NTNU
Norwegian University of
Science and Technology

# Simulations of Aluminum-Magnesium Alloys in a hexagonal close-packed Lattice based on the Cluster Expansion Method.

**Olve Lyngved Ødegård**

Master of Science in Physics and Mathematics
Submission date: June 2018
Supervisor: Jaakko Akola, IFY

Norwegian University of Science and Technology
Department of Physics

# Abstract

By using the density-functional theory and adopting the approach of Kohn and Sham, together with the cluster expansion method for Monte Carlo simulations, aluminum-magnesium alloys on a hcp lattice have been studied. The exchange-correlation energy functional mainly considered is the Perdew-Burke-Ernzerhof functional which belongs to the class of generalized gradient approximations. To perform calculations the open-source density-functional theory Python code named GPAW has been used. This in turn uses the projector-augmented wave method and the atomic simulation environment ASE. In conclusion, this thesis highlights the ability of the cluster expansion method combined with density-functional theory and Monte Carlo simulations to predict phases of an alloy on a given lattice. A cluster expansion model is developed for aluminum-magnesium alloys on a hcp lattice ranging in concentration from 25 % to 100 % magnesium. MC simulations reveal an $Mg_3Al$ phase emerging at temperatures below 300 K and its phase boundaries have been further studied in canonical and grand-canonical ensembles.

# Sammendrag

Ved å benytte tetthetsfunksjonal teori og følge fremgangsmåten til Kohn og Sham, sammen med cluster ekspansjons metoden og Monte Carlo simuleringer, har aluminium-magnesium legeringer arrangert i et hcp gitter blitt studert. Exchange-correlation energi funksjonalet som hovedsakelig blir brukt er Perdew-Burke-Ernzerhof funksjonalet som tilhører klassen generaliserte gradient approksimasjoner. For å gjøre beregninger benyttes den åpne kildekoden for tetthetsfunksjonal teori beregninger kalt GPAW som er skrevet i Python. Denne bruker igjen projector-augmented wave metoden og det atomistiske simuleringsverktøyet ASE. Kort oppsummert fremhever denne oppgaven cluster ekspansjons metodens evne kombinert med tetthetsfunksjonal teori og Monte Carlo simuleringer til å predikere faser av en legering i et gitt gitter. En cluster ekspansjons modell blir utviklet for aluminium-magnesium legeringer i et hcp gitter bestående av mellom 25 % og 100 % magnesium. Monte Carlo simuleringer avdekker en $Mg_3Al$ fase som dukker opp ved temperaturer under 300 K og dens fasegrenser har blitt studert i det kanoniske og semi-store kanoniske ensemblet.

# Preface

This document is a master thesis which is mandatory for all students taking a Master of Science in Applied Physics at the Norwegian University of Science and Technology (NTNU). It has been written during the spring semester of 2018, corresponds to 30 ECTS credits and can be viewed as a natural continuation of the mandatory 15 credit specialization project written during the fall semester of 2017. In the specialization project, the goal was to learn how to perform and get comfortable with density-functional theory calculations, utilizing GPAW and ASE. In this thesis, this has been used further to develop a cluster expansion model to perform simulations of aluminum-magnesium alloys on a hcp lattice. The project is the idea of my supervisor Prof. Jaakko Akola.

## Acknowledgements

First, I would like to sincerely thank my supervisor, Prof. Jaakko Akola, who has been paramount in planning and carrying out the work in this thesis. I would also like to give a special thanks to PhD Candidate David Kleiven for helping me out with performing calculations and running Monte Carlo simulations. Lastly, I thank my family, friends and girlfriend for encouragement and support throughout my studies.

# Table of Contents

# List of Tables

# List of Figures

# Abbreviations

| | | |
|---|---|---|
| CE | = | Cluster expansion |
| CV | = | Cross-validation |
| DFT | = | Density-functional theory |
| ECI | = | Effective cluster interactions |
| fcc | = | Face centered cubic |
| GGA | = | Generalized gradient approximation |
| hcp | = | Hexagonal close packed |
| HF | = | Hartree-Fock |
| KS | = | Kohn-Sham |
| LASSO | = | Least absolute shrinkage and selection operator |
| LDA | = | Local density approximation |
| MC | = | Monte Carlo |
| PBE | = | Perdew-Burke-Ernzerhof |
| SCF | = | Self-consistent field |
| xc | = | Exchange-correlation |

Chapter 1

# Introduction

In our time, metals are one of the most essential kind of materials used. This is because common metals have the necessary properties to be considered for various applications, ranging in size from constructions and large vehicles to electronics on the micro- or nano-scale. At the same time these metals can be manufactured on a large scale and at a low cost because of their natural occurrence on earth. The high demand for metals has developed the field of materials science into a huge industry always trying to improve for example the durability and strength of the different metals produced. A way of producing better materials is through alloying, combining a metal with other materials to obtain the desired properties for a specific application.

A metal of high interest for alloying is aluminum, mainly because of its light weight. In addition, recycling aluminum only requires a small percentage of the energy needed to manufacture it in the first place. By adding small concentrations of alloying elements to aluminum the properties of the resulting material can differ greatly from that of pure aluminum. Because the resulting material is mainly made up of aluminum alloying retains this light weight while being an easy way of altering the features of the material. One example of an alloying element that is frequently used is magnesium which when added can increase the strength of the alloy greatly because of its structure being hard to deform. It is also possible to create aluminum-magnesium alloys that are rich in magnesium. This is for example used in the automobile industry [1].

As explained, the field of materials science is highly focused on the

macroscopic properties of materials and their requirements in various applications. To better understand the behaviour of different materials and how to obtain their desired properties it is paramount to study the link between the macroscopic properties and physical and chemical processes at the electronic or atomic level. Not only does the chemical composition determine the macroscopic properties, the microstructure is also of great importance [2]. To study in detail the relationship between chemical composition, microstructure and macroscopic properties, computational methods are increasingly popular following the major escalation of computational power available during the last decades. The reason for this is that instead of employing a method of trial and error for developing a material in the lab, simulations can predict the outcome while saving both time and resources. The computational methods for this purpose rely mostly on density-functional theory (DFT) which is developing further and further as a way of predicting properties of materials and has thus become a very popular method [3].

The complexity of DFT calculations increase with the size of the system and enormous computational resources have to be available to obtain results that converge as the system size is increased. Moreover, the concentration-temperature phase diagram of an alloy is often desired. To obtain this the configurational space has to be sampled in great detail which can not be done by only applying DFT calculations, as this is too computationally expensive. Computational methods such as Monte Carlo (MC) or molecular dynamics simulations are thus needed.

To go from ab initio methods to simulations on atomic systems of considerable size the cluster expansion (CE) method is often employed, in effect linking DFT calculations together with MC simulations. This is done by first obtaining ground state energies for a given set of configurations using DFT. The energy of a configuration can then be expanded in different clusters of atoms and fitted to the calculated ground state energies with effective cluster interactions (ECI) as expansion coefficients. Such a CE model can be applied in large MC simulations of atomic systems as energies can be calculated more efficiently compared to DFT calculations of the same, large systems. Different characteristics of a material can then be inferred from the resulting energy calculations [4], such as phase diagrams of different types of alloys [5] [6].

In this thesis, a CE model is developed to combine initial DFT calculations with MC simulations to study aluminum-magnesium alloys. The

alloys studied here are rich in magnesium and have a crystal structure based on the hcp lattice. To begin with, chapter 2 gives the theoretical background for crystalline materials. In addition, it introduces several important concepts for the calculations such as metallic bonding along with terminology. Chapter 3 and 4 highlight the most relevant aspects of the theory behind DFT calculations and the CE method, respectively. In chapter 5 the computational methods are presented while the results and discussions are presented in chapter 6. In chapter 7 conclusions are drawn before the thesis ends with suggestions to future work in chapter 8.

Chapter 2

# Crystalline materials

In this chapter the theory behind different properties of crystal systems will be given. The focus will be on aluminum and magnesium starting out with their crystal structure as these are the materials studied in the thesis. The chapter is based on reference [7] which will function as the main reference for the entire chapter.

## 2.1 The structure of crystals

An ideal crystal is made by repeating a group of atoms infinitely, each group the same distance apart. The points in space where the groups are placed are called the lattice of the crystal whereas the group is called the basis. A three-dimensional lattice is defined by three vectors of translation, $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$, such that the crystal looks the same seen from the points $\mathbf{r}'$ and $\mathbf{r}$ defined by

$$\mathbf{r}' = \mathbf{r} + n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3, \tag{2.1}$$

where $n_1$, $n_2$ and $n_3$ are integers. In this way (2.1) defines the lattice. The crystal axes are defined as a suitable choice of translation vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ for the crystal, depending on its structure. Having chosen this set the position of the atoms in the basis can be defined for each lattice point given by

$$\mathbf{r}_i = x_i\mathbf{a}_1 + y_i\mathbf{a}_2 + z_i\mathbf{a}_3, \tag{2.2}$$

where $i$ can be any integer indicating that it is possible to have several atoms in the basis. The three crystal axes can be used to define a unit cell of the crystal with a volume given by

$$V_c = |\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3|. \tag{2.3}$$

The unit cell is called a primitive cell if it is the smallest cell possible to construct, containing a single lattice point. To the single lattice point several atoms can be attached at different positions, allowing more than one atom in the primitive cell. Moreover, the length of the edges of the unit cell, $a$, $b$ and $c$ not parallel to each other are denoted the lattice constants.

### 2.1.1 Face-centered cubic lattice for aluminum

The arrangement of the atoms in aluminum at room temperature and below is a face-centered cubic (fcc) lattice. This is recognized as a cubic structure with lattice points placed at each corner as well as in the middle of all the facets of the cube, giving rise to a total of 4 lattice points per unit cell. With a lattice constant at room temperature of $a = b = c = 4.05$ Å [7, p. 20] the unit cell of the aluminum structure is given by

$$\mathbf{a}_1 = (\frac{a}{2}, \frac{a}{2}, 0), \quad \mathbf{a}_2 = (\frac{a}{2}, 0, \frac{a}{2}) \quad \text{and} \quad \mathbf{a}_3 = (0, \frac{a}{2}, \frac{a}{2}), \tag{2.4}$$

while the basis consists of a single aluminum atom located at the origin.

### 2.1.2 Hexagonal close packed lattice for magnesium

The arrangement of the atoms in magnesium at room temperature and below is a hexagonal close packed (hcp) lattice. This is recognized as a hexagonal prism with lattice points placed at each corner. In addition, there is a lattice point in the center of each of the two hexagonal facets of the structure so that there are in total two lattice points per unit cell. The lattice constants of the magnesium structure at room temperature are $a = b = 3.21$ Å and $c = 5.21$ Å [7, p. 20] and the unit cell is given by

$$\mathbf{a}_1 = (\frac{a}{2}, -\frac{\sqrt{3}a}{2}, 0), \quad \mathbf{a}_2 = (\frac{a}{2}, \frac{\sqrt{3}a}{2}, 0) \quad \text{and} \quad \mathbf{a}_3 = (0, 0, c), \tag{2.5}$$

with a basis of two magnesium atoms, the first one at the origin and the second one located at $\mathbf{r}_2 = \frac{2}{3}\mathbf{a}_1 + \frac{1}{3}\mathbf{a}_2 + \frac{1}{2}\mathbf{a}_3$.

### 2.1.3 Reciprocal space

Because of the periodicity of the crystal lattice the different properties of the system will follow the same periodicity. An example is the electron density $n(\mathbf{r})$ which because the crystal is invariant under a translation by $\mathbf{T} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$ satisfies the relation

$$n(\mathbf{r}) = n(\mathbf{r} + \mathbf{T}). \tag{2.6}$$

This periodicity makes Fourier analysis a very powerful tool and it becomes useful to write the Fourier expansion

$$n(\mathbf{r}) = \sum_p n_p \exp(i\mathbf{G}_p \cdot \mathbf{r}), \tag{2.7}$$

where $p$ runs over all integers and the coefficients $n_p$ are complex numbers satisfying $n^*_{-p} = n_p$ to keep $n(\mathbf{r})$ real. The vectors $\mathbf{G}_p$ must then be a set of vectors such that (2.7) is invariant under all translations $\mathbf{T}$. This set can be constructed by the vectors of the reciprocal lattice, defined as

$$\begin{aligned} \mathbf{b}_1 &= 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3}, \quad \mathbf{b}_2 = 2\pi \frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3} \quad \text{and} \\ \mathbf{b}_3 &= 2\pi \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3}, \end{aligned} \tag{2.8}$$

where $\mathbf{b}_1, \mathbf{b}_2$ and $\mathbf{b}_3$ are primitive vectors of the reciprocal lattice if $\mathbf{a}_1, \mathbf{a}_2$ and $\mathbf{a}_3$ are primitive vectors for the lattice of the crystal. Also, the reciprocal vectors have the property

$$\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi\delta_{ij}, \tag{2.9}$$

where $\delta_{ij}$ is the Kronecker-delta symbol. Moreover, the reciprocal lattice is the Fourier transform of the crystal lattice. The set of vectors $\mathbf{G}$ and reciprocal lattice points are then given by the reciprocal lattice vectors

$$\mathbf{G} = m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_2\mathbf{b}_2, \tag{2.10}$$

where $m_1$, $m_2$ and $m_3$ are integers.

In reciprocal lattice the Brillouin zone is defined as the volume in reciprocal space around a reciprocal lattice point where the associated region in $\mathbf{k}$-space is closer to that point than any other reciprocal lattice point. An

equivalent definition exists for the crystal lattice in real space where the corresponding volume is called the Wigner-Zeits cell. The Brillouin zone for crystal structures has some critical points with high symmetry. These points are often of interest and can be summarized for a fcc and hexagonal lattice as follows:

**Table 2.1:** Critical points in the Brillouin zone of a fcc lattice.

| Symbol | Description |
|---|---|
| Γ | Center of Brillouin zone |
| K | Midpoint of an edge between two hexagonal surfaces |
| L | Center of a hexagonal surface |
| U | Midpoint of an edge between a square and a hexagonal surface |
| W | Corner point |
| X | Center of a square surface |

**Table 2.2:** Critical points in the Brillouin zone of a hexagonal lattice.

| Symbol | Description |
|---|---|
| Γ | Center of Brillouin zone |
| A | Center of a hexagonal surface |
| H | Corner point |
| K | Midpoint of an edge between two rectangular surfaces |
| L | Midpoint of an edge between a rectangular and a hexagonal surface |
| M | Center of a rectangular surface |

## 2.2 Electronic properties

Many of a materials properties stem from how its electrons behave. An example of this can be if the material is classified as a conductor or an insulator [7, p. 181]. Some important theoretical aspects of these electronic properties will now be presented.

### 2.2.1 The Fermi level

From the exclusion principle of Pauli [8] it is possible to have a maximum of two electrons in the same state, one with spin up and one with spin down. At temperatures approaching $T$ = 0 K the electrons will occupy the states with the lowest possible energy. At absolute zero for a non-interacting system of electrons the Fermi energy is defined as the kinetic energy of the highest occupied state. Increasing the temperature will add thermal energy to the electrons, giving them the possibility of reaching higher energy states. The Fermi-Dirac distribution $f(\epsilon)$, given by

$$f(\epsilon) = \frac{1}{e^{(\epsilon-\mu)/k_{\mathrm{B}}T} + 1} \tag{2.11}$$

then gives the probability of having an occupied state at the energy $\epsilon$, at thermal equilibrium. Here, $k_{\mathrm{B}}$ is the Boltzmann's constant and $\mu$ is the Fermi level, defined as the chemical potential for the electrons. The Fermi level can also be considered as an energy state for an electron so that the probability of finding an electron in this energy state, $\epsilon = \mu$, is fifty percent. In contrast to the Fermi energy, the Fermi level is defined for all temperatures and interacting systems as well.

### 2.2.2 Bloch-Waves

Because of the periodicity of crystal structures the potential that the electrons experience can be modeled as a periodic potential stemming from the positively charged ions, with the same periodicity as the crystal lattice. This potential can be expanded using a Fourier series written as

$$V(\mathbf{r}) = \sum_{\mathbf{G}} V_{\mathbf{G}} \exp(i\mathbf{G} \cdot \mathbf{r}), \tag{2.12}$$

where $\mathbf{G}$ is given by (2.10) and $V_{\mathbf{G}}$ are expansion coefficients. By solving the Schrödinger equation for a single electron,

$$\left[ -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}), \tag{2.13}$$

F. Bloch showed [7, p. 167] that the solutions must be on the form

$$\psi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r}) \exp(i\mathbf{k} \cdot \mathbf{r}), \tag{2.14}$$

called a Bloch wave, and that each of these solutions is an energy eigenstate. Here, $n$ denotes the energy state, $\mathbf{k}$ is the wave vector and $u_{n\mathbf{k}}(\mathbf{r})$ has the same periodicity as the crystal.

As shown, for each value of $\mathbf{k}$ there exists multiple energy states, denoted by $n$ and often called the band index. Each of these energy states have energies that vary smoothly with varying $\mathbf{k}$ which can be used to obtain a dispersion relation $E_n(\mathbf{k})$ for the electrons in a band $n$. The wave vectors $\mathbf{k}$ can be any point in the Brillouin zone. As the Brillouin zone is three-dimensional it is common to show $E_n(\mathbf{k})$ for values of $\mathbf{k}$ at straight lines between the critical points of the Brillouin zone. Representations like these are used to illustrate the electronic band structure of a crystal. The electronic band structure is among other things used to classify different materials. For metals the Fermi level is located inside an electronic band, while for semiconductors and insulators it lies inside a band gap where no electronic states occur.

### 2.2.3 Metallic bonding

For elements with one, two or three valence electrons the interaction between the nucleus and these valence electrons becomes weak. Putting together many such atoms in a bulk the valence electrons will become dislocated from their corresponding atom because of the weak bonding and contribute to a common cloud of electrons [9]. Losing their valence electrons the atoms end up being positively charged cations held together by the negatively charged electron cloud. Both aluminum and magnesium, with respectively three and two valence electrons, form metals through such a metallic bonding. Moreover, it is this bonding that orders aluminum and magnesium in their respective fcc and hcp structures and determines how much energy it takes to form the lattices. It should be noted that both aluminum and magnesium are simple metals in terms of having no semi-core states contributing to the valence electrons.

#### Cohesive energy

The cohesive energy of a system of atoms is defined as the energy gained by arranging the atoms in its crystal structure compared to keeping the atoms far apart. It is a measure of the strength of the bonds and thereby often

called the binding energy and can be given for a structure of $N$ atoms as

$$E_c = N \cdot E_t^s - E_b. \tag{2.15}$$

Here, $E_c$ denotes the cohesive energy, $E_t^s$ is the energy of an individual atom of type $t$ and $E_b$ is the total energy of the bulk of atoms. It is important to note that the energy $E_t^s$ is calculated for the single atom when it is separated far from any other atom and not a part of the bulk, denoted by the superscript $s$. The cohesive energy is often given per atom, denoted $\epsilon_c$, which for a bulk containing $N$ aluminum and $M$ magnesium atoms becomes

$$\epsilon_c = \frac{E_c}{N+M} = \frac{N \cdot E_{Al}^s + M \cdot E_{Mg}^s - E_b}{N+M}. \tag{2.16}$$

Here, $E_{Al}^s$ and $E_{Mg}^s$ denotes the energy of a single aluminum and magnesium atom, again kept far apart from any other atom.

