Ladd path.

The two pathways are summarized in Fig. 4.2, which highlights the difference between the two pathways.

4.4.1 Constraining the center of mass & finite-size corrections

In order to prevent a sharp peak in the integrand during thermodynamic integration as $\lambda \to 1$ and the harmonic potential goes to 0 it is recommended to fix the center of mass during Einstein crystal computations. This introduces two extra terms to the free energies,

$$F_C = F_E + \left(F_E^{CM} - F_E\right) + \Delta F^C M + \left(F_C - F_C^{CM}\right), \tag{4.33}$$

associated with introducing and removing the center-of-mass constraint, where F_E^{CM}, F_C^{CM} denotes the free energies of the Einstein crystal and solid with a fixed center mass. The additional term introduced to the free energy for a

single, for systems with a single spring constant k, was originally derived by Polson et al. [38]

$$\delta F = -\frac{1}{\beta} \ln \left[\left(\frac{\beta k}{2\pi \sum_{i=1}^{N} \mu_i^2} \right)^{3/2} \left(\frac{V}{N} \right) \right] \tag{4.34}$$

the reduced mass $\mu_i = m_i / \sum_i m_i$. More recently, Khanna et al. [39] generalized this result to the case in which the spring constants are no longer equal,

$$\delta F = -\frac{1}{\beta} \ln \left[\left(\frac{\beta k}{2\pi \sum_{i=1}^{N} \frac{\mu_i^2}{k_i}} \right)^{3/2} \left(\frac{V}{N} \right) \right]. \tag{4.35}$$

In addition to the finite-size correction contained in Eqn. 4.35, it was shown by Vega and Noya [40] that the Einstein crystal method has intrinsic finite-size effects, and proposed the finite-size scaling relation

$$f(N) = f(\infty) + \frac{1}{N} d \tag{4.36}$$

where $f(\infty)$ is the free energy per particle in the thermodynamic limit and d is the (arbitrary) gradient of the line.

PART II

Implementation, Results & Discussion

CHAPTER 5

Computing the Free Energy of Vashishta SiO_2 in LAMMPS

In this chapter we will look into how several variants of the Einstein crystal method may be implemented in the LAMMPS molecular dynamics package. Whilst the method, as well as its implementation in LAMMPS has been discussed in the literature on multiple occasions [34, 36, 39, 41, 42], yet many important details in the implementation is scattered making the barrier of entry to effectively utilizing the method somewhat high.

Given the end-goal of computing phase-diagrams, particular emphasis is put on efficient and automated processing and analysis of a large number of simulations. A crucial component in finding coexistence points, lest we spend significant time analyzing potentially thousands of simulations by hand.

We begin by giving a brief overview of the computational pipeline that has been developed, followed by detailed descriptions of the implementation and analysis of each step. LAMMPS scripts are provided in appendices C.1, C.2, C.3, and the python package that implements the workflow described in this chapter has been made available publicly on github[1].

Finally we study and compare the different implemented variants of the Einstein crystal method method by computing the Helmholtz free energy of α-quartz.

5.1 Computational Pipeline

A major challenge in this thesis has been the development of an efficient and automated computational pipeline that enables us to perform a large number of free energy calculations across isotherms or isobars in (P, T), as to locate points of coexistence between two phases. Furthermore, given the intrinsic finite-size effects of the Einstein crystal method [34], we should for each (P, T) perform simulations for several N employing the finite-size scaling relation by Vega and Noya [40] and extrapolating to an infinitely large crystal. Beyond just submitting and organizing a large number of molecular-dynamics simulations, we also wish to ensure that our insight is sound and that our results are robust. Hence, a complete and automated statistical analysis is performed for each free energy calculation providing robust estimates of the standard error.

[1] https://github.com/nicholaskarlsen/eincrystal

The first step in the pipeline is the so-called *initialization* step. This step takes an initial structure sourced from crystallographic data and runs molecular-dynamics to reach equilibrium in the NPT ensemble. Both the shape of the system and the positions of each individual atom is averaged in time, and a reference state for use in the subsequent NVT simulations is generated.

Following this, we have the second step of the pipeline where we compute the free energy using one of the several variations of the Einstein crystal method. Namely; direct equilibrium or non-equilibrium thermodynamic integration, or by free energy perturbation followed by thermodynamic integration. The three aforementioned methods are in principle equivalent, and should be interchangeable. However, they may differ in their applicability for different systems, as is found to be the case for Vashishta silica.

Note also that these steps are independent across different N, P, T and crystal structures, and may therefore be run parallel if one has access to a compute cluster, which is *strongly* advised if one wishes to apply the methods in this thesis to produce a phase diagram.

To summarize, we refer to the Fig. 5.1 which gives a graphical depiction of the pipeline.

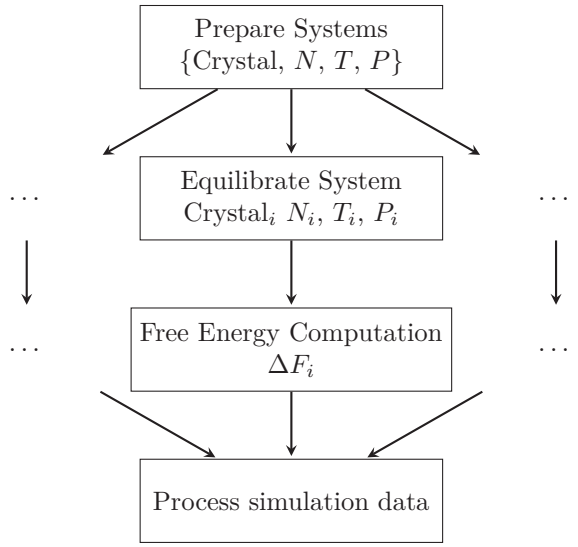


Figure 5.1: Outline of the general pipeline employed when computing free energies for a large number of different systems in parallel.

5.2 Crystal Initialization

The starting point for the free energy calculations is a set atomic positions, sourced from experimental crystallographic data or computational methods. These may be found in articles, or more readily in an online database of which there are several. Once a set of atomic positions have been obtained, the crystal must be equilibrated to the equilibrium NPT state according to the interaction potential, as this will general not correspond exactly to the crystal from sourced data. Either due to a difference in the interaction, the conditions at which the

data was captured, or both. Given that the quality this equilibration process sets the stage for all further calculations, it is rather important that it is done correctly.

5.2.1 Equilibrating the simulation box

The first step in the equilibration procedure is to obtain the correct shape of the simulation box, which in short is done by sampling the box shape in an NPT simulation and re-scaling to the average shape. A LAMMPS script that does this for Vashishta SiO_2 is provided in appendix C.1.

The process begins by initializing the velocities of each atom by sampling from the normal distribution $\mathcal{N}\left(0, \sqrt{\beta m}\right)$ individually for each component v_x, v_y, v_z, leading to a Maxwell-Boltzmann distribution of the particles speeds. To avoid roto-translation of the lattice, the net translational and rotational velocities are subsequently subtracted. Before starting the NPT simulation it is a good idea to run the system for a few steps in NVT. This is because the initial geometry may have an excess of potential energy when evaluated in the MD potential, which in an NPT simulation may lead to a rapid initial expansion of the crystal, bringing it far away from equilibrium and in the worst case cause a change in the crystal structure. A simple strategy for dealing with this is to run the system in NVT, where the box shape is fixed, for long enough to allow the potential energy to dissipate. An alternative approach utilized by Menon et al. [43] is to run NPT with an initial temperature $T/4$, gradually increasing to $T/2$ prior to the regular NPT simulation. Some limited testing of this approach yielded comparable results to the former, however I suspect this method may also suffer problems for high temperature, or particularly energetic initial configurations. Hence the former approach of running in NVT was preferred. To thermostat the system we utilized the Langevin thermostat, but in principle any thermostat that properly samples the canonical ensemble may be used for equilibration. In order to prevent the crystal from drifting, it is important to ensure that the thermostat does not act on the center of mass. For the Langevin thermostat, this is done by setting the net force acting on the system to zero. We also ensure that the temperature used in the thermostat excludes the center of mass. In LAMMPS, this is accomplished in the following way

```
compute Tcm all temp/com
fix 2 all langevin ${temp} ${temp} $(1000*dt) ${SEED} zero yes
fix_modify 2 temp Tcm
```

At this point we turn on the barostat. The barostat should be explicitly set to expand/contract centered on the center of mass, as this is not the case by default in LAMMPS. Furthermore, anisotropic pressure should be imposed by the barostat to ensure that the solid is under hydrostatic pressure and prevent the build up of stress in the crystals [34]. This is done by calling

```
compute COM all com
variable xcm equal c_COM[1]
variable ycm equal c_COM[2]
variable zcm equal c_COM[3]
fix 1 all nph aniso ${press} ${press} $(1000*dt) fixedpoint ${xcm} ${ycm} ${zcm}
fix_modify 1 temp Tcm
```

where we note the use of the `nph` fix style, which is used because we thermostat the system using the `langevin` fix. It is worth to note that this may also be accomplished in a similar way with the `npt` fix, which combines the Nose-Hoover thermostat [44, 45] with the Parinello-Rahman barostat [46] in a method by Shinoda, Shiga and Mikami [47].

The system should then be ran in NPT until equilibrium. Following this additional steps should be performed in the NPT-ensemble whilst sampling the shape of the simulation box, quantified by the length of each side. Once a sufficient number of samples have been obtained, the simulation box along with the atomic positions should be re-scaled to the average box shape.

5.2.2 Equilibrating the atomic positions

After the simulation box has been re-scaled, the system will then be in some random configuration which most likely differs a bit from an ideal crystal. It is not crucial that we obtain a fully ideal crystal as a reference state in Einstein crystal computations [34], however it is beneficial to get as close as possible. In Vega et al. [34] it is suggested to perform an energy minimization with gradient descent to relax the system into an ideal crystal. However, for many of the systems considered in this thesis this is problematic. The reason for this, which was observed when testing the methodology, is that gradient descent essentially relaxes the system into a state corresponding to a temperature of 0K, which may significantly change the structure of systems stable at higher temperatures. An example of this is β-quartz, which was observed to change into a non-periodic system after such an energy minimization. Hence we instead opt to run an additional NVT simulation whilst sampling the positions of each atom in the system individually and setting them to their average position. in LAMMPS, the position of each atom may be averaged by calling

```
compute xu all property/atom xu
fix avex all ave/atom 1 ${nsteps_pos_ave} ${nsteps_pos_ave} c_xu
```

and similarly for the y and z dimensions. One the sampling has ended, the positions of each atom may then be set to their average by

```
variable        avex atom f_avex
set             atom * x v_avex y v_avey z v_avez
```

This works for crystals under the assumption that all the atoms in the system approximately oscillate periodically about a point, and is free of any defects. Hence this methodology will only work for atomic solids and alterations must be made for molecular-solids, yet the main idea is likely still applicable.

5.3 Thermodynamic Integration

In this section, we will look into how to perform thermodynamic integration along the Frenkel & Ladd path [21] in LAMMPS. As a brief reminder, this involves computing the integral

$$\Delta F = \int_0^1 d\lambda \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda \tag{5.1}$$

with the the linear homotopy

$$U(\lambda) = \lambda U_C + (1 - \lambda)U_E \tag{5.2}$$

where U_C denotes the interaction potential, in this case Vashishta, and U_E a harmonic spring potential. The derivative of this potential is simply

$$\frac{\partial U(\lambda)}{\partial \lambda} = U_C - U_E. \tag{5.3}$$

Following Khanna et al. [39] we opt to compute Eqn. 5.1 numerically using a Gauss-Legendre quadrature method as

$$\Delta F \approx \sum_{i=0}^{n} w_i \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{\lambda_i} \tag{5.4}$$

where w_i are the weights of the n quadrature points on the interval $[0, 1]$. The integrand is evaluated by performing n separate MD simulations driven by the $U(\lambda)$ potential. A LAMMPS script performing these simulations is listed in appendix C.2. The primary components of which is the `fix spring` command, which applies harmonic springs to each atom tethering them to their initial positions, and the `fix adapt` command. This command allows us to scale the magnitude of U_C, but may unfortunately be somewhat difficult to utilize, especially for many-body potentials like Vashishta. This is discussed in depth in section 5.3.4.

5.3.1 Determining the Spring Constants of the Harmonic Potential

There exists several approaches [39] for setting the spring constants of the harmonic potential in the Einstein crystal reference state, but here we have opted to use mean-squared displacement (MSD) based springs [33] such that the MSD of the Einstein crystal reference state is equal to that of the Vashishta crystal. This is done by measuring the MSD, $\langle(\Delta \boldsymbol{r})^2\rangle$, of the Vashishta crystal in an NVT simulation, and setting the spring constants in the Einstein crystal to

$$k_i = \frac{3k_B T}{\langle(\Delta r_i)^2\rangle}. \tag{5.5}$$

We also employ the results from Khanna et al. [39] which generalizes the finite-size corrections of Polson et al. [38] to the case of multiple, different spring constants. We therefore measure the MSD of Si and O atoms individually, and assign them each different spring constants. For practical reasons this is done during the crystal initialization step, and thus the LAMMPS measurements are found towards the end in the script listed in appendix C.1.

5.3.2 Constraining the Center of Mass

In performing these simulations it is also important to keep the center of mass constrained in order to supress a weak divergence of the integrand Eqn. 5.1 at $\lambda = 0$ [34]. In the literature, it is suggested that this may be accomplished by ensuring that the thermostat excludes the center of mass and that the net force is zeroed [36, 39] which in LAMMPS translates to

```
fix fID group-ID langevin Tstart Tstop damp seed zero yes
compute cID all temp/com
fix_modify fID temp cID
```

and utilizing the recenter command[2] at each step [36, 39, 48].

```
fix fID all recenter INIT INIT INIT
```

However, it was found during the work on this thesis that these two constraints alone are not sufficient as they allow for drift in the momentum of the center of mass. For this reason, one should additionally impose an explicit constraint to the momentum of the center of mass by calling

```
fix fID all momentum 1 linear 1 1 1 rescale
```

This is studied in detail in section 5.4.1, when considering non-equilibrium thermodynamic integration. Yet we expect that the findings should also be applicable to thermodynamic integration.

5.3.3 Choice of Thermostat

In Einstein crystal computations, or indeed any computations driven by a harmonic potential, use of the Nose-Hoover thermostat [44, 45] is ill advised. Hoover [45] made a remark that it is not entirely clear if it is the canonical or microcanonical ensemble that is being sampled for harmonic systems being driven particularly small or large spring constants, k. Given that k is made progressively smaller during thermodynamic integration, we would thus enter into this questionable territory and should instead opt to use a different thermostat. Aragones, Valeriani and Vega [42] tested the use of the Langevin [49] and the Bussi-Donadio-Parrinello [30] thermostats in Einstein crystal calculations and found them to yield equivalent results. While both thermostats are available in LAMMPS, only the Langevin thermostat is available in a standard distribution of LAMMPS with the Bussi-Donadio-Parrinello thermostat being part of the EXTRA-FIX extension. For this reason, the Langevin thermostat was preferred.

5.3.4 Scaling the Vashishta potential in LAMMPS

In order to perform thermodynamic integration along the Frenkel and Ladd [21] pathway one needs to scale the strength of the interaction potentials of both the target and reference systems when performing thermodynamic integration. As a reminder, the potential used in this process is in the form

$$U(\lambda) = (1 - \lambda)U_E + \lambda U_C \tag{5.6}$$

where $\lambda \in [0, 1]$. This may be done using the `fix adapt` command[3] in LAMMPS which lets the user scale the value of certain parameters throughout a simulation. However, in order for an interaction potential to be scaled with this fix, an

[2] As a practical note, the `recenter` fix should be applied after all other fixes in the LAMMPS script. Else, the transform may be applied prior to the integration of the equations of motion rather than after.

[3] See https://docs.lammps.org/fix_adapt.html

`extract` method needs to be implemented for that potential in the .cpp file that defines it, usually named in the form `pair_<potential>.cpp`. To understand this, let us consider the LAMMPS implementation of the Leonard-Jones potential

$$U_{LJ}(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right], \tag{5.7}$$

contained in `pair_lj_cut.cpp` in the LAMMPS source code[4]. The following is an excerpt from this file, where the extract method is defined.

```
1 void *PairLJCut::extract(const char *str, int &dim)
2 {
3   dim = 2;
4   if (strcmp(str, "epsilon") == 0) return (void *) epsilon;
5   if (strcmp(str, "sigma") == 0) return (void *) sigma;
6   return nullptr;
7 }
```

Since scaling U_{LJ} by a factor λ is equivalent to scaling the parameter ε by λ, which enables us to scale the Leonard-Jones potential with the following command in LAMMPS

```
1 fix f_ID all adapt 0 pair lj/cut epsilon * * v_lambda
```

An `extract` method is relatively simple to implement for most potentials, and a list of the potentials for which it is already implemented is found in the LAMMPS documentation[5]. Unfortunately this is not the case for many-body potentials. Potentials in the MANYBODY package differ from regular two-body potentials in how their parameters are stored, and at the present time no extract method has been implemented for the Vashishta potential in LAMMPS.

In order to scale the Vashishta potential we have thus taken two separate approaches which are tested against each ofther for self-consistency. In the first approach we introduced a `scale` parameter in the `PairVashishta` class in a fork [6] of the LAMMPS source code. The scale parameter acts on the force acting on each atom, scaling it linearly. The scale parameter is also passed to the extract method, making it accessible by the fix adapt command in LAMMPS such that we may scale the Vashishta potential in the same way as described for Leonard-Jones, replacing `epsilon` with `scale`.

The second approach was to circumvent LAMMPS altogether and instead scaling the Vashishta potentials parameters directly in the parameter file. When we scale the Vashishta potential by a constant factor λ, the potential takes the form

$$\begin{aligned}\lambda \cdot \mathcal{U}(\boldsymbol{r}^N) = &\sum_i \sum_{j>i} \lambda \frac{H_{ij}}{r^{\eta_{ij}}} + \lambda \frac{Z_i Z_j}{r} e^{-r/r_{1s}} - \lambda \frac{D_{ij}}{r^4} e^{-r/r_{4s}} - \lambda \frac{W_{ij}}{r^6} \\ &+ \sum_i \sum_{j\neq i} \sum_{\substack{k>j \\ k\neq i}} \lambda \frac{B_{ijk}(\cos\theta_{ijk} - \cos\theta_0)^2}{1 + C_{ijk}(\cos\theta_{ijk} - \cos\theta_0)^2} e^{\frac{\xi}{r_{ij} - r_{0,ij}}} e^{\frac{\xi}{r_{ik} - r_{0,ik}}}\end{aligned} \tag{5.8}$$

[4] https://github.com/lammps/lammps/blob/develop/src/pair_lj_cut.h

[5] https://docs.lammps.org/fix_adapt.html

[6] https://github.com/nicholaskarlsen/lammps-scale-vashishta/blob/master/src/MANYBODY/pair_vashishta.cpp

Thus, we may scale the potential by making the following modifications to the parameters

$$\begin{aligned} H_{ij} &\to H'_{ij} = \lambda H_{ij} \\ D_{ij} &\to D'_{ij} = \lambda D_{ij} \\ Z_i &\to Z'_i = \sqrt{\lambda} Z_i \\ Z_j &\to Z'_j = \sqrt{\lambda} Z_j \\ W_{ij} &\to W'_{ij} = \lambda W_{ij} \\ B_{ijk} &\to B'_{ijk} = \lambda B_{ijk} \end{aligned} \tag{5.9}$$

to do this in an effective manner, we extend the genpot[7] python package adding a "scale" functionality which implements this transformation and writes the resulting parameters out in a parameter file that is then read by LAMMPS.

