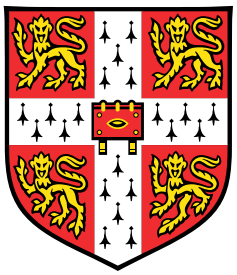
Fitting Interatomic Potentials to Reproduce Phase Transitions



Martin Schlegel

Department of Engineering
University of Cambridge

This thesis is submitted for the degree of
Doctor of Philosophy

King's College

March 2019

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee. Its length does not exceed $65,000$ words and it does not contain more than 150 figures as required by the Engineering Degree Committee.

Fitting Interatomic Potentials to Reproduce Phase Transitions

Martin Schlegel

Abstract

Due to the high computational demand of quantum mechanical simulations, researchers still rely on models for atomic interactions to simulate a large number of atoms. These so called "interatomic potentials" take a variety of forms ranging from the simple Lennard-Jones potential to machine learning potentials. They are usually fitted for specific structures as well as pressure and temperature realms and do not necessarily reproduce the full experimentally known solid and liquid phase diagram. There have been limited trials of fitting the parameters of interatomic potentials to phase transitions. In this work, I present improvements upon two already existing potentials from the literature – a CuAu potential and a Ti potential. These potentials are modified so that they reproduce experimental phase transitions better than before. For the CuAu potential, I fit the eutectic melting curve using my own implementation of the Nelder-Mead algorithm. The Ti potential is modified to predict a triple point which is in qualitative agreement with experimental predictions. To calculate phase transitions, I use the nested sampling algorithm which was already established in the literature. The Ti nested sampling results are verified with thermodynamic integration. My Nelder-Mead implementation allows for semi-automatic submission and analysis of the nested sampling runs. Both potentials are improved considerably. The new CuAu potential now accurately reproduces the melting curve for the alloy compositions used in the fitting procedure. The new Ti potential now – as opposed to before the modifications – predicts a triple point which is in qualitative accordance with experimental predictions. My work demonstrates how the transferability of potentials can be improved by just slight modifications to the original parameters using experimental data of the phase diagram. The Nelder-Mead implementation used for the CuAu potential could be used to improve other potentials. To fit more complex systems, the way the objective function is calculated could be upgraded in future.

Acknowledgements

I thank my supervisor Gábor Csányi for fruitful discussions and his support. Many thanks are due to all current and former members, affiliates and collaborators of the Csányi group with whom I had the chance to work for fruitful discussions and comments. In particular, I thank Noam Bernstein, for fruitful discussions, especially on the subject of nested sampling, and for his advice on nested sampling parameters, Lívia Bartók-Pártay for fruitful discussions, especially on the subject of nested sampling, Robert Baldock for fruitful discussions, especially on the subject of nested sampling, Volker Deringer for helpful discussions and comments and Thomas Daff for his advice on setting up the DFT simulations. I also thank Michael Finnis for a helpful comment regarding the zero-point energy, my colleagues Felix Mocanu and Jamie Wynn for useful discussions as well as my advisor Vikram Deshpande for helpful comments, mostly of a more general kind regarding the course of my PhD. Many thanks are due to Noam Bernstein, Albert Bartók-Pártay and Lívia Bartók-Pártay for their analysis scripts and programs and to Albert Bartók-Pártay for his input script for thermodynamic integration. Additionally, I thank Mark Johnson for giving me access to his phonon and free energy calculation program and fruitful discussions. I thank the EPSRC for funding my studies and King's College for additional funding. This work used the ARCHER UK National Supercomputing Service (http://www.archer.ac.uk) and the UK Research Data Facility (http://www.archer.ac.uk/documentation/rdf-guide). Additionally, I thank David Nicholson and Jamie Wynn for proofreading parts of this thesis. Finally, I would like to thank my friends and my family for enriching my life and especially my parents for their continued support.

Table of contents

List of figures

List of tables

Chapter 1

Introduction and approach

1.1 Introduction

Simulations have opened a new avenue of inquiry besides the traditional experimental and theoretical approaches. We can split the computational investigation into continuum and atomistic approaches. For the former we ignore that materials are made of atoms and treat the material as if it were continuous. For the latter, one simulates the materials as atoms or molecules.

The atomic interactions can in principle be calculated quantum mechanically. However, the more precise one wants to treat the system, the higher the computational costs become. This means that some methods are too computationally expensive for systems with a large number of atoms. For example, standard density functional theory is limited to ca 1000 atoms [1]. While improvements to over $16,000$ thousand atoms are possible, calculations with ca. $8,000$ atoms can already require a parallelization on over $60,000$ processors [1]. Hence, so called "interatomic potentials" are used instead if fast evaluations of energies are required. These are in principle mathematical functions which describe the potential energy of a system in dependence of the atomic positions of the system. They range from simple pair interaction models like the Lennard-Jones model [2, 3] over more complex models like the embedded atom method (EAM) [4, 5] and the modified embedded-atom method MEAM [6, 7] – which consider many-body interaction – to models which also consider the behaviour of interatomic bonds as for example the Tersoff potential [8] and ReaxFF [9]. Furthermore, a recent addition are machine learning potentials [10–12]. The interatomic potentials have parameters which need to be fitted. In this process, experimental or quantum mechanical data is usually used. It is far from trivial to create an acceptable interatomic potential [13]. Indeed, interatomic potentials are usually only fitted for a limited range of applications. Additionally, direct use of information which is gained from simulations or experiments at temperatures higher

than absolute zero as done for example in Ref. [14] and Ref. [15] is usually very limited [16–19]. This means that the potentials are only transferable to a limited extent failing to predict the correct crystal structures in some pressure and temperature ranges as demonstrated for example in Ref. [20] and Ref. [21]. The question is now, how one can improve the predictions of interatomic potentials for situations where they do not predict the correct phases and transition temperatures as found in experiments. A simple answer is to include the phase transition temperatures and pressures of interest directly in the fitting procedure. However, this requires the prediction of phase diagrams based on the intermediate potentials created in each step of the fitting procedure. The fact that methods for computing the stable phases were too computationally expensive and inconvenient was an impediment to this line of research [18, 17]. Though not yet explicitly including phase transition temperatures in a fitting routine, there have been recent efforts which consider more thermodynamic behaviour of the structures of interest while optimising potentials [17, 18, 22, 23, 16, 19]. Building on an interatomic titanium potential of Ref. [24], Ackland [16] considers, apart from how the thermodynamic behaviour of electrons might influence an interatomic potential, also modifying a parameter which is responsible for the behaviour of the potential when atoms are very close to each other. These modifications do not change the fact that the hcp phase is predicted by the potential to be the ground state [16]. However, the modifications significantly influence at which temperature the hcp and bcc phases melt [16]. Ackland hence proposes that modifying similar parts in already fitted interatomic potential functions may make these potentials more reliable [16]. Overall, I conclude that if one can find a suitable algorithm, optimising interatomic potentials from the literature to parts of the phase diagram is a promising endeavour. Indeed, recent progress in applying nested sampling [25, 26] to the constant pressure ensemble [27, 20, 28] may offer exactly the method needed for this.
As mentioned before, to fit to data of experimental phase transitions, one needs a suitable method of predicting phase transitions of interatomic potentials [18, 17]. It is common to present experimental data on phase transitions at constant pressure or for pressure-temperature phase diagrams, hence the simulations need to be conducted in such a manner that one can compare to these results. To calculate phase transitions, one has the choice between "blind search" and "informed search" methods [27]. For "informed search" methods, one needs to be aware of which phases are competing against each other at the temperatures and pressures of interest before starting the computations. This is not necessary for "blind search" methods which predict the stable phases and the transition temperatures and pressures themselves. In general, "informed search" methods require much less computational resources than "blind search" methods. However, if one fails to include a relevant phase in one's calculations the "informed search" method will lead to a wrong result. Given that I want to optimise

interatomic potentials to experimental phase transition data, it is crucial to avoid this with the potential parameters created during the fitting. Hence, if a "blind search" algorithm is available which is not too computationally expensive, this will be the best choice.
To fit interatomic potentials to experimental phase diagrams, one may need a high amount of phase diagram calculations based on potentials with different parameters. The method used for calculating the phase diagrams must therefore be as efficient and fast as possible. For calculating *P*-*T* phase diagrams, nested sampling [25, 26], Wang-Landau sampling [29] and parallel tempering [30, 31] are the "blind search" methods of interest [27, 20]. Of these three, nested sampling requires the least amount of computational resources [32, 27, 20]. Additionally, the calculation of phase diagrams is very convenient and phases can be identified in the analysis after the simulation has been conducted [27, 20, 28, 21]. Indeed, Baldock *et al.* in Ref. [20] and I in Ref. [33] see an interatomic potential fit with the help of nested sampling to phase transition data gathered by experiments as a possible research avenue. Gola *et al.* [34] optimise a CuAu EAM with the Nelder-Mead algorithm [35] – an algorithm which I use as well for my fitting in Chapter 3 – and test their melting curve predictions via nested sampling. To conclude, nested sampling, may offer an excellent way of conveniently and efficiently including information of experimental phase diagrams in the fitting procedure.
Nested sampling was invented in the field of statistics [25, 26], but has been used as modelling tool for among others Lennard-Jones particles [27, 20, 32, 28, 36], hard spheres [37], aluminium [27, 20], iron [21], the CuAu melting curve [28, 34], nickel-titanium alloy [20] and peptides[38]. I conducted preliminary research on titanium using nested sampling based on an atomistic model in Ref. [33].
Previous work which uses more than the usual amount of thermodynamic fitting data does usually either rely on other more conventional data as well, or usually modifies or proposes to modify potentials which have already been fitted [17, 18, 22, 23, 16]. Ref. [19] differs from these as the basis of the potential optimisation are zero temperature properties deduced from experiments at finite temperature. As I want to fit directly to phase transition temperatures, I follow the first approach and present work on modifying two interatomic potentials from the literature, a CuAu potential [39–41] and a Ti potential [15], to reproduce phase transitions by relying to a large extent on nested sampling. CuAu shows a eutectic melting curve and various different solid phases [42]. The eutectic curve alone is an interesting characteristic to study: a certain composition ratio of Cu and Au shows a melting temperature which is smaller than the melting temperature of pure Cu as well as the one of pure Au [42]. Also, there exist already nested sampling simulations on the CuAu melting curve showing that an investigation with nested sampling is feasible [28, 34]. Additionally, the nested sampling parameters used by Baldock *et al.* [28] for the CuAu EAM [39–41] which are available under

Ref. [43] serve as basis for my own calculations. Titanium is a good material to study as even though there exists a high amount of discussion regarding certain parts of the phase diagram, the qualitative behaviour is known, sufficient data for fitting is available and there exists a triple point which is interesting to model [44]. Additionally, fitting an interatomic potential for it is challenging [15], making it a good benchmark material for my optimisation approach. The potential created by Hennig *et al.* [15] serves as basis for my modifications.

The structure of this thesis is as follows: The following section elaborates further on the fitting approach for the different potentials. In Chapter 2, I consider some interatomic potentials, various methods of investigating phase transitions with atomistic calculations, the Nelder-Mead algorithm [35] as well as the materials and transitions of interest. Chapter 3 contains the results and technical details of the methodology of the CuAu potential optimisation. I conduct fitting to single melting transitions and fitting to three melting transitions simultaneously. The Ti results and technical details of the relevant methodology are given in Chapter 4. Here, I particularly focus on converging the results. I also conduct an explorative adjusting of the potential. In Chapter 5, I discuss first the CuAu results and then the Ti results. Finally, I end with a conclusion and an elaboration on future work in Chapter 6. I acknowledge programs, scripts and libraries used in this thesis in the appendix Section A.1 in which I also consider how the different interatomic potentials are evaluated.

1.2 Approach

The aim of this work is to test the viability of fitting interatomic potentials directly to phase transition temperatures. The approaches for the CuAu eutectic melting curve and the Ti solid phase P-T phase diagram are shown in Figure 1.1. For the CuAu system, I use an EAM [39–41] and for the Ti system a MEAM [15] as starting point of my fitting. After the system and an interatomic potential is chosen, I investigate them with nested sampling [28, 45] for different numbers of atoms. All thermodynamic methods used in this thesis use a simulation cell with periodic boundary conditions. The number of atoms N contained in this cell can have strong effects on the results. These effects are known as “finite size effects”. The number of atoms used per simulation cell in nested sampling is limited by the computational cost of the algorithm. For example for 64 iron atoms, the computational costs are between 640 and 8820 h depending on the transitions involved [21]. For too small numbers of atoms, atomistic computational methods for the predictions of phase diagrams could in principle predict phase diagrams which are far from those predicted for the thermodynamic limit of $N \to \infty$. Baldock *et al.* showed a lowering of the melting temperature predicted by an interatomic potential for aluminium [46] by more than 200 K when increasing the number of

atoms from 32 to 128 in Ref. [47]. In such a case, it would not make sense to directly use the results from nested sampling calculations for small numbers of atoms to fit to transition temperatures. Even modifications to change the qualitative nature of the phase diagram could be futile. Hence, it is important to first gain insight in the strength of the finite size effects for the different transitions.

In the case of CuAu eutectic, I start with the melting temperature of pure gold and repeat the procedure including the fitting for different transitions. If not already conducted in an earlier cycle, the melting transition temperature for $N \to \infty$ is determined through extrapolating the nested sampling predictions for melting temperatures of various different numbers of atoms per simulation cell N. The evaluation of the Ti P-T phase diagram is too expensive to allow for this. Hence, I use an informed search method, thermodynamic integration (specifically a nonequilibrium version [48, 49]), to investigate the behaviour of the system for a large number of atoms. For Ti, I also compare the energies predicted with the original interatomic potential [15] to the energies predicted with density functional theory (DFT) for some structures sampled with nested sampling. This is done to see how the potential performs for the different phases expected in the phase diagram [44, 50, 51], but is not strictly necessary for the fitting (signified by using a white box for the process in Figure 1.1).

The next step is to conduct the fitting procedure. For the CuAu eutectic, for which an extrapolation of the melting temperatures is possible, I approximate the melting temperature of a slight modification of an interatomic potential for $N \to \infty$ in the following manner: I assume that the ratio between the extrapolated melting temperature and the melting temperature calculated with nested sampling for a finite number of N atoms stays constant for such slight modifications to the interatomic potential. I refer to this ratio as "extrapolation constant". This ratio is calculated from the original potential [39–41]. I then use the Nelder-Mead algorithm [35] to fit to a single experimental melting temperature or several experimental melting temperatures. The evaluation of the melting temperatures predicted by the modified potentials is done via nested sampling using a small number of atoms N per simulation cell. The temperature for $N \to \infty$ is estimated via the aforementioned "extrapolation constant". If the resulting potential does not sufficiently reproduce the eutectic melting curve, the process is started again by changing to a more appropriate transition or set of transitions for the fitting. The aim is to increase the complexity (if necessary) by first fitting the gold melting, then the melting of CuAu at a 50:50 atomic ratio and finally fitting simultaneous the melting temperature of pure gold, copper and CuAu at a 50:50 ratio. The fitting of only pure gold is seen as trial which does not aim to reproduce the full eutectic curve.

For Ti, the fitting process is different. After verifying that the qualitative behaviour of the (solid phase) P-T phase diagram is not affected too much, I can use small numbers of atoms to

test different modifications with nested sampling. Here, the potential is modified with Fourier modes. The amplitude of this modification is modified in the qualitative fitting process. As I do not have a precise way of extrapolating the transition temperatures, the solid-solid transition phase diagram prediction of the promising candidate must be calculated for larger number of atom with informed search methods. This way, I can investigate the extent to which the new potential makes improved phase diagram predictions. In the specific case I use thermodynamic integration (a nonequilibrium version [48, 49] as before) and testing whether a phase is stable under a molecular dynamics (MD) simulation at certain pressures and temperatures. This allows to test in which respect the new potential has improved in the description of the experimental phase diagram. I investigate how the energy predictions of the new potential perform against the DFT results for the sample structures. This could help to understand the phase diagram predictions of the new potential. Again, I use a white box for the process in Figure 1.1 to denote the fact that this is not strictly necessary for the fitting. Finally, for the functions of the interatomic potential which have been changed, the modified potential functions are displayed together with the original ones for both approaches. Additionally, the predictions for elastic and lattice constants, the volume per atom of the new potentials for CuAu and Ti are compared to the predictions of the original potentials and experiments for important structures. This is done to show in which respect the new potentials have been changed and how it performs for basic properties in comparison to the original potentials.

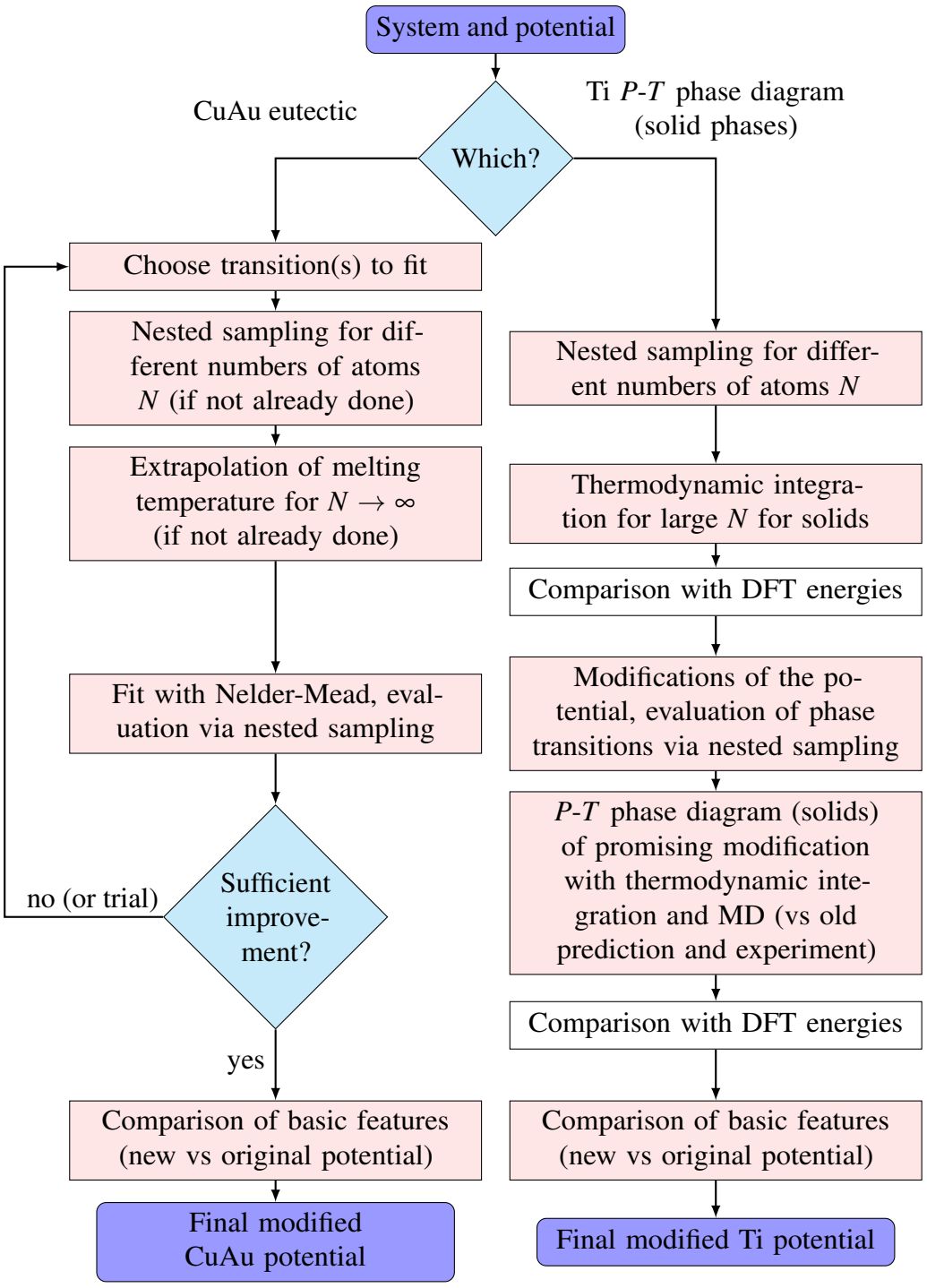


Fig. 1.1 A flowchart of the two approaches to fitting potentials presented in this work. It is explained in detail in the text. The basic steps are investigating the system with nested sampling and investigating of finite size effects. This is followed by the actual fitting procedure which needs to take account of the finite size effects or at least investigate them for the resulting potential. Eventually, the basic features of the new potentials need to be compared to the the original potentials.

Chapter 2

Literature review

The following chapter starts with a section on Molecular Dynamics, followed by short overview on a number of interatomic potential models. Here, I will also shortly present the MEAM potential of Hennig *et al.* [15] which I use in my work. Followed by this, I consider the Nelder-Mead algorithm [35], as well as a short summary of the relation between some thermodynamic quantities and the respective partition functions. Afterwards, I introduce the reader to the nested sampling algorithm for atomistic systems. Additionally, thermodynamic integration and the quasi-harmonic approximation are covered. I finish with a section on the phase diagram of titanium and a section on the melting curve of CuAu which I will both investigate in this thesis.

2.1 Molecular Dynamics

Molecular Dynamics (MD) is one of the most important methods for the simulation of materials on the atomistic or molecular level. The idea is to calculate the force on each particle via a method of choice. Based on Newton's equation of motion, one can approximate the new position of each particle after a time step Δt. However, it is to note that Molecular Dynamics is not an exact representation of reality. Even if we knew the exact positions of a number of particles, the velocities of all atoms to exact precision as well as the exact forces, we would not be able to predict the exact positions of the atoms over time using Molecular Dynamics. This is due to the fact that we need to numerically find a solution to Newton's equations of motion which leads to numerical errors [52]. Ref. [52] offer a more detailed analysis of this problem. In the following I cover the salient points: The aforementioned errors lead to a Lyapunov instability which means that the numerical prediction of the particle position, orientation and velocity differs from the hypothetical behaviour if the errors did not exist over a relatively short time. The size of these numerical errors is irrelevant for the

question whether the instability occurs or not. However, using an appropriate scheme to solve the equations of motion of an atomic system, Molecular Dynamic is useful for statistically computing the mean of its features over time [52]. I consider a simple but useful integration scheme in Section 2.1.1. Molecular Dynamics simulations can also be used to predict phase transitions as I explain in Section 2.1.2. Some examples of models used for evaluating the forces on the atoms are covered in Section 2.2.

2.1.1 The Verlet scheme

In most cases, the optimal scheme for conducting Molecular Dynamics is the Verlet scheme [52]. According to this scheme, the position of an atom with mass m after a time step $\Delta t > 0$ is given by [52]

$$\vec{r}(t+\Delta t) = 2\vec{r}(t) - \vec{r}(t-\Delta t) + \frac{1}{2}\frac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2}\Delta t^2 + O\left(\Delta t^4\right) \ . \tag{2.1}$$

Here, t is the time before the step. The time $t-\Delta t$ is the time before the last two steps. Given the force $\vec{F}$ on the atom, we can determine the acceleration of the atom $\vec{a} = \mathrm{d}^2\vec{r}/\mathrm{d}t^2 = \vec{F}/m$. If needed, the velocity is approximated as [52]

$$\vec{v}(t) = \frac{\vec{r}(t+\Delta t) - \vec{r}(t-\Delta t)}{2} + O\left(\Delta t^2\right) \ . \tag{2.2}$$

We ignore terms of order Δt^4 for the calculation of the atom position and terms of order Δt^2 for the velocity [52]. To reduce the error in Equation 2.2, another similar scheme is in principle available [52].
A derivation is of the Verlet scheme is for example given in Ref. [52], on which the following short derivation is based: First, we Taylor expand $\vec{r}(t+x)$ for small x:

$$\vec{r}(t+x) = \vec{r}(t) + \vec{v}(t)\,x + \frac{1}{2}\vec{a}x^2 + \frac{1}{6}\frac{\mathrm{d}\vec{a}}{\mathrm{d}t}x^3 + O\left(x^4\right) \tag{2.3}$$

Here, $\vec{v} = \mathrm{d}\vec{r}/\mathrm{d}t$ is the velocity of the atom. To eliminate the velocity, we now add the Taylor expansions $\vec{r}(t+\Delta t)$ and $\vec{r}(t-\Delta t)$. Rearranging the result leads to Equation 2.1. Equation 2.2 is gained similarly, just that we substract the two Taylor expansions from each other.

2.1.2 MD simulations for phase transitions

One of the simplest ways of investigating the phase transitions of materials is to just let a phase evolve at a certain temperature as well as pressure or volume. Hennig *et al.* [15] for

example use this to calculate their phase diagram. At certain pressures and temperatures the titanium bcc phase transforms into the hcp and the titanium ω [53] phase [15]. (I consider titanium and its phases [50] in more detail in Section 2.9.) As Hennig *et al.* however point out, the hcp phase does not transform in this situation to the ω phase when they would expect it to [15]. If a structure is at least metastable it can take a long time until it transforms to a more stable structure if we simulate a small number of atoms. This is the explanation Hennig *et al.* [15] give for the abovementioned lack of hcp to ω transformation. However, it is not clear how many MD steps one should conduct until one is sure of the stability of a phase. This means that even if a phase A transforms to a phase B, the phase B may be itself only metastable and transform to C after a certain time. Hence, it is difficult to determine whether B or C is stable with this method.
Another method is to create an interface between to phase A and B. Examples for the usage of this for liquid-solid transitions are Hennig *et al.* [15] and Ref. [54]. It is also possible to investigate solid-solid transitions in this manner as for example shown in Ref. [15]. Setting up solid-solid interfaces can be relatively complicated as one can see for example in the work of Hennig *et al.* [15] in which they set up the interface according to the likely phase transition path [55, 56]. This assumes that the interatomic potential actually favours this transition path which is in general not necessarily guaranteed. Hennig *et al.* however, include this path explicitly when optimising their potential [15]. Additionally, the lattice parameters between two phases may not match. This means that it is likely that the lattice parameters of at least one of the phases are not as they should be for the equilibrium at the simulated temperature and pressure or volume. This leads to a different Gibbs free energy of that structure which in turn could affect the stability of a phase. In the worst case, a phase which is stable may transform into an unstable one.

2.2 Interatomic potentials

Interatomic potential models are important in physics, materials science and engineering. There have been many developments for efficient but still applicable potential models. In the following, I want to give a short overview of a few potential forms.

2.2.1 Lennard-Jones potential

Lennard-Jones potential, dates back to 1924 [2, 3] and is one of the most basic and well known interatomic potentials. It describes the potential energy between a pair of atoms with

distance of r from each other. It is widely used in the following forms:

$$U_{LJ} = \varepsilon \left(\left(\frac{r_m}{r} \right)^{12} - 2 \left(\frac{r_m}{r} \right)^{6} \right) = 4\varepsilon \left(\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right) \tag{2.4}$$

The parameters ε and r_m – or for the second form ε and σ – need to be fitted for the system at hand. If the interatomic distance r is equal to r_m the force between the atoms becomes zero: We have reached the minimum with $U_{LJ} = -\varepsilon$. For $r = \sigma = r_m/2^{1/6}$, the potential energy is $U_{LJ} = 0$. The total potential energy is then calculated by the sum over all atom pairs:

$$E = \sum_{i<j} \varepsilon \left(\left(\frac{r_m}{r_{ij}} \right)^{12} - 2 \left(\frac{r_m}{r_{ij}} \right)^{6} \right) \tag{2.5}$$

Here, the distance between atom i and j is denoted as r_{ij}.

2.2.2 The embedded atom method

One way to include more than just the pair potential effects is the embedded atom method which, additional to pair interactions, uses an embedding energy $U\left(\rho_{h,i}\right)$ for each atom i [4, 5]. The other atoms give rise to an electron density, whose interaction with atom i the embedding function takes into account [4, 5]. An atom j at a distance r_{ij} from atom i leads to an electron density of $\rho_a\left(r_{ij}\right)$ at the coordinate of i which determines the so called "host density" [4, 5]:

$$\rho_{h,i} = \sum_{j \neq i} \rho_a\left(r_{ij}\right) \tag{2.6}$$

Adding the embedding function for each atom i to the pair potential contributions, the sum of $\phi\left(r_{ij}\right)$ over all atom pairs, we arrive at [4, 5]

$$E = \frac{1}{2} \sum_{i,j} \phi\left(r_{ij}\right) + \sum_i U\left(\rho_{h,i}\right) \tag{2.7}$$

for the potential energy of the system. The EAM can also be used for alloys of metals [57–59]. Here, one needs to consider which atom species interacts with which species. The function for the electron density becomes $\rho_{a,ij}\left(r_{ij}\right)$ and the pair interaction becomes $\phi_{a,ij}\left(r_{ij}\right)$ as now these functions themselves depend on the species of atoms i and j. Additionally, we now need to use an embedding function $U_i\left(\rho_{h,i}\right)$ whose form depends on the species of atom i. For two species A and B for example, assuming that the functions are symmetric under the exchange of i and j, we get eight different functions which make up the potential,

namely: $\phi_{a,AA}\left(r_{ij}\right)$, $\phi_{a,BB}\left(r_{ij}\right)$, $\phi_{a,AB}\left(r_{ij}\right) = \phi_{a,BA}\left(r_{ij}\right)$, $\rho_{h,AA}\left(r_{ij}\right)$, $\rho_{h,BB}\left(r_{ij}\right)$, $\rho_{h,AB}\left(r_{ij}\right) = \rho_{h,BA}\left(r_{ij}\right)$, $U_A\left(\rho_{h,i}\right)$ and $U_B\left(\rho_{h,i}\right)$.

2.2.3 The modified embedded atom method

The modified embedded atom method is a modification of the embedded atom method described in Section 2.2.2 [6, 7]. Baskes *et al.* use the term "modified EAM" or "MEAM" for the first time in Ref. [7]. However, in Ref. [7], Baskes *et al.* actually further improve on a potential published earlier in Ref. [6]. We therefore turn our attention to this earlier model: Shearing of Silicon cannot be described by a classical EAM [6]. This is due to the fact that the host density does not explicitly depend on the angle θ_{jik} enclosed by the bonds ij and ik [6]. The EAM hence ignores any effect of how the bonds are situated with regards to each other on the electron density contributed by them at the position of atom i [6]. Therefore, Baskes proposes to add the term [6]

$$\rho_{h,i,Correction} = -k_{angle} \sum_{j\neq i, k\neq i} \left(1 - 3\cos^2\left(\theta_{jik}\right)\right) \rho_a\left(r_{ij}\right) \rho_a\left(r_{ik}\right) \tag{2.8}$$

to the EAM host density [4, 5] given in Equation 2.7. Here, the bond length are again given by r_{ij} and r_{ik}. k_{angle} is an empirical parameter [6]. Extensive work was conducted in the literature to make the potential model more applicable [7, 60].
Let us now consider the MEAM potential for titanium by Hennig *et al.* [15] which I consider in this thesis. Their potential is based on a silicon MEAM potential [61]. The different functions of the potentials are described by splines [61, 15]. Due to this, the potential is very fast [15]. The host density takes the form of [15, 61]:

$$\rho_{h,i} = \sum_{j\neq i} \rho_a\left(r_{ij}\right) + \sum_{j\neq i, k\neq i} g\left(\cos\left(\theta_{jik}\right)\right) f\left(r_{ij}\right) f\left(r_{ik}\right) \tag{2.9}$$

Hennig *et al.* claim that opposed to other potentials, it can describe all usually expected titanium phases [50] (α, β and ω [53]) very well [15]. I consider the different titanium phases in Section 2.9. Note that the possibility to describe all phases with good accuracy is very important for my studies and is a important reason for choosing the potential. As I want to modify an already established potential to fit it to a phase diagram, I want to start with a potential which describes the phase diagram already very well.

2.3 The Nelder-Mead method

Many minimization methods use a gradient to minimize a function. This can be difficult if the gradient of the function to be minimized is not known analytically and would be very expensive to calculate numerically. Methods which do not use a gradient are a potential solution to this problem. The Nelder-Mead algorithm [35] is such a method. Ref. [62] gives a very detailed summary which I will rely on in the following. I present the Nelder-Mead algorithm for minimising the objective function while they use it for maximising. However, the steps are in principle the same. In the Nelder-Mead algorithm, $n+1$ vertices form a simplex which explores the n dimensional space. The Nelder-Mead algorithm has four parameters in its modern form: α, β, γ and δ [63]. In the original paper, the most success had been with the values $\alpha = 1$, $\beta = 1/2$, $\gamma = 2$ and $\delta = 1/2$ [35]. These values are kept in the vast majority of implementations [64]. The algorithm for minimising a function $f(\vec{x})$, the objective function, works as follows [62]:

1. Initialisation: We start with $n+1$ vertices. For each, we calculate the value of the objective function $f(\vec{x})$. Ref. [62] recommends to calculate these vertices as $\vec{x}_i = \vec{x}_1 + \Delta_i \cdot \vec{e}_i$ where $i \in \{2, 3..., n, n+1\}$. Here, $\vec{x}_1$ is chosen so that it is likely close to the minimum, Δ_i is a suitable parameter and the vectors $\vec{e}_i$ represent the unit vectors of our space [62]. Our iteration begins with a **reflection**.

2. Reflection: Now, we need to consider the ordering of the vertices with regards to $f(\vec{x})$. We take the vertex $\vec{x}_{max}$ with the greatest value $f_{max} = f(\vec{x_{max}})$, the vertex $\vec{x}_{2nd}$ with the second greatest value $f_{2nd} = f(\vec{x}_{2nd})$ and $\vec{x}_{min}$ with the smallest value $f_{min} = f(\vec{x}_{min})$. We calculate the centre coordinate $\vec{x}_{centre}$ of the vertices excluding the vertex x_{max}. The reflection vertex is then

$$\vec{x}_{refl} = \vec{x}_{centre} + \alpha(\vec{x}_{centre} - \vec{x}_{max}) \ . \tag{2.10}$$

   We choose one of the following three options:

   (a) We conduct an **expansion** in the case of $f_{refl} = f(x_{refl}) < f_{min}$.

   (b) We accept x_{refl}, discard x_{max}, and **determine whether the minimum has been reached** in the case of $f_{min} \leq f_{refl} < f_{2nd}$.

   (c) We conduct a **contraction** in the case of $f_{2nd} \leq f_{refl}$.

3. Expansion: We try an expansion of our simplex with a new vertex

$$\vec{x}_{expan} = \vec{x}_{centre} + \gamma(\vec{x}_{refl} - \vec{x}_{centre}) \ . \tag{2.11}$$

We now replace $\vec{x}_{max}$ in our simplex by $\vec{x}_{expan}$ in the case of $f_{expan} = f(\vec{x}_{expan}) < f_{refl}$ and by $\vec{x}_{refl}$ in all other cases. Afterwards, we **determine whether the minimum has been reached**.

4. Contraction: We first need to decide which point besides the centre point we use in our contraction. We use $\vec{x}_{refl}$ in the case of $f_{2nd} \leq f_{refl} < f_{max}$ (outer contraction) and $\vec{x}_{max}$ for all other cases (inner contraction). The new vertex is then

$$\vec{x}_{contr} = \vec{x}_{centre} + \beta \left(\vec{x}_{refl/max} - \vec{x}_{centre} \right) \ . \tag{2.12}$$

   Now, we replace $\vec{x}_{max}$ by $\vec{x}_{contr}$ if $f_{contr} = f(\vec{x}_{contr}) \leq f_{refl}$ for the outer contraction or if $f_{contr} < f_{max}$ for the inner contraction. If the respective conditions do not apply, we shrink the simplex. If they apply, we **determine whether the minimum has been reached**.

5. Shrink: We now shrink the simplex by replacing each vertex $\vec{x}_i$ except for $\vec{x}_{min}$ by

$$\vec{x}_i' = x_{min} + \delta \left(\vec{x}_i - \vec{x}_{min} \right) \ . \tag{2.13}$$

   This means that our best vertex is used as the centre of the shrinking procedure. Afterwards, we **determine whether the minimum has been reached**.

6. Determine whether the minimum has been reached: We terminate the algorithm if it is sufficiently converged. If this is not true, we conduct the next iteration starting with a **reflection**.

The Nelder-Mead method does not converge to a local minimum for all functions and starting positions, as there can be rare occasions when the vertex continues to shrink in size, but does not come closer to the minimum [65, 62]. Perhaps the easiest way to still find the minimum is to just change the initialisation vertices and start the algorithm anew, but a more sophisticated method of dealing with the problem, an "oriented restart" step [66], is in principle available [62]. The Nelder-Mead method does usually determine $f(\vec{x})$ roughly twice as many times as the total number of iterations [62]. Additionally, even the classical Nelder-Mead algorithm deals well with moderate statistical noise [63].

2.4 Partition functions and thermodynamic quantities

We want to shortly recapitulate a few thermodynamic quantities: For the canonical ensemble the internal energy U of a system, the average total energy $\langle E \rangle$ of the system, is related to

the partition function Z via:

$$U = -\frac{\partial \ln(Z)}{\partial \beta} \tag{2.14}$$

Here, β is defined as $\beta = 1/(k_B T)$ with the Boltzmann constant k_B and the temperature T. We can now calculate the heat capacity at constant volume:

$$C_V = \frac{\partial U}{\partial T} \tag{2.15}$$

One can easily calculate that the heat capacity C_V can also be written in dependence of the energy variance:

$$C_V = \frac{1}{k_B T^2}\left(\langle E^2\rangle - \langle E\rangle^2\right) \tag{2.16}$$

The Helmholtz free energy is:

$$F = -k_B T \ln(Z) \tag{2.17}$$

In the NPT-ensemble, the enthalpy H depends on the NPT-ensemble partition function Δ as follows:

$$H = -\frac{\partial \ln(\Delta)}{\partial \beta} \tag{2.18}$$

Similar to the canonical case, we arrive at the heat capacity at constant pressure via

$$C_P = \frac{\partial H}{\partial T} = \frac{1}{k_B T^2}\left(\langle H^2\rangle - \langle H\rangle^2\right) \ . \tag{2.19}$$

The Gibbs free energy is

$$G = -k_B T \ln(\Delta) \tag{2.20}$$

2.5 Nested sampling

The partition function of a system allows the calculation of important thermodynamic quantities which one can use to calculate phase diagrams as shown with success for example in Ref. [32], Ref. [20] and Ref. [21]. Nested sampling allows for an easy calculation of the partition functions [27, 32]. As already mentioned in Section 1.1, we do not need to supply the phases which could be expected in the phase diagram beforehand [27, 20]. This makes nested sampling an excellent method for computing phase diagrams if little information about the behaviour of the system is available [20, 21, 27, 28]. However, as also already considered in Section 1.1 such blind search algorithms are considerably more expensive than informed search methods. Even though nested sampling is more efficient than its competitors Wang-Landau sampling [29] and parallel tempering [30, 31] [20, 27], the interatomic potential is

still typically called on the order of 10^8 to 10^{10} times [20, 21, 27, 28]. This means that the number of atoms per simulation box is limited making the estimation of finite size effects more difficult than for informed search methods. Consider for example that finite size effects were considered for up to only 128 atoms in Ref. [47]. Indeed, nested sampling calculations for the calculation of phase transitions at constant pressure for metals and alloys are usually conduct with less than 100 atoms [20, 21, 27, 28, 34]. In the following, I want to introduce the reader to the basics concepts of nested sampling for atomic systems. More detailed work can be found in Ref. [27].

2.5.1 Nested sampling at constant volume

For simplicity, we will first consider nested sampling at constant volume. If we calculate the partition function of a system, we sum or integrate the relevant Boltzmann factor over all possible states of the system. In the canonical ensemble for example, the partition function Z of a system with N different microscopic states is well known as:

$$Z = \sum_{i=1}^{N} e^{-\beta E_i} \tag{2.21}$$

As before, we use $\beta = 1/(k_B T)$. E_i represents the energy of the state i.
In the case of the atomic systems with N particles, this sum transforms into an integral:

$$Z = \frac{1}{N! h^{3N}} \int \int e^{-\beta H(\vec{q},\vec{p})} \mathrm{d}\vec{q} \mathrm{d}\vec{p} \tag{2.22}$$

The Hamiltonian $H(\vec{q},\vec{p})$ describes the whole system and hence depends on all configurational coordinates $\vec{q}$ and all momentum coordinates $\vec{p}$. The factor of $1/N!$ accounts for the fact that the particles are indistinguishable. h is the Planck constant.
Nested sampling [25, 26] allows us to approximate this partition function as a sum over a number of sampled states i [32, 37]:

$$Z = C \cdot \sum_i \omega_i e^{-\beta E_i} \tag{2.23}$$

The constant C is irrelevant for the determination of phase transitions as we apply logarithms and/or derivatives when calculating the heat capacity, energy and free energy as seen in Section 2.4 [37]. Hence, we do not need to know its exact value. Each sample state i accounts for an infinite number of states with a similar energy. This is taken account for by the weight factor ω_i [32].

We now look at the algorithm itself [37, 27]: Salient to the algorithm are the sample systems (e.g. 64 atoms in a box with periodic boundary conditions) which are referred to as walkers in the following. The number of walkers is denoted as K. To initialize our algorithm, we distribute these walkers uniformly in phase space, limited only by a maximum walker energy ceiling, E_0. In the canonical case this means choosing uniformly random atom positions and momenta in our simulation box. We want to include ideal gas like states in our calculation which requires a sufficiently high energy ceiling. The iteration variable i starts at 0. The algorithm is as follows [37, 27]:

1. Cull the walker with the highest energy which we denote as E_{i+1} – i.e. delete it from the list of walkers.

2. Make a copy of a walker from the not culled walkers and ensure that it does not correlate anymore with the original using a suitable method. This walker must be randomly selected. The energy ceiling for this process is E_{i+1}. This newly created walker replaces the culled walker. The distribution of walkers should again be sufficiently uniform in phase space.

3. Repeat the whole loop with i set to $i+1$ until sufficient amount of data has been gathered.

Let us now consider the derivation given by Partay *et al.* [37] on which the remainder of this section is based. For each iteration i, we can calculate the average volume of the phase space which is still available for walkers fulfilling $E < E_{i+1}$: Every iteration, it decreases by a factor of $K/(K+1)$ [26, 32, 37]. We will call this volume remaining phase space volume and denote it by ζ. The remaining phase space volume in each iteration is then [26, 32, 37]:

$$\zeta_i = \zeta_0 \left(\frac{K}{K+1} \right)^i \tag{2.24}$$

A different way to see the remaining phase space volume ζ under a total energy limit E is to see it as an integral of the density of states $\Omega(E')$ over all energies $E' \leq E$:

$$\zeta(E) = \int_{-\infty}^{E} \Omega\left(E'\right) \mathrm{d}E' \tag{2.25}$$

It becomes clear now that the density of states can be written as derivative:

$$\Omega(E) = \frac{\mathrm{d}\zeta(E)}{\mathrm{d}E} \tag{2.26}$$

Similar to Ref. [37], where however only the configurational part of the integral is considered, instead of integration over all phase space coordinates as in Equation 2.22, we consider the space of total energies E:

$$Z = \frac{1}{N!h^{3N}} \int e^{-\beta E} \Omega(E)\, \mathrm{d}E = \frac{1}{N!h^{3N}} \int e^{-\beta E} \frac{\mathrm{d}\zeta(E)}{\mathrm{d}E} \mathrm{d}E \tag{2.27}$$

An easy substitution of E by ζ leads to an integral over remaining phase space volume. The nested sampling algorithm presented above has sampled energies and their remaining phase space volumes which allows for a numerical integration of the partition function [37, 27]:

$$Z = \frac{1}{N!h^{3N}} \int e^{-\beta E(\zeta)} \mathrm{d}\zeta \approx \frac{1}{N!h^{3N}} \sum_i (\zeta_i - \zeta_{i+1})\, e^{-\beta E(\zeta_i)} \tag{2.28}$$

Note that Ref. [37] and Ref. [27] do only consider the potential energy in the sum of Equation 2.28 and include the kinetic part in a factor before the sum. We can see that we have arrived at Equation 2.23 with the weight $\omega_i = \zeta_i - \zeta_{i+1}$. The phase space volume in between two consecutive energy ceilings i and $i+1$ can be interpreted as this weight [37]. In practice, we can set $\zeta_0 = 1$ as the resulting constant factor is then absorbed in the constant C in Equation 2.23 [37].

2.5.2 Constant pressure case

As Baldock showed in [27], we can calculate the partition function for the constant pressure P case in a similar way as the constant volume case in Section 2.5.1: We do not just sample over atom momenta and positions, but have to also include simulation box volume as well as shape in our decorrelation steps as well. I now shortly summarise how Baldock calculates the partition function based on the "fully flexible cell ensemble" [67] which can be used to approximate the NPT ensemble [27, 20]. Let us consider the linearly independent vectors $\vec{h}_1$, $\vec{h}_2$, $\vec{h}_3$ defining a parallelepiped. Together, they define our periodic simulation box. Further, we define:

$$\bar{\bar{h}} = \left(\vec{h}_1, \vec{h}_2, \vec{h}_3\right) \tag{2.29}$$

Using the volume V of the simulation cell, the normalised cell matrix $\bar{\bar{h}}_0$ is then given by:

$$\bar{\bar{h}}_0 = \frac{\bar{\bar{h}}}{V^{1/3}} \tag{2.30}$$

Here, V is given by $V = \left|\det\left(\bar{\bar{h}}\right)\right|$. The partition function is the given by [67, 68]:

$$\Delta \propto \frac{1}{N!h^{3N}} \int \mathrm{d}\bar{\bar{h}} \frac{1}{\det\left(\bar{\bar{h}}\right)^2} \int \mathrm{d}\vec{q} \int \mathrm{d}\vec{p} e^{-\beta H} \tag{2.31}$$

Here $H = E + PV$ is the enthalpy of our system. Baldock [27] uses the fractional coordinates $\vec{s}_i = \bar{\bar{h}}\vec{q}_i$ (i stands for atom i here) and presents the partition function in the following form [27, 67, 68]:

$$\Delta \propto \frac{1}{N!h^{3N}} \int_0^\infty \mathrm{d}V V^N \int \mathrm{d}\bar{\bar{h}}_0 \delta\left(\det\left(\bar{\bar{h}}_0\right) - 1\right) \int \mathrm{d}\vec{s} \int \mathrm{d}\vec{p} e^{-\beta H} \tag{2.32}$$

Here, $\delta(x)$ is the Dirac delta function.
Our algorithm in principle stays the same as in Section 2.5.1. However, now our phase space over which we sample also includes cell volume and cell shape and consider enthalpies H_i instead of energies [27]. Equivalently our partition function becomes [27, 20, 28]

$$\Delta = C \cdot \sum_i \omega_i e^{-\beta H_i} \ . \tag{2.33}$$

Here, C is again a constant. The remaining phase space volume is ζ_i and is determined in the same way as before leading to the weight $\omega_i = \zeta_i - \zeta_{i+1}$. Note that Ref. [27] and Ref. [20] do not consider the total enthalpy, but the enthalpy without the kinetic energy and include the kinetic part in C.

2.5.3 Parallelizing

There exist two ways of parallizing nested sampling in pymatnest [45] which I explain in the following.
First, instead of letting only one walker undergo the decorrelation scheme of length L, we can let $n_{walking}$ number of walkers undergo a decorrelation scheme of length $L/n_{walking}$ [27, 47].
The second option is to cull more than one walker per iteration [27, 69, 36].
In the case of $n_{cull} > 1$, we have to adjust our remaining phase space volume calculation [69, 36]: Again, we set $\zeta_0 = 1$ for simplicity. We cull the walkers consecutively. After culling the first walker, the remaining phase space volume will be by a factor of $K/(K+1)$ smaller [37, 26] as it is the case in Section 2.5.1. For the next walker, there are only $K-1$ walkers left to start with. Hence, after culling, the remaining phase space volume is again reduced by a factor of $(K-1)/(K-1+1)$. The next factor will be $(K-2)/(K-2+1)$ and so forth (assuming n_{cull} and K are sufficiently high. Denoting the modulo operation via %, we arrive

at the following formula for the remaining phase space volume after the n-th culling of a walker [36]:

$$\zeta_n = \prod_{i=0}^{n} \frac{K - i\%n_{cull}}{K - i\%n_{cull} + 1} \tag{2.34}$$

As before, the weight is $\omega_n = \zeta_n - \zeta_{n+1}$ [36].