### Formation energy of a vacancy

In addition to the cohesive energy the metallic bonding determines the formation energy of a vacancy, $E_{vf}$. This can be defined as the amount of energy needed to form a vacancy in a bulk of atoms and thought of as removing one of the atoms and breaking the bonds to its nearest neighbours in the process. The total energy of the bulk containing $N$ atoms is as mentioned denoted $E_b$ and can be given as

$$E_b = \sum_{i=1}^{N} E_i^b = N E_a^b. \tag{2.17}$$

Here, $E_i^b$ is the energy of a single atom when being a part of the bulk and $E_a^b$ denotes the average energy of a single atom in the bulk. In a similar manner $E_v$ denotes the total energy of the same bulk but with one of the atoms replaced with a vacancy, given by

$$E_v = \sum_{i=1}^{N-1} E_i^v = (N-1) E_a^v, \tag{2.18}$$

summing up the energy of each of the $N-1$ atoms given by $E_i^v$. Here, $E_a^v$ is the average energy of a single atom in the bulk with the vacancy. The

formation energy of a vacancy is then defined as

$$E_{vf} = (N-1)(E_a^v - E_a^b) = E_v - \frac{N-1}{N} E_b, \tag{2.19}$$

the difference in average energy per atom for the bulk with and without the vacancy times the number of atoms in the bulk with the vacancy.

Chapter 3

# The theory behind density-functional theory calculations

In this chapter the main ideas and a short theoretical background of DFT will be presented. Before this, however, adopting the idea of reference [10], the fundamental Schrödinger equation and the Hartree-Fock (HF) method is introduced. The main references for this chapter are references [11] and [12].

## 3.1 The Schrödinger equation

In electronic many-body structure calculations the Born-Oppenheimer approximation [13] is used. In this approximation the nuclei of the structure are assumed to be at fixed positions, giving rise to a static potential for the electrons. A many-body electronic wave function $\psi(\mathbf{r}_1, \ ... \ , \mathbf{r}_N)$ is then used to describe a stationary electronic state. The starting point of DFT calculations is the non-relativistic, time-independent many-electron Schrödinger equation [14] given by

$$\hat{H}\psi = \left[ -\frac{\hbar}{2m} \sum_{i=1}^{N} \nabla_i^2 + \sum_{i=1}^{N} V(\mathbf{r}_i) + \sum_{i=1}^{N} \sum_{j<i} U(\mathbf{r}_i, \mathbf{r}_j) \right] \psi = E\psi \quad (3.1)$$

for a system of $N$ electrons which the electronic wave function $\psi$ must satisfy. In this equation, $\hat{H}$ is the Hamiltonian, $\hbar$ is the reduced Planck constant or Dirac constant and $m$ is the mass of an electron. Moreover, $V(\mathbf{r}_i)$ is the potential stemming from the interaction between the atomic nuclei and an electron at position $\mathbf{r}_i$ and $U(\mathbf{r}_i, \mathbf{r}_j)$ is the potential resulting from the interaction between two electrons located at positions $\mathbf{r}_i$ and $\mathbf{r}_j$. $E$ denotes the energy of the ground state of the system of electrons. Important to note when later considering DFT is that the electronic wave function $\psi$ is depending on three coordinates for each electron, giving in total $3N$ variables for a system of $N$ electrons and making the exact solution of (3.1) computationally expensive for increasing $N$.

## 3.2 Hartree-Fock method

A strategy to reduce the complexity of finding the correct many-body wave function is to approximate it by what is known as the Hartree product [15]. The approximation starts out by considering wave functions for the individual electrons given by $\chi(\mathbf{x})$ and referred to as a spin-orbital. Here, $\mathbf{x}$ denotes the coordinates of the electron including both the spatial coordinates $\mathbf{r}$ and spin coordinate $s$ and the spin-orbitals are assumed to be orthonormal. The Hartree product is then defined as the product of the individual electron spin-orbitals,

$$\psi(\mathbf{x}_1, \mathbf{x}_2, \ \dots, \mathbf{x}_\mathrm{N}) = \chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2)...\chi_\mathrm{N}(\mathbf{x}_\mathrm{N}). \tag{3.2}$$

Although (3.2) reduces the complexity of the problem drastically, the Hartree product violates the Pauli exclusion principle and is thus not expected to give correct results. The violation is due to the fact that the expression is not antisymmetric when interchanging two of the electrons. This must be the case when considering fermions [8] as all states must be antisymmetric, which is an alternative way of expressing the exclusion principle. Formulated mathematically all wave functions must satisfy

$$\psi(\mathbf{x}_1, \mathbf{x}_2, \ \dots, \mathbf{x}_\mathrm{N}) = -\psi(\mathbf{x}_2, \mathbf{x}_1, \ \dots, \mathbf{x}_\mathrm{N}). \tag{3.3}$$

The obvious fault of the Hartree product not satisfying (3.3) can be fixed by taking the linear combination of both sides of the equation, giving

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}} \left[\chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2) - \chi_1(\mathbf{x}_2)\chi_2(\mathbf{x}_1)\right] \tag{3.4}$$

when considering a system of only two electrons. Generalized to a system of $N$ electrons it reads

$$\psi(\mathbf{x}_1, \mathbf{x}_2, \ \dots\ , \mathbf{x}_\mathrm{N}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \dots & \chi_\mathrm{N}(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \dots & \chi_\mathrm{N}(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_\mathrm{N}) & \chi_2(\mathbf{x}_\mathrm{N}) & \dots & \chi_\mathrm{N}(\mathbf{x}_\mathrm{N}) \end{vmatrix}, \tag{3.5}$$

which is known as the Slater determinant [16] and is the approximation of the electronic wave function in the HF method [17]. The advantage of this over the Hartree product is that it in addition to only considering spin-orbitals satisfies (3.3). Because of this it is expected to better describe the many-body electron system. By following this approximation (3.1) can be rewritten as a set of equations for the single-particle states [12, p. 71], given by

$$\left[ -\frac{\hbar}{2m}\nabla^2 + V(\mathbf{r}) + V_H(\mathbf{r}) - V_{Ex}(\mathbf{r}) \right] \chi_i(\mathbf{x}) = \epsilon_i \chi_i(\mathbf{x}) \tag{3.6}$$

and commonly referred to as the HF equations. Here, $\epsilon$ denotes the energy of the spin-orbital.

The potential from the interaction between the electrons is evidently replaced by a Hartree potential $V_H$ and an exchange term $V_{Ex}$. The Hartree potential is given by

$$V_H(\mathbf{r}) = e^2 \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathrm{d}\mathbf{r}', \tag{3.7}$$

where $e$ is the charge of an electron and $n(\mathbf{r}')$ is the electron density, defined by

$$n(\mathbf{r}) = \sum_{i=1}^{N} \chi_i^*(\mathbf{r}) \chi_i(\mathbf{r}) \tag{3.8}$$

where the explicit spin dependence has been omitted. The Hartree potential $V_H$ is the average Coulomb repulsion that a particle with charge $e$ would feel from the $N$ electrons in the system. It should be stressed that the electrons are assumed to move independently and thus, $V_H$ depends on the average position of the electrons. This averaged Coulomb repulsion, and not the instantaneous forces between the electrons, is what affects the electrons via

$V_H$. Another important point is the number of electrons which $V_H$ depends on, namely $N$. One might expect it only to depend on the average of $N-1$ electron charges as the electron does not interact with itself. This unphysical part of the Hartree potential is however eliminated by a corresponding part in $V_{Ex}$ [18, p. 88].

Regarding the exchange term one can note that it has no simple, physical interpretation. It appears solely because the Slater determinant is used as the approximation for the wave function. It has its name from the antisymmetric requirement of the wave function that one can exchange two electrons and only change the sign of the wave function. No further mathematical discussion of the exchange term will be given. An approximation for the spin-orbitals is needed to numerically solve (3.6). By expanding on a suitable and finite basis set of $M$ functions, $\{\phi\}$, with expansion coefficients $c$, the spin orbitals can be approximated by a linear combination

$$\chi_i(\mathbf{x}) = \sum_{j=1}^{M} c_{i,j}\phi_j(\mathbf{x}). \tag{3.9}$$

The numerical approach of the HF method is first to guess the form of the spin-orbitals. After this, (3.6) is solved iteratively until the spin orbitals are reproduced within an acceptable convergence criterion. Because of this way of approaching the solution, by iterating until the solution reproduces itself from the last step, the method belongs to what is called the self-consistent field (SCF) methods. As the HF method is not exact the energies obtained from the calculations are not expected to be correct and the difference between the exact ground-state energy and the HF ground-state energy is called the correlation energy [12, p. 70]. The name comes from the fact that the HF method assumes uncorrelated electrons. Ignoring the correlation can give results for the energy that are too low, leading to wrong results as for example a vanishing single-electron density of states near the Fermi level for a uniform electron liquid [12, p. 83]. Because of this and that the HF method can be computationally expensive for relatively large systems the focus should be turned towards a better approach. In other words, the stage is now set for DFT.

## 3.3 Density functional theory

The main idea behind DFT is to use the density of the electrons instead of the wave function to describe the system under consideration, hence the name of the theory. One of the advantages for numerical calculations is that this reduces the complexity of the system substantially since the electron density is a single three-dimensional object no matter what the number of electrons is. This chapter is largely based on reference [12] and will give a short theoretical introduction to the theory behind DFT calculations.

### 3.3.1 A variational principle for the density

A step of great importance in DFT is to obtain a variational principle for the density which is best shown with the Levy-Lieb constrained search algorithm [19]. The well-known Rayleigh-Ritz variational principle [20, p. 31] for the many-body wave function finds an upper energy-bound, $E$, for the energy of the ground-state, $E_0$, by searching for the lowest possible expectation value of $\hat{H}$ when altering $\psi$, written as

$$E = \min_{\psi} \langle \psi | \hat{H} | \psi \rangle \,. \tag{3.10}$$

The system considered is a $N$-electron system where the Hamiltonian is written as $\hat{H} = \hat{T} + \hat{H}_{e-e} + \hat{V}$. Here, $\hat{T}$ is the kinetic energy, $\hat{H}_{e-e}$ is the interaction between the electrons and $\hat{V}$ is the potential energy. The last term which stems from a local, external potential $V(\mathbf{r})$ can be written

$$\hat{V} = \int V(\mathbf{r}) \hat{n}(\mathbf{r}) \mathrm{d}\mathbf{r}, \tag{3.11}$$

where $\hat{n}(\mathbf{r})$ is the density operator. To formulate a variational principle for the density the first step is to choose a density $n(\mathbf{r})$ and minimize the energy via (3.10) for antisymmetric wave functions $\psi$ that give the same, chosen density, resulting in

$$E_V[n] = \min_{\psi \to n(\mathbf{r})} \langle \psi | \hat{T} + \hat{H}_{e-e} | \psi \rangle + \int V(\mathbf{r}) n(\mathbf{r}) \mathrm{d}\mathbf{r}. \tag{3.12}$$

The solution of (3.12) defines the functionals $E_V[n]$ and $F[n]$ where the last functional denotes the first term on the right hand side of the equation.

The second and last step of the algorithm is to also minimize with respect to the density, giving the required variational principle as

$$E = \min_{n(\mathbf{r})} \left[ F[n] + \int V(\mathbf{r}) n(\mathbf{r}) \mathrm{d}\mathbf{r} \right]. \tag{3.13}$$

If now for any positive number $\eta$ and variation in the density $\delta n(\mathbf{r})$,

$$\lim_{\eta \to 0} \frac{F[n + \delta n] - F[n]}{\eta} = \int D(\mathbf{r}) \delta n(\mathbf{r}) \mathrm{d}\mathbf{r} \tag{3.14}$$

is true, the functional derivative of $F[n]$ with respect to the density is $D(\mathbf{r})$. As $D(\mathbf{r})$ is independent of $\delta n(\mathbf{r})$ it is itself a functional of the density and can be written

$$D[n] = \frac{\delta F[n]}{\delta n(\mathbf{r})}. \tag{3.15}$$

Assuming the existence of $D[n]$, a small variation of $n(\mathbf{r})$ will give a variation of $E_V[n]$ as

$$E_V[n + \eta \delta n] - E_V[n] = \eta \int \left[ \frac{\delta F[n]}{\delta n(\mathbf{r})} + V(\mathbf{r}) \right] \delta n(\mathbf{r}) \mathrm{d}\mathbf{r} + \mathcal{O}(\eta^2). \tag{3.16}$$

As seen from (3.13) the first term on the right hand side of (3.16) vanishes for the density of the ground-state and this density must therefore satisfy

$$\frac{\delta F[n]}{\delta n(\mathbf{r})} = -V(\mathbf{r}), \tag{3.17}$$

showing that the functional derivative of $F[n]$ yields the potential $V(\mathbf{r})$.

### 3.3.2 The Hohenberg-Kohn theorems

The Hohenberg-Kohn theorems are fundamental to the starting point of DFT showing that, in principle, all properties of the ground-state of a many-body system can be determined by the ground-state electron density [21]. The two theorems concern any system of electrons moving in an external potential and can be given as follows [21]:

> **Theorem 1 -** *The external potential and hence the total energy, is a unique functional of the electron density.*

**Theorem 2 -** *The functional that gives the ground-state energy of the system, gives the lowest energy if and only if the input density is the true ground state density.*

In short, these theorems state a one-to-one correspondence between the electron density and the many-body wavefunction of the ground-state and a way of finding this density via a variational principle.

Since the ground state density in principle determines all properties of the system (3.13) and (3.17), taking the Hohenberg-Kohn theorems into consideration, provide a way to calculate them. However, the exact functional $F[n]$ is not known. It could be approximated entirely, but in some cases, one being that of the non-interacting electron system, the calculations can be done within a greatly simplified framework. The desired approach is thus the one that will give the correct results for the non-interacting system no matter which approximations are applied later.

### 3.3.3 The Kohn-Sham equations

The approach based on non-interacting electrons was established by Kohn and Sham (KS) in 1965 [22]. The first step is to use the density of the ground-state of a non-interacting system in some external, local potential $V_{KS}(\mathbf{r})$ to represent the density of the ground state of the interacting system. After this, the density is found by minimizing the energy-functional of the non-interacting system, denoted $E^{(0)}_{V_{KS}}[n]$, using a superscript $(0)$ which hereafter indicates that it is the ground-state. The minimization is given by

$$E^{(0)}_{V_{KS}}[n] = \min_{\psi \to n(\mathbf{r})} \langle \psi | \hat{T} | \psi \rangle + \int V_{KS}(\mathbf{r}) n(\mathbf{r}) \mathrm{d}\mathbf{r}, \tag{3.18}$$

where the first term on the right hand side, denoted $T_s[n]$, is the kinetic energy of the non-interacting system with ground-state density $n(\mathbf{r})$. This density must from (3.17) satisfy

$$\frac{\delta T_s[n]}{\delta n(\mathbf{r})} = -V_{KS}(\mathbf{r}). \tag{3.19}$$

The next step is to decompose the unknown functional $F[n]$ as

$$F[n] = T_s[n] + E_H[n] + E_{xc}[n], \tag{3.20}$$

introducing the Hartree energy functional

$$E_H[n] = \frac{e^2}{2} \int \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r} \tag{3.21}$$

and the exchange-correlation (xc) energy functional, denoted $E_{xc}[n]$ and defined by (3.20). By yet again using (3.17) one obtains

$$\frac{\delta T_s[n]}{\delta n(\mathbf{r})} = -V(\mathbf{r}) - V_H(\mathbf{r}) - V_{xc}(\mathbf{r}). \tag{3.22}$$

Here, $V_H(\mathbf{r})$ is the Hartree potential and $V_{xc}(\mathbf{r})$ is an xc potential given by

$$V_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})}. \tag{3.23}$$

Equation (3.22) can now be used together with (3.19) to give the KS potential $V_{KS}(\mathbf{r})$ as

$$V_{KS}(\mathbf{r}) = V(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}). \tag{3.24}$$

Since minimizing $E^{(0)}_{V_{KS}}[n]$ by definition of (3.18) corresponds to finding the density of the ground-state of the non-interacting system in $V_{KS}(\mathbf{r})$, this density is calculated by solving the KS equations given by

$$\left[ -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) \right] \phi_\alpha(\mathbf{x}) = \epsilon_\alpha \phi_\alpha(\mathbf{x}). \tag{3.25}$$

Here, $\alpha$ ranges from 1 to $N$ and $\phi_\alpha$ are known as the KS orbitals. It should be noted that this is an eigenvalue equation and that these orbitals do not have any formal physical interpretation, but still reflect the properties of real single-particle states remarkably well. They can be used to obtain the density of the ground-state as

$$n(\mathbf{r}) = \sum_{\alpha=1}^{N} \sum_{s} |\phi_\alpha(\mathbf{r}, s)|^2. \tag{3.26}$$

There are several advantages by taking this approach. The first is that, as earlier mentioned as a desired property, (3.25) gives correct results for a non-interacting system. The second advantage is that the kinetic energy $T_s[n]$ and the Hartree energy $E_H[n]$ can be calculated without much difficulty. Thus, only the term containing the xc energy, $E_{xc}[n]$, needs to

be approximated to calculate the functional $F[n]$ along with its functional derivative $V_{xc}(\mathbf{r})$. Lastly, the single-particle KS equations from (3.25) are less complicated to solve than the Schrödinger equation (3.1) for a system of many electrons.

Equations (3.25) are solved iteratively after making an approximation for the xc potential $V_{xc}(\mathbf{r})$. The first step is to make a guess for the density $n(\mathbf{r})$ which determines $V_{xc}(\mathbf{r})$. Then, (3.25) is solved using (3.26) to obtain a new density which can be used as a better guess and to recalculate the approximation of the xc potential. This cycle is repeated until the density is self-consistent, in effect reproduced within an acceptable convergence criterion. The final density is used to calculate ground-state properties. The fundamental property is the ground-state energy and to compute this it is important to note that the non-interacting functional for the kinetic energy is by construction given as

$$T_s[n] = \sum_{\alpha=1}^{N} \epsilon_\alpha - \int V_{KS}(\mathbf{r}) n(\mathbf{r}) \mathrm{d}\mathbf{r}. \tag{3.27}$$

By inserting this expression for $F[n]$ into (3.20) and using the result in the variational principle from (3.12) the expression for the ground-state energy becomes

$$E = \sum_{\alpha=1}^{N} \epsilon_\alpha - \frac{e^2}{2} \int\int \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathrm{d}\mathbf{r}' \mathrm{d}\mathbf{r} - \int n(\mathbf{r}) V_{xc}(\mathbf{r}) \mathrm{d}\mathbf{r} + E_{xc}[n]. \tag{3.28}$$

## 3.4 Approximating the exchange-correlation functional

If the xc functional was known exactly the KS approach would give exact results. However, as it still remains unknown the functional has to be approximated. There are several ways of doing this and this is a crucial step in terms of the correctness of DFT calculations. In this chapter two different types of approximations will be presented, known as the local-density approximation (LDA) and the generalized gradient approximation (GGA).

### 3.4.1 The local-density approximation

LDA denotes approximations to the xc energy functional that only depend on the electron density by assuming a slowly varying density across space. By this, approximations which for example also include derivatives of the density are excluded. LDA makes use of the fact that the form of the xc functional is known for the uniform electron gas. The real, inhomogeneous system under consideration is divided into small volumes of infinitesimal size, treating the electron density as a constant in each of these volumes. Then, the xc energy for the uniform electron gas at this density is used as an approximation to the true xc energy for each volume. This leads to an expression for the total exchange-correlation energy as [23]

$$E_{xc}^{LDA}[n] = \int n(\mathbf{r})\epsilon_{xc}(n(\mathbf{r}))\mathrm{d}\mathbf{r}. \tag{3.29}$$

Here, $\epsilon_{xc}(n(\mathbf{r}))$ is the xc energy of the uniform, interacting electron gas at density $n(\mathbf{r})$, per particle.