If the scaling has been implemented correctly, we should at least expect the two methods to be self consistent. A simple quantitative test of the implementations is then to measure the mean-squared displacement for various scale-factors in $[0, 1]$ in both methods. We expect the mean-squared displacement to grow to an increasingly larger plateau with a decreasing scalefactor, eventually changing to a linear function when we get diffusion according to Fick's law as the interactions due to the Vashishta potential vanish and the particles behave solely according to the Langevin thermostat resulting in what is essentially Brownian motion.

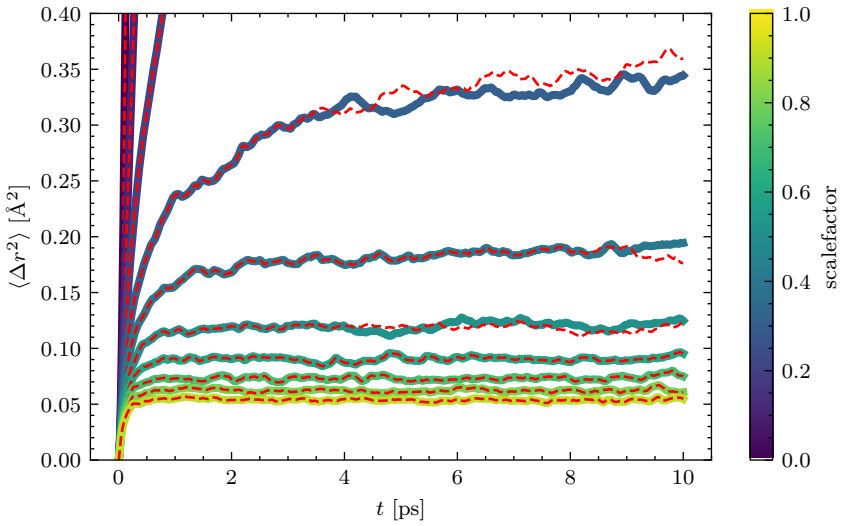


Figure 5.2: Mean-squared displacement of α-quartz scaled manually (coloured lines) and using fix adapt in LAMMPS (red dashed lines)

This is shown in Fig. 5.2, where we observe that both implementations overlap exceptionally well, as expected given all the MD simulations were ran with a fixed seed. The slight deviations observed after 4 ps in some of the

[7] https://github.com/evenmn/generate-parameter-files

lines may be attributed to slight numeric differences in the implementations, resulting in slightly different trajectories. As the scale-factor goes to zero we also observed a straight line indicating linear diffusion as expected. Hence the methods are deemed self-consistent, and for practical reasons the LAMMPS implementation chosen as one used in the subsequent TI calculations.

5.3.5 Detecting Equilibration

In computing Eqn. 5.4 an important aspect is to ensure that the trajectory we are averaging over is indeed an ensemble of *equilibrium* states. Furthermore, it is also desirable that we obtain statistically robust estimates of the standard error. We will in this subsection and the next tackle both of these topics, which together forms a method of fully automatic the analysis of our MD trajectories.

First, we consider the problem of detecting equilibration. Each step in the integrand of Eqn. 5.4 are evaluated from individual MD simulations, starting from an ideal reference state with velocities sampled from a Maxwell-Boltzmann distribution. This incurs a period of equilibration, during which we are not sampling from the canonical ensemble, and the trajectories from this period of equilibration should not be included in the calculation of any ensemble averages. As is often observed in MD, the equilibration time should also be expected to differ for differing conditions and system sizes. With the goal of automating as much of the statistical analysis as possible, we have considered several approaches. First there is the reverse cumulative averaging method by Yang, Bitetti-Putzer and Karplus [50], which looks for convergence in the function

$$f(i,n) = \bar{X}_i^{RCA} = \frac{1}{i}\sum_{j=n}^{n-1+1} X_j \tag{5.10}$$

to determine equilibration with respect to the normally distributed observable $\{X\}$. For correlated data-sets, such as MD trajectories, this requires corse-graining to de-correlate the dataset such that the central-limit theorem applies. However, the proposed method to accomplish this did not yield comparable results according to what was presented by the authors for our system and the method was therefore abandoned in favour of an alternative. We look to the method by Chodera [51], which was inspired by [50], but as noted by the author does not rely on the underlying distribution being Gaussian, as opposed to reverse cumulative averaging. In this method, we consider the effective number of uncorrelated points in a sub-sample of the trajectory $t = [t', \tau]$

$$N_{\text{eff}}(t') \equiv \frac{\tau - t' + 1}{g_t} \tag{5.11}$$

where τ is the total simulation time, and $g_t \equiv 1 + 2\tau_c$ the statistical inefficiency, derived from the correlation time τ_c. The method then works by setting $t' = \tau$, and gradually decreasing it with the expectation that N_{eff} should become larger as we include a larger portion of the trajectory. That is, until atypical data is eventually encountered as t' approaches 0 and the non-equilibrated portion of the trajectory enters into the calculation, at which stage we expect N_{eff} to sharply decrease. Therefore, the equilibration time, t_0, is defined simply as the t' for which N_{eff} is maximal. The approach was tested in multiple trajectories, one of which are shown in the top half of Fig. 5.3, where the non-equilibrated

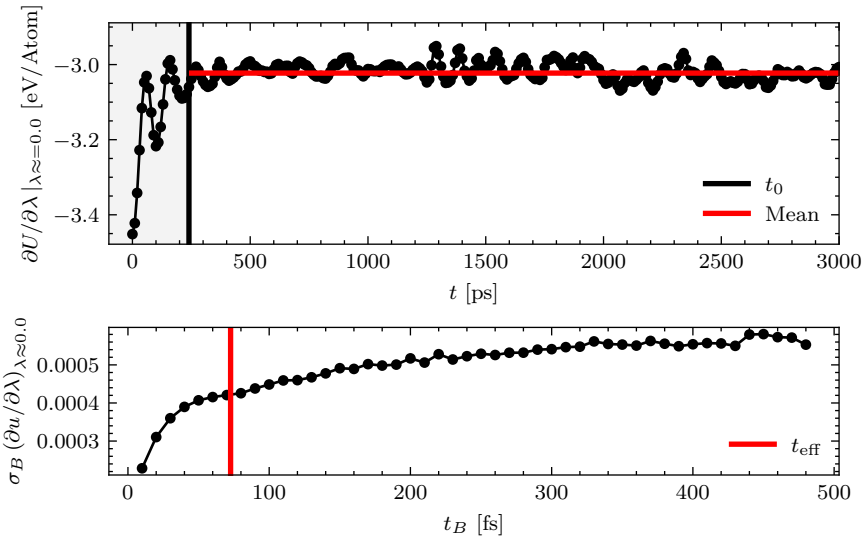


Figure 5.3: First few fs of a TI step with $\lambda = 0$ (pure Einstein crystal) along with the block-averaged estimate of the standard deviation of the mean where $t_{\text{eff}} = \Delta t \cdot N_s/N_{\text{eff}}$

region has been shaded in grey. The method was observed to perform well throughout the trajectories it was tested on, and the algorithm was implemented as part of the automated analysis.

5.3.6 Estimating the Standard Error

The error of the quadrature Eqn. 5.4 may be estimated by [39]

$$\sigma_{\langle\Delta f\rangle} = \sqrt{\sum_{i=1}^{n} w_i^2 \sigma_{\langle\partial u/\partial\lambda\rangle_i}} \tag{5.12}$$

where the standard deviation of the mean value of $\partial u/\partial\lambda_i$ is computed as

$$\sigma_{\langle\partial u/\partial\lambda\rangle_i} = \frac{\sigma_{\partial u/\partial\lambda_i}}{\sqrt{N_s}} \tag{5.13}$$

for N_s independent samples. It is however usually the case that molecular-dynamics trajectories are correlated with a non-zero correlation time τ_c. From this it follows that the central-limit theorem is no longer valid, and by extension Eqn. 5.13 is not directly applicable and will yield an under-estimate of the true standard deviation if employed on the entirety of the measured trajectory [52, 53]. The review article by Grossfield et al. [54] summarizes many of the best-practice methods for obtaining error estimates in MD as well as MC, but as they remark, there is no universal solution to this problem. In their textbook, Allen and Tildesley [24] also suggests several methods, one of which is the blocking method by Flyvbjerg and Petersen [55]. The blocking method is composed of what is essentially a series of renormalization group transformations,

$$x_i' = \frac{1}{2}(x_{2i-1} + x_{2i}). \tag{5.14}$$

The authors showed that the mean is invariant under this transformation, and that the standard deviation should converge to an estimate of its true value. The method is particularly appealing given is requires no additional parameters, and should in theory work for any equilibrated MD trajectory. The method was tested on some TI calculations for α-quartz at $T = 300\,\mathrm{K}$, $P = 1\,\mathrm{bar}$, but no clear convergence was observed for the trajectories it was tested on.

Another suggested approach is the so-called block-averaging method. In this method, we split our trajectory into a set of M non-overlapping blocks of size N_B. Each block $b = 1, \ldots, M$ then contains indices $(b-1)M < i \leq bM$. The mean of each block is then computed individually for each block b as

$$\langle \mathcal{O} \rangle_b = \frac{1}{N_B} \sum_{i \in b} \mathcal{O}(t_i). \tag{5.15}$$

For sufficiently large blocks the block-averages are uncorrelated, and thus their standard deviation

$$\sigma^2(\langle \mathcal{O} \rangle_{\mathrm{b}}) = \frac{1}{M-1} \sum_{b=1}^{m} (\langle \mathcal{O} \rangle_b - \langle \mathcal{O} \rangle)^2 \tag{5.16}$$

provides a good estimate of the true standard deviation of the mean. A notable challenge in using block-averaging with the intent of a fully automated statistical analysis is the determination of a suitable block-size, N_B. Allen and Tildesley [24] offers some helpful pointers, but in general some manual observation is advised. In Khanna et al. [39], they based their block-size on two

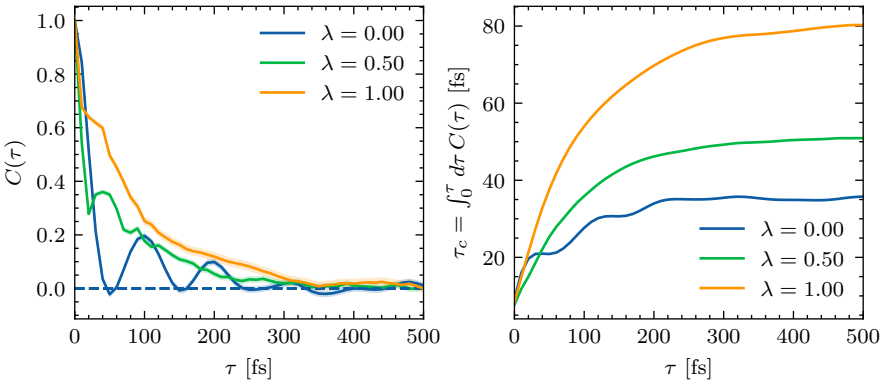


Figure 5.4: Auto-correlation function of $\partial u/\partial\lambda$, $C(\tau)$, for α-quartz consisting of $N = 1350$ atoms at $T = 293\,\mathrm{K}$ with a lattice equilibrated at $P = 1\,\mathrm{bar}$

different estimates of the correlation time. One of which was by integrating the auto-correlation function, $C(\tau)$, with respect to time. This was tested for some several systems, for example Fig. 5.4, but had a tendency to yield low correlation times. Furthermore, $C(\tau)$ did not always converge causing τ_C to fluctuate and was thus abandoned. However, if we recall from the previous subsection the method of Chodera [51] gives us N_{eff}, the number of effective independent samples out of the N_s samples contained in the trajectory. With this, we may efficiently determine a suitable block-size as N_s/N_{eff}. This was tested for different trajectories, one of which is depicted in Fig. 5.3 with the

automatically deduced block-time t_{eff} marked in red. Throughout testing this approach seemed well behaved for the systems being studied, and was thus implemented as the second part of the automated analysis.

So in summary, the automated analysis scheme for TI trajectories is composed of two stages. first, the un-equilibrated region is filtered out by the method of Chodera [51], then the standard error is approximated by block-averaging, with a block-size of $N_B = N_s/N_{\text{eff}}$.

5.4 Non-Equilibrium Thermodynamic Integration

Having thoroughly discussed the implementation of thermodynamic integration in LAMMPS, we now move on to the non-equilibrium variant. As a reminder non-equilibrium thermodynamic integration in the Einstein crystal method is done by computing the work done

$$W = \int_0^{t_s} dt \, \frac{\partial \mathcal{H}(\lambda_t)}{\partial \lambda_t} \frac{\partial \lambda_t}{\partial t} \tag{5.17}$$

in going from an ideal Einstein crystal to a Vashishta crystal in both directions, in a single dynamic process governed by the time-dependent potential

$$U(t) = \lambda(t) U_E + (1 - \lambda(t)) U_C \tag{5.18}$$

with switching function $\lambda(t)$. The change in free energy between the two systems is then given by

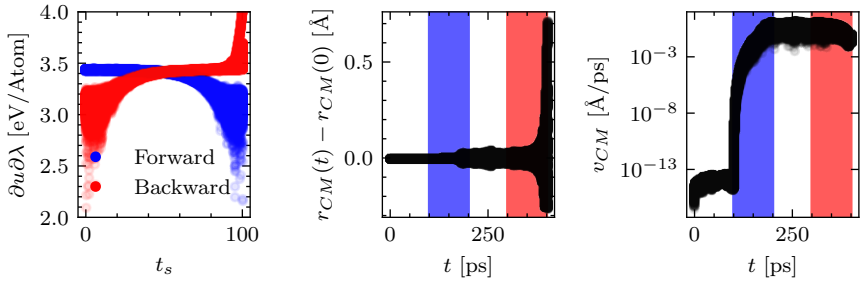
$$\Delta F = \frac{1}{2}(\langle W_{E \to C} \rangle + \langle W_{C \to E} \rangle). \tag{5.19}$$

averaged over multiple realizations of the process. In LAMMPS, this is made relatively easy by the `spring/ti` fix by Freitas, Asta and Koning [36]. This fix runs the process by transforming from a Vashishta crystal to an Einstein crystal, and back again with several possible switching functions $\lambda(t)$. Notably, this is done without requiring a `scale` parameter as discussed in section 5.3.4, making the approach more accessible by comparison.

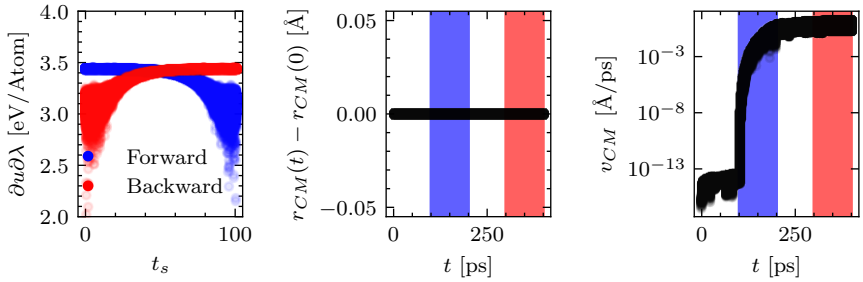
Many of the same practical considerations apply to both the regular and non-equilibrium TI, namely the choice of thermostat still matters (cf. section 5.3.3) and the center of mass must be constrained (cf. section 5.3.2). The latter of which we revisit in this section and perform a thorough study of the effects of different constraints. In regards to the switching function during the used during the switching process between the two systems, we follow the literature [56] and employ Eqn. 4.26.

5.4.1 The Effects of Different Center of Mass Constraints

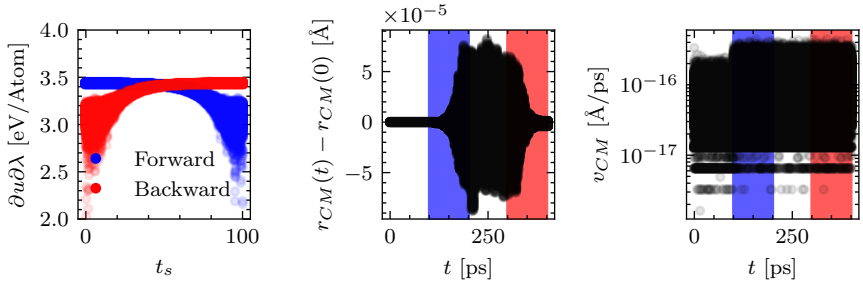
How to constrain the center of mass in Einstein crystal computations implemented in LAMMPS has been has been discussed several times in the literature [36, 39, 42]. However, in following these instructions a drift in the momentum of the center of mass (COM) was observed in our simulations of Vashishta SiO_2. We have therefore performed a comprehensive study of the effects of the different possible constraints to the COM in order to determine the optimal one. The results are summarized in Fig. 5.5, where we observe 30 realizations of the forward and backward switching indicated by blue and red



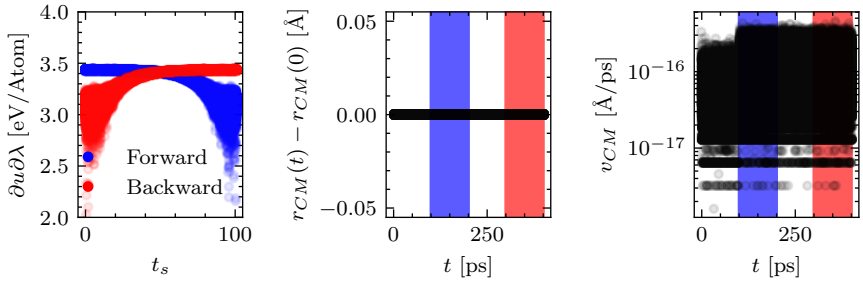
(a) No constraints



(b) Constrained COM with the `fix recenter` command



(c) Constrained COM momentum with the `fix momentum` command



(d) Both constraints

Figure 5.5: Non-Equilibrium thermodynamic integration from an Einstein crystal to Vashishta SiO_2 performed with different constraints.

respectively. The un-coloured, white regions in the r_{CM} and v_{CM} plots indicate the equilibration period between each switching process.

We start by considering the unconstrained process in 5.5a. In these simulations, we apply no explicit constraints on the center of mass yet still ensure that the thermostat does not act on the center of mass with the `zero yes` argument to the Langevin thermostat as described in section 5.3.2. We observe that once the switching process begins, the COM gains significant speed yet the COM is kept reasonably in place. This makes sense given the harmonic springs of the Einstein potential fixing the atoms about their initial positions. Then once the backward process begins, and the harmonic springs begin to relax we observe that the center of mass begins to shift quite dramatically. We will not spend much trying to find out why exactly this occurs, but a likely explanation may lie in the time dependent nature of the switching potential $U(\lambda_t)$. Say for example that at some point during the forward switching that the system is in a region of configuration space that is accessible to a Vashishta crystal, but not to an Einstein crystal, essentially that one of the springs are stretched a bit further than what it normally would. The system would then gain an unnatural jolt of energy. Say this occurs multiple time during the process, and thus the COM may gain some energy. It is worth to note that this is purely a speculation based on physical intuition on my part, and no in-depth analysis has been done.