2.6 Thermodynamic integration

Theoretically, we do not need to consider the different competing phases of our model or material when we set up the nested sampling algorithm. Opposed to this, informed-search methods require some knowledge about the competing phases. Examples for informed-search methods are thermodynamic integration, quasi-harmonic approximation and coexistence molecular dynamic simulations. The advantage of informed-search methods is that they are in general less computationally expensive. However, it is difficult to test all possible competing phases, especially as some atomic interaction models may stabilize phases which are not expected in nature. In the following I consider how to calculate the Helmholtz free energy of a system with thermodynamic integration. Besides not covering certain low temperature quantum mechanical contributions which we will cover in Section 2.7, thermodynamic integration also assumes that the phases which are investigated are at least metastable at the temperatures and pressures in question. As it does not sample the whole phase space, it requires much less computational resources than nested sampling or similar methods. This means larger numbers of atoms per simulation box can be considered than for the nested sampling algorithm from the previous Section 2.5. For example, Freitas *et al.* conducted calculations for 54000 atoms [48] while nested sampling calculations at constant pressure are rarely conducted with much more than 100 atoms [20, 21, 27, 28, 34, 47]

2.6.1 The equilibrium approach to thermodynamic integration

We will follow the explanation given in Ref. [52]. It is not usually possible to analytically calculate and impossible to directly measure the free energy of a system. Instead we need to calculate the free energy difference to a reference state, for example the Einstein crystal [70]. The Einstein crystal is a model in which the atoms are described as independent harmonic oscillators centred on their position in the crystal. It is useful for the calculation of the free energy of crystals. To calculate the free energy, we gradually change the potential governing the atomic interactions so that it describes the reference system at the end [71]. An example is

the switch from a crystal governed by an interatomic potential to the Einstein crystal [72] as for example used in Ref. [48] with nonequilibirum thermodynamic integration calculations. The idea is the following: A parameter λ is used to change the potential from state A to B by going from 0 to 1. The potential energy U_{pot} in dependence of λ is then given via the "Frenkel-Ladd path" [72, 52]:

$$U_{pot}(\lambda) = (1-\lambda)U_A + \lambda U_B \tag{2.35}$$

Here, U_A and U_B are the potential energies of state A and B respectively. Let us now use system A as reference. To get the free energy F_B of B we will integrate the derivative of the free energy:

$$F_B = \int_A^B \frac{\partial F}{\partial \lambda} \mathrm{d}\lambda + F_A \tag{2.36}$$

Here, F_A is the free energy of the reference system. The kinetic term E_{kin} of the Hamiltonian $H(\lambda) = E_{kin} + U_{pot}(\lambda)$ does not depend on the potential and hence also not on λ. Using the relation between canonical partition function Z and free energy F, we see that the following holds:

$$\frac{\partial F}{\partial \lambda} = \frac{\partial}{\partial \lambda}\left(-k_B T \ln(Z)\right) = -k_B T \frac{\frac{\partial Z}{\partial \lambda}}{Z} \tag{2.37}$$

If we now insert the definition of the canonical partition function,

$$Z = \frac{1}{N! h^{3N}} \int\int \exp\left(-\beta H(\lambda)\right) \mathrm{d}\vec{q}\mathrm{d}\vec{p}\ , \tag{2.38}$$

we arrive at

$$\frac{\partial F}{\partial \lambda} = \frac{\int\int \frac{\partial U_{pot}(\lambda)}{\partial \lambda} \exp\left(-\beta H(\lambda)\right) \mathrm{d}\vec{q}\mathrm{d}\vec{p}}{\int\int \exp\left(-\beta H(\lambda)\right) \mathrm{d}\vec{q}\mathrm{d}\vec{p}}\ . \tag{2.39}$$

This is nothing else than an ensemble average of $\partial U_{pot}/\partial\lambda$. To shorten it, we can write it the following way [52]:

$$\frac{\partial F}{\partial \lambda} = \left\langle \frac{\partial U_{pot}(\lambda)}{\partial \lambda} \right\rangle_\lambda \tag{2.40}$$

We can now numerically integrate Equation 2.36 and calculate the free energy of system B. Equation 2.36 becomes

$$F_B = \int_A^B \left\langle \frac{\partial U_{pot}(\lambda)}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda + F_A\ . \tag{2.41}$$

A problem with the thermodynamic integration method presented above is that we have to calculate

$$\left\langle \frac{\partial U_{pot}(\lambda)}{\partial \lambda} \right\rangle_{\lambda} \tag{2.42}$$

for every λ value required by our numerical integration by an appropriate sampling scheme. One may be able to speed up the simulations by using nonequilibrium methods [73, 49, 48]. We will consider a nonequilibrium scheme [49] which was recently implemented in LAMMPS [74] by Freitas *et al.* [48].

2.6.2 A nonequilibrium approach

In the previous section, we saw an equilibrium approach to thermodynamic integration. As mentioned before in Section 2.6, the number of calculations necessary to conduct the thermodynamic integration may be reduced through the use of nonequilibirum approaches [73, 49, 48]. In this Section, we will follow the explanation of a specific nonequilibrium approach [49, 48] given in Ref. [48].
Here, λ depends explicitly on the time in the simulation. Over the course of the time span t_{path}, $\lambda(t)$ changes from 0 to 1 switching system A to B. Let us reconsider the integral in Equation 2.41 for the Hamiltonian instead of the potential energy. This is the reversible work needed to go from state A to B [48]:

$$W_{reversible,A\to B} = \int_A^B \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda} \mathrm{d}\lambda \tag{2.43}$$

The irreversible work needed to go from state A to B without intermediate equilibration over a time t_{path} is given by [48]

$$W_{irreversible,A\to B} = \int_0^{t_{path}} \frac{\mathrm{d}\lambda}{\mathrm{d}t} \frac{\partial H(\lambda(t))}{\partial \lambda} \mathrm{d}t \ . \tag{2.44}$$

This integral is again calculated numerically. Now, we need to connect $W_{irreversible}$ to the free energy difference. The free energy difference is [48]

$$\Delta F = F_B - F_A = W_{reversible,A\to B} = W_{irreversible,A\to B} - E_{dissipation,A\to B} \ . \tag{2.45}$$

The energy $E_{dissipation,A\to B}$ which dissipates in that process from A to B is not generally the same as $E_{dissipation,B\to A}$ for the path from B back to A [49]. However, for large enough t_{path}, the equality $E_{dissipation,A\to B} = E_{dissipation,B\to A}$ is true [49]. This means that we can now easily calculate the free energy difference and the energy which dissipated in the process $A \to B$

and $B \rightarrow A$ [48, 49]:

$$\Delta F = \frac{W_{irreversible,A\rightarrow B} - W_{irreversible,B\rightarrow A}}{2} \tag{2.46}$$

$$E_{dissipation,A\rightarrow B} = E_{dissipation,B\rightarrow A} = \frac{W_{irreversible,A\rightarrow B} + W_{irreversible,B\rightarrow A}}{2} \tag{2.47}$$

To summarize, we now have a nonequilibrium method which for large enough t_{path} can calculate the free energy of system B using a reference system A via [48, 49]

$$F_B = F_A + \frac{W_{irreversible,A\rightarrow B} - W_{irreversible,B\rightarrow A}}{2} \ . \tag{2.48}$$

The remaining question is how exactly do we parametrize $H(\lambda)$. Ref. [48] uses a simple linear interpolation, the Frenkel-Ladd path [72, 52] as in Equation 2.35 (just with the Hamiltonian instead of the potential energy) and then proposes to describe $\lambda(t)$ through either

$$\lambda(t) = t \tag{2.49}$$

or [75, 76]

$$\lambda(t) = t^5\left(70t^4 - 315t^3 + 540t^2 - 420t + 126\right) \ . \tag{2.50}$$

2.7 Quasi-harmonic approximation

Another way to calculate the Gibbs free energy of a system is quasi-harmonic approximation. For this, we approximate the solid via non-interacting quantum harmonic oscillators. The phonons are assumed to not interact with each other and are treated as quantum harmonic oscillators. These assumptions are only valid for sufficiently low temperatures. Additionally, the phases must not have any phonon wave vector with imaginary phonon frequencies which would lead to unphysical imaginary contributions to the free energy. Such frequencies are a well known sign of instabilities. As we will see in the following, quasi-harmonic approximation does however consider quantum mechanical contributions which are relevant at low temperatures and ignored by purely classical approaches like thermodynamic integration and nested sampling.

Let us now start with the derivation of the Helmholtz free energy. The eigenvalues of the quantum harmonic oscillator are well known:

$$E_{\omega_{i,\vec{k}},n} = \hbar\omega_{i,\vec{k}}\left(n + \frac{1}{2}\right) \tag{2.51}$$

Here $\omega_{i,\vec{k}}$ is the frequency of phonon i and $n \in \mathbb{N}_0$ and $\vec{k}$ denotes the reciprocal vector in the first Brillouin zone. The canonical partition function of one oscillator with frequency $\omega_{i,\vec{k}}$ then becomes:

$$Z_{\omega_{i,\vec{k}}} = \sum_0^\infty e^{-\beta\hbar\omega_{i,\vec{k}}\left(n+\frac{1}{2}\right)} = e^{-\beta\frac{\hbar\omega_{i,\vec{k}}}{2}} \frac{1}{1-e^{-\beta\hbar\omega_{i,\vec{k}}}} \tag{2.52}$$

The Helmholtz free energy due to one phonon is

$$F_{\omega_{i,\vec{k}}} = -k_B T \ln\left(Z_{\omega_{i,\vec{k}}}\right) = \frac{\hbar\omega_{i,\vec{k}}}{2} + k_B T \ln\left(1-e^{-\beta\hbar\omega_{i,\vec{k}}}\right) . \tag{2.53}$$

The free energy is greater than 0 eV even at $T \to 0$ K due to the zero-point energy $\hbar\omega_{i,\vec{k}}/2$. As the phonons are not interacting with each other, the partition function part due to phonons for the whole system is

$$Z_{phonons} = \prod_i \prod_{\vec{k}} Z_{\omega_{i,\vec{k}}} \tag{2.54}$$

The free energy due to phonons is

$$F_{phonons} = -k_B T \ln\left(Z_{phonons}\right) \tag{2.55}$$

$$= \sum_{i,\vec{k}} \left(\frac{\hbar\omega_{i,\vec{k}}}{2} + k_B T \ln\left(1-e^{-\beta\hbar\omega_{i,\vec{k}}}\right)\right) . \tag{2.56}$$

The overall Helmholtz free energy is then

$$F = U_{pot} + F_{phonons} = U_{pot} + \sum_{i,\vec{k}} \left(\frac{\hbar\omega_{i,\vec{k}}}{2} + k_B T \ln\left(1-e^{-\beta\hbar\omega_{i,\vec{k}}}\right)\right) . \tag{2.57}$$

Here, U_{pot} is the potential energy of our structure in the ideal form. Now, we only need to sample the phonon frequencies sufficiently well. To compare to other free energy calculations we need to consider the difference in approach to phonons. For the quasi-harmonic calculations, we considered the phonons as independent quantum harmonic oscillators. However, MD simulations just solve Newton's Equation of motion. Hence, the system is purely classical. This means that to be able to compare to results gained from for example thermodynamic integration, we need to use the classic harmonic oscillator to approximate the phonons. The energy of the classic harmonic oscillator centred at $q = 0$ in one dimension in dependence of the spatial coordinate q and the momentum p is

$$E_{i,\vec{k}}(q,p) = \frac{1}{2} m \omega_{i,\vec{k}}^2 q^2 + \frac{1}{2}\frac{p^2}{m} . \tag{2.58}$$

Here, m is the mass of the oscillator, $\omega_{i,\vec{k}}$ is its angular frequency. The partition function then becomes the following:

$$Z_{\omega_{i,\vec{k}},\, classical} = \frac{1}{h} \int \int e^{-\beta \left(\frac{1}{2} m \omega_{i,\vec{k}}^2 q^2 + \frac{1}{2} \frac{p^2}{m} \right)} \mathrm{d}q \mathrm{d}p = \frac{1}{\hbar \omega_{i,\vec{k}} \beta} \tag{2.59}$$

The remaining steps are analogue to the steps for the quantum harmonic oscillator leading to the free energy for MD simulations of

$$F_{classical} = U_{pot} + \sum_{i,\vec{k}} \ln \left(\hbar \omega_{i,\vec{k}} \beta \right) \ . \tag{2.60}$$

Let us consider the free energy difference between a classical and a quantum oscillator. After a few steps one sees that for $T \to 0$, the difference becomes $\sum_i \hbar \omega_{i,\vec{k}}/2$, the sum of the zero-point energies of each oscillator mentioned before. We do however want to know which structures are dominant in the NPT ensemble. This means that we have to consider the Gibbs free energy of a structure. In the case of structures relaxed at a certain pressure P, the equation for the Gibbs free energy is well known:

$$G = F + PV \tag{2.61}$$

Here, V is the volume of the system and F is the Helmholtz free energy. In the case of a lattice which is a slightly strained version of a relaxed reference lattice with volume V_0, the following formula can be applied [77]:

$$G = F + PV - V \varepsilon'_{ij} \sigma'_{ij} \tag{2.62}$$

Here, ε'_{ij} are the elements of the deviatoric Cauchy's strain tensor and σ'_{ij} are the elements of the deviatoric stress tensor. In Equation 2.62, I sum over repeating indices. The elements of the deviatoric strain tensor ε'_{ij} and the deviatoric stress tensor σ'_{ij} are defined as

$$\varepsilon'_{ij} = \varepsilon_{ij} - \frac{V - V_0}{3V} \delta_{ij} \tag{2.63}$$

$$\sigma'_{ij} = \sigma_{ij} - \frac{P}{3} \delta_{ij} \ . \tag{2.64}$$

Here, ε_{ij} are the elements of Cauchy's strain tensor, σ_{ij} are the elements of the stress tensor and δ_{ij} is the Kronecker delta. We can sample over different possible lattices of a certain phase calculating the Gibbs free energy of the lattice with equation 2.62 provided we have a suitable reference structure so that the strain is sufficiently small. For example for a hcp phase

with lattice constants a and c, one could sample over the possible c/a ratios and volumes. Under the constraint that phase transitions are inhibited, the lattice with the minimum Gibbs free energy at the specific pressure and temperature is the one the phase will show. This means that the Gibbs free energy of a phase is the minimum Gibbs free energy of all possible lattices of that phase:

$$G(P,T) = min_{\text{lattice}}\left(G_{\text{lattice}}(P,T)\right) \tag{2.65}$$

Here, $min_{\text{lattice}}(...)$ denotes that we take the minimum overall possible lattices. The Gibbs free energy of a specific lattice is $G_{\text{lattice}}(P,T)$. As before, the stable phase is the phase with the lowest Gibbs free energy. This allows us now to describe solid-solid phase transitions at temperatures for which the quasi-harmonic approximation still holds and to predict *P-T* phase diagrams in this temperature realm.

2.8 Discussion of the use of the different methods of phase transition determination

In this section, I want to compare the different methods for determining phase transitions presented in this chapter.

Letting a phase evolve under MD simulations at certain temperatures and pressures does not work on metastable structures which could take an unknown and very long time to transform as explained in Section 2.1.2. However, as a simple MD simulation, the method is not very computationally expensive and finite size effects can be easily investigated. Additionally, it does not require that the investigated phase is metastable.

In the case of solid-solid transitions, setting up two competing phases in coexistence simulations (see Section 2.1.2) can lead to problems due to the interface. Due to different sizes of the crystals one may have to choose lattice constants which are not in accordance with the equilibrium lattice constants at the specific temperature and pressure. This can lead to artificially favouring one of the phases. This method also does not require a very large computational resources and can therefore be used to investigate transitions even for large number of atoms. For example, Smirnova *et al.* [78] conduct such simulations with their interatomic potential with 20250 atoms for the solid-liquid transition and with 16800 atoms for the transition between two crystal phases.

Thermodynamic integration (see Section 2.6) calculates the free energies of different phases which can then be compared to calculate the phase transitions. However, these phases need to be metastable as it is assumed that the phases do not transform during the MD simulations used for equilibration and sampling. Freitas *et al.* [48] calculated the free energy of systems

with up to 54000 atoms with thermodynamic integration which shows that very large systems can be investigated.
The free energies predicted by quasi-harmonic approximation (see Section 2.7) include quantum mechanical contributions that are ignored in the other methods mentioned in the current section. However, with increasing temperature, the assumption of non-interacting phonons becomes less valid making quasi-harmonic approximation increasingly unreliable. A phase with imaginary phonon frequencies for some phonon wave vectors cannot be calculated as the free energy would not be a real number anymore.
To calculate phase transitions with any of the above mentioned methods, one needs at least some information about the phases involved. This is not the case for nested sampling which predicts the phases by itself [27, 20]. As one does not need to provide the nested sampling algorithm with the phases involved in the transition, it is very useful for exploring unknown phase diagrams [27, 20, 28, 34, 21] and a possible method for fitting interatomic potentials [20]. Fitting interatomic potentials simultaneously to phase transition temperatures at various pressures or alloy compositions requires in principle the evaluation of unknown phase diagrams as a new iteration of the potential could predict a totally new phase. The other methods mentioned in this section are not as useful for the prediction of unknown phase diagrams – and in extension fitting interatomic potentials to experimental phase diagrams – as they could miss phases which one did not consider when setting up the calculations. This means that the phase diagram would be wrong, but one would not have any knowledge of this fact. Making matters worse, as explained above, these other methods do not work for all phases and have a different accuracy depending on the temperature. This means that one would have to use a combination of these methods. If one does not have much information of the different phases one would like to consider, one would not know which method to use for these. This means that a researcher would first need to spend significant amount of time testing which method might be applicable for each phase transition. However, nested sampling has high computational costs [21], which can make it hard to investigate finite size effects and would make the use of nested sampling even unnecessary if we could be sure about the phases involved in the phase diagram. Baldock *et al.* consider the effect of number of atoms up to 128 atoms in Ref. [47], but usually less than 100 atoms are used in the calculation of $P-T$ phase diagrams for metals and alloys [27, 20, 28, 34, 21]. The computational cost of nested sampling depends on the computational cost of evaluating the specific interatomic potential as well as on the transitions and systems in question [20, 28, 21]. Depending on the situation, an investigation of finite size effects might not be possible. Even if an investigation is possible, the error of the prediction for the thermodynamic limit may still be relatively large. This means that it is necessary to rely on informed search methods, like the other

methods of this section, as these can conduct investigations for larger numbers of atoms. The nested sampling results can indicate which phases need to be investigated. The choice of the method then depends on the phases involved. If the phases of interest are metastable in the temperature or pressure realm of interest, thermodynamic integration, quasi-harmonic approximation and coexistence MD simulations become candidates. The difficulty in creating a correct interface without stretching or compressing one phase inappropriately leads to errors. This can be avoided by the other two candidates, quasi-harmonic approximation and thermodynamic integration. For low temperatures, it would be advisable to use quasi-harmonic approximation as it covers quantum contributions like the zero-point energies. However, with increasing temperature, the assumption of phonons as non-interacting harmonic oscillators becomes less valid and the quantum mechanical contributions decrease in importance. Even though quasi-harmonic approximation does not immediately fail, thermodynamic integration will become more accurate and therefore preferable. Additionally, quasiharmonic approximation cannot calculate the free energy of structures with imaginary phonon modes which limits its usage. In the case of phases which are not metastable, we can consider the two MD simulations methods from the start of this section: Especially for solids, a simple MD simulation is more convenient to set up than an interface between phases.

2.9 Titanium phase diagram

There exist three well known titanium phases [44]: a bcc phase called β, an hcp phase called α and second hexagonal phase called ω [53]. Figure 2.1 shows the 2 atom cubic unit cell of the β phase marked by black lines. The darker blue spheres represent the atoms of the unit cell. The 2 atom unit cell of the α phase is displayed in Figure 2.2a in side view and from the top in Figure 2.2b. It is denoted with black lines. Blue and orange coloured spheres denote different layers. Darker colours in each layer signify the atoms of the unit cell. The c/a ratio is 1.588 at room temperature [79]. The ω phase has three atoms per unit cell [44]. The crystal structure and much detailed information about the ω phase can be found in Ref. [80]. It can be seen from the side in Figure 2.3a and from the top Figure 2.3b. As for the α phase, layers are distinguished by orange and blue colouring with darker colours in each layer denoting the atoms which are part of the unit cell. The unit cell is marked with black lines. The room temperature and 1 atm pressure c/a ratio is ca 0.608 according to the lattice constants given by Ref. [50]. There are indications of other phases at very high pressures [44]. Figure 2.4 depicts different experimental predictions for the equilibrium phase transition lines from the literature. Errandonea *et al.* [81] present 12 experimental melting temperature values for pressures up to ca 80 GPa. Their fit to these

presented in Ref. [81] (fit data from Ref. [82]), is denoted by an orange dashed line. The melting transition temperature curve fitted to experimental data by Stutzmann *et al.* [83] is given by a brown dashed line. I also present various experimental predictions for the solid part of the phase diagram with solid lines. The linear extrapolations of the α to ω transition lines are denoted by the dashed continuations. One sees the predictions according to Bundy [84] (cyan), Zil'bersteyn *et al.* [85] (magenta), Tonkov and Ponyatovsky [50] (blue) and Zhang *at al.* [51] (red). The experimental α to ω transition line slope found by Zhang *et al.* is 345 K/GPa [51]. This is very different from 93 K/GPa, the value reported by Tonkov [50]. Based on this new slope, Zhang *et al.* recalculate the phase diagram of Tonkov [86] as quoted by them [51] correcting the position of the triple point from 8.0 ± 0.7 GPa and 913 ± 50 K reported by Tonkov [86] as quoted by Zhang *et al.* [51] to 7.5 GPa and 913 K [51]. As explained in Ref. [44], an experimental consensus has been reached only for few parts of the solid part of the phase diagram. Especially, the onset of the ω-α transition is under discussion [44]. At room temperature, different experimental studies place the transition pressure in the relatively broad range from 2 GPa to 12 GPa [44].

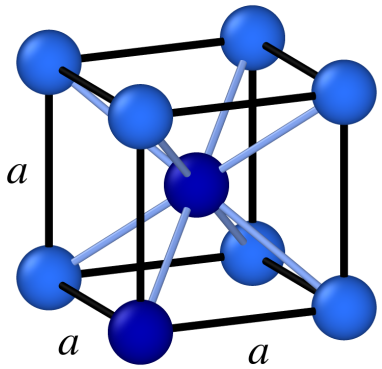
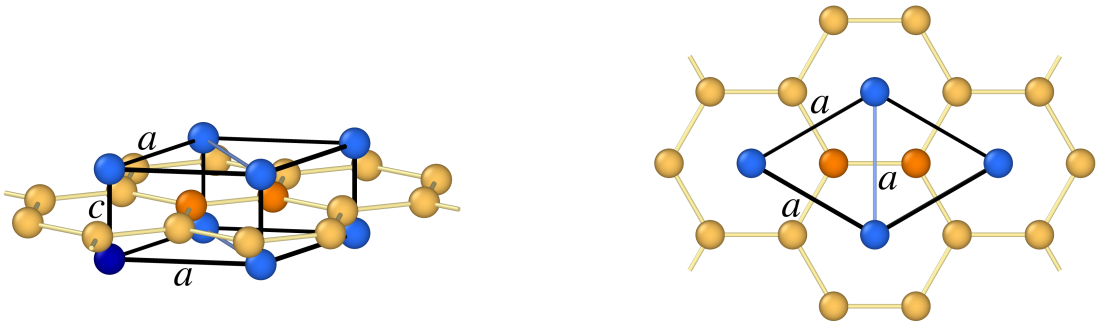


Fig. 2.1 Sketch of the titanium β phase unit cell (marked with black lines). The β phase has a bcc structure. The dark blue coloured atoms are part of the two atom cubic unit cell.

(a) (b)

Fig. 2.2 Sketches of the titanium α phase seen from the side (a) and from the top (b) lattices. The α phase has an hcp structure with a c/a ratio of 1.588 at room temperature [79]. The orange and blue coloured atoms denote different layers. The atoms coloured in a darker colours in each layer are part of the two atom unit cell which is denoted with black lines.



(a) (b)

Fig. 2.3 Sketches of the titanium ω phase lattice from the side (a) and from the top (b). The unit cell is marked with black lines. At room temperature, the ω phase has a c/a ratio of ca 0.608 based on the lattice constants given by Ref. [50]. As in Figures 2.2a and 2.2b, layers are distinguished by orange and blue colouring. The darker colouring in each layer denotes atoms within the unit cell.

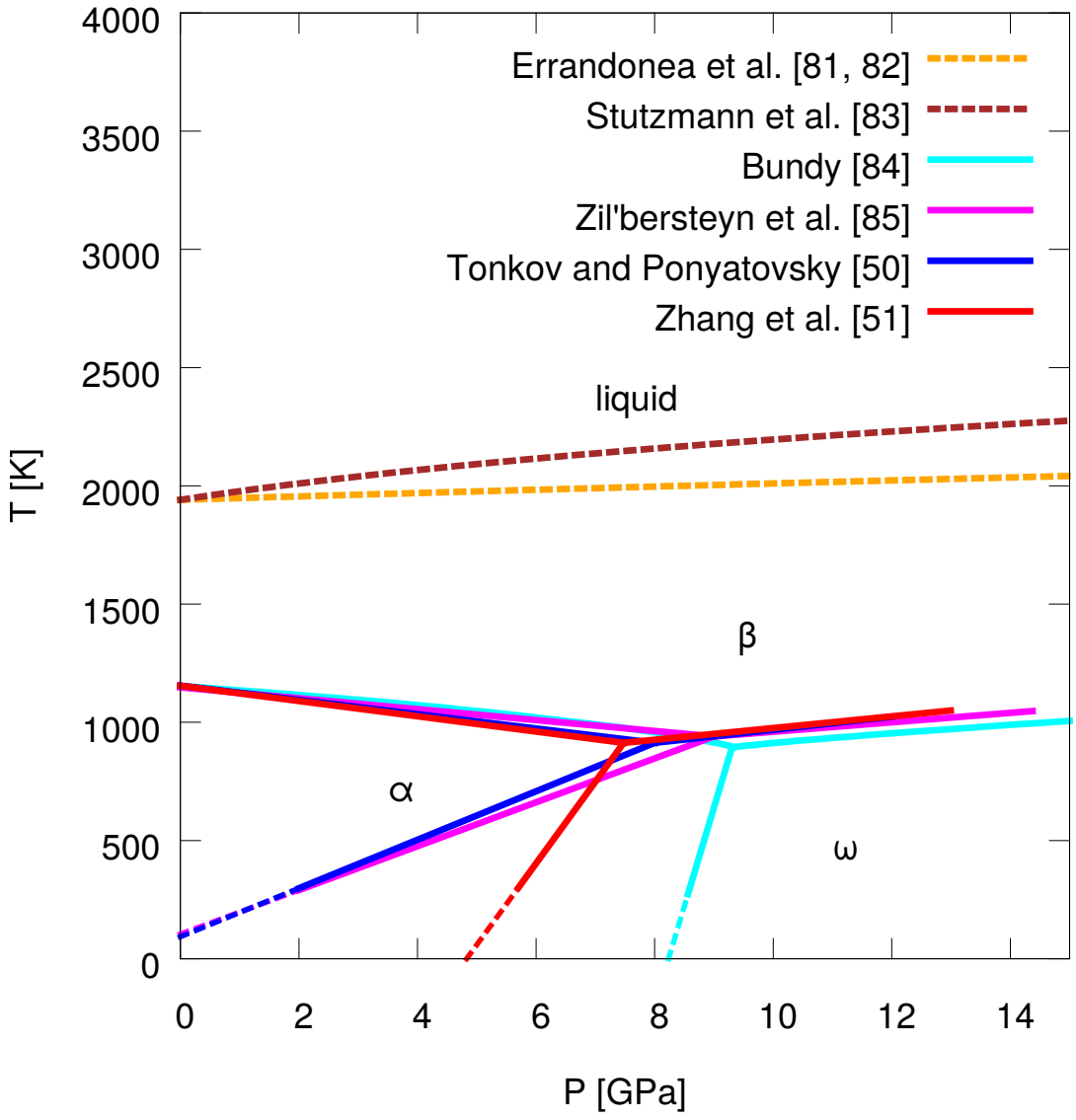


Fig. 2.4 Experimental titanium equilibrium phase transition line predictions. The melting line predicted by Errandonea *et al.* [81] provided via Ref. [82] is denoted by an orange dashed line. The melting line of Stutzmann *et al.* [83] is given by a brown dashed line. The other solid lines are predictions for the solid-solid transitions from Bundy [84] (cyan), Zil'bersteyn *et al.* [85] (magenta), Tonkov and Ponyatovsky [50] (blue) and Zhang *et al.* [51] (red). Dashed continuations of the solid-solid transition lines are linear extrapolations of the specific α to ω transition lines towards 0 K. The curves given for the melting are from fits to different experimental data [81–83]. The different titanium phases are labelled in the diagram. The slope of the α to ω phase is under discussion [51, 44].

2.10 The CuAu system – an example for eutectic behaviour

The CuAu phase diagram shows many different solid phases and an interesting melting curve [42]. The melting behaviour is shown in Figure 2.6. We see the liquidus and solidus temperature found by Bennett [87] quoted in Okamoto *et al.* [42] plotted against the atomic percentage of copper apart from the melting temperatures of pure Au and pure Cu which are from Ref. [88]. Above the liquidus temperature, the system is fully liquid. Conversely, below the solidus temperature, the system is fully solid. In between these two temperatures, solid phase and liquid phase coexist. The results for pure gold and copper are the ITS-90 definitions [88] and represented by magenta empty squares. The ITS-90, the "The International Temperature Scale of 1990", defines among others gold and copper as fixed points for the experimental determination of temperatures [88]. The solidus values found by Bennett [87] quoted in Okamoto *et al.* [42] are given by blue crosses. The liquidus values found by Bennett [87] quoted in Okamoto *et al.* [42] are denoted by downwards pointing orange triangles. The small dashes connecting the solidus points and the longer dashes connecting the liquidus points are only helplines. We can see the minimum resulting from the mixing at a Cu part of ca 44 % and at a temperature of ca 1183 K found by Bennett [87] quoted in Okamoto *et al.* [42]. This is an example of an eutectic system: The alloy melts at a lower temperature than its pure constituents. This is due to the different enthalpies of mixing for solids and liquids. More specifically, the enthalpy of mixing for the solid is greater than the one for the liquid [89]. In the opposite case when the enthalpy of mixing is smaller for solids than for liquids, we would see a maximum melting temperature for one of the intermediate alloy compositions [89]. The minimum is the so called "eutectic point". At this point the liquid transforms directly to the solid. However, if we either lower or increase the proportion of one of the alloy elements, one discovers regions where a liquid phase of the alloy coexists with a solid phase of the alloy. We see from Figure 2.6 that the phase of coexistence between the solidus and liquidus lines is however quite narrow in temperature for CuAu. The maximum difference between the solidus temperature and the liquidus temperature at any Cu percentage presented by Bennett [87] quoted in Okamoto *et al.* [42] is 23 K at 90.8 % atomistic percentage Cu. Directly below the solidus, CuAu at any atomic composition is in the fcc structure for which the species of the atoms is randomised, while it shows a variety of ordered phases at lower temperatures [42]. The fcc unit cell is shown in Figure 2.5. Dark green coloured atoms denote the atoms within the unit cell which is denoted with black lines.

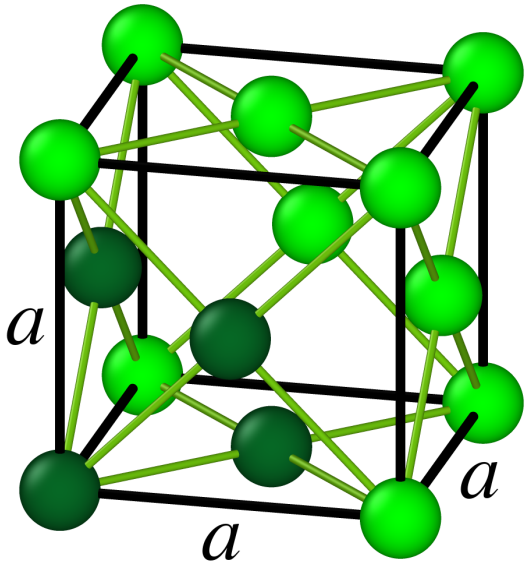


Fig. 2.5 The CuAu fcc unit cell below the solidus. The species of the atoms at each position is randomised throughout the lattice. Dark green atoms denote the atoms within the unit cell which is marked by black lines.

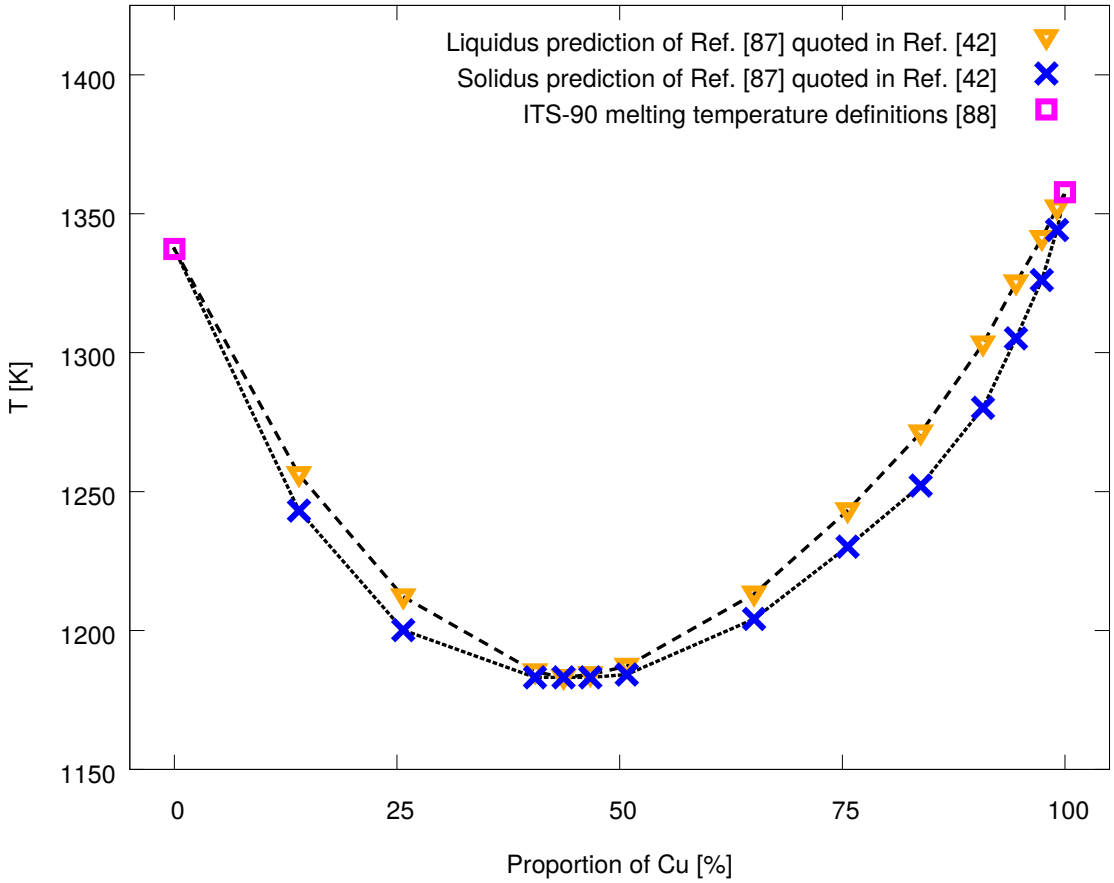


Fig. 2.6 The eutectic curve of CuAu according to experiments. The y-axis is the temperature in Kelvin. The x-axis denotes the atomic proportion of Cu in percent. Magenta empty squares denote the melting temperatures according the ITS-90 definition [88]. Downwards pointing orange triangles depict experimental values for the liquidus of Bennett [87] quoted in Okamoto *et al.* [42]. The solidus experimental results of Bennett [87] quoted in Okamoto *et al.* [42] are given by blue $\times$ symbols. The dashed black lines connecting the data points (short dashes for the solidus, longer dashes for the liquidus) are only a guide for the eyes.

Chapter 3

CuAu results

In this chapter, I present my work on improving a CuAu EAM [39–41] by fitting it with a Nelder-Mead algorithm [35] to reproduce the melting curve. The salient parameters for the nested sampling calculations are listed in Section A.3. I have elaborated on the approach in Section 1.2. The main idea is to start with a relatively easy target—fitting the pure Au melting temperature—and then increase the complexity by fitting anew to different transitions until the eutectic curve is sufficiently predicted by the fitted potential. However, before I cover the results of the fitting procedure for the interatomic potential, I present a number of technical details as well as some demonstration of the analysis of the nested sampling simulations. First, I present technical details of the EAM [39–41], of nested sampling with *pymatnest* [45] and of the analysis of the nested sampling calculation outputs in Section 3.1 to 3.4. Section 3.5 is a demonstration of identifying the phase below the melting temperature. Section 3.6 deals with the extrapolation of the melting temperature with respect to the number of atoms per simulation box in a nested sampling simulation. After covering details of the methodology of the actual fitting procedure in Section 3.7 to Section 3.10, we turn our attention to fitting of the pure Au part of the potential to the Au melting transition in Section 3.11. Section 3.12 covers experimental results used in the fitting considered after this section. Following this, the fitting of CuAu at a 50:50 ratio is covered in Section 3.13. Finally, in Section 3.14, the results for simultaneously fitting the melting temperatures of pure Au, pure Cu and CuAu at a 50:50 ratio are considered. I finish this section with a comparison of the latest fitted potential and original potential [39–41] as well as their performance on basic properties of the fcc structure for pure Cu and Au. The nested sampling parameters are based on the parameters used in Ref. [28] available in their supplementary material [43] and are only very slightly modified compared to theirs. I use a pressure of 0.1 GPa for all nested sampling runs presented in this chapter. To calculate a melting temperature, I use five independent nested sampling runs. If not denoted differently, I use the standard error of the

mean (SEM) from these five runs as its error. This error is then used in the error propagation and in the extrapolation procedure to estimate the standard errors (SE). However, five values lead to a large fluctuation in SE error estimate. To gain a more consistent estimate for the SE of the temperatures calculated with nested sampling runs (especially for determining the end of the fitting procedure), I use an average over the relative SEM gained from five independent runs from several slightly modified potentials. I discuss the validity of this in Section 5.1. I identify a disordered fcc phase just below the melting as expected from the experimental phase diagram [42] (see Section 2.10). My fitting procedure did not change this.

3.1 A short note on the CuAu potential used in the thesis

The CuAu EAM potential (referred to also as original potential) on which I base my fitting for CuAu presented in this thesis is itself a refit by Ward *et al.* [39] with a change to the forms of some functions of an earlier potential by Zhou *et al.* [41]. It was used in Ref. [28] to investigate the eutectic curve with nested sampling and can be found under Ref. [40]. The form of the potential is inspired by the EAM of Finnis and Sinnclair [90] which means that $\phi_{CuAu}(r) = \phi_{CuAu}(r)$, but $\rho_{a,CuAu}(r) \neq \rho_{a,AuCu}(r)$. This makes overall nine different functions which make up the potential.

3.2 Details of the *pymatnest* nested sampling implementation and basic terminology

My nested sampling calculations presented in this thesis are conducted with *pymatnest* [45] of Bernstein *et al.* which for my purposes relies on LAMMPS [74] to conduct the decorrelation steps. In this work, LAMMPS [74] is used to evaluate interatomic potentials for all nested sampling runs (see Section A.1 for more details on how interatomic potential are evaluated in this thesis). In the following, I want to elaborate on certain details of this implementation.

pymatnest [45] can conduct several kinds of decorrelation walks: an enthalpy based Markov chain Monte Carlo method (MCMC) applied on each atom separately [20], a Galilean Monte Carlo method (GMC) [91, 92] [28] and a Hamiltonian Monte Carlo method (HMC) [93] based method which considers the total enthalpy (TE-HMC) [28]. The first method conducts a random walk of a single atom and then evaluates the enthalpy. The

third method uses a MD walk to decorrelate the sample system, thereby moving all atoms at the same time. In the following, N is the number of atoms per simulation cell. Even when accounting for the fact that one would need N times as many MCMC steps to update every atom, it was shown that TE-HMC is more efficient in converging nested sampling [27, 28]. I therefore choose the TE-HMC options for all my Ti nested sampling calculations and keep it for my CuAu nested sampling parameters which are slight modifications of the nested sampling parameters used for CuAu [43] in Ref. [28]. The choice of HMC over GMC is justified by the indication that, compared to TE-HMC, GMC likely requires more computational resources for small number of atoms [28]. *pymatnest* offers to either conduct the decorrelation scheme in blocks with a set number of steps or without these blocks. These steps are a single MD step, volume step, cell shear step, cell stretch step and atom swap step (necessary for systems with different atom types). In the case of blocks, the number of each steps is in principle set while without blocks the moves are drawn at random. However, the ratio of the moves to each other stays on average ideally the same. Additional to the number of volume steps n_V, cell shear steps n_{shear}, cell stretch step $n_{stretch}$ and atom swap steps n_{swap}, one can also set the length of our MD trajectories L_{traj} and the number of MD trajectories per block. As mentioned before, even without using blocks, these numbers will then be used to determine the relative step ratios.

Again, n_{cull} denotes the number of walkers culled per iteration. In the following, I will use the expressions "calls per iteration". It is always implied that this is the number of calls per iteration per culled walker. In this case, "calls" denotes the total number L of potential calling steps. A MD trajectory of length 4 is counted as 4 calls. A cell step also counts as 1 call. The number of atoms is always counted per simulation cell.

One can also profile a process by setting the "profile" option to a number starting from 0 to $n_{proc} - 1$ where n_{proc} is the number of processes. *pymatnest* [45] allows us to analyse the resulting files via its script *ns_profile_analyse*. This is very helpful for estimating the wall time of a nested sampling run. It is also very useful for the investigation of the scaling with number of processors.

Let us now consider the output of a nested sampling run. The files with the suffix ".energies" incorporate the iteration number and the enthalpy of the respectively culled system. The data of the systems culled by process number #proc can be found in the files with the suffix ".traj.#proc.extxyz". These details include the atom position and cell vectors. The fraction of saved systems per file can be controlled via the input parameters. Similarly, there exist ".snapshot.#iter.#proc.extxyz" files where #iter represent the current iteration number. When the nested sampling run finishes or whenever a certain controllable iteration interval has passed, it saves the systems associated with each process to the respective files and deletes

the previous file. One can restart a nested sampling run from the iteration of the respective snapshot files by combining them to one file. The name of this file can then be used for the input parameter “restart_file”. For some runs I have used a *pymatnest* [45] version which I have modified slightly (more specifically the function “ns_run.py”) to reduce the output and to enable only the process number 0 to save the trajectories. This is however abandoned later. If not a single trajectory is used for analysis, I combine the trajectories via *ns_process_traj*, a script shipped with *pymatnest*.

3.3 Parallelization - implementation

The two methods for parallelizing nested sampling with *pymatnest* [45] have been explained in detail in Section 2.5.3. I now will consider their usage.
In all cases, *pymatnest* [45] decorrelates one walker per processor. I could confirm earlier findings [33] for the TE-HMC decorrelation that indicated that the number of calls per iteration can be kept relatively low for my calculations with the MEAM of Hennig *et al.* [15] and that increasing the number of walkers is a more important factor for convergence. This means, that the first method is limited as at some point the ratio $L/n_{walking}$ between total walk length L and number of walkers which are decorrelated per iteration $n_{walking}$ will become so small so that other parts of the program will dominate the computational cost. However, as we can see in Section 4.4, n_{cull} also has to be kept small to allow for convergence. The solution is a combination of both methods.

3.4 Identifying phase transitions

To identify phase transitions with nested sampling, I always calculate the heat capacity curve at constant pressure. Peaks in the heat capacity are usually a sign that a phase transition occurs. However, as for example demonstrated in Ref. [20], they may also indicate a Widom line [94, 95]. As I focus on solid-solid and solid-liquid transitions, this is not of concern for me. Due to the latent heat, the peak in heat capacity should be relatively clear for first order transitions. Second order transitions should show a discontinuity in the heat capacity [89].
As nested sampling is a statistical algorithm, the peak position has a statistical fluctuations. Hence, I estimate the transition temperature by averaging over a certain number of runs (usually five). The standard error of the mean calculated from these runs (assuming a Gaussian distribution for the fluctuations) is usually taken as an error estimate in this work. Note that the results need to be sufficiently converged for this. One important criterion is

that the independent runs for the same nested sampling parameters are relatively similar. The more similar the resulting curves become, the better is the convergence. However, it is possible to have similar runs which are not converged. For example, especially for a small number of runs, all runs may miss a transition if the number of walkers are too small [32]. The identification of the phases is done via XRD patterns. For this, I usually thermally average the trajectory samples with a weight from nested sampling. For this I created a program which is now part of *pymatnest* [45] (see Section A.1). This program has already been used to calculate thermally averaged XRD patterns to successfully identify phases in Ref. [21].
To average quantities like the XRD pattern via the nested sampling trajectory file, one must assign a probability to them. Let $i_{traj,current}$ be the iteration number of the system in question. $i_{traj,next}$ denotes the iteration number of the system following the current system in the trajectory file. Similar to Equation 2.33, I approximate the probability of the current system effectively as

$$p_{traj,current} \approx C \cdot (\zeta_{k_{current}} - \zeta_{k_{next}})\, e^{-\beta H_{traj,current}} \ . \tag{3.1}$$

Here $H_{traj,current}$, is the enthalpy of our current system. I also use $k_{current} = i_{traj,current} \cdot n_{cull}$ and $k_{next} = i_{traj,next} \cdot n_{cull}$ with n_{cull} equal to the number of culls per iteration. $\zeta_{k_{current}}$ and $\zeta_{k_{next}}$ are effectively the phase space volumes calculated during nested sampling. C is defined by normalising the sum over all probabilities of the trajectory systems to 1. Note that I assume that n_{cull} is considerably smaller than the interval at which culled walkers are saved.

3.5 A demonstration of phase transition identification for pure Cu

To demonstrate the identification of the fcc phase for the different CuAu composition calculations, I present thermally averaged XRD patterns averaged over 5 nested sampling runs for 36 Cu atoms at different temperatures in Figure 3.1 as explained in Section 3.4. This can only be done if the same phase is shown for all runs used in the averaging at the temperature which one considers. If a phase does not appear or appears at a different temperature, the mean over several runs cannot be used. (This averaging was not yet used for the pure Au fitting of Section 3.11, but it is usually used during the other fitting procedures presented in this Chapter.) I effectively use the original CuAu EAM [39–41] during the nested sampling runs (vertex 1 in Section 3.14). The temperatures for the XRD pattern predictions based on the nested sampling runs are 600 K (solid dark blue line), 700 K (solid cyan line), 800 K (solid light blue line) and 900 (solid orange line). I also compare it to the

XRD patterns of two reference structures: the fcc lattice (solid magenta line) and the ideal hcp lattice (dashed grey line). The volume used in the construction of the reference lattices is the thermally averaged volume at 600 K for run number 1 of the runs. The runs used for averaging are the same as used for calculating the Cu melting temperature of vertex 1 in the fitting of the melting curve in Section 3.14. The fcc XRD pattern becomes clearer with decreasing temperature.

3.6 Melting extrapolation

In the following, I investigate the dependence of the melting temperature predicted by nested sampling on the numbers of atoms N per simulation box. If not denoted differently, I use the original CuAu EAM [39–41]. I fit two functions to melting temperatures for different numbers of atoms to extrapolate the melting temperature for $N \to \infty$. To my knowledge, such fitting has not yet been conducted in this form in the literature, though Baldock *et al.* discuss the effects different numbers of atoms have on heat capacity peaks of phase transitions based on nested sampling calculation results in Ref. [47]. In Figure 3.2a, we see that the number of atoms has a strong effect on the melting temperature. The melting temperature from nested sampling T_m is shown in Kelvin against the number of atoms N. Error bars are statistical estimates from 5 independent runs. Red "+"-symbols show data which I will incorporate in an extrapolation scheme as I will explain in the following. The blue "×" symbols depict values which will be excluded from the extrapolation schemes. The fits from the different extrapolation schemes are signified with lines. Starting from 48 atoms, the melting temperature decreases with increasing number of atoms. However, this is not true for lower numbers of atoms. We see a decrease in melting temperature going from 24 atoms to 32 atoms. However, for 36 atoms the melting temperature increases again only to decrease for 40 atoms. For 48 atoms the melting temperature increases again. Following this, the melting temperature decreases monotonically. I attribute the aforementioned jumps in melting temperature for small numbers of atoms to finite size effects. The next step is the determination of the melting temperature value for $N \to \infty$. In the following I present two empirical ways of extrapolating the temperature values.

For numbers of atoms greater or equal than 48, I linearly fit an $k/N + h$ dependence as seen in Figure 3.2b. Here, I show $T_{m,nest}$ versus $1/N$. $T_{m,nest}$ is the average melting temperature of 5 nested sampling runs for a certain number of atoms N per simulation box. For $N \to \infty$, one gains an extrapolated value for the melting temperature $T_{m,\infty} = h$. We see that for $1/N \geq 1/40 = 0.25$, the behaviour is not in line with the fitted $k/N + h$ function. This is the realm of the jumps as mentioned before when describing Figure 3.2a. The fitting results are

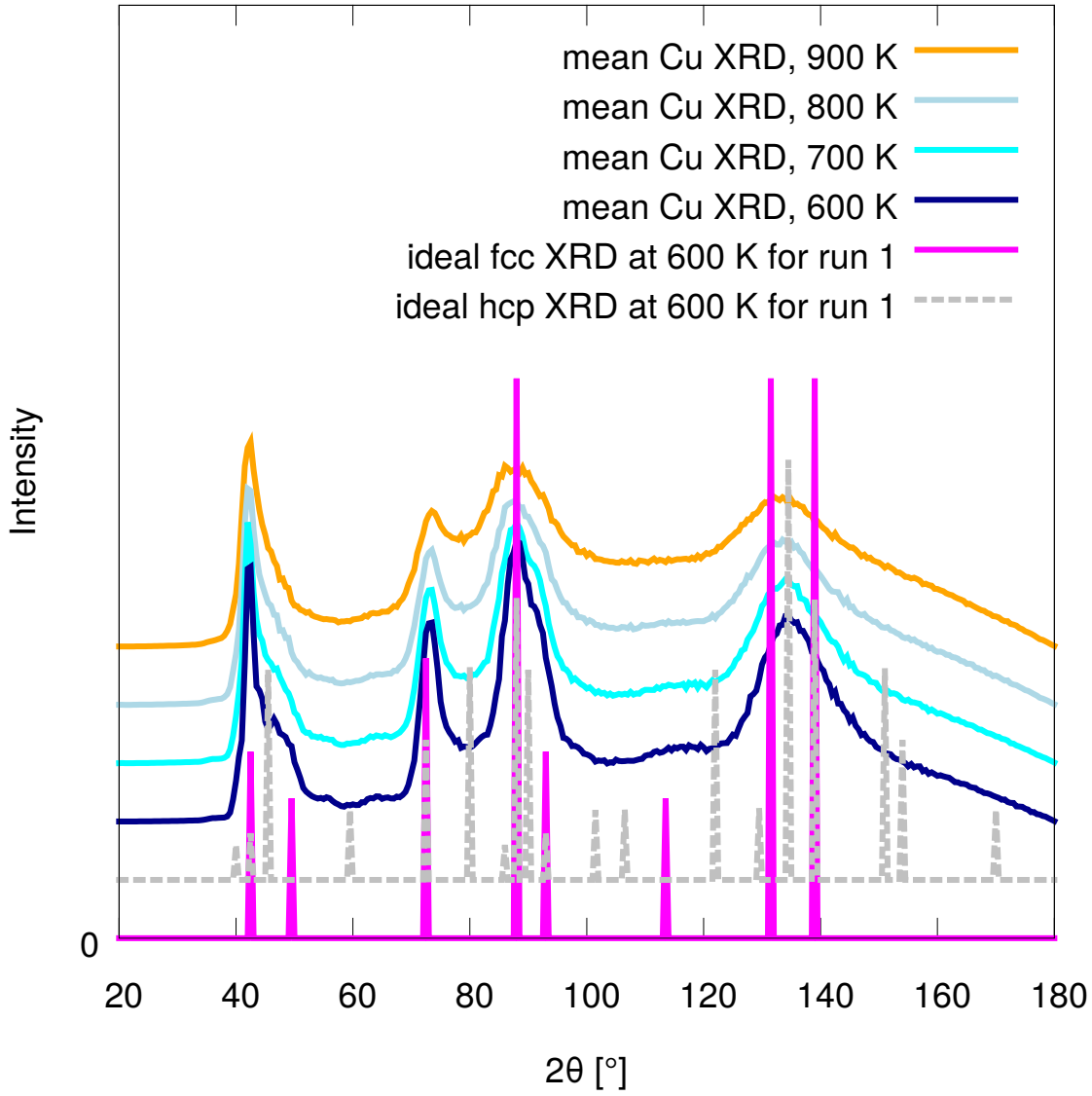


Fig. 3.1 Pure Cu XRD patterns averaged over 5 36 atoms nested sampling runs for 600 K (solid dark blue line), 700 K (solid cyan line), 800 K (solid light blue line) and 900 K (solid orange line). The runs are the same as the ones used for Cu for vertex 1 in Section 3.14. For pure Cu, these use effectively the original CuAu EAM [39–41]. The reference structures for the fcc reference XRD pattern (solid magenta line) and the hcp reference XRD pattern (dashed grey line) are calculated with the volume thermally averaged for 600 K for run number 1 of the 5 runs used for averaging. For the c/a ratio, I used the ideal value. We see that the fcc phase can be clearly identified.