Real systems can have large variations in the electron density and it is largely these variations that give materials their different properties, implying that LDA is of little value. However, the approximation can give reasonable results even with a rapidly varying density. One example is the separation distance between atoms at equilibrium which LDA reproduces with only a few percent of error [24]. Still, LDA is quite a naive approximation and to get better results for various types of calculations the generalized gradient approximation is necessary.

### 3.4.2 The generalized gradient approximation

The difference between LDA and GGA is that the latter includes information not only about the local density, but also about its local gradient. Containing more information this in some cases leads to better results compared to LDA, with slightly increased computational cost. There are many ways to include information about the gradient of the electron density, but all methods can be written as [12, p. 355]

$$E_{xc}^{GGA}[n] = \int f(n(\mathbf{r}), \nabla n(\mathbf{r}))\mathrm{d}\mathbf{r}. \tag{3.30}$$

Several forms of the function $f$ which depends on both the electron density $n(\mathbf{r})$ as well as its gradient $\nabla n(\mathbf{r})$ can give good results for properties of

different systems. Examples include total energies, excitation and ionization potentials where GGAs can be said to be superior to LDA [25].

The xc energy functional can be separated,

$$E_{xc}^{GGA}[n] = E_{x}^{GGA}[n] + E_{c}^{GGA}[n], \tag{3.31}$$

with the terms on the right hand side denoting the exchange and correlation part, respectively. With this separation the exchange part can be written as

$$E_{x}^{GGA}[n] = \int n(\mathbf{r})\epsilon_{x}^{LDA}(n(\mathbf{r}))F_{x}(s)\mathrm{d}\mathbf{r}. \tag{3.32}$$

Here, $\epsilon_{x}^{LDA}(n(\mathbf{r}))$ is the same as for LDA when separating into exchange and correlation parts for $\epsilon_{xc}(n(\mathbf{r}))$ in (3.29). The function $F_{x}(s)$ is called the enhancement factor over the LDA exchange. It is a function of the dimensionless, reduced gradient of the density, $s$, given as [26]

$$s = \frac{|\nabla n(\mathbf{r})|}{(24\pi^2)^{1/3}n(\mathbf{r})^{4/3}}. \tag{3.33}$$

The other part of the separation of $E_{xc}^{GGA}$, namely $E_{c}^{GGA}$, is very complex and not as easily presented. The forms of $E_{c}^{GGA}$ and the enhancement factor $F_{x}(s)$ vary for the different approximations belonging to the class of GGAs. One example of a popular GGA is the BLYP functional which has its name from the exchange part developed by Becke [27] and the correlation part developed by Lee, Yang and Parr [28]. Another example is the Perdew-Burke-Ernzerhof (PBE) functional [29]. This functional is used extensively in this thesis because calculations with this functional in general give good results for a wide range of physical systems [30]. For brevity, the functionals will not be given here and the references should be consulted for more details.

Chapter 4

# The cluster expansion method

The CE method will be presented in this chapter starting out with the theoretical framework based on reference [31] and [32]. After presenting the theory, a way of evaluating a given cluster expansion is discussed. More details about its implementation will be given in chapter 5 which concerns computational methods.

## 4.1 Theoretical framework

Consider a binary, crystalline material containing $N$ lattice points in total. To each point $p$ in the lattice a discrete configurational variable $\sigma$ is assigned, which take values $1$ or $-1$ for a binary system depending on the type of atom located at $p$. Thus, any configuration of the system can be described by the vector

$$\boldsymbol{\sigma} = \{\sigma_1, \sigma_2, ..., \sigma_{\mathrm{N}}\}. \tag{4.1}$$

Consider now a point cluster within the system containing only one lattice point, making the configurational space one-dimensional. Given two functions depending on the configuration of the point cluster, $f(\sigma)$ and $g(\sigma)$, their inner product in the configurational space is defined as

$$\langle f, g \rangle = \frac{1}{2} \sum_{\sigma=\pm 1} f(\sigma) g(\sigma). \tag{4.2}$$

The configurational variable $\sigma$ spans the one-dimensional space and an orthonormal and complete basis for this space can be constructed to represent the different possibilities at each lattice point of a binary system. This basis is given by two polynomials $\varphi_0(\sigma)$ and $\varphi_1(\sigma)$ as

$$\varphi_0(\sigma) = 1 \tag{4.3a}$$
$$\varphi_1(\sigma) = \sigma, \tag{4.3b}$$

where the subscript denotes the order of the polynomial. From these an orthogonal basis can be constructed for a cluster of any given size. This is done by noting that for a cluster $\alpha = \{p_1, p_2, ..., p_{n_\alpha}\}$, where $n_\alpha$ denotes the number of lattice points in the cluster, its configurational space corresponds to the product of all its subspaces [32]. The basis for a cluster $\alpha$ is thus given by

$$\Phi_\alpha(\boldsymbol{\sigma}_\alpha) = \prod_{i \in \alpha} \varphi_1(\sigma_i), \tag{4.4}$$

containing all possible products of the polynomials in (4.3). This is called the characteristic function for the cluster and its orthogonality follows from the orthogonality of the polynomials $\varphi_i(\sigma)$. The inner product between two functions of the configuration of cluster $\alpha$ then becomes

$$\langle f, g \rangle = \left(\frac{1}{2}\right)^{n_\alpha} \sum_{\sigma_1 = \pm 1} \cdots \sum_{\sigma_{n_\alpha} = \pm 1} f(\boldsymbol{\sigma}_\alpha) g(\boldsymbol{\sigma}_\alpha). \tag{4.5}$$

Now, any function of the configuration of the cluster can be expanded in the basis or characteristic function of the cluster. An important example is the configurational energy of the cluster which becomes

$$E(\boldsymbol{\sigma}) = \sum_\alpha V_\alpha \Phi_\alpha(\boldsymbol{\sigma}). \tag{4.6}$$

The coefficients $V_\alpha$ are the effective cluster interactions, given by the inner product of the characteristic function and configurational energy of the cluster,

$$V_\alpha = \langle \Phi_\alpha, E \rangle. \tag{4.7}$$

## 4.2 Evaluating the cluster expansion and selecting the effective cluster interactions

A crucial step in the CE method is to determine ECI by fitting them to calculated energies. A conventional way of doing this is by starting out with a few geometrically optimized input structures. Then, an optimal set of clusters within the structures is found by minimizing the cross-validation (CV) score [33], given by

$$CV = \frac{1}{n}\sum_{i=1}^{n}(E_i - \hat{E}_i)^2, \tag{4.8}$$

by fitting of ECI to the energy of the different configurations. Here, $\hat{E}_i$ is the energy of structure $i$, predicted by a fit to the $(n-1)$ other energies while $E_i$ is the calculated energy for that structure by DFT calculations. Historically, stepwise and ridge regression [34] has been used to a large extent for such regression purposes although they have some issues. One problem with stepwise regression is that it is not guaranteed to give better predictions except for some cases. Regarding ridge regression which aims to minimize overfitting by reducing the coefficients, this method does not set any ECI to zero and the model obtained can thus be complicated to understand. A method of regression that both guarantees better accuracy and performs a selection of the variables by setting some of them to zero is the least absolute shrinkage and selection operator (LASSO) method [35]. A sample of size $N$ is considered where each case $i$ in the sample has one outcome $y_i$. Furthermore, each case depends on $k$ variables given by the vector $x_i = (x_1, x_2, ..., x_k)^T$. The LASSO method can then be expressed as solving

$$\min_{\beta_0,\beta}\left[\frac{1}{N}\sum_{i=1}^{N}(y_i - \beta_0 - x_i^T\beta)^2\right] \tag{4.9}$$

subject to the condition

$$\sum_{j=1}^{k}|\beta_j| \leq t. \tag{4.10}$$

Here, $\beta$ is a vector containing $k$ elements which for the CE method corresponds to ECI that should be determined, while $t$ is the regularizing, free parameter. Using the LASSO method when determining ECI thus gives a

set of non-zero ECI possibly much smaller than the total number of ECI, making the CE model easier to interpret while at the same time minimizing the CV score. It can be useful to rewrite (4.9) and (4.10) into one equation. Let $y$ be a vector of the possible outcomes $y_i$ and $X$ be a matrix so that $X_{ij} = (x_i)_j$. Finally, by writing the $l^p$ norm as

$$||\beta||_p = \left( \sum_{i=1}^{N} |\beta_i|^p \right)^{\frac{1}{p}}, \tag{4.11}$$

the LASSO method can be written in a Lagrangian form as

$$\min_{\beta \in \mathbb{R}^k} \left[ \frac{1}{N} ||y - X\beta||_2^2 + \lambda ||\beta||_1 \right], \tag{4.12}$$

where $\lambda$ is a free parameter. This parameter is related to $t$, but their relationship depends on the data and is not given generally.

Chapter 5

# Computational methods

In this chapter, important computational details regarding GPAW and ASE as well as the CE and MC methods used will be presented. Additional theoretical aspects will also be given.

## 5.1 GPAW and ASE

As mentioned previously DFT has become a very popular choice for doing calculations on the atomic scale. There exists many ways to implement this formalism and to perform DFT calculations for this project the open-source DFT Python [36] code named GPAW [37] is employed. This program is based on the projector-augmented wave (PAW) method [38] and the atomic simulation environment ASE [39].

### 5.1.1 Density functional theory calculations

GPAW is based on the PAW method with element-specific functions and parameters from the GPAW library. By using the PAW method the KS orbitals, or wave functions, which oscillate rapidly near the nuclei are transformed to more smooth pseudo wave functions. GPAW is implemented within ASE. The way this is done is that GPAW utilizes ASE for setting up the domain of the calculation via the unit cell and the position of the different atoms as well as geometrically optimizing the structure. GPAW, being a calculator for DFT caculations perform several computational approxima-

tions to make it possible to calculate different quantities within an acceptable time frame. There is a tradeoff between the speed and the accuracy of a calculation and having too low accuracy in the calculations will typically give a systematic under- or overestimation of the energies. There are several parameters that can be specified for the calculations in GPAW affecting both the accuracy and computational cost, but the majority of them are kept at their default values in this thesis. One approximation is which xc functional to use which in this thesis is set to the PBE functional unless explicitly stated otherwise. Another one is the number of electronic bands to include for each spin, denoted $N_{\text{Bands}}$, with a default value equal to the number of atomic orbitals in the setup. Important to note is that if $N_{\text{Bands}}$ is negative the calculation will include a number of unoccupied bands in the calculation equal to the absolute value of that negative number in addition to all the occupied ones. The sampling of the Brillouin zone is also important where the default value is to only sample the $\Gamma$-point. In the calculations the Monkhorst-Pack [40] sampling is used by specifying the number of k-points in each direction. The smearing of the occupation numbers can also be controlled, following the Fermi-Dirac distribution given by (2.11) by setting the value for $k_{\text{B}}T$ in the equation with a default value of $k_{\text{B}}T$ = 0.1 eV when using periodic boundary conditions. It should be stressed that the fictitious temperature electrons corresponds to 0.1 eV, but the real temperature is 0 K in the calculations. Lastly, several convergence criteria can be set for the calculation cycles to be considered as self-consistent. For example, the default criterion for energy is that the difference over the three last iterations must be lower than 0.5 meV per valence electron whereas the difference in electron density has to be less than 0.0001 electrons per valence electron.

### 5.1.2 Plane-wave mode

GPAW represents the wave functions by using a real space grid of three dimensions. Although a real space grid implementation parallelizes very efficiently for large systems it is often faster for small systems to expand the wave functions in plane-waves in the $\mathbf{k}$-space. GPAW supports this plane-wave mode and as it is expected to work well for crystals because of their periodic structure this mode is used for the calculations in this project. Here, every quantity is represented by its corresponding Fourier transform where the Fourier transform is taken over the periodic unit cell. Moreover,

periodic boundary conditions are required giving bulk materials instead of molecules when using a small unit cell. The plane-waves used in the calculations are restricted by a kinetic energy cutoff, $E_{\text{Cut}}$, including only plane-waves with $|\mathbf{G}+\mathbf{k}|^2/2 < E_{\text{Cut}}$ where $E_{\text{Cut}}$ is 340 eV by default but can also be specified to other values.

### 5.1.3 Structure optimization

The different atomic structures are set up by using ASE and then geometrically optimized with GPAW and ASE. The algorithm used for optimizing structures in this project is called the preconditioned LBFGS algorithm. It is a linearized version of the BFGS algorithm [41] with limited memory. Moreover, it is a local optimization algorithm that converges when

$$\max_a |\mathbf{F}_a| < f_{max}. \tag{5.1}$$

Here, $a$ runs over all the atoms, $\mathbf{F}_a$ is the force experienced by atom $a$ and $f_{max}$ is a criterion set by the user which in the following calculations is set to $f_{max} = 0.05$ eV/Å. For each iteration of the algorithm it uses two quantities to determine where the atoms should be moved towards. The first quantity is the calculated force on each atom. The second is the matrix containing $\frac{\delta E^2}{\delta x_i \delta x_j}$, the second derivatives of the energy with respect to the coordinates of the nuclei called the Hessian matrix. For the LBFGS algorithm the inverse of the Hessian matrix, not the matrix itself, is updated at each step. By preconditioning [42] the optimization can be done in less time, where this advantage increases with the size of the system. It is possible to optimize both the unit cell as well as the atomic positions simultaneously or one at a time. Finally, it should be noted that optimizing the unit cell and atomic positions effectively changes the lattice constants of the structure and reduces the pressure to zero.

## 5.2 Cluster expansion

In this thesis the CE implementation developed at DTU is used [43]. In chapter 4 the cluster expansion method was introduced along with the evaluation of it and the selection of ECI. Here, the conventional way of developing the trial CE is given which is used in this project and based on [44].

### 5.2.1 Selection of additional structures for developing the cluster expansion

The structures used for DFT calculations and the clusters chosen within them determines the effectiveness and the precision of the CE method. Having obtained a trial CE which denotes the set of clusters chosen through the selected ECI, it should be evaluated using structures that do not belong to the initial set of input structures. This is because the trial CE gives better predictions for structures similar to the input structures as the CV score is calculated from this set. Hence, additional structures are needed for the evaluation of the trial CE. The conventional way of doing this is to predict near-ground and ground states of different compositions, optimize them and validate the trial CE by including the structures. The trial CE is developed in this way until the predicted ground states converge, adding the predicted near-ground and ground state structures to the initial set. The near-ground and ground states are predicted by simulated annealing [45] which is developed from the Markov chain MC method called the Metropolis-Hastings algorithm [46]. This method starts out with a structure at a given composition, random configuration and high temperature $T$. Using the analogy of annealing of metals the temperature is then decreased slowly. At each temperature $N_{MC}$ MC steps are performed where each step can be divided into two. The first is swapping the positions of two randomly selected atoms and computing the energy difference $\Delta E$ between the two different configurations by using the trial CE. The second is to accept the new state with a probability of

$$\min\left(1, e^{-\frac{\Delta E}{k_{\mathrm{B}}T}}\right). \tag{5.2}$$

In this way, by decreasing the temperature and repeating the process, the system will approach the near-ground or ground state of the desired composition as predicted by the trial CE.

## 5.3 Monte Carlo simulations

Having obtained a set of ECI from the CE method, MC simulations can be used for calculations. In this thesis, MC simulations are implemented in both the canonical and semi-grand canonical ensemble. These ensembles will be introduced here along with how to calculate enthalpy of formation,

free energy of formation, phase boundaries and free energy barriers. The ensembles are just briefly introduced as the focus is on their implementation in the MC simulations and further details are found in reference [47] and [48]. Most of the MC simulations in this thesis are performed based on the work by David Kleiven [49], except calculations of the free energy barriers. Hence, the algorithm for computing these barriers will be given in greater detail.

### 5.3.1 Canonical ensemble

A canonical ensemble is the statistical collection of the possible microscopic configurations of a system. Furthermore, the system is surrounded by a heat bath which the system is in thermal equilibrium with, at a given temperature $T$. For such a system both the volume $V$ and the number of particles $N$ are kept fixed. The temperature is the principal variable of the system, determining the probability of occupying each of the possible microscopic configurations and the fluctuations in the total energy of the system.

MC simulations can be performed in this ensemble to search for ground states and obtain a representative selection of configurations from the statistical distribution. This is done using the Metropolis-Hastings algorithm just as in simulated annealing for obtaining additional structures for the trial CE. It is thus necessary to highlight some important differences between obtaining new structures for the CE and performing MC simulations using the final CE. A key difference is that for simulated annealing only ground states are searched for, letting $T$ approach zero. For MC simulations using the final CE, not only the ground states at zero temperature can be approximated. In addition, as the Metropolis-Hastings algorithm is used, each MC step at each temperature makes the distribution of configurations approximate more and more the actual, desired distribution at that given temperature. Thus, both the ground state structures and and a representative selection of finite temperature simulations can be obtained, for example the total energy as a function of magnesium concentration in an alloy. Another difference is that for developing the trial CE, all additional structures should have the same total number of particles as the initial set and ECI are altered as each new set of structures is added. This is different from MC simulations using the final CE where the simulations can be performed for any desired number of particles as long as it is kept fixed. Also, ECI used

are given by the final CE and not altered as calculations are performed.

The last important difference is the number of MC steps, $N_{MC}$, performed at each temperature which is set to a certain value in simulated annealing. This is handled in another way for the subsequent MC simulations. In the canonical ensemble the energy is considered, taking a value $E_i$ at each MC step $i$. An estimator of the expectation value of the energy is the average over $M$ steps and given by

$$\overline{E}_{[1,L]} = \frac{1}{M} \sum_{i=1}^{M} E_i. \tag{5.3}$$

Then, $M$ is found by periodically estimating the variance of $\overline{E}_{[1,L]}$ [50] and end the simulation when this estimate satisfies

$$Var(\overline{E}_{[1,L]}) \leq \left( p\sqrt{\frac{2}{\pi}} \int_0^{1-\alpha} e^{-t^2} \mathrm{dt} \right)^2. \tag{5.4}$$

Here, $p$ is a precision in the energy set by the user and $\alpha$ is the confidence level. For example, wanting the precision $p$ with a probability of 99 % gives $\alpha = 0.01$. When stating the precision of a MC simulation in this thesis the parameter $p$ is meant.

### 5.3.2 Semi-grand canonical ensemble

The semi-grand canonical ensemble is similar to the canonical ensemble, but the crucial difference is that the chemical potential $\mu$ is allowed to vary. Thus, the volume and total number of particles are fixed at a given temperature, but the composition may change during a simulation. In practice the difference is that at each MC step, instead of swapping the atomic position of two atoms a randomly selected atom is replaced with an atom of another type. In a binary alloy of aluminum and magnesium this corresponds to changing one of the aluminum to a magnesium atom or the other way around. This new state is then accepted with a probability given by (5.2) as before. Doing MC simulations in this ensemble can as in the canonical ensemble both be used to search for ground states and obtain representative selections of configurations from the statistical distribution.

The number of MC steps needed for the simulation is determined in almost the same way as in the canonical ensemble. The only difference

is that in addition to the energy, the estimated variance of the chemical composition must satisfy the same requirement as for the energy.

