A Genetic Algorithm for Structure Prediction of Magnetic Materials

Edward James Higgins

Doctor of Philosophy

University of York

Physics

October 2018

Abstract

When considering global optimisation of magnetic crystal structures, it is important to consider both the atomic and spin degrees of freedom. This thesis presents a novel genetic algorithm for simultaneously optimising the magnetic and crystal structures of materials. The algorithm was first tested on a new magnetic interatomic potential presented in the thesis, and was shown to be capable of finding the correct atomic and magnetic structure. The algorithm was then used to study mixing the NiO(111)/MgO(111) interface, where the process behind the mixing was unknown. Results from this study suggest that mixing is driven by energetics of the system, rather than kinetic processes. Finally, the interface between the Heusler alloy CFAS and *n*-doped Ge, where experimental observations suggested an unknown interface phase, was studied. This work proposed the half Heusler structure for this phase, and predicted this to have unfavourable electronic properties.

Contents

List of Figures

Acknowledgements

I would like to thank everyone who has supported me through these past few years. Firstly, I would like to give a huge thank you to Matt Probert and Phil Hasnip for the time and effort they put into getting me through the PhD, and stimulating my interest in what I was doing.

Secondly, I would like to thank the people of N/102: Jack Shepherd, Aaron Hopkinson, Jacob Wilkins, Matt Hodgson and others who have come and gone throughout my time there. They are the people who made the office such an enjoyable place to be. I would especially like to thank Jack and Aaron for keeping me going when things got tough.

Finally I would like to thank Danielle Soper, my parents and the rest of my family, who always believed in me no matter what.

Declaration

I declare that the work presented in this thesis, except where otherwise stated, is based on my own research and has not been submitted previously for a degree in this or any other university. Parts of the work reported in this thesis have been published in:

N.L. Abraham and M.I.J. Probert. Erratum: Improved real-space genetic algorithm for crystal structure and polymorph prediction [Phys. Rev. B 77, 134117 (2008)]. *Phys. Rev. B*, 94:059904, Aug 2016

Zlatko Nedelkoski, Phil J Hasnip, Ana M Sanchez, Balati Kuerbanjiang, Edward Higgins, Mikihiko Oogane, Atsufumi Hirohata, Gavin R Bell and Vlado K Lazarov. The effect of atomic structure on interface spin-polarization of half-metallic spin valves: Co2MnSi/Ag epitaxial interfaces. *Applied Physics Letters*, 107(21):212303, 2015.

Chapter 1

Introduction

For thousands of years, humankind has relied on its ability to discover new materials and use them for ever more advanced technologies. Indeed many eras, such as the Iron and Bronze Ages, are defined by the materials they developed. Even in the modern day, discovery and subsequent development of semiconducting and magnetic materials have allowed computers to grow ever more powerful, changing the world dramatically over the past few decades.

Society's ability to develop new materials has come from a combination of repeated experimentation and a growth of knowledge of how materials are likely to behave. By noticing patterns in behaviour between experiments, theories can be developed about what is going on, and subsequently used to suggest directions for future experimentation. These theories can be anything from crude, qualitative approximations based on a small number of experiments, right up to rigorous mathematical descriptions based on many years of testing. As should be expected, the latter tend to be more accurate than the former, but take longer to develop. In the right amounts, this range of theories can prove invaluable to aid the discovery of new materials.

These days there are a range of theories on how materials behave, spanning classical and quantum physics and over a range of length and time scales [5]. Some methods, such as Quantum Monte Carlo, Hartree Fock, and Density Functional Theory, aim to find approximate solutions to Schrödinger's equation for the electrons in materials, allowing accurate calculation of their electronic properties. Others, such as Lennard

Jones[6] or EAM [7, 8] , classically model the interaction between atoms, including the electrons only implicitly. Similarly, course grained models such as the OxDNA model for DNA [9], the interaction between base units consisting of multiple atoms. These methods are able to simulate a larger number of atoms than atomistic methods, assuming they remain in the configuration defined by the course graining. Finally, continuum methods simulate materials such as concrete [10] as a continuous medium, with defined physical properties. This allows macro-scale simulations to be performed.

Using computers to model real-world systems using these theories, it is possible to predict, with remarkable accuracy, the behaviour of new materials without having to go through the trouble of fabricating them experimentally. In addition to this, knowledge gained from the details of a simulation can provide important insights, allowing improved fabrication and use of these materials [11].

1.1 Magnetic materials

One class of material of particular interest is that of magnetic materials. This includes conventional ferromagnets along with more exotic structures, such as antiferromagnets, ferrimagnets and spin glasses. These materials have been at the heart of many important technologies over the past century and continue to play an important role in current and developing technologies.

A widely used type of magnetic material is the ferromagnet. For many centuries, people have been using the ferromagnetic properties of iron and its compounds for navigation [12]. More recently, ferromagnets have provided a mechanism to convert electrical signals into mechanical movement and vice versa. In particular, motors and generators are widely used for providing and harnessing rotary motion, and loudspeakers and microphones use magnets to convert between oscillatory sound waves and electrical signals. Developments of new more powerful or robust ferromagnets are likely to benefit industries which rely on these technologies. In addition, many of the strong permanent magnets in use today rely on rare-earth elements [13] and there is concern over the sustainability of these elements' availability [14]. Therefore, developing ferromagnets made from more readily available materials is desirable.

Computing has also benefited greatly from magnetic materials. For much of the history of computers, magnetic media such as tapes, hard and floppy disks have used magnetic effects to store digital data without the need for a power source [15]. Even now, there is a good deal of research into novel materials such as for magnetic storage devices such as ferroelectrics for magnetic RAM devices [16] and FePt based materials for heat assisted magnetic recording (HAMR)[17]. In addition, the developing field of spintronics relies heavily on magnetic materials [18, 19]. For example, spin valves are typically made of two ferromagnetic layers such as Heusler alloys with a thin non-magnetic layer in between [19].

Given the prevalence of applications for magnetic materials in the modern world, it is very likely that the development of such materials will be of benefit. Using theoretical models to aid in the development of these materials can enable such materials to be developed faster and more cost effectively, as well as providing a deeper insight into how they work.

1.2 Thesis overview

This thesis outlines the development of an algorithm for prediction a materials crystal and magnetic structure, given only the constituent atoms. Chapters 2 to 4 outline the underlying theory and existing work involved in structure prediction. In particular chapter 4 describes a genetic algorithm (GA) originally designed by Abraham and Probert, that provides a starting point for the work outlined in the rest of the thesis, along with some improvements made over the course of this work. Chapter 5 describes the extension of the GA to optimise magnetic structures concurrently with the crystal structures.

Chapters 6 to 8 demonstrate the application of this algorithm on a range of systems. Firstly, chapter 6 describes and investigation into an idealised potential based on the Lennard-Jones model. Chapter 7 then uses the algorithm to investigate the origins of disorder nickel oxide/magnesium oxide interfaces. Finally, chapter 8 uses the algorithm to investigate the structure and effects of structures observed growing at the interface between Heusler alloys and germanium.

Chapter 9 concludes the thesis, along with suggesting possible extensions and improvements that could be made to the algorithm in the future.

Chapter 2

Materials simulation

2.1 Introduction

In the context of materials simulation, there are a number of different levels of theory that can be chosen. These range from continuum models, capable of simulating macroscopic sized objects [10], right down to quantum mechanical simulations, capable of simulating a few individual particles [20, 21], as demonstrated in figure 2.1. This chapter reviews some of the commonly used methods for simulating materials, including *ab initio* and atomistic simulations used in the rest of this work.

2.2 Potential energy surfaces

When simulating materials at the atomic level, a natural description is that of the positions of the atomic nuclei. For a system of N atoms with positions $\mathbf{R}_1 \ldots \mathbf{R}_N$, a $3N$ dimensional PES (Potential Energy Surface) can be defined:

$$E = E(\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_N) \tag{2.1}$$

Ignoring thermal effects, the global minimum of the PES corresponds to the stable configuration of these atoms, and other local minima correspond to metastable

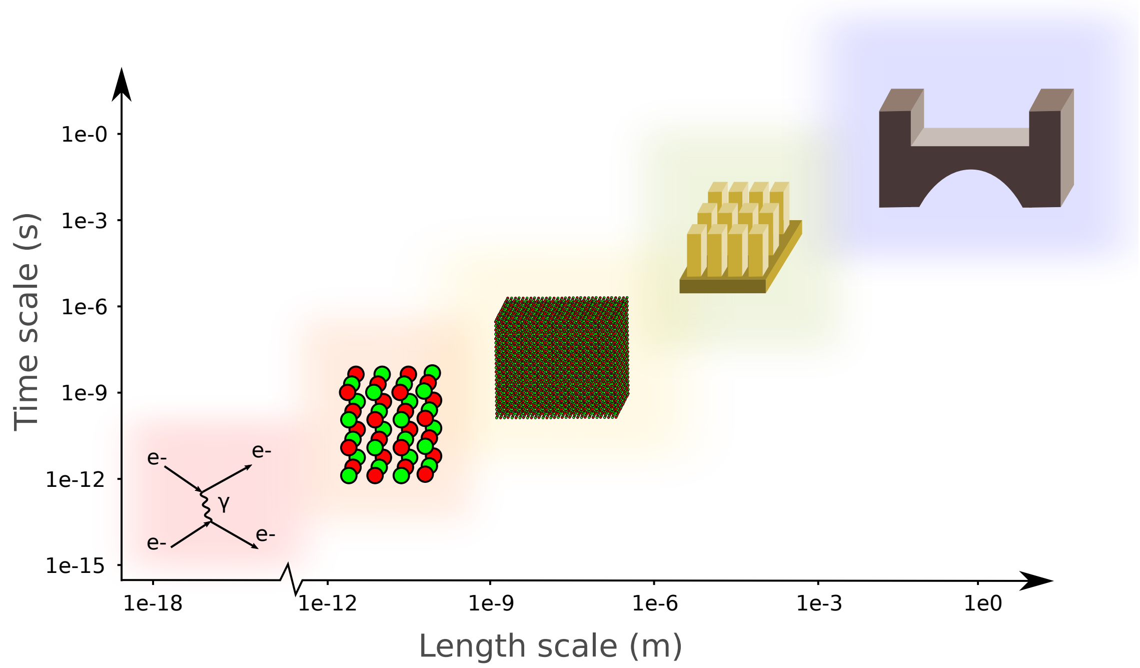

Figure 2.1: The range of length and time scales for materials simulations. From left to right: Quantum many-body, *ab initio*, atomistic, course grained and continuum simulations.

configurations. Typically, the stable configurations are the ones that will likely be found in nature; however if the local minima are in sufficiently deep wells, they can be considered thermodynamically stable. Figure 2.2 shows the slice of the PES for the H_2O molecule, where all dimensions are fixed except for the x-position of the oxygen atom. In this diagram, local and global minima can be seen when the oxygen is near the hydrogen atoms.

2.3 Periodic systems

Since many properties of a material are determined by the bulk of the material and not its surface, it would be useful to model this. However, it can take several thousand atoms or more to reproduce this behaviour. For crystals, this can be avoided by modelling a small repeating unit with PBCs (Periodic Boundary Conditions), as shown in figure 2.3.

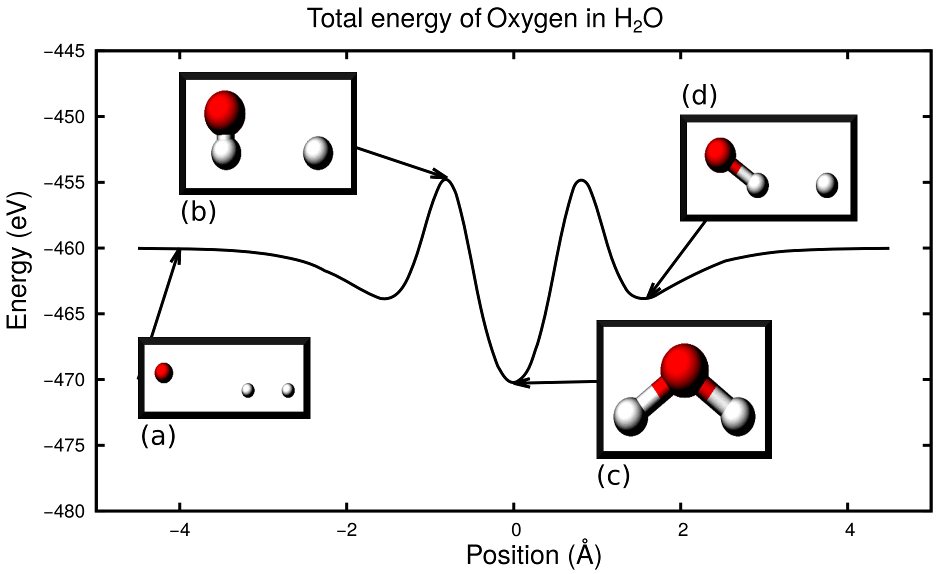

Figure 2.2: A 1-D Slice of the PES for H_2O from DFT simulations, showing: (a) The oxygen atom disassociated from the two hydrogen atoms, (b) A Maximum, where the hydrogen and oxygen atoms are at their closest, (c) The global minimum: a water molecule, and (d) A local minimum, with a hydrogen atom and a OH dimer

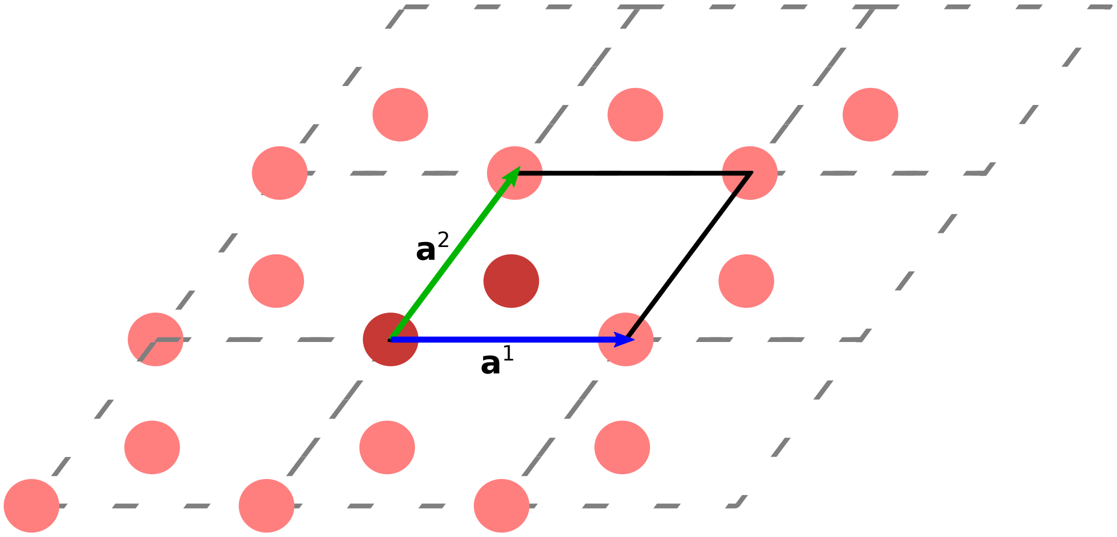

Figure 2.3: A 2D repeating system of 2 atoms forming a hexagonal lattice structure. The atoms (dark red circles), lattice vectors (blue and green lines) and unit cell (black solid lines) are shown, along with a number of periodic images (light red circles and grey dashed lines).

A crystal can be defined in terms of some *unit cell* vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ and the atoms within that cell, $\{\mathbf{R}_I\}, I = 1 \dots N$. Since this unit cell repeats indefinitely in all directions, the energy is invariant to translations of multiples of the lattice vectors and the PES has the restriction:

$$E(\mathbf{R}_1, \dots, \mathbf{R}_I, \dots, \mathbf{R}_N) = E(\mathbf{R}_1, \dots, \mathbf{R}_I + (l_1\mathbf{a}_1 + l_2\mathbf{a}_2 + l_3\mathbf{a}_3), \dots, \mathbf{R}_N) \tag{2.2}$$

Where l_1, l_2 and l_3 are integers.

More generally, systems can be defined as being periodic in any number of directions, for example, a graphene sheet would only repeat in 2 dimensions.

2.4 Density functional theory

Primarily, the crystal structure of a material is defined by the charge of the atomic nuclei and the nature of the electron density [22]. To demonstrate this, consider the Schrödinger equation for a set of atoms:

$$\begin{aligned} E\Psi(\mathbf{R}_{1\dots N}, \mathbf{r}_{1\dots n}) = \Bigg[& -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 - \sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 + \frac{1}{2}\sum_{I \neq J} \frac{e^2}{4\pi\epsilon_0}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \\ & + \frac{1}{2}\sum_{i \neq j} \frac{e^2}{4\pi\epsilon_0}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{I,i} \frac{e^2}{4\pi\epsilon_0}\frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} \Bigg] \Psi\left(\mathbf{R}_{1\dots N}, \mathbf{r}_{1\dots n}\right) \end{aligned} \tag{2.3}$$

Here, M_I and m_i denote the masses of the Ith nucleus and the ith electron respectively. $\mathbf{r}_i$ denotes the position of the ith electron, and ∇_I and ∇_i are the gradient operators applied to the Ith nucleus and the ith electron respectively. Ψ represents the total wave function of the system, $\hbar$ is the reduced Planck's constant and ϵ_0 is the permittivity of free space.

Due to the fact that the nuclei are much heavier than the electrons, (even for the lightest of atoms, $M_H \approx 2000 m_e$), the nuclear and electronic wave functions can be decoupled, allowing the electrons to be considered independently to the nuclei. In addition, the nuclei are extremely localised compared to the electrons, and can therefore be assumed to be classical point charges. Together, this set of approximations are commonly known as the Born-Oppenheimer approximation [22, 23].

As a result of this approximation, the first and third term in equation 2.3 can be considered constant, and can therefore be ignored. This only leaves the kinetic energy of the electrons and their interaction with each other and the stationary nuclei. This leaves the electronic Schrodinger equation:

$$E\Psi(\mathbf{r}_{1\ldots n}) = \left[-\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{4\pi\epsilon_0}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{I,i}\frac{e^2}{4\pi\epsilon_0}\frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} \right] \Psi\left(\mathbf{r}_{1\ldots n}\right) \tag{2.4}$$

In principle, the energy of any configuration of atoms can be calculated. In practice however, this is not tractable. To see this, consider the wave function for a single electron, $\Psi(\mathbf{r}_1)$. Since, in general, this will not have an analytic form, it has to be represented on, for example, a grid of points. Assuming a grid of 10 points in the $\mathbf{x}$ direction, 10 points in the $\mathbf{y}$ and 10 in the $\mathbf{z}$, this results in 1000 numbers to be computed. If a second electron were to be added to the system, the resulting wave function $\Psi'(\mathbf{r}_1, \mathbf{r}_2)$ will need an additional 1000 points for *each* point in Ψ, resulting in 1,000,000 numbers to compute. Similarly, 3 electrons would need 10^9 numbers, 4 electrons 10^{12} numbers, etc. This trend will continue to scale exponentially with the number of electrons and, for the 24 electrons in a diamond unit cell, a total of 10^{72} numbers need to be computed, requiring around 6×10^{58}TB of storage. Given that, in 2007, all of humankind was only able to store 2.9×10^8TB [24], this is clearly infeasible.

In an attempt to get past this, the system can be approximated as one of n fictional non-interacting electrons. The question is then what potential do these fictional electrons experience?