$k = 12700 \pm 2900$ K and $h = 1052 \pm 52$ K.
In Figure 3.2c, we can see the same principle as for Figure 3.2b, just that $1/N^2$ is used on the horizontal axis. I fit $T_{m,extrapol}(N) = k/N^2 + h$. Again, temperature values for N below or equal 40 and hence values for $1/N^2 \geq 1/40^2 = 6.25 \cdot 10^{-4}$ are not part of the fitting. We see however that for N values greater or equal 48, which is equivalent to $1/N^2 \leq 1/48^2 \approx 4.34 \cdot 10^{-4}$, the data follows the linear fit, the relation $k/N^2 + h$, relatively well. The results are $k = 427 \pm 49$ K and $h = 1138 \pm 17$ K. We can see that the $k/N^2 + h$ fit has much lower relative errors that the $k/N + h$ fit. This indicates that $k/N^2 + h$ is in principle a better fit.

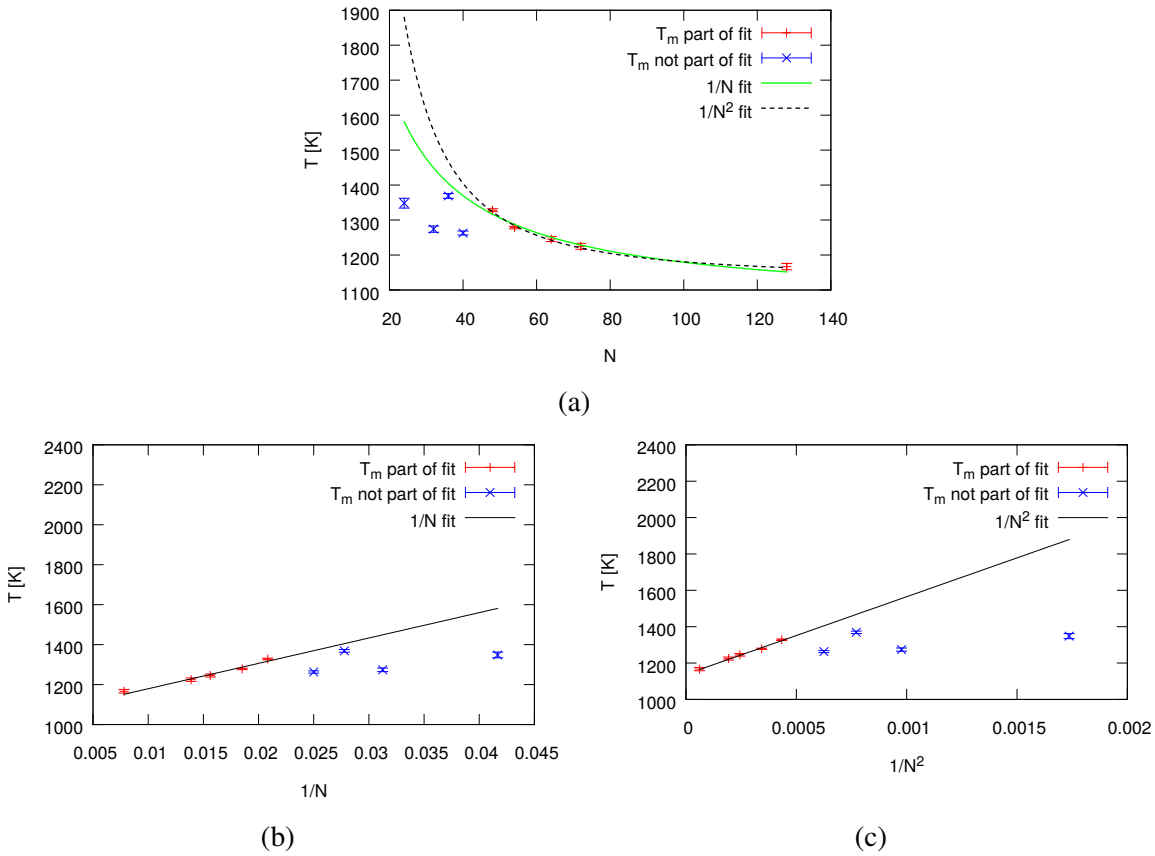
However, our potential fitting results which use the $k/N + h$ fitting from Sections 3.13.1 and 3.13.2 are still of use. The main focus of these potential fittings is to demonstrate that the fitting of potentials to experiments works in principle with my methodology. Recalculating the extrapolation of retroactively is relatively easy for single transition fitting as long as the resulting transition temperature is either always greater or smaller than the fitted value. This is conducted in Section 3.13.2.

We have seen how one can extrapolate a melting temperature from a number of nested sampling calculation. There is a relatively large finite size effect on the melting temperature calculated with nested sampling with different number of atoms N for the CuAu EAM [39–41] used in this thesis. Similarly strong dependence of the melting temperature and the boiling temperature were shown before for nested sampling for the Lennard-Jones potential [47]. However, it would be far too expensive to redo the aforementioned extrapolation for every vertex of our fitting algorithm. Instead, I make the assumption that, for a specific alloy composition, the ratio of the melting temperature predicted by a number of atoms N referred to as $T_{m,N}$ and the predicted melting temperature for $N \to \infty$, $T_{m,\infty}$, does not change much if one only slightly changes the parameters of the potential. This "extrapolation constant" is given by

$$k_{\text{extrapol},N} = \frac{T_{m,N}}{T_{m,\infty}} \tag{3.2}$$

and does depend on our extrapolation scheme as well as the number of atoms N used in the reference nested sampling calculation. The nested sampling simulations used to determine $T_{m,N}$ are independent from the ones used to fit the behaviour of the melting temperature with increasing number of atoms N. Indeed, the number of atoms used to calculate $T_{m,N}$ in this thesis is always lower than the minimum number of atoms used in the extrapolation. I am able to estimate the melting temperature predicted by a slightly modified potential by

$$T^{modified}_{m,\infty,estimate} = k_{\text{extrapol},N} \cdot T^{modified}_{m,N} \quad . \tag{3.3}$$



(a)

(b) (c)

Fig. 3.2 The melting temperature T_m in Kelvin predicted by nested sampling against the number of atoms N for pure Au (a), the inverse of the number of atoms $1/N$ (b) and the inverse of number of atoms squared $1/N^2$ (c) for pure Au. Blue "$\times$" symbols denote data which is not included in the fits. The data signified for red "$+$"-symbols is included in the fits. The different fits are given by lines as denoted in the figures. One can find a general extrapolation rule for $N \leq 48$ which corresponds to $1/N \geq 1/48 \approx 2.1 \cdot 10^{-2}$ in Figure b) and $1/N^2 \leq 1/48^2 \approx 4.34 \cdot 10^{-4}$ in Figure c). Finite size effects make the prediction increasingly difficult for very small numbers of atoms. I conduct linear fits for Figure b) and c), fitting the function $T_{m,extrapol}(N) = k/N + h$ and $T_{m,extrapol}(N) = k/N^2 + h$ respectively. The intersection of these fits with the vertical axis is the predicted melting temperature for $N \to \infty$. It is 1052 ± 52 K for the $1/N$ fitting and 1138 ± 17 K for the $1/N^2$ fitting. I use the original EAM [39–41] for these calculations.

Here, the superscript "modified" denotes that we consider the results of the slightly modified potential. This means that I only have to submit a batch of nested sampling calculations for a relatively small number of atoms per simulation box whenever I want to approximate the melting temperature of a slightly modified potential if I have already conducted an extrapolation for the unmodified potential for this transition. The assumption that the "extrapolation constant" does not change much for slight changes of the potential is supported by the data presented in Tables 3.1a and 3.1b. Here, I list various "extrapolation constants" for melting transitions with the original potential [39–41] and a modified version gained from fitting to the CuAu 50:50 melting transitions in Section 3.13.2 (vertex 28). Here, $k_{extrapol,original}$ denotes the constants gained from the calculations with the original potential and $k_{extrapol,modified}$ denotes the constants gained from the calculations with the modified potential. I use 36 atoms as the reference system for which I calculate $T_{m,N}^{modified}$ for all results presented in these tables. The two tables represent the different extrapolation schemes: Table 3.1a the $1/N$ scheme, Table 3.1b the $1/N^2$ scheme. The last column represents the difference between the extrapolation constants. For the $1/N$ extrapolation in Table 3.1a, the absolute value of the difference is smaller than the error for all systems. For $1/N^2$ in Table 3.1b, the errors of the differences as well as of the extrapolation constants are as expected to be smaller due to the smaller errors of $T_{m,\infty}$ due to the $1/N^2$ fitting as mentioned earlier. The absolute value of the difference is only smaller than the error for pure Cu. For CuAu 50:50 the absolute value of the difference is still smaller than twice the error. For Au however, with my current error estimate, the difference lies outside the 2 SE realm. I estimate the error for all melting temperatures used here from 5 nested sampling runs. However, as this error estimate is only based on 5 runs, it fluctuates considerably. If one were to take the averaged relative error from Section 3.14 for the 36 atoms calculations for the different transitions, the difference would be 0.016 ± 0.009. This would be within the 2 SE realm. Overall, I conclude that my assumption, that the extrapolation constant for a specific system stays the same for small changes to the potential, is usable for the fitting conducted in this chapter. Given the great disparity of my melting temperature predictions of the original EAM [39–41] from experimental data (see for example Ref. [96], Ref. [88], Ref. [97] and the results of Bennett [87] quoted in Okamoto *et al.* [42]) which is demonstrated for example in Figures 3.6 and 3.10 it is a suitable way to extrapolate the nested sampling calculations without conducting runs for different numbers of atoms for each new potential.

3.7 The objective function used in the fitting procedure

System	$k_{extrapol,original}$	$k_{extrapol,modified}$	$k_{extrapol,modified} - k_{extrapol,original}$
Au	0.757 ± 0.026	0.772 ± 0.02	0.015 ± 0.033
CuAu 50:50	0.731 ± 0.03	0.750 ± 0.038	0.019 ± 0.049
Cu	0.753 ± 0.03	0.719 ± 0.048	-0.034 ± 0.056

(a) "extrapolation constant" $k_{extrapol}$ for $1/N$ linear fits for different systems and EAM parameters as well as differences between them with regards to the modification of the potential.

System	$k_{extrapol,original}$	$k_{extrapol,modified}$	$k_{extrapol,modified} - k_{extrapol,original}$
Au	0.822 ± 0.006	0.838 ± 0.004	0.016 ± 0.007
CuAu 50:50	0.795 ± 0.013	0.821 ± 0.012	0.026 ± 0.017
Cu	0.806 ± 0.009	0.803 ± 0.015	-0.003 ± 0.018

(b) "extrapolation constant" $k_{extrapol}$ for $1/N^2$ linear fits for different systems and EAM parameters as well as differences between them with regards to the modification of the potential.

Table 3.1 The extrapolation constants for 36 atoms for different proportions of Cu and Au and for different potentials and extrapolation schemes. Additionally, the tables show the difference in extrapolation constants with regard to modification of the potential. I present the "extrapolation constants" for the original potential [39–41], $k_{extrapol,original}$, and the final accepted vertex of the CuAu 50:50 fitting from Section 3.13.2, $k_{extrapol,modified}$ (vertex 28). In Table (a), we see the extrapolation constants for the $1/N$ extrapolation scheme, while Table (b) portrays the results for the $1/N^2$ extrapolation scheme. The difference between "extrapolation constants" is within 1 SE for the $1/N$ extrapolation. For the $1/N^2$ extrapolation, the difference is within the 1 SE realm for pure Cu, within the 2 SE realm for CuAu 50:50 and just outside of the 2 SE realm for pure Au.

I choose the Nelder-Mead method [35] for fitting the potential. I have elaborated on this algorithm in Section 2.3. Firstly, I need an algorithm which does not need a derivative as these are not known. Secondly, I need an algorithm which is sufficiently stable when the data has a certain amount of statistical fluctuation, as all our phase transition temperature predictions have a certain statistical error (see Section 3.4). As one can see from the algorithm steps which are shown in Section 2.3, the order of the objective function values ordering is the critical part of the Nelder-Mead algorithm [35]. It deals relatively well with random noise, provided that this noise is not too strong, exactly due to this feature [63].
Inspired by least squares fitting, I choose the square root of the sum of quadratic differences as the objective function for the case of fitting N phase transitions:

$$f(\vec{x}) = \sqrt{\sum_{i}^{N} \left(T_{i,exp} - T_{i,extrapol}(\vec{x})\right)^2} \tag{3.4}$$

Here $\vec{x}$ represents the parameters of the model or the change from the original parameters of the model. *exp* denotes experimental values while *extrapol* denotes values extrapolated from

the nested sampling calculations. The sum is taken over the N phase transitions individually counted by the index i. These phase transitions i can be of different kinds, for example solid-solid or solid-liquid. Other parameters like alloy compositions or pressure can be changed, as well. In principle, one could also not use a squared root and just use the sum as for example done in Ref. [98] for a different kind of fitting data. The Nelder-Mead algorithm does not use any derivatives but only relies on the ordering of the objective function values [35]. Additionally, the function $\sqrt{x}$ is a strictly increasing function for values $x \geq 0$. Hence, the use of the square root does not make a difference as the order of objective functions of the different vertices is not changed. My choice of objective function can however be interpreted as a difference in temperature fitting space to the set of experimental temperatures. This viewpoint is also closer to my geometric interpretation of the stopping criterion elaborated on in Section 3.8.

3.8 Ending the Nelder-Mead algorithm

In this thesis, I use the abbreviation "SE" for the "standard error" and "SEM" for the "standard error of the mean". The standard error of the extrapolated melting temperature is an important part of my criterion for ending the Nelder-Mead algorithm [35] as I will explain in this Section.

In the original Nelder-Mead paper [35], a limit h_{lim} is chosen and the average $\overline{f}$ over all objective functions f_i of the current simplex is calculated. Nelder and Mead terminate the algorithm as soon as the estimate of the standard deviation σ_f for the distribution of objective functions fulfils the following condition [35]:

$$\sigma_f = \sqrt{\frac{\sum_i \left(\overline{f} - f_i\right)^2}{N}} < h_{lim} \tag{3.5}$$

Here we sum over all vertices i of the simplex. N is the number of dimensions of our problem. The overall number of vertices per simplex is $N+1$.

As my objective function values have a statistical fluctuation, I choose a different scheme. If one fits only one transition temperature, the objective function of Equation 3.4 becomes just the absolute value of the difference between experimental value and extrapolated transition temperature. The statistical error of the difference is only due to the calculation of the transition temperature with nested sampling. If the aforementioned absolute difference is equal to two statistical standard errors of the difference, I consider the extrapolated transition

temperature not to be significantly different from the experimental transition temperature. Hence, the algorithm is stopped in this case. For the number of transitions $N > 1$, we can consider the Mahalanobis distance [99]. The extrapolated transition temperatures are independent. If we now assume that the estimate for the mean transition temperature for each transition follows a Gaussian distribution with the SE which I calculate being equal to the standard deviation of this distribution, we get the following function for the Mahalanobis distance [100]:

$$d_M(\vec{x}) = \sqrt{\sum_i^N \left(\frac{T_{i,exp} - T_{i,extrapol}(\vec{x})}{\Delta T_i} \right)^2} \tag{3.6}$$

Here, the notation is the same as for Equation 3.4 and ΔT_i is the statistical error of $T_{i,extrapol}(\vec{x})$. For a multidimensional Gaussian, a Mahalanobis distance of 2 is equivalent to 2 standard deviations in one dimension [100]. Hence, I decide to accept a vertex as the final vertex if it is the first vertex for which $d_M < 2$ is fulfilled. If we use the condition $d_M < 2$ together with Equation 3.6, we arrive at the following equation which shall be my requirement to stop the Nelder-Mead algorithm:

$$2 > \sqrt{\sum_i^N \left(\frac{T_{i,exp} - T_{i,extrapol}(\vec{x})}{\Delta T_i} \right)^2} \tag{3.7}$$

Given that

$$\left(\frac{T_{i,exp} - T_{i,extrapol}(\vec{x})}{\Delta T_i} \right)^2 \geq 0 \ , \tag{3.8}$$

this is equivalent to

$$1 > \sum_i^N \left(\frac{T_{i,extrapol}(\vec{x}) - T_{i,exp}}{2\Delta T_i} \right)^2 \ . \tag{3.9}$$

In the following, the i-th element of the vector $\vec{T}_{extrapol}$ is defined as $T_{i,extrapol}$. The vector $\vec{T}_{exp}$ is defined equivalently. Let us now consider the N dimensional coordinate frame of all considered transitions. For $N = 2$, Equation 3.9 requires that the point defined by $\vec{T}_{extrapol}$ lies within an ellipse centred at the point defined by $\vec{T}_{exp}$. In this interpretation $2\Delta T_1$ and $2\Delta T_2$ are, depending on which one is larger, the semi-minor and semi-major axis of this ellipse. These axes are oriented parallel to the x-axis and y-axis respectively. Similarly, for $N = 3$, Equation 3.9, requires that $\vec{T}_{extrapol}$ lies within an equivalently defined ellipsoid.

3.9 Details of my own Nelder-Mead implementation

The Nelder-Mead algorithm [35] is used for fitting the melting temperatures of melting transitions for different alloy compositions. Even for calculating only the melting temperature of one of the CuAu compositions, a converged nested sampling run which samples down to a temperature sufficiently low to identify the solid phase below the melting can take very long, for example over 6 hours with the parameters employed in Section 3.14 for 36 atoms. To this, one also has to add a waiting time on the high performance computing cluster which fluctuates depending on the time of day and date of submission. This means that I cannot just use an existing implementation of the Nelder-Mead method and just calculate the objective function via nested sampling. Instead, I have created my own implementation of the Nelder-Mead method. The iteration is split in various substeps. If the objective function needs to be evaluated, the relevant submission files are created and nested sampling simulations are submitted. Then the program ends. These nested sampling simulations are conducted with "pymatnest" [45]. Once the nested sampling simulations have finished, the user needs to start the Nelder-Mead program again. The analysis is conducted in a similar fashion by my implementation of the Nelder-Mead method. During the analysis the heat capacity curves are calculated with the heat capacity calculation script of "pymatnest" [45]. XRD patterns are calculated with the XRD pattern creation script of "pymatnest" [45] which I created (see Section 3.4 and Section A.1). The XRD patterns are calculated for temperatures of the user's choice. Reference structures are automatically calculated as well. I use fcc and bcc as reference structures for the fitting presented in Section 3.11. For the other fittings presented in Chapter 3, I also use an ideal hcp structure. *Matplotlib* [101] is used to show the heat capacity curves and the XRD patterns. The user needs to confirm that the curves and the XRD patterns are acceptable before the algorithm can proceed. This is done to discover any convergence problems of nested sampling or the emergence of an unexpected phase during the fitting. The algorithm can not yet fit potentials if the transition it is supposed to fit is not predicted anymore by the modified potential. I also use various functions which I created from the collection of miscellaneous functions *misc_calc_lib.py* [45] (see Section A.1). I tested the Nelder-Mead part of my implementation using the Rosenbrock function [102] implementation of *SciPy* [103] against the *SciPy* [103] implementation of the Nelder-Mead algorithm. I chose the Rosenbrock function as Nelder and Mead demonstrate the behaviour of the Nelder-Mead algorithm among others with it [35]. During this process, new potential files need to be created and the original CuAu potential [39–41] file needs to be read in. The reading and writing in the required format, "eam/fs" [104], is done with the help of *matscipy* [105]. The modification of the potential parameters is explained in Section 3.10.

3.10 Fitting procedure

The CuAu potential [39–41] has 9 different functions with $\rho_{a,CuAu}(r) \neq \rho_{a,AuCu}(r)$ but with $\phi_{CuAu}(r) = \phi_{AuCu}(r)$ (see Section 3.1). I simplify the potential for my fitting by replacing $\rho_{a,CuAu}(r)$ with $\rho_{a,AuCu}(r)$ as soon as I start with a modification of the functions. In principle, one could refit the different parameters of the CuAu potential [39–41] functions [39, 41]. However, as the fitting is supposed to be a proof of principle, the modifications are kept simple. Hence, I only changed the parameters by multiplying the functions $U_{Au}(\rho_h)$, $\rho_{a,AuAu}(r)$, $\phi_{AuAu}(r)$, $U_{Cu}(\rho_h)$, $\rho_{a,CuCu}(r)$, $\phi_{CuCu}(r)$, $\rho_{a,AuCu}(r)$ and $\phi_{AuCu}(r)$ with parameters $x_{i,n}$. The index n in $x_{i,n}$ denotes the vertex number. The index i stands for the coordinate dimension. This defines the vector of the vertex $\vec{x}_n$. The mapping of the original functions [39–41] to the new once of vertex n is as follows:

$$\begin{pmatrix} U_{Au,n}(\rho_h) \\ \rho_{a,AuAu,n}(r) \\ \phi_{AuAu,n}(r) \\ U_{Cu,n}(\rho_h) \\ \rho_{a,CuCu,n}(r) \\ \phi_{CuCu,n}(r) \\ \rho_{a,AuCu,n}(r) \\ \phi_{AuCu,n}(r) \end{pmatrix} = \vec{x}_n^T \begin{pmatrix} U_{Au,orginal}(\rho_h) \\ \rho_{a,AuAu,orginal}(r) \\ \phi_{AuAu,orginal}(r) \\ U_{Cu,orginal}(\rho_h) \\ \rho_{a,CuCu,orginal}(r) \\ \phi_{CuCu,orginal}(r) \\ \rho_{a,AuCu,orginal}(r) \\ \phi_{AuCu,orginal}(r) \end{pmatrix} \tag{3.10}$$

Here, the subscript "*original*" denotes that the function is the function of the original potential [39–41]. The index n denotes that this is the function for the potential of vertex n. The superscript T denotes a transposition of the vector $\vec{x}_n$. For fitting pure Au, only the first three elements of the vectors are used. The original potential file [39–41] is used as basis for definition of the functions. It approximates the parameters via spline nodes which I just multiply to create the new potential. The parameter file format is "eam/fs" [104]. In general, to write to the CuAu potential file and to extract the function values of the functions which make up the potential, I use *matscipy* [105] (specifically, the "calculators" part). This is for example used in my implementation of the Nelder-Mead algorithm (see Section 3.9) and for the plotting of the functions in Section 3.14.1.

3.11 Au fitting

Let us now consider the first fitting trial. The complexity of the fitting will be increased in later sections. The Nelder-Mead fitting scheme is tested for simulations with 24 atoms by

fitting the melting temperature of pure gold through modifying only the functions $U_{Au}\left(\rho_{h,Au}\right)$, $\rho_{a,Au}(r)$ and $\phi(r)$ (see Section 2.2.2). Here, the respective functions are multiplied with their modification parameter as explained already in Section 3.10. Based on the results from Section 3.6 presented in Figure 3.2b, one arrives at $k_{extrapol,original}(N=24)=0.78\pm0.04$. The constant is based on a $1/N$ extrapolation. The reference melting temperature value for $N=24$ is 1348 ± 14 K. I use a value of 0.78 during the Nelder-Mead algorithm. I also use 1344 K as experimental value which I aim to fit. This is based on the formula for the melting temperature T_m from Ref. [96]:

$$T_m = 1339\left(1+\frac{P}{16.1\,\mathrm{GPA}}\right)^{0.57} \tag{3.11}$$

Here, P is the pressure. My nested sampling calculations are conducted at $P\approx0.1$ GPa which using Equation 3.11 leads to a melting temperature of ca 1344 K. The temperature one aims to fit is hence chosen as $T_{aim}=1344$ K. Possible errors from the experiments, from the fitting of Equation 3.11 as well as from the rounding of T_{aim} will be much smaller than the errors in extrapolating melting temperature for the potentials. The relative estimated error for the "extrapolation constant" used in this Nelder-Mead fit alone is ca 5 %. For a temperature extrapolated via this constant from nested sampling runs which has the same value as the above mentioned experimental melting temperature of 1344 K, the error due to the extrapolation constant alone is ca 68 K.

First trial runs of the scaling of $\phi_{Au,Au}(r)$ led to the conclusion that multiplying the function $\phi_{Au,Au}(r)$ with 1.04 is a good first modification. As first vertex of my Au melting temperature fitting, I take the original potential [39–41]. The other vertices are chosen so that each represents the multiplication of a different function by a factor of 1.04. The resulting starting simplex (iteration number $i=0$) is

$$S_0 = \left\{ \begin{pmatrix}1.0\\1.0\\1.0\end{pmatrix}, \begin{pmatrix}1.04\\1.0\\1.0\end{pmatrix}, \begin{pmatrix}1.0\\1.04\\1.0\end{pmatrix}, \begin{pmatrix}1.0\\1.0\\1.04\end{pmatrix} \right\} . \tag{3.12}$$

We see the fitting history in Figure 3.3. Here, the horizontal axis shows the iteration number i. The vertical axis represents the objective function values of the simplex after the i-th iteration. For the current procedure, the objective function simplifies to the absolute value of the difference between experimental value and the melting temperature extrapolated using $k_{extrapol,original}(N=24)$. The results for the aforementioned starting simplex are shown for iteration $i=0$. The algorithm needs only 3 iterations to arrive at a difference between the aimed temperature T_{aim} and the extrapolated temperature of vertex number 9 of

$T_{aim} - T_{m,extrapol,vertex\ 9} = -9 \pm 14$ K. The objective function is hence ca 9 K. The error here is based on using an average of the statistical relative errors of the melting temperature of all vertices calculated in this run and then multiplying this with the particular extrapolated melting temperature. This difference already fulfils the stopping condition as the difference between experimental value to which we fit and the extrapolated melting temperature is even lower than 1 SE. The modification parameters of the accepted vertex are

$$\vec{x}_{converged} = \begin{pmatrix} 1.0133333333333332 \\ 0.8666666666666661 \\ 1.1333333333333335 \end{pmatrix} . \tag{3.13}$$

In Figure 3.3, one also notices two expansion steps in iteration 1 and 3 which were both successful. They are responsible for the large decreases in the objective function value. The reflection in iteration 2 leads to a relatively low value but fails to trigger an expansion as it still slightly higher than the previous minimum. Using a not yet rounded value of the extrapolation constant, the predicted melting temperature for the model of vertex 9 is $T_{m,extrapol,vertex\ 9} = 1350 \pm 70$ K (rounded to closest 10 K). Here, I use the averaged relative error mentioned above for the melting temperature of vertex 9 as well as the error of the extrapolation constant (without any use of averaging) to calculate the overall error. After this successful trial refitting of the Au part of the original potential [39–41], I can turn my attention to fitting the full interatomic potential which I cover in the following sections.

3.12 Melting temperature values used in CuAu potential fittings

For the melting temperature fitted to for the CuAu 50:50 melting fit of Section 3.13, I used an estimate based on the data of Bennett [87] quoted in Okamoto *et al.* [42]. The closest alloy composition value of Bennett [87] quoted by Okamoto *et al.* [42] to a 50:50 CuAu atomic composition is a composition with 50.8 % Cu with a solidus temperature of 1184.15 K and a liquidus temperature of 1187.15 K. I take these as CuAu 50:50 values and estimate the melting temperature nested sampling would predict as the mean of both. This leads to a melting temperature of 1185.65 K. For the fitting I use the rounded value of 1186 K. Without any pressure stated for this data in Ref. [42], I assume that the pressure is ca 1 atm. I estimate the error of this melting temperature estimate to be half of the solidus and liquidus difference of the data of Bennett [87] quoted by Okamoto *et al.* [42] at 50.8 % Cu leading to an error of 1.5 K. For the simultaneous fitting of the pure Au, pure Cu and CuAu 50:50 melting

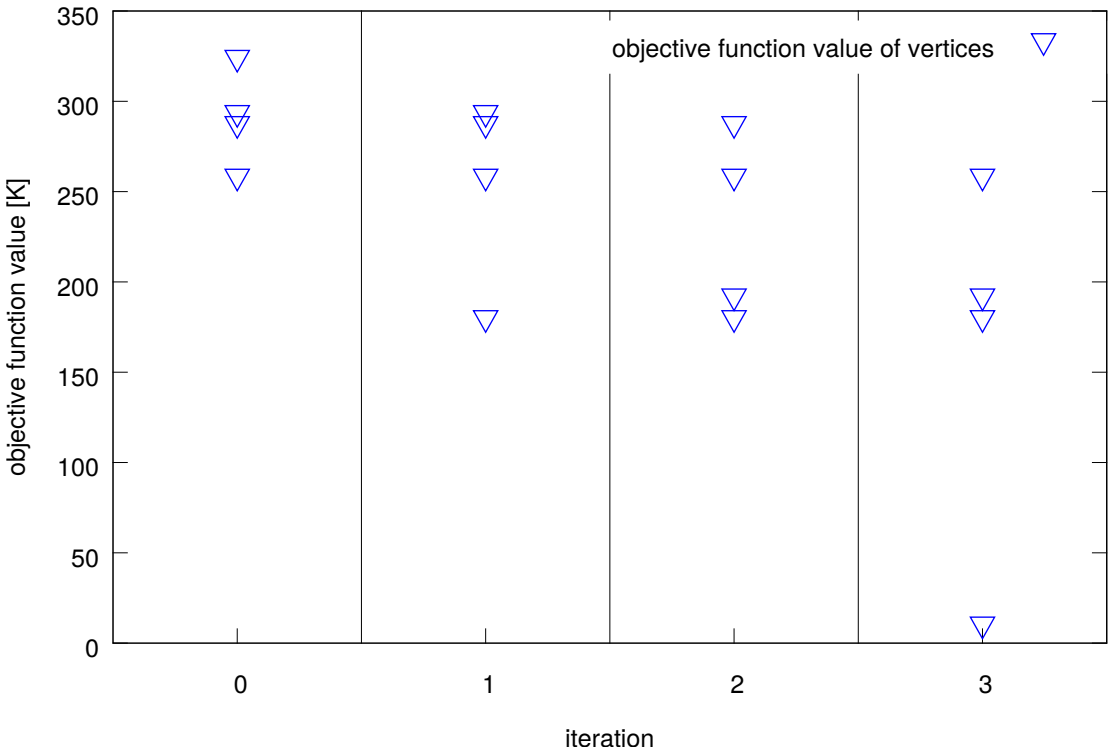


Fig. 3.3 The objective function values for the different vertices versus the iteration number i for fitting the pure Au melting temperature with 24 atoms. The iteration number i denotes the iteration which leads to the particular simplex. For $i = 0$, the objective functions values of the initialisation simplex are listed. In only 3 iterations, the algorithm arrives at a suitably low objective function value. We see the results of two successful expansion steps for iterations 1 and 3 which decrease the objective function value significantly. Iteration 2 does not lead to a new minimum in the reflection step, hence an expansion is not conducted. The expansion in iteration 3 leads to a objective function value of ca 9 K. The error of the extrapolated melting temperature of this vertex is 14 K as explained in the main text (using the average of relative SEMs of this fitting run). This puts the difference between aimed temperature and extrapolated temperature – the absolute value of this difference is the objective function in this case – clearly in the 1 SE realm of its own error. Hence the algorithm is stopped.

transitions of Section 3.14, I use the same estimate for the CuAu 50:50 melting transition as for fitting only the CuAu 50:50 melting in Section 3.13. For consistency, I also base my values for the melting temperature of pure Cu and pure Au of values under a pressure of 1 atm. The melting temperature of pure Au is then 1337.33 K under the ITS-90 definition [88] and the melting temperature of Cu is then 1357.77 K under the ITS-90 definition [88]. The melting temperature which I use in the actual fitting for Au is the rounded ITS-90 value for Au presented above, 1337 K [88]. The Cu value used in the fitting is 1356 K, the rounded value of the Cu melting temperature presented in Ref. [97]. Using 1358 K, the rounded value of the ITS-90 definition [88], should in general have only a minimal influence due to the large errors of the extrapolated melting temperature. Indeed, adjusting the objective functions appropriately shows that in the specific fitting run conducted for this thesis, the final vertex would have not been any different if I had used 1358 K, the rounded value of the ITS-90 value for Cu [88].

3.13 CuAu 50:50 fitting

3.13.1 24 atoms CuAu 50:50 fitting

Given the good results in Section 3.11, I now turn to fitting the CuAu potential to the CuAu 50:50 melting point which is close to the experimental eutectic point found for a ratio of 43.7:56.3 at a temperature of ca 1183 K by Bennett [87] quoted in Okamoto *et al.* [42]. Now, one needs to consider all eight different functions of the EAM as opposed to only three for pure gold. I use 1536 calls instead of 768 calls per iteration which are used for the pure Au fitting of Section 3.11 to ensure that the CuAu 50:50 results are converged. Given the success with the Au fitting, I use the same idea of multiplying the different functions by a factor of 1.04 to create my initial simplex as in Section 3.11. This initial simplex S_0 (iteration number $i = 0$) consists out of the first 9 vertices. This means that all vertices $\vec{x}_{1+l}$ with $l \in \{1,2,...,7,8\}$ have all elements equal to 1.0 except the l-th element which is set to 1.04. Vertex 1 is however the original potential with $\rho_{a,CuAu}(r) \neq \rho_{a,AuCu}(r)$ [39–41]. All other vertices have $\rho_{a,CuAu}(r) = \rho_{a,AuCu}(r)$ (see Section 3.10). The extrapolation constant is rounded to 0.73 and based on a $1/N$ extrapolation (see Table 3.1a). The Nelder-Mead method implementation is run until iteration 14. After this point, the algorithm is stopped as insufficient progress is made.

Figure 3.4 shows the objective functions values for all vertices of a simplex i after the new vertex was accepted in iteration i against the iteration number i. The average over all relative SEM of the nested sampling melting temperatures calculated for all vertices occurring in

this fitting run is ca 0.020. This leads to estimated statistical errors of the extrapolated transition temperature of ca 19 K to 21 K. To analyse the behaviour of the fitting, I calculate the difference between the minimum and the maximum melting temperature for each simplex. I use the above mentioned average relative error to estimate the error of these differences. The maximum difference occurs for the starting simplex ($i = 0$) with 81 ± 28 K. The minimum difference is 32 ± 29 K after iteration 14. The vertex accepted after iteration 2 is the last simplex for which the difference does not lie within 2 SE. I interpret it as a sign that the algorithm struggles to distinguish between the points. For example, in a single simplex, the vertex with the minimum objective function value and the vertex with the maximum value could have actually switched positions in their ordering with respect to the objective function value if one repeated the respective simulations. The Nelder-Mead algorithm however relies on the ordering of the vertices with respect to the objective function, hence small statistical errors can be tolerated but large ones can lead to problems [63]. This may be the reason why our Nelder-Mead method implementation converges very slowly compared to the fitting of pure Au in Section 3.11. To reduce the statistical error of the objective function, the algorithm is restarted with a higher number of atoms.

3.13.2 36 atoms CuAu 50:50 fitting

As mentioned before, to decrease the statistical error for the objective function, the number of atoms per simulation box are increased to 36. All other nested sampling parameters stay the same. The new extrapolation ratio for 36 atoms is 0.73 (see Table 3.1a). I also keep my starting vertices but not the melting temperatures from the 24 atom CuAu 50:50 fitting. In Figure 3.6, I present again the minimum objective function value of the vertices of the simplex i against the iteration number i. After iteration 16 the algorithm is stopped with a value of ca 7 K for vertex number 28 which satisfies the convergence criterion as explained below. The corresponding vertex is given by the following vector:

$$\vec{x}_{28} = \begin{pmatrix} 1.032057742628968 \\ 0.828184436903102 \\ 1.0615115089749447 \\ 0.9760511015017979 \\ 1.0180415643786564 \\ 1.0909789009427189 \\ 0.9828558369015812 \\ 1.0758726997708434 \end{pmatrix} \tag{3.14}$$

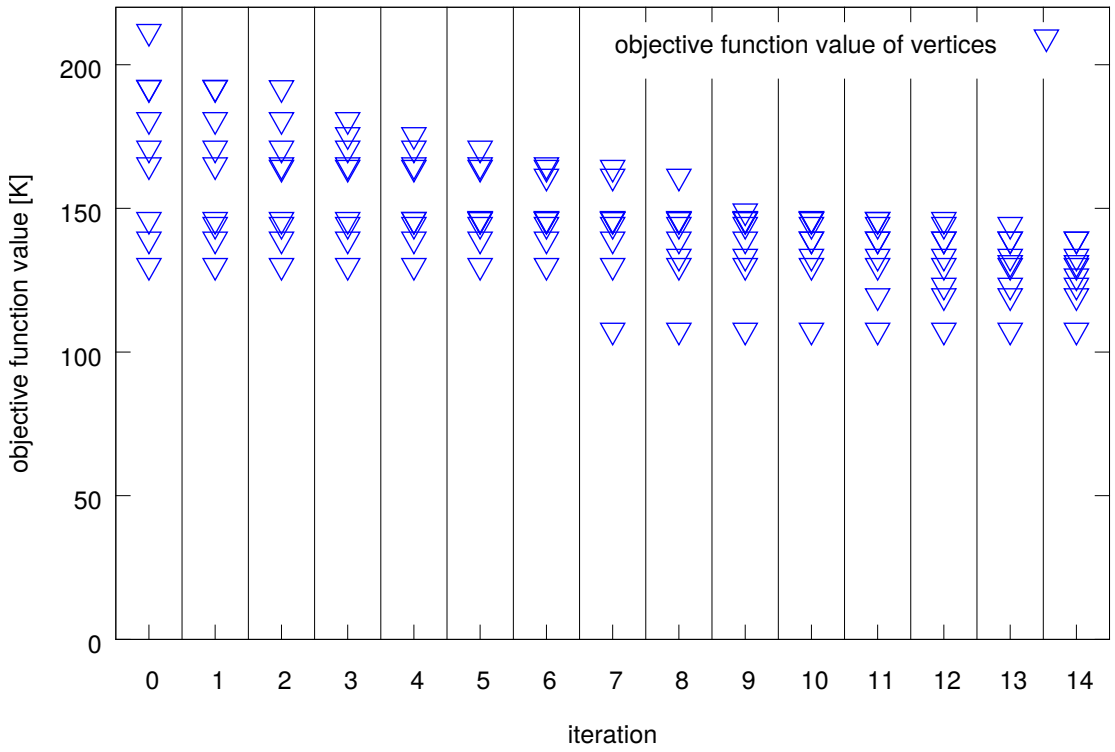


Fig. 3.4 The objective function values for the different vertices versus the iteration number for fitting CuAu 50:50 with 24 atoms. The algorithm is initialised for simplex 0. We only see one successful reflection at iteration 7. The expansion is however not successful. After iteration 7, the algorithm does not find a vertex with a lower objective function value for the time it is run. The spread, the difference between maximum and minimum, of the objective function values begins to shrink after the initialisation. The successful reflection in iteration 7 increases the spread but the shrinking continues afterwards. The spread in the last 6 iterations is below 50 K. The error for the difference between maximum and minimum melting temperature is however ca 29 K for these runs, based on the average relative error of ca 0.020 mentioned in the text. This puts the difference between the maximum and minimum objective function values of these simplices in the 2 SE realm.

The average relative statistical error of the melting temperature is ca 0.0074. The average is again taken over all vertices of this particular fitting run. This leads to a statistical error of the extrapolated melting temperature between ca 7 K and ca 9 K. As the objective function is simply the absolute difference between extrapolated melting temperature and experimental value, one can conclude that the difference of the accepted vertex is clearly within the 2 SE realm. The vertex 28 is created during a reflection. The expansion triggered by this increases however the objective function value and is hence ignored. The difference between minimum and maximum transition temperature in each simplex is greater than twice the error. The minimum difference is 36 ± 11 K at iteration 1 and the maximum difference is 114 ± 12 K at iteration 12. I take this as an indication that, as opposed to the 24 atoms calculations in Section 3.13.1, meaningful distinctions between the melting temperatures of the different vertices are much likelier to be drawn.

The shift of the melting temperature is demonstrated in Figure 3.6. My predictions using the original potential [39–41] are depicted with grey upwards pointing triangles and the predictions of the new potential are denoted by orange + symbols. Blue "×" symbols denote the ITS-90 definitions of Au and Cu [88]. The blue downwards pointing triangle is the approximation calculated in Section 3.12 based on the experimental data of Bennett [87] quoted in Okamoto *et al.* [42]. The error due to the way I estimate it (1.5 K, see Section 3.12) is much smaller than the symbol size and hence not shown. The melting temperature predictions for the different potentials were calculated by fitting the function $k/N + h$ as explained in Section 3.6. The CuAu 50:50 melting transition is fitted very well as could already be assumed from the 1 SE accordance of the melting temperature predicted during the fitting. It was increased by ca 206 K from 995 ± 39 K to 1201 ± 60 K. I approximate the value the expected experimental as 1186 K as explained in Section 3.12 based on the values of Bennett [87] quoted in Okamoto *et al.* [42]. This is the value to which the potential is fitted. It is ca 15 K below our $1/N$ extrapolation prediction. The pure gold value was increased by ca 280 K from 1044 ± 36 K to 1324 ± 35 K just ca 14 K below the value defined in the ITS-90 of 1337.33 K [88]. However, my new potential is not able to reproduce the characteristic eutectic melting curve of CuAu. The pure Cu melting temperature has been increased, but only by ca 20 K from 1095 ± 43 K to 1114 ± 73 K. This means that the change is clearly not statistically significant. This is ca 243 K below the experimental melting point of 1357.77 K which was defined at this value by ITS-90 [88].

We now consider how a different extrapolation scheme affects the results. To fit the melting temperature versus number of atoms curve of the nested sampling calculations, $T_{m,comp}(N) = k/N^2 + h$ may be more suitable as explained in Section 3.6. This leads to a new, in general greater value for $k_{extrapol,original}(N)$. In this specific case the value is 0.795 ± 0.013 (see

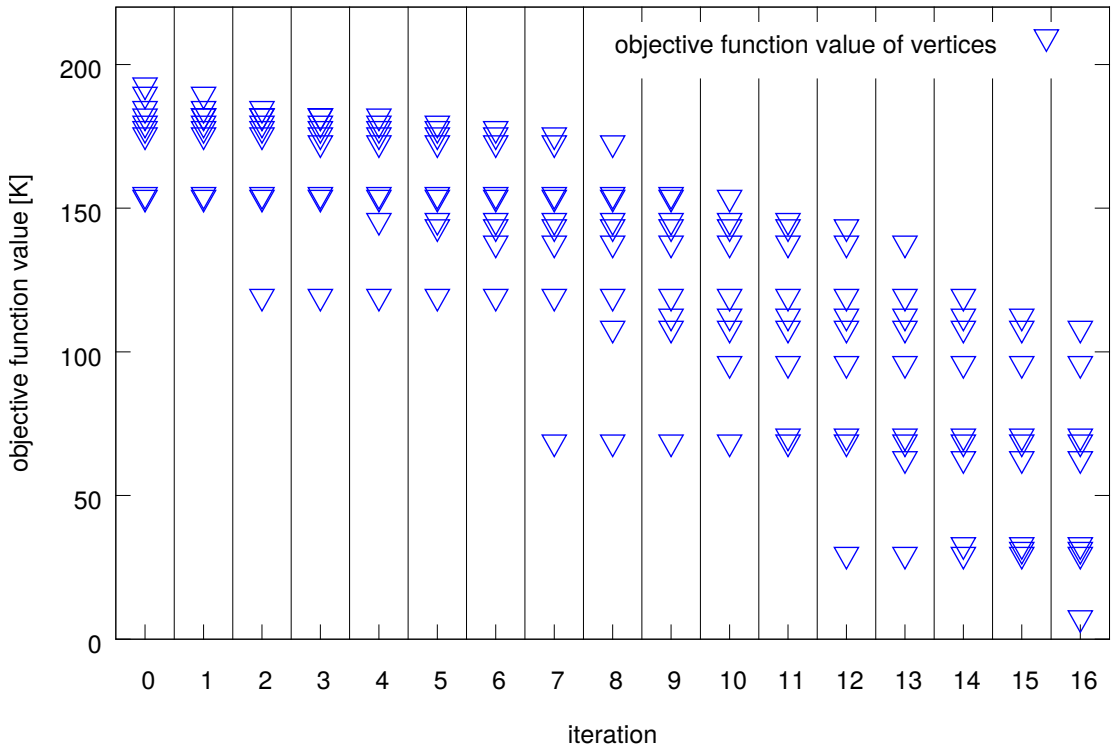


Fig. 3.5 The objective function values for the different vertices of each simplex versus the iteration number in which they are accepted for fitting CuAu 50:50 with 36 atoms using the $1/N$ extrapolation constant. Again, the algorithm starts with the initialisation simplex at iteration 0. We see expansions leading to considerably lower values of the objective function at iterations 2, 7 and 12. During iteration 16 an expansion is triggered as well, but the objective function increases due to it. Hence, just the reflection is used. The average relative error of the the melting temperature is ca 0.0074 leading to error expectations of ca 7 K to ca 9 K.

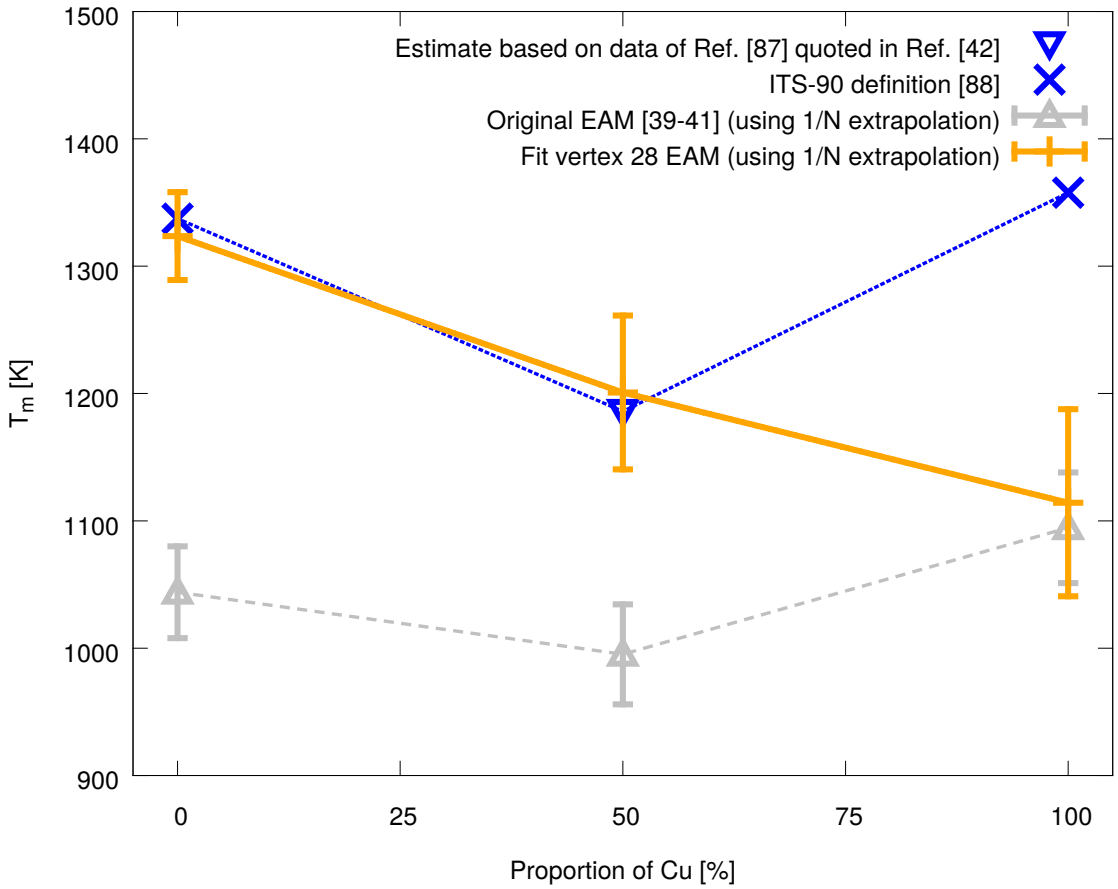


Fig. 3.6 The melting curve for the new CuAu potential (assuming a $1/N$ fitting) after fitting to the CuAu 50:50 melting transition (vertex number 28) in comparison to experimental and my results using the original EAM potential [39–41]. We see the melting temperature against the percentage of Copper. The blue "$\times$" symbols are the values for pure Cu and pure Au as defined by ITS-90 [88]. The blue downwards pointing triangle depicts my approximation for the value at 50:50 CuAu as explained in Section 3.12 based on the values of Bennett [87] quoted in Okamoto *et al.* [42]. The grey upwards pointing triangle depict the values which I calculated with the original potential [39–41]. Finally, the orange "+" symbols stand for results from the new modified potential. The melting temperatures of the potentials are extrapolated from results from different number of atoms with the $1/N$ scheme as described in Section 3.6. The lines are only meant to be guides and not to be seen as precise interpolation. The CuAu 50:50 value is in accordance with the estimate based on experiments. The pure Au value has also improved considerably. However, the Cu value has increased only very little so that it is now the minimum of the three presented values.