### 5.3.3 Enthalpy of formation

A property of interest when doing MC simulations is the enthalpy of formation which is a measure of how much energy is needed or released when an alloy is formed from its elements via a chemical reaction. It can thus be used to examine whether a given alloy is more stable than the pure phases. The enthalpy itself is given by

$$H = U + pV, \tag{5.5}$$

where $U$ is the internal energy, $p$ is the pressure and $V$ is the volume. The enthalpy of formation is then defined as the change in enthalpy when forming a substance from its elements. The enthalpy of formation for a given reaction can thus be expressed as

$$H_{\text{Formation}} = \sum H_{\text{Products}} - \sum H_{\text{Reactants}}, \tag{5.6}$$

in effect the difference in enthalpy between the reactants and products. By optimizing the structures involved the pressure becomes zero and it is possible to calculate the enthalpy of formation for a given reaction by simply using total DFT energies. Given an alloy of two different elements, $N_a$ atoms of type $a$ and $N_b$ atoms of type $b$, the enthalpy of formation can be expressed as

$$H_{\text{Formation}} = E_{\text{Alloy}} - \frac{N_a}{N} E_a - \frac{N_b}{N} E_b. \tag{5.7}$$

Here, $E_{\text{Alloy}}$ is the total energy of the alloy, $E_a$ and $E_b$ is the total energy of a bulk of $N$ atoms of respectively type $a$ and $b$ and $N = N_a + N_b$.

### 5.3.4 Free energy of formation

Another property considered here is the free energy of formation. Changes in free energy for a physical system help decide the direction of its reactions as at a constant temperature a material will minimize its free energy. It is therefore a useful quantity to calculate. The Gibbs free energy is expressed as

$$G = U + pV - TS. \tag{5.8}$$

Here, $S$ is the entropy of the system. If the structures used are optimized the pressure becomes zero and (5.8) emerges as the Helmholtz free energy,

$$F = U - TS. \tag{5.9}$$

Calculating the Helmholtz free energy for an alloy of $N_a$ atoms of type $a$ and $N_b$ atoms of type $b$, in total $N$ atoms, the free energy of formation is then expressed as

$$F_{\text{Formation}} = \sum F_{\text{Products}} - \sum F_{\text{Reactants}} = F_{\text{Alloy}} - \frac{N_a}{N} F_a - \frac{N_b}{N} F_b. \tag{5.10}$$

Here, $F_{Alloy}$ is the free energy of the alloy and $F_a$ and $F_b$ is the free energies of bulks of $N$ atoms of type $a$ and $b$, respectively. Since the configurational entropy is zero in the pure phases the second term in (5.9) vanishes and the free energy of formation can be expressed using total DFT energies for the pure phases, as

$$F_{\text{Formation}} = F_{\text{Alloy}} - \frac{N_a}{N} E_a - \frac{N_b}{N} E_b. \tag{5.11}$$

The remaining step is to calculate $F_{\text{Alloy}}$. Consider the partition function in the semi-grand canonical ensemble, $Z$, the sum over the possible microstates of the system [48]. This is related to the free energy via

$$e^{-\beta G} = Z, \tag{5.12}$$

where $\beta = \frac{1}{k_B T}$ and $G$ is equivalent to $F$ when the pressure is zero. By taking the natural logarithm of both sides, differentiating and using that $U = -\frac{\partial \ln Z}{\partial \beta}$ [51, p. 56] this can be rewritten as

$$\mathrm{d}(\beta F) = U \mathrm{d}\beta. \tag{5.13}$$

Integrating this equation on both sides yields

$$\beta_1 F_1 = \beta_0 F_0 + \int_{\beta_0}^{\beta_1} U \mathrm{d}\beta. \tag{5.14}$$

Letting $T_0$ approach infinity the second term in (5.9) for $F_0$ becomes dominant and $F_0$ can be approximated as $-T_0 S_0$. The quantity $S_0$ is given by

$$S_0 \xrightarrow{T_0 \to \infty} -k_B \sum_i x_i \ln x_i, \tag{5.15}$$

known as the entropy of mixing. Here, $x_i$ is the fraction of atoms of type $i$ given as the number of atoms of type $i$, $N_i$, divided by the total number of atoms in the alloy, $N$. Thus, $\beta_1 F_1$ can be expressed as

$$\beta_1 F_1 = \sum_i x_i \ln x_i + \int_0^{\beta_1} U \mathrm{d}\beta \tag{5.16}$$

which can be integrated numerically. Note that letting $T_0$ approach infinity makes the lower limit of the integral become zero. Finally, the free energy of formation becomes, combining (5.11) and (5.16),

$$F_{\text{Formation}} = k_B T_1 \left[ \sum_i x_i \ln x_i + \int_0^{\beta_1} U \mathrm{d}\beta \right] - \frac{N_a}{N} E_a - \frac{N_b}{N} E_b \tag{5.17}$$

which has to be evaluated at a given temperature $T_1$.

### 5.3.5 Phase boundary tracing

To determine the phase diagram of a system it is less computationally expensive to trace the boundaries of a phase compared to finding the whole region where this phase is stable. Here, a method of finding the boundary of a equilibrium of two phases when knowing a point where a first-order transition occurs is presented based on reference [50]. To begin with, the thermodynamic potential per atom in the semi-grand canonical ensemble is defined as

$$\phi(\beta, \mu) = -\frac{1}{\beta N} \ln \left( \sum_i e^{-\beta N (E_i - \mu x_i)} \right), \tag{5.18}$$

where $\mu$ is the chemical potential and $E_i$ is the energy per atom of type $i$. Assume that at a known value of $\beta$ and $\mu$ the system is separated in a mixture of the two phases $\alpha$ and $\gamma$. At this known point the thermodynamic potential given by (5.18) has to be the same for the two phases. The key idea of phase boundary tracing is then to find the change in $\mu$ needed to keep the thermodynamic potential of the two phases equal as $\beta$ increases by $d\beta$. Repeating this, $\mu(\beta)$ is approximated which again can be used to obtain the different concentrations in each phase at different temperatures. The equation which needs to be numerically solved is found by differentiating

$\beta\phi_\alpha = \beta\phi_\gamma$, which is true along the phase boundary. After taking the total differential the result can be rewritten as

$$\frac{d\mu}{d\beta} = \frac{E_\gamma - E_\alpha}{\beta(x_\gamma - x_\alpha)} - \frac{\mu}{\beta}, \tag{5.19}$$

where $x_\alpha$ and $E_\alpha$ is the concentration and energy evaluated at a given $\mu$ and $\beta$ for the phase $\alpha$, and likewise for the other phase $\gamma$. To numerically solve (5.19) MC simulations in the semi-grand canonical ensemble can be used to calculate all the necessary values.

### 5.3.6 Free energy barriers

For a given alloy the minimum in free energy as a function of the concentration of an atom, $G(x)$, indicates the most energetically favourable composition. It is possible to have local minima in free energy in addition to the global one, separated by higher levels of energy which are denoted free energy barriers. One can calculate the difference in free energy as a function of the concentration of an atom and thus the free energy barriers in the semi-grand canonical ensemble. In this ensemble the partition function is given by (5.12). As mentioned earlier it is also the sum over the possible microstates of the system, given by

$$Z = e^{-\beta G} = \sum_{\{\sigma\}} e^{-\beta(E(\sigma) - \mu x(\sigma))}. \tag{5.20}$$

Here, $x$ is the concentration of a chosen type of atoms and $\sigma$ is a given configuration of the system of atoms. The sum is thus performed over all possible configurations of the system. Now, the free energy of a given concentration $\tilde{x}$ can be calculated by taking the sum over all configurations giving this concentration,

$$e^{-\beta G(\tilde{x})} = \sum_{\{\sigma\} \in Q_{\tilde{x}}} e^{-\beta(E(\sigma) - \mu\tilde{x})} \tag{5.21}$$

where $Q_{\tilde{x}}$ is the set of configurations giving the concentration $\tilde{x}$. This can be rewritten as

$$e^{-\beta G(\tilde{x})} = \sum_{\{\sigma\}} \delta(x(\sigma) - \tilde{x}) e^{-\beta(E(\sigma) - \mu x(\sigma))} \tag{5.22}$$

where $\delta(x-y)$ is 1 for $x=y$ and 0 otherwise. This can again be rewritten as

$$e^{-\beta G(\tilde{x})} = Z\langle\delta(x(\sigma)-\tilde{x})\rangle \tag{5.23}$$

where $\langle ...\rangle$ denotes the thermodynamic average

$$\langle A\rangle = \frac{1}{Z}\sum_{\{\sigma\}} A(\sigma)e^{-\beta(E(\sigma)-\mu x(\sigma))}. \tag{5.24}$$

The difference in free energy between two concentrations $\tilde{x}_1$ and $\tilde{x}_2$ is thus found via

$$\frac{e^{-\beta G(\tilde{x}_1)}}{e^{-\beta G(\tilde{x}_2)}} = \frac{Z\langle\delta(x(\sigma)-\tilde{x}_1)\rangle}{Z\langle\delta(x(\sigma)-\tilde{x}_2)\rangle}. \tag{5.25}$$

Defining $\Delta G = G(\tilde{x}_1) - G(\tilde{x}_2)$ this becomes

$$e^{-\beta\Delta G} = \frac{\langle\delta(x(\sigma)-\tilde{x}_1)\rangle}{\langle\delta(x(\sigma)-\tilde{x}_2)\rangle} \tag{5.26}$$

which is independent on Z. Doing a MC simulation in the semi-grand canonical ensemble a histogram $\mathcal{H}$ counting the number of times the system visits the different concentrations can be obtained by analyzing the trajectory of the simulation. What is sampled in the simulation is thus $\langle\delta(x(\sigma)-\tilde{x})\rangle$ for all concentrations which can be used to calculate the differences in free energy via (5.26). In this manner, $\Delta G$ can be found for all concentrations.

The problem with this approach is that the system quickly goes towards the most favourable state and poor results are obtained for the regions of concentration where the probability of occupation is low. This is solved by applying the umbrella sampling method [52]. With this method, the total range of concentration of interest is divided into a set of $N$ concentration windows. Each window contains a given number of bins such that there is a one-to-one correspondence between every concentration and bin in that window. The first window at one end of the range of concentration contains $M$ bins. All the other $N-1$ windows contain $M+1$ bins, assigned to each concentration such that the first bin in each window corresponds to the last bin in the previous window. This is illustrated in Figure 5.1.

In each window $n$ a given number of MC steps is performed, obtaining a histogram $\mathcal{H}_n$ for calculating $\Delta G$ in that window. By requiring that $\Delta G$ is a continuous function of the concentration the overlapping concentrations/bins for every two following windows are used to merge $\Delta G$ between

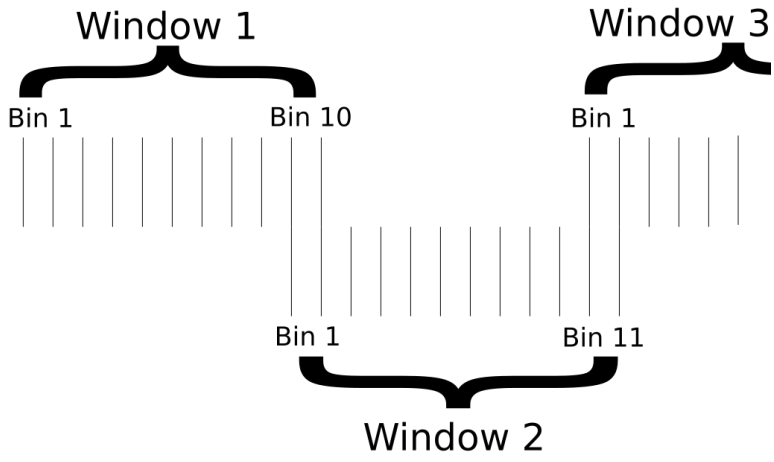



**Figure 5.1:** Schematic figure of the windows and bins in umbrella sampling.

them. This is illustrated in Figure 5.2. In this fashion $\Delta G$ can be obtained for all concentrations, also expecting good statistics in the ranges of concentration with a low probability if being occupied as the MC simulation is forced to perform a given number of steps inside each concentration window.

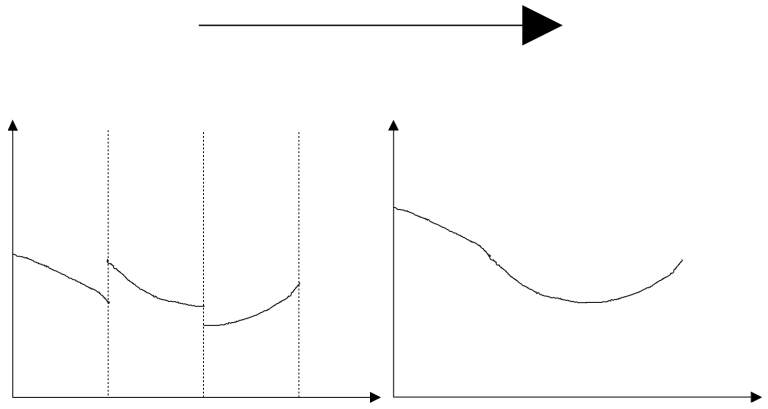


**Figure 5.2:** Illustration of the merging of $\Delta G$ between different windows.

Chapter 6

# Results and discussion

This chapter includes presentations and discussions of the results in the thesis. First, some results for the convergence of energy calculations using different values for the parameters in GPAW and calculations of the lattice constant of aluminum and magnesium using different xc functionals are presented. After this, calculations of cohesive energy and the formation energy of a vacancy follow. All of these were obtained for the specialization project, but are included here to explain the choice of parameters for the DFT calculations and show that they give adequate results. After this follows a presentation of the final CE developed for aluminum and magnesium in a hcp lattice. MC simulations including phase boundaries and free energy barriers based on the final CE are then presented and discussed.

## 6.1 Convergence of energy calculations

The purpose of this section is to find suitable choices for the different parameters in GPAW depending on the size of the system. This is done by calculating the total energy divided by the number of atoms varying each parameter at a time for one system of 1 aluminum atom and another one of 64 aluminum atoms. The structures are not optimized so as to speed up the calculations. It is often more interesting to look at energy differences, and moreover they tend to converge faster because of the systematic under-

or overestimation of the individual energy calculations. Here, individual energy calculations are used. This is because they are expected to converge more slowly than energy differences. Thus, the convergence criteria [1] they give for the different parameters are believed to be stricter than what is obtained for energy differences. An important point is when a calculation is said to be converged. The energy calculated at step $i$ is denoted $E_i$ where an increase in $i$ will increase the accuracy. Roughly speaking, the calculation will then be considered converged at step $i-1$ if the difference between $E_i$ and $E_{i-1}$ is negligible. The fixed values for the different parameters when checking the convergence in energy for a set of values for another parameter are summarized in Table 6.1a and Table 6.1b for a system of 1 and 64 aluminum atoms, respectively. When specifying the number of k-points the cube root of this will be the number of k-points in both the x-, y- and z-direction. The calculations are performed using the PBE functional.

**Table 6.1:** Fixed values for $E_{\text{Cut}}$, k-points, $N_{\text{Bands}}$ and $k_{\text{B}}T$ in the convergence tests.

**(a)** For a system of 1 aluminum atom.

| Parameter | Fixed value |
|---|---|
| $E_{\text{Cut}}$ [eV] | 340 |
| k-points | 512 |
| $N_{\text{Bands}}$ | $-2$ |
| $k_{\text{B}}$T [eV] | 0.1 |

**(b)** For a system of 64 aluminum atoms.

| Parameter | Fixed value |
|---|---|
| $E_{\text{Cut}}$ [eV] | 400 |
| k-points | 27 |
| $N_{\text{Bands}}$ | $-10$ |
| $k_{\text{B}}$T [eV] | 0.1 |

Figure 6.1 shows the total energy per atom of the two different structures for a set of cutoff energies where the other parameters are kept at their fixed values. From this, a value of $E_{\text{Cut}}$ = 800 eV for 1 atom and $E_{\text{Cut}} \simeq$ 500 eV for 64 atoms is considered to give highly accurate results. The systematic mismatch is understandable based on the difference in sampling of the k-space. Figure 6.2 shows the total energy per atom of the two different structures calculated with a varying number of k-points and the rest of the parameters at fixed values. Here, the number of k-points that must be used to have accurate results are found to be 512 and 64 for 1 and 64 atoms, respectively. One can see that with intensive sampling in k-space the same

[1] Not to be confused with the term convergence criteria in chapter 5.1.1.

result is obtained. Figure 6.3 shows the total energy per atom of the two different structures as a function of $k_{\mathrm{B}}T$ used in the Fermi-Dirac distribution given by (2.11), keeping the other parameters at fixed values. To get convergent results, $k_{\mathrm{B}}T$ = 0.1 eV for the smearing parameter seems to be good enough for both structures of 1 and 64 atoms.

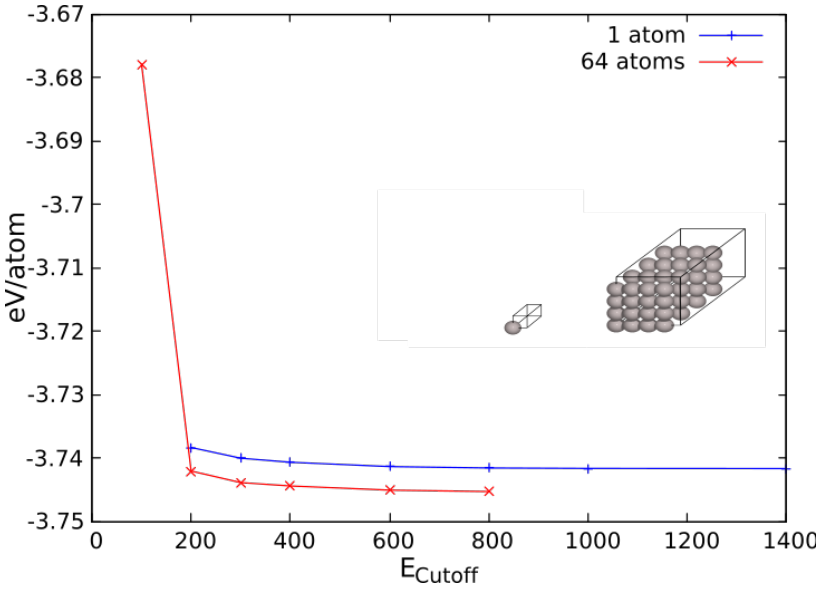



**Figure 6.1:** Total energy per atom as a function of $E_{\mathrm{Cut}}$ for a system of 1 and a system of 64 aluminum atoms.

Thus, a rule of thumb has been established for the value of the different parameters when performing DFT energy calculations depending on the size of the system. However, it should be noted that computing other quantities than energies may require a different set of parameters for fully convergent results.

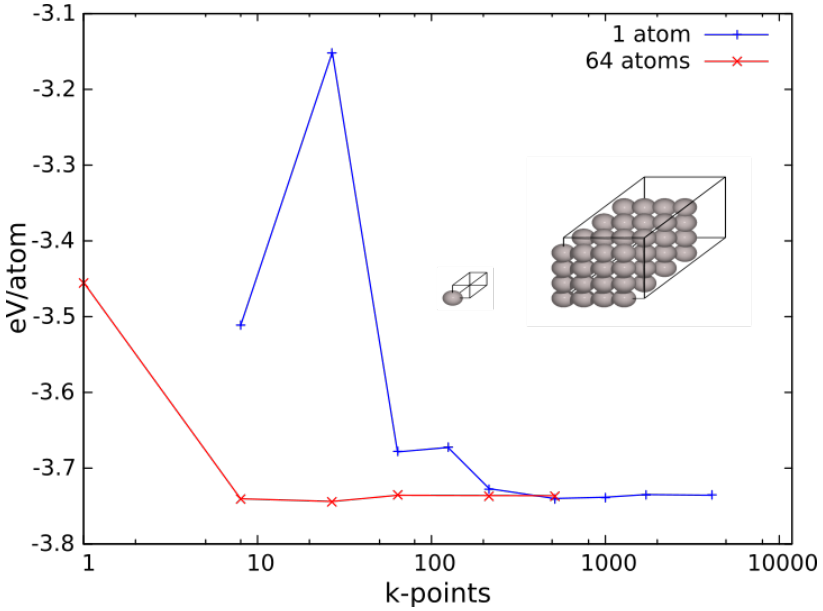



**Figure 6.2:** Total energy per atom as a function of the number of k-points for a system of 1 and a system of 64 aluminum atoms.