2.4.1 Hartree theory

One idea, first proposed by Hartree [25], is to have each electron feel the mean electrostatic repulsion of all the electrons.:

$$\varepsilon_i \psi_i(\mathbf{r}) = \left[-\frac{\hbar^2}{2m_e}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + V_{\text{Hartree}}(\mathbf{r}) \right] \psi_i(\mathbf{r}) \tag{2.5}$$

where

$$V_{\text{ext}}(\mathbf{r}) = -\sum_I \frac{e^2}{4\pi\epsilon_0} \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}|} \tag{2.6}$$

is the electron's interaction with the external potential from the nuclei, and

$$V_{\text{Hartree}}(\mathbf{r}) = \frac{e^2}{4\pi\epsilon_0} \int_{\text{Vol}} \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \tag{2.7}$$

is the electron's interaction with the density ρ of all the electrons in the system:

$$\rho(\mathbf{r}) = \sum_i \psi_i^* \psi_i \tag{2.8}$$

.

This model has some serious drawbacks. First of all, quantum many-body effects such as electron exchange, the repulsive force between electrons caused by the requirement for Ψ to be antisymmetric for exchanging electrons, have been ignored. Secondly, since $\rho(\mathbf{r})$ includes the electron i, it will interact with itself. This is known as self-interaction.

2.4.2 Kohn-Sham DFT

To fix this, an additional potential can be added to correct for these effects. However, it not obvious that this potential does not rely on all the electronic wave functions

$\psi_1 \ldots \psi_n$.

Fortunately, Hohenberg and Kohn [26] proved that the ground-state energy of a system of electrons in an external potential is uniquely determined by the electron density:

$$E[\rho] = \int_{\mathrm{Vol}} V_{\mathrm{ext}}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + F[\rho] \tag{2.9}$$

While they gave no indication as to what the functional F would be, Kohn and Sham [21] later showed that $F[\rho]$ could be written as:

$$F[\rho] = T_{\mathrm{KS}}[\rho] + E_{\mathrm{Hartree}}[\rho] + E_{\mathrm{xc}}[\rho] \tag{2.10}$$

where

$$E_{\mathrm{Hartree}}[\rho] = \int_{\mathrm{Vol}} V_{\mathrm{Hartree}}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \tag{2.11}$$

is the Hartree energy and $T_{\mathrm{KS}}[\rho]$ is the kinetic energy of a set of fictional non-interacting electrons ψ_i who's density exactly matches the density of the fully interacting electron system.

The exchange correlation energy

$$E_{\mathrm{xc}}[\rho] = \int_{\mathrm{Vol}} V_{\mathrm{xc}}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \tag{2.12}$$

contains all the effects neglected from the other terms. These include the electron exchange effects mentioned earlier as well as effects caused by the correlation of individual electrons due to the many-body nature of electron-electron interactions. E_{xc} also has the effect of correcting $T_{\mathrm{KS}}[\rho]$, since the fictional non-interacting electrons will not have the same kinetic energy as the real interacting electrons.

The eigen-energies and wave functions of these non-interacting electrons can be calculated by solving the single-particle Schrödinger equation for each state:

$$\varepsilon_i \psi_i(\mathbf{r}) = \left[-\frac{\hbar^2}{2m_e}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + V_{\text{Hartree}}(\mathbf{r}) + V_{\text{xc}}(\mathbf{r}) \right] \psi_i(\mathbf{r}) \tag{2.13}$$

Once the ψ_i have been computed, the kinetic energy

$$T_{\text{KS}}(\psi_{1\ldots n}) = -\frac{\hbar^2}{2m_e}\sum_i \langle \psi_i^* | \nabla^2 | \psi_i \rangle \tag{2.14}$$

can be evaluated.

2.4.3 The exchange-correlation energy

All that is unknown now is the exact form of E_{xc}. To do this, Kohn and Sham suggested that, for a smooth enough density, it would be sufficient to consider the density to be approximately uniform in space and, at each point in space, use the local energy density for a homogeneous electron gas of that density. This turns out to work fairly well, and is known as the Local Density Approximation (LDA)[21].

$$E_{\text{xc}}^{\text{LDA}}[\rho] = \int_{\text{vol}} \epsilon^{\text{HEG}}(\rho(\mathbf{r}))\, \rho(\mathbf{r}) d\mathbf{r} \tag{2.15}$$

Here, $\epsilon_{\text{xc}}^{\text{HEG}}$ is the exchange-correlation energy density of the homogeneous electron gas with density $\rho(\mathbf{r})$. This can be calculated using higher order levels of theory for a range of densities and analytic forms can be parameterised to these[27, 28].

In an attempt to improve on this, it is possible to consider not only the local density at each point in space, but also its gradient. This is known as the generalised gradient approximation (GGA).

$$E_{\text{xc}}^{\text{GGA}}[\rho] = \int_{\text{vol}} \epsilon^{\text{HEG}}(\rho(\mathbf{r}), \nabla\rho(\mathbf{r}))\, \rho(\mathbf{r}) d\mathbf{r} \tag{2.16}$$

One common GGA is the PBE functional [29].

GGAs often have the effect of softening the atomic bonds, and tend to give more accurate total energies and structural energy differences than LDA. They can however over correct for this, giving them a tendency to under-bind atoms [29]. There is also small additional cost to using a GGA over an LDA. Due to the implicit locality in local and semilocal functions such as the LDA and GGA, non-local effects such as the Van der Waals interaction are not included. As a result, a number of schemes have been proposed to include such effects [30, 31].

2.4.4 Spin density functional theory

When considering magnetic systems, it is essential to calculate the effects of spin in the model. This is done by separating the electronic states of opposing spins:

$$\{\psi_i\} = \Big\{\{\psi_i^{\uparrow}\}, \{\psi_i^{\downarrow}\},\Big\} \tag{2.17}$$

and subsequently the densities:

$$\rho^{\sigma}(\mathbf{r}) = \sum_i \psi_i^{\sigma\,*}\psi_i^{\sigma};\ \sigma = \{\uparrow, \downarrow\} \tag{2.18}$$

From this, the total density ρ and spin density σ can be calculated

$$\rho(\mathbf{r}) = \rho^{\uparrow} + \rho^{\downarrow} \tag{2.19}$$

$$\sigma(\mathbf{r}) = \rho^{\uparrow} - \rho^{\downarrow} \tag{2.20}$$

As a result of this, a different E_{xc} is required for the different spins [32]. Similar to before, E_{xc} can be calculated to a higher level of theory for a homogeneous electron gas, both polarised and non-polarised, and interpolated between them based on the value of $\frac{\sigma(\mathbf{r})}{\rho(\mathbf{r})}$ [27, 28]. This results in two sets of independent fictional electrons:

$$\varepsilon_i^\sigma \psi_i^\sigma(\mathbf{r}) = \left[-\frac{\hbar^2}{2m_e}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + V_{\text{Hartree}}(\mathbf{r}) + V_{\text{xc}}^\sigma(\mathbf{r})\right]\psi_i^\sigma(\mathbf{r}) \tag{2.21}$$

2.4.5 DFT+U

While LDA and GGA functionals correctly remove self-interaction effects for the homogeneous electron gas, they can still fail significantly where the electronic states are highly localised. This occurs frequently in atomic d-like or f-like electronic states, where there is a significant contribution from the Coulomb energy due to the charge density of electrons occupying those states. DFT+U [33, 34] attempts to correct for this by adding an energy term associated with the occupation of those states:

$$E_{\text{DFT+U}} = E_{\text{DFT}} + \frac{U}{2}\sum_I \sum_{m,\sigma}\left\{n_{mm}^{I\sigma} - \sum_{m'} n_{mm'}^{I\sigma} n_{m'm}^{I\sigma}\right\} \tag{2.22}$$

where

$$n_{mm'}^{I\sigma} = \sum_i \langle \psi_i^\sigma | \, P_{mm'}^I \, | \psi_i^\sigma \rangle \tag{2.23}$$

is the overlap between the single particle wave functions and $P_{mm'}^I$, the generalised projection operators of the localised states m and m' on atom I, and U is a parameter describing the on-site Coulomb exchange interaction [35]. Using $E_{\text{DFT+U}}$ has the effect of increasing the energy of partially occupied states to counteract the underestimation of these states in E_{Hartree}.

2.5 Plane waves and periodic lattices

Now that the problem has been described mathematically, the question of how to represent this computationally arises. Since the systems are periodic, a natural description would be a Fourier basis. In this description, every function is described as a sum of plane waves:

$$f(\mathbf{r}) = \sum_i c_i e^{i\mathbf{q}_i \cdot \mathbf{r}} \tag{2.24}$$

If the function f is periodic with the unit cell, then $\{\mathbf{q}\}$ should be chosen such that:

$$\mathbf{q} \cdot \mathbf{a}_i = 2\pi n,\ \forall n \in \{0 \ldots \infty\} \tag{2.25}$$

These special $\mathbf{q}$ vectors are usually given the symbol $\mathbf{G}$.

While it is true that all observable functions behave like this by definition, this is not the case in general. In particular, Bloch's theorem states that the wave functions need only have a $|\Psi|^2$ which is periodic [22, 36]. This allows them to take the form:

$$\psi_{\mathbf{k}} = u(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}} \tag{2.26}$$

where $u(\mathbf{r})$ is periodic with the unit cell and $\mathbf{k}$ is a vector in reciprocal space.

When considering an infinite crystal, each component of $\mathbf{k}$ can be considered continuous in the range $-\pi$ to π, since $\psi_k = \psi_{k+2\pi}$. Since there is no preferred $\mathbf{k}$, observables such as the total energy E or the density ρ are calculated by integrating over all allowed $\mathbf{k}$ in the Brillouin zone.

Computationally, this can be done as a finite sum over evenly spaced points in the allowed range [37]. It is important to make sure that the Brillouin zone is adequately sampled to achieve accurate results.

2.6 Interatomic potentials

Despite the improved efficiency of DFT verses full many-body QM, typical implementations scale cubically with the number of atoms and can only scale up to a few thousand atoms on today's largest computers. Methods exist which scale linearly with the number of atoms [38] which are able to simulate tens of thousands of atoms, however to extend beyond this, more approximations need to be made. One common approximation is to ignore the electrons entirely, and instead treat the atoms as point particles, interacting in some empirically derived potential. This is often described in terms of the number of atoms needed to calculate each term:

$$V = \sum_{I} V_{I}^{1-\text{body}} + \frac{1}{2!}\sum_{I,J} V_{IJ}^{2-\text{body}} + \frac{1}{3!}\sum_{I,J,K} V_{IJK}^{3-\text{body}} + \dots \tag{2.27}$$

In the absence of an external field, the 1-body term is usually constant for each atom, and is often ignored in fixed-N calculations.

2.6.1 2-Body potentials

For many systems, most notably these that are chemically inert, a good approximation is to assume that a 2-body potential exists between atoms. This is usually constructed with a long-range interaction describing the Van der Waals attractive force between the atoms, and a long-range repulsion. The Van der Waals interaction manifests as a $1/|\mathbf{r}_{IJ}|^6$ potential. The repulsive term is intrinsically more complicated, but is often assumed to be a higher-order polynomial or exponential function. Two common examples of 2-body potentials are the Lennard-Jones potential [6]:

For many systems, most notably these that are chemically inert, a good approximation is to assume that a 2-body potential exists between atoms. This is usually constructed with a long-range attraction mimicking the dipole-dipole interaction ,and a short-range repulsion between the atoms. Physically, the dipole-dipole term manifests as a $1/|\mathbf{r}_{IJ}|^6$ interaction. The repulsive term is intrinsically more complicated, but is often assumed to be a higher-order polynomial or exponential function. Two common examples of 2-body potentials are the Lennard-Jones potential [6]:

$$V_{IJ}^{\text{LJ}} = 4\varepsilon^{\text{LJ}}\left[\left(\frac{\sigma^{\text{LJ}}}{|\mathbf{R}_{IJ}|}\right)^{12} - \left(\frac{\sigma^{\text{LJ}}}{|\mathbf{R}_{IJ}|}\right)^{6}\right] \tag{2.28}$$

and the Buckingham potential[39]:

$$V_{IJ}^{\mathrm{Buck}} = Ae^{-B|\mathbf{R}_{IJ}|} - \frac{C}{|\mathbf{R}_{IJ}|^6} \tag{2.29}$$

where σ^{LJ}, $\varepsilon^{\mathrm{LJ}}$, A, B and C can be chosen to fit to either experimental data or higher order levels of theory.

In order to compare these potentials, they can be rewritten in the form:

$$V_{IJ}^{\mathrm{LJ}} = D\left[\left(\frac{R_{\mathrm{min}}}{R_{IJ}}\right)^{-12} - 2\left(\frac{R_{\mathrm{min}}}{R_{IJ}}\right)^{-6}\right] \tag{2.30}$$

$$V_{IJ}^{\mathrm{Buck}} = D\left[\left(\frac{6}{\zeta-6}\right)\exp\left(\zeta\left(1-\frac{R_{IJ}}{R_{\mathrm{min}}}\right)\right) - \left(\frac{\zeta}{\zeta-6}\right)\left(\frac{R_{\mathrm{min}}}{R_{IJ}}\right)^{6}\right] \tag{2.31}$$

Choosing $\zeta = 13.776$ leads to the two potentials having the same description near R_{min} [40].

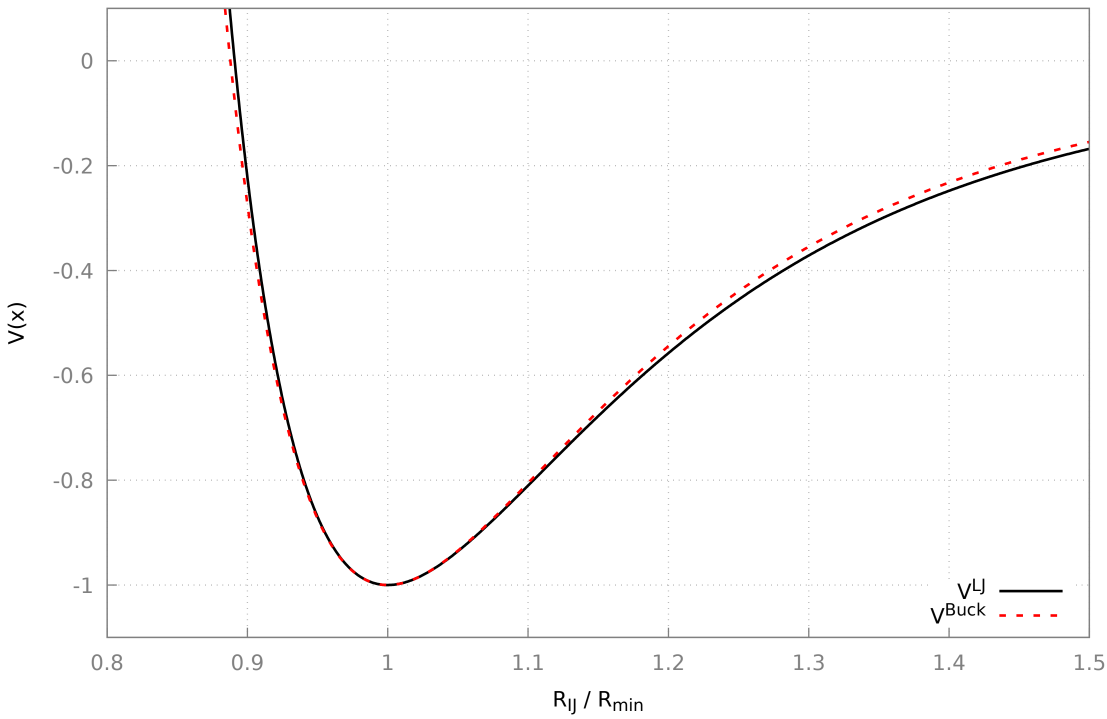

Figure 2.4: Comparison of the Lennard-Jones and Buckingham potentials in reduced units.

Figure 2.4 demonstrates that, near the minimum, both potentials have similar behaviour. However, V^{Buck} becomes unphysical near $|\mathbf{R}_{IJ}| = 0$ as the dominant $-C/|\mathbf{R}_{IJ}|^6$ term goes to $-\infty$.

2.7 Other potentials

For many materials, pair potentials do not provide a sufficiently accurate description of the system. Depending on the system in question, there are a number of ways in which to approach the problem.

For covalently bonded systems, it is often useful to include higher order terms of equation 2.27 explicitly, describing interactions between 3 or more bodies. Examples of this can be found in the BKS potential for Quartz [41] and Stillinger Weber potential for silicon [42]. These potentials are often parameterised to a particular system, and can behave unphysically beyond this. For example, Stillinger Webber is very good at describing bulk silicon but falls short when considering silicon surfaces [43]. As a result, they typically behave poorly for systems far from their parametrisations.

For metallic systems, a significant contribution to the energy comes from atomic interactions with delocalised electrons. For these systems, methods such as Finnis-Sinclair [44] and EAM [7, 8] include an interaction with some idealised electron density in addition to pair-wise interactions between the atoms.

In recent years, significant progress has been made in machine learning approaches to atomic potentials. In particular, Gaussian processes [45] and neural networks [46, 47] have both been used to describe the energetics of systems of atoms, demonstrating near-DFT accuracy, but on much larger systems of atoms, or on a much longer timescale.

These methods usually involve running many small-scale calculations with a higher accuracy method such as DFT, and fitting some generalised model system to them. A clever choice of model system allows the learnt systems to perform calculations infeasible for the original level of accuracy. For example, by approximating an atom's contribution to the energy based on its local neighbourhood, many more similar atoms could be modelled, each with energy contributions approximated from their neighbourhood, as shown in figure 2.5.

While machine learning approaches can provide large speed increases over higher

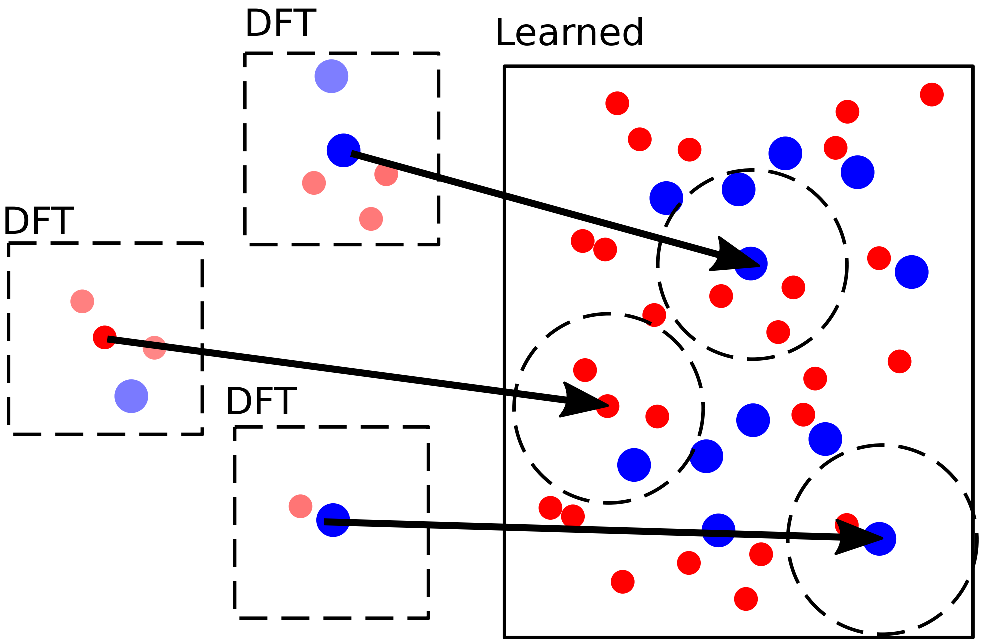

Figure 2.5: An illustration of high-accuracy training simulations of local environments (left) emerging in a larger learned system (right). The energy contribution of neighbouring atoms around a central atom is calculated for a range of possible configurations using DFT, and atomic energy contributions are approximated from these results in the larger system.

accuracy methods, they can only be relied upon simulating systems similar to those they were trained on. As a result, it can turn out to be more expensive to train the models sufficiently than it would have been to simply use a higher accuracy method to begin with.