Table 3.1b). As I extrapolate the melting temperature during the Nelder-Mead algorithm just by multiplying with $k_{extrapol,original}(N)$ one can use the same data as before to get the objective function values as if $T_{m,comp}(N) = k/N^2 + h$ were used all along. This works without a problem as long as the new extrapolated melting temperature is smaller or equal the experimental melting temperature. With the $1/N^2$ extrapolation scheme, the Nelder-Mead algorithm would have stopped after iteration 7 with vertex 17 as final vertex. The parameters are given by the following vector:

$$\vec{x}_{\text{converged, 17}} = \begin{pmatrix} 1.0106726074218755 \\ 0.950775146484375 \\ 1.033172607421875 \\ 0.9892272949218752 \\ 1.0050476074218757 \\ 1.033172607421875 \\ 0.9831726074218752 \\ 1.033172607421875 \end{pmatrix} \tag{3.15}$$

For this vertex, the difference between the melting temperature I aim for and the one predicted with the extrapolation constant of 0.795 is 11 ± 8 K. Here, the error is calculated by the average relative statistical error of ca 0.0074 which was estimated for the whole potential fitting procedure which relies on the $1/N$ extrapolation constant mentioned at the start of this subsection. In Figure 3.7, we see the melting curve predictions for the new potential parameters of vertex 17 in comparison to experimental predictions and the original parameters [39–41]. The new potential results are depicted with orange "+" symbols. The results for the original potential [39–41] are given by the grey upwards pointing triangles. The ITS-90 definition of the Au and Cu melting temperature [88] are shown with blue "×" symbols. The melting temperature for CuAu 50:50 is estimated from results of Bennett [87] quoted in Okamoto *et al.* [42] as shown in Section 3.12 and represented by the blue downwards pointing triangle. Lines are only guidelines. The melting temperatures using the original parameters [39–41] were predicted via fitting $T_m(N) = k/N^2 + h$ linearly as presented in Section 3.6. The melting temperature predictions using the modified potential parameters of vertex 17 for pure Cu and Au prediction were calculated via the $1/N^2$ scheme for nested sampling runs for different number of atoms N. The CuAu 50:50 melting predictions were calculated based on the extrapolation constant for 36 atoms gained from the original parameters $1/N^2$ extrapolation for CuAu 50:50 (see Table 3.1b). This extrapolation constant is then applied on the melting temperature calculated for 36 atoms for vertex 17 during the EAM fitting. Its error is calculated by using the SEM of the melting temperature for vertex 17 and the error of

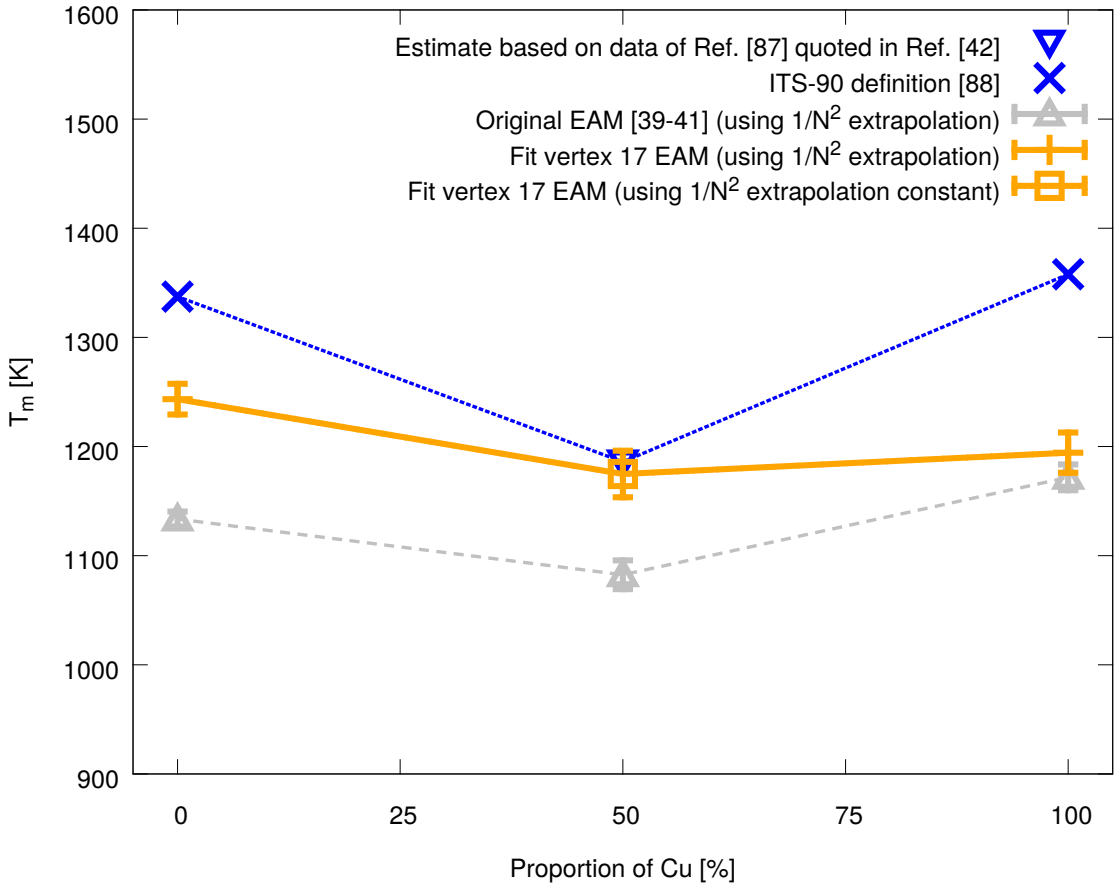


Fig. 3.7 The melting curve for the new CuAu potential (assuming a $1/N^2$ extrapolation) after fitting to the CuAu 50:50 melting transition in comparison to experimental values and my results from the original potential[39–41]. Again, the experimental values for pure Au and Cu [88] are denoted with blue "$\times$" symbols while the estimate for the CuAu 50:50 expected melting temperature is signified by a downwards pointing blue triangle estimated from the results of Bennett [87] quoted in Okamoto *et al.* [42]. Grey upwards pointing triangles stand for my predictions using the original potential[39–41]. The modified potential results (vertex 17) are represented by orange symbols. Orange "+" symbols represent values for which an $1/N^2$ extrapolation has been conducted. For the empty orange square, I use an extrapolation constant of ca 0.795 (see Table 3.1b, for the original potential). Note that the fitted potential is different from the one presented in Figure 3.6 as the $1/N^2$ extrapolation would have let the algorithm stop earlier. The CuAu 50:50 predictions of the new potential are in accordance with the experimental result as it was fitted to this value. The Au and Cu behaviour has also been improved and the eutectic behaviour has been preserved. However, the pure Au and Cu prediction is still significantly lower than the experimental values.

the extrapolation constant. The extrapolation constant is 0.795 ± 0.013 K as mentioned above. The melting temperature predicted by the EAM of vertex 17 is 1478 ± 12 K. This yields a predicted temperature of 1175 ± 21 K. The CuAu 50:50 melting prediction is in accordance with experiments, as it was fitted to the experimental value of 1186 K as estimated in Section 3.12 based on the results of Bennett [87] quoted in Okamoto *et al.* [42]. It was increased by ca 92 K. For pure Au, the vertex 17 potential predicts a ca 110 K higher value than the original increasing the melting temperature from 1134 ± 7 K to 1243 ± 14 K. This is still ca 94 K below the value defined in ITS-90 of 1337.33 K [88]. For pure Cu, we can see an increase of only ca 23 K from 1172 ± 12 K to 1194 ± 18 K. This is ca 163 K lower than the ITS-90 defined value for pure Au of 1357.77 K [88]. As for the potential of vertex 28 of the 36 atoms CuAu 50:50 fitting, the improvement of the Au temperature is not significant. Opposed to the earlier results presented in Figure 3.4, the eutectic curve was preserved. However, the pure metal melting temperature predictions are both still significantly below the experimental values. The preserving of the eutectic curve may have just been the fact that only 17 vertices were calculated in this case as opposed to the 28 vertices in the $1/N$ extrapolation based Nelder-Mead fitting. This means that the chance of changing the qualitative behaviour of the melting curve is just lower for vertex 17. To conclude, even though the fitting has led to a certain improvement of the prediction of the eutectic curve, the predicted melting curve is not in accordance with the experimental melting curve shown in Figure 3.7. Hence, I turn to fitting several transitions simultaneously in Section 3.14.

3.14 Fitting pure Cu, pure Au and CuAu 50:50 simultaneously

After having successfully fitted the CuAu melting transition for a 50:50 ratio alone, I now turn to fitting this melting transition as well as the pure copper and pure gold melting transitions at the same time. The hope is that fitting anew to more melting transitions yields an interatomic potential that reproduces the eutectic curve better than the one fitted for just the CuAu 50:50 transition. For the simultaneous fitting, the objective function f, is not as simple as in the previous sections. It is the general form defined in Equation 3.4, the square root of the sum over the quadratic differences:

$$f(\vec{x}) = \sqrt{\sum_{i}^{N} \left(T_{i,exp} - T_{i,comp}\right)^2} \tag{3.16}$$

I have updated my own Nelder-Mead implementation to deal with more than one transition and to calculate the above given objective function. To summarise, one just needs to submit additional nested sampling and analysis runs for each additional transition and calculate the resulting objective function. The Nelder-Mead algorithm itself does not need a modification. I use 36 atoms, 1152 walkers and $n_{cull} = 8$ to investigate all transitions. For the pure elements, 768 calls per iteration are used. For the CuAu 50:50 nested sampling calculations, I use 1536 calls. The extrapolation of the melting temperature is conducted as explained before in Section 3.6. This time, I use fitting of $k/N^2 + h$ instead of $k/N + h$ from the start to get the extrapolation constant. I chose the $1/N^2$ fit, as it yields a better fit as explained in Section 3.6. The resulting extrapolation constants are 0.822 ± 0.006 for Au, 0.795 ± 0.013 for CuAu 50:50 and 0.806 ± 0.009 for Cu (see Table 3.1b). I use these constants rounded to the third decimal place in the Nelder-Mead algorithm. I start with nearly the same vertices as in Sections 3.13.1 and 3.13.2. However, my first vertex represents not anymore the original potential [39–41]. Instead it represents a modification to it where $\rho_{a,CuAu}(r)$ is already replaced with $\rho_{a,AuCu}(r)$.

In Figure 3.8, we see the objective function values of the vertices of each simplex for each iteration. Iteration 0 shows our initialisation simplex values. There are successful expansion steps at iteration 1, 6 and 22. These expansions lead to considerably decreases of the objective function. During iteration 11, 12 and 14 expansions are tried by the algorithm, but are unsuccessful. Iteration 14, the decrease of the objective function is quite large, even though it is only due to a reflection step. Even though the algorithm fits to thrice as many values than it does to for the CuAu 50:50 ratio fitting, it takes only 22 iterations (not counting the initialisation) until the algorithm can be stopped. For the CuAu 50:50 ratio fitting, 16 iterations are needed.

The final accepted vertex is vertex number 38 after iteration 22 with a value of ca 11 K. It is given by:

$$\vec{x}_{38} = \begin{pmatrix} 0.936319511505266 \\ 0.879322040542473 \\ 1.0649616570066303 \\ 1.107900241381125 \\ 0.9674411022153587 \\ 1.114130915849804 \\ 1.0280473746351468 \\ 0.9825709539503822 \end{pmatrix} \tag{3.17}$$

This is the first vertex for which the 2 SE ellipsoid as defined in Section 3.8 contains the position vector of the extrapolated melting temperatures $\vec{T}_{m,extrapol}$. This is demonstrated

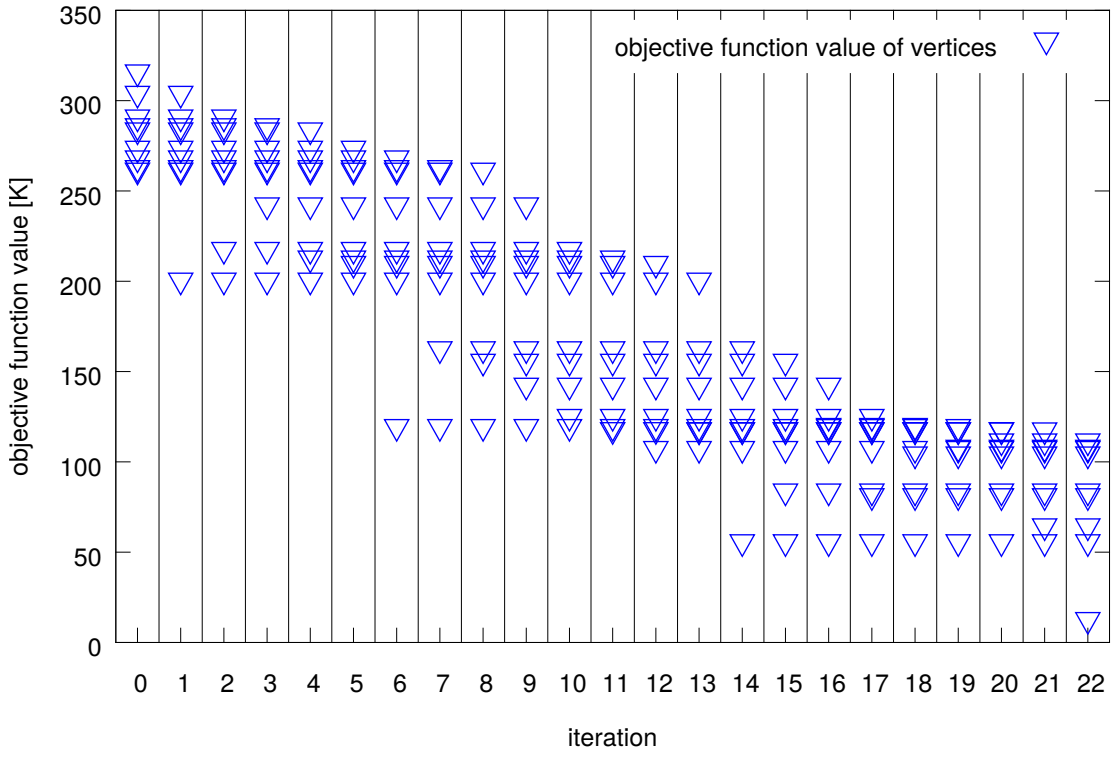


Fig. 3.8 The objective function values for the different vertices of each simplex versus the iteration number. The algorithm is initialised at iteration 0. In iteration number 1, the algorithm already conducts an expansion leading to a new vertex for which the objective function is considerably closer to 0 K. Further successful expansion steps during iteration 6 and 22 drive the value down further. The leap at iteration 14 is due to a reflection for which the subsequent expansion fails. There are also failed expansions st iteration 11 and 12. The algorithm arrives at a final value of ca 11 K after iteration 22.

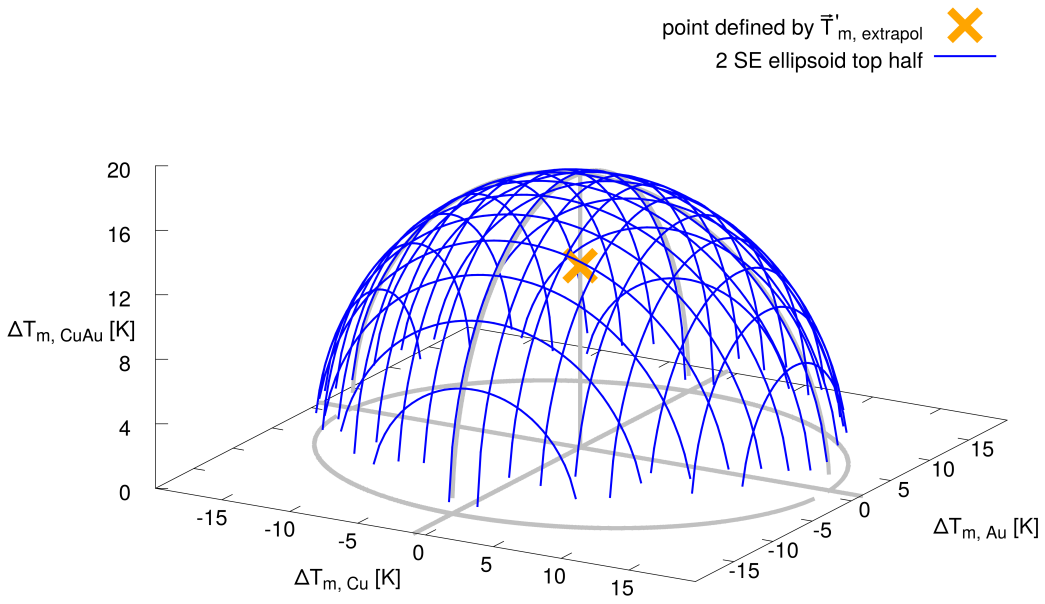


Fig. 3.9 Demonstration that vertex 38 of simultaneous fitting the Au, CuAu 50:50 and Cu melting temperatures is in the 2 SE ellipsoid based on the averaged statistical relative error. For better visibility, the coordinate system is shifted by $-\vec{T}_{m,exp}$ so that $\vec{T}'_{m,exp}$ is at the source of the coordinate frame. A blue mesh indicates the 2 SE ellipsoid. Using the grey auxiliary lines, one can see that the point defined by $\vec{T}'_{m,extrapol}$ (orange "$\times$" symbol) is within this ellipsoid.

in Figure 3.9. Here, I shifted the coordinate frame of melting temperatures by $-\vec{T}_{m,exp}$, the position vector of the melting temperatures to which I fit, and show the difference in melting temperatures defined as follows:

$$\Delta\vec{T}_m = \begin{pmatrix} \Delta T_{m,Cu} \\ \Delta T_{m,Au} \\ \Delta T_{m,CuAu} \end{pmatrix} = \begin{pmatrix} T_{m,Cu} - T_{m,exp,Cu} \\ T_{m,Au} - T_{m,exp,Au} \\ T_{m,CuAu} - T_{m,exp,CuAu} \end{pmatrix} \tag{3.18}$$

Here, $T_{m,exp,Cu/Au/CuAu}$ is the experimental melting temperature used in the fitting for the different compositions of the system. They are based on the values of Bennett [87] quoted in Okamoto *et al.* [42] for CuAu 50:50, on Ref. [97] for Cu and on Ref. [88] for Au as explained in Section 3.12. $T_{m,Cu/Au/CuAu}$ is an arbitrary melting temperature for Cu, Au or CuAu 50:50. In the new coordinate frame, the position vector $\vec{T}_{m,extrapol}$ becomes:

$$\vec{T}'_{m,extrapol} = \begin{pmatrix} T_{m,extrapol,Cu} - T_{m,exp,Cu} \\ T_{m,extrapol,Au} - T_{m,exp,Au} \\ T_{m,extrapol,CuAu} - T_{m,exp,CuAu} \end{pmatrix} \approx \begin{pmatrix} -0.8\,K \\ 1.2\,K \\ 11.3\,K \end{pmatrix} \tag{3.19}$$

$T_{m,extrapol,Cu/Au/CuAu}$ is the melting temperature for the different compositions of the system extrapolated from the 36 atoms nested sampling calculation via the extrapolation constant used during the fitting. The top of the ellipsoid is indicated by a blue mesh. Its principle semi-axes are equal to twice the statistical errors. To determine these statistical errors for the specific transitions, I use the average relative SEM of the heat capacity peak position of each vertex calculated during the fitting procedure. Via the grey helplines, we can see that the point defined by the position vector $\vec{T}'_{m,extrapol}$, denoted with the orange "$\times$" symbol, is in the ellipsoid. For a more analytically measure, we can use the Mahalanobis distance [99] as mentioned in Section 3.8. It is ca 1.3 showing clearly that $\vec{T}_{m,exp}$ and $\vec{T}_{m,extrapol}$ are not significantly different.

In Figure 3.10 we see my melting curve predictions for the modified potential (vertex 38), for the original potential [39–41], the melting temperature estimated for CuAu 50:50 based on experimental values of Bennett [87] quoted in Okamoto *et al.* [42] calculated as explained in Section 3.12 as well as the ITS-90 definitions for pure Cu and Au [88]. The $1/N^2$ extrapolation is used for the original potential predictions which are given by grey upwards pointing triangles. This allows for the calculation of the extrapolation constants shown in Table 3.1b. For the predictions for the vertex 38 potential, I use these extrapolation constants whose rounded values were used originally for the EAM fitting. The reference values are the melting temperatures calculated for vertex 38 during the EAM fitting procedure. The

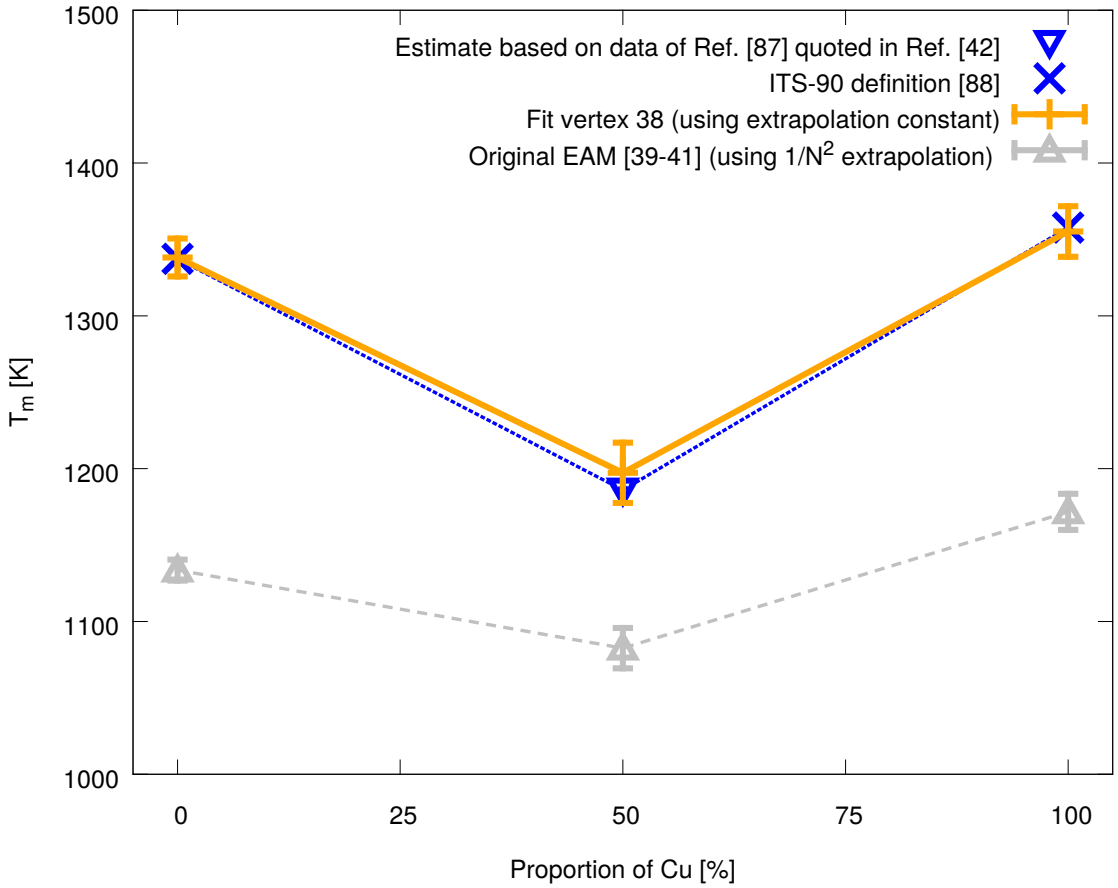


Fig. 3.10 The CuAu melting curve using the original potential [39–41], the modified potential (vertex 38) and experimental data in dependence of the atomic percentage of Cu. The results using the original potential are depicted through upwards pointing grey triangles. The blue "×" symbols signify the melting temperature for pure Au and Cu [88]. The downwards pointing blue triangle is my estimation based on the results of Bennett [87] quoted in Okamoto *et al.* [42]. The eutectic behaviour is preserved after the optimisation. We achieved very good accordance with the experiments increasing the melting temperature by ca 204 K for pure Au, by ca 115 K for CuAu 50:50 and by ca 183 K for pure Cu. The original potential [39–41] predictions and their error are calculated using the $1/N^2$ extrapolation. The results for the final accepted potential are calculated using the extrapolation constants gained in the aforementioned $1/N^2$ extrapolation for the original potential [39–41] (see Table 3.1b). The errors are based on the error of the extrapolation constant and the statistical SEM of the 5 nested sampling runs used as reference for each alloy composition. For each composition, the experimental melting temperature is within the standard error of the melting temperature predicted by the new potential.

resulting extrapolated values are signified by orange "+" symbols. Blue "×" symbols denote the experimental Au and Cu values [88]. The downwards pointing triangle signifies the melting temperature estimate based on the experimental data of Bennett [87] quoted in Okamoto *et al.* [42]. The errors of the original potential melting temperatures are based on the extrapolation as described in Section 3.6. Note that the error used to construct the ellipsoid in Figure 3.9 is the statistical error gained from averaging the relative errors. For the error calculation of the melting temperatures of the final vertex, I use the error originally estimated from the 5 different nested sampling runs per transition of vertex 38, which is different from the averaged relative error. I also included the error of the extrapolation constants. We see that the eutectic melting curve of the CuAu alloy can be reproduced very well for the investigated points. The pure Au melting temperature of the new potential was increased by ca 204 K from 1134 ± 7 K to 1338 ± 12 K. The melting temperature of the 50:50 CuAu melting transition is much higher as well for the new potential at 1197 ± 20 K. This is an increase of ca 115 K from the original prediction of 1083 ± 13 K. For the Cu melting temperature, the new potential predicts a melting temperature increased by ca 183 K from 1172 ± 12 K to 1355 ± 16 K. For the pure Au and Cu systems, the experimental values are the values defined in the ITS-90 for Au, 1337.33 K, and Cu, 1357.77 K [88]. The CuAu melting temperature was estimated as explained in Section 3.12 based on the experimental data of Bennett [87] quoted in Okamoto *et al.* [42] to be 1186 K. All experimental points are within one standard error of the new melting temperature estimates. Note that this is not very surprising as these values were fitted to values close to the experimental values presented in Figure 3.10 in the first place. As mentioned in Section 3.12 the values fitted to are based on values of Bennett [87] quoted in Okamoto *et al.* [42] and Ref. [97]. Additionally, the total error is greater than just the statistical error which was used to terminate the Nelder-Mead algorithm after reaching the 2 SE ellipsoid as define in Section 3.8. This is due to the fact that we need to include the error of the extrapolation constant.

3.14.1 The new functions, elastic constants, lattice parameters and volumes per atom

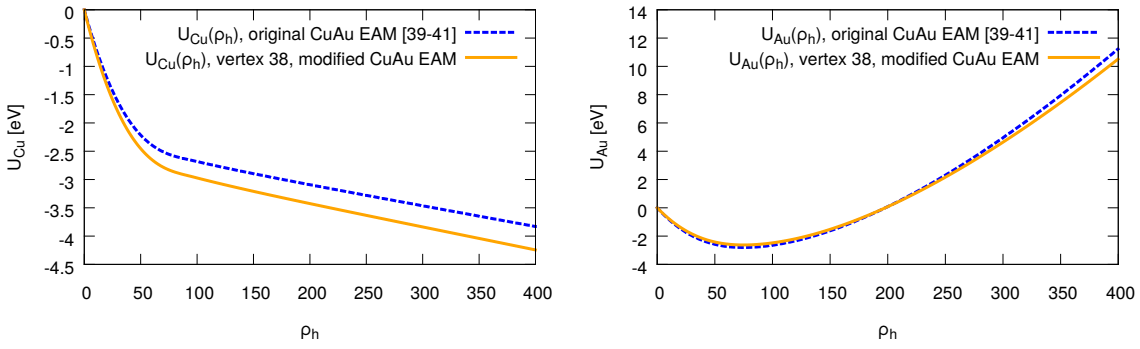
As the most recently created interatomic potential reproduces the eutectic curve very well, I turn to the last step of my approach of Section 1.2: comparing basic features of the new and original EAM [39–41]. This way, we can see how much the new potential deviates from the original potential and see how it performs now for basics properties which are not included directly in my fitting.

In the following, I turn my attention to the different functions which make up the original

CuAu EAM [39–41] and the EAM of vertex 38 from the simultaneous fitting of the melting temperature of Au, CuAu 50:50 and Cu. I compare the embedding functions in Figure 3.11 and the functions the different $\phi(r)$ and $\rho_a(r)$ functions which depend only on the distance r between a pair of atoms in Figures 3.12 and 3.13. In all Subfigures, I have used blue dashed lines to signify the functions for the original CuAu EAM [39–41] and orange solid lines to signify the functions of the modified potential of vertex 38. Interestingly, the embedding functions $U_{Cu}(\rho_h)$ shown in Figure 3.11a are very different from the one shown for Au in Figure 3.11b. The Au embedding functions show a minimum, while the Cu functions do not show a minimum. At some point between a host density between 50 and 100, they start to behave linearly with a negative slope. The effect of the modification on the Cu embedding function is stronger than for the Au embedding function, though both are clearly visible. The pair potential part of the EAM is shown in Figures 3.12a, 3.12b and 3.12c. In all cases, one can see a minimum and a repulsive behaviour for lower radii. For the Cu-Cu interaction (Figure 3.12a) the effect of the fit is strongest decreasing the minimum energy. The same happens in a lesser manner for the Au-Au interaction (Figure 3.12c). The change of the Au-Cu interaction modification (Figure 3.12a) is barely visible showing a very small increase in the energy value of the minimum. The pair functions $\rho_a(r)$ over which one sums to create the host density are portrayed in Figure 3.13a, 3.13b, 3.13c and 3.13d. The distance between the atoms of a pair is given again by r. Apart from a short period which is probably due to small numerical errors, the density pair functions do not have values below 0. The functions show a strong increase with decreasing number of atoms below ca 4 Å. The changes to the Cu-Cu and Au-Cu pair density functions shown in Figure 3.13a and 3.13c respectively, are barely visible. The Au-Au pair density function presented in Figure 3.13b changed visibly. The function value is lowered which means that high function values reached with the original potential [39–41] are now reached at a slightly lower atomic distance than before. The behaviour of $\rho_{CuAu}(r)$ is changed in a different manner as it is replaced by $\rho_{AuCu}(r)$. The overall change however is relatively small. For distances below ca 4.32 Å, the new potential pair density function is smaller than the original one. For distances above ca 4.32 Å the original one is smaller and decreases much more quickly to a value close to 0 for increasing r.

In Table 3.2, I show my results for the different elastic constants, the lattice constant and the volume per atom for the fcc lattice for Cu (Subtable 3.2a) and Au (Subtable 3.2b) using the original potential EAM [39–41] and the potential of vertex 38 of the simultaneous fitting of Cu, CuAu 50:50 and Au. Additionally, I include experimental values for comparison. The experimental values for the elastic constants of Cu are extrapolations to 0 K conducted in Ref. [106]. Similarly, for Au, the experimental elastic constant values were also extrapolated

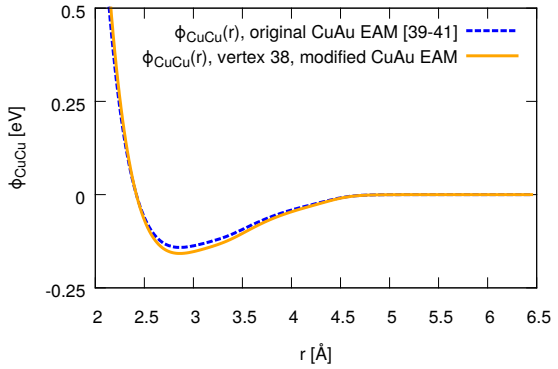
from experimental data to 0 K in Ref. [107]. I use finite difference to calculate the elastic constants at 0 K. For this, I use scripts by Aidan Thompson [108] which are part of *LAMMPS* [74, 109]. I adjusted them very slightly for my purposes. The starting structures are first relaxed. After the application of finite differences, the atomic positions are relaxed again. For these relaxations, a conjugate gradient scheme [110] is used. The lattice constants for the starting structures of Cu and Au (3.6 Å and 4.1 Å respectively) are rounded values of the results of Ref. [111]. For Cu, the values predicted by the EAM potentials are in agreement, except for the c_{11} value and to a lesser extent the c_{12} value. The c_{11} value of the newly fitted potential is ca 28 % lower than the one for the original potential. The c_{12} value of the newly fitted potential is ca 10 % higher than the one for the original potential. For Au, the changes are relatively strong. The modification leads to an increase of ca 30 % for c_{11}, ca 9 % for c_{12} and ca 30 % for c_{44}. The decrease for the lattice constant is only ca 1.5 %, leading to a volume per atom decrease of ca 4.6 %. This is only a slight change, but clearly more than for the Cu fcc phase, which shows decreased of ca 0.1 % for the lattice constant and and ca 0.3 % for the volume per atom. Overall, my new potential predicts elastic constants in good agreement with the experimental results [106, 107]: On average, the elastic constants are improved from the predictions of the original potential [39–41], especially for the c_{11} constant of Cu. The experimental lattice constant values for Cu [112] and Au [50] are represent both values for ca 298 K. In the case of Au, the deviation due to modification is much stronger than for Cu for which the deviations are very small. The modified potential value for the Au volume per atom is ca 5 % smaller than the volume per atom value based on the lattice constants value of Ref. [50].



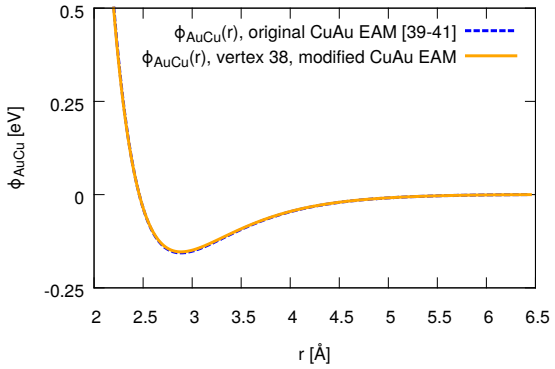
(a) Comparison of $U_{Cu}(\rho_h)$ plots

(b) Comparison of $U_{Au}(\rho_h)$ plots

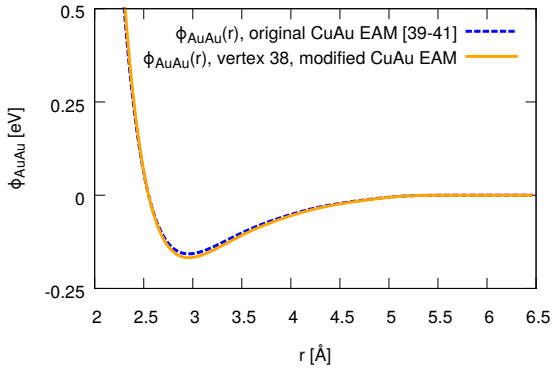
Fig. 3.11 Comparison of the embedding functions of the original CuAu [39–41] and the potential of vertex 38 from the simultaneous Au, CuAu 50:50 and Cu melting temperature fitting. Blue dashed lines represent the original functions while solid orange lines show the functions of my new potential



(a) Comparison of $\phi_{CuCu}(r)$ plots

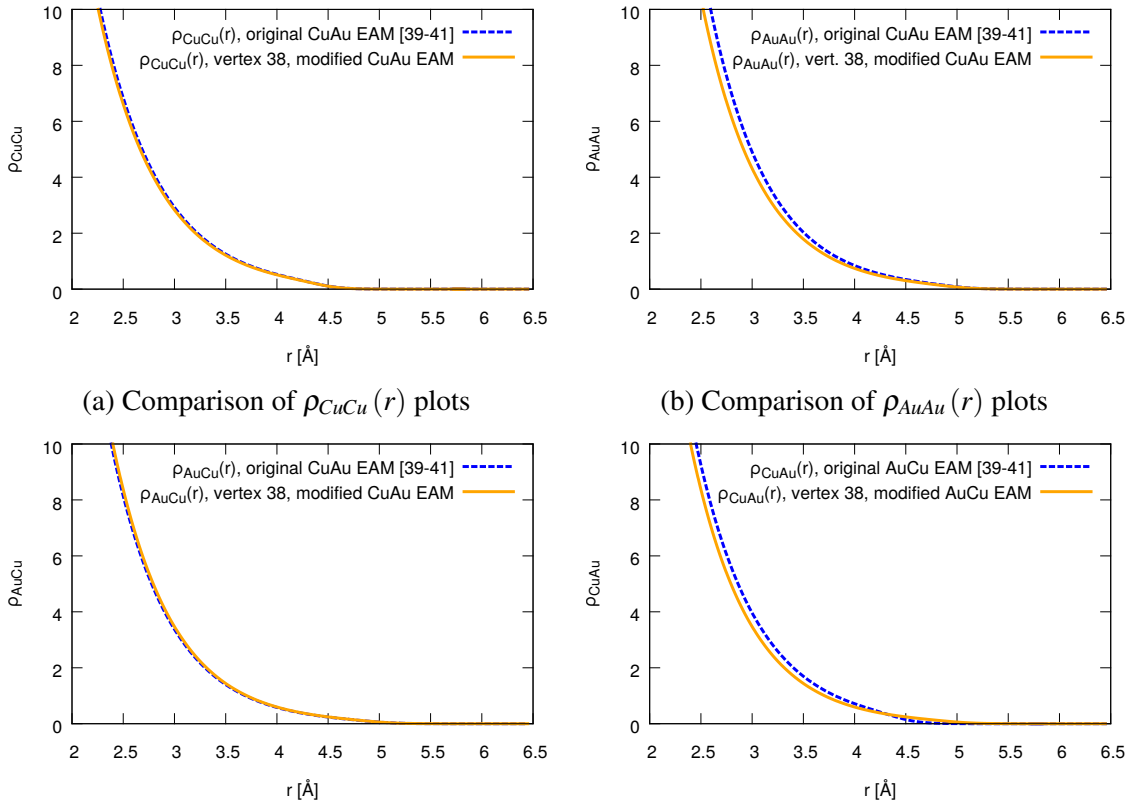


(b) Comparison of $\phi_{AuCu}(r)$ plots



(c) Comparison of $\phi_{AuAu}(r)$ plots

Fig. 3.12 The pair interaction part ($\phi(r)$) of the potential before and after the fitting. The distance between a pair of two atoms is given with r. Blue dashed lines denote the functions for the original potential [39–41]. The orange solid lines are the functions of the new potential of vertex 38. This vertex is gotten by the potential of vertex 38 from the simultaneous fitting of the Au, CuAu 50:50 and Cu melting temperatures.



(a) Comparison of $\rho_{CuCu}(r)$ plots

(b) Comparison of $\rho_{AuAu}(r)$ plots

(c) Comparison of $\rho_{AuCu}(r)$ plots

(d) Comparison of $\rho_{CuAu}(r)$ plots

Fig. 3.13 Comparison of the pair interaction parts of the host density($\rho_a(r)$ for the original CuAu potential [39–41] denoted by blue dashed lines and the potential of vertex 38 from the simultaneous Au, CuAu 50:50 and Cu melting temperature fitting denoted by solid orange lines. Here r is the distance between an atomic pair.

quantity	original CuAu potential [39–41]	vertex 38	experiments
c_{11} [GPa]	280	202	176
c_{12} [GPa]	137	150	125
c_{44} [GPa]	89	88	82
a [Å]	3.614	3.610	3.615
V_{at} [Å^3]	11.80	11.76	11.81

(a) Cu elastic constants (c_{11}, c_{12} and c_{44}), lattice constant (a) and volume per atom (V_{at}) for an fcc lattice which I calculated for the original CuAu potential [39–41] and the potential for vertex 38 of the three transitions fitting. The experimental values reported for the elastic constants are are extrapolations to 0 K from Ref. [106]. The experimental lattice constant value is from Ref. [112] for ca 298 K and used to calculate the volume per atom.

quantity	original CuAu potential [39–41]	vertex 38	experiments
c_{11} [GPa]	165	215	202
c_{12} [GPa]	156	170	170
c_{44} [GPa]	40	60	45
a [Å]	4.078	4.015	4.079
V_{at} [Å^3]	16.95	16.18	16.96

(b) Au elastic constants (c_{11}, c_{12} and c_{44}), lattice constant (a) and volume per atom (V_{at}) for an fcc lattice which I calculated for the original CuAu potential [39–41] and the potential for vertex 38 of three transitions fitting. The experimental values for the elastic constants are are extrapolations to 0 K from Ref. [107] based on their experiments. The experimental lattice constant value (at ca 298 K) is taken from Ref. [50]. I use the precise value of 4.0785 Å given in Ref. [50] to calculate the volume per atom.

Table 3.2 Comparison of several basic quantities of the fcc phase. I present values which I have calculated using the original CuAu potential [39–41] and the fitted potential based on it for vertex 38 gained in the simultaneous fitting of the Cu, CuAu 50:50 and Au melting temperatures. Experimental values are also included.

Chapter 4

Titanium results

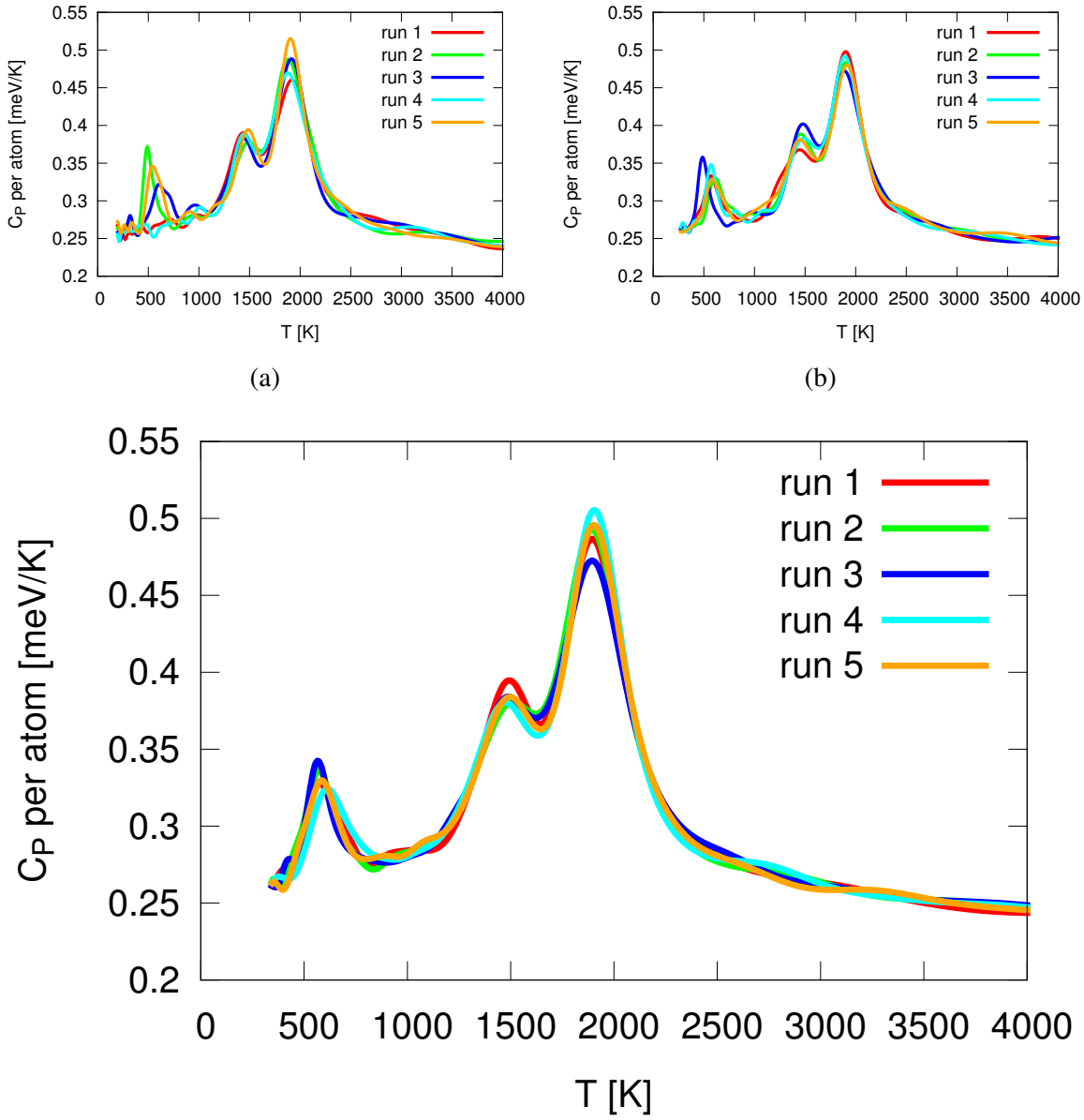
4.1 Titanium results introduction

In this chapter, I present my work on the titanium MEAM potential of Hennig *et al.* [15] as well as later my work relating to my modifications to it. The aim is to created a modification of the MEAM of Hennig *et al.* [15] which predicts the experimental phase diagram sufficiently well. I have listed the salient parameters for the nested sampling calculations in Section A.3. The most important technical details relating to the methodology of my nested sampling calculations with *pymatnest* [45] are already covered in Section 3.2 to Section 3.4. Hence, these are not repeated here. First, I determine which are the important factors for convergence of *pymatnest* [45] in this specific case. Afterwards, I start with the approach as given in Section 1.2. First, the original potential [15] predictions for the phase diagram need to be investigated. Hence, I turn my attention to the effects of different numbers of atoms on the phase diagram predicted with nested sampling. To get a better understanding of the phase diagram for a number of atoms per simulation box outside of the capabilities of *pymatnest*, I consider thermodynamic integration predictions in Section 4.6. In this section, I also elaborate on the technical details of the methodology of the nonequilibirum approach [48, 49] which I use. I also compare the energies predicted by the original MEAM [15] to DFT energies for a number of sampled structures in Section 4.7. For all these calculations using an interatomic potential which are presented in Sections 4.2 to 4.7, I only use the original potential of Hennig *et al.* [15]. Once an understanding of the phase diagram predicted by the original MEAM [15] for large numbers of atoms is reached, the actual modification of the original MEAM can be started in Section 4.8. Here, I cover different modifications to the potential which I test with nested sampling. As before, the finite size effects for large numbers of atoms need to be considered. Similar to my original potential [15] calculations, I consider the phase diagram of the final accepted version of the modified MEAM with thermodynamic integration and

MD simulations, as the cost of the nested sampling calculations would be too high for this. This is followed by testing how the new potential performs for the aforementioned sampled structures (see Section 4.9). In Section 4.10, I cover how much the potential function has changed and how the new interatomic potential performs for predicting basic quantities of the α, β and ω phase.
Noam Bernstein and Livia Bartók-Pártay found that culling more than one walker per iteration can lead to convergence problems [113], hence I use low numbers of culls and tested the influence of the number of culls in Section 4.4. From my results in Section 4.4, I deduce that the n_{cull} values of $n_{cull} = 120$ and $n_{cull} = 480$ used by myself in Ref. [33] are too high. It is therefore not clear in which sense the conclusions of Ref. [33] are still correct. Nevertheless, the input parameter values employed to gain the results of Ref. [33] are used as starting points for the parameters used for the work presented in this report. Additionally, I have to choose a number of atoms per box so that expected phases can be described [27]. This was not the case in Ref. [33]. Due to the problems of too high n_{cull} values and wrong number of atoms per simulation cell, it was necessary to repeat the convergence studies of Ref. [33]. If not denoted differently, the error of the transition temperature is the SEM estimated via the independent nested sampling runs with the same parameters.

4.2 Convergence with walker number

In this Section, I present convergence studies for the walker number: Figure 4.1 shows a batch of five runs for each 3840 (Figure 4.1a), 7680 (Figure 4.1b) and 15360 (Figure 4.1c) walkers. For all runs, I only culled one walker per iteration. The pressure is kept at 10 GPa and the number of calls per iteration is 840 for Figure 4.1a and Figure 4.1b and 672 for Figure 4.1c. The number of atoms is 54. The other parameters stay the same for all the runs. In the majority of curves, we can identify three peaks. Their average positions for each five runs batch are in between 500 and 600 K for the lowest peak (ω to α phase transition), in between 1450 K and 1500 K for the second peak (α to β phase transition) and at ca 1900 K for the highest displayed peak (β to liquid phase transition). The lowest peak does not appear in two out of the five runs for 3840 walkers (Figure 4.1a). However, for the higher number of walkers in Figure 4.1, it appears consistently. Additionally, the peaks align better with increasing number of walkers. I conclude, that as I did in Ref. [33] increasing the number of walkers improves the convergence.



(a) (b)

(c)

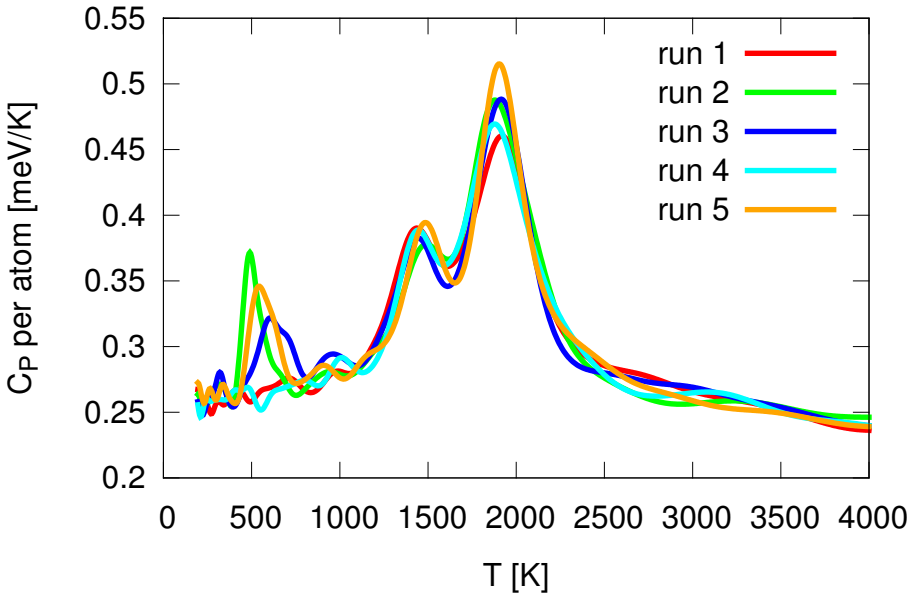
Fig. 4.1 The heat capacity against temperature curve convergence with number of walkers. We can see the heat capacity curves of each five different runs for 54 atoms, at 10 GPa, with 840 (a and b) and 672 (c) calls per iteration as well as 3840 (a), 7680 (b) and 15360 (c) walkers. I also use $n_{cull} = 1$. The increasing convergence with increasing number of walkers is apparent. With increasing numbers of walkers, the peaks become increasingly aligned. Additionally, (a) does not show the first peak in two out five runs.