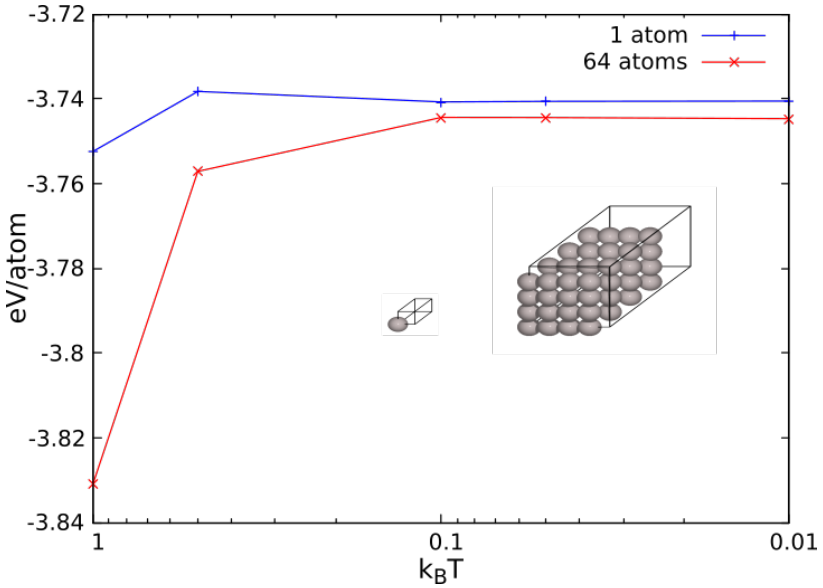



**Figure 6.3:** Total energy per atom as a function of $k_{\mathrm{B}}T$ for a system of 1 and a system of 64 aluminum atoms.

## 6.2 Lattice constants for different functionals

The lattice constants for bulk aluminum and magnesium are estimated using the LDA, BLYP and PBE xc functionals. This is done by optimizing the materials primitive unit cell using each of the functionals where the dimensions of the optimized cell are used to obtain the calculated lattice constant. The results for the three functionals as well as the experimental values previously introduced are given in Table 6.2a for aluminum and Table 6.2b for magnesium. The results agree well with previous calculations of the same type [53] which indicates that the set of parameters for systems of a few atoms found in section 6.1 gives accurate results also for optimizing structures and estimating lattice constants. Considering the LDA functional, a tendency to underestimate the lattice constants is observed which can be expected based on previous calculations with this functional [24]. Moreover, the size of the error can be explained by the fact that LDA is the simplest type of approximation for the xc functional. The BLYP and PBE functionals perform better, where the latter is seen to fit best with experimental values. Although the choice of the PBE functional for this thesis is not based on these results it is reassuring that this functional gives good results for the lattice constants.

**Table 6.2:** Lattice constants, experimentally and estimated using the LDA, BLYP and PBE exchange-correlation functionals.

**(a)** For bulk aluminum.

| | $a$ [Å] |
|---|---|
| Experimental | 4.05 [7] |
| LDA | 3.983 |
| BLYP | 4.112 |
| PBE | 4.04 |

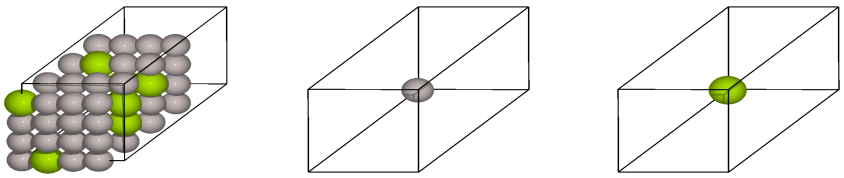
**(b)** For bulk magnesium.

| | $a$ [Å] | $c$ [Å] |
|---|---|---|
| Experimental | 3.21 [7] | 5.21 [7] |
| LDA | 3.126 | 5.124 |
| BLYP | 3.203 | 5.3 |
| PBE | 3.186 | 5.223 |

## 6.3 Cohesive energy

Returning to energy calculations where the convergence of the results depending on the different parameters has been studied thoroughly in chapter

6.1, the cohesive energy per atom is investigated. A calculation is done for a system of 56 aluminum and 8 magnesium atoms placed in a bulk fcc structure using the lattice constant of aluminum. The total energy of the bulk as well as for a single aluminum and magnesium atom is needed for the calculation. Important to note is that the single atom calculations should be performed in large unit cells to keep them separated from replicated copies. First, both the cell and atomic positions of the bulk of in total 64 atoms is optimized and the total energy is calculated. Reusing the same unit cell, the single atom is placed inside which ensures enough space between it and its nearest neighbours which are present because of periodic boundary conditions. The structures used are illustrated in Figure 6.4. Calculating the energy of both single-atom structures in addition to the bulk the desired cohesive energy per atom can be obtained from (2.16).



**(a)** The optimized bulk. **(b)** The single aluminum atom. **(c)** The single magnesium atom.

**Figure 6.4:** The structures used in the calculation of cohesive energy.

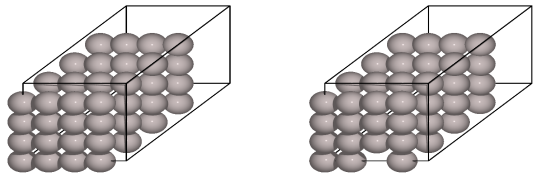
The theoretical values for bulk aluminum and magnesium are $\epsilon_c^{\text{Al}} = 3.39$ eV per atom and $\epsilon_c^{\text{Mg}} = 1.51$ eV per atom [7, p. 50]. A crude estimate of the calculation can be obtained via interpolation of the theoretical values giving $\epsilon_c^E = 3.16$ eV per atom, where the superscript $E$ denotes that it is an estimate. The calculated result for the structure under consideration is even closer to the value for aluminum, at $\epsilon_c = 3.32$ eV per atom. The positive sign which is expected shows that it is energetically favourable for the structure to be arranged in a fcc lattice as compared to separate atoms.

The calculation is repeated for another structure consisting of 48 aluminum and 16 magnesium atoms in a fcc bulk. Via interpolation the estimated cohesive energy is $\epsilon_c^E = 2.92$ eV while the calculated result is $\epsilon_c = 3.05$ eV. Again, the calculated value is higher than the crude estimate and the positive sign shows that the arrangement is energetically favourable.

The reason for both results here being closer to the theoretical value of aluminum is that both structures are rich in aluminum. Thus, both structures favour a fcc structure and the magnesium bulk reference value is far off in terms of structure and density. This makes the cohesive energy become closer to the theoretical value of aluminum than the linear interpolation. These results together with previous calculations for the cohesive energies of aluminum [54] and magnesium [55] with the PBE functional give us confidence that the DFT simulations give results of good quality.

## 6.4 Formation energy of a vacancy

Another total energy calculation performed is the formation energy of a vacancy. The calculation is done by first creating a bulk structure of the desired composition and optimizing both the unit cell and atomic positions before calculating the total energy. After this, the same structure is created, but one of the atoms is removed which creates a vacancy. This new structure is also optimized and the total energy is calculated. The formation energy of a vacancy can then be calculated via (2.19). This is done for a bulk fcc structure of aluminum as well as for a bulk hcp structure of magnesium. The method is illustrated for aluminum in Figure 6.5, showing the structures used where the vacancy can be observed at the bottom of Figure 6.5b.



**(a)** The optimized bulk. **(b)** The optimized bulk with a vacancy.

**Figure 6.5:** The structures used in the calculation of the formation energy of a vacancy.

The results are summarized in Table 6.3 including experimental results for aluminum and magnesium. As seen, the calculations are very good compared to previous experiments although a trend of obtaining slightly

higher values can be seen. These results shows that the parameters for GPAW found in section 6.1 are adequate for energy calculations.

**Table 6.3:** Formation energy of a vacancy in eV for bulk fcc aluminum and bulk hcp magnesium.

| | $E_{vf}$ | Experimental |
|---|---|---|
| Al | 0.67 | 0.64 [56] |
| Mg | 0.83 | 0.81 [57] |

## 6.5 Cluster expansion

A CE model is constructed from an initial set of structures of aluminum and magnesium of in total 64 atoms and varying concentration. The atoms are arranged in a hcp lattice corresponding to that of magnesium and the atomic positions are randomized. For all structures in both the initial and final set the composition varies from 25 % to 100 % magnesium and the structures as well as lattice constants are optimized. The reason for setting a limit at 25 % concentration of magnesium is that it improves the CV score of the CE. Also, as pure aluminum is arranged in a fcc lattice it does not fit well in the hcp lattice which is used here. The CE is developed from the initial set by repeatedly using simulated annealing to predict new sets of ground and near-ground states until converged. In addition, all possible configurations of a structure of 8 atoms replicated to 64 atoms are included, giving in total 101 structures. The DFT calculations of the energies are performed with $E_{\text{Cut}}$ = 500 eV, 64 k-points, $k_{\text{B}}T$ = 0.1 eV, $N_{\text{bands}}$ equal to 1.2 times the number of occupied bands and using the PBE functional. Finally, the configurational variable $\sigma$ takes values 1 for aluminum and -1 for magnesium.

The evaluation of the final cluster expansion is shown in Figure 6.6. Here, the predicted energy of each structure using the ECI of the final CE is shown with red circles together with a line indicating that predicted energies correspond to energies calculated via DFT. Red circles exactly at the black line thus show that CE and DFT calculations match excellently with a CV score of 3.2 meV/atom. The CV score and the root-mean-square error [58] is also included in the figure.

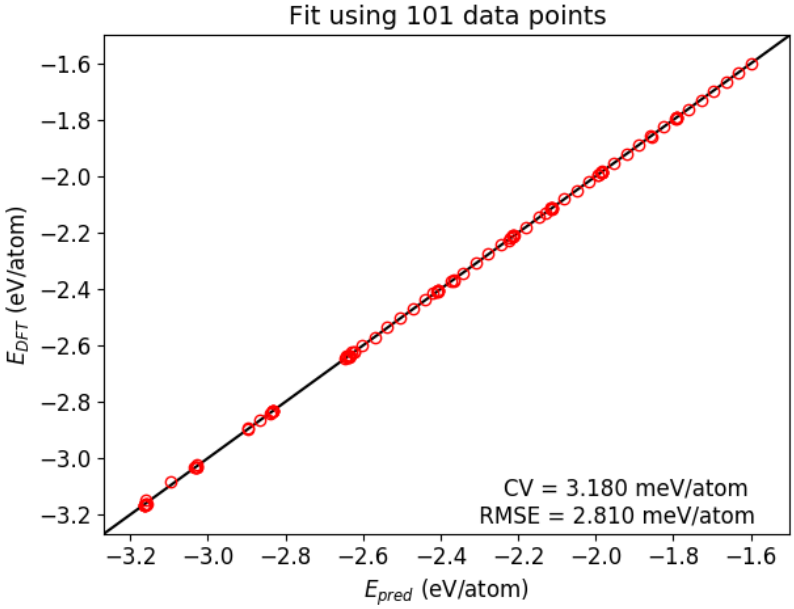



**Figure 6.6:** Evaluation of the final cluster expansion. The red circles are energies predicted by the CE and the black line indicates where $E_{\mathrm{DFT}} = E_{\mathrm{CE}}$.

The ECI obtained from this cluster expansion are given in Figure 6.7 while Figure 6.8 shows the same ECI, not including the zero- and one-body clusters. For detail, the ECI are also given numerically in Table 8.1 in Appendix A.

The zero- and one-body ECI are not interesting when looking at which configurations are energetically favourable because they are trivial. Also, the four-body ECI are hard to interpret. Thus, the most interesting ECI to examine further are the two-body ECI c2_1414_1_00 as well as the two different three-body ECI as these are the largest remaining ECI. Their clusters are shown in Figure 6.9 whereas the rest of the clusters are shown in Figure 8.1 in Appendix A. The two-body cluster shown in Figure 6.9a is made up of two atoms at a second nearest neighbour distance in the hcp lattice. The first three-body cluster, c3_1000_1_000 shown in Figure 6.9b, consists of three atoms all in nearest neighbour distance to each other in a hcp lattice. The other three-body cluster, c3_1414_1_000 shown in Figure 6.9c, consists of three atoms in a hcp lattice as well. Here, all interatomic distances corresponds to the nearest neighbour distance except one which corresponds

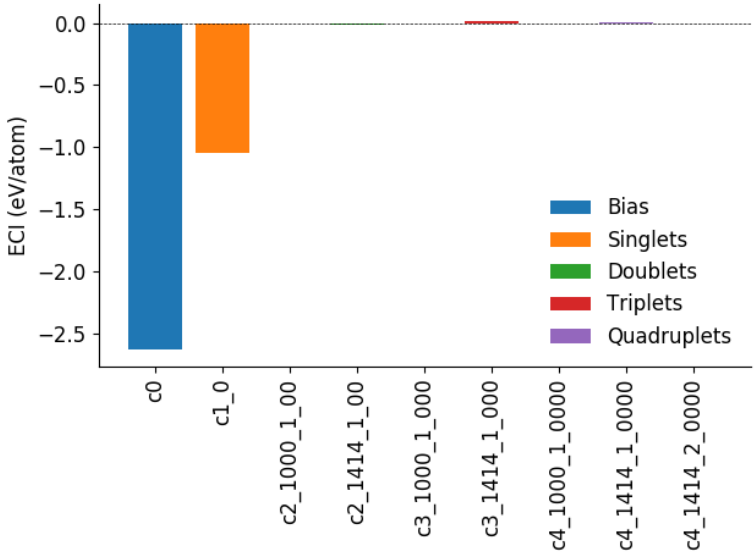



**Figure 6.7:** The effective cluster interactions obtained from the final cluster expansion.

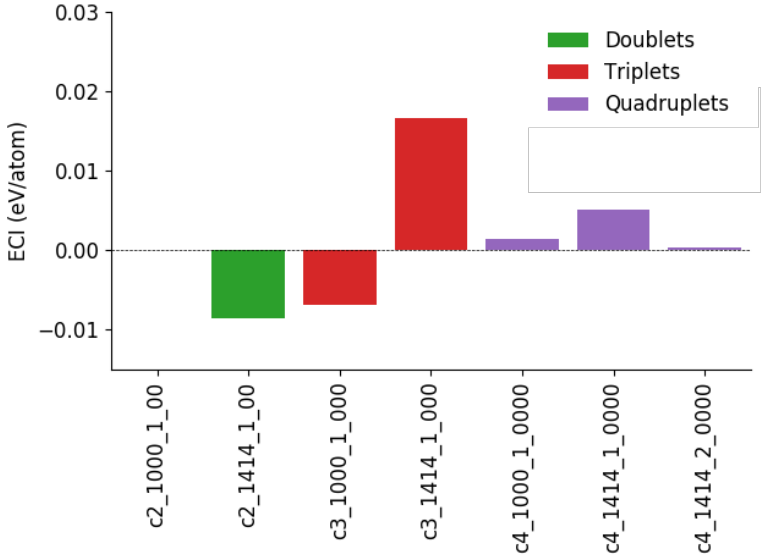
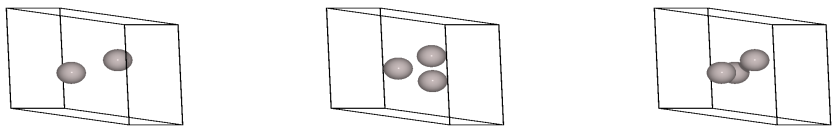



**Figure 6.8:** The effective cluster interactions obtained from the final cluster expansion, not showing the zero- and one-body clusters.

to the second nearest neighbour distance.



**(a)** c2_1414_1_00. **(b)** c3_1000_1_000. **(c)** c3_1414_1_000.

**Figure 6.9:** The three most important clusters of the ECI, illustrated using aluminum atoms.

The clusters of the three ECI highlighted here can be used to get an impression of which structures are energetically favorable for aluminum-magnesium alloys in a hcp lattice. This is done by combining (4.6) and the sign of the ECI for one cluster and deduce which configurations are favourable, as the system will try to minimize its energy. As mentioned, the configurational variable $\sigma$ is 1 for aluminum and -1 for magnesium and the ECI of the two-body cluster c2_1414_1_00 is negative. To get negative contributions to the configurational energy a positive characteristic function is needed which is only possible by putting equivalent atoms in the cluster. This is an indication that it is energetically favourable to have a second nearest neighbour distance between equal atoms, for both aluminum and magnesium atoms in this system. In the same manner, the three-body cluster c3_1000_1_000 gives negative contributions to the energy for three aluminum atoms or one aluminum and two magnesium atoms, all in a nearest neighbour distance. For the cluster c3_1414_1_000, negative energy contributions occur when either two aluminum and one magnesium atom or three magnesium atoms are assigned to the cluster.

The calculated lattice constants $a$ and $c$ from the optimization of the hcp structures in the DFT calculations for the CE are also analyzed. These are given in Figure 6.10. The red line represents a linear regression which is obtained from the experimental results reported in [59]. Moreover, the root-mean-square errors for the two different lattice constants when compared with the obtained regression line are indicated.

As seen, the calculated lattice constants depend linearly on the concentration just as in the experimental results, both decreasing with higher

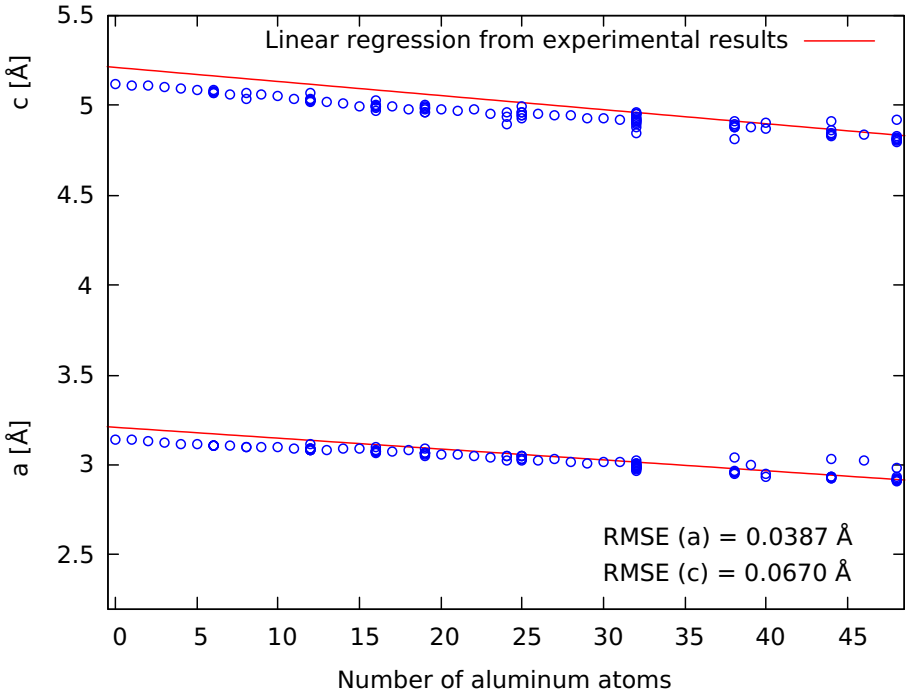



**Figure 6.10:** Calculated lattice constants $a$ and $c$ of the hcp lattice with root-mean-square errors, from the optimization of the structures used in the cluster expansion.

amount of aluminum. Also, the numerical results have root-mean-square errors that correspond to a 1.21 % and 1.29 % deviation, for $a$ and $c$ respectively, from the theoretical values of pure magnesium. This means that the numerical results match well with the experimental ones. A last observation is that both lattice constants seem to be slightly underestimated resulting in a overestimation of the density, especially for high concentrations of magnesium. This is not in agreement with previous results as given in Table 6.2b where the PBE functional is seen to overestimate the lattice constant $c$. This indicates a difference in the lattice constants calculated via optimization of structures when using a primitive or a larger unit cell. However, since periodic boundary conditions are applied the difference is believed to be due to the difference in input parameters, especially the sampling of k-space, for DFT calculations of small versus large structures.