With this in mind, we then move on to Fig. 5.5b, where `recenter` fix has been added to impose an explicit constraint on the COM. Whilst the center of mass is kept seemingly in place, we observe the exact same behaviour of v_{CM} as in the unconstrained case. This may be explained by the way the `recenter` works, which is essentially just performing a change of coordinates by setting r_{cm} equal to its initial value at each step[8], hence it is not sufficient for Einstein crystal calculations. What is instead advised is to use the `momentum` fix, with the re-scale argument which preserves the momentum net momentum in the system and thus enabling the thermostat to work as intended. As we observe in Fig. 5.5c, this has remedied the problem of the COM gaining speed, and does a reasonably good job at keeping the system centered. Lastly we see in Fig. 5.5d, where both `recenter` and momentum are applied and the small drift of r_{CM} is zeroed out.

Table 5.1: Change in free energy computed with different constraints

	Δf	$2\sigma_{\langle \Delta f \rangle}$	Abs. diff. in Δf from both fixes
no constraint	3.399124342	$8.9 \cdot 10^{-5}$	$3 \cdot 10^{-4}$
fix recenter	3.398827734	$7.6 \cdot 10^{-5}$	$1 \cdot 10^{-6}$
fix momentum	3.398826365	$7.2 \cdot 10^{-5}$	$1 \cdot 10^{-8}$
both fixes	3.398826355	$7.2 \cdot 10^{-5}$	–

To see the effects the different constraints has on the resulting change in free energy, we look to table 5.1. Taking the case of applying both fixes as the

[8] cf. this discussion on the LAMMPS mailing list https://matsci.org/t/center-of-mass-drift-in-long-npt-simulation/24197/9

ground truth, we observe that the difference falls within 2σ with 30 samples for 3 constraint method, with only the unconstrained method deviating. However, if one needs a higher level of precision and were to increase the number of samples by roughly an order of magnitude, the effects of the constraints would begin to matter as $\sigma \propto 1/\sqrt{N}$ and the error in bias from solely using fix recenter may exceed that due to variance.

5.5 Testing thermodynamic integration and non-equilibrium thermodynamic integration

In order to evaluate the two methods described so far, namely equilibrium and non-equilibrium variants of thermodynamic integration, we will apply the two methods by computing the free energy of α-quartz. In particular, we look at α-quartz equilibrated to $T = 300\,\mathrm{K}$, $P = 1\,\mathrm{bar}$ for system sizes $N = 720, 1350, 2268, 3969, 5760, 8019, 10800$, each of the initial states being constructed using the molecular-builder package [57]. The crystals were first allowed to relax in NVT for 5 ps followed by an equilibration period of 100 ps in NPT. The simulation box dimensions were then sampled for an additional 100 ps, and subsequently scaled to its average shape. The per-atom positions are then sampled for 50 ps, and set to the average position. This state is then stored, and to be used as the reference states for the subsequent TI calculations. The mean-squared displacement relative to the reference state, for Si and O atoms individually, is then measured for 100 ps, and the average is used to compute the spring constants of the Einstein crystal reference states by Eqn. 5.5.

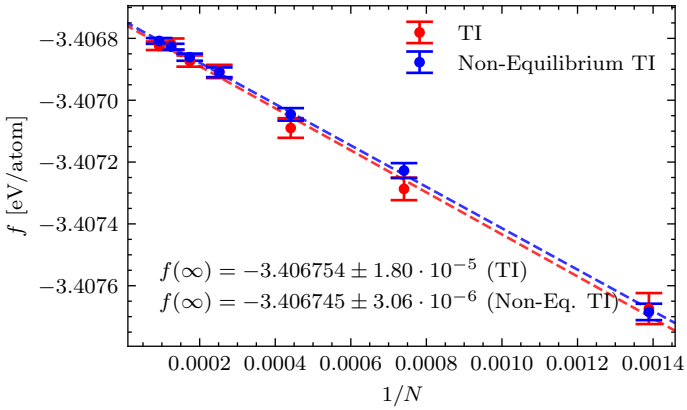


Figure 5.6: Finite-size scaling of the free energy in α-quartz at $T = 300\,\mathrm{K}$ with a lattice equilibrated at $P = 1\,\mathrm{bar}$ and the free energy calculated using the Einstein crystal method with thermodynamic integration as well as non-equilibrium thermodynamic integration. Weighted least-squares regression with inverse-variance weighting is employed, and twice the standard error is used as the uncertainty estimate in $f(\infty)$ in both cases.

In the equilibrium method, thermodynamic integration is performed along a 25 point Gauss-Legendre quadrature, sampling for 100 ps at each step. For the non-equilibrium variant we employ a switching time $t_s = 100\,\mathrm{ps}$ with 50 ps of equilibration prior to both the forward and backward switching processes.

Non-equilibrium thermodynamic is repeated 10 times, with the mean taken as the change in free energy.

The resulting free energies are shown in Fig. 5.6, where we have extrapolated to the thermodynamic limit, $f(\infty)$, according to the finite-size scaling relation Eqn. 4.36. We note the heteroskedastic[9] noise in $f(N)$, and weighted least-squares regression with inverse-variance scaled weights is employed in favour of ordinary least-squares as the best unbiased linear estimator [58].

We observe that both methods are consistent to the fourth decimal, with the non-equilibrium variant having errors in 10^{-6}, an order of magnitude lower than in the equilibrium variant. Note that for each N, a total 2.5 ns and 3 ns of simulation data went in to the equilibrium and non-equilibrium variants respectively. Hence, this extra order of magnitude in precision was obtained by only 20% additional simulation data, and by extension, computation time. Furthermore, we may arbitrarily add precision to the non-equilibrium result by running additional simulations. Where we should expect an extra order of magnitude of precision as we increase the number of simulations by a power of 10. In contrast, we may not as trivially increase the precision of the equilibrium TI integral. Whilst be may sample additional data at each point in the integrand, we may not arbitrarily add new points to the already computed Gauss-Legendre quadrature points. Hence, to increase the precision of the integral, it would have to be redone in its entirety for an entirely new set of quadrature points.

5.6 Free Energy Perturbation Followed By Thermodynamic Integration

Lastly we consider the variation of the Einstein crystal method as it is described in [34, 40, 41, 42], which we refer to as the Vega path, consisting of a free energy perturbation from an Einstein crystal to an interacting Einstein crystal followed by thermodynamic integration to the Vashishta crystal. Whilst the last to be presented, this was the first approach that was attempted by virtue of not requiring us to scale the Vashishta potential, meaning that it is accessible in LAMMPS without modification. As usual LAMMPS scripts implementing both steps are included in appendices C.4 and C.5, which should give sufficient details on how to implement this method.

As a reminder, the change in free energy is in this method computed in two steps. First the difference in free energy between the ideal Einstein crystal and a so-called interacting Einstein crystal, that is a crystal governed by the combined potential

$$U(\boldsymbol{r}^N) = U_E(\boldsymbol{r}^N) + U_C(\boldsymbol{r}^N), \tag{5.20}$$

by free energy perturbation (FEP). Following this, thermodynamic integration is performed to compute the difference in free energy between the interacting Einstein crystal and the Vashishta crystal along the linear homotopy

5.6.1 Free Energy Perturbation in LAMMPS

In free energy perturbation (FEP) we aim to compute

$$\Delta F_1 = \ln \left\langle e^{-\beta U_C} \right\rangle_E \tag{5.21}$$

[9] i.e the noise is not constant in $1/N$

which entails measuring the Vashista potential in a system governed solely by the Einstein crystals harmonic potential. In LAMMPS this may be done by first running a simulation consisting of an ideal Einstein crystal using the fix spring command, then re-running the resulting trajectory (dumpfile) and measuring the potential energy with the Vashishta potential (cf. appendix C.4). We again constrain the center of mass as described in section 5.3.2 and utilize the Langevin thermostat following the same rationale as in section 5.3.3. In regards to the choice of spring constants we deviate from the choices made for the previous two cases. Previously we opted for MSD based springs following the recently published results in [39], where the Einstein crystal method was generalized to the case of different spring constants. This was not yet available at the time [34, 42] was published, and thus we instead opted to follow their strong recommendation of artificially setting the mass of the atoms in the system to be the same. Vega et al. [34] states, without citation or argument, that this does not change the coexistence properties. We thus briefly explore this idea in appendix A.

The exponential $\exp\{-\beta U_C\}$ has a tendency to grow large, leading to numeric overflow [34, 41] and it is advised to rewrite the expression by adding and subtracting the potential energy of the ideal reference state, U_L, in the exponential yielding

$$\Delta F_1 = U_L - \frac{1}{\beta} \ln \left\langle e^{-\beta(U_C - U_L)} \right\rangle_E . \tag{5.22}$$

In regards to the spring constant, k, it is recommended [34] to set it such that

$$\ln \langle \exp\{-\beta(U_C - U_L)\} \rangle_E \approx 0.02\, N k_B T. \tag{5.23}$$

Since we do not know a priori which k satisfies this recommendation one must perform a search, either manually or by the use of some algorithm like for example a midpoint method.

5.6.2 Thermodynamic Integration in LAMMPS (The Vega Path)

In the Vega path we perform thermodynamic integration from the interacting Einstein crystal to the Vashishta crystal along the linear homotopy

$$U(\lambda) = (1 - \lambda)(U_E + U_C) + \lambda U_C \tag{5.24}$$

which has the derivative

$$\frac{\partial U(\lambda)}{\partial \lambda} = -U_E. \tag{5.25}$$

To further suppress the effects of the expected weak divergence in $\langle U_E \rangle$, it is recommended [34, 41] to rewrite the TI integral to the form

$$\Delta F_2 = -\int_{\ln(c)}^{\ln(k+c)} d(\ln(\lambda k + c))\, \langle U_E \rangle_\lambda\, (\lambda k + c) \tag{5.26}$$

which again is integrated numerically with a Gauss-Legendre quadrature method. Otherwise, all the usual concerns of constraining the center of mass (section 5.3.2) and thermostats (section 5.3.3) still apply.

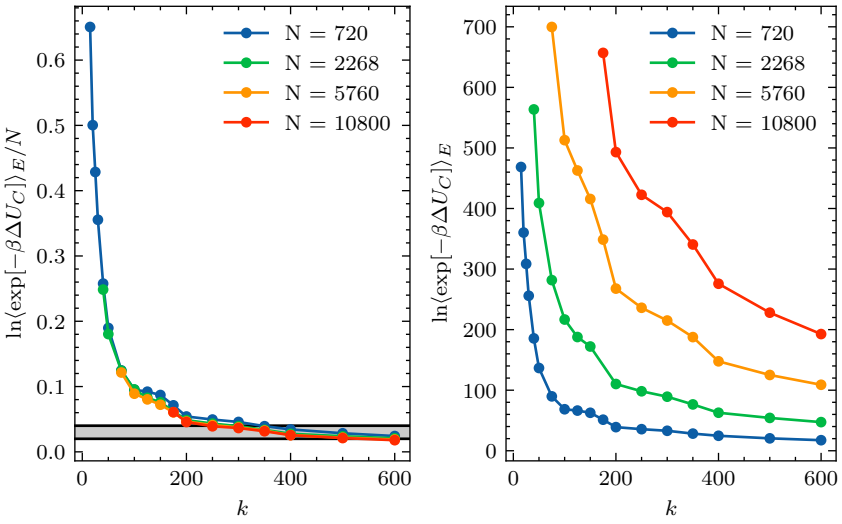


Figure 5.7: Part of the ΔF_1 term in the Vega variation of the Einstein crystal method for α-quartz equilibrated at $P = 1\,\text{bar}$, $T = 300\,\text{K}$, where the shaded region in (a) corresponds to the recommendation by [34]

5.6.3 Testing the Vega Pathway

In order to test the method, look at α-quartz equilibrated to $P = 1\,\text{bar}$, $T = 300\,\text{K}$ according to the methodology in section 5.2. Following [34], we artificially set the mass of both Si and O to $20\,\text{g/mol}$, approximately the average mass in SiO_2. We performed FEP for a selection of $k \in [1, 600]$ and several system sizes. We look at the behaviour of the left hand term in Eqn. 5.23 in Fig. 5.7 where in 5.7a we have shaded the interval $[0.02, 0.04]$. The first thing to note is the exclusion points greater than ~ 700 in Fig. 5.7b. These points were not willingly excluded, but rather due to numerical overflow in $\ln\langle\exp\{-\beta\Delta U_C\}\rangle$. Hence subtracting U_L on its own is not is not sufficient for Vashishta SiO_2, and the exponents still grow too large for numerical evaluation. Looking at Fig. 5.8 we observe how this may be a problem. Setting k to $200 - 300$ as required to avoid the numerical divergence means that the atoms are essentially locked in place to their rest positions, and we will in turn have to perform thermodynamic integration over an exceedingly long interval where very little happens. The first attempt to salvage this problem and numerically evaluate the exponential at lower k, was to expand in a series of cumulants

$$\ln\left\langle e^{-\beta\Delta U_C)}\right\rangle_E = \sum_{n=1}^{\infty} \kappa_n \frac{(-\beta)^n}{n!} \tag{5.27}$$

where κ_n denotes the n-th cumulant of ΔU_C. This trick worked quite well for $k > 5$, but produced some rather large numbers upwards of 10^{36} for $k < 2$. As

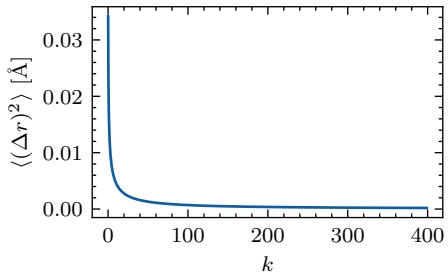


Figure 5.8: Mean-squared displacement of an interacting Einstein crystal with respect to the spring constant, k

a second attempt, we instead further re-write the expression

$$\begin{aligned}
\ln\left\langle e^{-\beta\Delta U}\right\rangle &= \ln\left(\frac{1}{M}\sum_{i=1}^{M} e^{-\beta\Delta U_i}\right) \\
&= \ln\left(\sum_{i=1}^{M} e^{-\beta\Delta U_i \cdot X/X}\right) - \ln M \\
&= \ln\left(\sum_{i=1}^{M} e^{X} e^{-\beta\Delta U_i/X}\right) - \ln M \\
&= \ln\left(\sum_{i=1}^{M} e^{-\beta\Delta U_i/X}\right) + X - \ln M
\end{aligned} \tag{5.28}$$

where we may set X as a large number, for example the maximum value in ΔU_i. This lets us evaluate Eqn.5.21 with minimal risk of numerical overflow even for small k. This leaves the question of how we should choose k. We note that the recommendation of setting k such that Eqn. 5.23 is satisfied does not seem viable for Vashishta SiO_2 which as we see in Fig. 5.7 this corresponds to setting k such that the atoms are essentially tethered to their rest positions as observed in Fig. 5.8.

We therefore opted to redo the simulations with mean-squared displacement based springs for Si and O individually, using the more recently published finite-size correction for Einstein crystals with different spring constants [39]. We also utilize Eqn. 5.28 to avoid numerical overflow in the exponential when computing Δf_1.The resulting free energies are plotted in Fig. 5.9, where we according to the finite-size scaling relation [40] expect $f(N)$ to be linear in $1/N$. As we see in Fig. 5.9 this is not the case, even though the free energies are quite close to what we expect numerically based on our previous TI calculations. Given that we already have two working methods, we will not spend the time to perform an in depth analysis of why the method fails. The error may likely be attributed to the degree at which the phase-space of Vashishta SiO_2 and the Einstein crystal fails to overlap. When this is the case, there is a large bias contribution to the mean-squared error in FEP, cf. Chipot and Pohorille [31]. There also exists methods to systematically determine the magnitude of this effect [59].

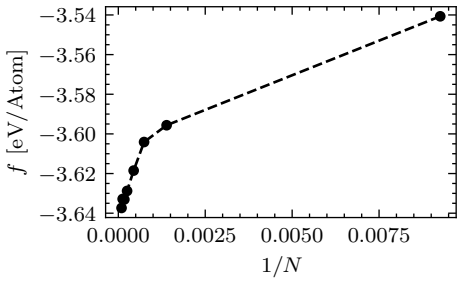


Figure 5.9: Free energy of α-quartz at $P = 1\,\text{bar}, T = 300\,\text{K}$ computed with the the Einstein crystal method using FEP followed by TI (The vega path)

5.7 Summary

We have in this section given an outline of an efficient workflow that aims to enable the computation of free energies using the Einstein crystal method for a large number of systems across N, P and T. We have further given detailed descriptions on how to implement three different variations of the Einstein crystal in LAMMPS, namely by TI and non-equilibrium TI along the Frenkel & Ladd path, as well as by FEP followed by TI in the Vega path. Upon evaluation of the three methods, we discovered self-consistency in the free energies computed by TI and non-equilibrium TI, with the latter offering some practical benefits whilst also yielding higher precision for a comparable simulation time. We also discovered that the Einstein crystal method along the Vega path did not yield results according to our expectations for α-quartz simulated in the Vashishta potential.

Hence, for the reasons above we opt to employ non-equilibrium thermodynamic for all further free energy calculation as we search for coexistence points in silica polymorphs.

CHAPTER 6

Finding points of coexistence between Silica polymorphs

In this chapter we aim to find points of coexistence between different phases of crystalline Silica. Silica has exceptionally complex phase behaviour [1] with a large number of metastable states [4, 5], and we will thus limit ourselves to only the ones which are known to be thermodynamically stable, namely α-quartz, β-quartz, β-tridymite[1], coesite and stishovite. The latter of which, stishovite, is also dropped as it is not held stable by the Vashishta potential.

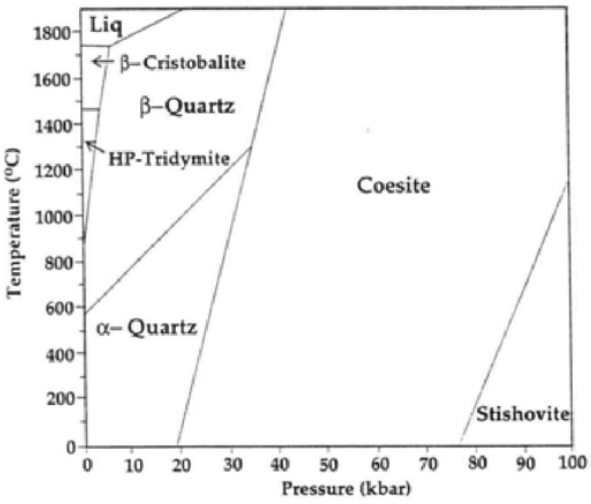


Figure 6.1: Experimentally determined phase diagram of Silica [60]

Our approach in finding the coexistence points follows that in Vega et al. [34]. We compute Gibbs free energy along isotherms and isobars looking for points in which it is equal for two different polymorphs. As a guide, we look at an existing experimental phase diagram, depicted in Fig. 6.1, which should give us a rough idea where to look. It should however be noted that whilst qualitative features may be captured, the phase diagram is not necessarily accurately reproduced by classical interaction potentials [61]. Hence, our search space along isotherms

[1] Sometimes referred to as HP-tridymite

and isobars should thus be sufficiently broad and somewhat lacking in bias towards existing phase diagrams.