4.3 The effects of different walk lengths

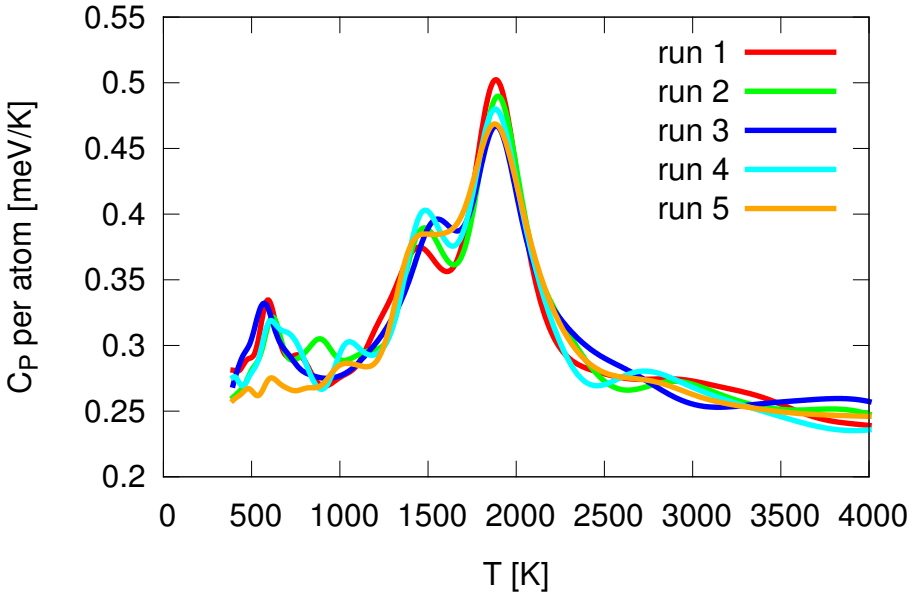
We now turn to the effect of various number of calls have on the appearance of the low temperature solid-solid transition. I was not able to consistently show the appearance of the lowest heat capacity peak by only increasing the number of calls. To demonstrate this, I refer to Figure 4.2 where I show independent runs for different number of calls at 10 GPa where I culled 1 walker per iteration. I have reprinted Figure 4.1a as Figure 4.2a. Instead of using 840 calls as in Figure 4.2a, I used 3840 calls per iteration in for these runs to create 4.2b. We can see that overall convergence did not visibly improve. The lowest peak (ω to α phase transition) is not visible for run 5 in Figure 4.2b. The increase in appearance ratio may just be statistical fluctuation and can not be used to infer that this is due to a better convergence. It is interesting to note, that just doubling the number of walkers leads to a doubled runtime and cost. Ignoring the improvement of parallelization for longer walk length, doubling the number of calls leads to a doubled runtime and cost as well. However a more than quadrupled number of calls per iteration did not lead to convergence while doubling the number of walkers did. It is computationally inefficient to quadruple the number of calls if doubling the number of walkers already suffices. I have conducted studies with low number of calls (120 calls and 240 calls) for $n_{cull} = 8$ and number of walkers $K = 7680$ at 10 GPa. For each 120 calls is conducted 5 runs and for 240 calls 6 runs. These runs have a different relative ratio of the different steps kinds than the runs presented previously in this thesis for titanium. Instead of using 1 MD trajectory with 4 steps, and 1 steps of each kind of cell step, I use 1 MD trajectory of 8 steps, 32 volume steps as well as 8 steps for shear and stretch moves each for these new runs. Nevertheless, they show that one could go as low as 240 calls per iteration and still reliably find the ω phase while 120 calls can already lead to the disappearance of the ω phase. I conclude as already in Ref. [33] that the numbers of walkers is the defining parameter once a minimum number of calls has been chosen for the Ti MEAM potential of Hennig *et al.* [15].

4.4 Culling more than one walker per iteration

As we have seen in Sections 4.2 and 4.3, increasing the number of walkers helps resolve the convergence issues whereas just increasing the number of calls per iteration does not lead to a consistent appearance of the ω phase and low temperatures. However, parallelization via distribution of decorrelation steps on other walkers is limited by low numbers of calls per iteration for the decorrelation walks, as explained in Section 3.3. Hence, evaluation times of the phase diagrams become very long. Therefore, I turn to the second way of parallelizing



(a)



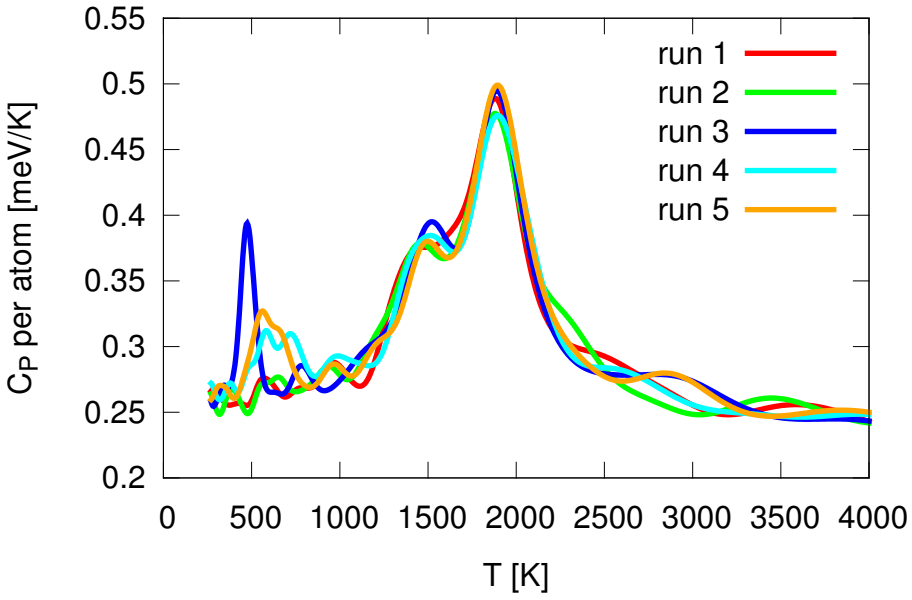
(b)

Fig. 4.2 The heat capacity against temperature curve convergence with different number of calls per iteration at 10 GPa. In both cases, we used 3840 walkers with 54 atoms and 10 GPa. The number of calls are 840 (a) and 3840 (b). n_{cull} was set to 1. We can see that convergence does not improve visibly with increase in the number of calls.

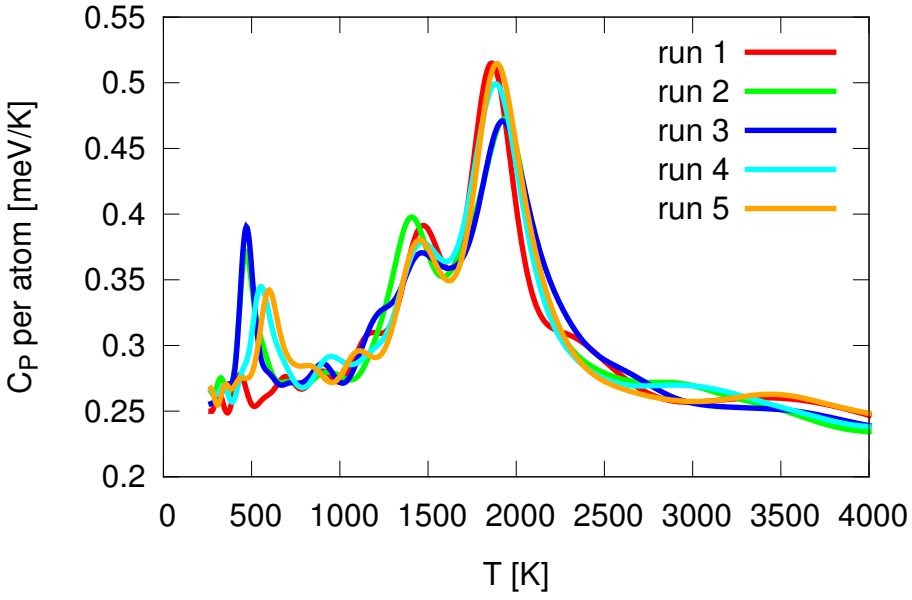
nested sampling in *pymatnest* [45]: Increasing the number of systems culled per iteration [69, 36]. I start with low numbers of n_{cull} as Noam Bernstein and Livia Bartók-Pártay have experienced convergence problems when using $n_{cull} \neq 1$ [113].
We see the results for two different n_{cull} values for 3840 walkers in Figure 4.3. Figure 4.3a shows five heat capacity curves for $n_{cull} = 4$ as well number of walkers $K = 3840$. We can see that the ω phase does not appear in all runs as the relevant peak (lowest temperature peak in the other runs) is missing for run 1 and 2. The ω to α transition peak of run 4 and 5 are very broad. Run 4 even shows a double peak and run 5 a shoulder at the ω to α transition point. Setting $n_{cull} = 8$ (Figure 4.3b) does not lead to a convergence either as run 1 does not show an ω to α transition. Figure 4.4a shows C_P curves for 7680 walkers while all other input parameters are kept the same. We can see that the ω phase appears in all runs. I have also doubled n_{cull} for Figure 4.4b and found that the ω phase appears in all runs. Finally, I show the case of $n_{cull} = 32$ in Figure 4.4c. Here, the convergence is severely affected. The runs 1 and 7 do not show the ω phase. For the runs presented in Figure 4.3a the ratio K/n_{cull} is 960 while it is 480 for the ones of Figure 4.4b. Even though that I cull more walkers per total number of walkers K per iteration in Figure 4.4b than in Figure 4.3a, Figure 4.4b shows convergence while Figure 4.3a does not show it. I conclude that the ratio K/n_{cull} is not the dominant factor defining whether the runs converge.

4.5 Phase diagrams with different number of atoms

We now consider the convergence of the phase diagram with nested sampling. For this, I conducted several nested sampling runs at different pressures for 24 and 54 atoms. After finding the phase transitions, we can use this data to draw a phase diagram [27, 20]. One can then investigate trends related to increasing the number of atoms. Additionally to the several 24 and 54 atoms runs, I conducted several runs for 108 atoms at 12 GPa. For 24 and 54 atoms, the high temperature solid phase is not as clear to identify via the XRD pattern as for 108 atoms. Some indication of a potential β phase can be found for 54 atoms. I extrapolate that this high temperature solid phase is indeed the β phase from my calculations for 108 atoms (see Section 4.5.2).
For each investigated combination of pressure and number of atoms, I conduct five independent runs. The phase transition temperatures are then estimated by the mean peak position of the relevant heat capacity peak. Additionally, I use the half width at half maximum (HWHM) to gain a measure of the error of our nested sampling calculation [47]. Sometimes, a phase transition does not occur for all runs. Also, for 4 GPa of Figure 4.5, an ω to α transition was discovered for all runs, but due to restart file corruption for one run (number 1), the peak

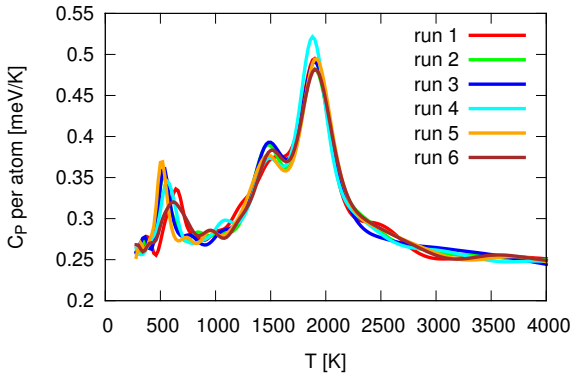


(a)

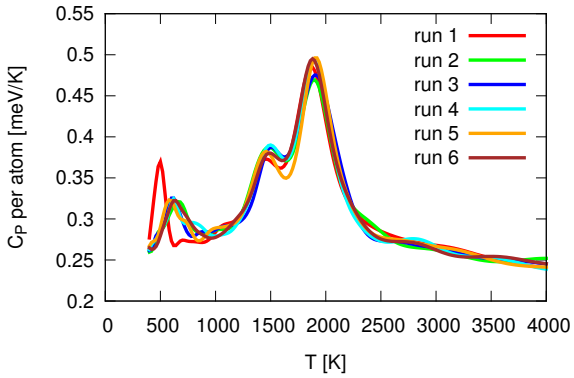


(b)

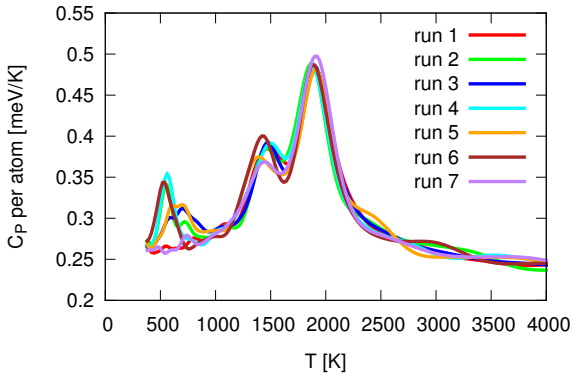
Fig. 4.3 n_{cull} tests for 3840 walkers and 54 atoms at 10 GPa. We see independent heat capacity curves for 240 calls per iteration and 3840 walkers for $n_{cull} = 4$ (a) and $n_{cull} = 8$ (b). Some runs show the ω to α phase transition in the vicinity of 500 K while others do not. This means that the runs are not converged.



(a)



(b)



(c)

Fig. 4.4 n_{cull} tests for 7680 walkers and 54 atoms at 10 GPa. Independent runs for $n_{cull} = 8$ (a), $n_{cull} = 16$ (b) and $n_{cull} = 32$ (c) are displayed in this figure. I use 240 calls per iteration. The convergence becomes worse with increasing n_{cull}. While Figure (a) and (b) always show the ω to α phase transition, Figure (c) does not display it for 2 of its 7 runs. If the ω to α transition appears, this is on average a little above 500 K. We can see that increasing n_{cull} can lead to full loss of convergence as shown for the case of $n_{cull} = 32$ (c).

position cannot be precisely determined for this specific run, as the C_P curve would need to be determined at lower temperatures. In these cases, the presented mean transition temperatures only take into account the occasions when the transition temperature can be determined. For comparison, I insert the prediction of Hennig *et al.* from Figure 5 of Ref. [15] into the phase diagrams.

Figure 4.5 shows the predictions for 24 atoms. I use 840 calls per iteration, 7680 walkers and

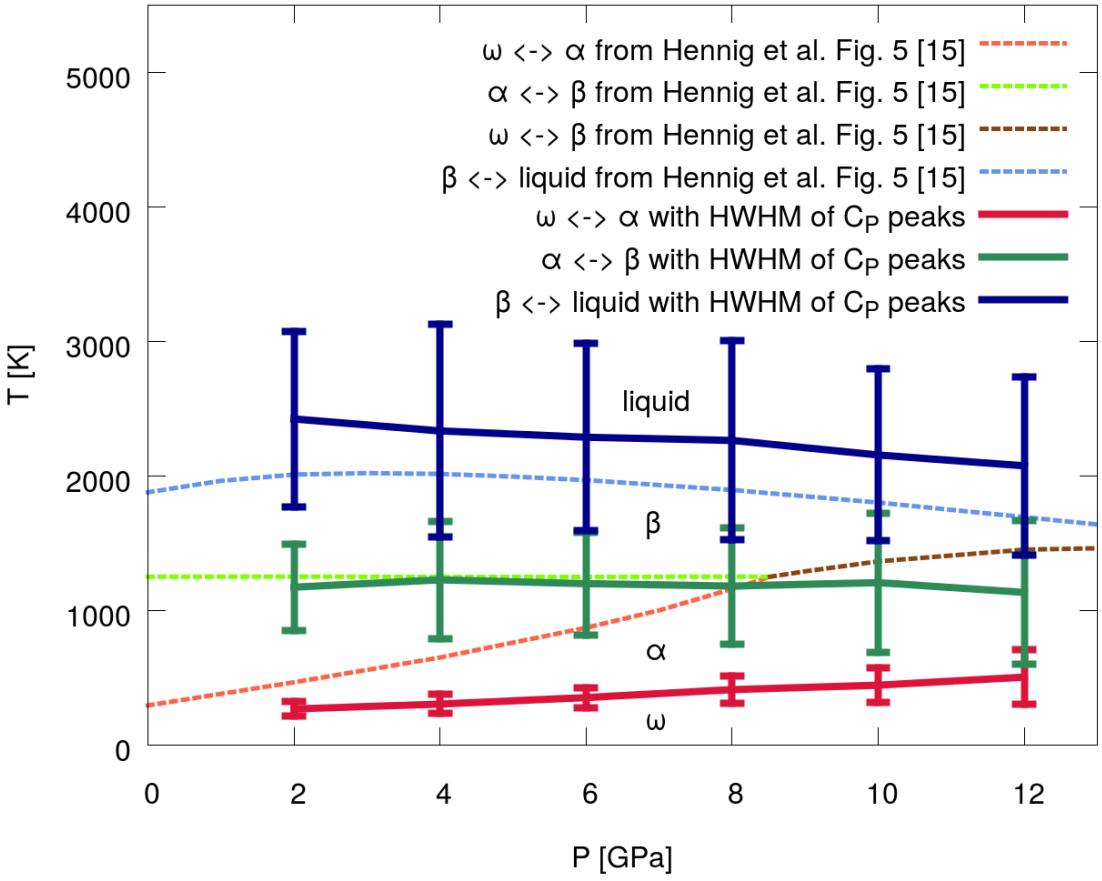


Fig. 4.5 The phase diagram for 24 atoms. The x-axis depicts the pressure in *GPa* while the y-axis shows the temperature in K. I use 840 calls per iteration, 7680 walkers and $n_{cull} = 1$ for all pressures but 2 GPa. For 2 GPa, I use 240 calls per iteration, 15360 walkers as well as $n_{cull} = 8$. I present the data of Figure 5 of Ref. [15] (dashed lines) together with my own results (solid lines). The HWHM is given by the error bars. The labels α, β, ω and "liquid" are for the phases of my nested sampling results. One run at 4 GPa was excluded from the determination of the ω to α peak position as explained in the text. In Ref. [15], Hennig *et al.* show a triple point which I can not find with my calculations.

culled one walker per iteration for the pressures greater than 2 GPa. For 2 GPa, convergence is not achieved with these parameters. One run out of the batch of five runs does not show the

ω phase but predicts the α phase at low temperatures for the above given parameters. For the converged results shown in Figure 4.5, I use 15360 walkers, 240 calls per iteration and culled 8 walkers per iteration. Solid lines show my own predictions while dashed lines represent the predictions of Hennig *et al.* [15]. The liquid to solid transition of our calculations (solid blue line) occurs at an offset of ca 300 K higher than we would expect from Hennig *et al.* [15]. We can see a slight decrease for the melting temperature from ca 2370 K at 2 GPa to ca 2080 K at 12 GPa for the nested sampling results. Apart from the aforementioned offset my prediction for the melting curve follows the predictions of Hennig *et al.* [15] considerably well. I predict (based on the 108 atoms calculations as explained at the beginning of this section) the β phase as highest temperature solid phase which again is in line with the calculations of Hennig *et al.* From 2 GPa to 8 GPa our α to β phase transition temperature (solid green line) for 24 atoms follows the predictions of Ref. [15] of a constant temperature of 1250 K well. However, I can not find a triple point at ca 1250 K and ca 8.5 GPa as Hennig *et al.* do in Figure 5 of Ref. [15]. Instead of a ω to β phase transition (dashed brown line) whose transition temperature is increasing with pressure, we continue to see an α to β phase transition (solid green line) with a relatively constant transition temperature. After my calculations, the ω phase sets in at ca 270 K at 2 GPa. This is much lower temperature than the ca 470 K at 2 GPa predicted by Hennig *et al.* One can see a slight increase in the ω to α transition temperature of my results up to 500 K at 12 GPa. The resulting slope of ca 24 K/GPa is visibly smaller than the one predicted by Ref. [15].

We now consider the case of 54 atoms. As the calculations are not fully converged for 2 GPa in the case of 24 atoms and 7680 walkers and needed 15360 walkers to converge at this pressure, I focus on higher pressures. As we can see in Figure 4.6, increasing the number of atoms has a limited effect on the phase diagram. Again my predictions are represented by solid lines, whereas dashed lines stand for the results of Ref. [15]. The different colours depict the same transitions as in Figure 4.5. My melting transition temperature (solid blue line), has shifted considerably and is now more in line with the predictions of Hennig *et al.* (dashed blue line). We can see an increasing difference between our predictions and their results with increasing pressure. This increases the difference from the predictions from Ref. [15] from ca 20 K at 4 GPa to ca 120 K at 12 GPa. The order of my own predicted phase transitions is the same as for 24 atoms. Again, I can not find the expected solid phase triple point of Hennig *et al.* [15]. Interestingly, the α to β transition for 54 atoms rose by more than 250 K from the one calculated for 24 atoms and predicted by Hennig *et al.* It is also not constant anymore, but shows a slight but clear decrease in transition temperature when approaching 12 GPa. Again, a ω to β phase transition does not exist. The ω phase does not appear at 4 GPa in my nested sampling runs. However, given that it only appears in one out of

five runs for 6 GPa and three out of five runs for 8 GPa, it could well be that a higher number of walkers can show the occurrence of the phase. Overall, the ω to α phase transition for 54 atoms increased slightly compared to 24 atoms. At 12 GPa the 24 atom calculations predict 504 ± 16 K while the 54 atom calculations predict 633 ± 12 K. (Note that the errors given here are the standard errors of the mean calculated from the respective five independent runs and are not the HWHM.)

Figure 4.7 shows five independent runs for 54 atoms (all black) and five independent runs

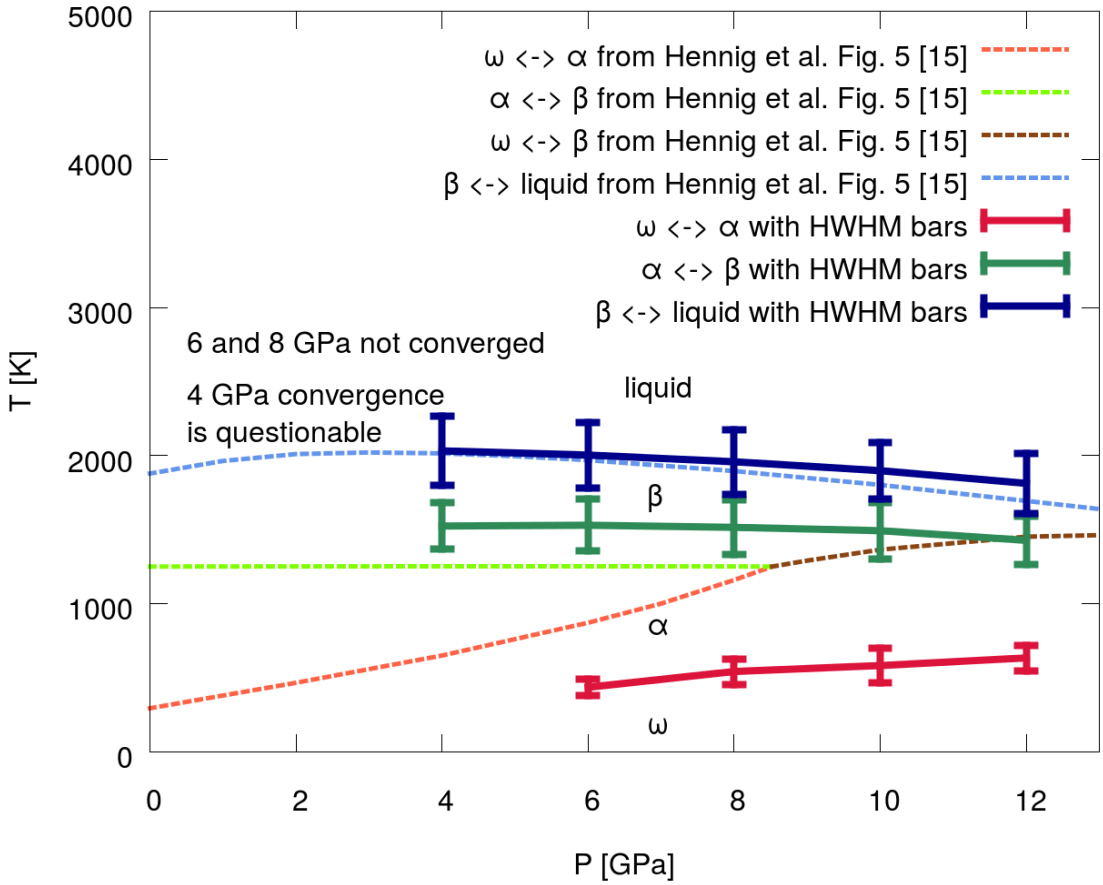


Fig. 4.6 The phase diagram for 54 atoms. Solid lines are the results calculated with nested sampling while dashed lines represent phase transition lines shown by Hennig *et al.* in Figure 5 of Ref. [15]. I use 840 calls per iteration for the 4, 6 and 12 GPa runs as well as 672 for the remaining pressures. n_{cull} is set to 1. All runs except the ones at 12 GPa use 15360 walkers. For 12 GPa, I use 7680 walkers. The convergence problems for 4, 6 and 8 GPa are explained in the text. The phase labels α, β ,ω and "liquid" represent again the nested sampling predictions. The triple point is missing as for 24 atoms in Figure 3.4

for 108 atoms (different colours) at 12 GPa. The 108 atom melting curve peak appears to be well converged at 1792 ± 9 K. This is in accordance with the melting curve value of 1812 ± 8

for the 54 atom runs. Again, I use the standard error of the mean of the five independent runs for the error estimates of these two melting temperatures. It is also interesting to note that the melting peak has become much sharper and higher for 108 atoms compared to 54 atoms. Similarly, the HWHM of the α to β transition is visibly smaller for 108 atoms. The average peak position for 108 atoms of 1410 ± 50 K is in line with my 54 atoms prediction of 1428 ± 11 K. However, is is important to note that the spread of individual peak positions (and heights) is much bigger than for 54 atoms as we can see from the error estimates for the average positions. This indicates convergence issues. Only run 1 and 5 show the ω to α transition with peaks at ca 835 K and ca 735 K respectively. For this case, the HWHM for 108 atoms are comparable with the ones for 54 atoms. This transition value is much greater than the average peak position found for this transition of 634 ± 12 K for 54 atoms. Here, again, I use the standard error of the mean.

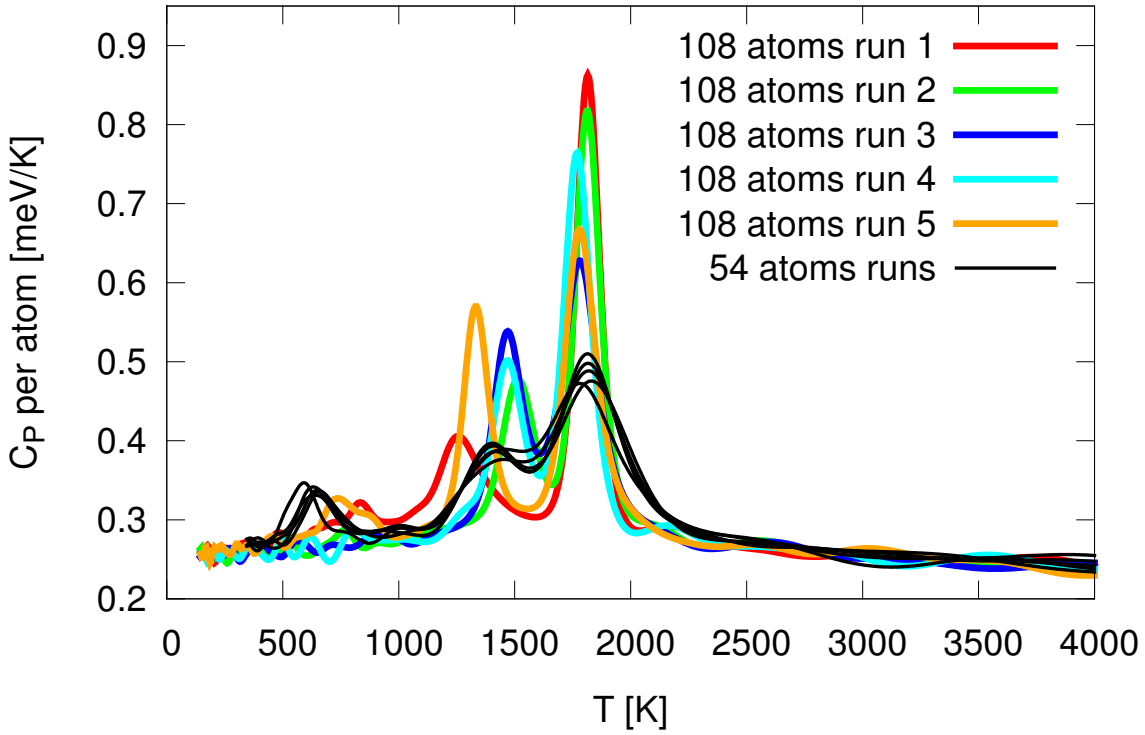


Fig. 4.7 Heat capacity versus temperature for 54 and 108 atoms at 12 GPa. For the 54 atoms runs (black), I use 840 calls and 7680 walkers with $n_{cull} = 1$ while I use 240 calls, 7680 walkers and $n_{cull} = 8$ for the 108 atoms runs (coloured runs).

4.5.1 Using MD to average atomic positions of trajectory frames

For high temperatures, I have found that phases are not always easy to define with the thermally averaged XRD pattern. For this case, I conduct what I call "MD averaging" in this thesis for the structures of the nested sampling trajectory file which have a nested sampling derived probability above a certain threshold. For each of these structures, I conduct a MD simulation with *QUIP* [114] at a temperature controlled by a Langevin thermostat [115] in the NVT ensemble. The lattice positions are averaged via an exponential moving average (EMA) conducted by *QUIP* [114] during the MD simulation. The average for the starting positions is just the start positions. For each following iteration i, the average of a quantity x becomes:

$$\langle x \rangle_{EMA,i} = \langle x \rangle_{EMA,i-1} \cdot e^{\frac{-\Delta t}{\tau}} + x_i \cdot \left(1 - e^{\frac{-\Delta t}{\tau}}\right) \tag{4.1}$$

Here Δt is the timestep which is set to 1 fs. x_i is the value of quantity x at iteration i. τ controls how quickly old quantities loose in importance for the average. It is set to 1000 fs for the calculations presented in Section 4.5.2 for which I also use 2000 MD steps. This process estimates the lattice positions around which the atoms fluctuate at high temperatures for each "MD averaged" system of the trajectory file.

4.5.2 Identification of phases for 108 atoms

In the following, I present an analysis for run 5 of the 108 atoms runs presented in Figure 4.7. In Figure 4.8, I show the clear identification of the ω and α phase as well as an indication of the β phase. The average volume predicted at 1600 K for run number 5 and the c/a ratios of Hennig *et al.* [15] for the ω and α phase define the structures for which the reference XRD patterns are calculated. Below the lowest temperature C_P peak of run 5, the system is in the ω phase. The XRD reference pattern of the ω phase (grey dashed lines) agrees very well with the the 400 K XRD pattern predicted for the system by run 5 (solid dark blue line). The 1000 K XRD pattern for the system (solid cyan line) is in general in accordance with the α XRD reference pattern. However, there are small peaks – for example one slightly below 60 °– which I attribute to a remaining ω phase while the α phase is dominant. For the 1600 K XRD pattern prediction (solid orange line) for the system, the identification is not as easy. The first three peaks of the β reference pattern (dashed magenta line) can be clearly seen. The forth peak coincides with an feature the is similar to a shoulder. At higher values of 2θ than this, there is no clear identifiable feature. Overall, it is still a good indication of the β phase. To show more clearly that the structure is indeed the β phase, let us consider Figure 4.9. Here, I present the XRD pattern at 1600 K for "MD averaged" trajectories as explained in Subsection 4.5.1 for run 5 (solid orange line) in comparison with the not "MD averaged"

trajectory XRD pattern (dashed blue line) and the reference β pattern (dashed grey line). I used MD trajectories of length 1000. The reference pattern and the not "MD averaged" pattern (both for 1600 K) are already shown in Figure 4.8. With the "MD averaging", the β phase can be clearly identified. All but the last peak of the reference structure slightly above 150 °can be identified.

4.6 Thermodynamic integration results with the original Ti MEAM

The problem with my nested sampling calculations in Section 4.5 is that they are conducted with relatively small numbers of atoms. Finite size effects could potentially explain the deviation of the phase diagram of Hennig *et al.* [15]. To exclude this I need to investigate the system with a larger number of atoms. In the following, I use thermodynamic integration to investigate the titanium phase diagram and verify my nested sampling results. As before, I use the MEAM of Hennig *et al.* [15] to describe the atomic interactions. I investigate the validity of the thermodynamic integration with quasi-harmonic approximation in Subsection 4.6.4. Before I present the results, I want to present the technical details including the parameters of both methods in the following. The parameters and methodology stays the same throughout this thesis.

4.6.1 Thermodynamic integration methodology

The thermodynamic integration method has been elaborated on in Section 2.6. Freitas *et al.* [48] elaborated on their implementation of a nonequilibrium version [49] in LAMMPS [74, 116]. In this thesis, I use this implementation following the instruction of Freitas *et al.* [48] for my calculations and setting up of my LAMMPS MD script which I use in the following. As they do, I combine a barostat [117, 67, 118, 119] with a Langevin thermostat [115] and do not allow cell shape changes. During this MD run, the cell shape is fixed. I have automated the process for calculating the Gibbs free energy of a structure via a python script. The number of runs, the different structures to be investigated, a list of temperatures and a list of pressures can be defined. I create the structure of interest via "quippy" [120] in python and write it as ".extxyz" file [121]. The starting structures used for the calculations of Chapter 4 are constructed using the volume and lattice constant ratios of energy minimised structures of Hennig *et al.* [15]. The conversion of ".extxyz" file to the LAMMPS format is done via a script by Albert Bartók-Pártay [122]. LAMMPS is then able to read the structure. I start a LAMMPS molecular dynamics (MD) *NPT* [115, 118] calculations to first equilibrate the

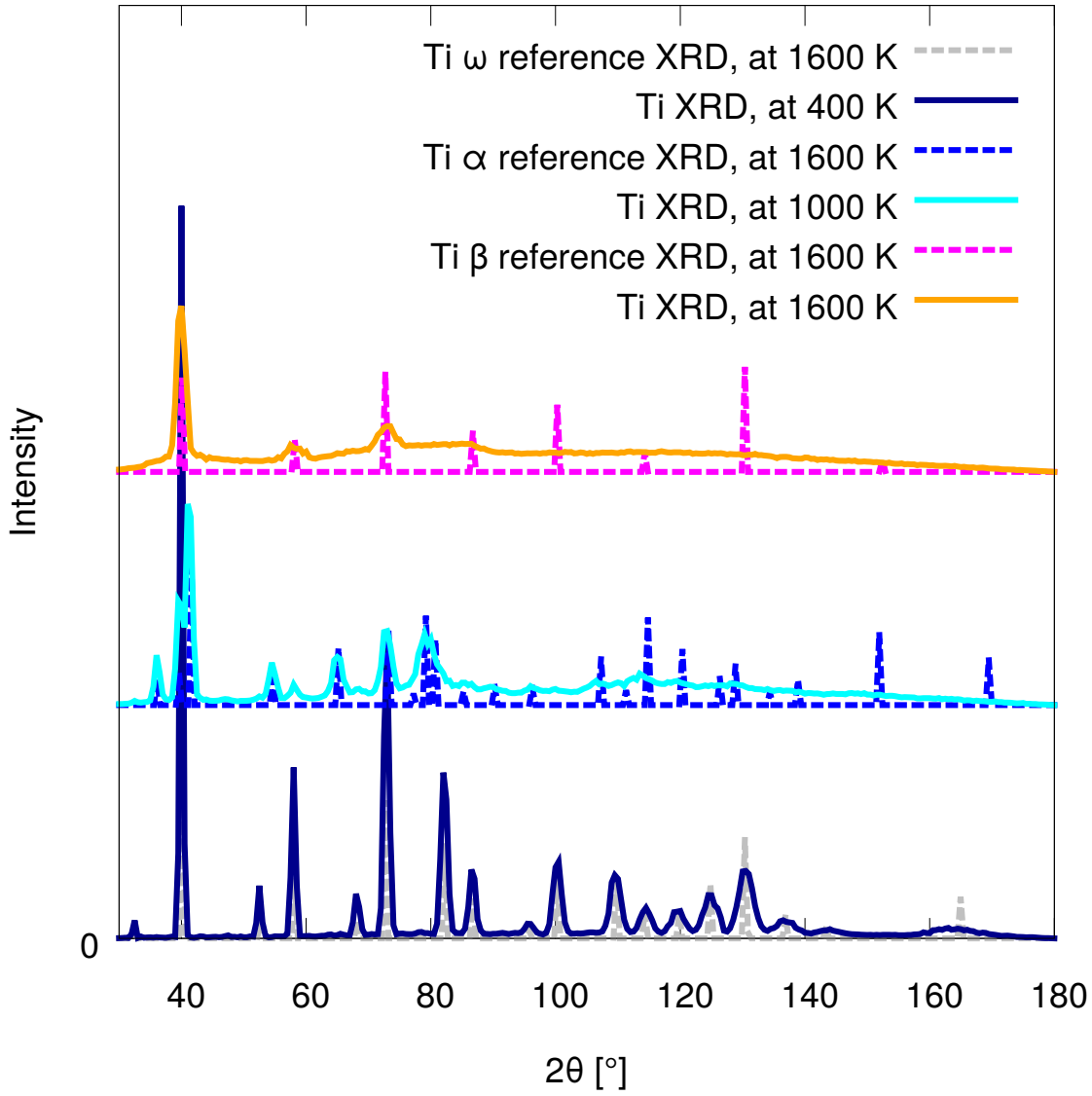


Fig. 4.8 The XRD pattern for the 108 atoms run number 5 whose C_P curve is shown in Figure 4.7 shown with differently coloured solid lines for 400 K (dark blue), 1000 K (cyan) and 1600 K (orange). The reference XRD pattern are denoted by dashed lines. Grey dashed lines represent the ω phase XRD reference pattern, blue dashed lines represent the α phase XRD reference pattern and magenta dashed lines show the β phase XRD pattern. An offset in intensity is introduced for the patterns. This makes it easier to determine the phases. When applying the offset, the XRD patterns of the different temperatures have been matched with the reference patterns which describe the phase. The reference patterns are created from structures with the average volume predicted by the run number 5 trajectory file at 1600 K. The c/a ratios used for the α phase and ω phase are the zero temperature and zero pressure values Hennig *et al.* present for their MEAM in Ref. [15].

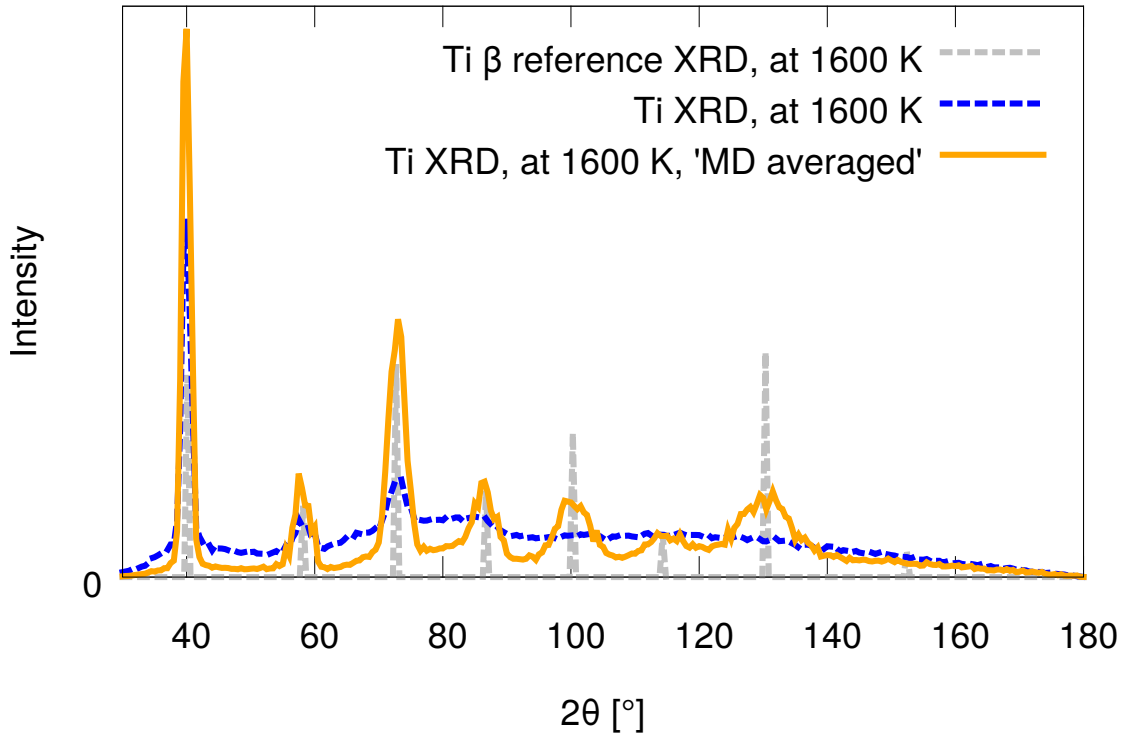


Fig. 4.9 “MD averaged” XRD pattern for 108 atoms Ti nested sampling run number 5 whose C_P curve is presented in Figure 4.7 at 1600 K in comparison to the XRD pattern of the same run at the same temperature without “MD averaged” and a β phase (bcc) reference pattern for 1600 K. The XRD pattern of the “MD averaged” trajectory file is denoted with solid orange lines. Dashed blue lines represent the XRD pattern without “MD averaging” as shown already in Figure 4.8. The β XRD reference pattern for the average volume at 1600 K according to run 5 is represented by grey dashed lines. The “MD averaging” leads to a much clearer XRD pattern which is in very good accordance with the reference pattern.

structure at the temperature and pressure of interest and after sufficient time has passed get the lattice constants for this temperature and pressure by averaging over a number of systems from the following MD trajectory. Then, my python script reads the averages from a file, creates the structure of interest with the averaged lattice parameters and prints it as ".extxyz" file using "quippy". This file is again transformed by using the transformation script of Albert Bartók-Pártay [122] again. In the following LAMMPS MD *NPT* [115, 118] simulation, the mean squared displacement (MSD) $\langle \|\vec{r}(t) - \vec{r}_0\|^2 \rangle$ is calculated after an equilibration time. Here, $\vec{r}(t)$ is the position vector of an atom over time t. The starting position vector of the atom is given by $\vec{r}_0$. "$\langle ... \rangle$" defines an average over all atoms in the simulation. The Mean Squared Difference (MSD) is used to calculate the spring constant $k_{Einstein}$ of the Einstein crystal [70] as in Ref. [48]:

$$k_{Einstein} = \frac{3k_BT}{\langle \|\vec{r}(t) - \vec{r}_0\|^2 \rangle} \tag{4.2}$$

I use the spring constant for a LAMMPS script of Albert Bartók-Pártay [123] which conducts the thermodynamic integration. I have slightly adapted the script to enable automation. This script conducts the thermodynamic integration along the Frenkel-Ladd path [72, 52] as described by Freitas *et al.* [48]. It is conducted in a NVT ensemble using a Langevin thermostat [115] as in Ref. [48]. The parameter $\lambda(t)$ from Equation 2.36 is changed with function number "2" [75, 76] (see Equation 2.50) which is elaborated on in Ref. [48]. In all LAMMPS simulations mentioned in this section, 1 fs long timesteps are used.
The results are integrated using a small program of Lívia Bartók-Pártay [124] to get the irreversible work according to Ref. [48]. The Einstein crystal is our reference state A. The system we are actually investigating is system B. With these notations, based on Equation 2.48 [48, 49], the definition of the Gibbs free energy (see Equation 2.61), the Free energy of the Einstein crystal as well as a correction term $F_{correction}$ [125, 48], the Gibbs free energy is then given by [48]

$$G = 3Nk_BT\ln\left(\hbar\sqrt{\frac{k_{Einstein}}{m}}\beta\right) + \frac{W_{irreversible,A\to B} - W_{irreversible,B\to A}}{2} + F_{correction} + PV \tag{4.3}$$

Here, I use that the Gibbs free energy G for a system relaxed at an isotrop pressure P is related to the Helmholtz free energy F by $G = F + PV$. Here, V is the volume. The free energy of the harmonic oscillator has been covered before in Section 2.7. Using this, the free energy of system A is

$$F_A = 3Nk_BT\ln\left(\hbar\sqrt{\frac{k_{Einstein}}{m}}\beta\right) \tag{4.4}$$

which leads to Equation 4.3. Here, I have used that the angular frequency of a harmonic is $\omega = \sqrt{\frac{k_{Einstein}}{m}}$ with m being the mass of the atom. N is the number of atoms in the simulation. Here, I use a mass of $m = 47.88\,\mathrm{u}$ [126]. The centre of mass is fixed during the MD calculations and hence one needs to introduce $F_{correction}$ [125] as explained by Freitas *et al.* [48] which is [125, 48]

$$F_{correction} = k_B T \ln\left(\frac{N}{V}\left(\frac{2\pi k_B T}{N k_{Einstein}}\right)^{3/2}\right) \tag{4.5}$$

I conduct the scheme described in this section at the same temperature and pressure several times. The mean of the Gibbs free energy and its standard error are calculated from these runs. To determine the transition temperature for a certain pressure, I calculate the difference ΔG_{at} in Gibbs free energy per atom between two competing phases calculated as in Equation 4.3 for a number of temperatures. If two so created data points which follow each other in temperature straddle the $\Delta G_{at} = 0\,\mathrm{eV}$ line, the phase transition occurs between them. I hence conduct a linear interpolation in between these two points. The intersection of the linear interpolation with $\Delta G = 0\,\mathrm{eV}$ line yields the transition temperature. The error is calculated by propagating the errors of the data points.
For the α and ω phase results presented in this thesis, I use $10,000$ steps each for the various equilibrations and $10,000$ steps each for each direction of the switching procedure. The thermal averages used in the procedure are calculated using values of the MD simulation after equilibration calculated every 10 steps during the last $10,000$ steps of the MD calculation. For the β phase, I use $1,000$ steps for the equilibration, $10,000$ steps for one switching direction each, as well as values for averaging every 10 steps over $40,000$ steps after equilibration. I use 8x8x8 supercells of the respective phases. For the α phase I use a unit cell with 2 atoms. For the ω phase, the cell contains 3 atoms. An 8x8x8 supercell of the cubic 2 atom unit cell is used for the β phase. I use 5 independent runs for each temperature.

4.6.2 Quasi-harmonic approximation methodology

The quasi-harmonic approximation is used to test the validity of my thermodynamic integration calculations. To calculate the phonons and free energies via quasi-harmonic approximation, I used *Caesar* [127]. It checks for symmetries of the lattice using *SPGLIB* [128] to lower the amount of calculations. It uses *QUIP* [114] via a wrapper to evaluate the forces to calculate the force constants necessary for the phonon calculation. I have added the classical calculations of the phonon free energy for a given single phonon frequency to my personal version of *Caesar*. The phonon frequencies are then sampled randomly to

calculate the free energy of the system. For the sampling, I use $100,000$ samples. The mass of titanium is set to 47.867 u [129]. To approximate the Gibbs free energy of the α and ω phase, I construct a grid of volumes V and lattice constant ratios c/a in the following way: I relax structures at 24 pressures during two different runs. First, I start at 0 GPa decreasing in 1 GPa steps down to -6 GPa. Then, I start at 1 GPa increasing in 1 GPa steps until I finish the relaxation for 17 GPa. For 0 GPa and 1 GPa, I conduct the relaxation starting from structures constructed with the volumes and c/a ratios calculated by Hennig *et al.* [15] in their relaxation at 0 GPa and 0 K for the Ti MEAM [15]. For the next pressure in each run, I use the result of the previous relaxation as starting value. The relaxation is done via "quippy" [120] using a limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method [130, 131]. Each of the V values and c/a values gained from this relaxation marks a position on the volume and c/a axis. Together, these axes form the aforementioned grid. In my specific case that means that it is a 24×24 grid. For each point of the grid, I create the structure in question using "quippy" [120] via the functions I created from *misc_calc_lib.py* [45] (see Section A.1 for more general details of that library including the partial re-implementation of a function originally written by Noam Bernstein from *pymatnest* [45]). Then, I relax the atomic positions using an ASE [132] implementation of a Broyden-Fletcher-Goldfarb-Shanno (BFGS) method [62] for atomic relaxation while not changing the lattice. Having prepared the structures like this, I use *Caesar* [127] to get the free energies. The supercell is calculated in such a way that the minimum distance between an atom and its periodic images is longer than twice the effective cutoff for force calculations leading to $8 \times 8 \times 8$ supercells for the 2 atom hcp unit cell and $7 \times 7 \times 7$ supercells for the 3 atom ω cell. For this, I use the aspect ratio function of *misc_calc_lib.py* which is a partial re-implementation of a similar function by Noam Bernstein found in *pymatnest* [45] (see Section A.1 for more details). The Gibbs free energy is then calculated via Equation 2.62 as explained in Section 2.7. The reference system is always the system of the relaxed systems at various pressures from the start with the same volume V_0 as the current volume V. The free energy predictions are however different depending on whether one uses the classical or quantum oscillator. Hence, the Gibbs free energy has to be calculated separately for the classical and quantum treatment.