## 6.6 Monte Carlo simulations

Having developed a final CE model the ECI obtained are further used to do MC simulations in both the canonical and semi-grand canonical ensemble.

### 6.6.1 Cooling down at constant chemical potential

A set of MC simulations is performed by cooling down a system of 1000 atoms to 50 K at a constant chemical potential in the semi-grand canonical ensemble. This is done for 41 different values of the chemical potential $\mu$, ranging between -0.99 and -1.06 eV/atom and with a precision of $10^{-3}$ effectively determining the simulation length via (5.4). These limits are used because for these values pure aluminum and pure magnesium are formed, respectively, as the simulation approaches 0 K. The results are shown in Figure 6.11.

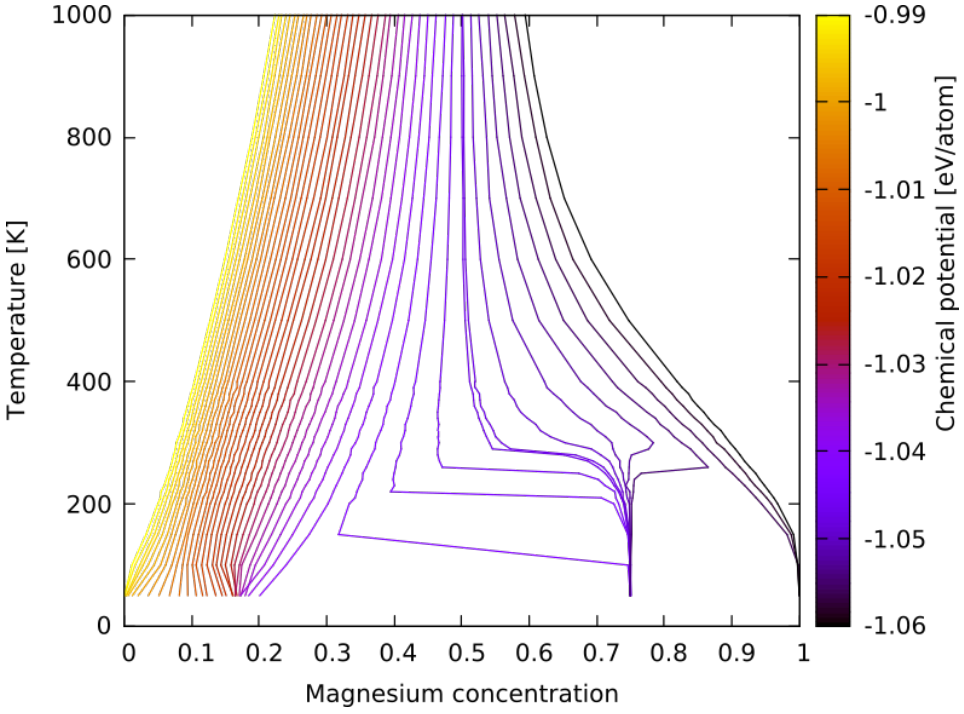



**Figure 6.11:** Monte Carlo simulations in the semi-grand canonical ensemble, cooling down a system of 1000 atoms to 50 K for different chemical potentials.

The magnesium concentration ranges from 0 to 1, indicating the molar fraction in the system. These results show that for high or low enough values of $\mu$ the system tends to go towards pure aluminum and pure magnesium, respectively. For many of the chemical potentials close to -0.99 the system goes towards an alloy rich in aluminum, but still containing a bit of magnesium. This is probably an artifact of the final CE model where the ECI are obtained from structures of 25 % magnesium or more. In addition to pure aluminum and magnesium, for some chemical potentials the system goes towards a structure of 75 % magnesium which indicates that this is another stable phase for aluminum and magnesium alloys in a hcp lattice. This transition to the $Mg_3Al$ phase occurs slightly below 300 K for chemical potentials between -1.0546 and -1.0420 eV/atom. All final structures from the MC simulation which end up at 75 % magnesium are equal. This structure is shown in Figure 6.12 and it is compatible with some of the implications derived earlier for the different ECI. Magnesium atoms are placed, although not exclusively, at second nearest neighbour distances which is an implication from the two-body cluster. Moreover, one can observe three-body clusters of nearest neighbours consisting of one aluminum and two magnesium atoms which is consistent with the three-body cluster c3_1000_1_000. As indicated by the three-body cluster c3_1414_1_000, clusters of three magnesium atoms with all distances equal to the nearest neighbour distance except one being that of the second nearest neighbour are observed as well. A final observation is that at high temperatures the magnesium concentration tends to go towards 50 %, showing that there is no clustering and aluminum and magnesium are mixed together randomly.

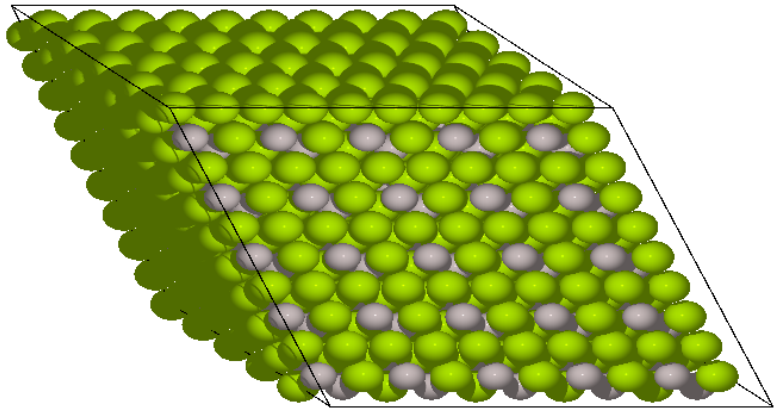


**Figure 6.12:** The final structure for Monte Carlo simulations in the semi-grand canonical ensemble ending up at 75 % magnesium.

### 6.6.2 Cooling down at constant concentration of magnesium

A set of MC simulations is also performed in the canonical ensemble. In each simulation a system of 1000 atoms is cooled down to 100 K at a constant concentration of magnesium. This is done for 17 different concetrations ranging from 25 % to 100 % magnesium as these are the limits in the final CE model. The precision used is $10^{-5}$. From these simulations the enthalpy of formation as well as the free energy of formation can be calculated at each temperature and concentration. The results for the enthalpy of formation are shown in Figure 6.13. Again, the magnesium concentration indicates the molar fraction in the system.

As mentioned earlier, the enthalpy is zero for pure elements, also shown in these results for magnesium. At high temperatures, as the system seeks to minimize its enthalpy, a mixture of pure aluminum and pure magnesium

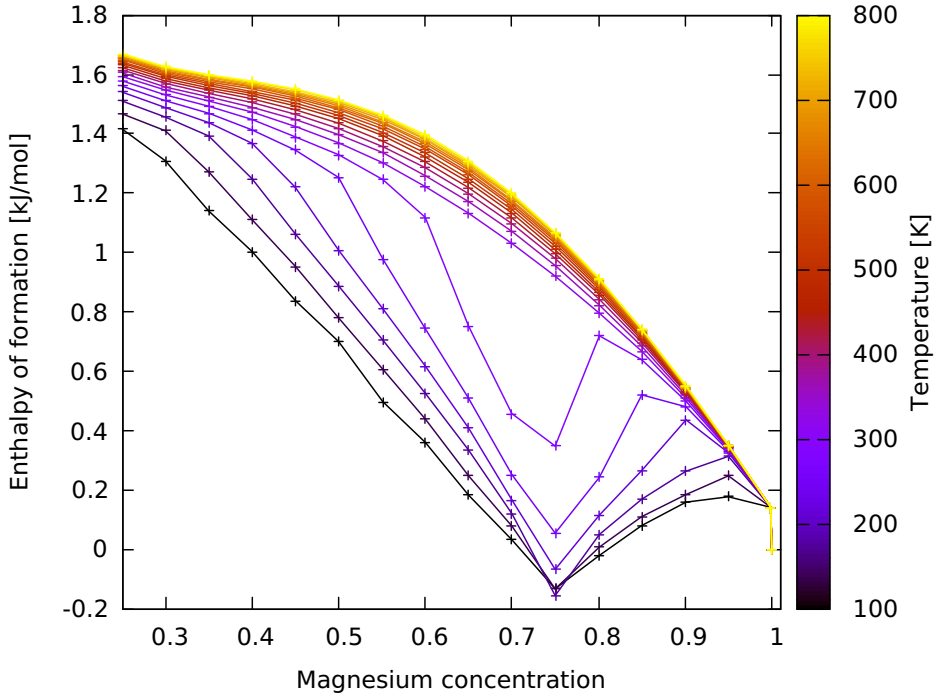



**Figure 6.13:** Enthalpy of formation calculated using Monte Carlo simulations in the canonical ensemble, cooling down a system of 1000 atoms to 100 K for different concentrations of magnesium.

in segregated phases is thus what can be expected based on these results. This is because structures of other concentrations are not favourable to form in terms of enthalpy. However, as the temperature is lowered a transition occurs slightly below 300 K. Here, structures containing 75 % suddenly get a significantly lower enthalpy of formation. Decreasing the temperature further, approaching 200 K, the enthalpy of formation becomes negative and in effect lower than for the pure elements. Thus, for temperatures below 200 K the system will try to form distinct regions containing a concentration of 75 % magnesium in addition to the pure elements. Cooling the system down to 100 K the final structure obtained at 75 % magnesium is exactly the same as the one given in Figure 6.12. Thus, the MC simulations in the semi-grand canonical ensemble and the canonical ensemble both predict a stable phase for low temperatures in addition to the pure elements, namely $Mg_3Al$.

The data from the MC simulations in the canonical ensemble is also used to calculate the free energy of formation via (5.17). The results are shown in Figure 6.14.

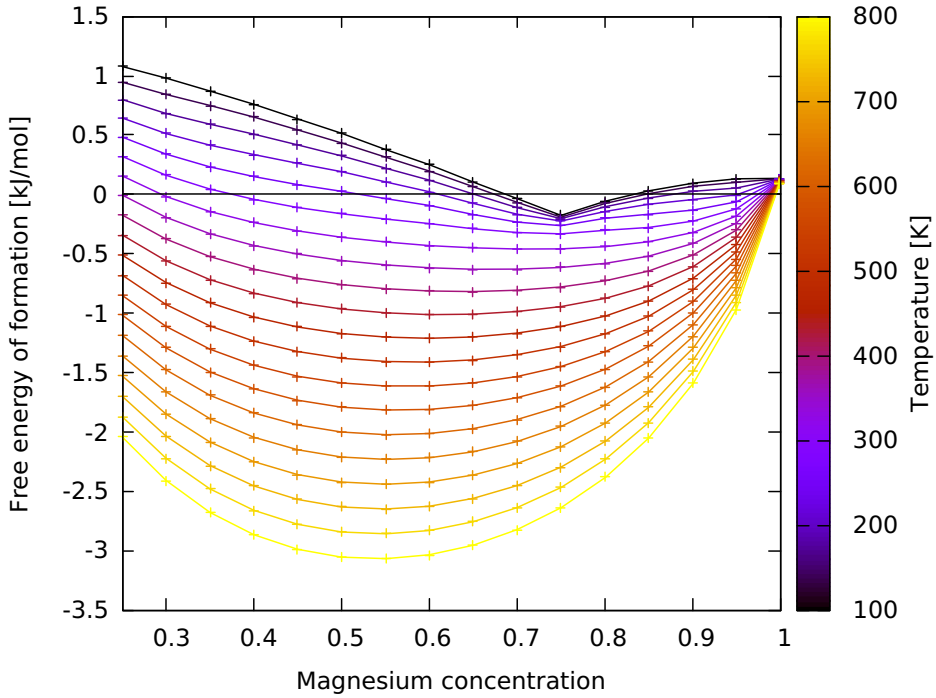



**Figure 6.14:** Free energy of formation calculated using Monte Carlo simulations in the canonical ensemble, cooling down a system of 1000 atoms to 100 K for different concentrations of magnesium.

As shown here the phase $Mg_3Al$ is favourable in terms of free energy of formation. This is the case for low temperatures up to about 250 K, indicated by the minimum in the free energy of formation at 75 % magnesium. Increasing the temperature, the minimum at 75 % magnesium disappears and the minimum shifts gradually towards a concentration of 50 %. This is more clear in the high temperature limit which is shown in Figure 6.15 where the free energy of formation is calculated from a MC simulation at 20 000 K. Here, the minimum is clearly located at 50 % magnesium as expected based on entropy of mixing. Hence, these results are consistent with the previous ones giving the phase $Mg_3Al$ at temperatures lower than around 250 K and a mixture of aluminum and magnesium above the same

temperature. Here however, aluminum and magnesium form an alloy at a finite temperature. This is in contrast to the conclusions based on enthalpy which misses the contribution from the entropy of mixing. The alloy obtain when considering the free energy and not the segregated, pure elements is the correct physical picture.

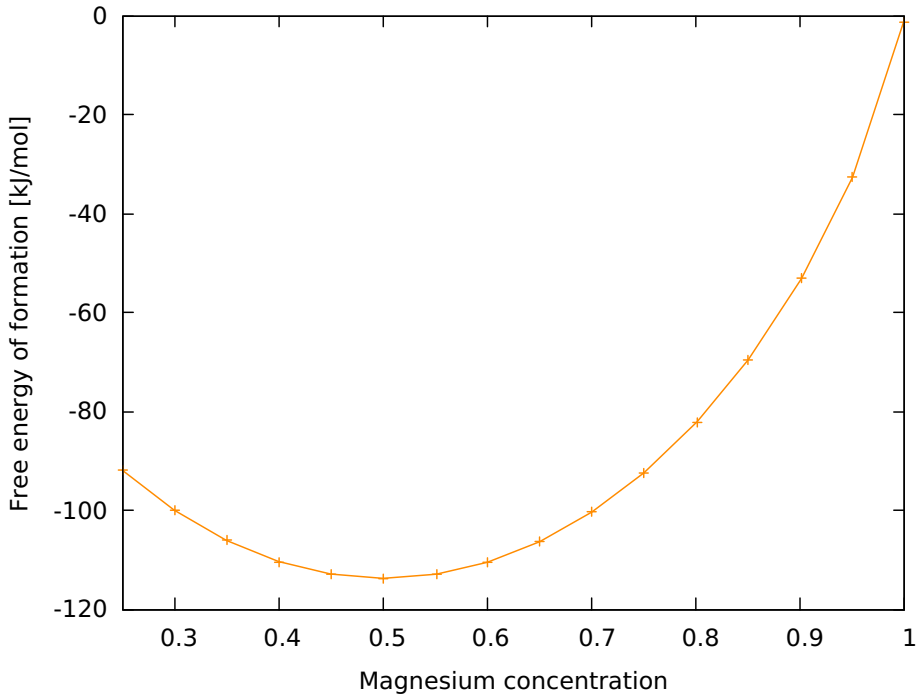



**Figure 6.15:** Free energy of formation calculated using Monte Carlo simulations in the canonical ensemble, for a system of 1000 atoms at 20 000 K for different concentrations of magnesium.

### 6.6.3 Phase boundary tracing

By performing MC simulations in the semi-grand canonical ensemble all quantities necessary to trace the boundaries of the phase $Mg_3Al$ via (5.19) are calculated. This is done 100 times with a precision of $10^{-3}$ as higher precision would be too expensive in terms of computational time. Each simulation performed takes about 6-8 hours on a standard computer. In Figure 6.16 the average temperature as a function of $\Delta\mu$ is shown together

with its standard deviation [58], starting at $\mu = -1.058$ eV/atom and $T$ = 100 K. Along this line the two phases $Mg_3Al$ and pure magnesium coexist in the hcp lattice.

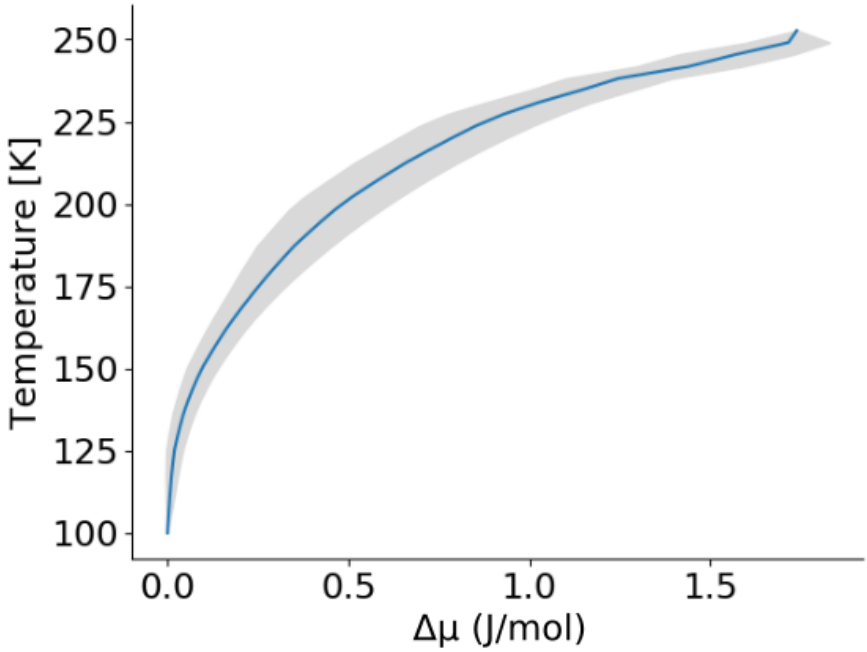



**Figure 6.16:** Average temperature as a function of $\Delta\mu$ for the phase boundary tracing. The grey area indicates the standard deviation.

In Figure 6.17 the average temperature as a function of magnesium concentration is shown together with its standard deviation for the two phase boundaries. Between these boundaries one can expect a co-existence of the pure magnesium phase, with aluminum atoms placed randomly around, and regions richer in aluminum corresponding to the phase $Mg_3Al$. On the right side of the blue line, in effect for very high concentrations of magnesium, the phase $Mg_3Al$ is not expected to be seen. It is not possible to draw any conclusions about what happens to the left of the orange line, in effect at magnesium concentrations lower than 75 %. This is because the algorithm used finds the boundary between pure magnesium and $Mg_3Al$ in the $\mu$-temperature plane corresponding to the two phase boundaries in the concentration-temperature plane. The phase boundaries observed here do not necessarily coincide with experimental phase diagrams. The reason

for this is that the phases considered here are restricted to a hcp lattice and could thus be metastable. As experimental phase diagrams only show stable and not metastable phases, the phase boundaries can differ significantly from what is observed here. In fact, it is possible that the phase $Mg_3Al$ does not show up at all. Another point is that the results are relying on a fixed lattice model whereas real materials vibrate and expand with temperature.