2.8 Simulating magnetic effects

While quantum-based approaches can naturally include magnetic effects, they are often too complex to include explicitly in atomic potentials, and their effects included through the mechanical and thermal properties used to parameterise them [48]. There are some potentials however that try to explicitly include magnetic effects [48, 49, 50] and some machine learning models have been trained to deal with magnetic systems [51].

Another approach is to assume that the atoms are fixed on a lattice. In this case, the distance-dependence of the energy can be considered constant and only the spin-dependence considered. When considering the spin-dependence of a system of atoms with atomic spins $\mathbf{S}_I$, there are 4 main terms which contribute to the energetics of the system [52]:

$$V_{IJ}(\mathbf{S}_I, \mathbf{S}_J) = V^{\mathrm{ex}} + V^{\mathrm{di}} + V^{\mathrm{ani}} + V^{\mathrm{Ze}} \tag{2.32}$$

where

$$V_{IJ}^{\mathrm{ex}} = -2J_{IJ}\mathbf{S}_I \cdot \mathbf{S}_J \text{ (+ higher order terms)} \tag{2.33}$$

is the Heisenberg exchange interaction,

$$V_{IJ}^{\mathrm{di}} = \frac{1}{2}\frac{\mu_B^2\mu_0}{4\pi\mathbf{R}_{IJ}^3}\left[\mathbf{S}_I \cdot \mathbf{S}_J - \frac{3}{|\mathbf{R}_{IJ}|^2}(\mathbf{S}_I \cdot \mathbf{R}_{IJ})(\mathbf{S}_J \cdot \mathbf{R}_{IJ})\right] \tag{2.34}$$

is the magnetic dipole interaction,

$$V_I^{\mathrm{Ze}} = -\mu_B \mathbf{H}_{\mathrm{app}} \cdot \mathbf{S}_i \tag{2.35}$$

is the Zeeman energy, describing interactions with an external field $\mathbf{H}$, and V_{IJ}^{ani} is some anisotropic interaction depending on the crystal environment of the atom, usually through the effect of spin-orbit coupling.

As in atomic potentials, the parameters for magnetic simulations can be determined from a combination of *ab initio* simulation and experiments.

Chapter 3

Local & global optimisation

3.1 Introduction

Physically useful systems tend to represent minima in the energy landscape. In order to determine these structures, a range of optimisation algorithms can be used. This chapter reviews a number of optimisation algorithms used in the field of computational materials simulation, both to find locally and globally stable structures.

3.2 Minima and Maxima

Given a mathematical description of a system of atoms, the question arises of which systems are physical and useful, and which are unphysical or unstable. As was eluded to earlier, physical systems tend to lie at the minima of their energy landscapes. This can be explained simply by looking at what would happen if this were not the case, demonstrated in figure 3.1.

Knowing this then, it would be very useful to be able to calculate the location of these minima, i.e. the locally stable atomic configurations of the system. This is where the idea of function optimisation comes in. The field of function optimisation aims to find the extrema (maxima and minima) of functions. This work will focus on minimisation since interesting structures lie at energy minima. However it should

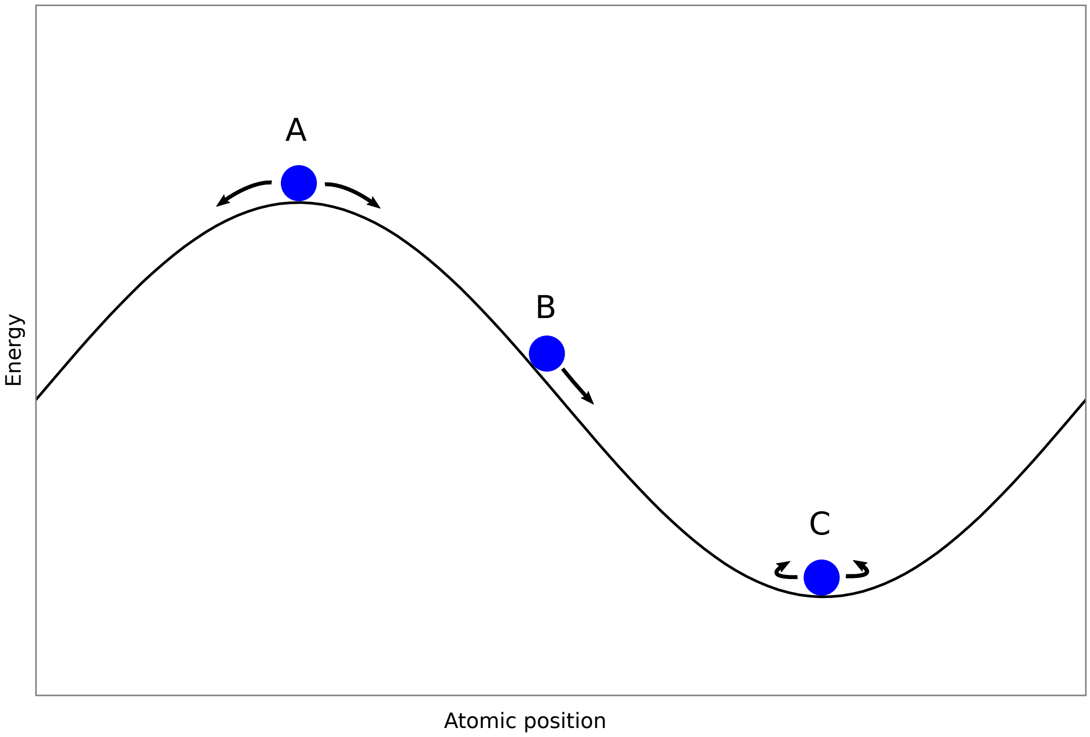

Figure 3.1: Dynamics of an atom on a PES at: A: Zero force, but will lower energy if any force applied, B: Nonzero force towards the minimum, C: Zero force, with restoring force if any external force applied.

be noted that the methods described below apply to other optimisation problems as well.

3.3 Local verses global optimisation

When considering function optimisation, there are 2 main classes of problems, local and global optimisation. Local optimisation aims to find minima nearby to some starting point, often the minimum of the well containing that point. Global optimisation on the other hand aims to find the lowest overall minimum across the entire search space. This is demonstrated in figure 3.2.

In the case of structure optimisation, local optimisation is useful when there is some idea of the structure of interest, but need to calculate positions to a higher level of precision. Global optimisation is more useful when either there is little or no information about the structure of interest interest. Global optimisation is also

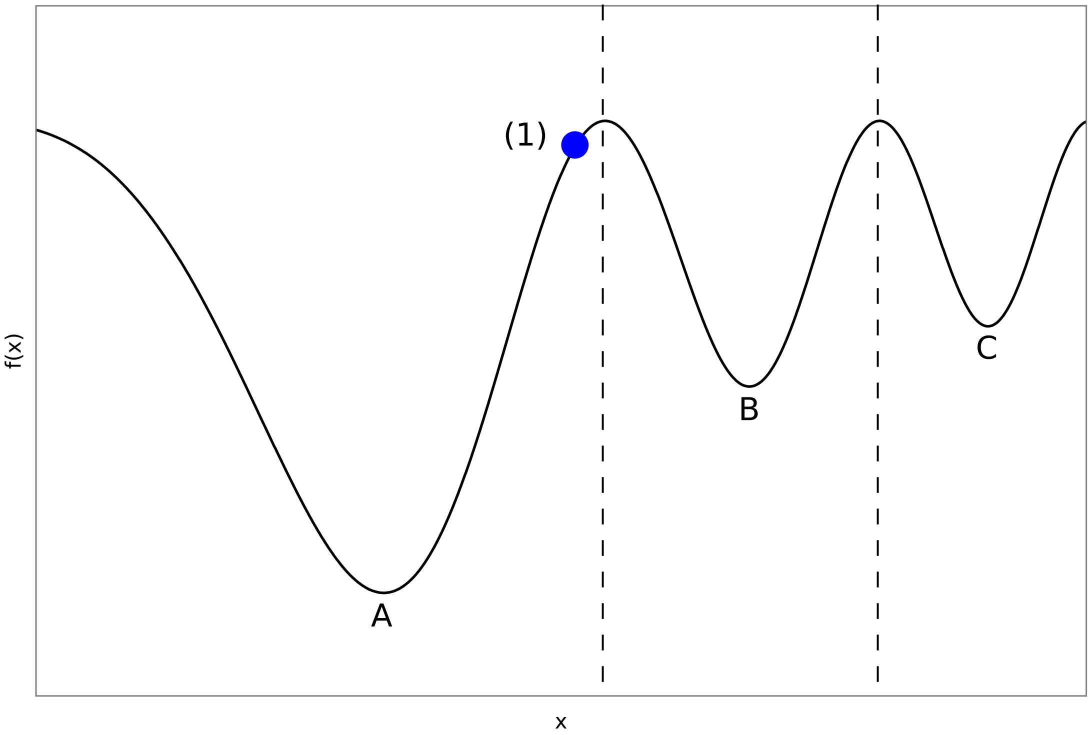

Figure 3.2: An example of global (A) and local (B, C) minima, with their basins of attractions highlighted. Note: point (1) is closer in x to B, but is in A's basin of attraction

useful when there are a range of structures of interest, since these algorithms tend to sample a large portion of the search space.

3.4 Local optimisation algorithms

3.4.1 Steepest descent methods

Local minimisation algorithms start off with some input vector $\mathbf{x}$, and aim to find some improved vector $\mathbf{x}^{\text{opt}}$ near $\mathbf{x}$ such that the objective function $f(\mathbf{x}^{\text{opt}})$ is a minimum. There are a number of approaches to this problem, but a popular class of solution involves the idea of gradient descent [53]. This involves taking the information about the gradients $\mathbf{g}$, where

$$g_i(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial x^i} \tag{3.1}$$

in order to modify $\mathbf{x}$ in such as way as to reduce the value of f. This process is repeated until the gradient is zero in all directions, at which point a minimum has been reached.

The simplest example of a gradient descent algorithm is that of steepest descent. In this method, $\mathbf{x}$ is improved by adding a small vector directly opposite the gradient each iteration since, for an infinitesimally small change, this direction will most significantly reduce $f(\mathbf{x})$. The value of $\mathbf{x}$ at iteration $i+1$ can be calculated as:

$$\mathbf{x}^{i+1} = \mathbf{x}^{i} - \lambda \mathbf{g}^{i} \tag{3.2}$$

where λ is either some small constant, or an optimisable parameter. In this case, a 1D optimisation of λ is performed, known as a line search. The value of λ used for each iteration is the value which minimises f at $\mathbf{x}^{i+1}$.

Even with a sufficiently small λ, the steepest descent algorithm is guaranteed to converge for continuous functions and can take many iterations to get near the minimum [54]. On the other hand if a line search is performed, each step is guaranteed to be orthogonal to the previous step, which can again result in many steps being needed to the minimum. This is a particular problem if $f(\mathbf{x})$ is highly anisotropic, as can be seen in figure 3.3.

3.4.2 Quasi-Newton methods

In order to improve the steepest descent algorithm, information from more higher order derivatives can be used. Newton's method calculates the updated positions $\mathbf{x}$ as:

$$\mathbf{x}^{i+1} = \mathbf{x}^{i} - \mathbf{H}^{-1} f^{i} \tag{3.3}$$

where $\mathbf{H}_{ij} = \frac{\partial^2 f}{\partial \mathbf{x}_i \partial \mathbf{x}_j}$ is the matrix of second derivatives known as the Hessian.

In order to compute the full Hessian matrix for a vector of n elements, $O(n^2)$ second

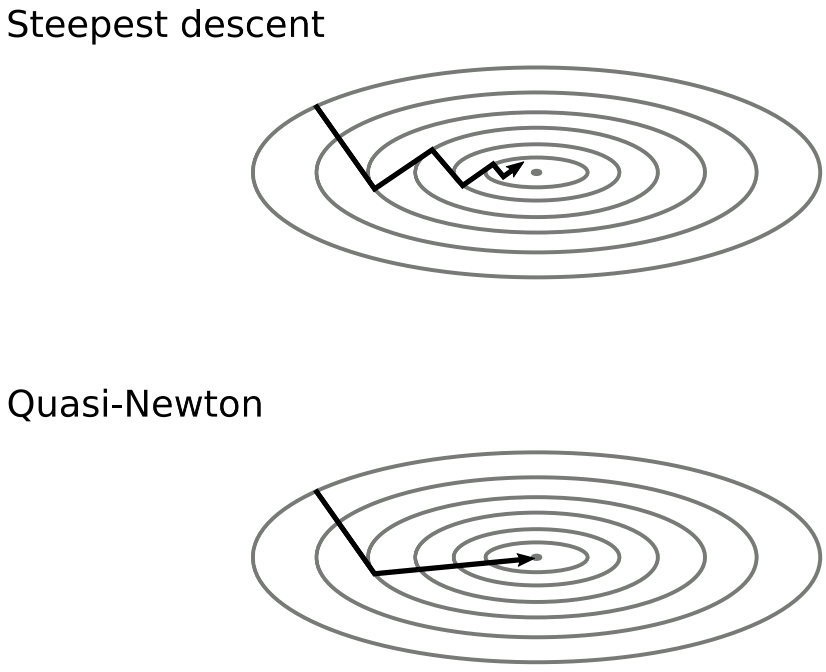

Figure 3.3: Illustration of steepest descent vs. a quasi-Newton method for the 2D anisotropic harmonic well.

order derivatives need to be computed, which can get very expensive. As a result, a number of quasi-Newton methods have been developed. These methods either use information from multiple previous iterations to implicitly include second order effects, or explicitly calculate an approximate Hessian matrix.

One algorithm which does the former is the two point steepest descent (TPSD) algorithm [55]. TPSD aims to use information of the positions and gradients from the previous iteration in order to predict an improved step size for the current iteration:

$$\mathbf{x}^{i+1} = \mathbf{x}^{i} - \alpha^{i}\mathbf{g}^{i} \tag{3.4}$$

where

$$\alpha^{i} = \frac{\Delta\mathbf{x} \cdot \Delta\mathbf{g}}{\Delta\mathbf{g} \cdot \Delta\mathbf{g}}, \tag{3.5}$$

and

$$\Delta \mathbf{x} = \mathbf{x}^i - \mathbf{x}^{i-1}, \tag{3.6}$$

$$\Delta \mathbf{g} = \mathbf{g}^i - \mathbf{g}^{i-1} \tag{3.7}$$

Another quasi-Newton method is the BFGS method [53], named after Broyden [56], Fletcher [57], Goldfarb [58] and Shanno [59], who independently discovered it. This method explicitly stores and updates an approximate Hessian $\mathbf{B}$ each iteration. BFGS is updated to $\mathbf{x}$ is:

$$\mathbf{x}^{i+1} = \mathbf{x}^i - \lambda \left(\mathbf{B}^{-1}\right)^i f^i \tag{3.8}$$

and $\mathbf{B}$ is updated as:

$$\mathbf{B}^{i+1} = \mathbf{B}^i + \frac{(\Delta f^i)(\Delta f^i)^\top}{(\Delta f^i)^\top (\Delta \mathbf{x})^i} - \frac{\mathbf{B}^i (\Delta \mathbf{x}^i)(\Delta \mathbf{x}^i)^\top (\mathbf{B}^i)^\top}{(\Delta \mathbf{x}^i)^\top \mathbf{B}^i (\Delta \mathbf{x}^i)} \tag{3.9}$$

To improve computational efficiency, $\mathbf{B}^{-1}$ is often stored and updated rather than $\mathbf{B}$. This way, the matrix does not need to be inverted every iteration.

For the first iteration, $\mathbf{B}$ is usually chosen to be the identity $\mathbf{I}$, or some initial guess for the Hessian, if one exists.

While BFGS performs very well in the quadratic region (i.e. where the function is approximately parabolic), its explicit reliance on the approximate Hessian mean that, if the curvature at the starting point and near the minimum are significantly different, the algorithm can struggle to converge. This is particularly noticeable in crystal structure optimisation where, near convergence, the energy landscape is approximately quadratic. However, far from any minima this is not the case, and methods such as TPSD perform better.

3.4.3 Preconditioning

The rate of convergence of both steepest descent and quasi-Newton methods are affected by how isotropic the search space is. This is because all dimensions are considered equivalently when calculating step sizes. This is not always the case, for example a variable cell geometry optimisation may be be simultaneously optimising the lattice vectors and the atomic positions within the lattice, which will behave differently.

If some knowledge of the system is known in advance, each dimension in the search space can be scaled such that they are all approximately equal. For example, methods exist for geometry optimisation [60], as well as wave function minimisation [61] and density mixing [62] in density functional theory calculations.

3.4.4 Other local minimisers

In the field of crystal structure optimisation, another well known local energy minimisation algorithm is damped molecular dynamics (DMD) [63, 64]. This involves evolving the dynamics of a system of atoms with some additional force included to damp the dynamics close to the minima. This can be viewed as a quasi-Newton algorithm using the first derivatives of the positions, rather than the second [63].

$$\mathbf{x}^{i+1} = \mathbf{x}^{i} + \mathbf{v}^{i}\Delta t\,, \tag{3.10}$$

$$\mathbf{v}^{i+1} = \mathbf{v}^{i} + \left(\frac{1}{1+\gamma}\right)\left(\frac{\mathbf{a}^{i}\Delta t}{2m} - \gamma\mathbf{v}^{i}\right) \tag{3.11}$$

where γ is some damping parameter. Careful choice of γ will result in critical damping of the system [63], whereas choosing $\gamma = 0$ results in standard NVE MD using Euler's integration scheme. Damped MD can similarly be applied to higher order MD integration schemes, such as Verlet integration.

While these techniques are generally less efficient than second order quasi-Newton methods, they can be useful if, for example, computing the second derivatives is too costly.

3.5 Global optimisation algorithms

Unlike local minimisation algorithms which only search for a minimum in the region of their initial starting state, global minimisation algorithms search for the vector $\mathbf{x}^{\text{opt}}$, or the set of vectors $\{\mathbf{x}\}^{\text{opt}}$ where the function has its lowest value. In other words,

$$\mathbf{x}_i \in \{\mathbf{x}\}^{\text{opt}} \text{ iff } f(\mathbf{x}_i) \leq f(\mathbf{x}') \,\forall\, \mathbf{x}' \tag{3.12}$$

Often, by symmetry or coincidence, there are multiple global minima for a given function. Depending on the use-case, a global minimiser's task can be to find a single global minimum, the set of global minima, or the set of all minima, including the global minima.

It can be shown that, for the general global optimisation problem, no single global optimisation algorithm is superior. This is known as the no free lunch (NFL) theorem [65]. However, with some domain-specific knowledge of the optimisation problem, different optimisation algorithms can be shown to have advantages [66].

3.6 Metaheuristic algorithms

Due to the nature of global optimisation, metaheuristic algorithms are often used. These algorithms use more general algorithmic frameworks than exact mathematical methods, and their designs are often inspired by natural processes.