4.6.3 MD equilibration to test the stability of phases

I ran MD simulations at a number of temperatures to determine the stability of the ω phase for some pressures. This is mainly relevant to Subsections 4.8.1 and 4.8.2, but is also considered briefly in Subsection 4.6.4. Such calculations are conducted for bcc by Hennig *et al.* [15]. In the following I will refer to this as "MD stability tests". I did this by running MD simulation

for 10^6 MD steps of 1 fs. I use the LAMMPS MD script mentioned in Section A.1 which follows the instructions on how to do the MD simulations by Freitas *et al.* [48]. Accordingly, I apply a Langevin thermostat [115] in combination with a barostat [117, 67, 118, 119] as they do [48]. Now however, I allow the cell shape to change during the simulations. The XRD pattern is averaged over samples at the end of the equilibration and used to identify the phase. I use *QUIP* [114] for calculating the XRDs called via *misc_calc_lib.py* [45] (see Section A.1). c/a values for the MEAM of Hennig *et al.* [15] from Ref. [15] are used in the calculation of the reference structures. The starting structure is created using the parameters of the relaxed structure given in Ref. [15] with the original MEAM [15]. I use a mass for Ti of 47.88 u [126].

4.6.4 Comparison to quasiharmonic approximation

To validate my thermodynamic integration results, I compare it to results of quasiharmonic approximation calculations using Caesar [127] which calls QUIP [114] for the potential evaluation (see Subsection 4.6.2). The parameters and technical details of the thermodynamic integration have been explained in Subsection 4.6.1 (for parameters see the end). As mentioned earlier in Section 2.7, the thermodynamic integration method used in this thesis does treat the system as fully classical system. This means that to compare thermodynamic integration to quasiharmonic integration, one has to approximate the phonons via the classical harmonic oscillator and not the quantum mechanical harmonic oscillator. In Figure 4.10, we see the α phase Gibbs free energy per atom of thermodynamic integration (orange crosses), of the quasiharmonic calculations based on the quantum harmonic oscillator (dashed magenta line) and of the quasiharmonic calculations based on the classical harmonic oscillator against the temperature (solid blue line). The temperature is given in Kelvin while the Gibbs free energy is given in eV. The pressure is 12 GPa. The statistical error for the thermodynamic integration results is calculated from 5 independent runs and is between $6.4 \cdot 10^{-6}$ eV and $1.4 \cdot 10^{-4}$ eV for the temperature values up to including 800 K. This makes the error smaller than the line size of the symbols. We see that the results of the quasiharmonic calculations based on the classical harmonic oscillator are in very good accordance with the thermodynamic integration predictions for low temperatures. For 200 K, the difference is still below 1 meV. For 250 meV it is bigger than 1 meV. The quantum mechanical treatment is important for low temperatures as we can see from the visible difference to the thermodynamic integration and classical harmonic oscillator quasi-harmonic approximation results at 150 K and lower. However, the difference between the two quasiharmonic results becomes small relatively quickly with increasing temperature for low temperature though the decrease is flattening up again relatively soon. At 400 K the quantum mechanically treated quasiharmonic

calculations is already less than 2 meV higher than the classically treated ones. However, at 800 K the quantum mechanical quasiharmonic Gibbs free energy is still more than 0.5 meV higher. In comparison, at 400 K, the thermodynamic integration Gibbs free energy is already more than 0.5 meV higher than the quantum mechanical quasiharmonic approximation treatment. At 800 K the thermodynamic integration Gibbs free energy is already more than 5 meV higher than the quantum mechanical treated quasiharmonic integration. This shows the problems of the quasiharmonic approximation for even intermediate temperatures.

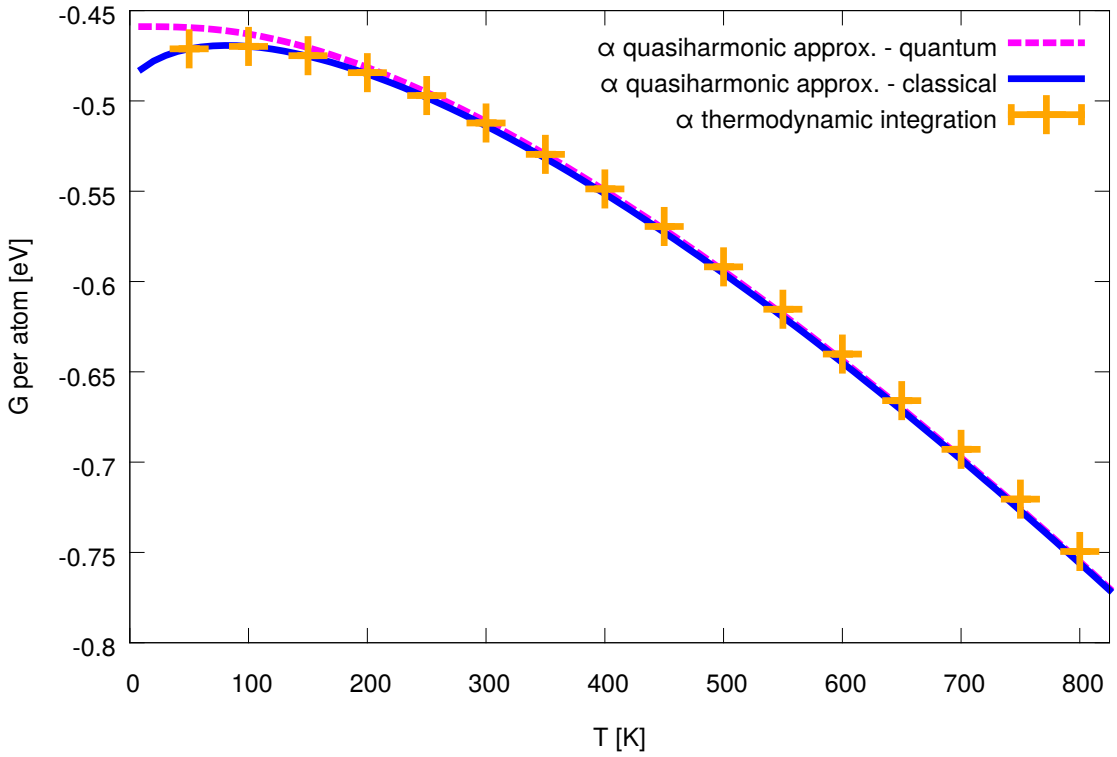


Fig. 4.10 The Gibbs free energy per atom of the α phase versus temperature using different methods at 12 GPa. The results of the quasiharmonic calculation are given with magenta dashed lines and blue solid lines. The former are the results for modelling the phonons as quantum harmonic oscillator and the latter the results for modelling the phonons as classical harmonic oscillators. The thermodynamic integration results are given by the orange crosses. The statistical errors of the thermodynamic integration are smaller than the line width of the symbols. We see that for low temperatures, the thermodynamic integration results are in accordance with the quasiharmonic calculations using a classical harmonic oscillator. The neglect of the zero-point energies and other quantum contributions leads to a significant error for low temperatures, as seen by the difference to the quasiharmonic calculations which do not neglect these contributions.

The previous results in Section 4.5 show that the competing phases are β, α and ω for the used MEAM [15] as also stated in Ref. [15]. In Figure 4.11, we see the difference between

the Gibbs free energy per atom of the α (G_α) and ω phase (G_ω), $\Delta G = G_\alpha - G_\omega$, at 12 GPa. The parameters used are the same as the ones given at the end of Subsection 4.6.1. An MD simulation with 10^6 MD steps of 1 fs with LAMMPS [74] as explained in Subsection 4.6.3 shows that an initially set up ω phase transforms at 1400 K to the β phase while it is still in the ω phase after 10^6 steps at 1350 K. These MD simulations are started with an ω phase with the same volume per atom and c/a ratio as the volume per atom and the c/a ratio found by Hennig *et al.* [15] for the original Ti MEAM [15]. Hence, I only show thermodynamic integration results until 1350 K in Figure 4.11. For temperatures below 731 ± 12 K, we can see that the ω phase has the lower Gibbs free energy of the two phases and is hence the stable phase. At 731 ± 12 K the ω phase transforms into the α phase. I calculate the transition temperature as explained in Subsection 4.6.1. Starting from this transition, until including 1350 K the α has a lower Gibbs free energy than the ω phase. For the next temperature value of 1400 K, the ω phase is already unstable under an MD simulation as mentioned before and hence not a candidate for the dominant phase. The thermodynamic integration method used in this thesis can not calculate its Gibbs free energy anymore in this case. One notes that the absolute difference in Gibbs free energy between the α and the ω phase is very small throughout the temperature range for which the α phase is stable. The maximum absolute Gibbs free energy difference per atom difference for the simulations presented in Figure 4.11 during the temperature range over which the α is stable is 4.03 ± 0.15 meV at 1150 K. This is lower than the absolute difference to either form of the quasiharmonic calculations at 800 K for the α phase presented in Figure 4.10. This indicates that one needs to conduct thermodynamic integration to get reliable results in this temperature realm.

4.6.5 Thermodynamic integration phase diagram for the original Ti MEAM

In Figure 4.12, I compare the phase diagram predicted by Hennig *et al.* [15] in Figure 5 of Ref. [15] with the solid phase diagram predicted for their potential with the help of thermodynamic integration (TI). Grey lines and labels depict the results of Hennig *et al.* The orange upwards pointing triangles depict the α to β phase transition. The parameters for the thermodynamic integration are given at the end of Subsection 4.6.1. The blue stars signify the ω to α phase transition. The results of Hennig *et al.* differ significantly from mine even though they were conducted with the same potential. I cannot reproduce the triple point which they show at ca 8.5 GPa and ca 1250 K in their phase diagram [15]. I do not find any ω to β transition in my phase diagram prediction. The thermodynamic integration results show the following ordering of phases with increasing temperature: ω, α and β. This ordering

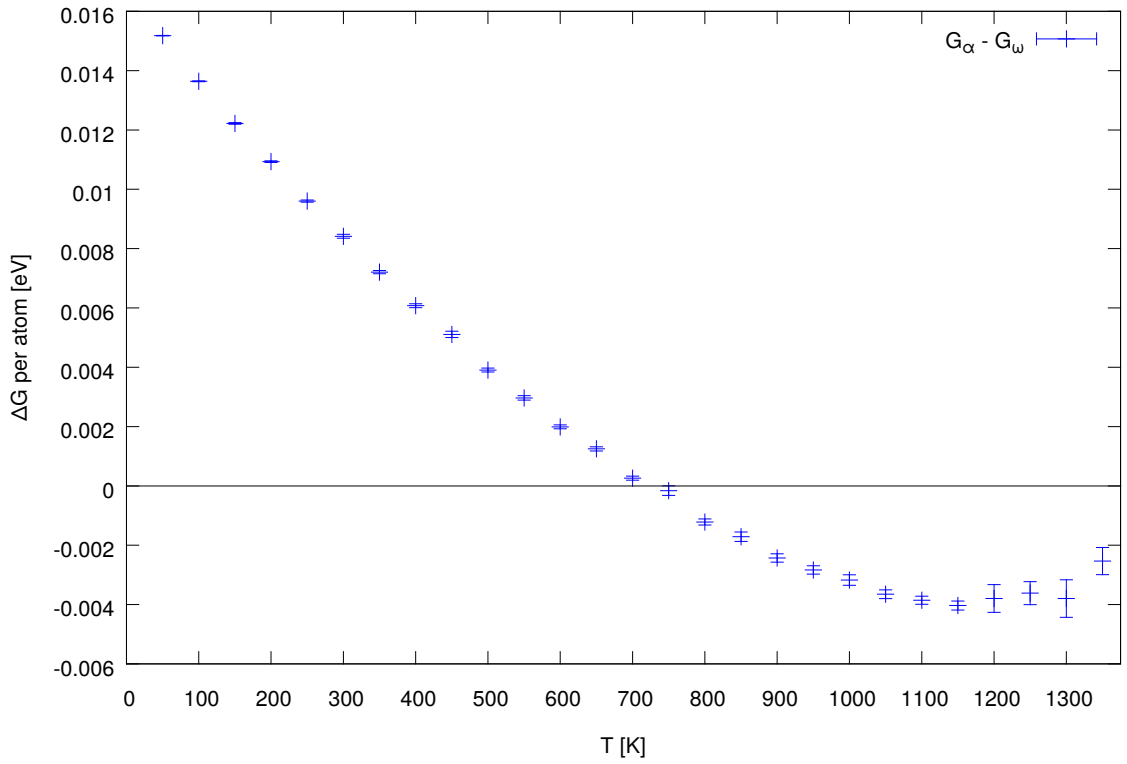


Fig. 4.11 The difference between Gibbs free energy per atom $\Delta G_{atom} = G_{\alpha,\,atom} - G_{\omega,\,atom}$ for the α and ω structures against the temperature at 12 GPa. $G_{\alpha,\,atom}$ and $G_{\omega,\,atom}$ are the Gibbs free energies per atom of the α and the ω phase respectively. Thermodynamic integration can only be used as long as the structure are stable under molecular dynamics simulations. Hence, the difference in Gibbs free energy is only given until 1350 K. We see that the ω phase is stable for low temperatures. At 731 ± 12 K, titanium transforms to the α phase.

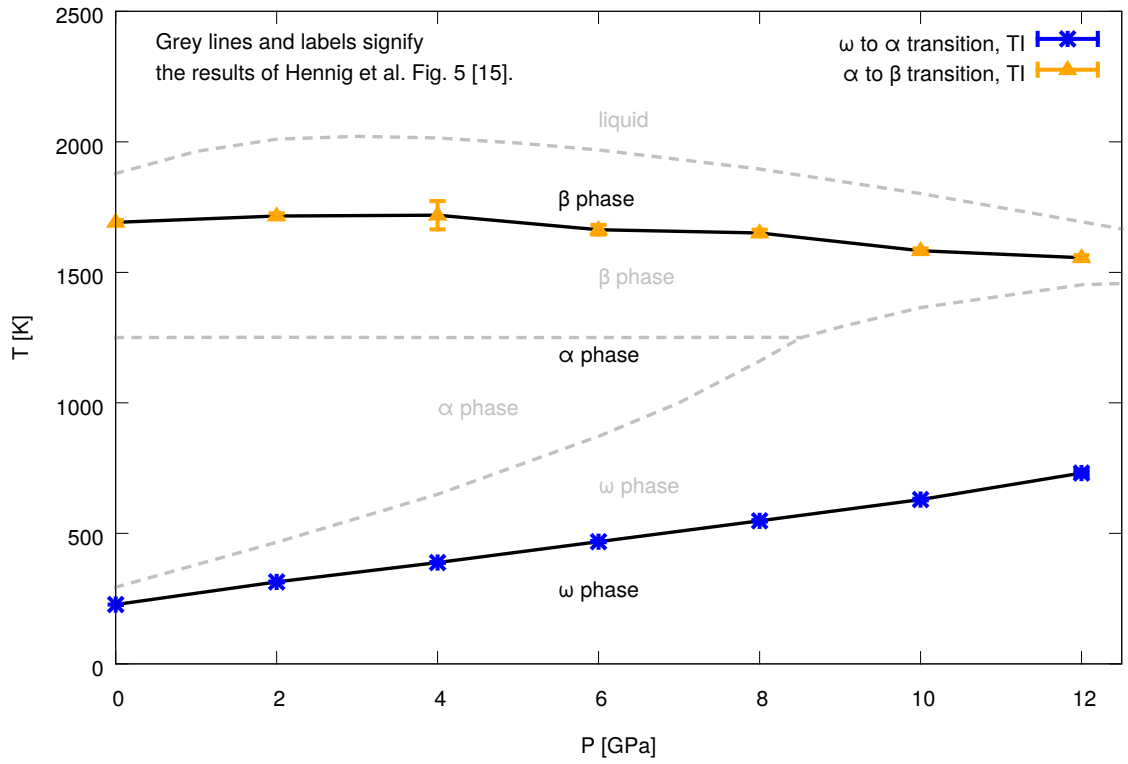


Fig. 4.12 The solid-solid phase transitions for the original MEAM [15] predicted by my thermodynamic integration calculations in comparison to the predictions by Hennig *et al.* from Figure 5 of Ref. [15]. The grey labels and lines depict the results of Hennig *et al.* [15]. The other data points are gained from my own thermodynamic integration calculations (TI). Blue stars are used for the ω to α phase transition. Orange upwards pointing triangles are used for the α to β transition. We see significant differences. Most importantly, my thermodynamic integration results do not show a triple point.

is at least in line with the result of Hennig *et al.* [15] for pressures below ca 8.5 GPa. Note that Hennig *et al.* state a pressure of ca 8 GPa and a temperature of ca 1200 K in the text of their paper for the triple point [15]. For 0.0 GPa, Hennig *et al.* show an ω to α transition temperature close to ca 300 K in their diagram. This is not that far from the thermodynamic integration result of 227.1 ± 0.7 K. However, the slopes of the ω to α transitions are very different which explains in part why one never sees a triple point in the depicted pressure realm. The α to β transition temperature is relatively constant for my predictions showing an only slight decrease with pressure after 4 GPa. This is similar to the constant transition temperature behaviour found by Hennig *et al.* However, I predict a transition temperature which is considerably higher: From 0 GPa to 8 GPa the values range between ca 1650 and 1720 K for the thermodynamic integration results while Hennig *et al.* found a value of ca 1250 K [15].

4.6.6 Comparison of thermodynamic integration results with the original MEAM to experimental results

We have seen in Section 4.6.5, that I predict a phase diagram which is very different from the one predicted by Hennig *et al.* in Ref. [15]. Let us now compare the thermodynamic integration predictions to predictions made by Zhang *et al.* [51] based on experimental data. Both are depicted in Figure 4.13. The phase transition lines of Zhang *et al.* [51] are represented by dashes line with grey short dashes. Grey labels depict the different solid phases. The thermodynamic integration results are denoted by blue stars for the ω to α transition and orange upwards pointing triangles for the α to β transition. I have also added the melting curve fitted to experiments by Errandonea *et al.* [81] (fit data from Ref. [82]). This melting curve is denoted by a dashed line with long dashes. We see that the potential of Hennig *et al.* [15] misses the triple point according to the thermodynamic integration calculations. As opposed to the experimental prediction of Zhang *et al.* [51], the α phase is stable over at least a certain temperature range for all investigated pressures. A transition from ω to β is therefore not predicted by thermodynamic integration. Additionally the α to β transition line of Zhang *et al.* is considerably lower than my thermodynamic integration based predictions. The difference in slope is not as strong. The experimental predictions according to Zhang *et al.* [51] show an only slightly negative slope. The thermodynamic integration results predict a relative constant value until 4 GPa from which they too show a slightly negative slope. Zhang *et al.* predict the α phase to be stable at low temperatures and pressures. This is in opposition to my thermodynamic integration predictions with the original MEAM [15] and the predictions of Hennig *et al.* [15] which both find the ω to be

stable at low temperatures and pressures. Additionally the ω to α transition line predicted by the MEAM of Hennig *et al.* [15] is much steeper.

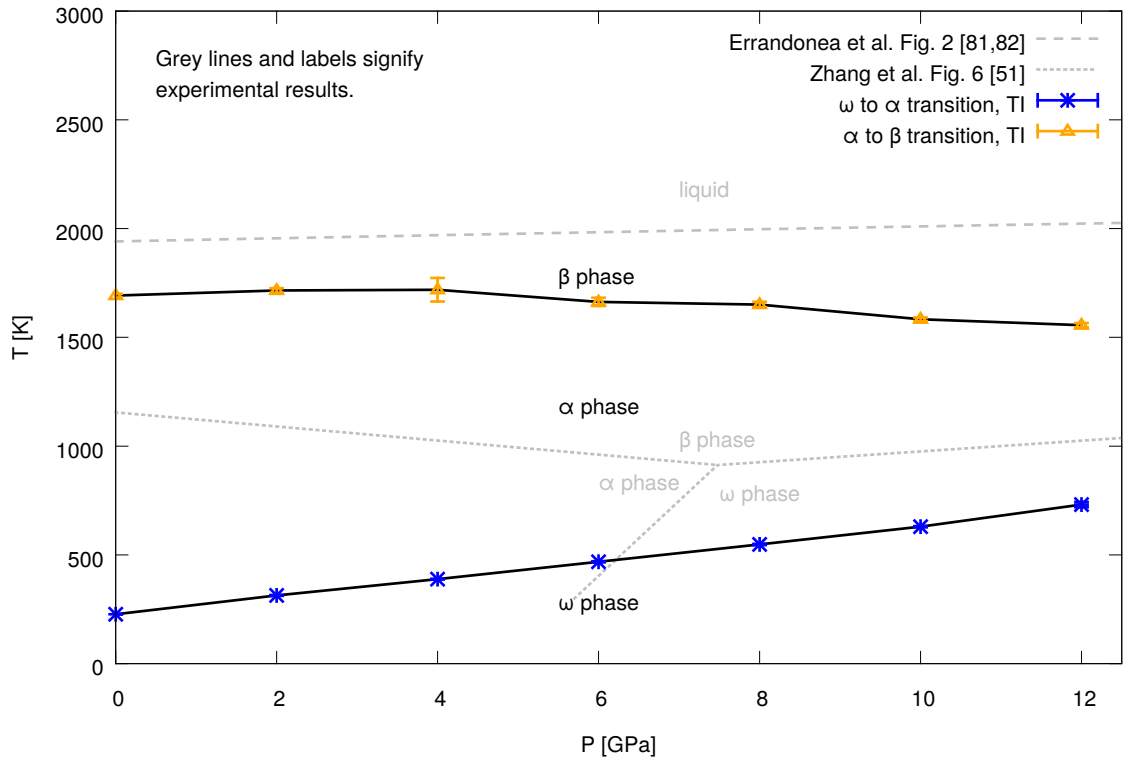


Fig. 4.13 The solid-solid transition phase diagram of the original MEAM [15] based on my thermodynamic integration calculations in comparison to experimental results. The solid-solid predictions of Zhang *et al.* [51] are given by the grey short dashes. The melting curve of Errandonea *et al.* [81, 82] is signified by the grey long dashes. The grey labels denote the phases predicted by the experiments. The blue stars are the data points of the ω to α phase transition calculated with thermodynamic integration. I also use thermodynamic integration for the α to β transition for which I use orange upwards pointing triangles. The triple point was not predicted by my calculations.

4.7 Comparison of original Ti MEAM energies to DFT results

We now consider how the original MEAM potential [15] compares to DFT calculations with CASTEP [133] for the different structures of the phase diagram with the "PBE" generalized gradient approximation (GGA) functional [134, 135]. More details of the DFT calculations can be found in Section A.2. This comparison could potentially shed light on why my

predictions for the solid part of the phase diagram with the original potential of Hennig *et al.* [15] are so different from the experimental predictions of Zhang *et al.* [51]. The evaluation of the MEAM is done using *quippy* [120]. However, instead of comparing to ideal structures and defects as in Ref. [15], I consider snapshots of thermalised structures which I gain from nested sampling runs. In detail, for 2 GPA, 6 GPA and 10 GPA, I take the 10 most likely structures from the trajectory file of one nested sampling run for each pressure with converged parameters at four different temperatures. The trajectory files contain every 100-th culled system. Here, I use 24 atoms per simulation box, $n_{cull} = 8$ and 240 calls per iteration for all runs. For 2 GPa, I use 15360 walkers. For the other pressures, I use 7680 walkers. For

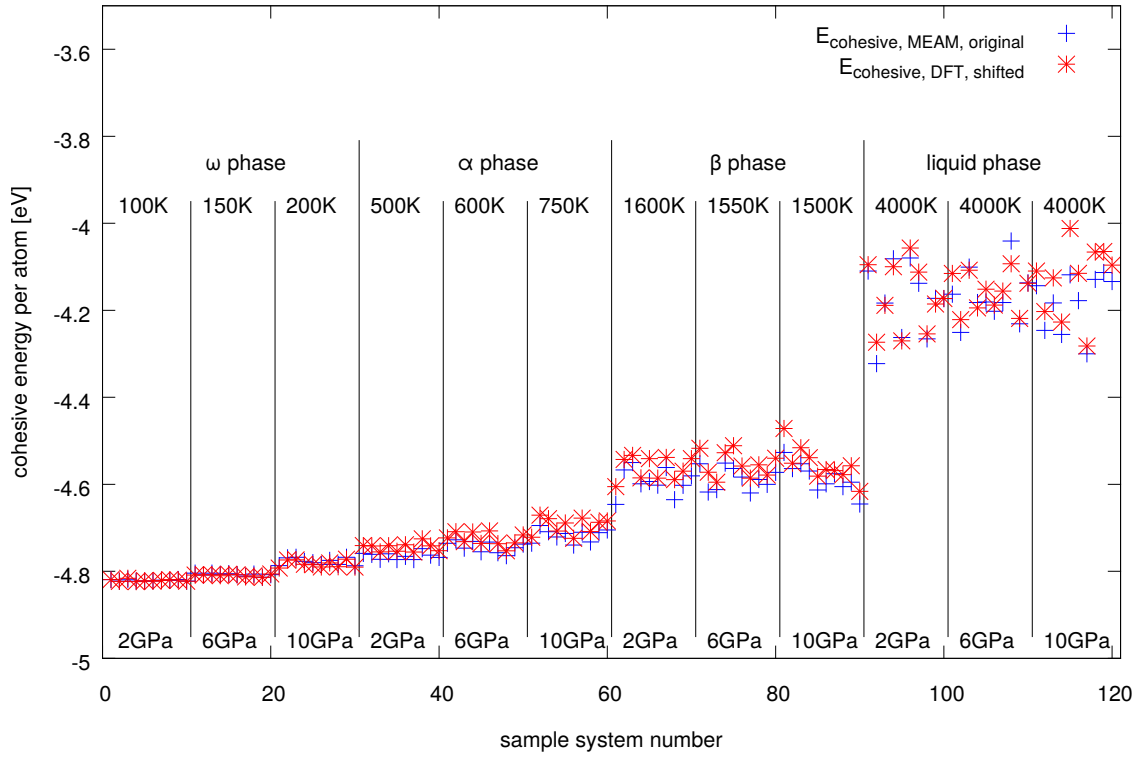


Fig. 4.14 The cohesive energies per atom for a selected number of structures at different temperatures and pressures calculated with DFT and with the original MEAM of Hennig *et al.* [15]. The DFT energies are denoted by red stars. Blue crosses describe the MEAM energies. A value of ca 1.7354 eV is added to the DFT energies. This is the difference between the average energy of the ω structures at 2 GPa for the DFT energies and the MEAM energies. This way, one is able to compare the different results. Qualitatively, they are in good agreement. However, there are some slight deviations which I cover in Figure 4.15.

the temperatures, I choose 100 K for 2 GPa, 150 K for 6 GPa and 200 K for 10 GPa for the ω phase structures. For the α phase structures, I choose 500 K for 2 GPa, 600 K for 6 GPa and 750 K for 10 GPa. For the β phase structures, I choose 1600 K for 2 GPa, 1550 K for 6 GPa

and 1500 K for 10 GPa. For the liquid phase I choose 4000 K for all pressures. The cohesive energy results for the MEAM and the DFT calculations are presented in Figure 4.14. Blue crosses show the results for the MEAM of Hennig *et al.* [15]. Red stares signify the results using DFT. The DFT parameters are given in Section A.2. I add a value of ca 1.7354 eV to the DFT energies so that the average energy for the ω phase structures are the same for both DFT and MEAM calculations for 2 GPa. Qualitatively, the results of the MEAM calculation are similar to the DFT results. The energies of the structures as well as the width of the energy distributions for each structure increase while one goes from the low temperature ω phase to the liquid. Qualitatively, the MEAM calculations are in overall good agreement with the DFT results, though there are some slight deviations.

Figure 4.15 shows the difference between the cohesive energies using the MEAM of Hennig *et al.* [15]

$$E_{cohesive,MEAM,original} \tag{4.6}$$

and the shifted DFT cohesive energies

$$E_{cohesive,DFT,shifted} \tag{4.7}$$

presented in Figure 4.14. The vertical axis shows

$$\Delta E = E_{cohesive,MEAM,original} - E_{cohesive,DFT,shifted} \ . \tag{4.8}$$

The difference for the ω phase is on average only a slight overestimation for the structures gained for 6 and 10 GPa. The fact that the difference is very small is not surprising as the DFT energies have been shifted so that there is not any difference for the average ω phase energy for 2 GPa. From this point of view, the α and the β phase are underestimated by the MEAM. The liquid comparison is not as clear for 2 GPa and 6 GPa as the solid comparison due to the high spread of the energy value distribution. On average, the values point to a slight underestimation for 2 GPa and 6 GPa and a stronger as well as clear underestimation for 10 GPa. Given that the choice of reference value for the DFT value is in a certain way arbitrary, one could argue that the ω phase energy is overestimated by the original MEAM [15] compared to the other solid phase structures. This may explain in parts the missing triple point which one would expect from the results of Hennig *et al.* [15]. The ω phase may just not be stable enough compared to the α phase to supplant it fully at high pressures. However, it is to note that this explanation neglects the influence of phonons and the pV term of the Gibbs free energy which could counteract the overestimation of the ω phase energies.

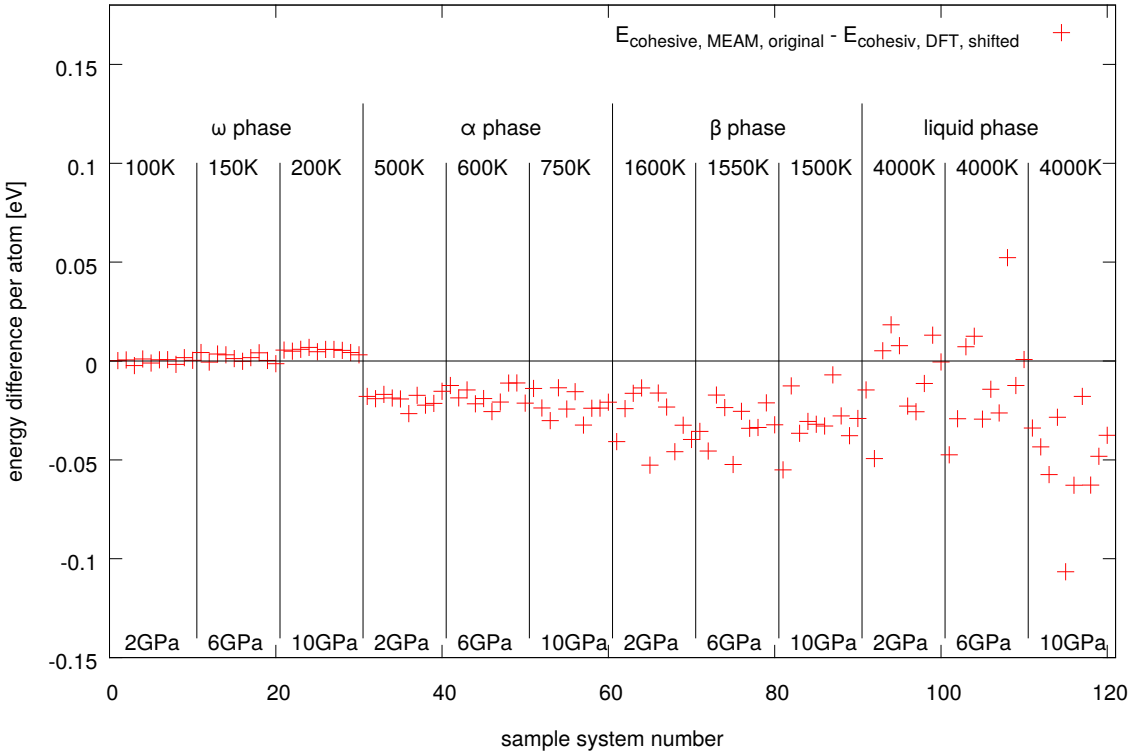


Fig. 4.15 The differences $\Delta E = E_{cohesive,MEAM,original} - E_{cohesive,DFT,shifted}$ between the cohesive energies when using the MEAM of Hennig *et al.* [15] and DFT cohesive energies ($E_{cohesive,MEAM,original}$ and $E_{cohesive,DFT,shifted}$ respectively) from Figure 4.14 sorted by phase. The ω structure energies of 2 GPa are used as reference point as explained in the text. The MEAM overestimates slightly the DFT calculations for the higher pressure ω structures. With this reference point, the MEAM underestimates the α and β phase. The liquid phase is not as clear as the spread of energy values is relatively large. The data points on average towards an overall slight underestimation by the MEAM.

4.8 Modifying the angle dependent part of the MEAM

I have shown that the Ti MEAM potential of Hennig *et al.* [15] does not correctly predict the triple point of titanium [51]. I now conduct small modifications to the original MEAM of Hennig *et al.* [15] with the aim of creating a modified potential which does predict the triple point relatively close to the experimental prediction [51]. The modification to the spline parameters of the MEAM of Hennig *et al.* [15] is conducted with Fourier modes to reduce the amount of fitting parameters while still changing the function over the relevant realm. In principle, more Fourier modes could approximate more complex changes in the future. The modifications presented in this thesis are meant to explore the possibility of fitting. Hence the changes are only limited to the functions $f\left(r_{ij}\right)$ and $g\left(\cos\left(\theta_{jik}\right)\right)$ of electron density function definition of the titanium MEAM of Hennig *et al.* [15, 61]:

$$\rho_{h,i} = \sum_{j \neq i} \rho_a\left(r_{ij}\right) + \sum_{j \neq i, k \neq i} g\left(\cos\left(\theta_{jik}\right)\right) f\left(r_{ij}\right) f\left(r_{ik}\right) \tag{4.9}$$

The reasoning is that only slight changes to the potential may be necessary to reproduce the experimental phase diagram as the potential does in some aspects already describe parts of the experimental phase diagram well as I show in Section 4.6.6. The overall behaviour of the potential is not meant to be changed much. Hence, I only change the part of the potential which describes the effect of how bonds between atoms interact.
In detail, I add the respective values of the Fourier modes to the original values of the functions f and g [15] at the specific spline node points. For this process I use *MATLAB* [136]. I base all parameters on the parameter file shipped with *LAMMPS* [74, 137]. This is not exactly adding Fourier modes, as the spline interpolation will lead to difference from the values one would get by adding the Fourier inspired modes to the respective functions. For the function $g\left(x\right)$, I use the the modes:

$$m_g\left(x\right) = A\left(\cos\left(\frac{2\pi x}{L_g}\right) + \cos\left(\frac{4\pi x}{L_g}\right)\right) \tag{4.10}$$

and for $f\left(x\right)$, I use:

$$m_f\left(x\right) = A\left(\sin\left(\frac{2\pi\left(x - x_{max,f}\right)}{L_f}\right) + \sin\left(\frac{4\pi\left(x - x_{max,f}\right)}{L_f}\right)\right) \tag{4.11}$$

x is a place holder for the respective variables used in the functions f and g. The parameter A determines how strong the modifications are. Also, I define $L_f = x_{max,f} - x_{min,f}$. $x_{min,f}$ represents the x-position of the start node f. Similarly, $x_{max,f}$ represent the x-position

of the end nodes of f. I use $L_g = L_f$. I use the values from the parameter file shipped with *LAMMPS* [74, 137] for the MEAM [15] of $x_{min,f} = 2.055801767399996649$ Å and $x_{max,f} = 4.410000000000000142$ Å. This means that $L_f \approx 2.3541982326$ Å. For a future fitting, one could experiment with modifying the position of the spline start and end nodes. However, for simplicity, the start and end node positions are not moved in the exploratory study presented in this thesis.

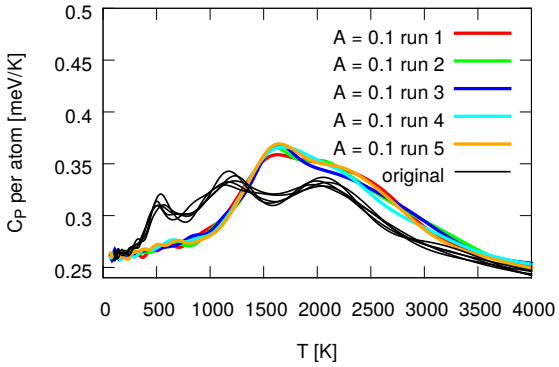
To explore how the Ti MEAM potential of Hennig *et al.* [15] reacts to small changes to the function f and g of Equation 2.9, I conduct five runs each with a variety of values of the aforementioned modification factor A at 12 GPa. We can see the results for some of the parameters for 24 atom in Figure 4.16. They are sorted with decreasing A, starting at $A = 0.01$. Each subfigure also includes a set of five heat capacity curves for $A = 0$, the original MEAM [15], for reference. The reference curves are black. For the curves for the modified potentials, I use colours to make it easier to distinguish between different runs.

I use 240 calls per iteration and set $n_{cull} = 8$ for all runs. For $A = 0.1$, $A = 0.05$ and the reference runs, 15360 walkers are employed. The runs for $A = -0.02$, $A = -0.04$ and $A = -0.08$ are conducted with 7680 walkers. It is very hard to identify the high temperature solid phase for 24 atoms with the usual XRD pattern analysis. Higher numbers of atoms yield in general a more reliable result. Hence, I refrain from making statements about the exact nature of the high temperature solid phase in the following. In Subfigure 4.16a, for $A = 0.1$, the ω to α transition peak has vanished while the other peaks have moved together so much that they can barely be distinguished. Instead of the ω phase, only the α phase can be identified at low temperatures in the investigated realm. The results for $A = 0.05$ (Subfigure 4.16b) show three peaks, apart from run 3 in blue which shows more of a plateau. Most prominently, the ω to α transition peak has shifted from ca 530 K found for the original to ca 140 K. The shift to higher temperatures for the other visible peaks was smaller but still considerable.

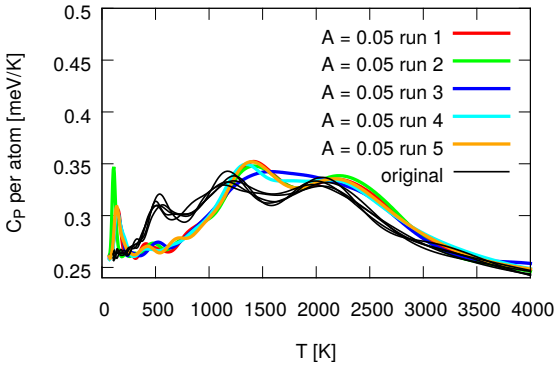
It appears that a small positive modification factor might lower the ω to α transition temperature. The question is now whether a small positive multiplication factor increases the ω to α transition temperature. If one increases the ω to α melting temperature enough, one might even loose the α phase fully so that the ω phase directly transforms to the β phase, assuming that this phase is indeed preserved. Note that I do not prove the existence of the β phase for 24 atoms. I will investigate this for 54 atoms at 12 GPa for a promising modification candidate ($A = -0.04$). One can indeed see the increase in temperature up to which the ω phase is table for small negative values of A as shown in the following. For $A = -0.02$ (Subfigure 4.16c) instead of two clear peaks below 1500 K for each run, we see a less clear picture. It is not possible to clearly determine the peak position anymore, as some

of the smaller fluctuations may just be artefacts. This can be interpreted as the two peaks moving together. The change in melting temperature is not significant: It changes slightly from 2033 ± 23 K from the reference calculations of the original potential to 1996 ± 17 K. The errors here are the SEM of the C_P curve peak positions presented in Subfigure 4.16c. For $A = -0.04$ (Subfigure 4.16d), we see what resembles even more a single peak than for $A = -0.02$ instead of the two peaks below 1500 K for the original potential. There are small shoulders on the low temperature side of the peak which may be indications of the existence of two phase transitions whose peaks are just very close together. The mean temperature of the temperatures of the peak maxima is 952 ± 17 K. The new melting temperature is 1964 ± 36 K. Again, the errors are the SEM of the heat capacity peak positions. For $A = -0.08$ (Subfigure 4.16e), we now have a single clear peak at 889 ± 3 K followed by the melting peak at 1862 ± 16 K. The errors are again the SEMs. The first peak is however quite high in comparison to the melting peak. This might be an indication that this peak represents actually two phase transitions which are very close to each other.
The potential with the modification factor $A = -0.08$ appears to be a very promising candidate for a modified potential that reproduces a triple point. However, nested sampling runs for 4 GPa and 8 GPa for $A = -0.08$ with 24 atoms per simulation box continue to show only one solid-solid peak with the ω phase as low temperature phase. This makes the existence of a triple point above 4 GPa very unlikely. These calculations are done with 240 calls and 7680 walkers. Calculations for $A = -0.04$, for 24 atoms at 2 GPa with 240 calls and 15360 walkers show two solid-solid peaks with the ω phase as the low temperature phase followed by the α phase at higher temperatures. This means that even for high number of atoms, a triple point may exist somewhere between 2 GPa and 12 GPa.

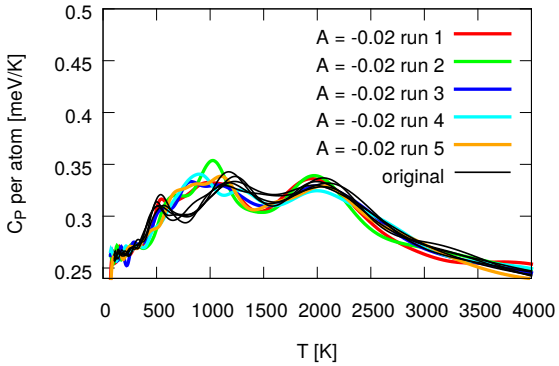
The peak below 1500 K for the MEAM with modification factor $A = -0.04$ in Subfigure 4.16d is not as clear as the one for $A = -0.08$ in Subfigure 4.16e. However, we have seen how increasing the number of atoms per simulation box leads to an increase in the ω to α transition. An increase in number of atoms is therefore hoped to make the peak below 1500 K clearer. Indeed, Figure 4.17 shows a much clearer peak for 54 atoms. Here, we see five heat capacity curves from nested sampling calculations for 24 atoms (blue lines) and 54 atoms (orange lines) for $A = -0.04$ at 12 GPa. For both 24 atoms and 54 atom, 240 calls, 7680 walkers and $n_{cull} = 8$ is used. The temperature of the lower peak increased considerably while the higher peak, which stands for the melting, is now found at a lower temperature. We see that the peaks have become more defined, however, this could also be a finite size effect due to the larger number of atoms used [47]. Nevertheless, the existence of the peak alone indicates again the possibility of a ω to β transition. The XRD pattern analysis shows an



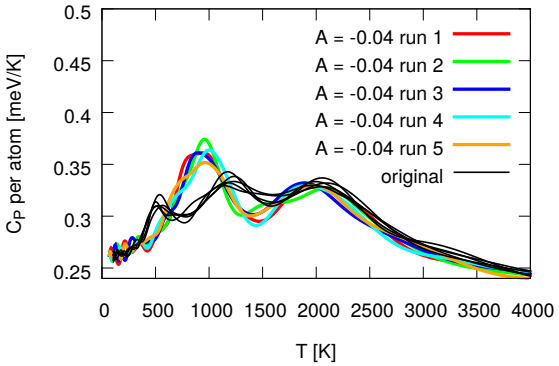
(a) $A = 0.1$



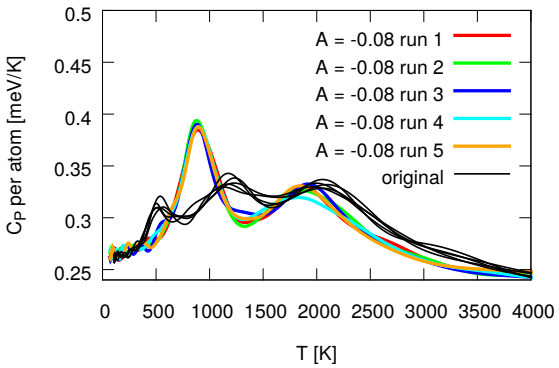
(b) $A = 0.05$



(c) $A = -0.02$



(d) $A = -0.04$



(e) $A = -0.08$

Fig. 4.16 C_P versus temperature curves for various values of the modification factor A in comparison to my nested sampling results for the original MEAM of Hennig *et al.* [15]. The results for the original potential are given in black while the results for the different runs with the modified potentials are given in colours to ease identification of shoulders and peak positions. The modification factors A are given in the labels below the respective graphs. A negative modification factor A pushes the peaks below 1500 K together while a positive pushes them apart from each other.

indication for a potential β phase at high temperatures and that the low temperature phase is the ω phase.

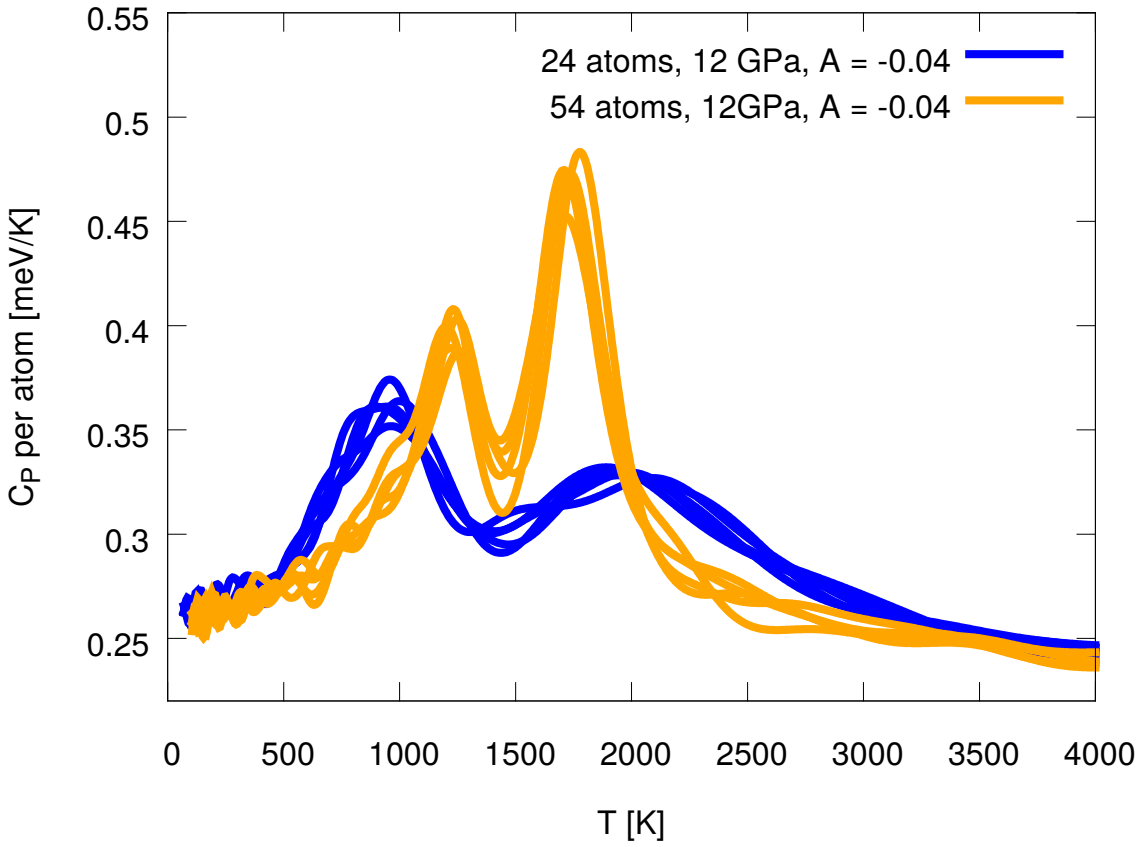


Fig. 4.17 Heat capacity for the modified MEAM with $A = -0.04$ at 12 GPa for different numbers of atoms. For each number of atoms, five independent nested sampling runs are conducted. The 24 atoms are denoted by blue lines while the 54 atoms are given by orange lines and have clearer defined peaks. The lower peak position was raised in temperature while the higher peak was lowered. For 54 atoms, the system is in the ω phase for low temperatures. For temperatures above the lowest peak position, a phase which shows an indication of the β XRD pattern is dominant for 54 atoms until the system melts. For 24 atoms we also find the ω phase at low temperatures. The peak at higher temperature denotes the melting. For both number of atoms, I use 240 calls, 7680 walkers as well as as $n_{cull} = 8$.

To summarise, the 54 atoms calculations with $A = -0.04$ showed a relatively clear peak which indicates a potential ω to β transition at 12 GPa. Additional calculations (with 15360 walkers) show the ordering ω phase, α phase and an indication of a potential β phase with increasing temperature for 6 GPa and 54 atoms. Hence, the next step is to conduct thermodynamic integration to calculate the phase diagram at higher number of atoms to minimise the finite size effects. I cover these in the Subsections 4.8.1 and 4.8.2.

4.8.1 Phase diagram with the modified Ti MEAM in comparison to one done with the original Ti MEAM

As explained in Section 1.2, the finite size effects for the solid-solid transition phase diagram need to be minimised by using larger numbers of atoms for the promising potential candidate with $A = -0.04$ to check to which extent its phase diagram prediction has been improved. Before we compare directly to an experimental phase diagram in Subsection 4.8.2, let us first consider how the phase diagram predictions of the modified MEAM compare to my predictions using the original MEAM of Hennig *et al.* [15]. I show phase diagram predictions of both in Figure 4.18. For both potentials I use the parameters as described at the end of Subsection 4.6.1 for the thermodynamic integration. The stability of the ω phase against the β phase is calculated via MD as explained in Subsection 4.6.3 by using 10^6 equilibration steps similar to the calculations done by Hennig *et al.* [15] with the β phase. However, I start with the ω phase. I distinguish between the results by using grey labels for the phases predicted by the original potential [15] and grey dashed lines for the phase transition lines predicted by the original potential [15] via my thermodynamic integration calculations. All other labels and data points (whether connected to lines or not) are from the calculations using the modified MEAM. For the ω to α transition data points, I use blue stars. The α to β transition is given by orange upwards pointing triangles. The α to ω transition which only appeared for the modified MEAM is denoted by a magenta diamond. All these transitions are calculated using thermodynamic integration (TI). The ω to β transition is not possible to calculate with the thermodynamic integration method used in this thesis as one of the phases always transforms under MD simulations at the investigated temperatures. Hence, I resort to the MD calculations mentioned at the beginning of this subsection. The resulting ω to β transition data points are signified by downward pointing cyan triangles. I do not calculate any melting curve. The modification of the MEAM leeds to a decrease in the β transition temperature of values between ca 270 K and ca 310 K for pressures from 0 GPa to 8 GPa. Additionally, it leeds to the occurrence of the ω to β transition at 10 GPa and 12 GPa. The transition from the α or ω phase to the β phase are roughly at a constant temperature ranging between ca 1370 K and 1410 K. Additionally, the modified potential shows an increased ω to α transition temperature. The increase is 260.7 ± 1.5 K from a transition temperature of 227.1 ± 0.7 K to one of 487.8 ± 1.4 K at 0.0 GPa. This difference increases with pressure. At 8.0 GPa the ω to α transition temperature of 971 ± 10 K for the modified MEAM is 424 ± 11 K higher than the transition temperature of the original MEAM of 547.9 ± 2.8 K. If we take the transition temperature values at 0 GPa and 8 GPA for the modified MEAM, we get an estimate of the slope of ca 60 K/GPa. If we do the same for the original MEAM for 0 GPA and 12 GPa, we get a slope of ca 40 K/GPa (rounded to the nearest 10 K/GPa). For

12 GPa the modified MEAM does not predict an ω to α transition. At 10 GPa the ω to α phase transition temperature of the modified MEAM of 1183 ± 14 K is higher than expected if we consider the slope of the lower pressures. At 1210 ± 12 K thermodynamic integration predicts a transformation back to the ω phase. The difference, and hence the temperature range where an α would exist, is however only 27 ± 25 K. The error is so large because the α phase is only represented by one value which is used for the calculation of both transition temperatures. Given that this difference is smaller than twice its standard error and the α phase only appeared for one transition temperature at this pressure, it is possible that the α phase does not even occur at this pressure and that it only occurs in the calculations due to statistical fluctuations. This means that the triple point could potentially lie between 8 and 10 GPa. This is a very big difference to my thermodynamic integration predictions of the original potential [15] which do not show any signs of a triple point up to including 12 GPa.