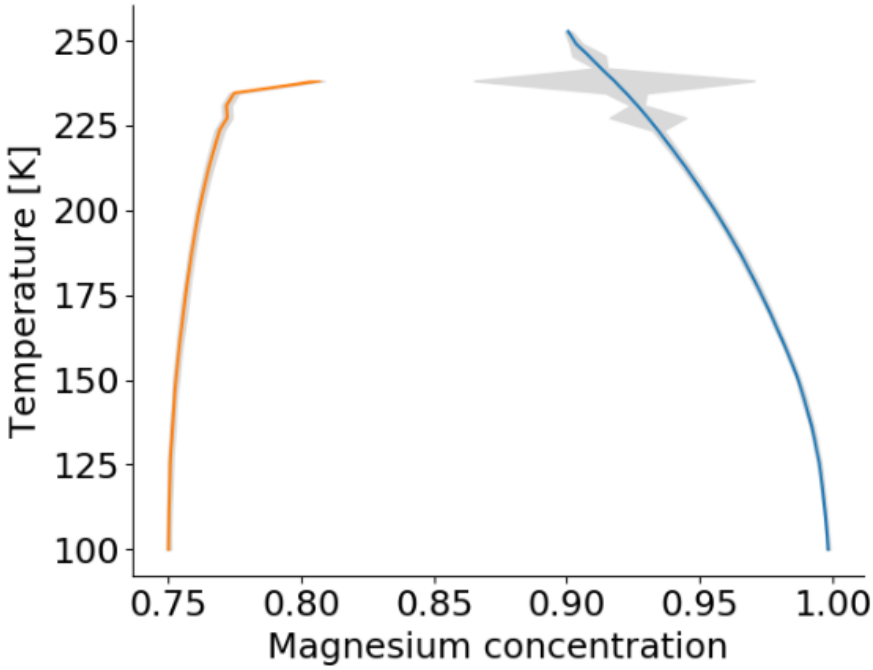



**Figure 6.17:** Average temperature as a function of magnesium concentration for the two phase boundaries from phase boundary tracing. The grey area indicates the standard deviation.

The phase boundaries are also plotted together with the results from the cooling down at constant chemical potential in the semi-grand canonical ensemble. This is shown in Figure 6.18 with red lines indicating the phase boundaries and the grey lines corresponding to those of cooling down in the semi-grand canonical ensemble. Here, one of the grey lines lie inside the phase boundaries, corresponding to a constant chemical potential of $\mu = -1.056$ eV/atom. This line is indicated by an arrow. This is not as expected, as all lines following a constant chemical potential should be located outside the phase boundaries. In other words, this chemical potential

is expected to give the phase $Mg_3Al$. In summary, the result indicates that the phase boundaries found numerically are correct except the one line inside the phase boundaries which should be expected to end up in the phase $Mg_3Al$. This discrepancy will be studied further in the following section.

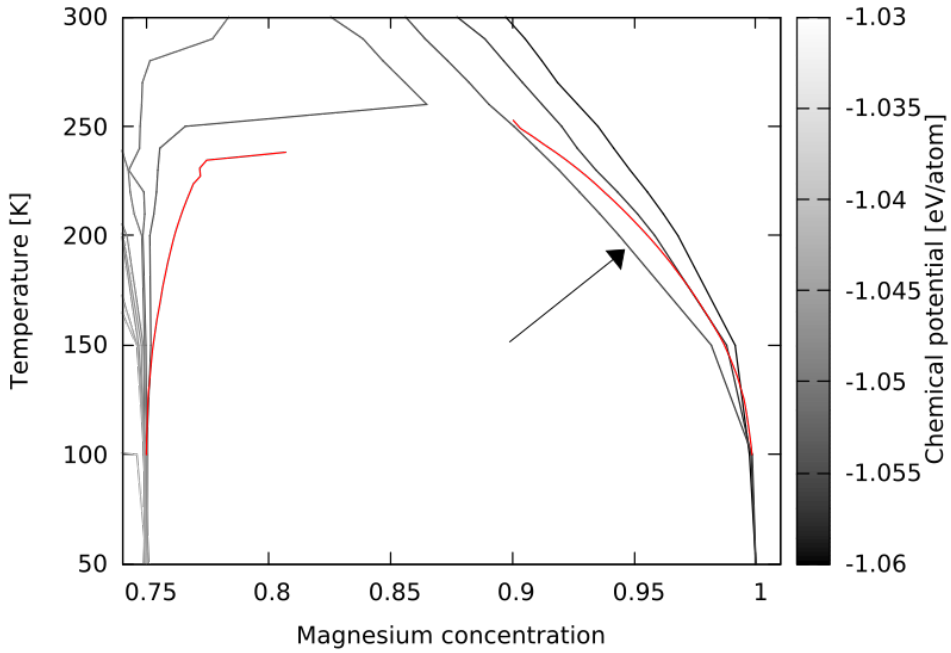



**Figure 6.18:** The phase boundaries, indicated by the red lines, shown together with some of the results from the cooling down at constant chemical potential in the semi-grand canonical ensemble.

### 6.6.4 Free energy barriers

By using the method of umbrella sampling, the free energy as a function of magnesium concentration is calculated via (5.26). This is done for several temperatures at a chemical potential of $\mu = -1.056$ eV/atom for a system of 1 000 atoms. The magnesium concentration ranges from 100 % to 70 % with 30 concentration windows. Each window contains 11 bins, except the first which contains 10 bins, to get a one-to-one correspondence between every bin and concentration while maintaining an overlap between each window. In each concentration window 2 000 000 semi-grand canonical

MC steps are performed. The results for $\beta\Delta G$ as a function of magnesium concentration are shown in Figure 6.19 and Figure 6.20 for high and low temperatures, respectively.

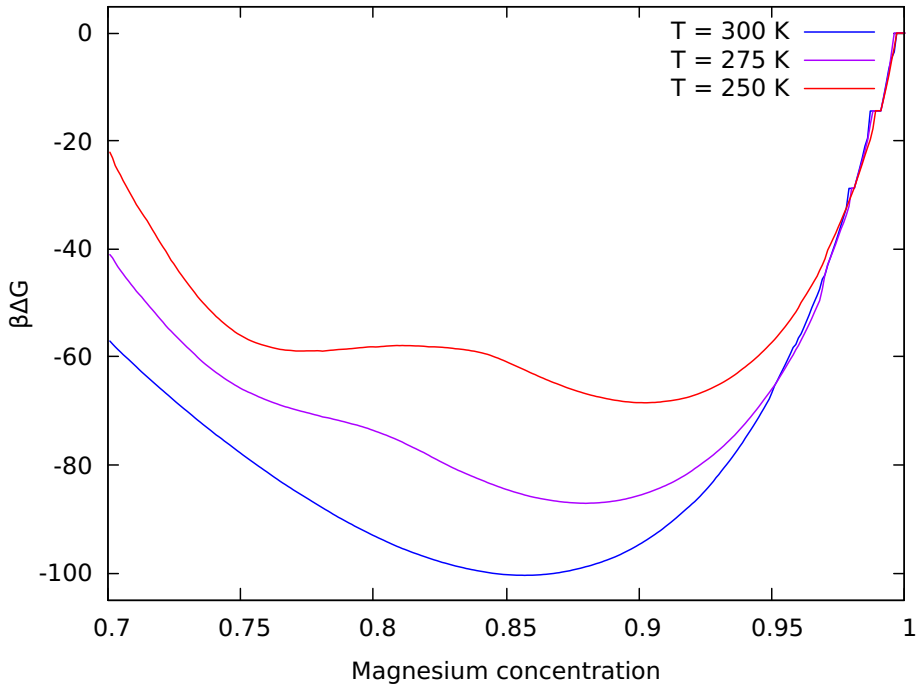



**Figure 6.19:** $\beta\Delta G$ as a function of magnesium concentration for $T$ = 300 K, $T$ = 275 K and $T$ = 250 K.

For high temperatures as shown in Figure 6.19, the global minimum in free energy agree well with the line at $\mu = -1.056$ eV/atom from Figure 6.18 just as expected. When going from $T$ = 275 K to $T$ = 250 K a local minimum in free energy appears at a lower concentration of magnesium and a higher value of $\beta\Delta G$ compared to the global minimum.

For lower temperatures as shown in Figure 6.20, a transition occurs. As the temperature reaches the range where clustering of $Mg_3Al$ happens, two separate minima in free energy are seen just as for $T$ = 250 K in Figure 6.19. The difference is that now, at temperatures of $T$ = 225 K and below the global minimum is located at a lower concentration of magnesium approaching 75 % as the temperature is reduced further. This global minimum is separated from a local minimum at higher concentrations of

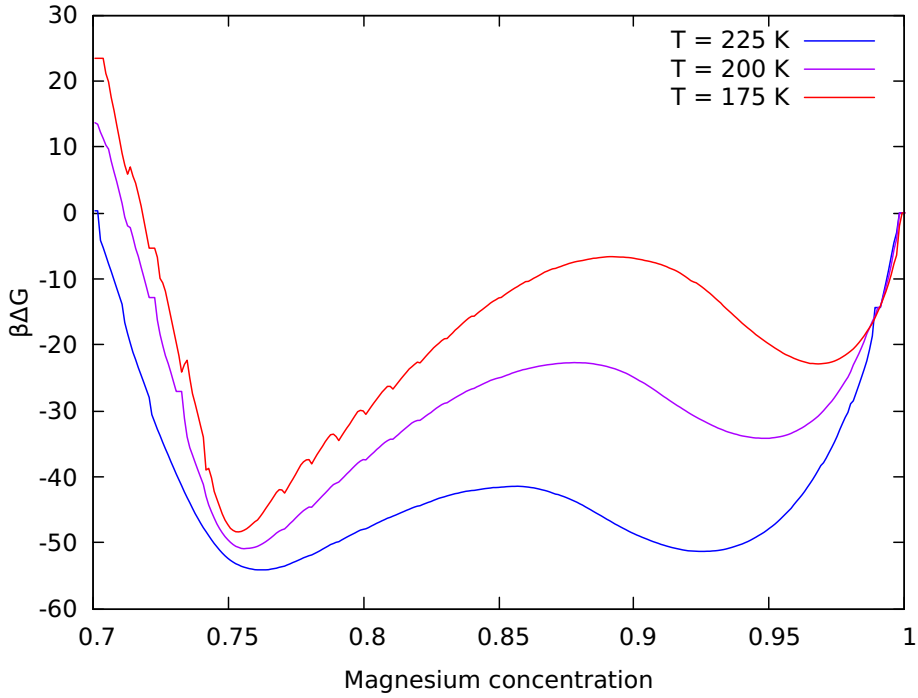



**Figure 6.20:** $\beta\Delta G$ as a function of magnesium concentration for $T$ = 225 K, $T$ = 200 K and $T$ = 175 K.

magnesium by a barrier in free energy. Thus, the global minimum in free energy as a function of concentration has shifted from high to lower concentrations of magnesium indicating that $Mg_3Al$ becomes the most energetically favourable phase for temperatures at $T$ = 225 K and below. This agrees well with the previous results and the global minimum at each temperature lies outside the phase boundaries shown in Figure 6.18, as they should. One can also note that the local minimum at higher concentrations of magnesium moves further and further towards pure magnesium as the temperature is decreased.

The reason for the system cooling down at $\mu = -1.056$ eV/atom not reaching the most energetically favourable phase, but instead approaching pure magnesium and staying inside the phase boundaries can be explained from these free energy barriers. At $T$ = 250 K the system is at a concentration of about 90 % magnesium as expected from the global minium in free energy at this temperature. When reaching a temperature of $T$ = 225

K the global minimum is located at about 75 % magnesium. The system is then instead of going towards $Mg_3Al$ trapped at the local minimum at a higher concentration of magnesium because of the free energy barrier between the two minima. In other words, the system actually favours a different phase, but it does not have the necessary energy or time to cross the free energy barrier and reach the phase $Mg_3Al$. The system stays at the local minimum in free energy at high concentrations of magnesium as the temperature is decreased further as lowering the temperature only decreases the energy of the system, making it even harder to cross the free energy barrier. Moreover, it is these local minima that correspond to the line lying within the phase boundaries as seen when comparing with Figure 6.18. Thus, the discrepancy of the previous section has been explained and the phase boundaries found previously have one less reason to be doubted.

An important point is that redoing the cooling down in the semi-grand canonical ensemble at $\mu = -1.056$ eV/atom the system could might as well go to the correct phase $Mg_3Al$ instead of being trapped at the local minimum in free energy at higher concentrations of magnesium. It could be an artifact of the discretization of the temperature range and for further calculations the steps in temperature should be smaller. This is especially important in the range where the transition from pure magnesium to $Mg_3Al$ occurs. Another point is that these results have not been seen in previous calculations of free energy in this thesis. This is because before the property considered was the free energy of formation and not the free energy. Finally, as the temperature is lowered one sees that $\beta\Delta G$ becomes a less smooth function of the magnesium concentration. This is because at lower temperatures the probability of accepting a new state after a MC step decreases as evident from (5.2). Having less MC steps accepted the statistical results from the MC simulation get worse. Thus, the number of MC steps needed for getting a smoothly varying $\beta\Delta G$ increases as the temperature is lowered.

The quantity $\beta\Delta G$ is also calculated along the phase boundary between pure magnesium and $Mg_3Al$ using some of the values for $\mu$ and $T$ from the results shown in Figure 6.16. The results using 2 000 000 semi-grand canonical MC steps in each of in total 30 concentration windows are shown in Figure 6.21.

Along the boundary in the $\mu$-$T$ plane separating $Mg_3Al$ and pure magnesium the two phases should coexist. In other words, the free energy as a function of magnesium concentration should have two local minima at dif-

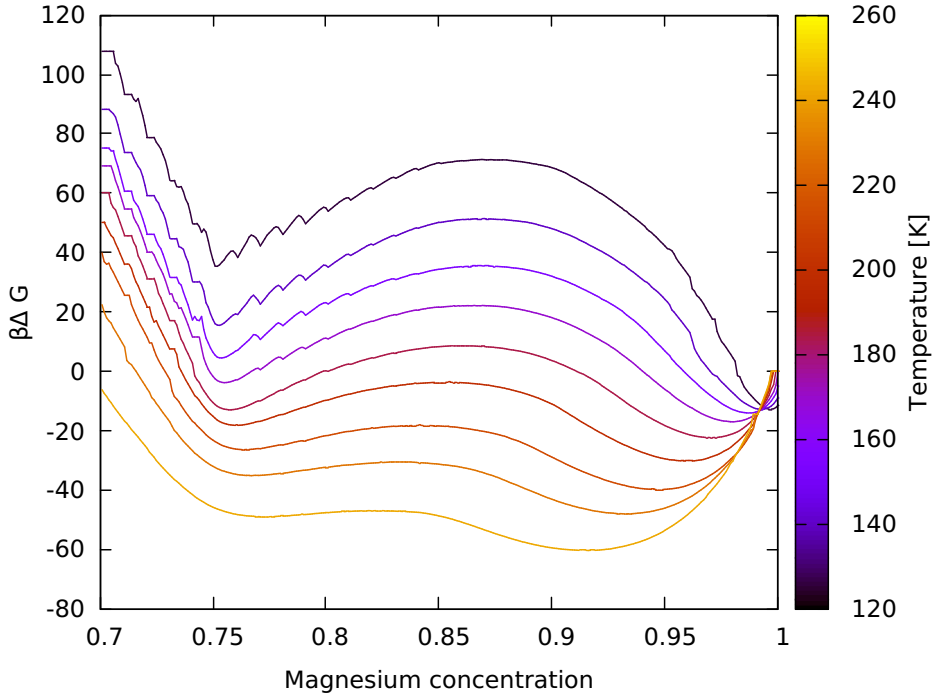



**Figure 6.21:** $\beta\Delta G$ as a function of magnesium concentration along the phase boundaries of $Mg_3Al$.

ferent concentrations. One of them should move towards 75 % magnesium and the other towards 100 % magnesium as the temperature is lowered. This is also what is seen in Figure 6.21. However, the two local minima should be located at equal values of $\beta\Delta G$ as along the phase boundary both phases are equally favoured. This is not the case here, where the minimum at high concentrations of magnesium is lower, corresponding to pure magnesium being preferred and indicating that there may be a systematic error when calculating phase boundaries due to the chosen precision.

As explained earlier, (5.19) is used to obtain the phase boundaries. This equation is hard to solve numerically and the calculations are computationally expensive resulting in a precision of $10^{-3}$. Recalculating $\beta\Delta G$ for a given temperature and varying the chemical potential until the local minimum at approximately 75 % magnesium becomes the global one the error in $\mu$ when calculating the phase boundaries at the precision of $10^{-3}$ can be estimated. This is done for $T$ = 212.72 K, corresponding to the third

line from below in Figure 6.21 where the calculated chemical potential is $\mu \simeq$ -1.0576 eV/atom. The result is shown in Figure 6.22 using 1 000 000 semi-grand canonical MC steps for $\mu$ = -1.0569 eV/atom and $\mu$ = -1.0568 eV/atom.

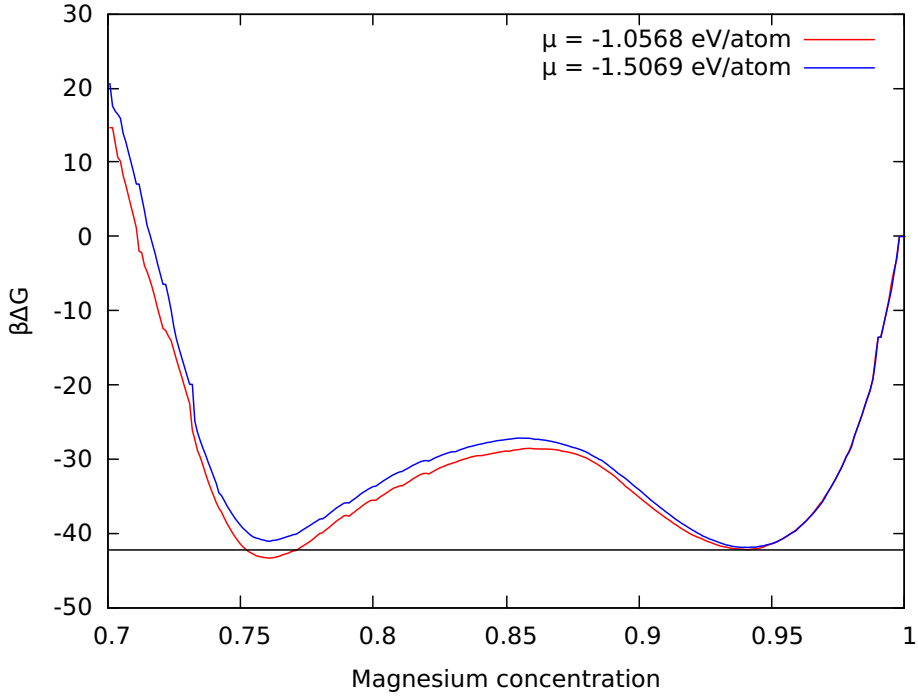


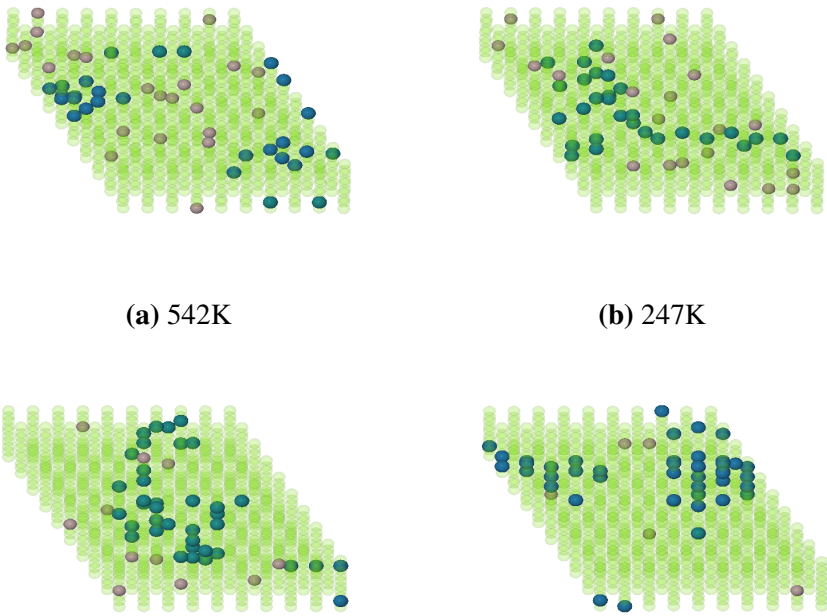

**Figure 6.22:** $\beta\Delta G$ as a function of magnesium concentration at $T$ = 212.72 K with $\mu$ = -1.0569 eV/atom and $\mu$ = -1.0568 eV/atom.