We will replicate our analysis for three different parameterizations of the potential. First there is the the original 1990 parameterization by Vashishta et al. [14], then there is the 1997 parameterization by Broughton et al. [15] which was tuned to study quartz oscillators. Lastly, we study the unpublished parameterization by Wang et al. [19] which was fit to experimental values of the bulk modulus, equilibrium volume and melting temperature of silica. There exist an additional widely available parameterization by Nakano, Kalia and Vashishta [62] which was excluded from this study as it was a modified version of the 97 parameters aimed primarily at modeling amorphous silica, and is therefore not particularly relevant to our study of crystalline silica.

6.1 Coexistence points along an isobar

Based on the experimentally determined silica phase diagram [4] we expect four phase transitions to occur along an isobar at $P = 1\,\text{bar}$. At low temperatures, α-quartz will is the thermodynamically stable polymorph until it undergoes a displacive transition to β-quartz due to a slight shift in its structure characterized by the mean Si-O-Si bond angles increasing from 143.61° at room temperature to 150.9° at 846 K [1, 4]. At 1140 K β-tridymite becomes the thermodynamically stable phase, yet upon heating β-quartz will bypass this transition and remain meta-stable until it undergoes a reconstructive phase transition to cristobalite at $\sim 1323\,\text{K}$ [4, 63]. β-cristobalite then becomes the thermodynamically stable phase at 1743 K until the melting point at 1978 K.

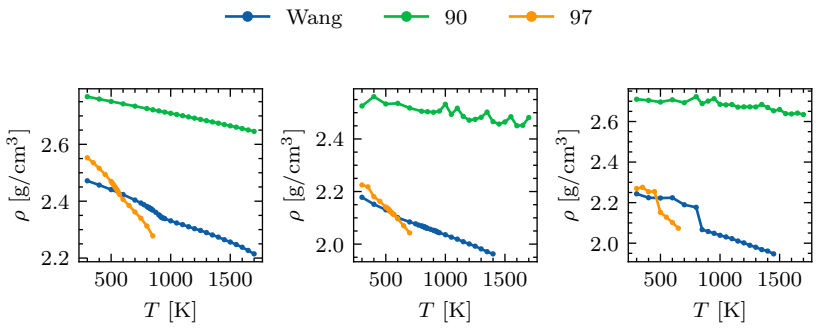
6.1.1 Procedure

We look for these phase transitions by computing free energies of each structure along an isobar $P = 1\,\text{bar}$. In order to generate reference states for the Einstein crystal calculations, we follow the procedure described in section 5.2, using experimental data for α-quartz, β-tridymite [64] and β-cristobalite [65] as initial structures. Each structure allowed to equilibrate in NPT for 100 ps followed by an additional 100 ps of sampling the simulation box, which is then re-scaled to its average shape. We then run 50 ps in NVT whilst sampling the positions of each atom and set them to their average value. This configuration is then stored to be used as the reference state in the Einstein crystal calculations. We then measure the mean-squared displacement of Si and O atoms individually 50 ps use this to compute the spring constants used in the Einstein crystal reference state. We then perform non-equilibrium thermodynamic integration from the Einstein crystal to the Vashishta crystal with a switching time $t_s = 100\,\text{ps}$ and 25 ps of equilibration prior to both the forward and backward switching. The non-equilibrium thermodynamic integration is repeated 10 times for each NPT, and the mean is produced as the change in free energy. The resulting Helmholtz free energy is then used to compute the Gibbs free energy by the relation $G = F - PV$, where P is taken as the pressure at which the Vashishta crystal is equilibrated to during the initialization procedure. For each phase, pressure and temperature this procedure is repeated for a selection of 5 crystal

sizes with $N \in [1 \cdot 10^3, 2 \cdot 10^4]$, and the finite-size scaling relation [40]

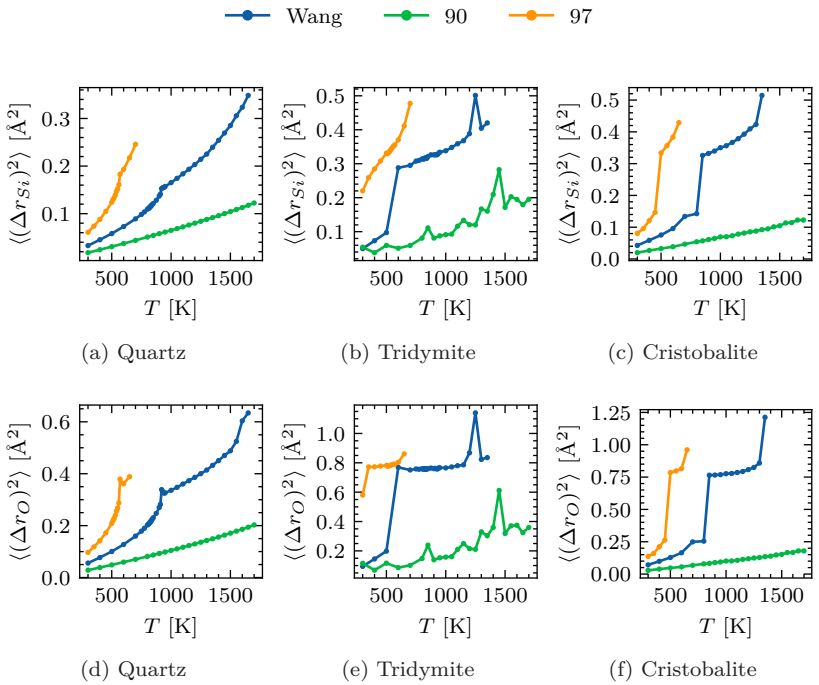
$$g(N) = g(\infty) + \frac{d}{N} \tag{6.1}$$

is used to estimate the free energy per particle of an infinitely large system by inverse-variance scaling least-squares regression, taking the intercept as $\lim_{N\to\infty} g(N)$.



(a) Quartz (b) Tridymite (c) Cristobalite

Figure 6.2: Density profile of the equilibrated reference states generated during the initialization procedure for all simulations along an isobar at $P = 1\,\text{bar}$



(a) Quartz (b) Tridymite (c) Cristobalite

(d) Quartz (e) Tridymite (f) Cristobalite

Figure 6.3: Mean-squared displacement of Si and O atoms in Vashishta SiO_2 along an isobar at $P = 1\,\text{bar}$

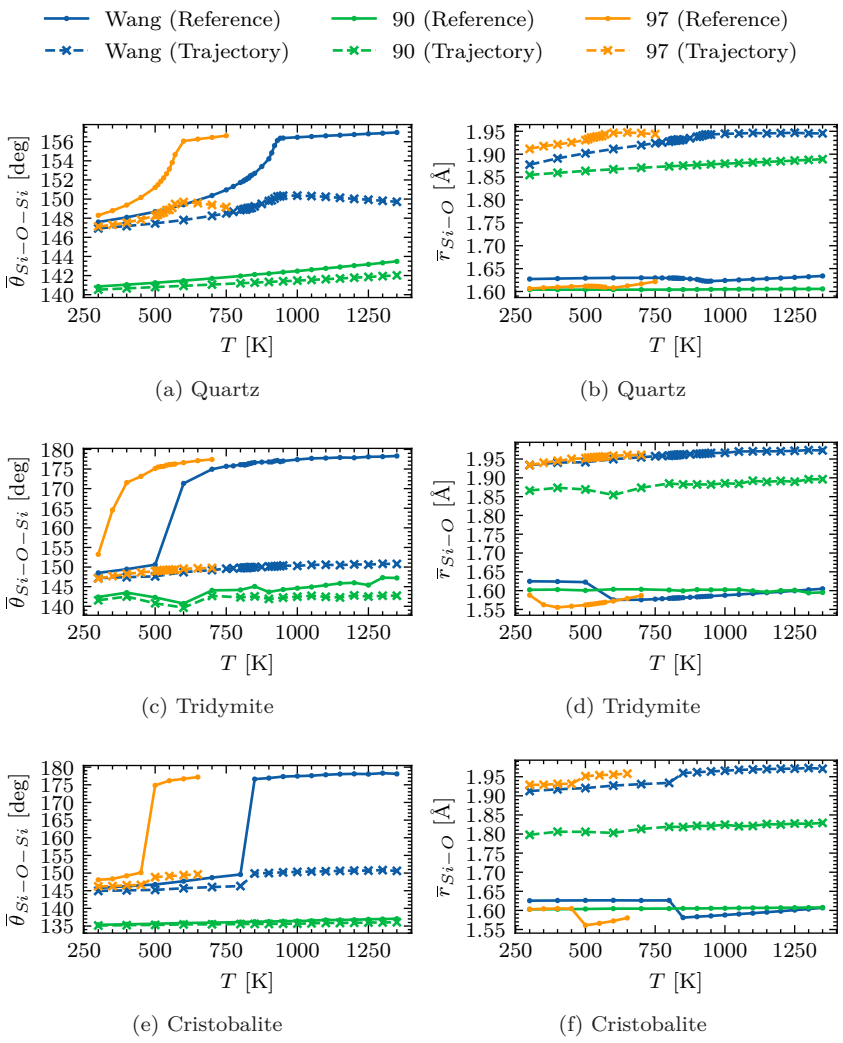
Using the experimental phase diagram as a guide, we perform a broad search in T to infer about the behaviour of the different parameterizations, setting the upper bound in T to when we observe melting as indicated by a large drop in the density ρ.

6.1.2 Structure along the isobar

To understand how the crystals behave along the isobar, and perhaps more importantly that our desired structure is obtained during equilibration, we begin our analysis by looking at the structure of the crystals. While there exists many sets of order parameters to quantify the structure of molecular systems, notably the Steinhardt order parameters [66], there is to my knowledge none that quantify the structural changes in Silica polymorphs. Hence to gain an overview of the different structures produced in our simulations we rely on 4 measures. The density of the equilibrated reference states, ρ, which we expect to be a continuous function in P and T for equilibrium states. The mean-squared displacement, $\left\langle (\Delta \boldsymbol{r})^2 \right\rangle$ of Vashishta SiO_2, quantifies the movement in our system. As well as the average Si-O-Si bond angles and Si-O bond distances in both the equilibrated reference states and averaged over several independent configurations of Vashishta SiO_2 in MD simulations, frequently used measures when inferring about the structure of SiO_2 [1, 4]. The measurements are shown in Figs. 6.2, 6.3 and 6.4 for quartz, tridymite and cristobalite and all 3 parameterizations of the Vashishta potential considered.

Cristobalite

We begin our analysis by looking at the densities Fig. 6.2 and the mean-squared displacement of the Si atoms 6.3. Looking first at β-cristobalite, we note the rather apparent discontinuity in both the density and mean-squared displacement in the Wang and 97 parameterizations. Looking at a selection of reference structures at both ends of the discontinuity, two of which are depicted in Fig. 6.5, we see quite clearly that the crystal structure of cristobalite is only present at temperatures following the drop in density, where the structure at $T = 500\,\mathrm{K}$ is unstructured, lacking periodicity and thus more than likely not cristobalite. We then look to the bond angles and distances in Figs. 6.4e, 6.4f, which may require some explanation to be fully understood. First, we note that the "Reference" measurements are made in the single, static configuration that is generated using the initialization step (section C.1). The "Trajectory" measurements on the other hand are made by sampling the bonds in 20 independent configurations along a dynamic trajectory governed by the Vashishta potential. So while related, the discrepancy between the two is not surprising as they measure fundamentally different things. As we recall from section C.1, the per-atom positions in the reference states were set by averaging their x, y and z positions individually in time, which runs the risk of slightly shifting them away from their true trajectory for any atoms that are not oscillating freely, and uniformly in each direction of x, y, z. For β-cristobalite, this process yields a structure that is similar to the ideal structure proposed by Wyckoff [65] with 180° Si-O-Si bond angles. However, it was argued by Nieuwenkamp [67] that β-cristobalite exhibits a disordered structure in which oxygen atoms move freely in a about a circular plane normal to vectors between



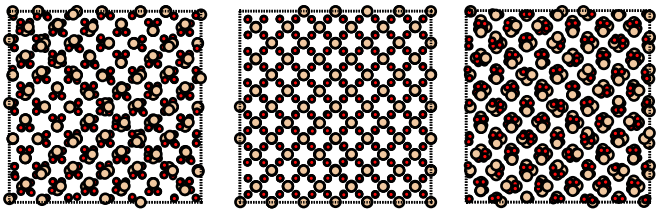
(a) Quartz (b) Quartz

(c) Tridymite (d) Tridymite

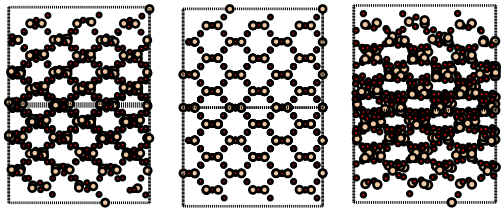
(e) Cristobalite (f) Cristobalite

Figure 6.4: Arithmetic mean of Si-O-Si bond angles and Si-O bond lengths in Quartz, Tridymite and Cristobalite reference states and trajectories along an isobar at $P = 1$ bar for the Wang [19], 90 [14] and 97 [15] parameterizations of the Vashishta potential

pairs of silicon atoms, rather than being fixed to their midpoint as sketched in Fig. 6.6. With this in mind, the discrepancy between the Si-O-Si bond angles in the trajectory and reference state is unsurprising, as the way in which we average the atomic positions during when constructing the reference states will naturally produce a structure akin to that proposed by Wyckoff [65] with bond angles of 180, yet in motion the true structure is expected to have Si-O-Si bond angles of $\sim 148°$ [4], which is in line with what we observe. It should however be noted that we will never measure the exact "ideal" bond angle in a trajectory, as the atomic positions will always deviate somewhat from the ideal structure, and the measured bond angle is in essence a function of the mean-squared



(a) $T = 500\,\mathrm{K}$ (Wang) (b) $T = 1000\,\mathrm{K}$ (Wang) (c) $T = 1000\,\mathrm{K}$ (90)



(d) $T = 500\,\mathrm{K}$ (Wang) (e) $T = 1000\,\mathrm{K}$ (Wang) (f) $T = 1000\,\mathrm{K}$ (90)

Figure 6.5: Cristobalite reference states generated at $P = 1\,\mathrm{bar}$ with temperatures (a,d) $T = 500\,\mathrm{K}$ and (b,e) $T = 1000\,\mathrm{K}$ using the Wang parameterization of the Vashishta potential, and (c,f) $T = 1000\,\mathrm{K}$ using the 90 parameterization. Each structure is viewed from two different perspectives.

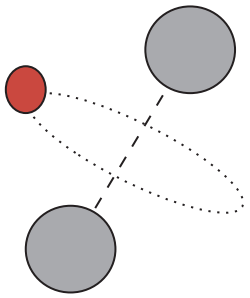
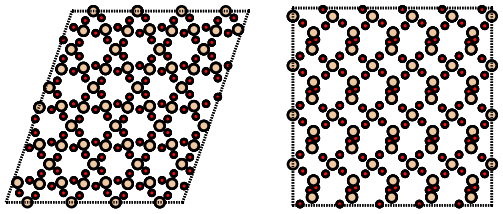


Figure 6.6: Visual depiction of how Oxygen (red) moves in relation to pairs of Silicon atoms (grey) in β-cristobalite

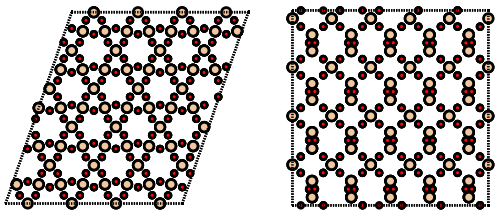
displacement of the atoms in each time-step. Yet it nonetheless serves as a useful insight. Similar observations is made for the 97 parameterization, but at lower temperatures. For the 90 parameterization we may by observing the noise in $\rho(T)$ seen in Fig. 6.2 and angles in Fig. 6.4e question if β-cristobalite is held stable. By observing a selection of the atomic structures it was indeed concluded that is not, as the generated reference states did not correspond to cristobalite. See for example Fig.6.5c

Quartz

For quartz we do not, in contrast to cristobalite, observe discontinuities in the densities and mean-squared displacement (Figs. 6.2,6.3) but rather distinctive



(a) α-Quartz,$T = 300\,\mathrm{K}$ (b) α-Quartz,$T = 300\,\mathrm{K}$



(c) β-Quartz,$T = 1000\,\mathrm{K}$ (d) β-Quartz,$T = 1000\,\mathrm{K}$

Figure 6.7: α-quarts (a,b) and β-quartz reference structures of the Wang parameterization of Vashishta SiO_2, each viewed from two different perspectives

shifts in the gradients, in the Wang and 97 parameterizations. Furthermore, in both these parameterizations, the Si-O-Si bond angles exhibit behaviour that is distinctive to the displacive phase transition between α/β-quartz. That is, we expect the bond angles in α-quartz to gradually increase until the β-quartz structure essentially locks in place [1, 4, 68, 69]. By observing the reference states, Fig. 6.7, one may with a trained eye indeed confirm that the α/β transition has occurred. Note in particular the additional horizontal symmetry of the oxygen atoms (red) present in Fig. 6.7d compared to Fig. 6.7b. The same may also be observed in Figs. 6.7a, 6.7c, but is a bit harder to spot[2]. This behaviour is notably observed regardless of whether we use α or β quartz as the initial structure during the initialization procedure (section 5.2), and bond angles measured at the same (P, T) in both cases overlapped completely, indicating that the phase transition occurs during the initialization procedure. This means that will not be able to find the free energy difference between the two phases using the Einstein crystal method and by extension, not locate the coexistence points. However, given this lack of meta-stability we may yet obtain estimates of the coexistence line simply by locating the temperatures at which the Si-O-Si bond angles flattens out, and by repeating this along several isobars we may get a reasonable estimate of the coexistence line. We do this for wide range of isobars for both the Wang and 97 parameterizations, with the former shown in Fig. 6.8.