4.8.2 Comparison of modifications to experiments

In the following I consider how the modified Ti MEAM potential predictions compare to the experimental phase diagram. In Figure 4.19, we see the experimental solid-solid transitions according to Zhang *et al.* [51] denoted by grey dashed lines with small dashes. The long grey dashes are used to depict the melting transition results of Errandonea *et al.* [81] (fit data from Ref. [82]). Grey labels signify the respective experimental phases. My thermodynamic integration predictions for the $A = -0.04$ modification to the MEAM of Hennig *et al.* [15] are shown in comparison. Blue stars are used to depict the ω ot α transition. The α to ω transition is given by a magenta diamond. The α to β phase transition is denoted by orange upwards pointing triangles. All these aforementioned transitions predicted by the modified MEAM are calculated with thermodynamic integration (TI). The ω to β transition is calculated with molecular dynamic stability tests (MD) as explained before and signified by downwards pointing cyan triangles. The modified MEAM phase diagram has been described already in Section 4.8.1. At 10 GPa the α phase may not appear at all and hence a triple point might exist between 8 GPa and 10 GPa. Under this premise, the modified MEAM does predict a phase diagram which is in much better accordance with the experimental phase diagram than the original Ti MEAM [15] which did not predict a triple point at all. Zhang *et al.* predict 7.5 GPa and 913 K for the triple point coordinates [51]. For the pressure value at least, this is relatively close to the pressure realm for which I expect a triple point. The β phase still becomes stable for the modified MEAM at a temperature which is considerably higher than the experiment. However, this temperature has already been lowered considerably by the modification as seen in Section 4.8.1. The ω to α phase transition is however not as well reproduced. The modification leads to an increase in ω to α transition temperature

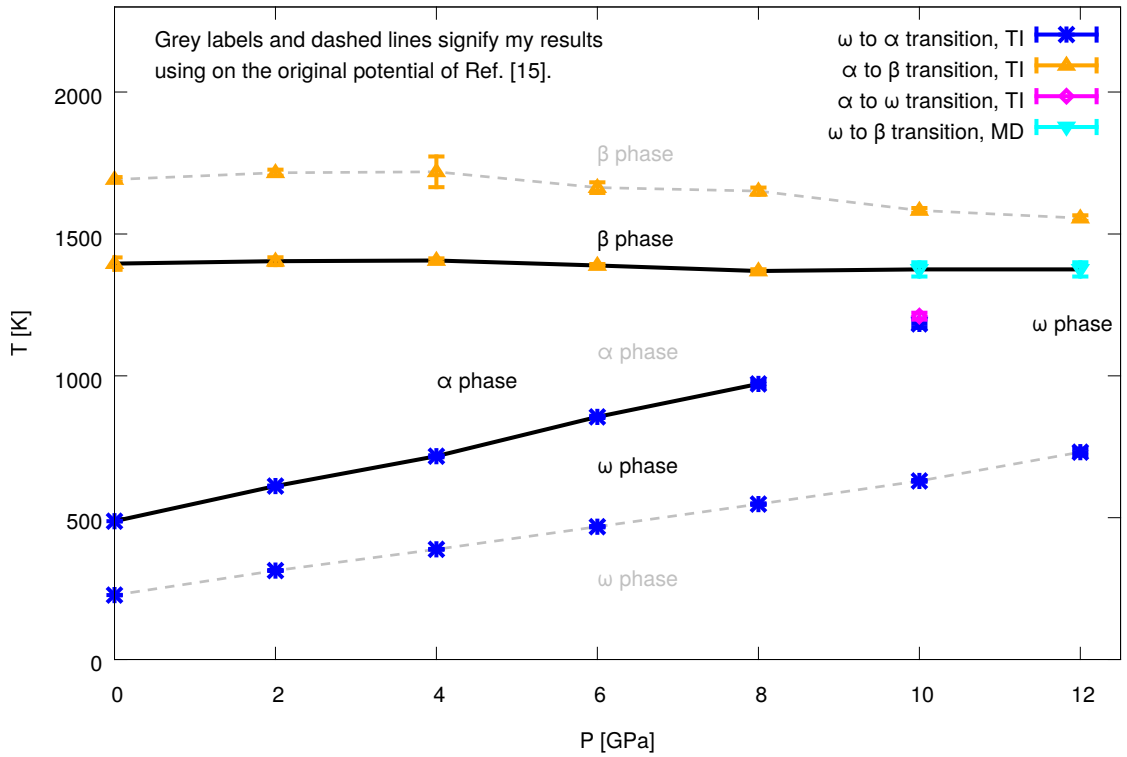


Fig. 4.18 Comparison of original and modified MEAM solid phase diagram predictions with thermodynamic integration and MD simulations. The dashed grey lines and the grey labels represent the predictions using the original MEAM of Hennig *et al.* [15]. All other symbols, labels and lines represent the results for the modified MEAM. The blue stars represent the ω to α transition. The α to ω transition is represented by a magenta diamond. Orange upwards pointing triangles signify the α to β transitions. Finally, the ω to β transition is given by cyan downwards pointing triangles. All transitions apart from the ω to β transition were calculated using thermodynamic integration (TI). The ω to β transition was calculated via "MD stability tests". The modified MEAM predict a considerably different phase diagram. The α to β transition temperature is lowered while the ω to α transition temperature is increased. Additionally, the phase diagram shows an indication of the previously missing triple point between 8 GPa and 10 GPa as explained in the text.

overall. If one takes the aforementioned triple point of Zhang *et al.* and uses that their slope prediction is 345 K/GPa [51], one arrives at at a transition pressure of ca 4.9 GPa for a pressure-driven transition from the α to ω phase at 0 K. At 0 K and 0 GPa the phase should hence be the α phase not the ω phase. In this realm the modification actually leads to a worse prediction as the ω phase is now even stable up to higher temperatures. However, it is to note that the original Ti MEAM did predict the ω phase as 0 K phase as well for this pressure range.

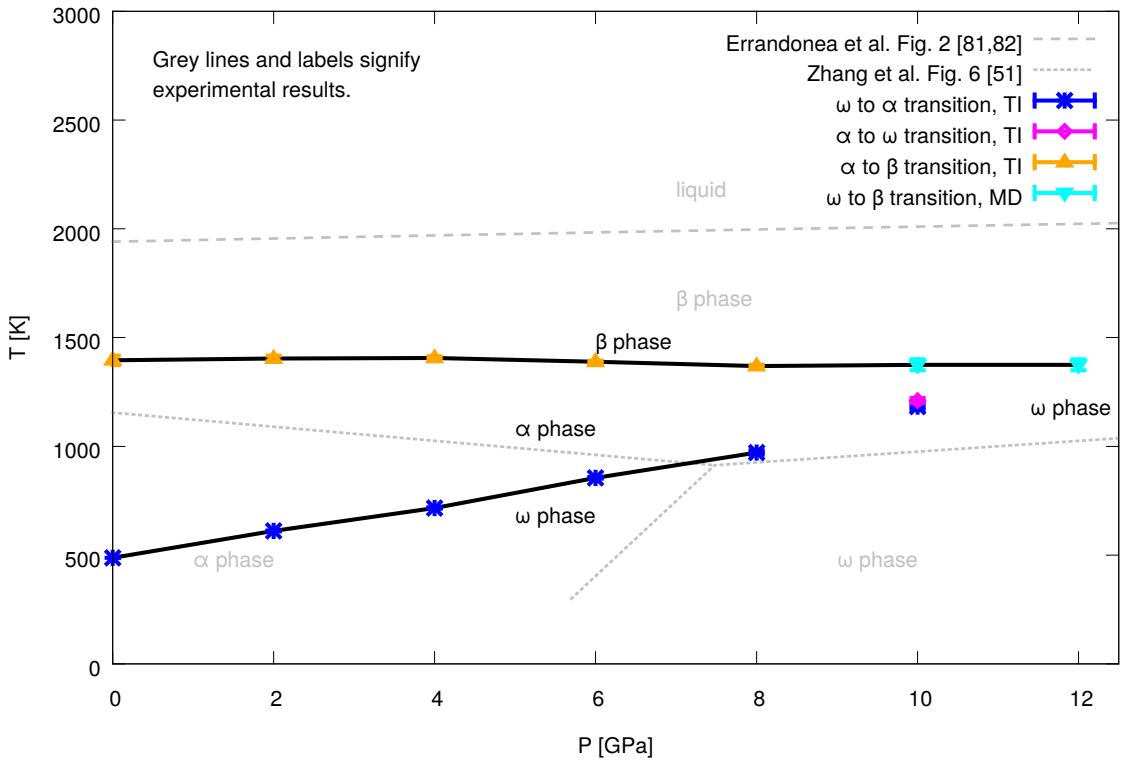


Fig. 4.19 The solid-solid transition phase diagram of the modified MEAM in comparison to experimental data. We see the phase transitions shown in comparison to experimental results from Zhang *et al.* [51] and Errandonea *et al.* [81, 82]. Again, the ω to α transition is given with blue stars. The upwards pointing orange triangles denote the α to β transition, while the ω to β transition is denoted by downwards pointing cyan triangles. The magenta diamond symbol stands for an α to ω transition at 10 GPa. The occurrence of the α phase before this transition was questioned in Section 4.8.1. Hence, a triple point might be found between 8 GPa and 10 GPa. All of my predictions apart from those for the ω to β transition are conducted with thermodynamic integration (TI). The ω to β is calculated by MD stability tests (MD). Grey labels and grey dashed lines represent the experimental data. Here, short dashes denote the solid-solid transition results of Zhang *et al.* [51]. The long dashes are used to depict the melting curve results of Errandonea *et al.* [81, 82].

4.9 Comparison of energies predicted by the modified Ti MEAM to DFT results

After seeing how the phase diagram prediction have improved using the modified version of the MEAM of Hennig *et al.* [15], we now turn to the DFT energies presented earlier in Section 4.7. One might assume that the modified potential will represent the α and β phase structures better than the original as an indication of a triple point is found. Figure 4.20 shows the difference between cohesive energy predictions of the DFT calculations and the modified MEAM (blue "+" symbols) as well as the original MEAM (red "×" symbols) of Hennig *et al.* [15] for the structures sampled at different temperatures and pressures as explained in Section 4.7. As before in Section 4.7, I add ca 1.7354 eV to the DFT energies so that the average energy predicted for the ω phase structures by the original MEAM at 2 GPa is the same as the average energy predicted by the DFT calculations. The changes to the energies due to the modification are relatively small. Overall, we can still say that in relation to the other solid phases, the modified MEAM still overestimates the energy of the ω phase. To see the changes between the MEAM potentials, we now consider Figure 4.21. Here, the differences for the cohesive MEAM energies per atom

$$\Delta E_{MEAM} = E_{cohesive,MEAM,A=-0.04} - E_{cohesive,MEAM,original} \tag{4.12}$$

are shown for the different structures. Here, $E_{cohesive,MEAM,A=-0.04}$ is the cohesive energy predicted by the modified MEAM with $A = -0.04$ and $E_{cohesive,MEAM,original}$ is the cohesive energy predicted when using the original potential of Hennig *et al.*[15]. For the modified MEAM the ω, β and the liquid phase cohesive energies are on average decreased while the α phase energy is on average increased. There is however one outlier at 10 GPa. To investigate the structure, I calculated the XRD pattern as before via *QUIP* [114], called via *misc_calc_lib.py* [45] and supported these calculations via a visualisation with *OVITO* [138] using its "bond-angle analysis" [139, 140] tool. The ω phase shows a characteristic colouring pattern. For the reference structure, I use again the c/a ratio of Hennig *et al.* [15] for their MEAM [15]. Both methods suggest that this is an ω phase rather than an α phase which explains the behaviour as the energy of the ω phase was decreased. As mentioned before in Section 4.7, one needs to consider the Gibbs free energy to make a clear conclusion regarding a phase's stability. However, this shift in energy may explain at least in part why the area in the phase diagram of the α phase shrinks, as this makes it energetically less favourable in relation to the other phases. It is however noteworthy that even though the phase diagram has

improved considerably with the modified potential, the difference from the DFT energies has only been slightly improved when it comes to difference between ω and α phase.

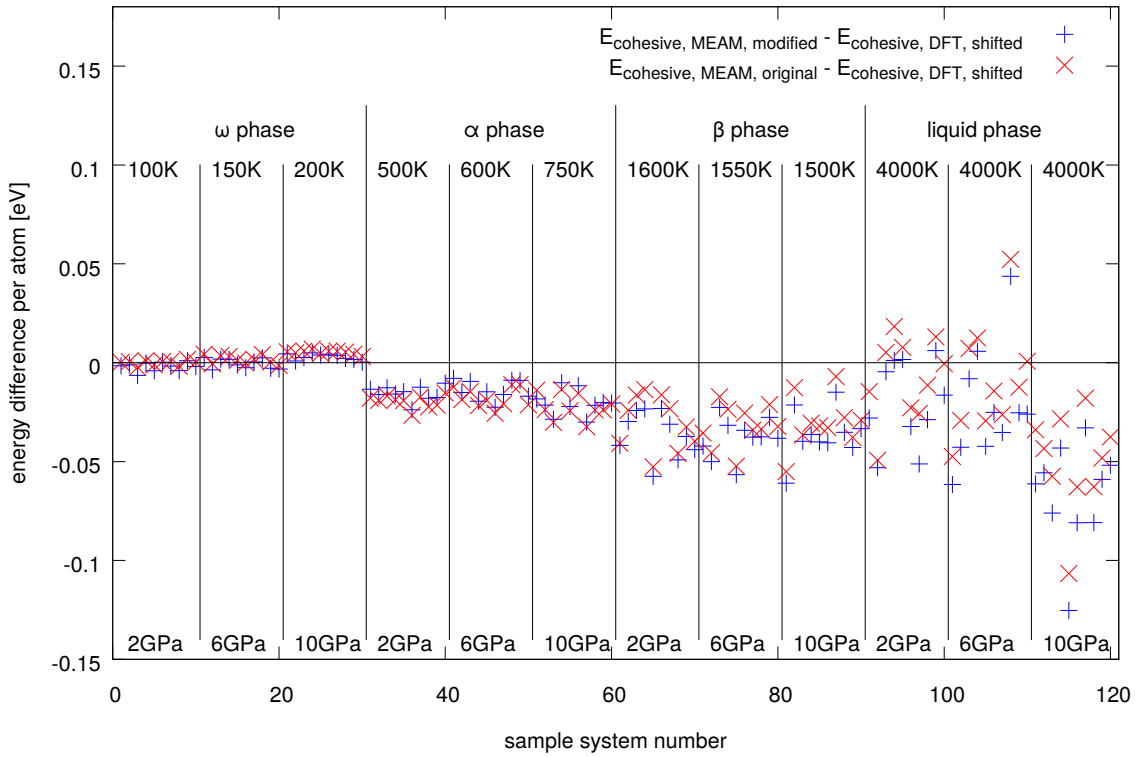


Fig. 4.20 The difference between the MEAM cohesive energies and DFT cohesive energies per atom for a selection of structures for the original MEAM of Hennig *et al.*[15] and the modified MEAM with $A = -0.04$. The energy difference of the original MEAM and the DFT energies are the same as in Figure 4.15. The structures are as explained in Section 4.7. Blue "+" symbols are the differences to the modified MEAM energies. Red "×" symbols are the difference for the original potential. The changes are relatively small. Overall, relative to the other solid phases, the ω phase energy is still overestimated by the modified potential. As mentioned before, I add ca 1.7354 eV to the DFT energies to have the same average energy values for the ω phase at 2 GPa as the original MEAM. This is simply to make comparison easier. The difference between original and modified MEAM energies are shown in Figure 4.21.

4.10 Functions of the modified MEAM and basic parameters of the stable structures

As the final step of my overall procedure (see Section 1.2), I investigate how strongly the new potential differs from the original potential of Hennig *et al.* [15] and how it performs for

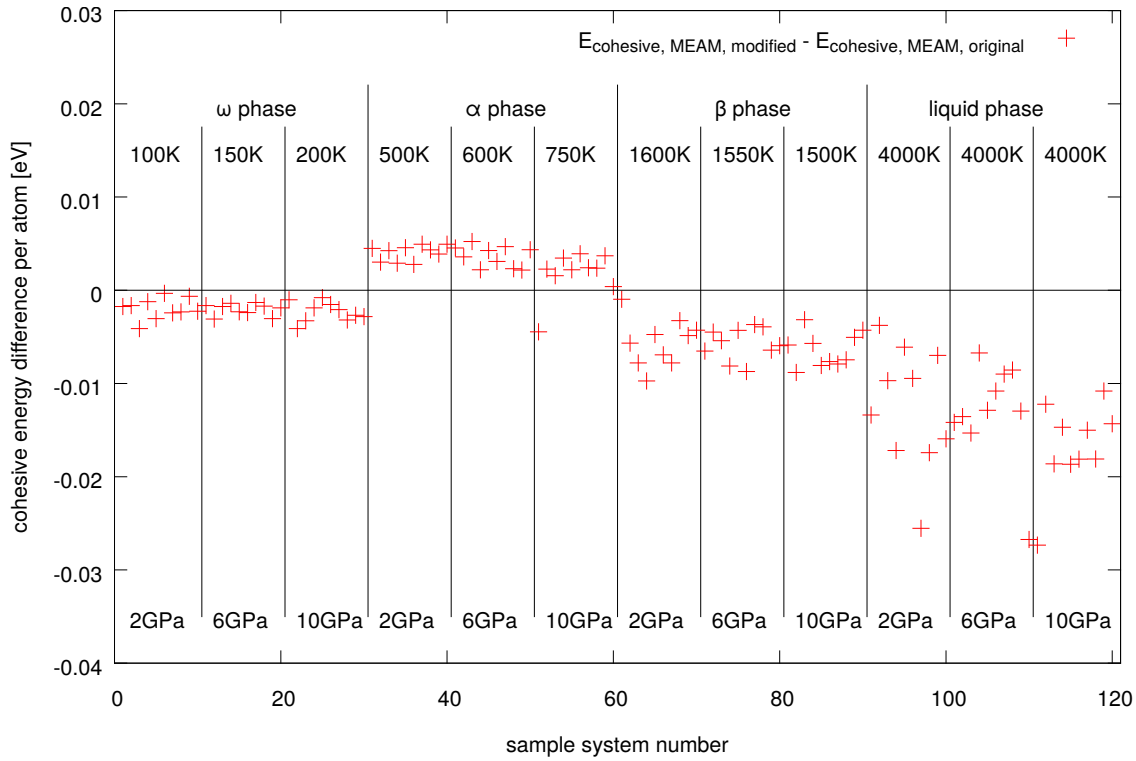
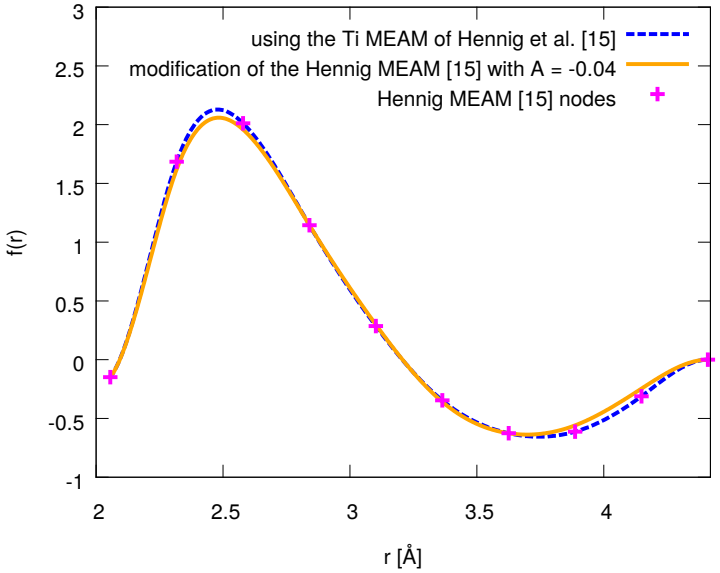


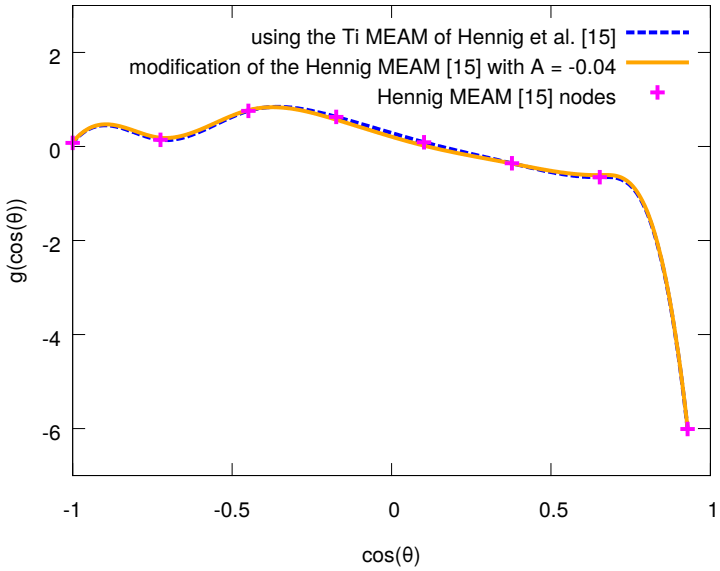
Fig. 4.21 The difference between modified and original MEAM of Hennig *et al.* [15] cohesive energies per atom for the structures specified in the text. For the ω phase, for the β phase and for the liquid, the energy is lowered by the modification of the MEAM. The α phase energies are however increased apart from one outlier at 10 GPa which XRD pattern analysis and *OVITO* [138] "bond-angle analysis" [139, 140] suggest is likely an ω structure. In summary, the α phase energy is increased compared to the energies of the other shown phases.

basic parameters of the phases expected for the pressures which I investigated [44, 50, 51]. Let us now consider how the functions which make up the MEAM of Hennig *et al.* [15] are changed. In Figure 4.22, one can see how the new functions for $f(r)$ (Subfigure 4.22a) and $g(\cos(\theta))$ (Subfigure 4.22b) of the new potential with $A = -0.04$ compares to the original functions of the potential of Hennig *et al.* [15]. The original potential function values are denoted by the dashed blue curve. The nodes of this function are given by the magenta "+" symbols. The modified potential function values are denoted by the solid orange curve. The changes to the functions are in general very small. For $f(r)$ the peak at ca 2.5 Å is slightly lowered with little change to its position. The minimum value between 3.5 Å and 4 Å is slightly increased while its radial position is decreased. This means that the strength of the characteristic features of the function is lowered by a small amount. For $g(\cos(\theta))$, the changes are not as characteristic as for $f(r)$. Most notably, the position of the maximum was shifted by a small amount to lower values of $\cos(\theta)$ and barely decreased in the function value.

In Tables 4.1, 4.2 and 4.3, we see the elastic constants, lattice constants, ratio of lattice constants (where applicable) and volume per atom for the phases which occur in the Ti phase diagram in pressure realm investigated in this thesis [50, 44] predicted by my calculations for the original MEAM [15] and for my modification of it with $A = -0.04$ as well as some experimental values. The calculation of the elastic constants is conducted in the same manner as in Section 3.14.1. The starting structures of Ti which are then relaxed are based on the volumes per atom and c/a ratios of relaxed structures at zero pressure presented for the original MEAM [15] in Ref. [15]. The different quantities predicted by the modified MEAM change for the most part only slightly compared to the ones predicted by myself for the original MEAM [15]. Overall, the elastic constants change more then the lattice constants and lattice constant ratios as well as the volume per atom. For the elastic constant I even find deviations of over 10 % between original and modified potential. These occur all for the α phase. The modification leads to a reduction of c_{11} by ca 13 %, an increase of c_{12} by ca 15 % and a reduction of c_{44} by 14 %. The quantities I present for the original MEAM in Tables 4.1, 4.2 and 4.3, agree, if they were given, with the once presented by Hennig *et al.* in Ref. [15] in Table II and Table III of their paper in all cases, except for the values for a of the α phase (found in their text) and c_{11} and c_{12} of the ω phase. The a value they report for the α phase is only 0.001 Å larger than the value I report. In light of rounding and potentially different tolerances for finishing the relaxation this is negligible. For the elastic constants of the ω phase however, Hennig *et al.* predict $c_{11} = 191$ GPa and $c_{12} = 78$ GPa [15]. This is very different from my values of $c_{11} = 177$ GPa and $c_{12} = 93$ GPa. Interestingly, the sum of these parameters (c_{11} and c_{12}) are nearly the same for Hennig *et al.* [15] and me. Potentially,



(a) Comparison of $f(r)$ plots



(b) Comparison of $g(\cos(\theta))$ plots

Fig. 4.22 Comparison of the modified functions with modification parameter $A = -0.04$ to the functions of the original potential of Hennig *et al.* [15]. Dashed blue lines are the results with the original functions. The magenta "+" symbols depict the spline nodes of the original potential. The orange solid line represent the function values of the modified potential. Subfigure 4.22a shows the $f(r)$ functions and nodes. Subfigure 4.22b plots the $g(\cos(\theta))$ functions and nodes.

quantity	original MEAM [15]	modified MEAM, $A = -0.04$	experimental
c_{11} [GPa]	177	174	
c_{12} [GPa]	93	93	
c_{13} [GPa]	64	68	
c_{33} [GPa]	233	230	
c_{44} [GPa]	48	48	
a [Å]	4.606	4.627	4.625
c [Å]	2.815	2.789	2.81
c/a	0.611	0.603	0.608
V_{at} [Å^3]	17.24	17.24	17.35

Table 4.1 Ti elastic constants (c_{11}, c_{12}, c_{13}, c_{33} and c_{44}), lattice constants (a and c), lattice constant ratio (c/a) as well as the volume per atom (V_{at}) of the ω lattice which I calculate for the original Ti MEAM [15] and modified MEAM with $A = -0.04$. The experimental values of the lattice constants (at room temperature) are taken from Ref. [50] and are used to calculate the c/a value and the volume per atom.

Hennig *et al.* use a different unit cell for their ω phase or a different orientation of their ω unit cell than I. All experimental values given here except for the elastic constant for the α phase at 4 K [141] in Table 4.2 represent experimental values for temperatures far away from 0 K [50, 79, 142] (see captions of Tables 4.1, 4.2 and 4.3 for details) which makes it harder to compare the results. In light of this fact, the values predicted by the original [15] and the modified MEAM are in relative good accordance with experiments for these intermediate to high temperature values [50, 79, 142]. For the elastic constants at 4 K of the α phase [141] (see Table 4.2) the agreement is still acceptable, the modified potential worsens the predictions compared to the original potential [15] in most cases.

quantity	original MEAM [15]	modified MEAM, $A = -0.04$	experimental
c_{11} [GPa]	174	151	176
c_{12} [GPa]	95	109	87
c_{13} [GPa]	72	71	68
c_{33} [GPa]	188	181	191
c_{44} [GPa]	58	50	51
a [Å]	2.930	2.928	2.951
c [Å]	4.678	4.693	4.686
c/a	1.596	1.603	1.588
V_{at} [Å^3]	17.40	17.43	17.67

Table 4.2 Ti elastic constants (c_{11}, c_{12}, c_{13}, c_{33} and c_{44}), lattice constants (a and c), lattice constant ratio (c/a) as well as the volume per atom (V_{at}) of the α lattice which I calculate for the original Ti MEAM [15] and modified MEAM with modification factor $A = -0.04$. I also added experimental results for comparison. For these, the results for the elastic constants (at 4 K) are from Ref. [141]. The lattice constants – which are also used to calculate the volume per atom – and the c/a ratio are from experiments at ca 301 K [79].

quantity	original MEAM [15]	modified MEAM, $A = -0.04$	experimental
c_{11} [GPa]	95	100	134
c_{12} [GPa]	111	108	110
c_{44} [GPa]	53	49	36
a [Å]	3.272	3.273	3.307
V_{at} [Å^3]	17.51	17.53	18.07

Table 4.3 Ti elastic constants (c_{11}, c_{12} and c_{44}), lattice constant (a) as well as the volume per atom (V_{at}) of the two atom cubic bcc (β phase) cell which I calculate for the original Ti MEAM [15] and the modified MEAM with modification parameter $A = -0.04$. Experimental results are given for comparison. For these, experiments on phonons at ca 1238 K are the basis of the reported elastic constants [142]. The experimental lattice constant value (at ca 1173 K) is from Ref. [50]. I calculate the volume per atom using the precise value 3.3065 Å given in Ref. [50].

Chapter 5

Discussion

5.1 CuAu calculations – Discussion

Baldock *et al.* [28] show results using the original CuAu EAM [39–41] for 64 atoms per simulation box at 0.1 GPa and predict a eutectic curve with some deviations from experimental results [42] to which they compare. However, I have shown in Section 3.6 that 64 atoms are not enough to achieve fully converged results for the melting temperature for the used CuAu EAM [39–41]. Indeed, an extrapolation with $1/N^2$ is necessary to gain a reliable result. The 64 atom results significantly overestimate the melting temperature. The extrapolation shows that the original CuAu EAM [39–41], even though it predict a eutectic curve, underestimates experimental melting predictions from Ref. [88] and estimates based on the results of Bennett [87] quoted in Okamoto *et al.* [42] as explained in Section 3.12 considerably as shown in Figures 3.6, 3.7 and 3.10.

The main problem of the fitting for the CuAu melting curve presented in this thesis is that it assumes that only the melting transition temperature changes during fitting and that the solid below the liquid does not change to a different phase. A way to consider phase changes has not yet been included in the calculation of the objective function. However, I have made sure that the solid phase below the melting temperature did indeed stay the same throughout the fitting by checking the XRD pattern of the solid phase just below the melting temperature at one or more suitable temperatures. I cover a potential upgrade of the objective function calculation in Section 6.2.

A important limitation of the fitting is the statistical error. It is likely that a too high statistical error leads to the slow convergence seen in the 24 atoms CuAu 50:50 temperature fitting shown in Section 3.13.1. Convergence issues similar to this were noted in Ref. [63]. A 36 atoms fit with lower errors does not show this behaviour as seen in Section 3.13.2. However, using more runs to calculate the mean and its error would also be a possible alternative to

lower the statistical error of the temperature prediction. Depending on the size of the steps usually taken during the Nelder-Mead algorithm in relation to the statistical error there may be a systematic bias in the fitting. Let us assume one fits only one temperature. If we start at a lower temperature than the temperature we aim to fit, it is in principle more likely that the first vertex which is within the 2 statistical SE realm is in the lower half than in the upper half of this realm. The statement is reversed if we start at higher temperatures. If the statistical SE is relatively small in comparison to the step size in temperature, this should not have too much of an effect. However, the effect will become larger if the ratio of average statistical error to average absolute difference between the objective function of the replaced vertex and the objective function of the replacing vertex increases. In principle, this behaviour can be generalised for fitting to more than one transition. However, the fact that not all transition temperatures to which one fits are necessarily approached equally fast will diminish this effect as the temperature prediction of some transitions may already have overshot the value which one aims to fit during the fitting. Additionally, given the statistical and systematic errors in the extrapolation of the melting temperature as well as the systematic error due to modelling with a pressure which is not the normal pressure, I assume that the effect is small in comparison to these errors. I consider the errors of the melting temperature in more detail at a later point in this section. In addition, the algorithm requires that statistical errors in the melting temperature are kept relatively low [63] – as I have elaborated upon already in this section with regards to the 24 atoms CuAu 50:50 fitting – which should also lower the effect of the problem.

Additionally, the error estimation for the average peak positions needs to be considered. During the fitting, only 5 runs are used. This leads to fluctuations in the standard deviation estimates. This most likely is responsible for most of the fluctuation in the error estimate during fitting. This makes stopping the algorithm via the 2 SE ellipsoid and the Mahalanobis distance [99] as explained in Section 3.8 less reliable as I rely on the error estimates for this. To gain a more dependable statistical error estimate, I use the average of the relative statistical error of the melting temperature predicted by several modified potentials (separately for each CuAu composition). However, it is not clear whether the relative error is indeed a constant when the potential parameters are slightly changed as during the fitting. To compare the results to a different assumption, I calculate the average of the SE of the melting temperatures of the vertices used in the simultaneous fitting of the Cu, CuAu 50:50 and Au melting temperatures separately for each different percentage of Cu. For vertex 38, the eventually accepted vertex, the averaged relative error leads to an error estimate in the melting temperature for 36 atoms of ca 10.9 K for Cu, of ca 10.7 K for CuAu 50:50 and ca 9.2 K for Au. The averages of the SEM of all vertices are ca 10.0 K for Cu, ca 10.4 K for

CuAu 50:50 and ca 8.6 K for Au. We see that the difference between the expected standard errors of the melting temperatures for the different percentages of Cu is very small. For the fitting run conducted for the simultaneous fitting of melting temperatures for this thesis, the different approach to error averaging does not change the results at all and will most likely only have very minor effects on reruns. This increases the confidence in my results.
The constant pressure nested sampling algorithm used in this thesis does not specifically consider the liquid-solid coexistence regions found in experiments for CuAu [42] mentioned in Section 2.10. Hence, one would not expect it to be seen for small number of atoms. The system would need to be large enough to allow for some form of solid-liquid interfaces. Judging from experimental data of Bennett [87] quoted in Okamoto *et al.* [42] presented in Figure 2.6, the region is small for all compositions of CuAu. It does not exist at the eutectic point and for pure Cu and pure Au. That means that, if the potential (modified or original [39–41]) models CuAu well enough, the region should be very small for the 50:50 ratio as this one is close to the eutectic point atomic percentage of Cu of 43.7% as determined by Bennett [87] quoted in Okamoto *et al.* [42]. For 50.8 atomic % Cu, Bennett's results [87] quoted in Okamoto *et al.* [42] show a solidus to liquidus difference of 3 K. The maximum solidus to liquidus difference of Bennett's results [87] quoted in Okamoto *et al.* [42] is found at an atomistic percentage of Cu of 90.8 % with 23 K. In light of the big underestimation of the experimental melting temperatures, the issue of the solid-liquid coexistence is of subordinate importance.
The usage of the extrapolation constant to take the finite size effects into account introduces two kinds of systematic errors. First, it is not clear whether the system actually follows the extrapolation scheme. Secondly, once a fitting scheme is chosen, the extrapolated melting temperature for $N \to \infty$ has an error. Additionally, the reference temperature calculated with nested sampling for the reference number of atoms has an error as well. These errors lead to the error of the extrapolation constant. The $1/N$ fitting shows a much larger error for both parameters than the $1/N^2$ fitting as I demonstrate in Section 3.6. I therefore rely on the $1/N^2$ fitting based extrapolation constant during my final fitting to three transitions in Section 3.14. For the fitting of solely the Au melting temperature which relies on the $1/N$ extrapolation in Section 3.11, I have demonstrated that the error due to our choice of experimental values is negligible in light of error of the extrapolation constant alone. Let us now consider how the case is for the $1/N^2$ fitting to the Au, CuAu 50:50 and Cu melting temperatures from Section 3.14. For this, using the averaged relative error in all error estimates for 36 atoms, I retroactively expect relative error of the temperature predicted with the extrapolation constant of ca 0.010 for the Au transition, 0.016 for the CuAu 50:50 transition and 0.014 for the Cu transition. The extrapolation constants are the same as the ones in Table 3.1b. I used

rounded values of these in the fitting of Section 3.14. For hypothetical melting temperatures extrapolated with the aforementioned constants which are equal to the experimentally based temperature values used during fitting based on the results from Ref. [88], Ref. [97] and Bennett [87] quoted in Okamoto *et al.* [42] as explained in Section 3.12, this leads to errors of ca 13 K, 19 K and 19 K, respectively. The nested sampling calculations are conducted at 0.1 GPa. Hence, technically, the fitting should have been conducted with the melting curve at this pressure. For gold, the Equation 3.11 fitted in Ref. [96] yields a value of ca 1344 K. This is ca 7 K higher than the value used in the fitting of 1337 K based on the Au melting temperature of Ref. [88]. The fit of Ref. [143] predicts a Cu melting temperature value of 1362 K which is 6 K higher than the temperature used in the fitting of 1356 K based on the value of Ref. [97]. Let us now assume that the change in expected melting temperature for CuAu 50:50 is in the same realm. Also, I estimate the additional error due to the CuAu 50:50 melting temperature estimate to be ca 1.5 K as explained in Section 3.12. Overall, I concluded that, given the aforementioned errors of the $1/N^2$ extrapolation and the fact that it is not sure that this is indeed the best extrapolation formula possible, the effect of the pressure of 0.1 GPa being used for the nested sampling calculations and the approximations and choice of values in Section 3.12 is likely negligible.

The single fitting of the CuAu 50:50 melting temperature has not been successful in also raising the melting temperatures of pure Cu and pure Au sufficiently as shown in Figures 3.6 and 3.7. To ensure this, the pure Cu and pure Au melting temperatures are included in the fitting as presented in Section 3.14. I present the lattice constants, the volume per atom and the elastic constants for the Cu fcc lattice and the Au fcc lattice for a temperature of 0 K and a pressure of 0 GPa calculated using the original EAM [39–41] as well as using the potential of vertex 38 from the three transition fitting in Subsection 3.14.1. Apart from the value for c_{11}, the new potential changes for Cu only slightly. Apart from the c_{11} value, the change in values is overall stronger for Au. Overall the new values are broadly in agreement with the experimental data. Given that these values are not part of the fitting procedure, even the changes for Au are still acceptable.

To summarise, I have shown that a Nelder-Mead algorithm is useful for optimising the phase transition predictions of an existing EAM potential [39–41] by fitting directly to melting temperature values based on experiments (Ref. [88], Ref. [97] and Bennett [87] quoted in Okamoto *et al.*) as explained in Section 3.12. This is especially interesting considering that the fitting scheme is very simple: I just replace $\rho_{a,CuAu}(r)$ with $\rho_{a,AuCu}(r)$ and multiply the functions which make up the potential (see Section 2.2.2) with the respective parameters as explained in Section 3.10. For the specific run for the fitting of three transition temperatures shown in this thesis, the algorithm takes only 22 iterations. As the objective function has a

statistical error, the number of iterations necessary to achieve convergence has a statistical distribution as well. As the algorithm itself has more parameters which are fitted than melting temperatures which are optimised, it is likely that there exists a set of parameters which predict melting temperatures similarly close to the melting temperatures aimed for during the fitting, but is very different from the vertices found in this work. This means that it is possible that there is potential which changes the original parameters less while reproducing the melting curve equally well as the results presented in this work. Such a set of parameters may even lead to smaller changes in lattice constants, volumes per atom and elastic constants. However, the cost of rerunning the Nelder-Mead algorithm several times are prohibitively high. Hence, the range of possible potentials has not been investigated in this work.

5.2 Ti calculations – Discussion

We have seen in Sections 4.2, 4.3 and 4.4 the difficulty in converging simulations for the Ti MEAM potential [15] for a given number of atoms. A high number of walkers is necessary. The basin for the ω structure is likely very small during the iterations where the α phase is dominant so that it is very likely do be lost or not even found if the number of walkers are small [32]. The fact that the ω phase appears only at very low temperatures increases the likelihood of the basin being lost as as the system has to undergo more culls to reach the iterations when the ω basin increases in size relative to the α basin. This, together with the computational expensiveness of a MEAM compared to an EAM leads very long calculation times limiting convergence studies with regard to the number of atoms. I could not find any indication of a triple point with converged runs up to including runs with 54 atoms. For 108 atoms runs presented in this thesis, which still show convergence issues, there is not any indication of an ω to β transition. According to a batch of five 216 atom runs with 3840 walkers an ω to β transition and an α to β are both possible. However, I have not been able to converge simulations for 216 atoms due to the high computational cost. Hence, the heat capacity curves are not included in this thesis. It is not possible to draw conclusions whether a triple point exists from the nested sampling runs if they are not converged. Hence, I use thermodynamic integration based on the structures which I found with the nested sampling algorithm with 1024 atoms for the β and α phase and 1536 atoms for the ω phase to get a reliable result. My phase diagram predictions for the original MEAM [15] using thermodynamic integration are first shown in Section 4.6. These results show that the triple point does not exist at pressures below 12 GPa and confirm my converged nested sampling results for 24 and 54 atom qualitatively. These means that I do not reproduce the triple point found for the phase diagram for the Ti MEAM of Hennig *et al.* [15]. Hennig *et al.* analyse the

NPT [144] MD trajectory of a β phase structure at various points of the pressure-temperature space while applying a Langevin thermostat [15]. They consider the phase in which the system settles to be the stable phase for the individual pressure-temperature point. This allows them to infer the positions of solid-solid transitions. The initial simulation boxes are cubes. They use 432 atoms and a maximum of 10^6 time steps. The time step length is 1 fs. However, the result of this method may be influenced by the energy barrier of the transition path. If the energy barrier for the β to α phase is much higher than the β to ω barrier, the system may end up in a metastable ω phase even though the α phase is more favourable. In the case that the α phase is favourable long enough simulation times would see the ω phase transform to the α phase. However, it is never clear what the appropriate simulation time is. Indeed, Hennig *et al.* find that 1 ns is not enough to let an initial ω phase transform in to the α phase and vice versa when they conduct *NVT* MD simulations with a Langevin thermostat [15]. The pressure which is roughly equivalent to the volume they used in these simulations is 10 GPa [15]. The temperature of these simulations was 1200 K [15]. Hennig *et al.* also conduct interface simulations in this *NVT* ensemble with the Langevin thermostat, on the basis of which they conclude that the α phase is stable at ca 0 GPa and the ω phase is stable at ca 15 GPa [15]. Hennig *et al.* set the temperature to 1200 K for both simulations. As the phases have different volumes per atom for the MEAM of Hennig *et al.* [15] at 0 K and 0 GPa as already shown by Hennig *et al.* themselves [15], these interface simulations may have however issues with interface stresses. The thermodynamic integration does not suffer from the issues of the two methods utilised by Hennig *et al.* [15] to investigate solid-solid transitions. Given the small difference in Gibbs free energy between the α and the ω phase for temperatures after the transition temperature at 12 GPa in Figure 4.11, it is possible that the phase diagram prediction of Hennig *et al.* [15] is not fully correct. My thermodynamic integration calculations show inaccuracies at low temperatures due to the classical treatment of the system as covered in Sections 2.7 and 4.6. However, this is also the case for the methods for phase transition determination of Hennig *et al.* as they also rely on the classical MD [15]. Additionally, I am mainly interested in the question whether a triple point is predicted by the original MEAM of Hennig *et al.*[15]. The phase diagram of Hennig *et al.* shows it at ca 1250 K and ca 8.5 GPa while they estimate a value of ca 8 GPa and ca 1200 K in their text [15]. These temperatures are far away from the realm where one would expect any significant error due to the classical treatment of the system as demonstrated for 12 GPa in Section 4.6 (see Figure 4.10). Hence, I am confident in my assumption for the fitting that my thermodynamic integration predictions are correct and that the experimentally seen triple point I try to achieve with my later modifications at a pressure of 7.5 GPa and a temperature of 913 K [51] is not predicted by the original MEAM of Hennig *et al.* [15].

To improve the Ti MEAM of Hennig *et al.* [15], I made changes to the functions $f(r)$ and $g\left(\cos\left(\theta_{jik}\right)\right)$ via Fourier modes as explained at the beginning of Section 4.8. These functions control directly and indirectly how the angle defined by the atomic bonds ij and ik influences the potential energy. I do not modify other functions as I want to change the original potential [15] only slightly to preserve much of its original features reported in Ref. [15] and as I want to limit the number of influences to the modifications for my first trials for this potential. The modification to $f(r)$ increases or decreases the characteristic behaviour depending on its single parameter A (see the beginning of Section 4.8) as demonstrated in Section 4.10 (see Subfigure 4.22a). The modification to $g\left(\cos\left(\theta_{jik}\right)\right)$, leads to a symmetric change of the spline nodes and is also only controlled by A. In principle, one could have taken even simpler modifications using single modes, testing various of these separately on one function at a time. Additionally, one could define the periodic range as 2, the range of the function $\cos\left(\theta_{jik}\right)$. However, the modifications shown in this thesis are already very promising so that these options are not explored. The current modifications are sufficient to prove that a successful modification is possible. An additional point to make is that I only modify the spline nodes of the Ti MEAM of Hennig *et al.* This is not the same as adding the Fourier modes to the original functions directly. This limits which Functions can be approximated via adding the Fourier modes leading perhaps to unexpected behaviour of the splines if strong changes were necessary. However, it might also be useful in smoothing out some of the high frequency Fourier modes which might be beneficial. Given that the modifications are relatively small and only few modes already lead to very good results, these problems can most likely be neglected. Out of several possible parameter values, I have settled on $A = -0.04$ as elaborated on in Section 4.8. Thermodynamic integration calculations and "MD stability tests" for the ω phase when necessary as presented in Sections 4.8.2 and 4.8.1 show a clear improvement in predicting the P-T phase diagram compared to my predictions for the original Ti MEAM of Hennig *et al.* [15]. Most importantly, my calculations indicate the existence of a triple point, likely between 8 and 10 GPa. This is close to the experimental triple point pressure estimate of 7.5 GPa by Zhang *et al.* [51]. The "MD stability tests" for the ω to β transition are the same principle as the single phase MD simulations with which Hennig *et al.* [15] calculate their phase diagram, just that I start with an ω phase in a hexagonal simulation cell instead of a cubic β phase. In principle, it has the same limitations which I have already elaborated on in this section. Even though the modified MEAM with $A = -0.04$ does now predict a triple point as well as an α to β transition temperature which is much closer to the experimental results of Ref. [51], and is hence a clear improvement on the original MEAM of Hennig *et al.* [15] in this respect, the potential could be modified to reproduce the results of Zhang *et al.* [51] even better. According to

Zhang *et al.* [51], the ω phase should not be stable at any temperature for low pressures as discussed already in Section 4.8.2. The modified MEAM with $A = -0.04$ predicts however exactly that. The modification even increases the temperature up to which it should be stable for low pressures compared to my results for the original MEAM of Hennig *et al.* [15]. It is however to note that before the work of Zhang *et al.* [51] a much lower value of the slope of the α to ω transition of 93 K/GPa was assumed [50]. Using this slope and the previously assumed triple point positions of ca 8 GPa and ca 913 K [50], I extrapolate a transition temperature of the ω phase to the α phase at 0 GPa at ca 169 K. Given the struggles in determining the α to ω transition pressure at a given temperature experimentally [50], this is naturally not a proof that there is a stable ω phase for some low temperatures at 0 GPa. However, given the hysteresis which limits experiments on whether the ω or α phase is stable at a given pressure and temperature greatly [50], a qualitative reproduction of the experimentally found α to ω transition slope of Zhang *et al.* [51] is sufficient. In this respect, the modification of the original MEAM [15] actually worsened the phase diagram prediction as it increased the ω to α phase transition temperature for low pressures compared to the original MEAM as discussed in Section 4.8.1. However, in light of the fact that the modified MEAM with $A = -0.04$ now reproduces the triple point, this is of lesser importance. The lattice constants, lattice constant ratios and elastic constants which I have calculated with the modified MEAM with $A = -0.04$ are overall still in good agreement with the parameters which I have calculated using the original MEAM of Hennig *et al.* [15] as shown in Section 4.10 as well as in relatively good agreement with the presented experimental results. For the original MEAM [15], this is in most parts a reproduction of the values calculated by Hennig *et al.* [15]. The different values which they find for the c_{11} and c_{12} elastic constants might be due to different unit cell definitions as I already noted in Section 4.10. The quantities tested for the new MEAM are surely not exhaustive, but nevertheless give an indication how different the new MEAM is compared to the original MEAM of Hennig *et al.* when it comes to modelling the α, β and ω phases. The mostly slight changes to the lattice constants and elastic constants combined with the considerable improvement the phase diagram prediction indicates that the potential has been improved while possibly only slightly changing its previously fitted behaviour when it comes to other parts.

It is also to note that I did not investigate how the melting curve prediction changed due to the modification. This could be included in a future fitting of titanium.

It is also interesting to note that the $A = 0.04$ MEAM energies of the sampled structures whose DFT results are presented in Section 4.7 did not change considerably overall as covered in Section 4.7. However, the DFT Gibbs free energy has not been calculated in this work. It is in principle possible that the modified MEAM Gibbs free energy is much closer

to the DFT Gibbs free energy then the original MEAM [15] Gibbs free energy for at least some temperatures.
These good results achieved with only slight modifications are a great motivation to apply Nelder-Mead fitting as I have done in this thesis for the original CuAu potential [39–41] to the Ti MEAM of Hennig *et al.* [15]. I elaborate on this further in Section 6.2.

Chapter 6

Conclusion and Outlook

6.1 Conclusion

In this thesis I presented modifications of two potentials – a CuAu EAM potential [39–41] and the Ti MEAM potential of Hennig *et al.* [15] – with the aim of reproducing experimentally found phase transitions. For both, I use the nested sampling [25, 26] implementation *pymatnest* [45] to determine phase transitions. For the Ti calculations, I also rely on the additional use of informed search methods.