As seen here, changing the chemical potential to $\mu$ = -1.0569 eV/atom makes the minimum at approximately 75 % magnesium become the global one. Thus, the correct chemical potential along the phase boundary is located somewhere between $\mu$ = -1.0569 eV/atom and $\mu$ = -1.0568 eV/atom. The estimate of the error in $\mu$ when calculating phase boundaries via (5.19) at a precision of $10^{-3}$ thus becomes 0.07 %. Although the difference in $\beta\Delta G$ at the calculated chemical potential for the phase boundaries differ by certain amount, the error is relatively small. The obtained phase boundaries are thus believed to be correctly calculated.

### 6.6.5 Visualization of cluster formation

As shown above, the $Mg_3Al$ phase starts to form at a critical temperature range between 200 K and 300 K in a hcp lattice. It is helpful to visualize this process of clustering which will be done here. This is done by performing a MC simulation in the canonical ensemble for cooling down a system of 1000 atoms in a hcp lattice. The magnesium concentration is set to 95 % such that it is impossible for the whole system to transform into $Mg_3Al$. At each temperature 100 000 MC steps are performed instead of setting a precision as this simulation just serves as an illustration of the clustering process. At every 5 steps the largest connected cluster of aluminum atoms is found. After performing all MC steps at a temperature the largest connected cluster found at this temperature is stored, with periodic boundary conditions in mind. The results for four different temperatures are shown in Figure 6.23. The temperatures chosen are one temperature lying well above (542 K), two temperatures inside (247 K and 210 K) and one temperature well below (136 K) the critical temperature range where clustering starts happening. The green, almost transparent atoms are magnesium atoms. The grey and dark blue atoms are aluminum atoms where the latter indicates that they belong to the largest connected cluster of aluminum atoms.

As seen here, at 542 K the aluminum atoms are placed quite randomly around in the lattice. The largest connected cluster contains approximately half of the available aluminum atoms. Lowering the temperature a transition occurs and more and more aluminum atoms become clustered together to form the phase $Mg_3Al$, just as expected for 247 K and 210 K. The formation of $Mg_3Al$ can be seen by the majority of aluminum atoms clustering together at distances corresponding to the double of the nearest neighbour distance just as in Figure 6.12. Finally, at 136 K almost all the aluminum atoms are connected in one large cluster also with distances between aluminum atoms equal to the double of a nearest neighbour distance. Thus, $Mg_3Al$ has been formed.



**(a)** 542K **(b)** 247K

**(c)** 210K **(d)** 136K

**Figure 6.23:** Largest cluster of aluminum at four different temperatures during a Monte Carlo simulation in the canonical ensemble, highlighted by blue atoms.

Chapter 7

# Conclusion

To summarize, aluminum and magnesium alloys on a hcp lattice rich in magnesium have been studied. The convergence of total energy calculations is ensured to obtain a reasonable set of input parameters for different system sizes in DFT calculations. Where comparisons are possible the results agree well with either theory, experimental results or previously performed calculations. An example of this is the calculation of the lattice parameters for aluminum and magnesium. These calculations also show that the PBE functional performs better than the BLYP and LDA functionals, at least in these cases. Also, the calculated formation energy of a vacancy in bulk aluminum and magnesium are as mentioned earlier in good agreement with previous results. Thus, DFT calculations with the PBE functional are believed to work well for studying electronic and structural properties of metals and alloys and the set of input parameters obtained can be considered as sufficiently accurate.

A CE model is developed for aluminum and magnesium on a hcp lattice from 25 % to 100 % magnesium by fitting ECI to the total energies of DFT calculations resulting in a CV score of 3.18 meV/atom. By interpreting the ECI the model is seen to favour equal atoms in a second nearest neighbour distance. Moreover, clusters of three atoms all in a nearest neighbour distance are favoured with either three aluminum or one aluminum and two magnesium atoms. The CE model predicts a new phase $Mg_3Al$ which starts to form somewhere below 300 K. This is seen from the calculated minima in enthalpy of formation and free energy of formation. The boundaries of this phase are also found and they agree well with the canonical and semi-

grand canonical MC simulations performed in this thesis, apart from the one simulation with a fixed chemical potential of $\mu = -1.056$ eV/atom in the semi-grand canonical ensemble. However, this discrepancy is explained by calculating the free energy barrier separating $Mg_3Al$ and pure magnesium. The new phase $Mg_3Al$ could be hard to find experimentally considering that it starts to form below 300 K. One could imagine a scenario where one heats up an alloy of aluminum and magnesium to a very high temperature before cooling it down below 300 K almost instantly. This would leave the alloy unstable with many defects where $Mg_3Al$ could start to form spontaneously, easier than if the alloy has an ideal hcp lattice. Still, it is no guarantee that $Mg_3Al$ will appear as this thesis only examines a fixed hcp lattice and the phonon and thermal expansion contribution are not considered. Other configurations in another lattice could be more energetically favourable, in which case $Mg_3Al$ is just a metastable phase.

Finally, the results show that the CE method combined with DFT calculations and MC simulations makes a powerful tool for examining alloys on different lattices. It should be noted that for other types of lattices the CE model needs to be trained all over again. The estimate of the error in $\mu$ when calculating phase boundaries is small at a value of 0.07 %.

Chapter 8

# Future work

As mentioned, the phase $Mg_3Al$ could be metastable and only appearing when the hcp lattice is the only lattice considered. A natural way forward would thus be to study alloys of aluminum and magnesium on other lattices, for example the fcc lattice. As pure aluminum forms in a fcc lattice this would be expected to cover more aluminum rich parts of the concentration range. Other lattices should also be examined as aluminum and magnesium combined could have stable phases on other types of lattices than what is found for their pure phases. One can in principle obtain information about the whole phase diagram by considering the different lattices and combining this knowledge. It is also possible to study metastable phases as here to get information about nucleation and precipitation. Another possibility is to introduce a third component in the alloy.

Experimentally, one could study whether $Mg_3Al$ is stable or not by heating an aluminum and magnesium alloy up and rapidly cooling it down as mentioned earlier. If $Mg_3Al$ does not start to form after this process, this would implicate that it is only stable on a hcp lattice.

# Bibliography


[1] A. I. Taub and A. A. Luo, "Advanced lightweight materials and manufacturing processes for automotive applications," *MRS Bulletin*, vol. 40, no. 12, pp. 1045–1054, 2015.

[2] G. Lütjering, "Influence of processing on microstructure and mechanical properties of ($\alpha$+$\beta$) titanium alloys," *Materials Science and Engineering: A*, vol. 243, no. 1, pp. 32–45, 1998.

[3] A. Pribram-Jones, D. A. Gross, and K. Burke, "DFT: A theory full of holes?" *Annual Review of Physical Chemistry*, vol. 66, no. 1, pp. 283–304, 2015.

[4] Q. Wu, B. He, T. Song, J. Gao, and S. Shi, "Cluster expansion method and its application in computational materials science," *Computational Materials Science*, vol. 125, no. 0927-0256, pp. 243 – 254, 2016.

[5] J. Teeriniemi, J. Huisman, P. Taskinen, and K. Laasonen, "First-principles modelling of solid Ni–Rh (nickel-rhodium) alloys," vol. 652, no. 1, pp. 371–378, 2015.

[6] J. Teeriniemi, P. Taskinen, and K. Laasonen, "First-principles investigation of the Cu–Ni, Cu–Pd, and Ni–Pd binary alloy systems," *Intermetallics*, vol. 57, no. 1, pp. 41–50, 2015.

[7] C. Kittel, *Introduction to Solid State Physics*, 8th ed. New York: John Wiley & Sons, Inc, 2005.

[8] W. Pauli, "Nobel lecture: Exclusion principle and quantum mechanics," *Nobel Lectures, Physics*, pp. 1942–1962, 1946.

[9] P. F. Lang and B. C. Smith, "Metallic structure and bonding," *World Journal of Chemical Education*, vol. 3, no. 2, pp. 30–35, 2015.

[10] K. Solevåg-Hoti, "Formation of intermediate bands in silicon by transition metal doping - A DFT/DFTB computational study," Master's thesis, Norwegian University of Science and Technology, Trondheim, 2016.

[11] D. S. Sholl and J. A. Steckel, *Density functional theory - A practical introduction*. Hoboken: John Wiley & Sons, Inc, 2009.

[12] G. F. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid*. New York: Cambridge University Press, 2005.

[13] M. Born and R. Oppenheimer, "Zur quantentheorie der molekeln," *Annalen der Physik*, vol. 389, no. 20, pp. 457–484, 1927.

[14] E. Schrödinger, "Quantisierung als eigenwertproblem; von Erwin Schrödinger," *Annalen der Physik*, vol. 384, no. 4, pp. 361–377, 1926.

[15] D. R. Hartree, "The wave mechanics of an atom with a non-coulomb central field. Part I. Theory and methods," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 24, pp. 89–110, 1928.

[16] J. C. Slater, "The theory of complex spectra," *Physical Review*, vol. 34, no. 2, pp. 1293–1322, 1929.

[17] V. Fock, "Näherungsmethode zur lösung des quantenmechanischen mehrkörperproblems," *Zeitschrift für Physik*, vol. 61, no. 1-2, pp. 126–148, 1930.

[18] H. S. Friedrich, *Theoretical Atomic Physics*, 2nd ed. Berlin: Springer, 1998.

[19] E. H. Lieb, "Density functionals for coulomb systems," *International Journal of Quantum Chemistry*, vol. 24, no. 3, pp. 243–277, 1983.

[20] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry*. New York: McGraw-Hill, 1989.

[21] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Physical Review*, vol. 136, no. 3B, pp. B864–B871, 1964.

[22] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Physical Review*, vol. 140, no. 4A, pp. A1133–A1138, 1965.

[23] B. Santra, "Density-functional theory exchange-correlation functionals for hydrogen bonds in water," Ph.D. dissertation, Technischen Universität Berlin, Berlin, 2010.

[24] P. Haas, F. Tran, and P. Blaha, "Calculation of the lattice constant of solids with semilocal functionals," *Physical Review B*, vol. 79, no. 8, p. 085104, 2009.

[25] G. Ortiz and P. Ballone, "Pseudopotentials for non-local-density functionals," *Physical Review B*, vol. 43, no. 8, pp. 6376–6387, 1991.

[26] I. Y. Zhang and X. Xu, *A New-Generation Density Functional: Towards Chemical Accuracy for Chemistry of Main Group Elements*. Berlin: Springer Science & Business Media, 2013.

[27] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," *Physical Review A*, vol. 38, no. 6, pp. 3098–3100, 1988.

[28] C. Lee, W. Yang, and R. G. Parr, "Development of the colle-salvetti correlation-energy formula into a functional of the electron density," *Physical Review B*, vol. 37, no. 2, pp. 785–789, 1988.

[29] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Physical Review Letters*, vol. 77, no. 18, pp. 3865–3868, 1996.

[30] L. He, F. Liu, G. Hautier, M. J. T. Oliveira, M. A. L. Marques, F. D. Vila, J. J. Rehr, G.-M. Rignanese, and A. Zhou, "Accuracy of generalized gradient approximation functionals for density functional perturbation theory calculations," *Physical Review B*, vol. 89, no. 6, p. 064305, 2014.

[31] J. M. Sanchez, "Cluster expansions and the configurational energy of alloys," *Physical Review B*, vol. 48, no. 18, pp. 14 013–14 015, 1993.

[32] J. Sanchez, F. Ducastelle, and D. Gratias, "Generalized cluster description of multicomponent systems," *Physica A: Statistical Mechanics and its Applications*, vol. 128, no. 1, pp. 334–350, 1984.

[33] A. van de Walle and G. Ceder, "Automating first-principles phase diagram calculations," *Journal of Phase Equilibria*, vol. 23, no. 4, p. 348, 2002.

[34] T. P. Ryan, *Modern Regression Methods*, 2nd ed. New York: Wiley-Interscience, 2008.

[35] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[36] G. van Rossum, *Python tutorial, Technical Report CS-R9526*, Centrum voor Wiskunde en Informatica, Amsterdam, 1995.

[37] J. J. Mortensen, L. B. Hansen, and K. W. Jacobsen, "Real-space grid implementation of the projector augmented wave method," *Physical Review B*, vol. 71, p. 035109, 2005.

[38] P. E. Blöchl, "Projector augmented-wave method," *Physical Review B*, vol. 50, no. 24, pp. 17 953–17 979, 1994.

[39] S. R. Bahn and K. W. Jacobsen, "An object-oriented scripting interface to a legacy electronic structure code," *Computing in Science & Engineering*, vol. 4, no. 3, pp. 56–66, 2002.

[40] H. J. Monkhorst and J. D. Pack, "Special points for brillouin-zone integrations," *Physical Review B*, vol. 13, no. 12, pp. 5188–5192, 1976.

[41] C. G. Broyden, "The convergence of a class of double-rank minimization algorithms 1. General considerations," *IMA Journal of Applied Mathematics*, vol. 6, no. 1, pp. 76–90, 1970.

[42] D. Packwood, J. Kermode, L. Mones, N. Bernstein, J. Woolley, N. Gould, C. Ortner, and G. Csányi, "A universal preconditioner

for simulating condensed phase materials," *The Journal of Chemical Physics*, vol. 144, no. 16, p. 164109, 2016.

[43] J. Chang, *A Python library for working with atoms*, 2018 [accessed May 22, 2018]. [Online]. Available: https://gitlab.com/jinchang/ase

[44] A. Seko, Y. Koyama, and I. Tanaka, "Cluster expansion method for multicomponent systems based on optimal selection of structures for density-functional theory calculations," *Physical Review B*, vol. 80, no. 16, p. 165122, 2009.

[45] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[46] W. K. Hastings, "Monte Carlo sampling methods using markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.

[47] J. W. Gibbs, *Elementary Principles in Statistical Mechanics*. New York: Charles Scribner's Sons, 1902.

[48] J. G. Briano and E. D. Glandt, "Statistical thermodynamics of polydisperse fluids," *The Journal of Chemical Physics*, vol. 80, no. 7, pp. 3336–3343, 1984.

[49] D. Kleiven, *Monte Carlo package targeted at systems studied with the Cluster Expansion*, 2018 [accessed May 21, 2018]. [Online]. Available: https://github.com/davidkleiven/CEMC

[50] A. van de Walle and M. Asta, "Self-driven lattice-model Monte Carlo simulations of alloy thermodynamic properties and phase diagrams," *Modelling and Simulation in Materials Science and Engineering*, vol. 10, no. 5, p. 521, 2002.

[51] J. O. Andersen, *Introduction to Statistical Mechanics*, 1st ed. Trondheim: Tapir Academic Press, 2012.

[52] D. Chandler, *Introduction to Modern Statistical Mechanics*, 1st ed. New York: Oxford University Press, 1987.

[53] L. M. Almeida, C. Fiolhais, and M. Causà, "Properties of simple metals beyond the local density approximation of density functional theory," *International Journal of Quantum Chemistry*, vol. 91, no. 2, pp. 224–229, 2003.

[54] G. I. Csonka, J. P. Perdew, A. Ruzsinszky, P. H. T. Philipsen, S. Lebègue, J. Paier, O. A. Vydrov, and J. G. Ángyán, "Assessing the performance of recent density functionals for bulk solids," *Physical Review B*, vol. 79, no. 15, p. 155107, 2009.

[55] M. Pozzo and D. Alfè, "Structural properties and enthalpy of formation of magnesium hydride from quantum Monte Carlo calculations," *Physical Review B*, vol. 77, no. 10, p. 104103, 2008.

[56] K. G. Lynn and P. J. Schultz, "Vacancy formation energy measurements in single crystal aluminum using a variable-energy positron beam," *Applied Physics A*, vol. 37, no. 1, pp. 31–36, 1985.

[57] C. Mairy, J. Hillairet, and D. Schumacher, "Energie de formation et concentration d'équilibre des lacunes dans le magnésium," *Acta Metallurgica*, vol. 15, no. 7, pp. 1258 – 1261, 1967.

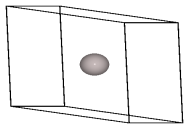
[58] N. J. Salkind, *Encyclopedia of Research Design*. Thousand Oaks: SAGE Publications, Inc, 2010.

[59] J. L. Murray, "The Al–Mg (aluminum-magnesium) system," *Journal of Phase Equilibria*, vol. 3, no. 1, p. 60, 1982.

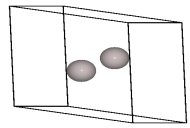
# Appendix A

**Table 8.1:** Numerical values of the effective cluster interactions from the final cluster expansion.

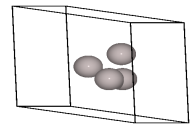
| ECI | Value [eV/atom] |
|---|---|
| c0 | $-2.6357104336220036$ |
| c1_0 | $-1.0477578743973592$ |
| c2_1000_1_00 | $1.9704034169076734 \times 10^{-5}$ |
| c2_1414_1_00 | $-0.0086207715025969599$ |
| c3_1000_1_000 | $-0.0068774873140020742$ |
| c3_1414_1_000 | $0.01663724158871387$ |
| c4_1000_1_0000 | $0.001447359159530979$ |
| c4_1414_1_0000 | $0.0051126797326320134$ |
| c4_1414_2_0000 | $0.00034941577797427485$ |



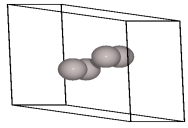
**(a)** c1_0.



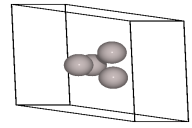
**(b)** c2_1000_1_00.



**(c)** c4_1000_1_0000.



**(d)** c4_1414_1_0000.



**(e)** c4_1414_2_0000.

**Figure 8.1:** The remaining clusters of the effective cluster interactions.