In regards to the 90 parameterization, no such behaviour is observed and the crystal retains the structure of α-quartz all throughout the isobar exhibiting no signs of a phase transition or structural changes of any kind in Figs.6.2, 6.3, 6.4. This was also reaffirmed by looking at the structures by eye, and not observing

[2]Unless you happen to have spent the better part of a year starting at silica polymorphs

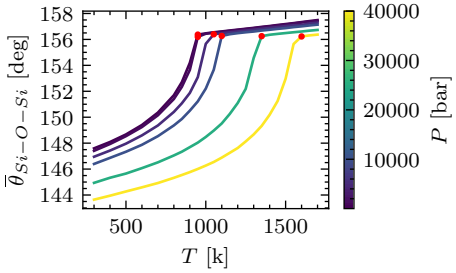
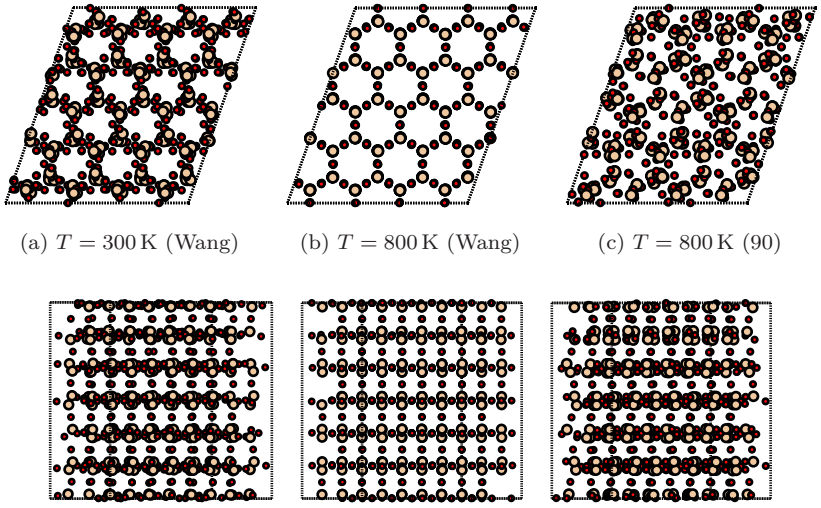


Figure 6.8: Mean Si-O-Si bond angles in reference structures for the Wang parameterizations along several isobars, where the red dots indicates the estimated location of the α/β-quartz transition



(a) $T = 300\,\mathrm{K}$ (Wang) (b) $T = 800\,\mathrm{K}$ (Wang) (c) $T = 800\,\mathrm{K}$ (90)

(d) $T = 300\,\mathrm{K}$ (Wang) (e) $T = 800\,\mathrm{K}$ (Wang) (f) $T = 800\,\mathrm{K}$ (90)

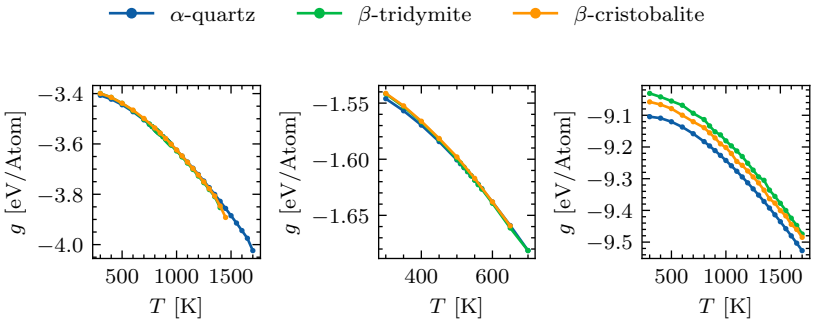
Figure 6.9: β-tridymite reference states generated by the Wang parameterization of the Vashishta potential at (a,d) $P = 1\,\mathrm{bar}$ and (b,e) $T = 300\,\mathrm{K}, 800\,\mathrm{K}$, additionally (c,f) the 90 parameterization at $T = 800\,\mathrm{K}$. The structures are viewed from two perspectives, highlighting the stacking hexagonal structure of β-tridymite.

a symmetry similar to the one seen in Fig. 6.7.

Tridymite

In tridymite we again observe discontinuities in mean-squared displacement in Figs. 6.2, 6.3 in the 97 and Wang parameterizations, indicative of a structural change. The 90 parameterization being noisy throughout, hinting that the structure may not be held stable. The bond angles in Fig. 6.4 tells the same story. Much the same as in β-cristobalite, the ideal structure of β-tridymite was originally proposed to have 180° Si-O-Si bond angles [4], yet the oxygen atoms were later shown to be slightly displaced from their ideal positions resulting

in Si-O-Si bond angles of 149.2° [4]. As was the case for β-cristobalite, the discrepancy between the reference and trajectory bond angles in Fig. 6.4 are consistent with these observations. As an additional check, we again look at a selection of the structures, a few of which are shown in Fig. 6.9. As expected, the structures at temperatures before the structural change in the Wang and 97 parameterizations are not stable in contrast to the ones at higher temperatures. Further, we observe that the 90 parameterization is unable to keep tridymite in a stable form.



(b) Wang (c) 97 (d) 90

Figure 6.10: Gibbs free energy for T along an isobar at $P = 1$ bar

6.1.3 Coexistence points along the isobar

Now that we have an understanding of how the structures of the different polymorphs evolve along the isobar, we turn to Fig. 6.10 which shows the computed free energies of all 3 systems, including the regions in T where they were determined not to correspond to silica polymorphs. In isolation these are seemingly well behaved throughout the entire range in T satisfying the condition for thermodynamic stability

$$\left(\frac{\partial^2 G}{\partial T^2}\right)_P \leq 0. \tag{6.2}$$

Hence, additional analysis of the structures similar to what we did in the previous subsection is paramount when attempting to locate coexistence points using the Einstein crystal method such that one does not draw false conclusions.

Given our structural analysis, and coincidentally also Fig. 6.14c we may conclude that the 90 parameterization of of the Vashishta potential does not undergo any phase-transitions along $P = 1\,\text{bar}$, and we thus reduce further analysis to just the Wang and 97 parameters. Turning our attention to Fig. 6.11 we have the difference in free energy between all 3 polymorphs along the $P = 1\,\text{bar}$ isobar for the Wang and 97 parameterizations. Consistent in both of these, we observe a phase transition between α-quartz and β-tridymite at $T_c = 818.9\,\text{K}, 543.1\,\text{K}$ for the Wang and 97 parameterizations respectively. Notably prior to the expected α/β-quartz transition in both cases. No coexistence is observed between β-tridymite and β-cristobalite prior to melting, yet coexistence between the meta-stable β-quartz and β-cristobalite

is observed in both parameterizations in the region where β-tridymite is the thermodynamically stable state, that is, the polymorph with the lowest free energy.

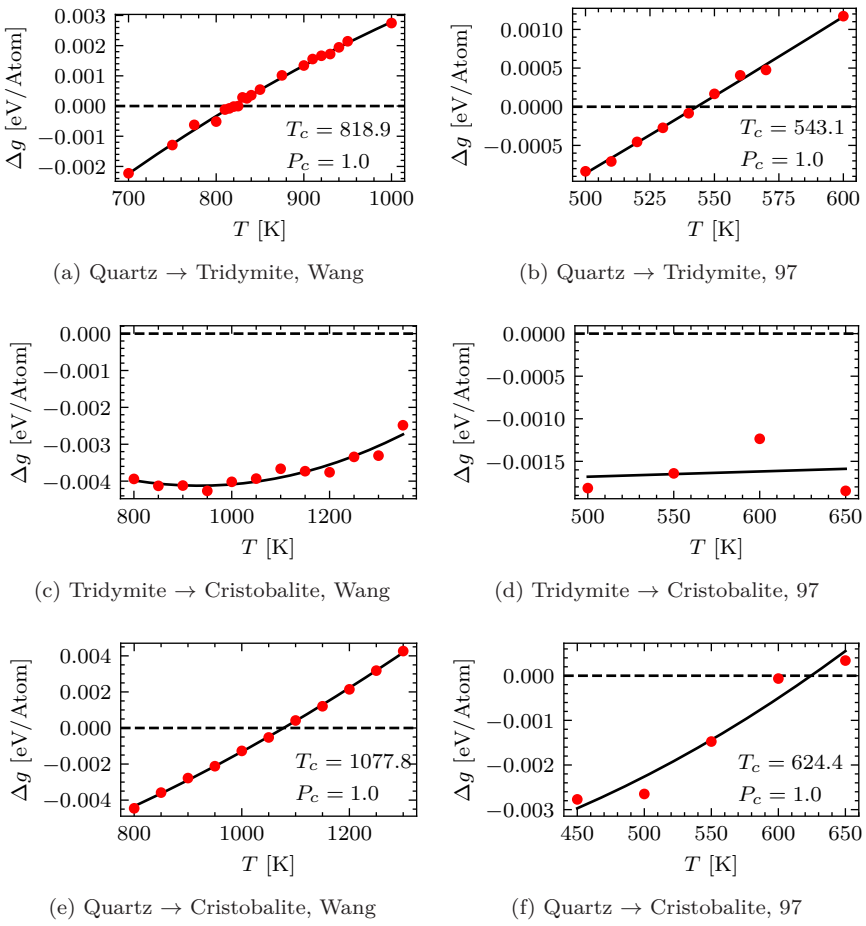
Looking at Fig. 6.11a we note the signs of noise in Δg. The computed uncertainties in the Einstein crystal free energy calculations were for the free energies making up this plot all in an order of 10^{-10} to 10^{-11}, thus the accounted for errors are vanishingly small, and does not explain the observed variation in Fig. 6.11a, or indeed any of these plots. We may question whether we have accounted for all of the statistical noise, i.e variance, in our analysis. As of now we account for the variance in non-equilibrium thermodynamic integration over the 10 trials that go in to these computations, as well as the uncertainty in the $f(\infty)$ estimator in the linear fit during finite-size scaling. However it should be noted that we also impose an implicit bias in our calculation by our choice of reference state. That is, the one that is produced by the initialization scheme. One of the assumptions in the Einstein crystal method is that our reference states are "close enough", with no strict requirements that they correspond exactly to the target structure [34]. It is however a requirement that there are no phase-transitions when integrating the linear homotopy in thermodynamic integration. Can we really be so sure that this is the case? As we have observed for β-cristobalite in particular, it is quite difficult to determine the structures we actually obtain during the initialization scheme. Furthermore, as seen in quartz, transitions between the α/β structures will readily during initialization. Nevertheless, these effects are likely rather small as suggested by the predictable, systematic nature in Fig. 6.11. Furthermore, since we lack a set of order parameters that quantify the exceptionally rich phase behaviour of Silica polymorphs, it is rather difficult to be certain if this is indeed a source of the observed noise.

6.2 Coexistence points along an isotherm

In addition to the coexistence points we have already located, we also expect a transition between α-quartz and coesite. In contrast to the coexistence lines we have previously looked at, the α-quartz/coesite transition intersects with P axis at $T = 0$, as opposed to intersecting with T at $P = 0$. Hence, the most effective strategy in locating a coexistence point in this case is to compute free energies along as isotherm. Given that we know α-quartz is stable at 300 K, we choose this as our isotherm and compute free energies along a wide span in P. We repeat the same procedure as for the isobars, described in section 6.1.1 and obtain the free energies depicted in Fig. 6.14. Here we observe that the stability criterion for Gibbs free energy

$$\left(\frac{\partial^2 G}{\partial P^2}\right)_T \leq 0 \tag{6.3}$$

is satisfied. We also observe the continuity of $\rho(P)$ in Fig. 6.13, again being indicative that there are no significant structural changes along the therm. As opposed to the other polymorphs we have considered, coesite is not classified by the average Si-O-Si bond angle, but in stead 4 separate Si-O-Si bond angles [5]. Hence in the interest of brevity, we simply our analysis somewhat and look at the a selection of the reference structures, some of which are shown in Fig. 6.15.



(a) Quartz → Tridymite, Wang

(b) Quartz → Tridymite, 97

(c) Tridymite → Cristobalite, Wang

(d) Tridymite → Cristobalite, 97

(e) Quartz → Cristobalite, Wang

(f) Quartz → Cristobalite, 97

Figure 6.11: Difference in Gibbs free energy along isobars at $P = 1$ bar for quartz, tridymite and cristobalite modeled by the Wang and 97 parameterizations of the Vashishta potential, with coexistence temperature (T_c) and pressure (P_c) indicated in the figures.

The observed structures were according to expectation for all of the computed pressures, including the both ends of the isotherms in all 3 potentials. Hence, we are quite satisfied that what we are looking at is indeed coesite, and that it is well behaved. For quartz, a similar analysis as for the isobar was performed and we observed clear indications that quartz remained as α-quartz throughout. This is also reaffirmed by the Si-O-Si bond angles in Fig. 6.12.

We see the difference in Gibbs free energy along the isotherms in Fig. 6.16. Here we observe the α-quartz/coesite transition in all 3 parameterizations at $P_c = 112.6, 48.6$ and 4.5 kbar for the Wang, 97 and 90 parameterizations respectively.

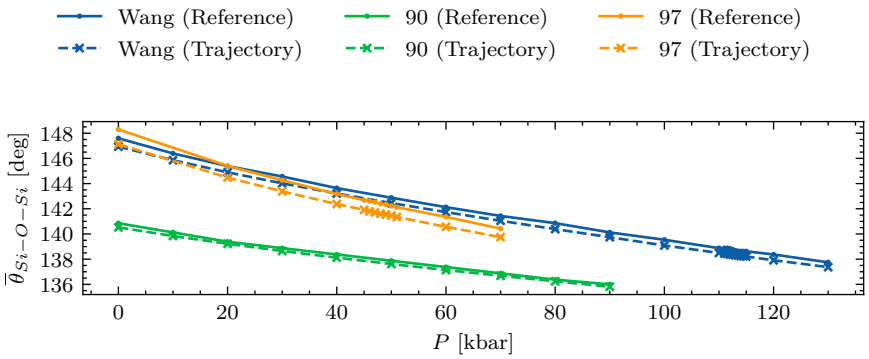
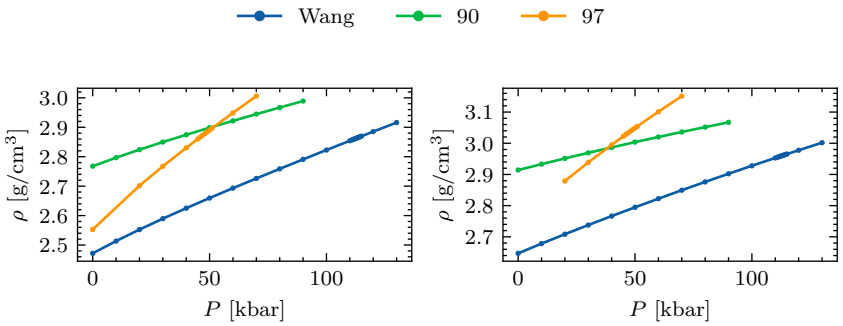
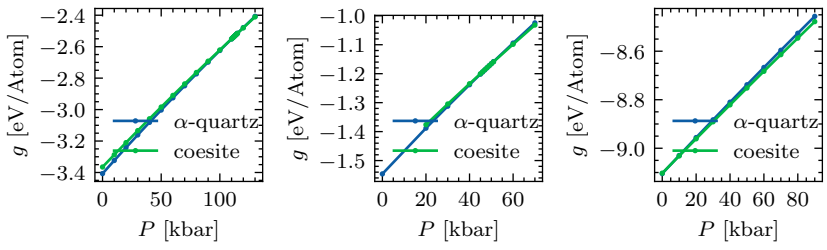


Figure 6.12: Arithmetic mean of Si-O-Si bond angles and Si-O bond lengths in α-quartz reference states and trajectories along an isotherm at $T = 1\,\mathrm{K}$ for the Wang, 90 and 97 parameterizations of the Vashishta potential



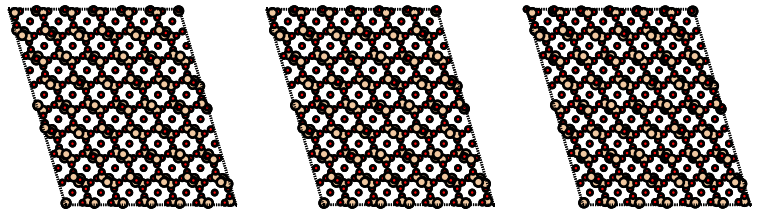
(a) Quartz (b) Coesite

Figure 6.13: Density, ρ, for an isotherm along $T = 300K$ for α-quartz and coesite for the Wang, 90 an 97 parameterizations of the Vashishta potential.

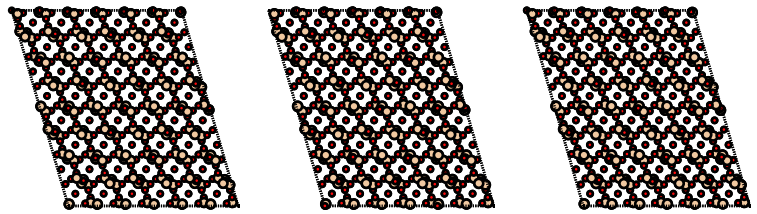


(a) Wang (b) 97 (c) 90

Figure 6.14: Gibbs free energy per particle, g, for an isotherm along $T = 300$ K for the Wang, 90 and 97 parameterizations of the Vashishta potential.

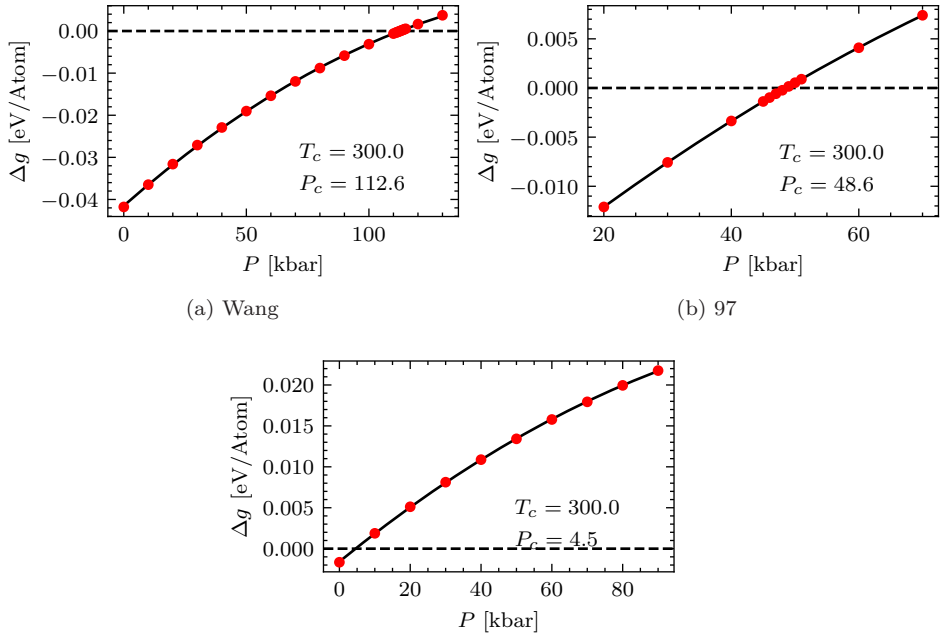


(a) $P = 20\,\mathrm{kbar}$ (Wang) (b) $P = 20\,\mathrm{kbar}$ (97) (c) $P = 20\,\mathrm{kbar}$ (90)



(d) $P = 60\,\mathrm{kbar}$ (Wang) (e) $P = 60\,\mathrm{kbar}$ (97) (f) $P = 60\,\mathrm{kbar}$ (90)

Figure 6.15: Coesite reference states at $T = 300\,\mathrm{K}$ and $P = 20\,\mathrm{kbar}, 60\,\mathrm{kbar}$ generated by the (a,b) Wang, (b,e) 97 and (c,f) 90 parameterizations of the Vashishta potential.



(a) Wang (b) 97

(c) 90

Figure 6.16: Free energy difference between α-quartz and coesite along an isotherm at $T = 300\,\mathrm{K}$

CHAPTER 7

Computing the phase diagram of Silica

Having obtained a set of coexistence points, we may now employ the Gibbs-Duhem integration method by Kofke [22] to numerically solve the Clausius–Clapeyron equation, a first-order differential equation

$$\frac{\mathrm{d}P}{\mathrm{d}T} = \frac{\Delta h}{T\Delta v} \tag{7.1}$$

where $\Delta h, \Delta v$ denotes the difference enthalphy and volume per particle between the two coexisting phases. We will in this chapter begin to describe how the method may be implemented in LAMMPS. Following this, we present the phase diagram of SiO_2 modelled by the Wang [19] as well as the 97 [15] parameterizations of the Vashishta potential.