For the CuAu potential, I wrote an implementation of the Nelder-Mead algorithm [35] to iteratively fit the potential to various melting points. For each evaluation of the objective function, I used several nested sampling runs to calculate the melting temperature for the current version of the EAM. I was able to create a potential which reproduces the experimental melting curve for the compositions considered in the fitting in only 22 iterations and after calculating only 38 objective functions.

Additionally to nested sampling, I used thermodynamic integration and when needed "MD stability tests" of the ω phase to determine the phase diagram for both the original Ti MEAM of Hennig *et al.* [15] and my modification to this Ti MEAM, as it is too computationally expensive to extrapolate the phase diagram with nested sampling for high number of atoms as done for the CuAu potential for this system and potential. I showed that the original MEAM of Hennig *et al.* [15] does not show a triple point for pressures $P \leq 12\,\mathrm{GPA}$. I made several changes to the Ti MEAM potential using Fourier modes. Eventually, I produced a potential (modification parameter $A = -0.04$) which shows an indication of a triple point with thermodynamic integration. This means that one can now qualitatively reproduce the experimental solid-solid transition phase diagram to a large extent with this modified MEAM.

Apart from these main results, I also would stress the importance of checking for convergence with regards to the number of atoms per simulation box when investigating the MEAM of

Hennig *et al.* [15] and my modifications to it with nested sampling. As seen in Sections 4.5 and 4.8, the Ti phase transitions showed a strong dependence on the number of atoms per simulation box. Additionally, the melting curve of the CuAu potential [39–41] also showed a strong dependence on the number of atoms as seen in Section 3.6. Indeed, the potential performs far worse than assumed in Ref. [28] in comparison to experimental data. (Ref. [28] quote data from Ref. [42]. I use data from Ref.[88] and Bennett [87] quoted in Okamoto *et al.* [42] for my comparison.) For the melting temperature T_{melt}, I was able to establish that $T_{melt}(N) = T_{m,\infty} + k/N^2$ where N is the number of atoms per simulation box is a good fit. The melting temperature for $N \to \infty$ can then be determined via fitting to melting temperatures calculated with nested sampling for different numbers of atoms. The parameter k is another value to be fitted. Additionally, I was able to show that the ratio $T_{m,\infty}/T_{melt,N_{ref}}$ of the specific system, where $T_{melt,N_{ref}}$ is the melting temperature for a reference number of atoms N_{ref}, does not change too much for small modifications to the potential (see Section 3.6). This allows fast calculation of the melting temperature during my fitting routine in the first place, as I otherwise would need to redo the nested sampling calculations at different numbers of atoms and repeat the fitting for each new modified potential.

For Ti, I could show that the combination of nested sampling with thermodynamic integration is very powerful. Nested sampling, even if not fully converged, yields the relevant phases which compete against each other. Thermodynamic integration – or any other reliable informed search method – is then able to calculate the phase diagram to a higher precision. This allows for a precise calculation of the phase diagram while at the same time saving computational resources as one does not need to conduct expensive nested sampling runs for high number of atoms. If all relevant phases are indeed discovered by nested sampling, this combination of methods also avoids the problem of neglecting a relevant phase which is inherent in informed search methods.

It is also interesting to note that the phases appear relatively stable in regard to small changes to the potential parameters both for the modifications to the CuAu potential [39–41] and for the modifications to the Ti MEAM [15]. I did not find any appearance of phases which were not already shown in the nested sampling calculation of the original potentials in the investigated temperature, pressure and alloy composition realm. This allows for limiting the number of nested sampling calculations during fitting as one could potentially use informed search methods like thermodynamic integration or coexistence simulations to replace some nested sampling calculations during the fitting procedure.

To summarise, I successfully used nested sampling and as well as nested sampling and thermodynamic integration to show the previous predictions of the melting curve [28] for a CuAu potential [39–41] and the solid-solid transitions [15] for a Ti MEAM [15]

are not fully correct and show considerable differences from experimental results. I was able to successfully improve these interatomic potentials by modifying them to reproduce experimental phase transition data while preserving much of the elastic constants and lattice constants predictions of the original potentials. For the original Ti MEAM [15] the parameters I calculated were mostly a recalculation of parameters calculated in Ref. [15]. The Ti MEAM [15] was modified by choosing from various values of the parameter A which controls the modification strength. However, for fitting the CuAu melting curve, I created a Nelder-Mead [35] implementation which allows for nearly automatic scheduling and analysis of nested sampling runs with only limited human supervision necessary. This algorithm could in principle already fit solid-solid transitions provided that new phases do not need to appear during the fitting and existing phases do not need to vanish during the fitting. The objective function would need to be updated though to deal with more complex situations as explained in Section 6.2. Concluding, I have shown that improving interatomic potentials by fitting to experimental phase diagram data is possible and enlarged the range of transferability of the CuAu EAM [39–41] and the Ti MEAM [15]. This opens up the possibility of improving other existing potentials from the literature in the same manner.

6.2 Outlook

From my work, it appears that small modifications to the potentials do not easily stabilise a previously unconsidered phase. It follows that a fitting scheme which combines nested sampling with informed search methods may be faster and cheaper than the scheme used for the CuAu potential fitting in this thesis: One could conduct nested sampling calculations with the original potential and verify the result with informed search methods based on the competing structures found with nested sampling. A certain number of iteration could then be conducted with the informed search methods. Afterwards, one could again do a nested sampling calculation to see if there is any new relevant phase and from then do new calculations with informed search methods based on the relevant phases found with the most recent nested sampling calculations. However, if a phase occurred which would not have been considered in the informed search method, one might have to restart the algorithm. In the worst case scenario, it might actually be more efficient to apply a combination of nested sampling and informed search methods for each vertex. The addition of informed search methods to the scheme may be especially useful for a rigorous refitting of the titanium potential of Hennig *et al.*[15] using my implementation of the Nelder-Mead algorithm. Here one would need to consider different phase transitions at different pressures instead of different alloy compositions. The heat capacity peaks in the heat capacity curves produced

with nested sampling have a finite width. This means that for example Ti transitioning at different temperatures from ω to α to β can be mistaken for showing a direct ω to β transition if the difference between the ω to α transition temperature and the α to β transition temperature is too small. In this case, the two transition peaks would add up to one peak. Combined with the finite size convergence problem of nested sampling shown in my Ti calculations, this means that Gibbs free energy calculations as for example thermodynamic integration are more precise when employing the same amount of computational resources. The fitting of the Ti MEAM [15], requires however the creation of a new phase transition, the ω to β transition to reproduce the triple point found for titanium [51]. This means that disappearing of the α phase needs to be considered somehow in the fitting. One measure is to introduce some kind of punishment for the missing of a transition and the existence of other transitions which are not predicted by experiments for each pressure one considers. The squared difference between phase transition temperatures enclosing a phase which needs to vanish – for Ti the ω to α and the α to β transition temperature at pressures above the triple point pressure – could also be included in the objective function calculation. This could be a measure on how much of the unwanted phase is still predicted by the potential.
However, the method given above assumes that one does not need to create a totally new phase during the potential fitting. Additionally, it necessitates potential specific input on which transitions are wanted and which should be avoided. If a new phase appeared for one of the modifications of the original potential, the objective function calculation would need to be adjusted. A more general way which also would be much more convenient to automatise the evaluation of the objective function could be the following: The phase diagram is calculated for a vertex, for example with the combination of nested sampling and informed search methods. Then a grid is created over the phase diagram. For each point of the grid, one could calculate a measure of similarity between the most likely structures of the nested sampling trajectory at this point to the structure predicted from experiments for this point. A thermal average similar to the thermal average of the XRD pattern used in this thesis could then be conducted. The work of Ref. [145] and Ref. [146] could be a good starting point for constructing a similarity measure. This measure would need to be a number ≥ 0 which tends to 0 the more similar two structures become. The average of all these points would then be a measure for how close the phase diagram predicted by the potential is to the experimental one. It would be interesting to research if this could then be a suitable objective function to fit.

Appendix A

Appendix

A.1 Programs

For this thesis, I use a variety of scientific programs, scripts and libraries. In the following I want to mention the main ones and acknowledge the programs of others if not done in other parts of the thesis. *QUIP* [114] and *quippy* [120] are used at various points. *quippy* allows us to call *QUIP* as it wraps *QUIP* routines and functions. *quippy* is also necessary for some features of *pymatnest* [45]. *quippy* was especially useful for creating atomic structures and is hence used by my collection of miscellaneous functions *misc_calc_lib.py* [45]. For the function "calc_aspec_ratio(at)" of this library, I re-implemented the calculation of the aspect ratios from Noam Bernstein's "min_aspect_ratio(vol, cell)" in "ns_run.py" which is part of *pymatnest* [45]. For the nested sampling calculations I use *pymatnest* [45] which I elaborate on further in Section 3.2. For some calculations of the heat capacity curves, I use "analysis_for_pymatnest.f90" from my previous work, Ref. [33]. "analysis_for_pymatnest.f90" itself is a re-implementation of Noam Bernstein's "ns_analyse.py" from *pymatnest* [45] as already explained in Ref. [33]. Later calculations of the heat capacity are done directly with "ns_analyse.py" from *pymatnest* [45]. *pymatnest* [45] has further scripts for profiling runs and combining trajectories which I use. Additionally, I created a program that allows for the calculation of thermally averaged radial distribution functions (RDF, disabled by default and not used in this thesis) and the X-ray diffraction pattern (XRD pattern) using *QUIP* [114] which is now part of *pymatnest* [45]. Additionally I included my library of miscellaneous functions *misc_calc_lib.py* of which some are used in the XRD-averaging program in *pymatnest* [45]. The XRD (and RDF) calculation is done via *QUIP*. In the current version it is called over my aforementioned library of miscellaneous functions. For some calculations, I use a bash script which consists of parts of a analysis script by Noam Bernstein [147] instead. The XRD calculation is used also to identify the phase of systems gotten

during LAMMPS [74] MD simulations described in Subsection 4.6.3.
I use my own implementation of the Nelder-Mead algorithm [35] to fit the melting temperatures. I cover the most important details of the implementation in Section 3.7. For my thermodynamic integration calculations, I use a LAMMPS [74] implementation [48, 116]. Further details can be found in Subsection 4.6.1. LAMMPS [74] is also used for MD calculations to check the stability of a phase. The LAMMPS script for this is the same as used for the *NPT* [115, 118] equilibration of thermodynamic integration and follows the description on how to perform the simulations in LAMMPS of Freitas *et al.* [48] (see Subsection 4.6.1). To calculate the elastic constants via LAMMPS [74, 109] and analyse the results, I use scripts [108] shipped with LAMMPS [74, 109] which I slightly adjust for my purposes.
The original CuAu EAM [39–41] used in this thesis and its modifications are always evaluated via LAMMPS [74]. This is done either directly with LAMMPS or indirectly with LAMMPS via *pymatnest* [45]. The MEAM of Hennig *et al.* [15] and its modifications is evaluated usually via LAMMPS [74, 137] directly or evaluated indirectly with LAMMPS [74, 137] via *pymatnest* [45]. LAMMPS [74, 137] has its own parameter file (added by Alexander Stukowski). Hennig *et al.* [15], though have also published their original parameters [148]. For use with *QUIP* [114] and *quippy* [120], I created a new potential parameter file [149] during the time leading up to my MPhil thesis [33] using very slightly adjusted programs created originally for Ref. [150]. For this, parameters are taken directly from Ref. [148]. Exceptions to the usage of *LAMMPS* are the calculations related to *Caesar* [127] where *QUIP* [114] and *quippy* [120] are used, the calculations of the MEAM energies (original MEAM [15] and modified MEAM) for comparison to the DFT values in Sections 4.7 and 4.9 for which I use *quippy* to call the *QUIP* implementation and the "MD averaging" (see Subsection 4.5.1) for which I use *QUIP*. The modifications for the modified MEAM to the *QUIP* potential file [149] and to the LAMMPS potential file [74, 137] are based on the LAMMPS parameters [74, 137]. I use *MATLAB* [136] for the modifications of the splines (see Section 4.8) and printing of the spline interpolations presented in Section 4.10. Whenever necessary, a script by Albert Bartók-Pártay [122] is used to transform an ".extxyz" format file to the format which *LAMMPS* [74] reads. The scripts [122–124, 48] used in thermodynamic integration are listed and elaborated on in Subsection 4.6.1. I also use *SciPy* [103] and *numpy* [151, 152] at various points during my work. Additionally, *OVITO* [138], sometimes in conjunction with its "bond-angle analysis" [139, 140], is used to investigate some structures visually. *OVITO* is also used to create the sketches of the unit cell of the different phases in Sections 2.9 and 2.10.

A.2 DFT parameters

I use *CASTEP* [133] to calculate reference DFT [153–155] energies presented in Sections 4.7 and 4.9. For the DFT reference calculations, I use the "PBE" generalized gradient approximation (GGA) functional [134, 135], a cutoff energy of 700 eV, Gaussian smearing (see for example [156]) with a smearing width of 0.1 eV, a "k-point spacing" of 0.015 Å^{-1} for the k-point grid [157], a "finite basis spacing" of 10 eV, a maximum of 300 SCF cycles, 30 % additional bands (*CASTEP* parameter "PERC_EXTRA_BANDS") and the Broyden mixing scheme [158, 159] considering the last 10 electron densities with the *CASTEP* parameter "MIX_CHARGE_AMP" set to 0.1. The energy I use is calculated by *CASTEP* as average of free energy and energy which limits the effect of the smearing [160]. The finite basis set corrections [161] are not included in the values reported in this thesis. To approximate the energy of an isolated atom for the cohesive energy, I use a single atom in a cubic simulation box with an edge length of 20 Å with the same parameters as in the reference calculations.

A.3 Salient parameters of the nested sampling runs

In this section I present the parameters which are most important for the nested sampling calculations used or mentioned in this thesis together with a reference to the point in the thesis where the specific calculations are used. The values of n_{cull}, the number of culls per iteration, are kept low as Noam Bernstein and Livia Bartók-Pártay discovered convergence problems when using some n_{cull} values greater than 1 [113]. My trials in Section 4.4 are however a good indication that small n_{cull} values can be used. The CuAu nested sampling parameters are based on the ones used by Baldock *et al.* [28] for their CuAu calculations using the original CuAu potential [39–41] which can be found under Ref. [43]. My parameters are only very slight modifications to theirs. I use a pressure of ca 0.1 GPa for all my Au, Cu and CuAu 50:50 calculations and batches of 5 runs. Further parameters are given in Table A.1.

Table A.2, shows these parameters for Ti for figures and arguments relying solely on the original MEAM of Hennig *et al.* [15]. Table A.2 on the other hand is used for the parameters for Ti used for figures and arguments considering the modification of the original MEAM of Hennig *et al.* [15].

Reference to usage	atomic comp.	n_{cull}	L	K	step ratios
Fig. 3.1	36:0	8	768	1152	8:16:8:8:8
Fig. 3.2	24, 32, 36, 40	8	768	1152	8:16:8:8:8
(Not part of fit)					
Fig. 3.2	48, 54, 72	8	768	1152	8:16:8:8:8
(Part of fit)	64, 128	1	768	1152	8:16:8:8:8
Tab. 3.1	0:36	8	1536	1152	8:16:8:8:8
(Reference melting	18:18	8	1536	1152	8:16:8:8:8
temperatures)	36:0	8	1536	1152	8:16:8:8:8
Tab. 3.1, Fig. 3.6, 3.7, 3.10	48	8	1536	1152	8:16:8:8:8
(Extrapolation of	54	8	1536	1152	8:16:8:8:8
melting temperatures,	64	8	1536	1152	8:16:8:8:8
total number of atoms	72	8	1536	1152	8:16:8:8:8
given)	128	8	1536	1152	8:16:8:8:8
Sec. 3.11	0:24	8	1536	1152	8:16:8:8:8
(Fitting of pure Au)					
Sec. 3.13.1	12:12	8	1536	1152	8:16:8:8:8
(Fitting of CuAu 12:12)					
Sec. 3.13.2	18:18	8	1536	1152	8:16:8:8:8
(Fitting of CuAu 18:18)					
Sec. 3.14	0:36	8	768	1152	8:16:8:8:8
(Fitting of Au,	18:18	8	1536	1152	8:16:8:8:8
CuAu 18:18 and Cu)	36:0	8	768	1152	8:16:8:8:8

Table A.1 Salient parameters used in the various Cu, Au and CuAu 50:50 nested sampling runs. The label "atomic comp." stands for "atomic composition". It is either given as N_{Cu}:N_{Au}, where N_{Cu} and N_{Au} are the numbers of atoms per simulation box of Cu and Au respectively, or given as the total number of atoms $N_{tot} = N_{Cu} + N_{Au}$. In the latter case, it means that all different compositions with this total atom number referred to, have the same parameters. Further parameters are the number of culls per iteration n_{cull}, the total number of calls expected during the MD scheme L (see Section 3.2), the number of walkers K and the step ratios (see Section 3.2). The order of different steps for the step ratio is L_{traj}, n_V, n_{shear}, $n_{stretch}$, n_{swap}. I use a pressure of ca 0.1 GPa. A single MD trajectory of length L_{traj} is used.

Reference to usage	N	P [GPa]	n_{cull}	L	K	step ratios
Fig. 4.1a, Fig. 4.2a	54	10	1	840	3840	4:1:1:1
Fig. 4.1b	54	10	1	840	7680	4:1:1:1
Fig. 4.1c	54	10	1	672	15360	4:1:1:1
Fig. 4.2b	54	10	1	3840	3840	4:1:1:1
Fig. 4.3a	54	10	4	240	3840	8:32:8:8
Fig. 4.3b	54	10	8	240	3840	8:32:8:8
End of Sec. 4.3	54	10	8	120	7680	8:32:8:8
	54	10	8	240	7680	8:32:8:8
Fig. 4.4a	54	10	8	240	7680	8:32:8:8
Fig. 4.4b	54	10	16	240	7680	8:32:8:8
Fig. 4.4c	54	10	32	240	7680	8:32:8:8
Fig. 4.5	24	2	8	240	15360	8:32:8:8
	24	4	1	840	7680	4:1:1:1
	24	6	1	840	7680	4:1:1:1
	24	8	1	840	7680	4:1:1:1
	24	10	1	840	7680	4:1:1:1
	24	12	1	840	7680	4:1:1:1
Sec. 4.5	24	2	1	840	7680	4:1:1:1
Fig. 4.6	54	4	1	840	15360	4:1:1:1
	54	6	1	840	15360	4:1:1:1
	54	8	1	672	15360	4:1:1:1
	54	10	1	672	15360	4:1:1:1
	54	12	1	840	7680	4:1:1:1
Fig. 4.7	108	12	8	240	7680	8:32:8:8
	54	12	1	840	7680	4:1:1:1
Fig. 4.8, Fig. 4.9	108	12	8	240	7680	8:32:8:8
Figs 4.14, 4.15, 4.20, 4.21	24	2	8	240	15360	8:32:8:8
	24	6	8	240	7680	8:32:8:8
	24	10	8	240	7680	8:32:8:8
Sec. 5.2	216	12	8	240	3840	8:32:8:8

Table A.2 Salient parameters used in the nested sampling calculations mentioned in this thesis for Ti using the original MEAM of Hennig *et al.* [15]. Here, N is the number of atoms per simulation box, P is the pressure in GPa, n_{cull} is the number of walkers culled per iteration, L is the overall number of calls as elaborated on in Section 3.2, the number of walkers is given with K and "step ratios" denotes the ratio the different step kinds have to each other. In this case, the ordering is L_{traj}, n_V, n_{shear}, $n_{stretch}$ with steps as explained in Section 3.2. The length of the MD trajectory given in this ratio is made from one single trajectory which is not broken up.

Reference to usage	N	P [GPa]	n_{cull}	L	K	step ratios	modification
Fig. 4.16	24	12	8	240	15360	8:32:8:8	$A = 0.1$
	24	12	8	240	15360	8:32:8:8	$A = 0.05$
	24	12	8	240	7680	8:32:8:8	$A = -0.02$
	24	12	8	240	7680	8:32:8:8	$A = -0.04$
	24	12	8	240	7680	8:32:8:8	$A = -0.08$
	24	12	8	240	15360	8:32:8:8	original
Sec. 4.8	24	4	8	240	7680	8:32:8:8	$A = -0.08$
	24	8	8	240	7680	8:32:8:8	$A = -0.08$
Sec. 4.8	24	2	8	240	15360	8:32:8:8	$A = -0.04$
Sec. 4.8	54	6	8	240	15360	8:32:8:8	$A = -0.04$
Fig. 4.17	24	12	8	240	7680	8:32:8:8	$A = -0.04$
	54	12	8	240	7680	8:32:8:8	$A = -0.04$

Table A.3 Salient parameters used in the nested sampling calculation when discussing modifications to the original MEAM of Hennig *et al.* [15]. The definitions of the labels are the same as in Table A.2, apart from the addition of the column "modification". This column denotes which potential is used for the calculation. The original MEAM [15] is denoted by the label "original". For modifications of this original MEAM, I list the modification parameter A which is explained in Section 4.8. Again, the length of the MD trajectory given for the "step ratio" is again from a single trajectory.

References

[1] A. Thiess, R. Zeller, M. Bolten, P. H. Dederichs, S. Blügel, Massively parallel density functional calculations for thousands of atoms: KKRnano, Physical Review B 85 (23) (2012) 235103.

[2] J. E. Jones, On the determination of molecular fields.—I. From the variation of the viscosity of a gas with temperature, Proceedings of the Royal Society A 106 (738) (1924) 441–462.

[3] J. E. Jones, On the determination of molecular fields.—II. From the equation of state of a gas, Proceedings of the Royal Society A 106 (738) (1924) 463–477.

[4] M. S. Daw, M. Baskes, Semiempirical, quantum mechanical calculation of hydrogen embrittlement in metals, Physical Review Letters 50 (17) (1983) 1285–1288.

[5] M. S. Daw, M. I. Baskes, Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals, Physical Review B 29 (12) (1984) 6443–6453.

[6] M. I. Baskes, Application of the embedded-atom method to covalent materials: A semiempirical potential for silicon, Physical Review Letters 59 (23) (1987) 2666–2669.

[7] M. I. Baskes, J. S. Nelson, A. F. Wright, Semiempirical modified embedded-atom potentials for silicon and germanium, Physical Review B 40 (9) (1989) 6085–6100.

[8] J. Tersoff, New empirical approach for the structure and energy of covalent systems, Physical Review B 37 (12) (1988) 6991–7000.

[9] A. C. Van Duin, S. Dasgupta, F. Lorant, W. A. Goddard, ReaxFF: A reactive force field for hydrocarbons, The Journal of Physical Chemistry A 105 (41) (2001) 9396–9409.

[10] A. P. Bartók, M. C. Payne, R. Kondor, G. Csányi, Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons, Physical Review Letters 104 (13) (2010) 136403, GAP is available for non-commercial use from www.libatoms.org.

[11] A. P. Bartók, R. Kondor, G. Csányi, On representing chemical environments, Physical Review B 87 (18) (2013) 184115, GAP is available for non-commercial use from www.libatoms.org.

[12] J. Behler, Perspective: Machine learning potentials for atomistic simulations, The Journal of Chemical Physics 145 (17) (2016) 170901.

[13] J. A. Martinez, D. E. Yilmaz, T. Liang, S. B. Sinnott, S. R. Phillpot, Fitting empirical potentials: Challenges and methodologies, Current Opinion in Solid State and Materials Science 17 (6) (2013) 263–270, Frontiers in Methods for Materials Simulations.

[14] M. Mendelev, S. Han, D. Srolovitz, G. Ackland, D. Sun, M. Asta, Development of new interatomic potentials appropriate for crystalline and liquid iron, Philosophical Magazine 83 (35) (2003) 3977–3994.

[15] R. G. Hennig, T. J. Lenosky, D. R. Trinkle, S. P. Rudin, J. W. Wilkins, Classical potential describes martensitic phase transformations between the α, β, and ω titanium phases, Physical Review B 78 (5) (2008) 054121.

[16] G. J. Ackland, Temperature dependence in interatomic potentials and an improved potential for Ti, Journal of Physics: Conference Series 402 (1) (2012) 012001.

[17] G. Bonny, R. C. Pasianot, L. Malerba, Fitting interatomic potentials consistent with thermodynamics: Fe, Cu, Ni and their alloys, Philosophical Magazine 89 (34–36) (2009) 3451–3464.

[18] R. C. Pasianot, L. Malerba, Interatomic potentials consistent with thermodynamics: The Fe–Cu system, Journal of Nuclear Materials 360 (2) (2007) 118–127.

[19] J. Winey, A. Kubota, Y. Gupta, A thermodynamic approach to determine accurate potentials for molecular dynamics simulations: Thermoelastic response of aluminum, Modelling and Simulation in Materials Science and Engineering 17 (5) (2009) 055004.

[20] R. J. N. Baldock, L. B. Pártay, A. P. Bartók, M. C. Payne, G. Csányi, Determining pressure-temperature phase diagrams of materials, Physical Review B 93 (17) (2016) 174108.

[21] L. B. Pártay, On the performance of interatomic potential models of iron: Comparison of the phase diagrams, Computational Materials Science 149 (2018) 153–157.

[22] G. Bonny, R. C. Pasianot, N. Castin, L. Malerba, Ternary Fe–Cu–Ni many-body potential to model reactor pressure vessel steels: First validation by simulated thermal annealing, Philosophical Magazine 89 (34–36) (2009) 3531–3546.

[23] G. Bonny, R. Pasianot, L. Malerba, Fe–Ni many-body potential for metallurgical applications, Modelling and Simulation in Materials Science and Engineering 17 (2) (2009) 025010.

[24] G. J. Ackland, Theoretical study of titanium surfaces and defects with a new many-body potential, Philosophical Magazine A 66 (6) (1992) 917–932.

[25] J. Skilling, Nested Sampling, in: R. Fischer, R. Preuss, U. von Toussaint (Eds.), Bayesian inference and maximum entropy methods in science and engineering, American Institute of Physics Conference Proceedings, Vol. 735 (1), AIP Publishing, 2004, pp. 395–405.

[26] J. Skilling, Nested sampling for general Bayesian computation, Bayesian Analysis 1 (4) (2006) 833–860.

[27] R. J. N. Baldock, Classical statistical mechanics with nested sampling, PhD thesis, University of Cambridge (2015).

[28] R. J. N. Baldock, N. Bernstein, K. M. Salerno, L. B. Pártay, G. Csányi, Constant-pressure nested sampling with atomistic dynamics, Physical Review E 96 (4) (2017) 043311.

[29] F. Wang, D. P. Landau, Efficient, multiple-range random walk algorithm to calculate the density of states, Physical Review Letters 86 (10) (2001) 2050–2053.

[30] R. H. Swendsen, J.-S. Wang, Replica Monte Carlo simulation of spin-glasses, Physical Review Letters 57 (21) (1986) 2607–2609.

[31] D. D. Frantz, D. L. Freeman, J. D. Doll, Reducing quasi-ergodic behavior in Monte Carlo simulations by J-walking: Applications to atomic clusters, The Journal of Chemical Physics 93 (4) (1990) 2769–2784.

[32] L. B. Pártay, A. P. Bartók, G. Csányi, Efficient sampling of atomic configurational spaces, The Journal of Physical Chemistry B 114 (32) (2010) 10502–10512.

[33] M. Schlegel, Investigating the temperature driven phase transitions in titanium with nested sampling, MPhil thesis, University of Cambridge (2015).

[34] A. Gola, L. Pastewka, Embedded atom method potential for studying mechanical properties of binary Cu–Au alloys, Modelling and Simulation in Materials Science and Engineering 26 (5) (2018) 055006.

[35] J. A. Nelder, R. Mead, A simplex method for function minimization, The Computer Journal 7 (4) (1965) 308–313.

[36] S. Martiniani, J. D. Stevenson, D. J. Wales, D. Frenkel, Superposition enhanced nested sampling, Physical Review X 4 (3) (2014) 031034.

[37] L. B. Pártay, A. P. Bartók, G. Csányi, Nested sampling for materials: The case of hard spheres, Physical Review E 89 (2) (2014) 022302.

[38] N. S. Burkoff, R. J. Baldock, C. Várnai, D. L. Wild, G. Csányi, Exploiting molecular dynamics in nested sampling simulations of small peptides, Computer Physics Communications 201 (2016) 8–18.

[39] L. Ward, A. Agrawal, K. M. Flores, W. Windl arXiv:1209.0619 [cond-mat.mtrl-sci].

[40] The website https://atomistics.osu.edu/eam-potential-generator/index.php used by Ref. [28] is not working anymore, the parameter can now be found as 'AuCu.setfl.fs' in Ref. [43], the supplementary material to Ref. [28]].

[41] X. W. Zhou, R. A. Johnson, H. N. G. Wadley, Misfit-energy-increasing dislocations in vapor-deposited CoFe/NiFe multilayers, Physical Review B 69 (14) (2004) 144113.

[42] H. Okamoto, D. J. Chakrabarti, D. E. Laughlin, T. B. Massalski, The Au-Cu (gold-copper) system, Bulletin of Alloy Phase Diagrams 8 (5) (1987) 454–473.

[43] R. J. N. Baldock, N. Bernstein, K. M. Salerno, L. B. Pártay, G. Csányi, Supplementary material to Ref. [28], found under https://journals.aps.org/pre/abstract/10.1103/PhysRevE.96.043311#supplemental.

[44] N. Velisavljevic, S. MacLeod, H. Cynn, Titanium alloys at extreme pressure conditions, in: A. K. M. Nurul Amin (Ed.), Titanium Alloys - Towards Achieving Enhanced Properties for Diversified Applications, InTech, Rijeka, Croatia, 2012, pp. 67–86, available from http://www.intechopen.com/books/titanium-alloys-towards-achieving-enhanced-properties-for-diversified-applications/titanium-alloys-at-extreme-pressure-conditions accessed on 19 Sept. 2017.

[45] N. Bernstein, R. J. N. Baldock, L. B. Pártay, M. Schlegel, J. R. Kermode, T. D. Daff, A. P. Bartók, G. Csányi, pymatnest, https://github.com/libAtoms/pymatnest.

[46] X.-Y. Liu, F. Ercolessi, J. B. Adams, Aluminium interatomic potential from density functional theory calculations with improved stacking fault energy, Modelling and Simulation in Materials Science and Engineering 12 (4) (2004) 665–670.

[47] R. J. N. Baldock, L. B. Pártay, A. P. Bartók, M. C. Payne, G. Csányi, Determining pressure-temperature phase diagrams of materials: Supplementary information, Physical Review B, supplementary material to Ref. [20], available at http://journals.aps.org/prb/supplemental/10.1103/PhysRevB.93.174108.

[48] R. Freitas, M. Asta, M. de Koning, Nonequilibrium free-energy calculation of solids using LAMMPS, Computational Materials Science 112 (A) (2016) 333–341.

[49] M. de Koning, Optimizing the driving function for nonequilibrium free-energy calculations in the linear regime: A variational approach, The Journal of Chemical Physics 122 (10) (2005) 104106.

[50] E. Y. Tonkov, E. Ponyatovsky, Phase transformations of elements under high pressure, CRC press, Boca Raton, 2005.

[51] J. Zhang, Y. Zhao, R. S. Hixson, G. T. Gray III, L. Wang, W. Utsumi, S. Hiroyuki, H. Takanori, Experimental constraints on the phase diagram of titanium metal, Journal of Physics and Chemistry of Solids 69 (10) (2008) 2559–2563.

[52] D. Frenkel, B. Smit, Understanding molecular simulation: From algorithms to applications, Elsevier, San Diego, 2002.

[53] J. C. Jamieson, Crystal structures of titanium, zirconium, and hafnium at high pressures, Science 140 (3562) (1963) 72–73.

[54] E. Asadi, M. A. Zaeem, S. Nouranian, M. I. Baskes, Two-phase solid–liquid coexistence of Ni, Cu, and Al by molecular dynamics simulations using the modified embedded-atom method, Acta Materialia 86 (2015) 169–181.

[55] D. Trinkle, R. Hennig, S. Srinivasan, D. Hatch, M. Jones, H. Stokes, R. Albers, J. Wilkins, New mechanism for the α to ω martensitic transformation in pure titanium, Physical Review Letters 91 (2) (2003) 025701.

[56] D. Trinkle, D. Hatch, H. Stokes, R. Hennig, R. Albers, Systematic pathway generation and sorting in martensitic transformations: Titanium α to ω, Physical Review B 72 (1) (2005) 014105.

[57] S. Foiles, Calculation of the surface segregation of Ni-Cu alloys with the use of the embedded-atom method, Physical Review B 32 (12) (1985) 7685–7693.

[58] S. M. Foiles, M. I. Baskes, M. S. Daw, Embedded-atom-method functions for the fcc metals Cu, Ag, Au, Ni, Pd, Pt, and their alloys, Physical Review B 33 (12) (1986) 7983–7991.

[59] R. A. Johnson, Alloy models with the embedded-atom method, Physical Review B 39 (17) (1989) 12554–12599.

[60] M. I. Baskes, Modified embedded-atom potentials for cubic materials and impurities, Physical Review B 46 (5) (1992) 2727–2742.

[61] T. J. Lenosky, B. Sadigh, E. Alonso, V. V. Bulatov, T. D. de la Rubia, J. Kim, A. F. Voter, J. D. Kress, Highly optimized empirical potential model of silicon, Modelling and Simulation in Materials Science and Engineering 8 (6) (2000) 825–841.

[62] G. H. Givens, J. A. Hoeting, Computational Statistics, John Wiley & Sons, Hoboken, 2013.

[63] R. R. Barton, J. S. Ivey Jr, Modifications of the Nelder-Mead simplex method for stochastic simulation response optimization, in: Proceedings of the 23rd conference on Winter simulation, IEEE Computer Society, 1991, pp. 945–953.

[64] J. C. Lagarias, J. A. Reeds, M. H. Wright, P. E. Wright, Convergence properties of the Nelder–Mead simplex method in low dimensions, SIAM Journal on Optimization 9 (1) (1998) 112–147.

[65] K. I. M. McKinnon, Convergence of the Nelder–Mead simplex method to a nonstationary point, SIAM Journal on Optimization 9 (1) (1998) 148–158.

[66] C. T. Kelley, Detection and remediation of stagnation in the Nelder–Mead algorithm using a sufficient decrease condition, SIAM Journal on Optimization 10 (1) (1999) 43–55.

[67] G. J. Martyna, D. J. Tobias, M. L. Klein, Constant pressure molecular dynamics algorithms, The Journal of Chemical Physics 101 (5) (1994) 4177–4189.

[68] T.-Q. Yu, J. Alejandre, R. López-Rendón, G. J. Martyna, M. E. Tuckerman, Measure-preserving integrators for molecular dynamics in the isothermal–isobaric ensemble derived from the Liouville operator, Chemical Physics 370 (1–3) (2010) 294–305.

[69] N. S. Burkoff, C. Várnai, S. A. Wells, D. L. Wild, Exploring the energy landscapes of protein folding simulations with Bayesian computation, Biophysical Journal 102 (4) (2012) 878–886.

[70] A. Einstein, Die Plancksche Theorie der Strahlung und die Theorie der spezifischen Wärme, Annalen der Physik 327 (1) (1907) 180–190.

[71] J. G. Kirkwood, Statistical mechanics of fluid mixtures, The Journal of Chemical Physics 3 (5) (1935) 300–313.

[72] D. Frenkel, A. J. C. Ladd, New Monte Carlo method to compute the free energy of arbitrary solids. Application to the fcc and hcp phases of hard spheres, The Journal of Chemical Physics 81 (7) (1984) 3188–3193.

[73] M. Watanabe, W. P. Reinhardt, Direct dynamical calculation of entropy and free energy by adiabatic switching, Physical Review Letters 65 (26) (1990) 3301–3304.

[74] S. Plimpton, Fast Parallel Algorithms for Short-Range Molecular Dynamics, Journal of Computational Physics 117 (1) (1995) 1–19, [Please also see http://lammps.sandia.gov].

[75] L.-W. Tsao, S.-Y. Sheu, C.-Y. Mou, Absolute entropy of simple point charge model water by adiabatic switching processes, The Journal of Chemical Physics 101 (3) (1994) 2302–2308.

[76] M. de Koning, A. Antonelli, Einstein crystal as a reference system in free energy estimation using adiabatic switching, Physical Review E 53 (1) (1996) 465–474.

[77] P. Goyal, C. W. Monroe, New foundations of Newman's theory for solid electrolytes: Thermodynamics and transient balances, Journal of The Electrochemical Society 164 (11) (2017) E3647–E3660.

[78] D. E. Smirnova, S. V. Starikov, V. V. Stegailov, Interatomic potential for uranium in a wide range of pressures and temperatures, Journal of Physics: Condensed Matter 24 (1) (2012) 015702.

[79] R. R. Pawar, V. T. Deshpande, The anisotropy of the thermal expansion of α-titanium, Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography 24 (2) (1968) 316–317.

[80] S. K. Sikka, Y. K. Vohra, R. Chidambaram, Omega phase in materials, Progress in Materials Science 27 (3–4) (1982) 245–310.

[81] D. Errandonea, B. Schwager, R. Ditz, C. Gessmann, R. Boehler, M. Ross, Systematics of transition-metal melting, Physical Review B 63 (13) (2001) 132104.

[82] D. Errandonea, private correspondence, fit data for the Ti melting curve presented in Ref. [81] (2019).

[83] V. Stutzmann, A. Dewaele, J. Bouchet, F. Bottin, M. Mezouar, High-pressure melting curve of titanium, Physical Review B 92 (22) (2015) 224110.

[84] F. P. Bundy, Formation of new materials and structures by high-pressure treatment, in: Irreversible Effects of High Pressure and Temperature on Materials, ASTM Special Technical Publication No. 374, ASTM, Philadelphia, 1965, pp. 52–67.

[85] V. A. Zil'bershteyn, N. P. Chistotina, A. A. Zharov, N. S. Grishina, E. I. Estrin, Alpha-omega transformation in titanium and zirconium during shear deformation under pressure, The Physics of Metals and Metallography 39 (2) (1976) 208–211, translated from Russian by V. Alford [Fizika Metallov i Metallovedenie 39 (2) (1975) 445–447].

[86] E. Y. Tonkov, High pressure phase transformations: A handbook, Vol. 2, Gordon and Breach Science Publishers, Philadelphia, 1992.

[87] H. E. Bennett, The solidification curves of the gold-copper system, Journal of the Institute of Metals 91 (1962-1963) 158.

[88] H. Preston-Thomas, The international temperature scale of 1990 (ITS-90), Metrologia 27 (1) (1990) 3–10.

[89] D. A. Porter, K. E. Easterling, M. Y. Sherif, Phase Transformations in Metals and Alloys, CRC Press, Boca Raton, Florida, 2009.

[90] M. W. Finnis, J. E. Sinclair, A simple empirical N-body potential for transition metals, Philosophical Magazine A 50 (1) (1984) 45–55.

[91] J. Skilling, Bayesian computation in big spaces-nested sampling and Galilean Monte Carlo, AIP Conference Proceedings 1443 (1) (2012) 145–156.

[92] M. Betancourt, Nested sampling with constrained Hamiltonian Monte Carlo, AIP Conference Proceedings 1305 (1) (2011) 165–172.

[93] S. Duane, A. D. Kennedy, B. J. Pendleton, D. Roweth, Hybrid Monte Carlo, Physics Letters B 195 (2) (1987) 216–222.

[94] V. V. Brazhkin, Y. D. Fomin, A. G. Lyapin, V. N. Ryzhov, E. N. Tsiok, Widom line for the liquid-gas transition in Lennard-Jones system, The Journal of Physical Chemistry B 115 (48) (2011) 14112–14115.

[95] V. V. Brazhkin, A. G. Lyapin, V. N. Ryzhov, K. Trachenko, Y. D. Fomin, E. N. Tsiok, Where is the supercritical fluid on the phase diagram?, Physics-Uspekhi 55 (11) (2012) 1061–1079.

[96] D. Errandonea, The melting curve of ten metals up to 12 GPa and 1600 K, Journal of Applied Physics 108 (3) (2010) 033517.

[97] W. C. Roberts-Austen, K. T. Rose, On certain properties of the alloys of the gold-copper series, Proceedings of the Royal Society of London 67 (435-441) (1901) 105–112.

[98] G. Bonny, A. Bakaev, D. Terentyev, Y. A. Mastrikov, Interatomic potential to study plastic deformation in tungsten-rhenium alloys, Journal of Applied Physics 121 (16) (2017) 165107.

[99] P. C. Mahalanobis, On the generalized distance in statistics, Proceedings of the National Institute of Science of India 2 (1936) 49–55.

[100] R. De Maesschalck, D. Jouan-Rimbaud, D. L. Massart, The Mahalanobis distance, Chemometrics and Intelligent Laboratory Systems 50 (1) (2000) 1–18.

[101] J. D. Hunter, Matplotlib: A 2D graphics environment, Computing in Science & Engineering 9 (3) (2007) 90–95.

[102] H. Rosenbrock, An automatic method for finding the greatest or least value of a function, The Computer Journal 3 (3) (1960) 175–184.

[103] E. Jones, T. Oliphant, P. Peterson, et al., SciPy: Open source scientific tools for Python, http://www.scipy.org/ [accessed 29 March 2019] (2001–).

[104] Definition of various potential formats, https://lammps.sandia.gov/doc/pair_eam.html, [Last accessed 29 March 2019].

[105] matscipy, https://github.com/libAtoms/matscipy.

[106] W. C. Overton, J. Gaffney, Temperature variation of the elastic constants of cubic elements. I. Copper, Physical Review 98 (4) (1955) 969–977.

[107] J. Neighbours, G. Alers, Elastic constants of silver and gold, Physical Review 111 (3) (1958) 707–712.

[108] A. Thompson, scripts for calculating elastic constants in LAMMPS [74], shipped with LAMMPS [74], found in "/examples/ELASTIC" (git version of 5 October 2017).

[109] Website on elastic constant calculation with LAMMPS [74] mentioning scripts [108] available in "LAMMPS/examples/ELASTIC", see https://lammps.sandia.gov/doc/Howto_elastic.html [Last accessed on 29 March 2019].

[110] E. Polak, G. Ribiere, Note sur la convergence de méthodes de directions conjuguées, ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique 3 (R1) (1969) 35–43.

[111] W. P. Davey, Precision measurements of the lattice constants of twelve common metals, Physical Review 25 (6) (1925) 753–761.

[112] M. Straumanis, L. Yu, Lattice parameters, densities, expansion coefficients and perfection of structure of Cu and of Cu–In α phase, Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography 25 (6) (1969) 676–682.

[113] L. B. Pártay, private conversation (2016).

[114] QUIP, http://libatoms.github.io/QUIP/.

[115] T. Schneider, E. Stoll, Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions, Physical Review B 17 (3) (1978) 1302–1322.

[116] Website of the thermodynamic integration feature of *LAMMPS* [74] of Ref. [48], https://lammps.sandia.gov/doc/fix_ti_spring.html, implementation described in Ref. [48]. [Published online on LAMMPS webpage; last accessed on 29 March 2019].

[117] W. Shinoda, M. Shiga, M. Mikami, Rapid estimation of elastic constants by molecular dynamics simulation under constant stress, Physical Review B 69 (13) (2004) 134103.

[118] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, Journal of Applied Physics 52 (12) (1981) 7182–7190.

[119] M. E. Tuckerman, J. Alejandre, R. López-Rendón, A. L. Jochim, G. J. Martyna, A Liouville-operator derived measure-preserving integrator for molecular dynamics simulations in the isothermal–isobaric ensemble, Journal of Physics A: Mathematical and General 39 (19) (2006) 5629–5651.

[120] quippy, http://libatoms.github.io/QUIP/.

[121] Extended XYZ format, https://libatoms.github.io/QUIP/io.html#extendedxyz, [Last accessed on 29 March 2019].

[122] A. P. Bartók, script to convert ".extxyz" files [121] to *LAMMPS* [74] readable format.

[123] A. P. Bartók, Script to do thermodynamic integration according to Freitas [48] in LAMMPS [74, 48].

[124] L. B. Pártay, Program to integrate the results of Ref. [123].

[125] J. M. Polson, E. Trizac, S. Pronk, D. Frenkel, Finite-size corrections to the free energies of crystalline solids, The Journal of Chemical Physics 112 (12) (2000) 5339–5342.

[126] J. Noyes, J. Asher, O. Jones, G. Phillips, Tables of physical and chemical constants and some mathematical functions, 16th Edition, Longman Scientific & Technical, Harlow, 1995, [Originally compiled by G. W. C. Kaye and T. H. Laby. Now prepared under the direction of an editorial committee].

[127] M. Johnson, R. J. Needs, in preparation.

[128] A. Togo, I. Tanaka, SPGLIB: A software library for crystal symmetry search, arXiv preprint arxiv:1808.01590.

[129] J. Meija, T. B. Coplen, M. Berglund, W. A. Brand, P. De Bièvre, M. Gröning, N. E. Holden, J. Irrgeher, R. D. Loss, T. Walczyk, T. Prohaska, Atomic weights of the elements 2013 (IUPAC technical report), Pure and Applied Chemistry 88 (3) (2016) 265–291.

[130] D. C. Liu, J. Nocedal, On the limited memory BFGS method for large scale optimization, Mathematical Programming 45 (1–3) (1989) 503–528.

[131] J. Nocedal, Updating quasi-Newton matrices with limited storage, Mathematics of Computation 35 (151) (1980) 773–782.

[132] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, K. W. Jacobsen, The atomic simulation environment—a Python library for working with atoms, Journal of Physics: Condensed Matter 29 (27) (2017) 273002.

[133] S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. I. J. Probert, K. Refson, M. C. Payne, First principles methods using CASTEP, Zeitschrift für Kristallographie-Crystalline Materials 220 (5/6) (2005) 567–570.

[134] J. P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Physical Review Letters 77 (18) (1996) 3865–3868.

[135] J. P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple [Phys. Rev. Lett. 77, 3865 (1996)], Physical Review Letters 78 (1997) 1396.

[136] MATLAB, version 8.3.0.532 (R2014a), The MathWorks Inc., Natick, Massachusetts, 2014.

[137] pair_style meam/spline, http://lammps.sandia.gov/doc/pair_meam_spline.html, [Published online on LAMMPS webpage; accessed on 19 Sept. 2017].

[138] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool, Modelling and Simulation in Materials Science and Engineering 18 (1) (2010) 015012, ovito website: http://ovito.org/, accessed on 29 August 2016.

[139] OVITO bond-angle anlysis tool, http://www.ovito.org/manual/particles.modifiers.bond_angle_analysis.html, accessed on 29 August 2016.

[140] G. J. Ackland, A. P. Jones, Applications of local crystal structure measures in experiment and simulation, Physical Review B 73 (5) (2006) 054104.

[141] E. S. Fisher, C. J. Renken, Single-crystal elastic moduli and the hcp $\rightarrow$ bcc transformation in Ti, Zr, and Hf, Physical Review 135 (2A) (1964) A482–A494.

[142] W. Petry, A. Heiming, J. Trampenau, M. Alba, C. Herzig, H. R. Schober, G. Vogl, Phonon dispersion of the bcc phase of group-IV metals. I. bcc titanium, Physical Review B 43 (13) (1991) 10933–10947.

[143] P. W. Mirwald, G. C. Kennedy, The melting curve of gold, silver, and copper to 60-kbar pressure: A reinvestigation, Journal of Geophysical Research: Solid Earth 84 (B12) (1979) 6750–6756.

[144] M. Parrinello, A. Rahman, Strain fluctuations and elastic constants, The Journal of Chemical Physics 76 (5) (1982) 2662–2666.

[145] M. Ceriotti, G. A. Tribello, M. Parrinello, Demonstrating the transferability and the descriptive power of sketch-map, Journal of Chemical Theory and Computation 9 (3) (2013) 1521–1532.

[146] S. De, A. P. Bartók, G. Csányi, M. Ceriotti, Comparing molecules and solids across structural and alchemical space, Physical Chemistry Chemical Physics 18 (20) (2016) 13754–13769.

[147] N. Bernstein, bash script to analyse trajectory files.

[148] R. G. Hennig, T. J. Lenosky, D. R. Trinkle, S. P. Rudin, J. Wilkins, MEAM potential for titanium, http://ftp.aip.org/epaps/phys_rev_b/E-PRBMDO-78-104829/EPAPS-MEAM-Potential.dat, [last accessed 23 March 2019].

[149] M. Schlegel, Parameter file for Ti MEAM [15, 148] in *QUIP* [114], unpublished, created during the course of my MPhil in Scientific Computing at the University of Cambridge (2015).

[150] M. Schlegel, Application of an embedded atom model, unpublished, submitted as small project report for the MPhil in Scientific Computing at the University of Cambridge, Cambridge (2015).

[151] T. E. Oliphant, A guide to NumPy, Trelgol Publishing USA, 2006.

[152] S. Van Der Walt, S. C. Colbert, G. Varoquaux, The NumPy array: A structure for efficient numerical computation, Computing in Science & Engineering 13 (2) (2011) 22–30.

[153] P. Hohenberg, W. Kohn, Inhomogeneous electron gas, Physical Review 136 (3B) (1964) B864–B871.

[154] W. Kohn, L. J. Sham, Self-consistent equations including exchange and correlation effects, Physical Review 140 (4A) (1965) A1133–A1138.

[155] M. C. Payne, M. P. Teter, D. C. Allan, T. Arias, J. D. Joannopoulos, Iterative minimization techniques for ab initio total-energy calculations - molecular-dynamics and conjugate gradients, Reviews of Modern Physics 64 (4) (1992) 1045–1097.

[156] A. De Vita, The energetics of defects and impurities in metals and ionic materials from first principles., Ph.D. thesis, University of Keele (1992).

[157] H. J. Monkhorst, J. D. Pack, Special points for Brillouin-zone integrations, Physical Review B 13 (10) (1976) 5188–5192.

[158] G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Physical Review B 54 (16) (1996) 11169–11186.

[159] C. G. Broyden, A class of methods for solving nonlinear simultaneous equations, Mathematics of Computation 19 (92) (1965) 577–593.

[160] M. Gillan, Calculation of the vacancy formation energy in aluminium, Journal of Physics: Condensed Matter 1 (4) (1989) 689–711.

[161] G. P. Francis, M. C. Payne, Finite basis set corrections to total energy pseudopotential calculations, Journal of Physics: Condensed Matter 2 (19) (1990) 4395–4404.