3.6.1 Simulated annealing

A good starting point for finding the crystal structures of materials is to look at how it is done experimentally. The idea of annealing has been around for centuries as a mechanism for reducing stresses, and removing dislocations and defects from materials. The process involves heating the material to around the temperature at which recrystallisation can occur and letting it slowly cool, allowing the constituent atoms to fall into more energetically favourable configurations [67].

Simulated annealing involves a similar principle [68], by which the dynamics of a system are propagated through time. While doing this, the simulation temperature is controlled in order to allow the system to explore a wide search space (high T), while exploring the minima of potential wells (low T). When run for a sufficiently long time, and with appropriate control over the temperature, many approximate global minima can be found.

While simulated annealing has been used for structure optimisation [69], it has the disadvantage that only local steps can be taken in the configuration space, so if the system gets stuck in a deep local minimum, it can become difficult to escape.

3.6.2 Basin hopping

Basin hopping resolves this pitfall by performing larger random jumps in the space of possible configurations using a Monte Carlo algorithm, then using locally optimising the resultant structure to find the minimum of that particular basin [70].

Minima hopping [71, 72, 73] improves upon this idea by accepting or rejecting each hop based upon the energy difference it creates. These jumps can either be performed in regular Cartesian coordinates, or in a more chemically motivated coordinate system, as done by Panosetti *et al.* [73].

3.6.3 Metadynamics

Metadynamics provides an alternative solution to the drawbacks of simulated annealing [74]. As the simulation evolves, repulsive Gaussians are added to the function at points in the configuration space as they are being explored. This results in basins of attraction that have been searched slowly filling, directing the search into other minima. This allows the system to locate new minima as is done in simulated annealing, without getting stuck in regions of configuration space which have already been explored.

This technique has the advantage that the free energy surface can be recovered, and the stability of finite temperature structures can be investigated [75]. However like simulated annealing, movement through the configuration space can only be done locally, having to explore each minimum it finds. This means that if the search does not start near the global minimum, it can take a long time to reach it. In addition, the performance of metadynamics is highly dependent on the shape of the Gaussians used to fill the wells, with performance slowed significantly if the applied Gaussians are too narrow or too wide.

3.6.4 Population based methods

While methods like SA and metadynamics will explore all of configuration space if given enough time, they are limited early on to regions near their starting position. This is because, at each iteration, they can only make local movements. One popular solution to this is to use a number of multiple starting positions. Each can then independently be explored in an attempt to find local minima. Members of this population can then be updated in order to explore other regions of the configuration space. Due to the independent nature of these members, population based methods often have more scope for parallelism than other methods.

3.6.4.1 Random search

The simplest population based method is the random search, where a large population of independent members is used to sample the configuration space. While this will eventually find the global solution, there is no information built up about the space, therefore configurations of interest are as likely to be found by the first member as the last. A variation on this technique has been to great effect by Pickard and Needs [76], who choose populations of ’random sensible structures’. This choice can be based on physical augments such as ensuring bond lengths are in physically sensible ranges or imposing symmetry on the system, or on chemical intuition such as building structures out of molecular base units.

3.6.4.2 Evolutionary algorithms

Another popular class of population based methods for structure prediction is that of evolutionary algorithms [77, 78, 79, 80]. In these algorithms, the ‘fitness’ of members of a population is evaluated in some way, in an attempt to decide which of them are most likely to find solutions of interest. The population is then updated in some way, based on information learned from these members, and the process is repeated.

Most common amongst evolutionary algorithms are the genetic algorithms (GA). GAs perform this population update by selecting favourable members of the population, and generating new population members from combinations of these members. Depending on the problem at hand, a variety of methods exist for the fitness evaluation, member recombination and selection stages. This work focusses on a GA designed for crystal structure prediction, which will be detailed in the next chapter.

When considering how to represent an optimisation problem for a genetic algorithm, a choice needs to be made about how the search vector is to be represented for these operations. A simple GA may treat the input vector simply as a 1-D array of real numbers or as a bit string representing that vector as illustrated in figure 3.4. However, it is often the case that a more physically meaningful representation of the search vector can improve the convergence characteristics of the algorithm.

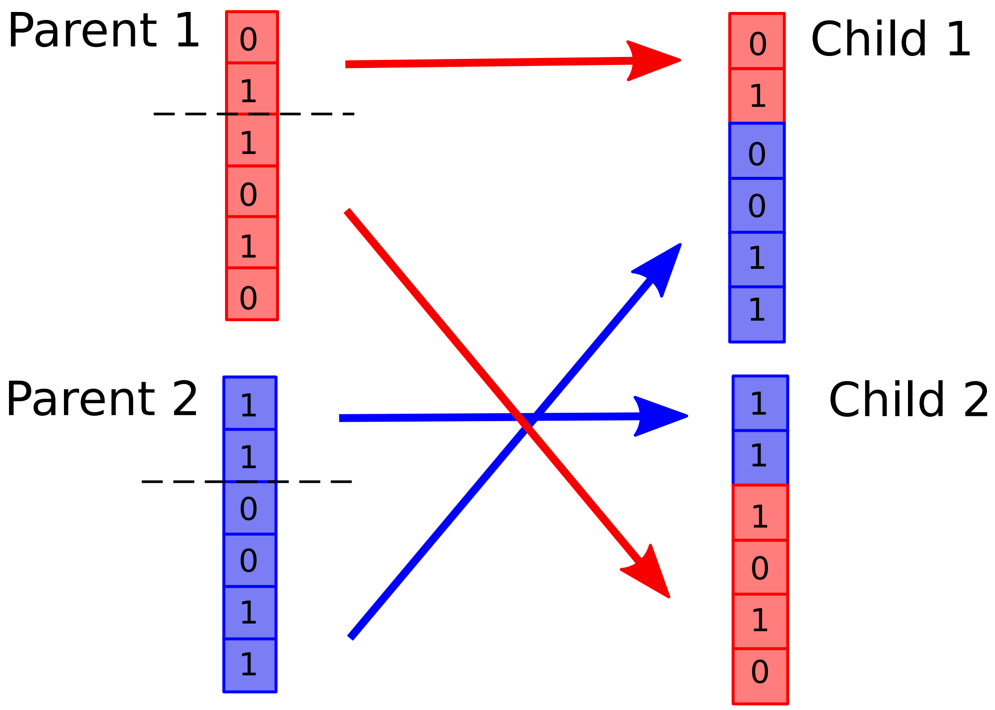

Figure 3.4: Demonstration of binary encodings being shared during crossover of two members represented as bit-strings.

In addition to pure GAs, a number of hybrid GAs exist, incorporating ideas from other algorithms. For example, memetic algorithms allow each member of the population to perform local searches [81, 82]. Additionally, ideas such as metadynamics or tabu searches can be applied to these local searches.

3.6.5 Tabu search

In addition to the optimisation algorithms above, a number of algorithms exist to dissuade a calculation from exploring regions which have already been explored. These methods are known as tabu search methods [83]. The main idea of these methods is to somehow keep a memory of which regions of configuration space have already been explored, and guide the search away from these.

There are three kinds of memory which can be utilised in a tabu search:

- Short-term: A list of states which have been recently visited
- Intermediate-term: Intensification rules to focus the search on some region of

interest, and

- Long-term: Diversification rules which drive the search into new regions.

Any or all of the above memories can be used to improve the performance of other optimisation algorithms.

Chapter 4

A Realspace genetic algorithm for structure prediction

4.1 Introduction

This chapter reviews the GA for crystal structure prediction first developed by Abraham and Probert [78, 84]. Ideas specific to the domain of structure prediction are discussed, such as how the crystal is represented in the GA and how the various GA operations are interpreted in this context. In addition updates to the GA which have been developed over the course of this work are discussed. These include the use of niching for structural diversity and a corrected structure factor fingerprint.

4.2 Overview of the GA

The work presented in this thesis is based upon a GA for crystal structure prediction written by Abraham and Probert [78, 84] as part of the CASTEP code [85], known from hereon in as `castep_GA`. CASTEP is a materials modelling program written in Fortran, primarily for DFT simulations of periodic systems. The `castep_GA` program provides a GA for structure prediction based on the energies and forces calculated by CASTEP, and utilises a number of CASTEP features, such as energy evaluation and local geometry optimisation.

In `castep_GA`, each member is represented by a unit cell described by its cell vectors, and a list of atoms described by their positions relative to the unit cell. It takes as an input an example unit cell, along with a list of calculation parameters, both for the GA and for the underlying energy calculations. The process for optimising the crystal structure of the material is performed as follows:

1) First, `castep_GA` reads in a CASTEP-style cell and parameter file, and creating a population of initially *mutated* structures based on this cell.
2) The geometry of these members of the population are optimised locally using CASTEP's internal methods.
3) The *fitness* of the converged members is evaluated.
4) Members are *selected* for *crossover* based on their *fitness*, and new members are generated.
5) Members of the population are *mutated.*
6) Any members in the population who are new or who have been modified are then geometry optimised, as performed in step 2).
7) The *fitness* of these members is re-evaluated.
8) Members are *selected* for the next generation based on their *fitness.*
9) Steps 4-8 are then repeated until the calculation has *converged.*

In order to explain how `castep_GA` works, it is important to look at the GA operations in the context of atomic structure prediction.

4.3 Crossover

In the crossover step, two members, commonly known as the parents, are combined in some way to produce some number of new members, known as children. Since the members are represented by atomic positions in some unit cell, it makes sense to choose some atoms from one parent and some from another to produce a child.

One simple way of doing this is to take analogous planar cuts through the unit cells

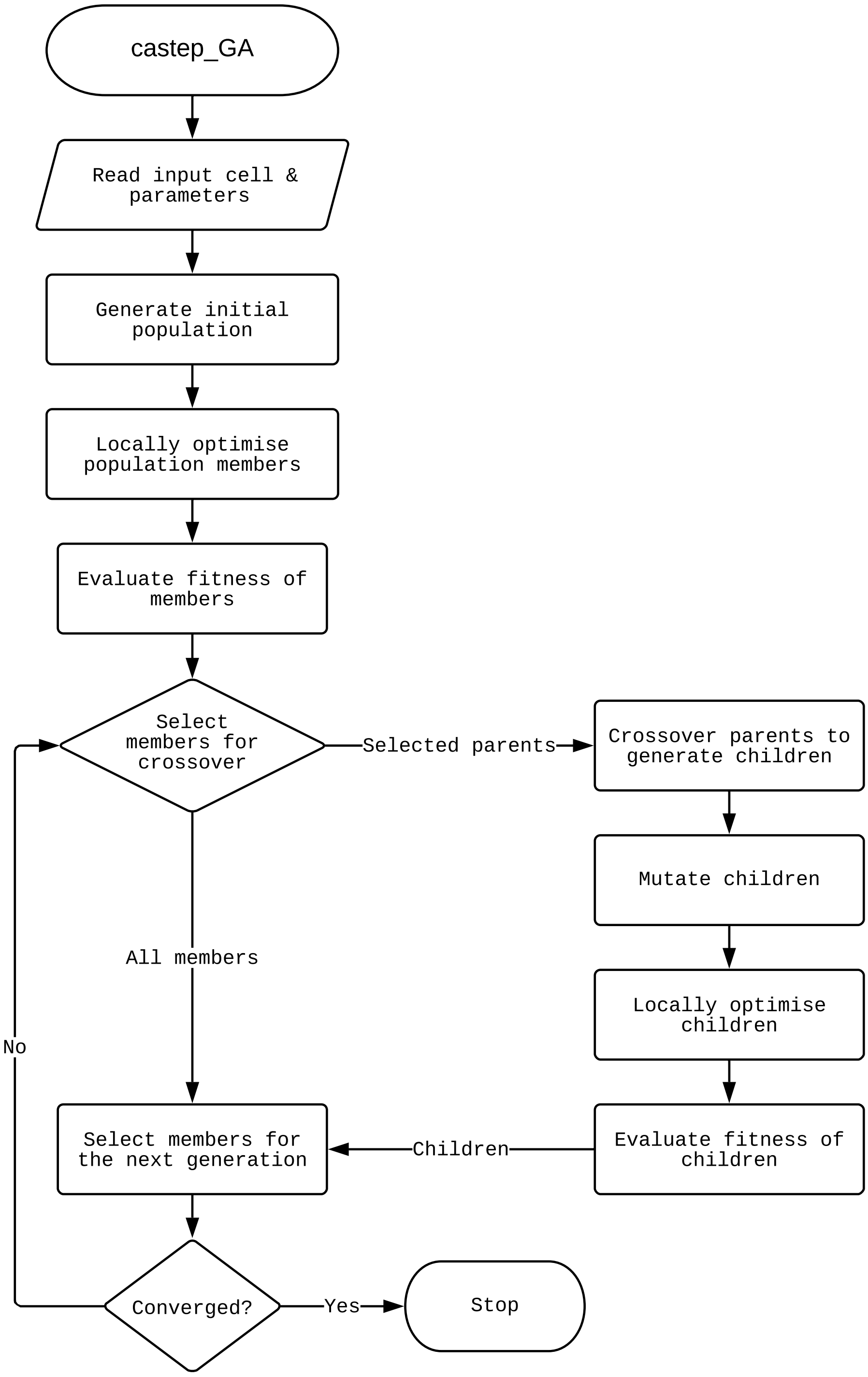

Figure 4.1: Program flow of `castep_GA`

of each parent. One child can then be formed by taking all the atoms from one side of the cut from one parent and all the atoms from the other side from the other. A second child can then be formed from the remaining atoms. In fact, has been shown [78] that instead of using planar cuts, using a plane-wave shaped cut performs significantly better. This is because discontinuities caused by planar cuts produce extra work for the local optimiser.

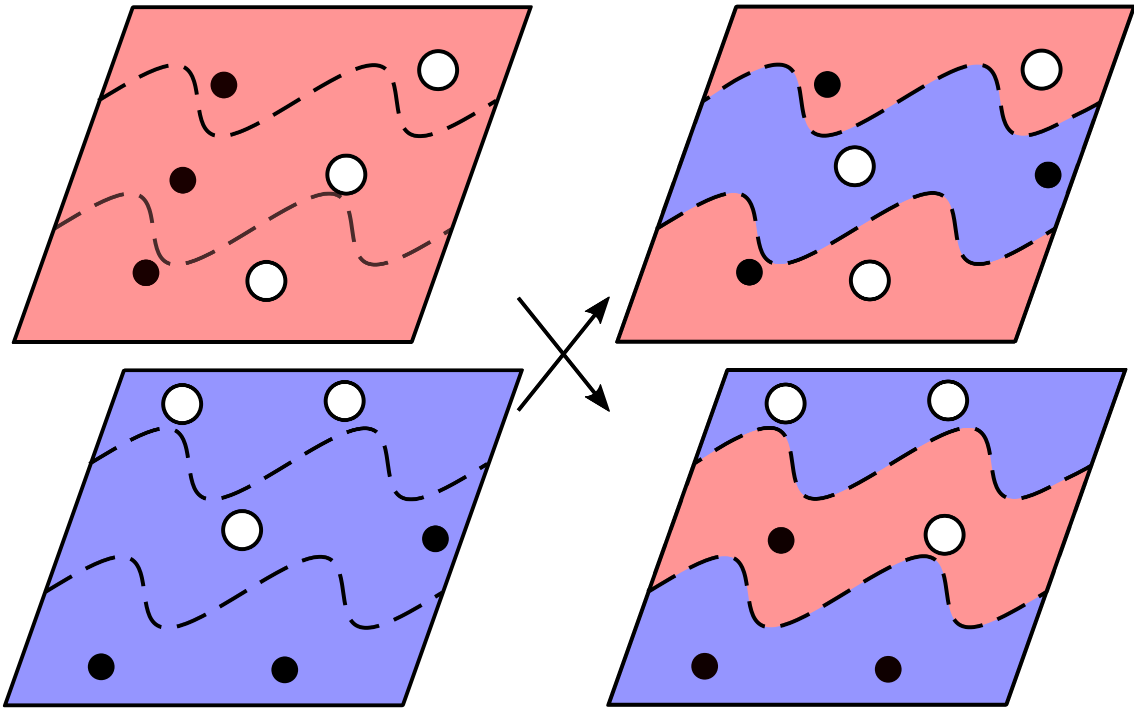

Figure 4.2: Real-space crossover of 2 6-atom cells in 2 dimensions using the periodic cuts. Atoms from the two parents (left) are chosen based on whether they are within the cut or not, and copied into the appropriate children (right).

When performing crossover, it is possible that the two parents will have different lattice vectors. In this case, a choice needs to be made between performing the crossover in absolute or fractional coordinates. Performing the crossover in fractional coordinates will preserve the periodic nature of structures, but absolute distances and angles between atoms is not guaranteed to be conserved. Performing crossover in absolute coordinates will preserve local distances between atoms, but may not be commensurate with the resultant unit cell, losing the periodic nature of the structure. For these reasons, the crossover in `castep_GA` is performed in fractional coordinates.

4.4 Mutation

As mentioned in the previous chapter, it is useful to mutate the members in order to more rapidly cover the search space. In `castep_GA`, there are three mutation operations that can be performed:

- Cell vector perturbation, where the unit cell vectors are modified from their current values,
- Atomic perturbation, where a chosen atom is moved from its current position, and
- Atomic permutation, where two atoms of different species are swapped with each other.

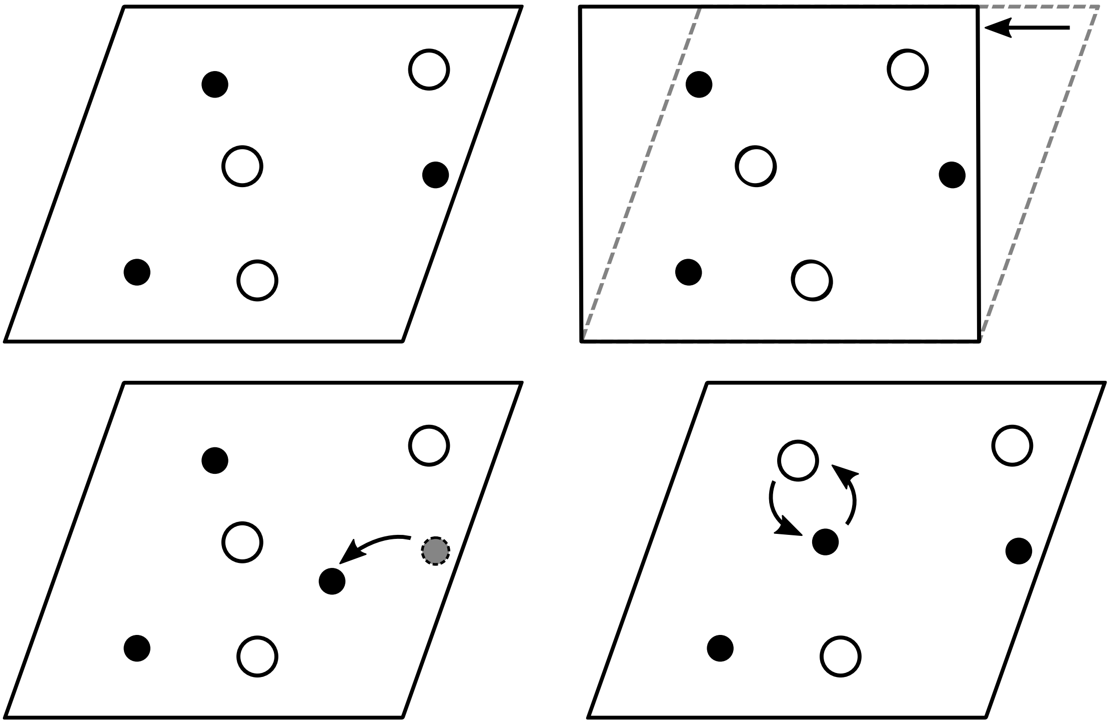

Figure 4.3: Example of mutation operations on a unit cell in 2 dimensions. The original cell (top left) can be mutated by cell vector permutation (top right), atomic perturbation (bottom left) or atomic permutation (bottom right).