7.1 Gibbs-Duhem Integration in LAMMPS

In the context of molecularm dynamics simulations, Eqn. 7.1 takes the form

$$\frac{\mathrm{d}P}{\mathrm{d}T} = \frac{\langle h\rangle_{\mathrm{II}} - \langle h\rangle_{I}}{T(\langle v\rangle_{\mathrm{II}} - \langle v\rangle_{\mathrm{I}})} \tag{7.2}$$

where $\langle\cdot\rangle_{\mathrm{I}}, \langle\cdot\rangle_{\mathrm{II}}$ denotes ensemble averages in systems I and II respectively. In solving Eqn. 7.2 evaluations of the source term is not simply a function call, but rather a the result of MD simulations of phases I and II separately at the same (P, T). Hence, evaluations of Eqn. 7.2 relatively expensive and usage of a standard solver such the fourth order Runge-Kutta method is both time and compute intensive. However, given that the coexistence line is a smooth [22], usually monotone [23] function it is advised [22, 24, 33] to employ predictor-corrector (PECE) methods [70]. In cases where the coexistence line is particularly steep in $P(T)$, it may be useful to instead take the inverse of Eqn. 7.1 and solve for $\mathrm{d}T/\mathrm{d}P$ [34].

The MD simulations required to integrate Eqn. 7.2 are themselves quite straight forward, consisting of standard NPT simulations whilst sampling the volume and enthalpy. The challenge in implementing Gibbs-Duhem integration instead lies in combining these NPT simulations with a differential equation solver which requires a lot of scripting, and at present no publicly available tools exist to aid this purpose. Hence i have developed a flexible python

package that should work for a wide range of different systems, including Vashishta SiO_2, and made it openly available on Github [71]. The package currently supports Forward-Euler, Runge-Kutta 4 and a PECE method for solving Eqn. 7.2. Additionally, the equilibration detection method by [51] is utilized to ensure that the averages in h, v are indeed sampled from equilibrium states.

7.1.1 Predictor-Corrector Method

In the interest of minimizing the number of simulations required to produce the coexistence line we employ the following predictor-corrector scheme [24]. Let y denote the dependent variable, P or T, and f the source term in the corresponding Clauseus-Clapeyron equation For the first step we employ the trapezoid PECE method

$$\begin{aligned} \tilde{y}_1 &= y_0 + hf_0 \\ y_1 &= y_0 + \frac{h}{2}(\tilde{f}_1 + f_0) \end{aligned} \tag{7.3}$$

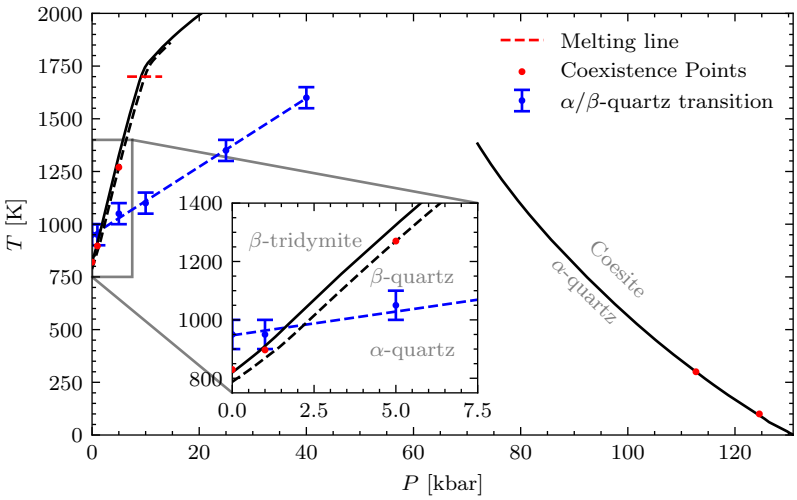
obtaining y_0, y_1, requiring two function evaluations. To obtain the next two points $i = 2, 3$ we then employ the midpoint PECE method to obtain

$$\begin{aligned} \tilde{y}_{i+1} &= y_{i-1} + 2hf_i \\ y_{i+1} &= y_{i-1} + \frac{h}{3}\left(\tilde{f}_{i+1} + 4f_i + f_{i-1}\right) \end{aligned} \tag{7.4}$$

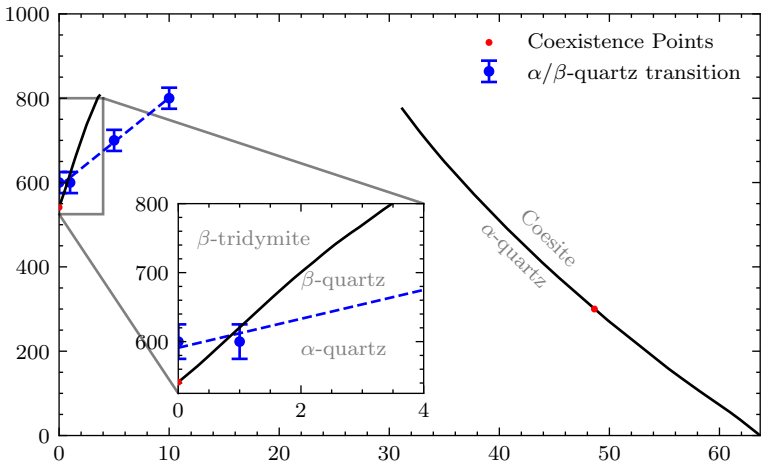
requiring two new function evaluations at each step. Having obtained the 4 points, $i = 0, 1, 2, 3$. We may continue with Adams predictor corrector

$$\begin{aligned} \tilde{y}_{i+1} &= y_i + \frac{h}{24}(55f_i - 59f_{i-1} + 37f_{i-2} - 9f_{i-3}) \\ y_{i+1} &= y_i + \frac{h}{24}\left(9\tilde{f}_{i+1} + 19f_i - 5f_{i-1} + f_{i-2}\right) \end{aligned} \tag{7.5}$$

for the remainder of the steps. The benefit of this methodology being that we only need a two set of MD simulations per integration step for 4-th order accuracy [70] compared to a 4-th order Runge-Kutta method which would require four sets of MD simulations run in sequence for each step.



(a) Wang Parameterization



(b) 97 Parameterization

Figure 7.1: Phase diagrams for the solid phases in the Wang [19] and 97 [15] parameterizations of Vashishta SiO_2, where the red dots are coexistence points found in free energy calculations. In (a) Gibbs-Duhem integration was performed from $P_c = 112.6\,\text{kbar}, T_c = 300\,\text{K}$ for the quartz/coesite line and $P_c = 1\,\text{bar}, T_c = 820\,\text{K}$ (solid), $P_c = 5\,\text{kbar}, T_c = 1270\,\text{K}$ (dashed) for the quartz/tridymite line. In (b) we started from $P_c = 1\,\text{bar}, T_c = 541\,\text{K}$ for quartz/tridymite and $P_c = 48.6\,\text{kbar}, T_c = 300\,\text{K}$. In both cases, the α/β-quartz line is estimated by looking at change in Si-O-Si bond angles along isobars, where the uncertainty is taken as the resolution in ΔT.

7.2 Phase diagram of Vashishta Silica

We perform Gibbs-Duhem integration solving Eqn. 7.2 for $P(T)$ with a step size $\Delta T = 1\,\mathrm{K}$ using the PECE method described in section 7.1.1. This was done for coexistence points between quartz/tridymite and quartz/coesite in both the 97 and Wang parameterizations of the Vashishta potential. Prior to each Gibbs-Duhem integration we copied the closest available reference state in (P, T) obtained during the free energy calculations in chapter 6. The systems were equilibrated for 200 ps and the resulting configuration was stored as the initial state for the Gibbs-Duehm integration. For each evaluation of the source term in Eqn. 7.2, 100 ps of MD was performed and averages were computed from the equilibrated parts of the trajectories, selected automatically according to the algorithm by Chodera [51] (cf. section 5.3.5). The resulting phase diagrams are shown in Fig. 7.1

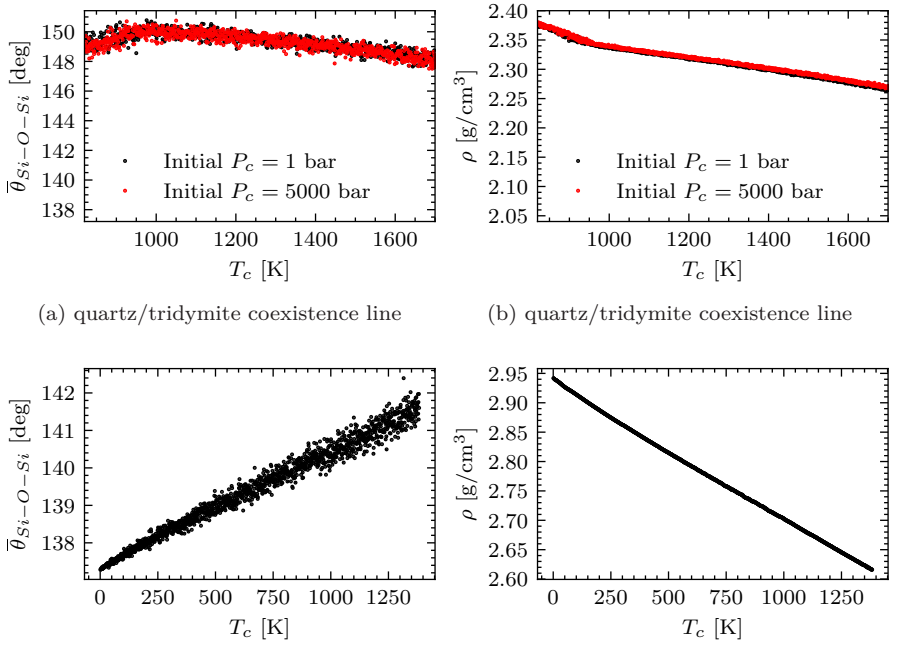
For the α-quartz/coesite transition in the Wang parameterization we integrate starting from the coexistence point at $P_c = 112.6\,\mathrm{kbar}, T_c = 300\,\mathrm{K}$. The systems used were composed of $N = 2269, 2880$ atoms for α-quartz and coesite respectively. We also located an additional coexistence point from free energy calculations following the procedure as described in chapter 6. This additional coexistence point is located at $P_c = 124.5\,\mathrm{kbar}, T_c = 100\,\mathrm{K}$, and is used to verify the correctness of the coexistence line computed by Gibbs-Duhem integration. In regards to the quartz/tridymite transition, we expect a phase transition from α-quartz to β-quartz somewhere along the quartz/tridymite coexistence line based on what we observed in section 6.1.2. Hence, we integrate from the α-quartz/β-tridymite coexistence point at $P_c = 1\,\mathrm{bar}, T_c 820\,\mathrm{K}$ as well as a β-quartz/β-tridymite coexistence point that was located at $P_c = 5\,\mathrm{kbar}, T_c = 1270\,\mathrm{K}$. In both cases systems composed of $N = 2269, 2016$ atoms were used for α/β-quartz and β-tridymite respectively. We also located an additional point at $P_c = 1\,\mathrm{kbar}, T_c = 897\,\mathrm{K}$ with free energy calculations, again, to verify the method.

For the 97 parameterization we have been a bit more sparse in our usage of computational resources, and integrate starting from the coexistence point $P_c = 48\,\mathrm{kbar}, T_c = 300\,\mathrm{K}$ for the α-quartz/coesite transition, and at $P_c = 1\,\mathrm{bar}, T_c = 541\,\mathrm{K}$ for the α-quartz/β-tridymite transition. The same system sizes have been utilized as in the Wang calculations.

In addition to the coexistence lines by Gibbs-Duhem integration, we have also included the line at which the α/β quartz transition is estimated to occur as inferred by the Si-O-Si bond angles in the reference systems. cf. Fig. 6.8 and the discussion section 6.1.2.

7.2.1 Discussion and comparison with existing phase diagrams

The first, and perhaps most interesting thing to note about the phase diagrams in Fig. 7.1 is that they are both qualitatively very similar, differing in scale by roughly a factor 2. Both potentials produce a quartz/tridymite line that is of the correct qualitative shape, even though in contrast to the experimental phase diagram (Fig. 6.1) the α/β-quartz transition intersects with the line. To confirm that the α/β-quartz transition happens during Gibbs-Duhem integration, we measure Si-O-Si bond angles as well as the densities in the trajectories as seen in Fig. 7.2 for the Wang parameterization. This is in line with what we observed



(a) quartz/tridymite coexistence line

(b) quartz/tridymite coexistence line

(c) quartz/coesite coexistence line

(d) quartz/coesite coexistence line

Figure 7.2: Mean Si-O-Si bond angles and density in quartz during Gibbs-Duhem integration along the quartz/tridymite (a-b) and the quartz/coesite (c-d) coexistence lines in the Wang parameterization of the Vashishta potential.

in our structural analysis in section 6.1.2. Note also that the overlap in bond angles and densities in the coexistence lines that were found in $P_c = 1\,\text{bar}$ and $P_c = 5\,\text{bar}$, showing that the α/β-quartz transition will readily occur in a simulation trajectory. Note that the trajectories in used Gibbs-Duhem integration are continuous in the sense that the full configuration $(\boldsymbol{r}^N, \boldsymbol{v}^N)$ are carried over between each step. Similar behaviour was also observed in the 97 parameterization. This means that in both parameterizations, there is a triple-point between α-quartz, β-quartz and β-tridymite along the quartz/tridymite coexistence line. Next we move on to look at the melting line marked in Fig. 7.1a. As remarked by Frenkel and Smit [33], Gibbs-Duhem integration has no built-in diagnostics, and additional checks should always be made. This is one such case, wherein the coexistence line may not in isolation look problematic. However, as we observe in Fig. 7.3 tridymite melts at $T \approx 1700\,\text{K}$, thus any results above this melting line should be disregarded. It should further be noted that this is not necessarily the point at which molten silica becomes thermodynamically stable as this may occur at a lower temperature, but likely serves as a reasonable upper bound estimate.

We then look to the α-quartz/coesite transition. In both cases the transition occurs at rather high pressures compared to what is expected experimentally, within the region where we would expect a coesite/stishovite transition. We also note that the coexistence line is tilted in the opposite direction to that of the experimental phase diagram, as well as the β-quartz/coesite coexistence line

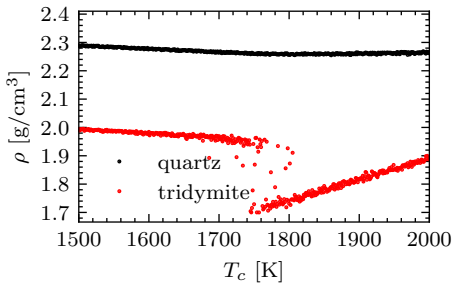


Figure 7.3: Density of quartz and tridymite along the coexistence line integrated from $P_c = 1\,\text{bar}, T_c = 820\,\text{K}$ in the Wang parameterization of the Vashishta potential.

was determined for the BKS interaction potential by Saika-Voivod et al. [61]. In a recently published article, the α-quartz/coesite coexistence line estimated by Erhard et al. [72] using a quasi-harmonic approximation method [73] for several silica models. Among these was the 97 parameterization of the Vashishta potential. Their coexistence line intersect with $T = 0$ at $P \approx 50\,\text{kbar}$, where as our coexistence line intersects at $P = 63.7\,\text{kbar}$. Hence there is roughly a $10\,\text{kbar}$ discrepancy between the two methods. Similar to our coexistence line, theirs also curve in the same direction. It should be noted that their methods are to my understanding approximate, though they do not go into much detail as to how much so. However, they did not locate the α-quartz/coesite coexistence line in the 90 parameterization of the Vashishta potential, which they also tested, listing coesite as sole stable state. Whilst we did not prioritise drawing this coexistence line with Gibbs-Duhem integration, we were able to capture its existence at $P_c = 4.5\,\text{kbar}, T_c = 300\,\text{K}$, which might be indicative that their method of drawing coexistence lines is not as accurate when compared with the free energy methods we have used.

PART III

Conclusion

CHAPTER 8

Summary and Outlook

8.1 Summary

The aim for this thesis has been to evaluate the phase behaviour of crystal polymorphs Silica modeled by the Vashishta potential, and to that end, we have been successful. To compute free energies of the crystal structures we implemented three variations of the Einstein crystal method, evaluating and comparing their performance in detail. We found that the thermodynamic pathway proposed in Vega et al. [34], consisting of free energy perturbation followed by thermodynamic integration was not suitable for Vashishta SiO_2 and instead utilized the Frenkel and Ladd [21] path, in particular the non-equilibrium variant which was observed to be the most computationally efficient. During our analysis, we also uncovered a subtle issue in how the center of mass is usually [36, 39, 48] constrained in Einstein crystal calculations done in LAMMPS, and found a way to remedy this by imposing and additional, explicit constraint on the momenta.

To aid our hunt for coexistence points, we developed an efficient and automated computational workflow for determining free energies with the Einstein crystal method. This enabled us to concurrently and efficiently study three separate parameterizations of the Vashishta potential, namely the 90 [14], 97 [15] and Wang [19] parameterisations. The computational workflow has been documented in detail, and may in principle be applied to study other crystalline beyond Vashishta SiO_2.

We located the coexistence point between α-quartz and β-tridymite for the 1997 and Wang parameterizations of the Vashishta potential, as well as indications of the α/β-quartz transition. The latter of which could not be determined by the Einstein crystal method, as the transition occurs quite readily during the initialization of the crystal reference states. In regards to the 90 parameterizations, we found that in the isobar at 1 bar, only α-quartz was stable, and did not observe any signs of the α/β-quartz transition. Additionally, we were also able to locate the coexistence point between α-quartz and coesite in all three parameterizations of the Vashishta potential.

Moving on, we then implemented a general python package to aid with Gibbs-Duhem integration in LAMMPS. The package was then used to draw out the phase diagrams of the 97 and Wang parameterizations of the Vashishta potential. We observed reasonable qualitative correspondence with the experimental phase diagram in both cases for the quartz/β-tridymite. We did however see signs

of signs of a α-quartz/β-quartz/β-tridymite triple point, that is not present in the true phase diagram. In contrast, the α-quartz/coesite coexistence lines did not exhibit the expected qualitative behaviour in that their gradient in $P(T)$ is opposite of that which is expected in the experimental phase diagram (Fig. 6.1). However, for the 97 parameterization the observed line is reasonably close to the α-quartz/coesite line shown in [72], also computed with the 97 parameterization of the Vashishta potential, but not by the use of free energy methods.

8.1.1 Outlook

As is usually the case in science, there is much yet to be done. Silica exhibits exceptionally rich phase behaviour [1, 4, 5], and evaluation of the thermodynamically stable polymorphs is only scratching the surface. Evaluation of the many metastable states of Vashishta Silica poses an interesting, and likely complicated journey. Furthermore, we have not concerned ourselves with the melting line of Silica, which may evaluated by similar methods to what has been done here. Namely, by computing free energy differences from a suitable reference state and performing Gibbs-Duhem integration with respect to the solid Silica polymorphs.