For a pure GA, mutations are essential to cover the entire search space, since there is no other mechanism for component values to change from their original value from the first generation. However since `castep_GA` uses a local optimiser, the role of the mutation step is different. Mutations allow a member to move into different basins of attraction, allowing the local optimiser to explore them. As a result, cell and atomic perturbation mutations need to be sufficiently large such that the local optimiser will fall into a different minimum.

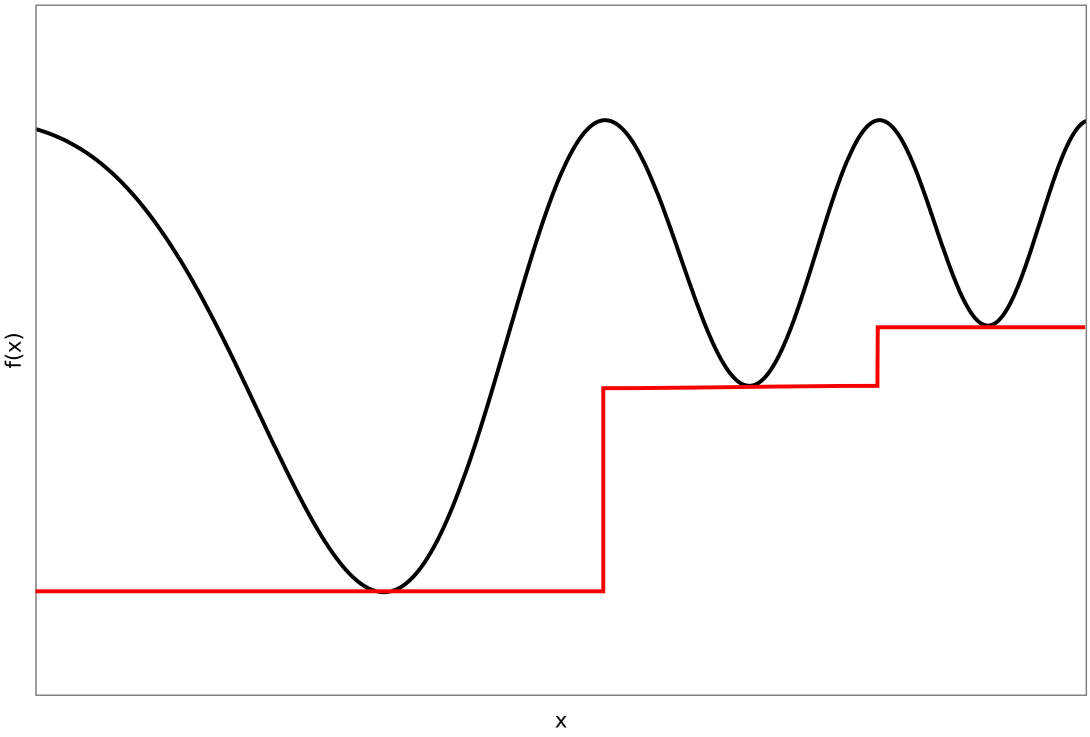

Figure 4.4: The Effect of local optimisation on the apparent landscape of the GA. Since the local optimiser returns the local minimum if (black line) for all points in that minimum's basin of attraction, the GA is effectively optimising x with respect to the red line. This transformation has also been discussed by Wales[1].

4.5 Selection

At a number of points in the GA, a subset of members need to be selected from the population. In particular, this is done when selecting parents for crossover and selecting members to advance to the next generation. Naively, one might assume that simply selecting the fittest members will allow the algorithm to converge most rapidly. However this can lead to a phenomenon known as stagnation, where all members of a population represent the same solution. A better way of selecting members is by taking a weighted random sample of the population, where the weight of each member being selected is related to its fitness. This is commonly known as roulette selection.

4.6 Fitness evaluation

In order to compare members of the population, each member is assigned with a value, between 0 and 1, saying how likely that member is to improve the search chances in future generations. The most important factor in this is the energy associated with that structure. However it is also useful to include a contribution relating to how diverse that member is with respect to the rest of the population. In order to do this, pairs of members need to be compared in in order to find out how similar they are. This is where the idea of structure fingerprinting comes in useful.

4.6.1 Structural fingerprinting

Defining a comparison between two physical structures is not a simple problem. Naively, one might be tempted to take the atoms from one structure and the atoms from another and simply sum the displacements made by each individual atom. However, there are many situations where this method will give nonzero values for physically identical structures. One obvious example of this would be if all the atoms were rigidly shifted by 1 Å. The structures would be calculated to be N Å for an N atom system, but the physics of that system would be the same. Another issue is that of permutations. For a system where any number of atoms can be considered indistinguishable from each other, for example two carbon atoms in a diamond structure, any permutation of these atoms would result in a physically identical structure, but a nonzero comparison.

To solve this, the idea of structural fingerprinting is used. A fingerprint is some mathematical description of the atomic structure, which is unchanged by such translations or permutations of the individual atoms. There are a number of options which fit these parameters, such as the radial distribution function or the crystallographic structure factor [84], as well as other heuristics based on bond orderings or more abstract interpretations [86, 87].

In the case of `castep_GA`, the spherically averaged structure factor $\Lambda(k)$ is used as a structural fingerprint. Physically, this describes the amplitude and phase of a wave with wave number k diffracted from the crystal lattice, averaged over all incident directions. This is calculated for a range of values of k for each structure.

In the original version of `castep_GA` had an incorrectly derived form of this[84, 88]. Originally the form of $\Lambda(k)$ used was:

$$\Lambda(k) = N\sum_{I=1}^{N}\rho_I^2 + 2\sum_{I=1}^{N}\sum_{J>I}^{N}\rho_I^2\rho_J^2 \times J_0\left(\sqrt{3}\pi k|\mathbf{R}_{IJ}|\right) \tag{4.1}$$

Where ρ_I is the charge density of the nucleus of atom I, defined to be the charge of the nucleus Z_I at the position of the nucleus $\mathbf{R}_I$, and zero everywhere else. J_0 is a Bessel function.

This work re-derived and implemented the correct form:

$$\Lambda(k) = \sum_{I=1}^{N}\rho_I^2 + 2\sum_{I=1}^{N}\sum_{J>I}^{N}\rho_I\rho_J \times j_0\left(k|\mathbf{R}_{IJ}|\right) \tag{4.2}$$

where j_0 is the Bessel function of the first kind. The derivation of this is presented in appendix A.

To compare two structures, denoted $\mathbf{x}$ and $\mathbf{x}'$, the Pendry R-factor (reliability factor) [89] can be used. The R-factor can be calculated as:

$$R_{\mathbf{x}\mathbf{x}'} = \frac{\sum_{k_r}|\Lambda_{\mathbf{x}}(k_r) - \Lambda_{\mathbf{x}'}(k_r)|}{\sum_{k_r}\Lambda_{\mathbf{x}}(k_r)} \tag{4.3}$$

Once the value of $R_{\mathbf{x}\mathbf{x}'}$ between all pairs of structures in a generation are known, structures which are too similar can have their fitness adjusted to dissuade all of them from being selected. There are a number of ways in which this can be done. Originally, `castep_GA` used a hybrid fitness, where each structure was compared against the best structure in the generation. The modified fitness was then calculated as a linear combination of the energetic fitness and the fingerprint fitness:

$$f'(\mathbf{x}) = \alpha f(\mathbf{x}) + (1-\alpha)R_{\mathbf{x}\mathbf{x}^{\text{best}}} \tag{4.4}$$

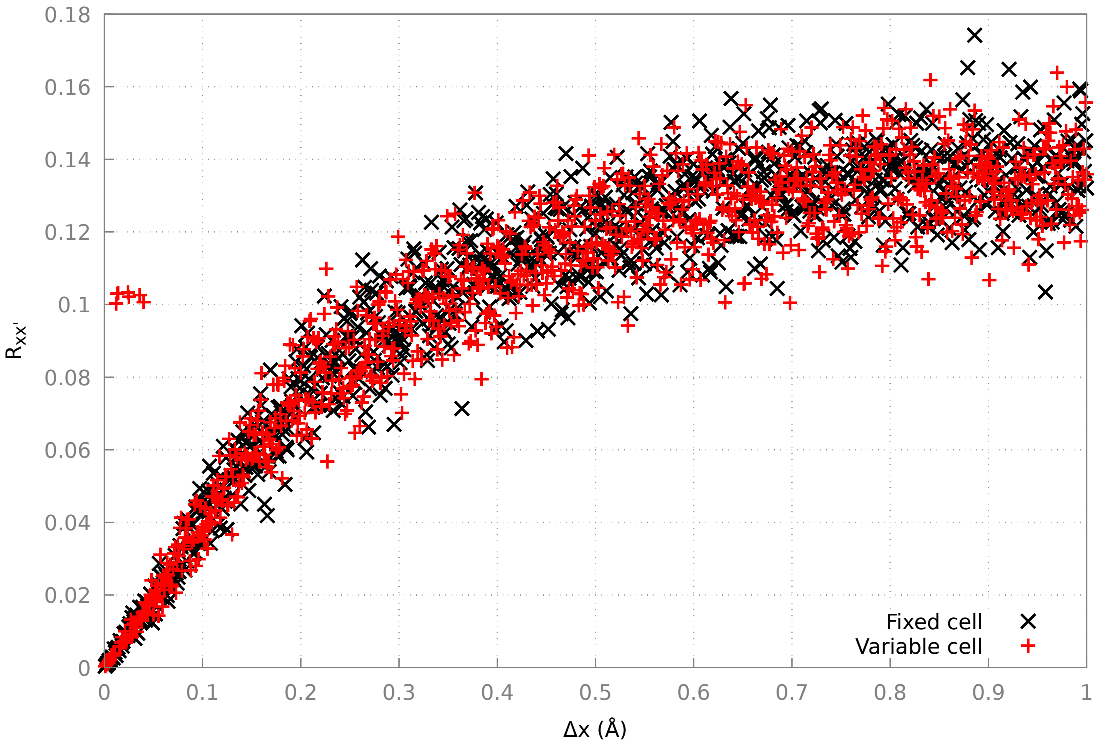

Figure 4.5: R-factor between perturbed structures in the Si_8 unit cell. For a given random perturbation on the atomic positions (black and red) and optionally the cell vectors (red), this graph shows the R-factor between that structure and the original structure.

While this improved the algorithm, there were some flaws with it. Firstly, the technique would only penalise structures similar to the best member. It could be that a number of other members all represent the same structure, but none of them are the lowest energy. This would allow all of them to be selected as if they were unique. Secondly, the value α has no obvious physical analogue, so the value of this can only be calculated from experimentation with the algorithm and may not be transferable.

An alternative method, implemented as part of this work, is to use the concept of niches [90, 91]. This idea is inspired by ecological niches, where members of a population adapt based on the environment they are in. In the GA, each member's fitness is adjusted based on the similarity of all other members in the population, to mimic the concept of competition for resources. This is done by adjusting the fitness function to:

$$f'(\mathbf{x}) = \frac{f(\mathbf{x})}{\sum_{\mathbf{x}'} s(R_{\mathbf{x}\mathbf{x}'})} \tag{4.5}$$

where $s(R)$ is known as the sharing function, and is defined as

$$s(R) = \left(1 - \frac{R}{R_{\text{scale}}}\right),\ R < R_{\text{scale}};\, 0\,\text{otherwise.} \tag{4.6}$$

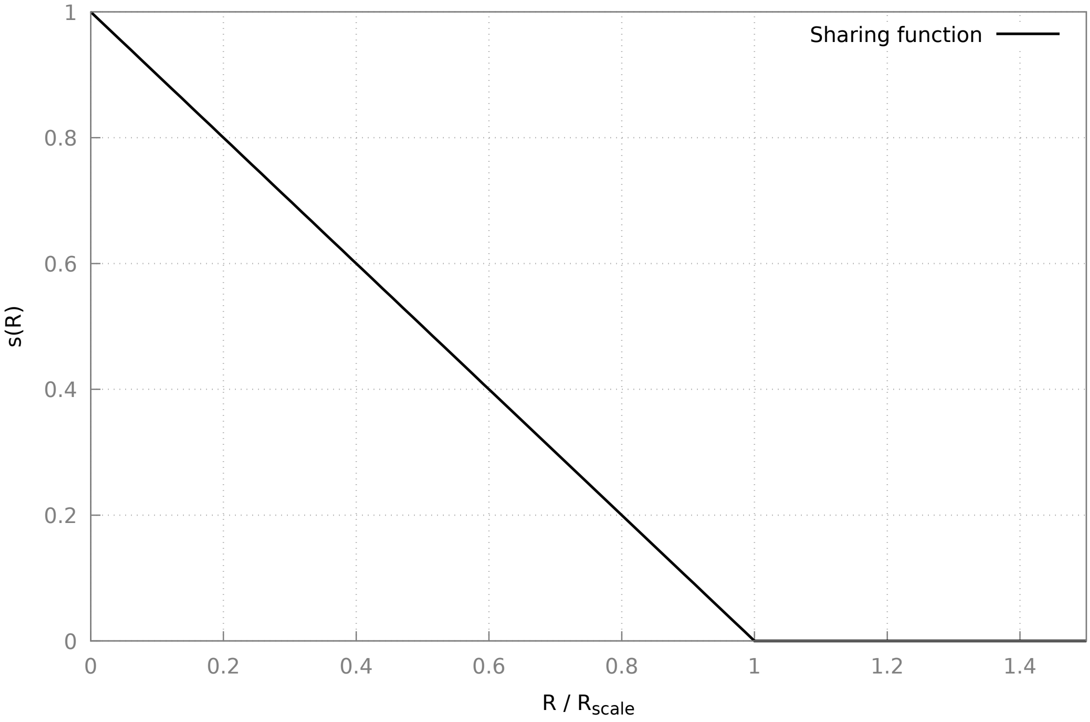

Figure 4.6: Example of the sharing function

The effect of this is that if two structures are identical, then $s(R) = 1$, and the fitness of each structure is maximally penalised. However if the two structures have a high $s(R)$, then neither will be penalised. Two similar structures, i.e. two structures with $0 < R < R_{\text{scale}}$ will be somewhat penalised, though not as much as if they were identical.

The question naturally arises of what R_{scale} should be. Since there is an approximate mapping between the size of R and the amount a structure has been perturbed from its original position, physical arguments can be made to choose an R_{scale} which corresponds to perturbations of a structure significant enough as to force it out of its basin of attraction. For this work, the value of R_{scale} was chosen to be 0.03, as this corresponds to around 0.25 Å, as is shown in figure 4.5.

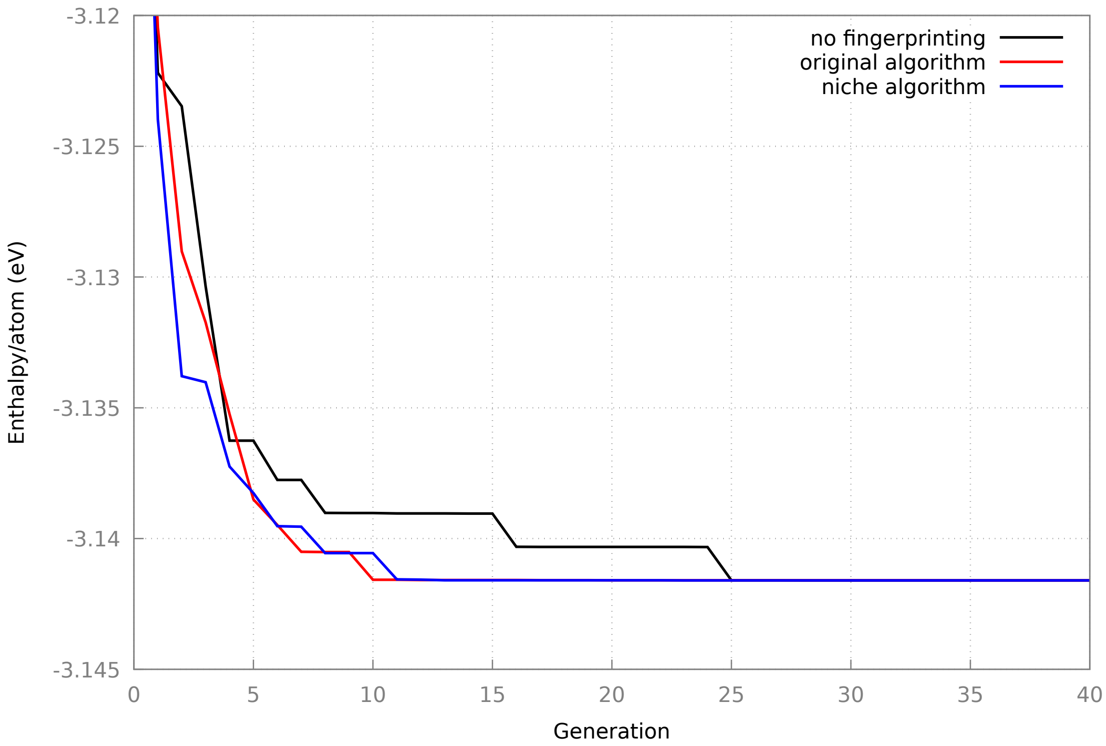

Figure 4.7: Lowest energy structures found at each generation of GA runs on a high-pressure Si_8 system, averaged over 10 runs.

The effects of niching algorithm can be seen in figures 4.7 and 4.8 for GA calculations of a high-pressure Si_8 cell. Each algorithm was run 10 times and the results averaged to remove errors. Figure 4.7 shows the accelleration that fingerprinting provides using both the original hybrid fitness algorithm and the new niching algorihtm. Both of these algorithms find the ground state structure in around 10 generations, compared to the 25 generations taken by the run without fingerprinting.

Figure 4.8 shows the diversity of structures found in the final generation of each GA. It can be seen that the runs without fingerprinting and, to a lesser extent, the runs with the original hybrid fitness algorithm both find the ground state structure multiple times. The new niching algorthm however is able to explore a wider range of higher energy minima due to the increased diversity.