Here we outline a few additional ways in which this work may be extended:

- We were unable to experiment with the non-equilibrium variant of Gibbs-Duhem integration, so-called dynamic Clauseus-Clapeyron integration [74]. The method draws the entirety of a coexistence line in a single, dynamic, forward and backward sweep. Reminiscent of the way non-equilibrium thermodynamic works. The method was recently implemented in LAMMPS by Cajahuaringa and Antonelli [75] via a LAMMPS extension, however attempts to compile the code were unsuccessful. I suspect this is simply because the LAMMPS developers have made minor changes to the source code since the extension was written, something a also suffered from in my own modification of LAMMPS. Successfully applying this method has the potential speeding up the process of drawing the phase diagrams significantly.

- The free parameters in the Vashishta potential has shown to have a significant effect on the phase-diagram. The workflow defined in this thesis presents good opportunity to apply the genetic optimization algorithm by Chan et al. [76]. Furthermore, one may compute the explicit dependence in the coexistence line of the parameters in the interaction potential by generalized Gibbs-Duhem integration. cf. [34].

- Classifying silica structures was the most manually laborious part of the work presented in this thesis. We performed some experimentation with using SOAP descriptors [77] to quantify the local geometry of Silica, and generate a machine-learning model to classify Silica polymorphs. The model was trained on Einstein crystals with rest positions based of experimental silica structures, and whilst the resulting SOAP descriptors showed good separation in when looking at their principal components in 2D in a PCA analysis, the early performance of the model was poor when tested on Vashishta silica structures, and the venture was not pursued

further. A similar approach was taken in the recently published article by Chung et al. [78], but with different parameters used in the classification model.

Appendices

APPENDIX A

Mass independence of coexistence points

In Vega et al. [34] it is stated without further elaboration that the coexistence properties of a system is independent of the mass of its constituents, instead varying only as a function of the parameters of the interaction potential. In a similar manner, it is stated in [41] that the coexistence properties are independent of the thermal wavelength, and as such may be arbitrarily set to a convenient value. In this section i will briefly explore this idea.

Consider a system governed by the Hamiltonian

$$\mathcal{H} = \sum_{i}^{N} \frac{\boldsymbol{p}_i^2}{2m_i} + U(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \tag{A.1}$$

where U is the potential energy due to a conservative force field, depending only on the positions of the particles, $\{\boldsymbol{r}\}$. In the canonical ensemble, we may then write the partition function as

$$Z(N,V,T) = \int \frac{d\boldsymbol{p}^N d\boldsymbol{r}^N}{h^{3N}} e^{-\beta\mathcal{H}} = \frac{1}{\prod_{i=0}^{N} \Lambda_i^3} \int d\boldsymbol{r}^N \, e^{-\beta U} \tag{A.2}$$

where Λ_i denotes the thermal de Broglie wavelength, due to the integral over the momenta which may be expressed as

$$\Lambda_i = \frac{h}{\sqrt{2\pi m_i k_B T}} \tag{A.3}$$

With this, we may write the Helmholtz Free Energy, F, as a function of the partition function as

$$-\beta F = \ln Z = \ln \int d\boldsymbol{r}^N \; e^{-\beta U} - \sum_{i=0}^{N} \ln \Lambda_i^3 \tag{A.4}$$

In this form, we can see quite clearly that the explicit mass dependence of the free energy is contained in the second term, which is a constant factor along any given isotherm. Hence, changing the masses of one, or multiple particles would only shift the free energy by some constant term. Thus, the location of a coexistence point does not depend on the mass, or indeed the thermal wavelength as a whole provided it is held constant along the isotherm.

APPENDIX B

Vashishta parameters

Here we provide the 1990 [14], 1997 [15] and the Wang [19] parameterizations of the Vashishta potential for SiO_2.

Table B.1: 1990 Parameterization of the Vashishta potential [14]

	Si	O
Z_i [eV]	1.6	-0.8
$r_{1s} = 999$	$r_{4s} = 4.43$	$r_c = 10.0$

Two-Body Interaction

	Si-Si	Si-O	O-O
η_{ij}	11	9	7
H_{ij} [eV Å$^{\eta}$]	0.82023	163.859	743.848
D_{ij} [eV Å^4]	0.0	44.2357	22.1179
W_{ij} [eV Å^6]	0.0	0.0	0.0

Three-Body Interaction

	B_{ijk} [eV]	$\cos(\theta_0)$	C_{ijk}	ξ [Å]	r_0 [Å]
O-Si-O	5.0365	-0.333333	0.0	1.0	2.6
O-Si-O	20.146	-0.777146	0.0	1.0	2.6

Table B.2: 1997 Parameterization of the Vashishta potential [15]

	Si	O
Z_i [eV]	0.7872	-0.3936
$r_{1s} = 4.43$	$r_{4s} = 2.5$	$r_c = 5.5$

Two-Body Interaction

	Si-Si	Si-O	O-O
η_{ij}	11	9	7
H_{ij} [eV Å$^{\eta}$]	0.155	30.923	140.38
D_{ij} [eV Å^4]	0.0	10.701	5.3504
W_{ij} [eV Å^6]	0.0	0.0	0.0

Three-Body Interaction

	B_{ijk} [eV]	$\cos(\theta_0)$	C_{ijk}	ξ [Å]	r_0 [Å]
O-Si-O	4.993	-0.333333	0.0	1.0	2.6
O-Si-O	19.972	-0.805930	0.0	1.0	2.6

Table B.3: Wang Parameterization of the Vashishta potential [19]

	Si	O
Z_i [eV]	1.2	-0.6
$r_{1s} = 4.43$	$r_{4s} = 2.5$	$r_c = 5.5$

Two-Body Interaction

	Si-Si	Si-O	O-O
η_{ij}	11	9	7
H_{ij} [eV Å$^{\eta}$]	0.39246	78.3143	355.5263
D_{ij} [eV Å^4]	0.0	24.882782	12.441391
W_{ij} [eV Å^6]	0.0	0.0	0.0

Three-Body Interaction

	B_{ijk} [eV]	$\cos(\theta_0)$	C_{ijk}	ξ [Å]	r_0 [Å]
O-Si-O	4.993	-0.333313	0.0	1.0	2.6
O-Si-O	19.972	-0.777146	0.0	1.0	2.6

APPENDIX C

LAMMPS Scripts

In this appendix you will find all the LAMMPS scripts necessary to perform the free energy calculations in this thesis. The same files are also hosted on github at https://github.com/nicholaskarlsen/eincrystal/ along with many useful python scripts to aid with Einstein crystal calculations en masse.

C.1 Crystal Initialization

Listing C.1: `in.initbox`

```
# Author: Nicholas Karlsen
#
# Command  line variables:
# temperature
# nsteps_thermostat_relax   - Number of steps to let the crystal initially relax in NVT
# nsteps_npt_eq             - Number of steps to equilibriate in NPT
# nsteps_npt_sample         - Number of steps to sample box shape in NPT
# nsteps_pos_ave            - Number of steps to sample atomic positions in NVT
# nsteps_msd                - Number of steps to sample MSD in NVT

# Prepare the system
boundary        p p p
units           metal
atom_style      atomic
pair_style      vashishta
read_data       init.data
mass            1 15.999400 # O
mass            2 28.085501 # Si
pair_coeff      * * SiO.vashishta O Si

group           O type 1
group           Si type 2

velocity        all create ${temperature} ${SEED}
velocity        all zero linear
velocity        all zero angular

# Keep track of the COM to ensure that its kept constant.
compute         COM all com
variable        xcm equal c_COM[1]
variable        ycm equal c_COM[2]
variable        zcm equal c_COM[3]
compute         Tcm all temp/com

# ---------------------------------------------------------------------------- #
```

```
36 # NVT: Relax the thermostat
37 #
38 # Begin the simulation in NVT in order to let the thermostat relax and avoid an initial
39 # rapid expansion of the lattice that would otherwise occur in NPT.
40 # ------------------------------------------------------------------------------------ #
41 fix             1 all nve
42 fix             2 all langevin ${temperature} ${temperature} $(1000*dt) ${SEED} zero yes
43 fix_modify      2 temp Tcm
44
45 thermo_style    custom step time pe ke press temp density vol v_xcm v_ycm v_zcm lx ly lz
46 log             log_initbox.lammps
47 thermo          10
48 run             ${nsteps_thermostat_relax}
49
50 # ------------------------------------------------------------------------------------ #
51 # NPT: Find the equilibrium shape of the simulation box
52 #
53 # NOTE: Unfix and refix recenter because it needs to be applied AFTER the fix which
54 # contains the integrator such that the recentering is done AFTER the timestep
55 # ------------------------------------------------------------------------------------ #
56 unfix           1
57 fix             1 all nph aniso ${pressure} ${pressure} $(1000*dt) fixedpoint ${xcm} ${ycm} ${zcm}
58 fix_modify      1 temp Tcm
59
60 # Let the system equilibrate
61 run             ${nsteps_npt_eq}
62
63 # Reset the timestep to avoid funny behaviour with fix ave/time
64 reset_timestep  0
65 dump            df all custom 10000 dump.lammpstrj id type xs ys zs
66
67 # System is now equilibriated; start measuring the box sizes to find the equilibrium shape
68 # Average the simulation box shape to find the equilibrium shape for the given pressure
69 variable        lx equal lx
70 variable        ly equal ly
71 variable        lz equal lz
72 fix             avelx all ave/time 1 ${nsteps_npt_sample} ${nsteps_npt_sample} v_lx
73 fix             avely all ave/time 1 ${nsteps_npt_sample} ${nsteps_npt_sample} v_ly
74 fix             avelz all ave/time 1 ${nsteps_npt_sample} ${nsteps_npt_sample} v_lz
75
76 run             ${nsteps_npt_sample}
77
78 # Change the size of the box to its equilibrium shape
79 change_box all x final 0 $(f_avelx) y final 0 $(f_avely) z final 0 $(f_avelz) remap
80
81 unfix avelx
82 unfix avely
83 unfix avelz
84
85 # ------------------------------------------------------------------------------------ #
86 # NVT: Average the positions of the atoms
87 # ------------------------------------------------------------------------------------ #
88 unfix           1
89 fix             1 all nve
90 undump df
91 reset_timestep  0
92
93 # Record the average position of each atom
94 compute         xu all property/atom xu
95 compute         yu all property/atom yu
96 compute         zu all property/atom zu
97
```

```
fix             avex all ave/atom 1 ${nsteps_pos_ave} ${nsteps_pos_ave} c_xu
fix             avey all ave/atom 1 ${nsteps_pos_ave} ${nsteps_pos_ave} c_yu
fix             avez all ave/atom 1 ${nsteps_pos_ave} ${nsteps_pos_ave} c_zu

run             ${nsteps_pos_ave}

# Set the position of each atom to the average position (equilibrium position)
variable        avex atom f_avex
variable        avey atom f_avey
variable        avez atom f_avez
set             atom * x v_avex y v_avey z v_avez

write_data      optim_config.data

unfix           avex
unfix           avey
unfix           avez

# ---------------------------------------------------------------------------- #
# NVT: Sample the mean-squared displacement
#
# Used to compute the spring constant for an ideal Einstein crystal that replicates
# the same MSD, given by: k = 3 k_B T / (2 * <dr^2>)
# ---------------------------------------------------------------------------- #

compute         MSD_O O msd
compute         MSD_Si Si msd

thermo_style    custom step time c_MSD_O[4] c_MSD_Si[4]
log             log_msd.lammps
run             ${nsteps_msd}
```

C.2 Thermodynamic Integration

Listing C.2: `in.ti`

```
# Author: Nicholas Karlsen
#
# Performs Thermodynamic integration from an Einstein crystal to Vashishta SiO2
#
# IMPORTANT: This file requires a modified version of LAMMPS that has a "scale" parameter
# in the Vashishta pair-style such that fix_adapt may be used.
# see the changes made in the fork: https://github.com/nicholaskarlsen/lammps-scale-vashishta
#
# Command line variables:
# T             Temperature
# N_STEPS       Number of timesteps
# lambda        Integration variable/scaling parameter
# index         Index corresponding to the integration step (ie I = \sum_i^N w_i * f(x_i))

variable        lambda  equal ${lamb}
variable        k_O     equal ((1-v_lambda)*${spring_const_O})
variable        k_Si    equal ((1-v_lambda)*${spring_const_Si})

# Prepare the system
dimension       3
boundary        p p p
units           metal
atom_style      atomic
pair_style      vashishta
```

```
25
26 read_data       optim_config.data
27
28 mass            1 15.999400 # O
29 mass            2 28.085501 # Si
30
31 group           1 type 1
32 group           2 type 2
33
34 pair_coeff      * * SiO.vashishta O Si
35
36 velocity        all create ${temperature} ${seed}
37 velocity        all zero linear
38 velocity        all zero angular
39
40 # Set NVT conditions
41 fix             fix_nve all nve
42 fix             fix_langevin all langevin ${temperature} ${temperature} $(100*dt) ${seed} zero yes
43 fix             fix_adapt all adapt 0 pair vashishta scale 1 1 v_lambda
44 if "${k_O} != 0" then &
45     "fix            fix_spring_O 1 spring/self ${k_O}" &
46     "fix            fix_spring_Si 2 spring/self ${k_Si}"
47 # Recenter COM after every step
48 fix             fix_momentum all momentum 1 linear 1 1 1 rescale
49 fix             fix_recenter all recenter INIT INIT INIT
50
51 variable        UE_O equal (f_fix_spring_O)
52 variable        UE_Si equal (f_fix_spring_Si)
53
54 compute         MSD_O 1 msd
55 compute         MSD_Si 2 msd
56
57 #dump            df all custom 1000 dump_${index}.lammpstrj id type element x y z
58 #dump_modify     df element O Si
59 thermo          10
60 # Note: pe corresponds to U_C
61 thermo_style    custom step time ke pe v_UE_O v_UE_Si c_MSD_O[4] c_MSD_Si[4]
62 log             log_ti_${index}.lammps
63 run             ${nsteps}
```

C.3 Non-Equilibrium Thermodynamic Integration

Listing C.3: `in.neq_ti`

```
1  # Author: Nicholas Karlsen
2  #
3  # Performs Non-Equilibrium Thermodynamic integration from an Einstein crystal to
4  # Vashishta SiO2 along the homotopy:
5  #
6  #       U(lambda) = (1 - lambda) * U_Vashishta +  lambda * U_Einstein
7  #
8  # Thus the simulation is follows the pathway:
9  #   -> Equilibriate pure Vashishta crystal
10 #   -> Integrate from Vashishta to Einstein
11 #   -> Equilibriate pure Einstein crystal
12 #   -> Integrate from Einstein to Vashishta
13 #
14 # Command line variables:
15 # seed                  Seed used in the random number generator
16 # nsteps_eq             Number of steps to equilibriate during the before the fw and bw processes
```

```
# nsteps_ti             Number of steps to perform the backward and forward processes
# spring_const_O        Spring constant of the Oxygen atoms in the ideal einstein crystal
# spring_const_Si       Spring constant of the Silicon atoms in the ideal einstein crystal
# temp                  Temperature imposed by the Langevin thermostat

log             neq_ti_${seed}.lammps

# Prepare the system
dimension       3
boundary        p p p
units           metal
atom_style      atomic
pair_style      vashishta
read_data       optim_config.data
mass            1 15.999400 # O
mass            2 28.085501 # Si
group           O type 1
group           Si type 2
pair_coeff      * * SiO.vashishta O Si
velocity        all create ${temp} ${seed} mom yes rot yes dist gaussian

fix             NVE all nve
fix             ti_O O ti/spring ${spring_const_O} ${nsteps_ti} ${nsteps_eq} function 2
fix             ti_Si Si ti/spring ${spring_const_Si} ${nsteps_ti} ${nsteps_eq} function 2
fix             thermostat all langevin ${temp} ${temp} $(100*dt) ${seed} zero yes

compute         Tcm all temp/com
fix_modify      thermostat temp Tcm

# Energies
variable        U_O equal f_ti_O
variable        U_Si equal f_ti_Si
variable        dU equal (pe-v_U_O-v_U_Si)

# Integration variable
variable        lambda_O equal f_ti_O[1]
variable        lambda_Si equal f_ti_Si[1]
variable        dlambda equal f_ti_O[2]

# Fix the center of mass by manually removing any drift of the COM after integration
fix             fmomentum all momentum 1 linear 1 1 1 rescale
fix             frecenter all recenter INIT INIT INIT

dump            df all custom 5000 dump_neq_ti_${seed}.lammpstrj id type x y z
thermo_style    custom step time pe ke v_dU v_U_O v_U_Si v_dlambda v_lambda_O v_lambda_Si
thermo          10

run ${nsteps_eq}    # Equilibriate Vashishta
run ${nsteps_ti}    # Vashishta -> Einstein
run ${nsteps_eq}    # Equilibriate Einstein
run ${nsteps_ti}    # Einstein -> Vashishta
```

C.4 Free Energy Perturbation (Vega Path)

```
# Run simulation with atoms governed by harmonic potential centered on the initial positions
dimension       3
boundary        p p p
units           metal
atom_style      atomic
pair_style      none
```

```
7
8  read_data       optim_config.data
9
10 mass            1 20 # O
11 mass            2 20 # Si
12
13 velocity        all create ${temp} ${seed}
14 velocity        all zero linear
15 velocity        all zero angular
16
17 # Set NVT conditions
18 fix             fix_langevin all langevin ${temp} ${temp} $(100*dt) ${seed}
19 fix             fix_nve all nve
20
21 # Harmonic potential located at init positions of each atom
22 fix                 fix_spring all spring/self ${k}
23
24 # Adds spring potential to the PotEng computation
25 fix_modify      fix_spring energy yes
26
27 # Fix the center of mass
28 fix             fix_recenter all recenter INIT INIT INIT
29
30 dump            dumpfile all custom 10 dump_id_${index}.lammpstrj id type x y z
31 run             ${N_STEPS}
32
33 clear
34
35 # rerun the dump file to sample the PE in the Vashishta potential
36 dimension       3
37 units           metal
38 atom_style      atomic
39 pair_style      vashishta
40
41 read_data       optim_config.data
42
43 pair_coeff      * * SiO.vashishta O Si
44
45 thermo          1
46 thermo_style    custom step time pe
47 log             log_real_${index}.lammps
48 rerun           dump_id_${index}.lammpstrj dump x y z
```

C.5 Thermodynamic Integration (Vega Path)

```
1  # Author: Nicholas Karlsen
2  #
3  # Performs Thermodynamic integration from an an interacting Einstein crystal
4  # to a crystal by integrating along the homotopy:
5  #
6  #       U(lambda) = (1 - lambda) * (U_Vashishta + U_Einstein) +  lambda * U_Vashishta
7  #
8  #   -> Integrate from interacting Einstein crystal to Vashishta crystal
9  #
10 # Command line variables:
11 # seed                  Seed used in the random number generator
12 # temperature           Temperature imposed by the Langevin thermostat
13 # nsteps                Number of simulation timesteps
14 # k                     spring constant
15 # idx                   Index of the computation (useful in post processing)
```

```
16
17 dimension       3
18 boundary        p p p
19 units           metal
20 atom_style      atomic
21 pair_style      vashishta
22
23 read_data       optim_config.data
24 mass            1 20 # O
25 mass            2 20 # S
26
27 pair_coeff      * * SiO.vashishta O Si
28 velocity        all create ${temp} ${seed}
29
30 fix             fix_langevin all langevin ${temp} ${temp} $(100*dt) ${seed} zero yes
31 fix             fix_nve all nve
32
33 if "${k} != 0" then &
34     "fix            fix_spring all spring/self ${k}" &
35     "fix_modify     fix_spring energy yes"
36
37 fix             fix_recenter all recenter INIT INIT INIT
38
39 compute         MSD all msd
40
41 dump            dumpfile all custom 100 dump_${idx}.lammpstrj id type x y z
42 thermo          1
43 thermo_style    custom step time ke pe atoms c_MSD[4]
44 log             log_${idx}.lammps
45 run             ${nsteps}
```

Bibliography

[1] Cohen, R. E. 'First-principles theory of crystalline SiO 2'. In: *Reviews in Mineralogy and Geochemistry* vol. 29, no. 1 (Jan. 1994), pp. 369–402.

[2] Beall, G. H. 'Industrial applications of silica'. In: *Reviews in Mineralogy and Geochemistry* vol. 29, no. 1 (Jan. 1994), pp. 469–505.

[3] Clarke, F. W. *Analysis of rocks and minerals from the laboratory of the United States Geological Survey, 1880 to 1908*. Tech. rep. 419. Publication Title: Bulletin. Govt. Print. Off., 1910.

[4] Heaney, P. J. 'Structure and chemistry of the low-pressure silica polymorphs'. In: *Reviews in Mineralogy and Geochemistry* vol. 29, no. 1 (Jan. 1994), pp. 1–40.

[5] Hemley, R. J., Prewitt, C. T. and Kingma, K. J. 'High-pressure behavior of silica'. In: *Reviews in Mineralogy and Geochemistry* vol. 29, no. 1 (Jan. 1994), pp. 41–81.

[6] Broughton, J. Q. et al. 'Concurrent coupling of length scales: Methodology and application'. In: *Physical Review B* vol. 60, no. 4 (July 1999). Publisher: American Physical Society, pp. 2391–2403.

[7] Hu, W. et al. *Extreme-Scale Density Functional Theory High Performance Computing of DGDFT for Tens of Thousands of Atoms using Millions of Cores on Sunway TaihuLight*. arXiv:2003.00407 [physics]. Apr. 2020.

[8] VandeVondele, J., Borštnik, U. and Hutter, J. 'Linear Scaling Self-Consistent Field Calculations with Millions of Atoms in the Condensed Phase'. In: *Journal of Chemical Theory and Computation* vol. 8, no. 10 (Oct. 2012). Publisher: American Chemical Society, pp. 3565–3573.

[9] Dolino, G. and Vallade, M. 'Lattice dynamical behavior of anhydrous silica'. In: *Reviews in Mineralogy and Geochemistry* vol. 29, no. 1 (Jan. 1994), pp. 403–431.

[10] Gibbs, G. V., Downs, J. W. and Boisen, M. B. 'The elusive SiO bond'. In: *Reviews in Mineralogy and Geochemistry* vol. 29, no. 1 (Jan. 1994), pp. 331–368.

[11] Beest B.W.H., v., Kramer, G. and Santen R.A., v. 'Force fields for silicas and aluminophosphates based on ab initio calculations'. In: *Physical Review Letters* vol. 64, no. 16 (1990), pp. 1955–1958.

[12] Tsuneyuki, S. et al. 'First-Principles Interatomic Potential of Silica Applied to Molecular Dynamics'. In: *Physical Review Letters* vol. 61, no. 7 (Aug. 1988). Publisher: American Physical Society, pp. 869–872.

[13] Duin, A. C. T. van et al. 'ReaxFFSiO Reactive Force Field for Silicon and Silicon Oxide Systems'. In: *The Journal of Physical Chemistry A* vol. 107, no. 19 (May 2003). Publisher: American Chemical Society, pp. 3803–3811.

[14] Vashishta, P. et al. 'Interaction potential for SiO2: A molecular-dynamics study of structural correlations'. In: *Physical Review B* vol. 41, no. 17 (June 1990). Publisher: American Physical Society, pp. 12197–12209.

[15] Broughton, J. Q. et al. 'Direct atomistic simulation of quartz crystal oscillators: Bulk properties and nanoscale devices'. In: *Physical Review B* vol. 56, no. 2 (July 1997). Publisher: American Physical Society, pp. 611–618.

[16] Pham-Ba, S. and Molinari, J.-F. 'Creation and evolution of roughness on silica under unlubricated wear'. en. In: *Wear* vol. 472-473 (May 2021), p. 203648.

[17] Rountree, C. L. et al. 'A unified study of crack propagation in amorphous silica: Using experiments and simulations'. en. In: *Journal of Alloys and Compounds*. Proceedings of the 12th International Symposium on Metastable and Nano-Materials (ISMANAM-2005) vol. 434-435 (May 2007), pp. 60–63.

[18] Johnson, S. E. et al. 'The quartz phase transition: Does it drive damage and reaction in continental crust?' en. In: *Earth and Planetary Science Letters* vol. 553 (Jan. 2021), p. 116622.

[19] Wang, W. et al. 'Interatomic potentials for molecular dynamics simulations of hydrolysis and stress corrosion cracking of silica glass'. In: *Unpublished* ().

[20] Eike, D. M. and Maginn, E. J. 'Atomistic simulation of solid-liquid coexistence for molecular systems: Application to triazole and benzene'. In: *The Journal of Chemical Physics* vol. 124, no. 16 (Apr. 2006). Publisher: American Institute of Physics, p. 164503.

[21] Frenkel, D. and 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'. In: *The Journal of Chemical Physics* vol. 81, no. 7 (Oct. 1984). Publisher: American Institute of Physics, pp. 3188–3193.

[22] Kofke, D. A. 'Direct evaluation of phase coexistence by molecular simulation via integration along the saturation line'. In: *The Journal of Chemical Physics* vol. 98, no. 5 (Mar. 1993). Publisher: American Institute of Physics, pp. 4149–4162.

[23] Swendsen, R. H. *An Introduction to Statistical Mechanics and Thermodynamics*. 2nd ed. Oxford University Press, 2020.

[24] Allen, M. P. and Tildesley, D. J. *Computer Simulation of Liquids*. eng. 2nd ed. Oxford: Oxford University Press, 2017.

[25] Jones, J. E. and Chapman, S. 'On the determination of molecular fields.—I. From the variation of the viscosity of a gas with temperature'. In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* vol. 106, no. 738 (Oct. 1924). Publisher: Royal Society, pp. 441–462.

[26] Buckingham, R. A. and Lennard-Jones, J. E. 'The classical equation of state of gaseous helium, neon and argon'. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* vol. 168, no. 933 (Oct. 1938). Publisher: Royal Society, pp. 264–283.

[27] Ryckaert, J.-p., Ciccotti, G. and Berendsen, H. J. C. 'Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes'. In: *J. Comput. Phys* (1977), pp. 327–341.

[28] Berendsen, H. J. C. et al. 'Molecular dynamics with coupling to an external bath'. In: *The Journal of Chemical Physics* vol. 81, no. 8 (Oct. 1984). Publisher: American Institute of Physics, pp. 3684–3690.

[29] Braun, E., Moosavi, S. M. and Smit, B. 'Anomalous Effects of Velocity Rescaling Algorithms: The Flying Ice Cube Effect Revisited'. In: *Journal of Chemical Theory and Computation* vol. 14, no. 10 (Oct. 2018). Publisher: American Chemical Society, pp. 5262–5272.

[30] Bussi, G., Donadio, D. and Parrinello, M. 'Canonical sampling through velocity rescaling'. In: *The Journal of Chemical Physics* vol. 126, no. 1 (Jan. 2007). Publisher: American Institute of Physics, p. 014101.

[31] Chipot, C. and Pohorille, A. *Free Energy Calculations: Theory and Applications in Chemistry and Biology*. 1st ed. Springer Series in chemical physics. Berlin, Heidelberg: Springer, 2007.

[32] Lelievre, T., Stoltz, G. and Rousset, M. *Free Energy Computations: A Mathematical Perspective*. English. 1st ed. Imperial college press, 2010.

[33] Frenkel, D. and Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*. en. Ed. by Frenkel, D. and Smit, B. 2nd ed. San Diego: Academic Press, Jan. 2002.

[34] Vega, C. et al. 'Determination of phase diagrams via computer simulation: methodology and applications to water, electrolytes and proteins'. en. In: *Journal of Physics: Condensed Matter* vol. 20, no. 15 (Mar. 2008). Publisher: IOP Publishing, p. 153101.

[35] Zwanzig, R. W. 'High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases'. In: *The Journal of Chemical Physics* vol. 22, no. 8 (Aug. 1954). Publisher: American Institute of Physics, pp. 1420–1426.

[36] Freitas, R., Asta, M. and Koning, M. de. 'Nonequilibrium free-energy calculation of solids using LAMMPS'. en. In: *Computational Materials Science* vol. 112 (Feb. 2016), pp. 333–341.

[37] Koning, M. de and Antonelli, A. 'Einstein crystal as a reference system in free energy estimation using adiabatic switching'. In: *Physical Review E* vol. 53, no. 1 (Jan. 1996). Publisher: American Physical Society, pp. 465–474.

[38] Polson, J. M. et al. 'Finite-size corrections to the free energies of crystalline solids'. In: *The Journal of Chemical Physics* vol. 112, no. 12 (Mar. 2000). Publisher: American Institute of Physics, pp. 5339–5342.

[39] Khanna, V. et al. 'Free energies of crystals computed using Einstein crystal with fixed center of mass and differing spring constants'. In: *The Journal of Chemical Physics* vol. 154, no. 16 (Apr. 2021). Publisher: American Institute of Physics, p. 164509.

[40] Vega, C. and Noya, E. G. 'Revisiting the Frenkel-Ladd method to compute the free energy of solids: The Einstein molecule approach'. In: *The Journal of Chemical Physics* vol. 127, no. 15 (Oct. 2007). Publisher: American Institute of Physics, p. 154113.

[41] Reddy Addula, R. K., Veesam, S. K. and Punnathanam, S. N. 'Review of the Frenkel–Ladd technique for computing free energies of crystalline solids'. In: *Molecular Simulation* vol. 47, no. 10-11 (July 2021). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/08927022.2020.1775221, pp. 824–830.

[42] Aragones, J. L., Valeriani, C. and Vega, C. 'Note: Free energy calculations for atomic solids through the Einstein crystal/molecule methodology using GROMACS and LAMMPS'. In: *The Journal of Chemical Physics* vol. 137, no. 14 (Oct. 2012). Publisher: American Institute of Physics, p. 146101.

[43] Menon, S. et al. 'Automated free-energy calculation from atomistic simulations'. In: *Physical Review Materials* vol. 5, no. 10 (Oct. 2021). Publisher: American Physical Society, p. 103801.

[44] Nosé, S. 'A molecular dynamics method for simulations in the canonical ensemble'. In: *Molecular Physics* vol. 52, no. 2 (June 1984). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00268978400101201, pp. 255–268.

[45] Hoover, W. G. 'Canonical dynamics: Equilibrium phase-space distributions'. In: *Physical Review A* vol. 31, no. 3 (Mar. 1985). Publisher: American Physical Society, pp. 1695–1697.

[46] Parrinello, M. and Rahman, A. 'Polymorphic transitions in single crystals: A new molecular dynamics method'. In: *Journal of Applied Physics* vol. 52, no. 12 (Dec. 1981). Publisher: American Institute of Physics, pp. 7182–7190.

[47] Shinoda, W., Shiga, M. and Mikami, M. 'Rapid estimation of elastic constants by molecular dynamics simulation under constant stress'. In: *Physical Review B* vol. 69, no. 13 (Apr. 2004). Publisher: American Physical Society, p. 134103.

[48] Noya, E. G., Conde, M. M. and Vega, C. 'Computing the free energy of molecular solids by the Einstein molecule approach: Ices XIII and XIV, hard-dumbbells and a patchy model of proteins'. In: *The Journal of Chemical Physics* vol. 129, no. 10 (Sept. 2008). Publisher: American Institute of Physics, p. 104704.

[49] Schneider, T. and Stoll, E. 'Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions'. In: *Physical Review B* vol. 17, no. 3 (Feb. 1978). Publisher: American Physical Society, pp. 1302–1322.

[50] Yang, W., Bitetti-Putzer, R. and Karplus, M. 'Free energy simulations: Use of reverse cumulative averaging to determine the equilibrated region and the time required for convergence'. In: *The Journal of Chemical Physics* vol. 120, no. 6 (Feb. 2004). Publisher: American Institute of Physics, pp. 2618–2628.

[51] Chodera, J. D. 'A Simple Method for Automated Equilibration Detection in Molecular Simulations'. In: *Journal of Chemical Theory and Computation* vol. 12, no. 4 (Apr. 2016). Publisher: American Chemical Society, pp. 1799–1805.

[52] Smith, E. B. and Wells, B. H. 'Estimating errors in molecular simulation calculations'. In: *Molecular Physics* vol. 52, no. 3 (June 1984). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00268978400101481, pp. 701–704.

[53] Nicholls, A. 'Confidence limits, error bars and method comparison in molecular modeling. Part 1: The calculation of confidence intervals'. en. In: *Journal of Computer-Aided Molecular Design* vol. 28, no. 9 (Sept. 2014), pp. 887–918.

[54] Grossfield, A. et al. 'Best Practices for Quantification of Uncertainty and Sampling Quality in Molecular Simulations [Article v1.0]'. In: *Living journal of computational molecular science* vol. 1, no. 1 (2018), p. 5067.

[55] Flyvbjerg, H. and Petersen, H. G. 'Error estimates on averages of correlated data'. In: *The Journal of Chemical Physics* vol. 91, no. 1 (July 1989). Publisher: American Institute of Physics, pp. 461–466.

[56] Koning, M. de and Antonelli, A. 'Adiabatic switching applied to realistic crystalline solids: Vacancy-formation free energy in copper'. In: *Physical Review B* vol. 55, no. 2 (Jan. 1997). Publisher: American Physical Society, pp. 735–744.

[57] Sveinsson, H. A. *molecular-builder*. Language: eng. Nov. 2021.

[58] Aitken, A. C. 'IV.—On Least Squares and Linear Combination of Observations'. en. In: *Proceedings of the Royal Society of Edinburgh* vol. 55 (1936). Publisher: Royal Society of Edinburgh Scotland Foundation, pp. 42–48.

[59] Kofke, D. A. 'On the sampling requirements for exponential-work free-energy calculations'. In: *Molecular Physics* vol. 104, no. 22-24 (Nov. 2006). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00268970601074421, pp. 3701–3708.

[60] Klein, C. and Hurlbut, C. S. *Manual of Mineralogy*. 21st ed. Wiley, 1993.

[61] Saika-Voivod, I. et al. 'Phase diagram of silica from computer simulation'. In: *Physical Review E* vol. 70, no. 6 (Dec. 2004). Publisher: American Physical Society, p. 061507.

[62] Nakano, A., Kalia, R. K. and Vashishta, P. 'First sharp diffraction peak and intermediate-range order in amorphous silica: finite-size effects in molecular dynamics simulations'. en. In: *Journal of Non-Crystalline Solids* vol. 171, no. 2 (Aug. 1994), pp. 157–163.

[63] Mosesman, M. A. and Pitzer, K. S. 'Thermodynamic Properties of the Crystalline Forms of Silica'. In: *Journal of the American Chemical Society* vol. 63, no. 9 (Sept. 1941). Publisher: American Chemical Society, pp. 2348–2356.

[64] Kihara, K., Matsumoto, T. and Imamura, M. 'Structural change of orthorhombic-I tridymite with temperature: A study based on second-order thermal-vibrational parameters'. de. In: *Zeitschrift für Kristallographie - Crystalline Materials* vol. 177, no. 1-4 (Dec. 1986). Publisher: De Gruyter (O), pp. 27–38.

[65] Wyckoff, R. W. G. 'IX. Die Kristallstruktur von -Cristobalit SiO2 (bei hohen Temperaturen stabile Form)'. de. In: *Zeitschrift für Kristallographie - Crystalline Materials* vol. 62, no. 1-6 (Dec. 1925). Publisher: De Gruyter (O), pp. 189–200.

[66] Steinhardt, P. J., Nelson, D. R. and Ronchetti, M. 'Bond-orientational order in liquids and glasses'. In: *Physical Review B* vol. 28, no. 2 (July 1983). Publisher: American Physical Society, pp. 784–805.

[67] Nieuwenkamp, W. 'Über die Struktur von Hoch-Cristobalit'. de. In: *Zeitschrift für Kristallographie - Crystalline Materials* vol. 96, no. 1-6 (Nov. 1937). Publisher: De Gruyter (O), pp. 454–458.

[68] Xie, X. et al. 'Study on - quartz phase transition and its effect on dielectric properties'. In: *Journal of Applied Physics* vol. 111, no. 10 (May 2012). Publisher: American Institute of Physics, p. 104116.

[69] Kihara, K. 'An X-ray study of the temperature dependence of the quartz structure'. In: *European Journal of Mineralogy* vol. 2, no. 1 (Jan. 1990), pp. 63–77.

[70] Press, W. H. et al. *Numerical Recipes: The Art of Scientific Computing*. 3rd ed. Cambridge University Press, 2007.

[71] Karlsen, N. *github.com/nicholaskarlsen/gibbs_duhem_integration*.

[72] Erhard, L. C. et al. 'A machine-learned interatomic potential for silica and its relation to empirical models'. en. In: *npj Computational Materials* vol. 8, no. 1 (Apr. 2022). Number: 1 Publisher: Nature Publishing Group, pp. 1–12.

[73] Togo, A. et al. 'First-principles phonon calculations of thermal expansion in ${\text{Ti}}_{3}{\text{SiC}}_{2}$, ${\text{Ti}}_{3}{\text{AlC}}_{2}$, and ${\text{Ti}}_{3}{\text{GeC}}_{2}$'. In: *Physical Review B* vol. 81, no. 17 (May 2010). Publisher: American Physical Society, p. 174301.

[74] Koning, M. de, Antonelli, A. and Yip, S. 'Single-simulation determination of phase boundaries: A dynamic Clausius–Clapeyron integration method'. In: *The Journal of Chemical Physics* vol. 115, no. 24 (Dec. 2001). Publisher: American Institute of Physics, pp. 11025–11035.

[75] Cajahuaringa, S. and Antonelli, A. 'Non-equilibrium free-energy calculation of phase-boundaries using LAMMPS'. en. In: *Computational Materials Science* vol. 207 (May 2022), p. 111275.

[76] Chan, H. (et al. 'Machine Learning Classical Interatomic Potentials for Molecular Dynamics from First-Principles Training Data'. English. In: *Journal of Physical Chemistry. C* vol. 123, no. 12 (Jan. 2019). Institution: Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States). National Energy Research Scientific Computing Center (NERSC) Publisher: American Chemical Society.

[77] Bartók, A. P., Kondor, R. and Csányi, G. 'On representing chemical environments'. In: *Physical Review B* vol. 87, no. 18 (May 2013). Publisher: American Physical Society, p. 184115.

[78] Chung, H. W. et al. 'Data-centric framework for crystal structure identification in atomistic simulations using machine learning'. In: *Physical Review Materials* vol. 6, no. 4 (Apr. 2022). Publisher: American Physical Society, p. 043801.