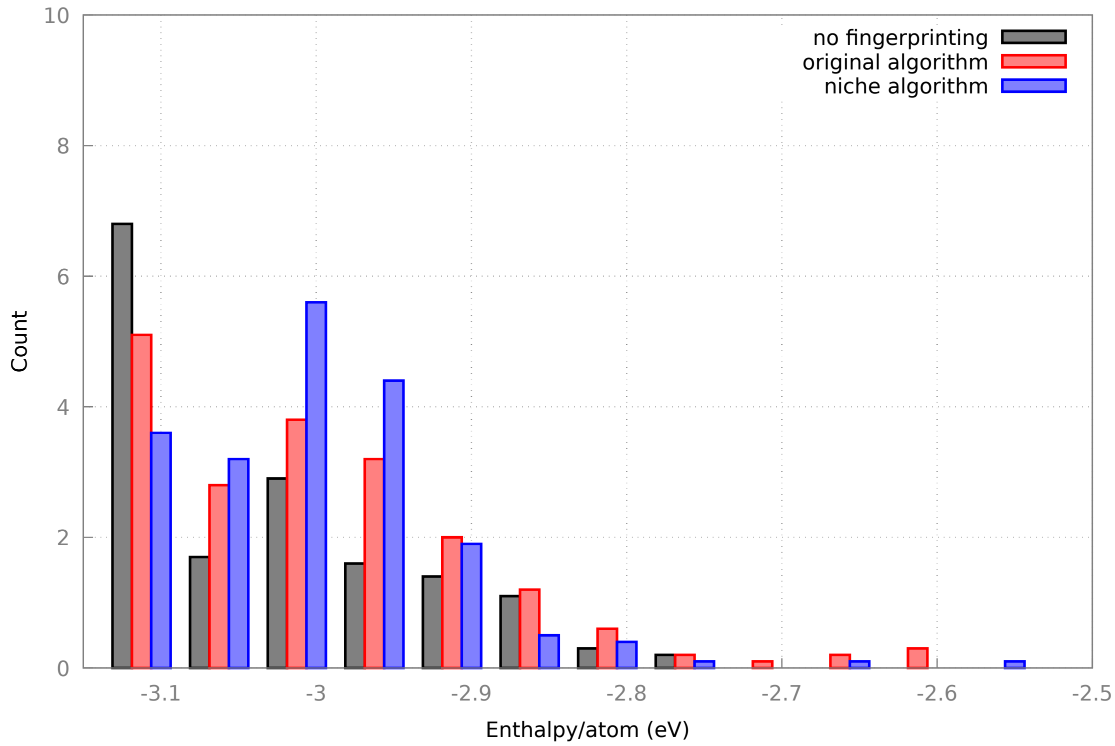

Figure 4.8: The effect of hybrid fitness vs. niching on the diversity of structures in the final generation of GA calculations high-pressure Si_8 system, averaged over 10 runs.

Chapter 5

Extending the GA to magnetic structures

5.1 Introduction

This chapter discusses the development of the GA to handle magnetic degrees of freedom as well as simply structural. Concepts introduced in the previous chapter are extended where necessary to handle magnetic state optimisation. Finally, a new magnetic structure fingerprint is introduced, inspired by spin polarised neutron scattering theory, as a natural extension to the structure factor fingerprint discussed before.

5.2 Problems with a traditional GA for structure prediction

When considering magnetic systems, the energy depends not only on the atmoic positions, but also the spin of the system. In the atomic basis, this term can be approximated as the local spin density around each atom, i.e. an atomic spin. This can be described as a scalar value aligned along some axis (collinear spins) or as a 3D vector (non-collinear spins). This raises the complexity of the system from

3^N degrees of freedom for non-magnetic systems to 4^N for collinear systems, or 6^N degrees of freedom for non-collinear systems.

Since the original `castep_GA` has no knowledge of the atomic spins and is only aware of the atomic positions, the energy landscape becomes multivalued. In other words, a single crystal structure can produce two or more physical systems with different energies. In the case of DFT calculations, the system which is actually found by `castep_GA` is left entirely up to `CASTEP`'s internal electronic minimiser. Since this only locally minimises the electronic wave functions from random values, there is no guarantee that the energy calculated for that structure is the physically correct one. This will hinder the efficiency of `castep_GA` since information about the spin configurations of previous members is lost.

One example of a system which can suffer from this is bulk iron. As can be seen in figure 5-1, iron's crystal structure, BCC, is only the lowest energy structure when the magnetic structure is ferromagnetic. If, over the course of a GA calculation, `castep_GA` finds antiferromagnetic iron or nonmagnetic iron, both of which are unphysical systems, it will falsely decide that BCC iron is an unfavourable structure.

In addition to this, efforts to dissuade similar crystal structures will mean that structures with similar crystal structures but significantly different magnetic structures will be dissuaded from being selected for future generations. This has the effect of hiding the true solutions from the GA.

5.3 Addition of spin to the GA

In order to negate these effects, spin degrees of freedom were included in `castep_GA`. In order to do this, a number of operations had to be extended to be spin-aware.

5.3.0.1 Spin propagation

Since `castep_GA` already represents the system as a set of atoms along with a unit cell, extending this to include spins is relatively simple. One need simply extend the

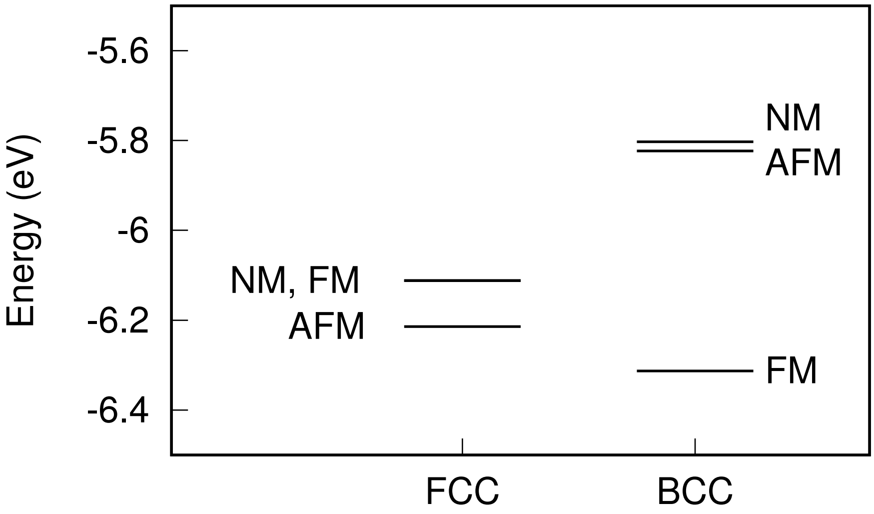

Figure 5.1: Comparison of FCC and BCC energies for different magnetic configurations of Fe unit cells [2].

description of an atom to include its spin along with its species and position.

In order for `castep_GA` to use and manipulate these atomic spins, there needs to be some mechanism for passing these to the individual CASTEP calculations. Passing an initial atomic spin configuration into CASTEP can be specified by an initial magnetic moment as either a scalar or a vector in the `.cell` file. This allows CASTEP to initialise the electronic wave function in such a way as to localise electronic spins around atoms. This encourages local wave function minimisation to start near a converged state with the magnetic properties which are desired.

At the end of the calculation, CASTEP has information about the spin of the system in the electron density. In order to project this onto an atomic basis, a number of options are available, such as a Wannier basis, or using Mulliken or Hirshfeld analysis to assign regions of the election density to specific atoms. For this work, Hirshfeld analysis [92] was chosen since this associates the entire charge and spin density to the atoms, rather than just that around the atoms as is the case in Mulliken analysis. This process involves allocating a fraction of the densities at each point to nearby atoms, based upon the ratio of the density of equivalent free-atom densities for those atoms. Information about the atomic spins on the atoms can then be passed back to `castep_GA`, allowing spins to be locally minimised by CASTEP in a similar manner to the positions.

5.3.0.2 Spin mutations

In order for `castep_GA` to efficiently optimise the spins, mutation operations need to be defined. This work defines two such mutations: perturbation and permutation, similar to the atomic positions.

In the case of perturbations, a distinction needs to be made between collinear and non-collinear spin systems. For collinear spins systems, there is only one scalar value per atom to optimise. Since CASTEP only takes in an atomic spin to initialise the electron density, it effectively acts as a local optimiser for the atomic spin. The perturbation therefore takes the same role as the atomic position perturbations, where the role of the perturbation is to move the system from one basin of attraction to another. Since there is a maximum value these spins can take, i.e. the total number of electrons associated with the atom, issues arise trying to add an additional perturbation to an existing spin. For example, an atom with saturated spin given a positive perturbation will become unphysical. As a result, the spin perturbation is evaluated as a uniform random number between `-spin_max` and `spin_max`, where `spin_max` is a user parameter, set to the maximum spin value expected on any atom. For example, for Fe_2O_3, `spin_max` should be set to around $5\hbar/2$, since an ionised iron atom is capable of having a spin of $5\hbar/2$. This perturbation allows atomic spins to spontaneously magnetise, demagnetise or flip, irrespective of the initial value.

In the case of non-collinear spins, a similar procedure is performed. In this case, the atomic spin is set to a random vector within the sphere of radius `spin_max`. Again, this allows any spin state to be found by the perturbation within the range specified, without the risk of being stuck in a state based on the structure's history. As with the collinear spin situation, local optimisation of the spins is performed within CASTEP in order to find the spin minima.

For permutations, the situation is analogous to swapping the atomic positions of different species. However, since atomic spin is closely related to the atomic structure, it only makes sense to swap atomic spins of atoms of the same species. For example, swapping the spin of a nickel and oxygen atom in NiO makes no physical sense, since all the spin exists on the Ni atoms, and the O atoms generally have zero spin.

5.3.0.3 Magnetic fingerprinting

As mentioned before, the original fingerprint will be unable to distinguish different magnetic structures for materials with similar atomic structures. As a result, the niching algorithm described in the previous chapter will incorrectly penalise distinct magnetic structures.

To overcome this, the original fingerprint has been modified to include a magnetic term. Since the original fingerprint is based on scattering amplitudes, it makes sense to look to spin-sensitive scattering experiments. For magnetic neutron scattering [93], the scattering intensity can be calculated as

$$F_{\text{neutron}}^2 = F_{\text{nucl}}^2 + q^2 F_{\text{magn}}^2 \tag{5.1}$$

Here,

$$F_{\text{nucl}}^2 = \left| \sum_{N \text{ atoms}} b_N \exp(2\pi i \mathbf{k} \cdot \mathbf{r}) \right|^2 e^{-2W} \tag{5.2}$$

is the non-magnetic scattering intensity, where b_N is the nuclear scattering intensity for atom N;

$$F_{\text{magn}}^2 = \left| \sum_{N \text{ atoms}} p_N \exp(2\pi i \mathbf{k} \cdot \mathbf{r}) \right|^2 e^{-2W} \tag{5.3}$$

is the magnetic contribution to the scattering, where $p \propto S_N$; and

$$q^2 = \sin^2 \alpha \tag{5.4}$$

where α is the angle between the scattering and magnetisation vectors.

By analogy, the structure factor fingerprint can be extended to differentiate magnetic structures

$$\Lambda = \Lambda_{\text{orig}} + q^2 \Lambda_{\text{magn}} \tag{5.5}$$

where

$$\Lambda_{\text{magn}}(k_r) = \sum_{n=1}^{N} S_n^2 + 2 \sum_{n=1}^{N} \sum_{m>n}^{N} S_n S_m j_0 \left(k_r |\mathbf{r}_n - \mathbf{r}_m| \right) \tag{5.6}$$

and q is some scaling parameter, allowing differences in magnetic structure to be more or less sensitive with respect to crystal structure.

Figure 5.2 demonstrates how the modified fingerprinting function behaves for structures as the magnetic structure of a crystal changes.

For this work, a value of $q^2 = 5.0$ was chosen in equation 5.5. This was done so that structures which vary by 1 $\hbar/2$ per atom or more appear distinct to the GA, as can be seen in figure 5.2.

5.4 Summary

Using a traditional GA on magnetic structures can lead to poor performance, and can be steered away from finding the global spin state by higher energy spin structures.. Redefining crossover, mutation and permutation operations to make them spin-aware enables both the magnetic and crystal structure of materials to be searched concurrently. In addition, the extension of the structure factor fingerprint allows the GA to distinguish between identical crystal structures with distinct magnetic structures.

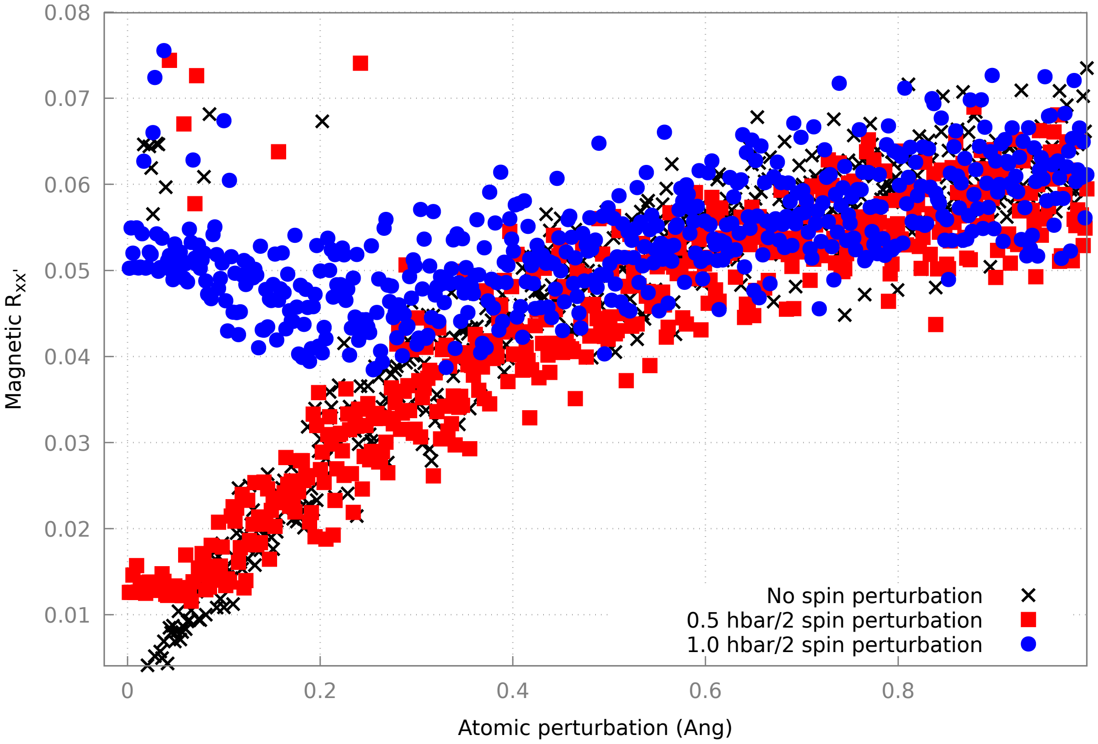

Figure 5.2: R-factor calculated for a Fe_6 unit cell. As well as a perturbation being applied to the atomic positions, each atom is given up to 0.0 (black), 0.5 $\hbar/2$ (red) or 1.0 $\hbar/2$ perturbation to its atomic spin. For small atomic perturbations, it can be seen that systems with a small spin perturbation still appear similar, whereas systems with a larger spin perturbation can now be distinguished.

Chapter 6

LJ+S: A magnetic pair-potential

6.1 Introduction

In this chapter, a new magnetic pair potential is proposed for the use of testing the GA's performance on magnetic structures. The potential is first parameterised to the FM iron dimer, and then to an AFM analogue. The GA is used to search for the lowest energy configurations of these two new potentials.

6.2 The magnetic Lennard-Jones model

As discussed in chapter 2, magnetic effects can be calculated from first principles methods such as DFT. However, since a significant portion of these effects are intrinsically quantum mechanical or many-body in nature, it is difficult to capture all of these effects in empirical potentials without limiting them to some specific regime. One example of this is the Heisenberg model, where atoms are placed on a lattice, and *only* the magnetic effects are considered. In addition to this, magnetic moments on isolated atoms tend to arise from partially filled electronic states, as dictated by Hund's rules. As a result, atoms with nonzero magnetic moments tend to be either metallic or chemically active, causing an additional challenge to empirical pair potentials.

Given all this, attempting to describe real materials using a pair-potential is beyond the scope of this work. However, it is possible to combine the spatial dependence on the energy from non-magnetic empirical potentials with the magnetic dependence on the energy from lattice-based methods like the Heisenberg model to create a potential which behaves somewhat like a magnetic material. This work defines a magnetic Lennard-Jones potential:

$$V_{ij}^{LJ+S} = V_{ij}^{LJ} - A\mathbf{S}_i \cdot \mathbf{S}_j f(\mathbf{R}_{ij}) + \frac{B}{|\mathbf{R}_{ij}|^3}\left[\mathbf{S}_i \cdot \mathbf{S}_j - \frac{3}{|\mathbf{R}_{ij}^2|}(\mathbf{S}_i \cdot \mathbf{R}_{ij})(\mathbf{S}_j \cdot \mathbf{R}_{ij})\right] \tag{6.1}$$

This potential is made up of three terms:

- The unmodified Lennard-Jones potential,
- An exchange-like term, and
- A magnetic dipole interaction term

The Lennard-Jones, dipole and anisotropy term are included as stated in chapter 2. However the Heisenberg exchange term is usually defined on a lattice, only including pairs of atoms with discrete relative lattice positions. As a result, no spatial dependence for this term exists. For this work, an exponential decay term was used to mimic the short term nature of Pauli repulsion:

$$f(\mathbf{R}_{ij}) = e^{-\alpha|\mathbf{R}_{ij}|} \tag{6.2}$$

However, other effects such as the RKKY interaction could be included here.

By tuning the values of A, B and α, a range of magnetic effects can be described.

6.3 Parameterisation of the magnetic Lennard-Jones potential

For real-world magnetic materials, it is often the exchange term which dominates the magnetic interaction [94]. Because of this, the work presented in this chapter will use a simplified version of the LJ+S potential, ignoring the dipole and anisotropy terms (i.e. choosing $B = 0$).

In this case, only 2 parameters need to be chosen, A and α. As a result of this, the effect of the magnetic modifications to the Lennard-Jones potential are to raise and/or lower the height of aligned and anti-aligned spins, depending on the sign of A. If A is positive, pairs of aligned spins will act to lower the energy and anti-aligned spins will raise it. In this case, ferromagnetic magnetic structures are expected to have the lowest energy, since aligned spins would act to decrease the total energy of the system. If A is negative, pairs of anti-aligned spins will lower the energy and pairs of aligned spins will raise it. In this case, antiferromagnetic structures are expected to have the lowest energy, since aligned spins would act to increase the total energy of the system.

In order to get physically sensible values for these, they can be parameterised in a number of ways. The way which has been chosen here is to parameterise the values to the equilibrium distance R_{min} of an atomic dimer, along with the difference in energy at R_{min} and $2R_{\text{min}}$ between aligned and anti-aligned dimers, as demonstrated in figure 6.1.

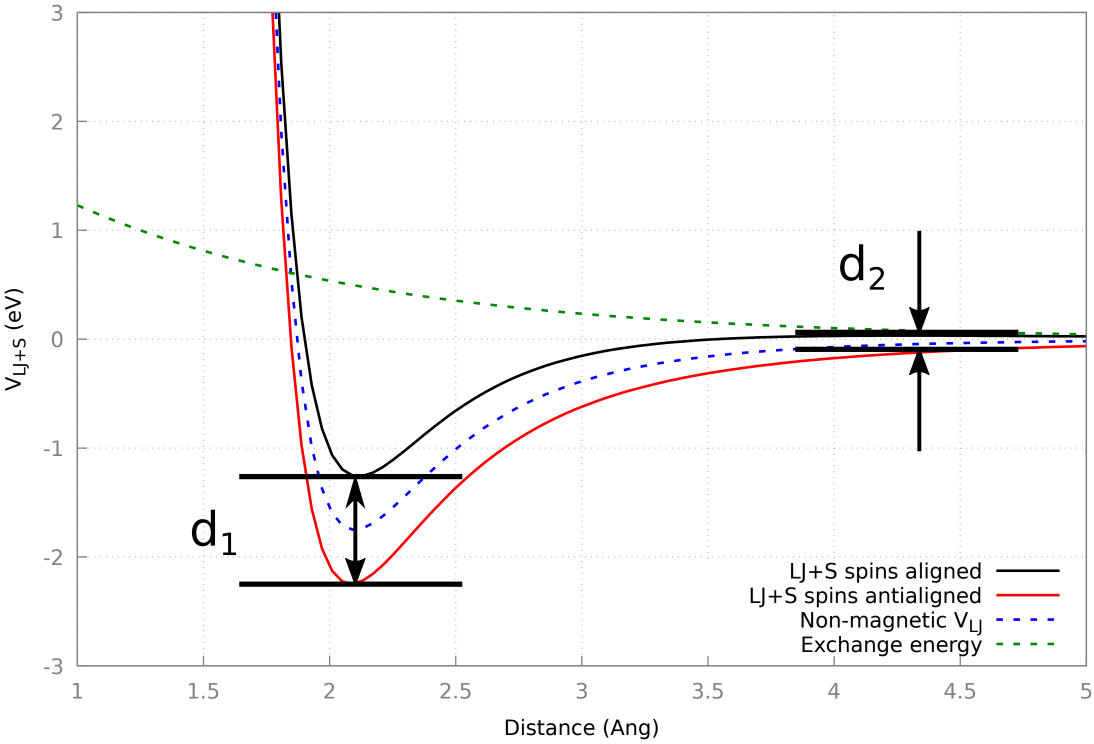

Figure 6.1: The form of the Lennard-Jones + Spin potential, showing V_{IJ}^{LJ+S} for pairs of atoms with both aligned and anti-aligned spins (solid lines), as well as the forms of the constituent terms (dashed lines). The energy differences d_1 and d_2 at R_{min} and $2R_{\mathrm{min}}$ are used to parameterise the potential.

.

6.4 A ferromagnetic parameterisation

Parameter	**Value**
$\varepsilon^{\mathrm{LJ+S}}$	1.75 eV
$\sigma^{\mathrm{LJ+S}}$	1.87 Å
A	-2.82 eV
α	0.83 Å^{-1}

Table 6.1: Table listing the parameters for the FM parameterisation of the LJ+S potential.

For a ferromagnetic parameterisation of the LJ+S potential, DFT calculations of Fe dimers were performed. Figure 6.2 shows the DFT energies, along with the parameterised LJ+S for a range of spin orientations. It can be seen that, although the width of the well has not been captured accurately, the equilibrium distance and bond energy are fairly close to the values for the iron dimer. In addition, the difference in energy between aligned and anti-aligned spins is similar, and that difference decays at a similar rate. The LJ+S function was fitted to the DFT data in the range 1.8Å to 4.5Å. This is because below 1.8Å, the Fe pseudopotential cores overlap to produce unphysical results. This can be seen by the discontinuities in the gradients at this point.

6.4.1 Performance of the GA

The GA was run on a 6-atom system of the LJ+S atoms with the FM parameterisation, optimising the atomic positions, spins and lattice parameters. The GA was run with a population size of 24, and ran for a total of 50 generations.

Using the local geometry optimisation algorithms built into CASTEP, each member was locally optimised using the TPSD algorithm for 50 generations to get into the quadratic region, then finished off with the BFGS algorithm. The structures were converged to a tolerance of 1meV. Since the atomic spins were specified explicitly

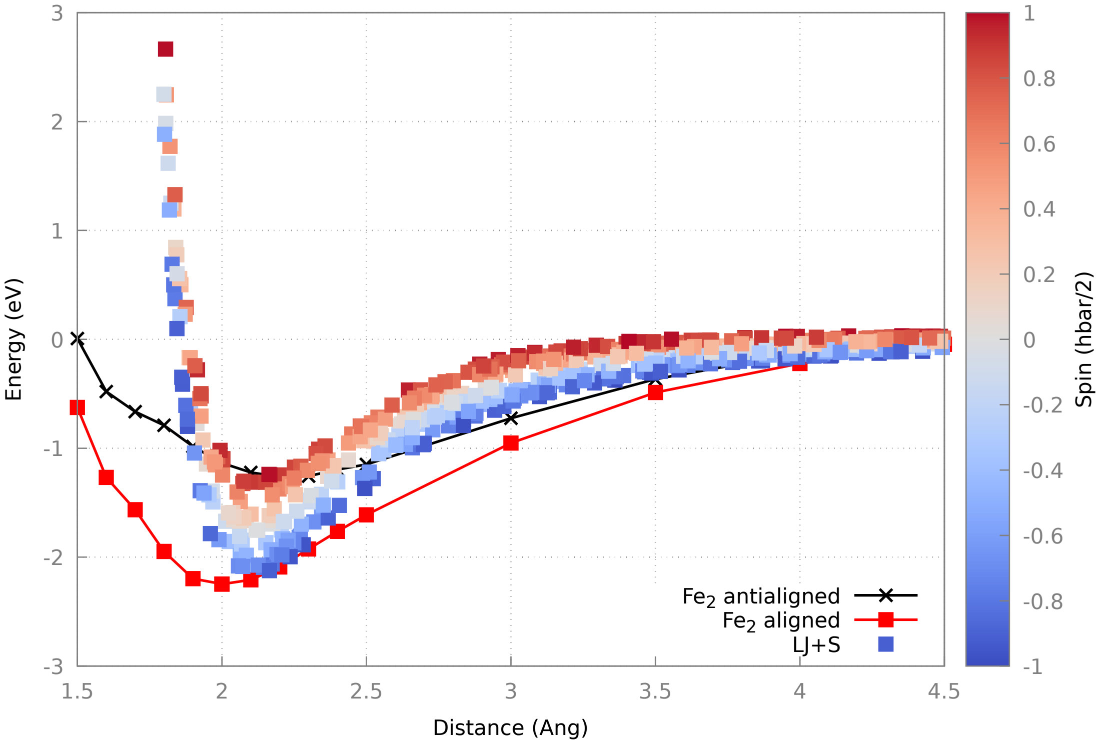

Figure 6.2: The FM parameterised LJ+S energies for atoms with randomly oriented spins of unit magnitude (points), along with the DFT data used to parameterise them (lines). The color of the points indicates the value of $\mathbf{S}_I \cdot \mathbf{S}_J$ for the atomic pair.

and were not a result of DFT densities, there was no mechanism to locally optimise the spins within CASTEP. In order to account for this, the spin mutation operation was modified to be a normally distributed rotation around the surface of the unit sphere, centring on the spin's previous position.

A mutation amplitude of 2 Å was used for the ions, at a rate of 0.03 mutations per atom. This was chosen to allow on average one atom every member to mutate, but for each jump to be significant enough to hop lattice sites. The spin mutation rate was 0.06 per atom, since the mutations in the spin would be less dramatic.

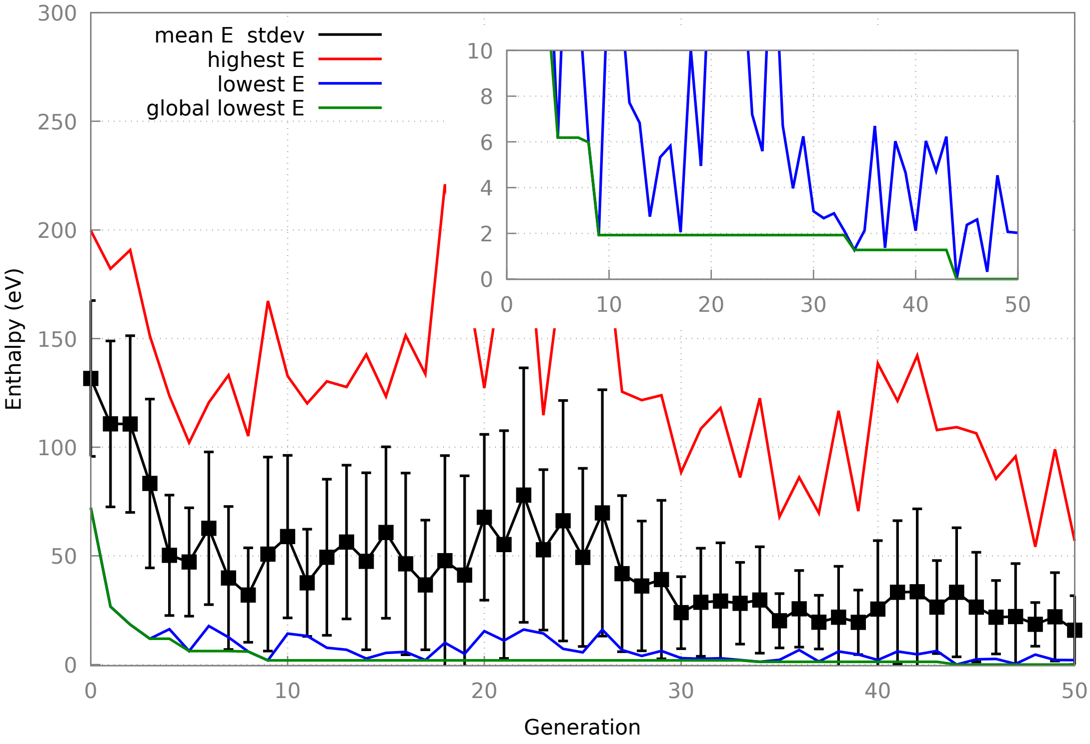

Figure 6.3: The range of enthalpies found per generation for the FM-parameterised LJ+S model.

Figure 6.3 shows the range of enthalpies found per generation. The GA significantly reduces the average enthalpy of the generations in the first 10 generations, before settling off. Figure 6.4 suggests that this is because the difference in energy between different spin structures is more significant than that of different crystal structures. As a result, the GA is able to rapidly identify the magnetic structure as FM early on in the calculation, while the atomic positions continue to be optimised further into the calculation.

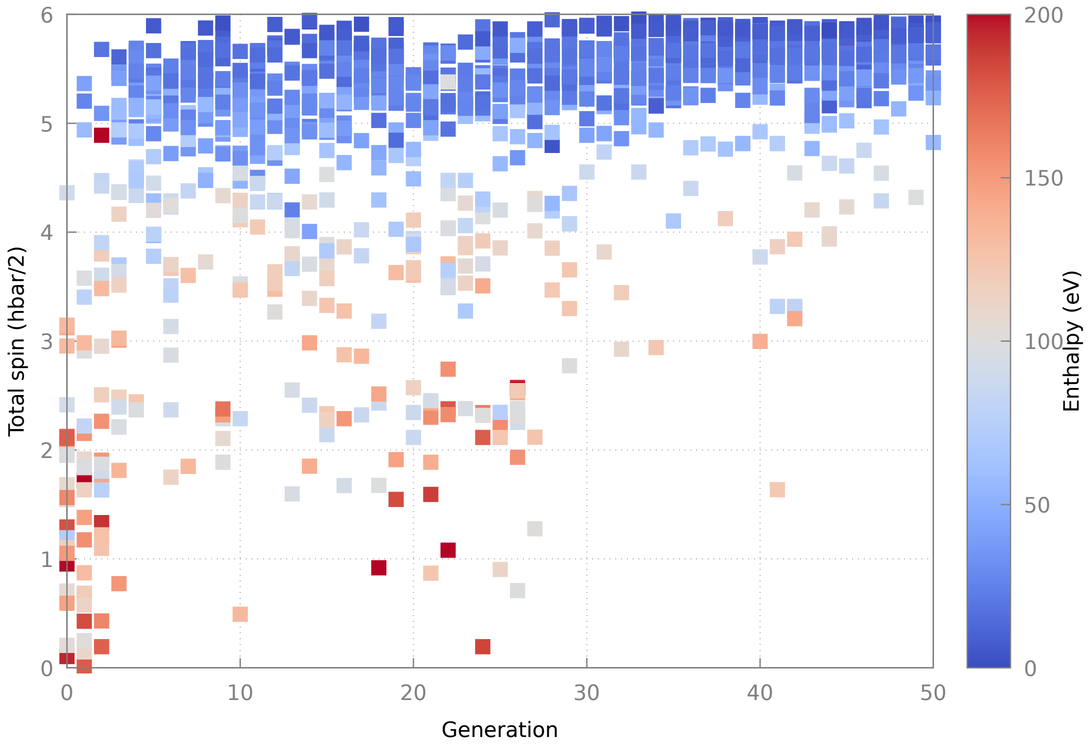

Figure 6.4: The range of spins found per generation for the FM-parameterised LJ+S model. It can be seen that the GA favours searching in the high spin region, where the lowest enthalpy structures are found.

6.4.2 Final structure

The final lowest enthalpy structure is shown in figure 6.5. Common neighbour analysis by the software visualisation package Ovito reveals this structure to be FCC.

6.5 An antiferromagnetic parameterisation

By reversing the sign of A in the magnetic Lennard-Jones potential, an AFM parameterisation can be formed. This was done, and the GA was run on it to search for the ground-state magnetic and crystal structure. The same parameters were used as in the FM parameterisation, and it was left to run for 80 generations.

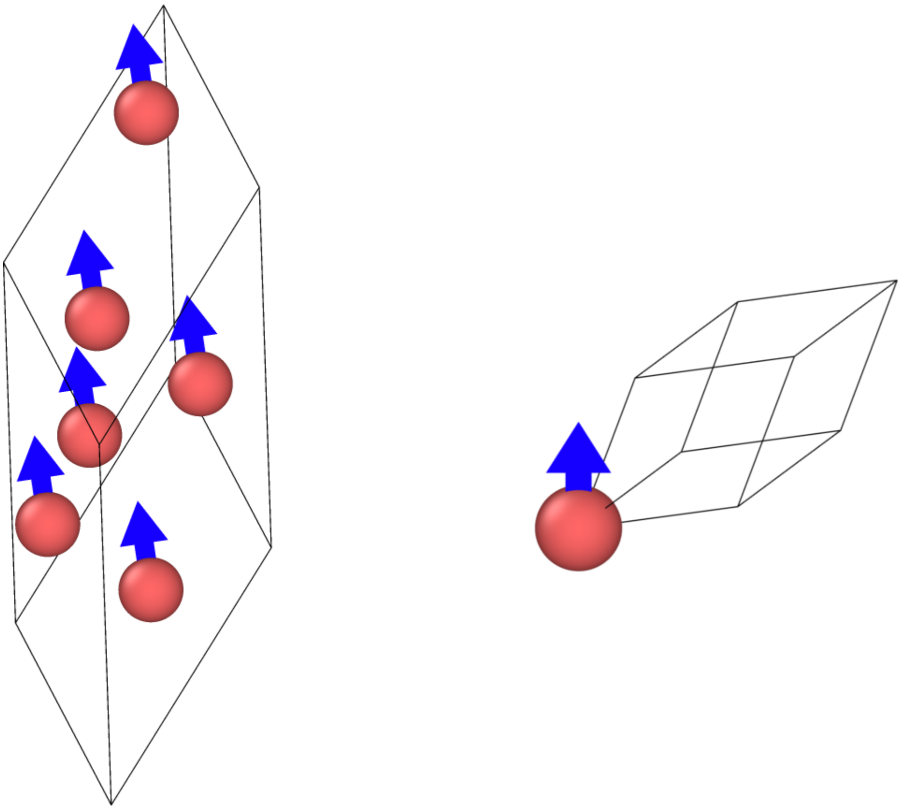

Figure 6.5: Lowest enthalpy structure found by the GA showing the spins aligned (left) and the associated FCC primitive cell (right). The arrows represent the unit spin vectors on each atom.

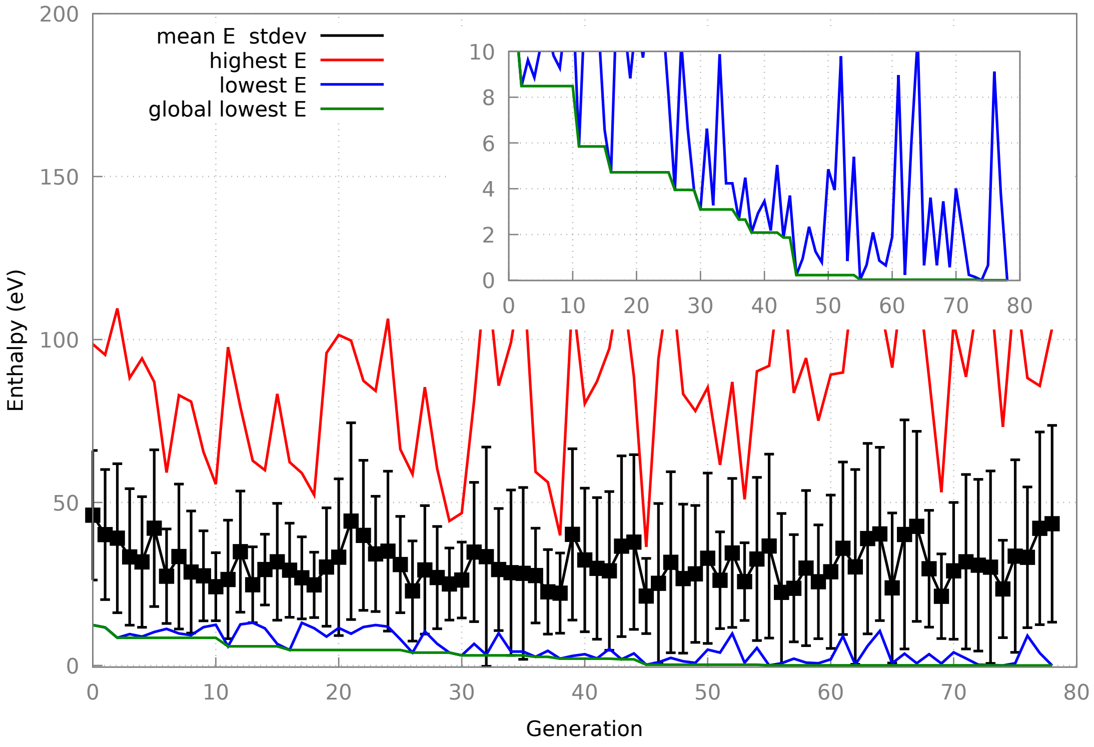

Figure 6.6: The range of enthalpies found per generation for the AFM-parameterised LJ+S model.

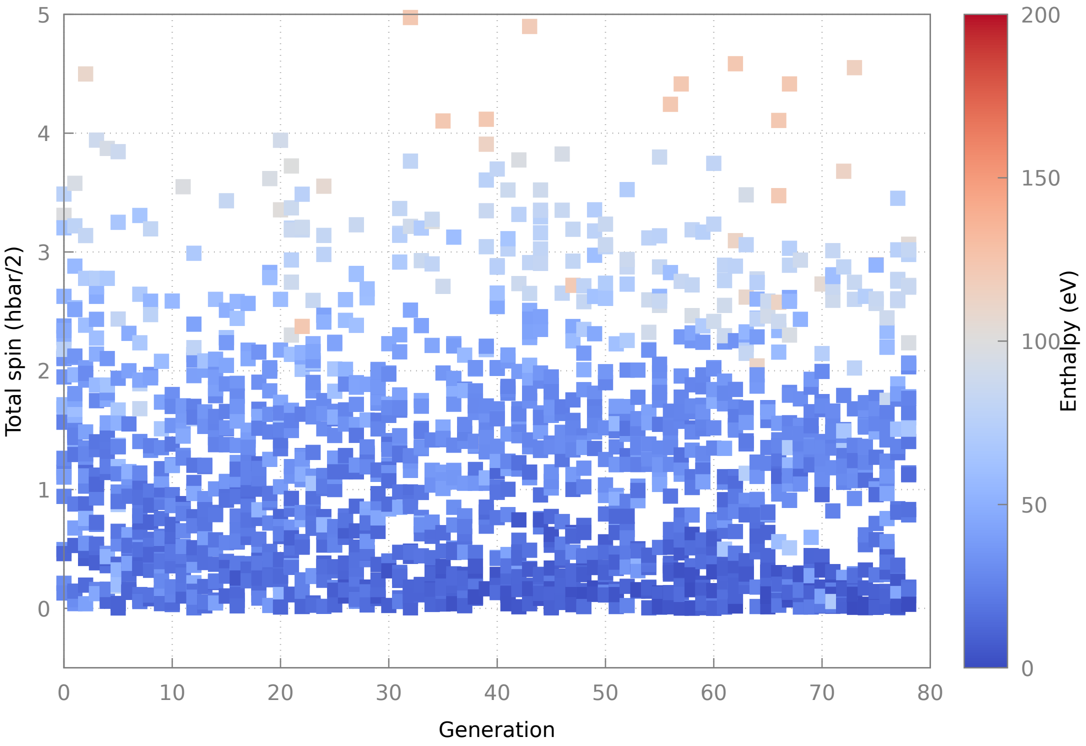

Figure 6.7: The range of spins found per generation for the AFM-parameterised LJ+S model. It can be seen that the GA favours searching in the low spin region, where the lowest enthalpy structures are found.

6.5.1 Final structure

The final lowest enthalpy structure is shown in figure 6.8. Common neighbour analysis by the software visualisation package Ovito reveals this structure to be BCC, with opposite spins on the (0,0,0) and $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ sites of the conventional BCC unit cell.

6.6 Summary

A new model potential was proposed as a way of testing the performance of the magnetic GA. The potential was based upon the Lennard-Jones potential for noble gasses and the Heisenberg model for lattice spins. This potential was parameterised to the aligned and anti-aligned iron dimers, and the ground state structures explored by the GA.

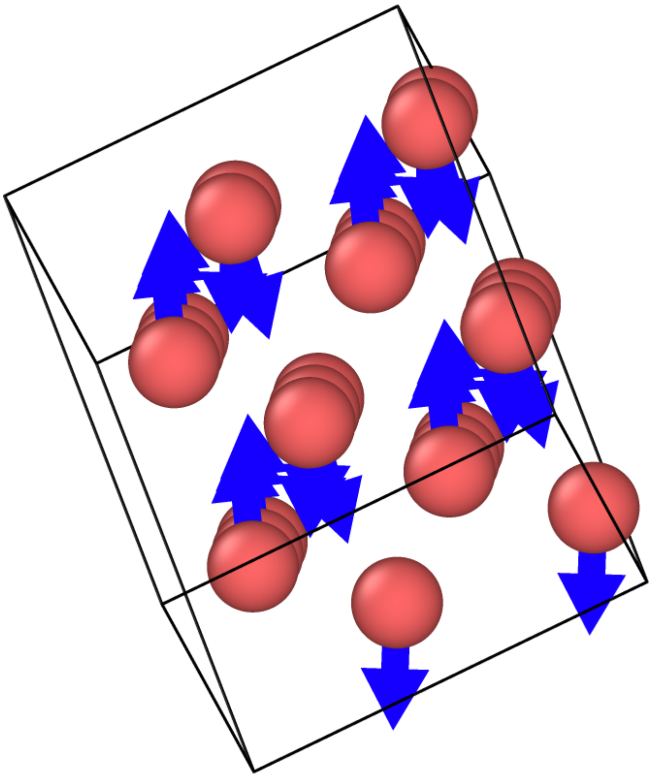

Figure 6.8: 2x2x1 unit cells of the lowest enthalpy structure found by the GA for the AFM parameterised LJ+S potential. Common neighbour analysis reveals this structure to be BCC.

The GA was able to successfully focus it search in the right portion of the spin state, favouring FM structures for the FM parameterisation and AFM structures for the AFM parameterisation. The FM parameterisation was found in the lowest configuration in an FCC lattice, with all spins aligned. The AFM parameterisation however was found in a BCC structure, with spins anti-aligned between the two conventional BCC atomic sites.

Chapter 7

Polar oxide interfaces

7.1 Introduction

In this chapter, the GA is used to investigate disorder at the NiO(111)/MgO(111) interface, where incremental interface mixing of nickel and magnesium is observed. The GA predicts that gradual mixing of nickel and magnesium as seen in experiment is energetically favourable. In addition, the nickel atoms retain the AFM ordering of the bulk, although both magnetic and atomic disorders are both discovered within thermal energy of the lowest enthalpy predicted structure.

7.2 Polar oxide interfaces

In recent years, polar oxides have been widely studied due to the wide variety of properties which they can exhibit. In particular, MgO has been studied for its high tunnelling magnetoresistance [95] making it a good candidate for magnetic tunnel junctions and NiO has attracted interest as an antiferromagnetic insulator. In addition to their interesting bulk properties, polar oxide interfaces can also exhibit interesting electronic surface properties. It is thought that the building up of electrostatic potential in thin film perovskite polar oxide interfaces leads to a 2D electron gas [96].

Traditional models of polar oxide interfaces predict that, due to this building up of electrostatic potential, atomic mixing was needed to stabilise the interface. However more recent research shows that effects such as electronic reconstructions or screening of the interface polarity can allow clean stable interfaces to form [97]. It is expected that oxides with strong ionic bonding such as NiO and MgO will experience the most significant effects of the polar discontinuity [3].

7.3 The NiO/MgO interface

Work by Lazarov *et al.* looked at the effects of growing NiO both on atomically clean MgO(100) and MgO(111) substrates [3]. In the case of both the polar NiO(111)/MgO(111) and the non-polar NiO(100)/MgO(100) interface, interfacial mixing occurred over around 8 cation planes, with nickel content increasing in approximately linear increments until the full NiO phase was formed, as seen in figures (7.1) and (7.2). In the case of the NiO(100)/MgO(100) interface, HAADF imaging reveals a zigzag structure at the interface, suggesting that nickel diffuses along the {111} planes, as seen in figure (7.1).

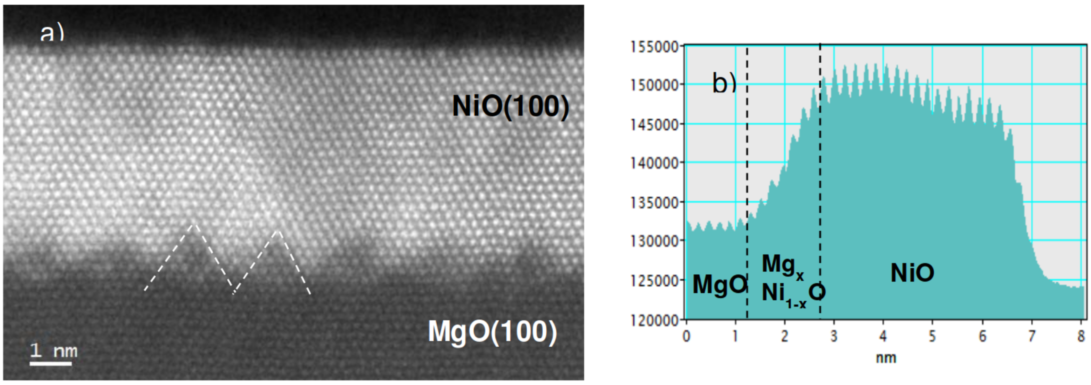

Figure 7.1: Experimental structure of the NiO(100)/MgO(100) interface. HAADF imaging (a) show the formation of zigzags along the {111} planes (white dashed lines). (b) shows the intensity profile across the interface. Pictures from [3]

While it is known from cluster-expansion calculations that NiO and MgO will mix in the bulk [98], it is not known why this particular type of mixing at the (111) interface is observed in these experiments. Similar mixing in the (100) interface suggests that the mixing is not due to the polar nature, however it is not known whether

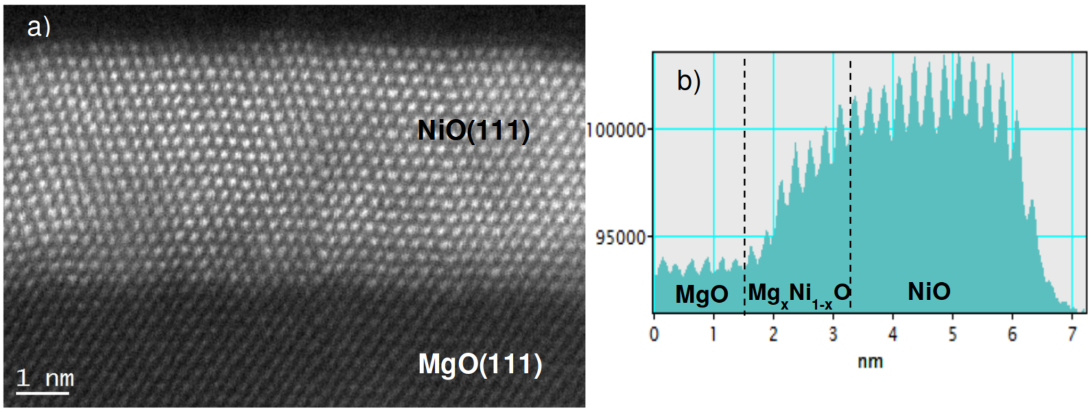

Figure 7.2: Experimental structure of the NiO(111)/MgO(111) interface. Intensity profile (b) shows incremental increase in nickel content across approximately 8 Ni/Mg layers. Pictures from [3]

it is driven by the energetics of the interface or by the dynamics of the growth method. In addition, it is not known if the NiO retains its AFM spin structure at the interface, or if any magnetic reordering occurs. For these reasons, the structure of the NiO(111)/MgO(111) interface was investigated using the magnetic GA.

7.4 Interface structure search with the GA

The calculation was performed on a 2x2x3 supercell of MgO on top of a 2x2x3 supercell of MgO. The clean interface is shown in figure 7.3. In order to preserve coupling to the bulk, the top 4 atomic layers of MgO and the bottom 4 atomic layers of NiO were fixed to the bulk structure, and not optimised by the GA.

Since both the pure MgO and NiO structures and the experimental interface grow in a rock salt structure, the atoms were restricted to these lattice sites in the GA. In order to achieve this, the mutation rate and amplitude were chosen to be 0. In addition, permutations were only performed on the Ni/Mg sub-lattice sites since there is no evidence of mixing between the two sites experimentally. The permutation rate was chosen to be 0.05 per atom was used for the first 20 generations, reducing this to 0.02 per atom afterwards. This was chosen to promote mixing early on with an average of 1-2 swaps per member, the reducing this down to ~0.5 swaps per member once the population was sufficiently diverse.

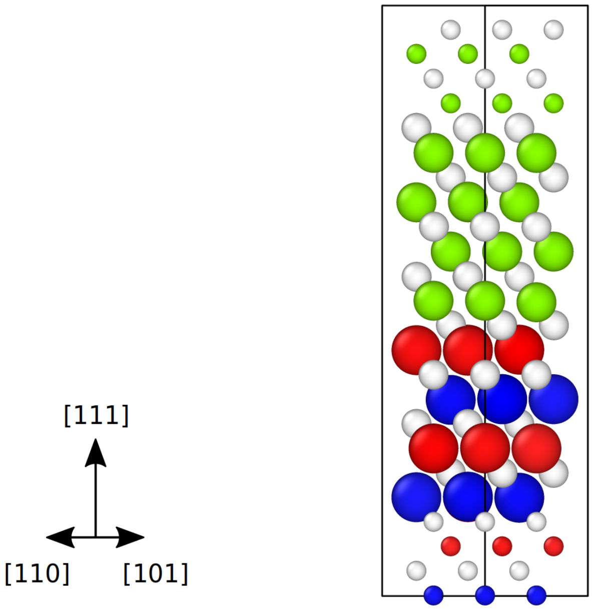

Figure 7.3: Structure of the clean MgO/NiO interface showing the oxygen (white), magnesium (green) and nickel (red=spin up, blue=spin down) atoms. Smaller atoms at the top and bottom represent atoms who's position and spin were constrained in the geometry optimisation.

Since the region of interest was at the NiO/MgO interface, a single periodic cut was made for crossover through the centre of the interface, with an amplitude of up to 6 atomic layers in the z direction.

The structures were allowed to locally optimise the atomic positions to a tolerance of 1 meV to reduce the effect of atomic displacements from enthalpy differences between structures. Since these structures grown on significant amounts of MgO, the lattice vectors were fixed for the calculations. Since there is only a 1% lattice mismatch between NiO and MgO [99], this is a reasonable approximation.

For each structure, the total energies were computed using DFT simulations in the CASTEP code. The calculation used a cutoff energy was 700 eV and k-point sampling on a 6x6x2 Monkhurst Pack grid. The Nickel atoms had a Hubbard U value of 1.0 eV. This was chosen to allow direct comparison with previous work done by Hasnip [100].

To confirm that the parameters reproduced the correct physical results, the properties of bulk NiO were calculated. NiO was calculated to be AFM with a net spin of 0.32×10^{-12} $\hbar/2$ and a spin of 1.40 $\hbar/2$ on each of the nickel atoms. The Kohn Sham

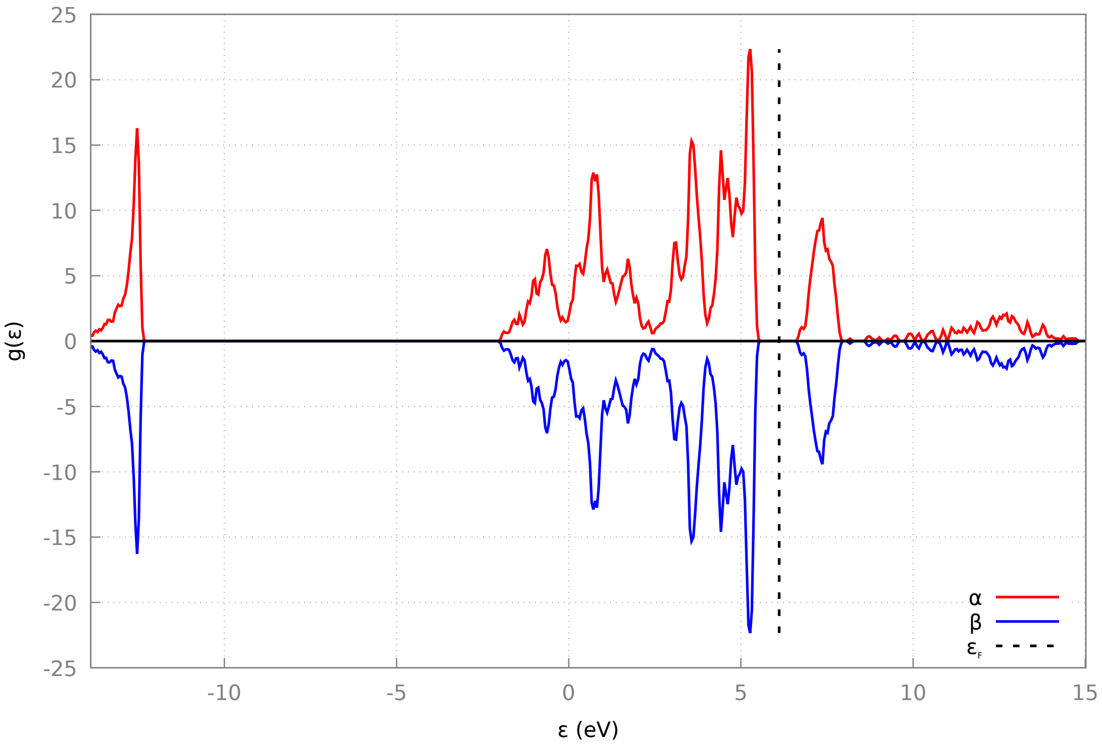

Figure 7.4: Calculated density of states for NiO using the simulation parameters.

band gap of NiO was calculated to be 0.48 eV, this can be seen in the density of states in figure 7.4. While this is below the experimental value of around 4 eV, it correctly captures the insulating nature of NiO. A more accurate band gap could have been attained with a higher value of U, however this can have the effect of unphysical flattening of the Ni d states, leading to incorrect results [101].

7.5 Performance of the GA

7.5.1 Convergence

The GA was run for 51 generations, at which point no new best structure had been found for 17 generations, and the overall lowest enthalpy structure had been rediscovered in 4 different generations. Figure 7.5 shows the distribution of energies of the final selected population at each generation and figure 7.6 shows the same data for the children generated each generation.

As can be seen in figure 7.5, the vast majority of the structures found by the GA

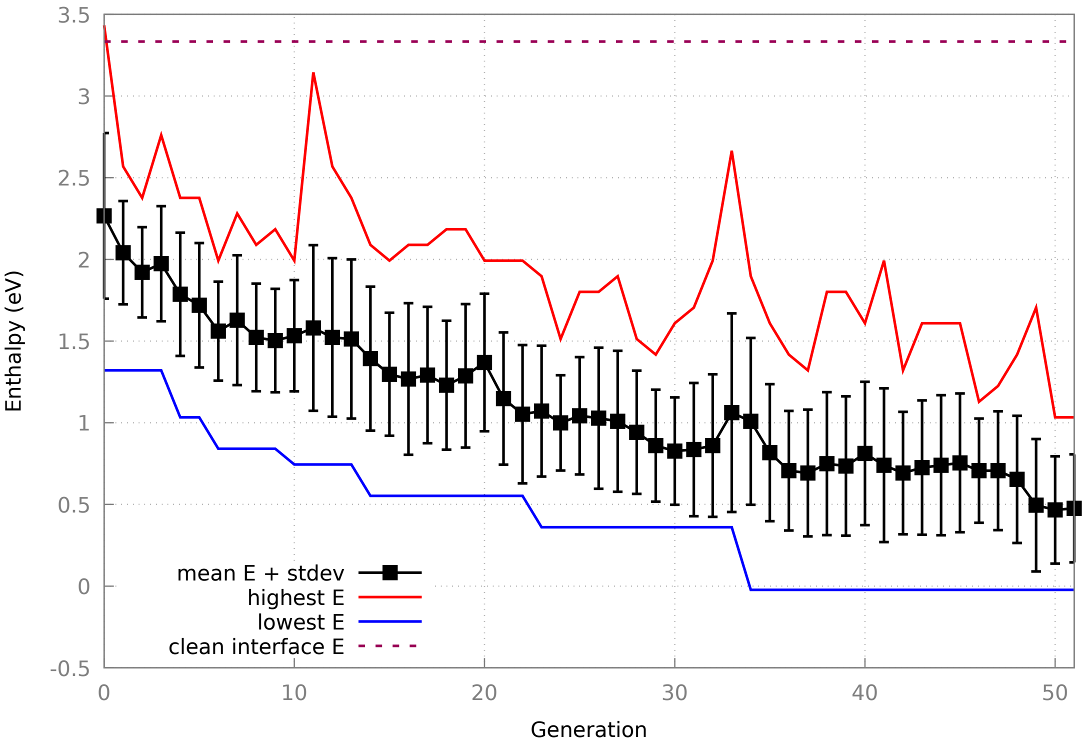

Figure 7.5: Highest (red), mean (black) and lowest (blue) enthalpies of the population against generation number for NiO/MgO. The dashed lined represents the enthalpy of the clean interface.

were significantly lower in energy than the clean interface, suggesting that interfacial mixing is strongly favoured. In addition, after around generation 20, the average energy of the structures in a given population was below the lowest enthalpy structure of the first 3 generations. Throughout the calculation, neither the range or standard deviation in the enthalpies of the population or the children dropped significantly, demonstrating that the population was not stagnating.

7.5.2 Population spin distribution

Figure 7.7 shows a plot of the enthalpy of each structure against the total spin of that structure. It can be seen that all structures found had atoms of nonzero spin, with the sum of the atomic spin magnitudes ranging from around 35 $\hbar/2$ to around 40 $\hbar/2$. It can be seen that the lowest enthalpy structures resulted in a total spin of 0, demonstrating that the lowest energy structures found by the GA are all AFM. It is important to note that the GA preferentially sampled the low ($< 5\ \hbar/2$) structures

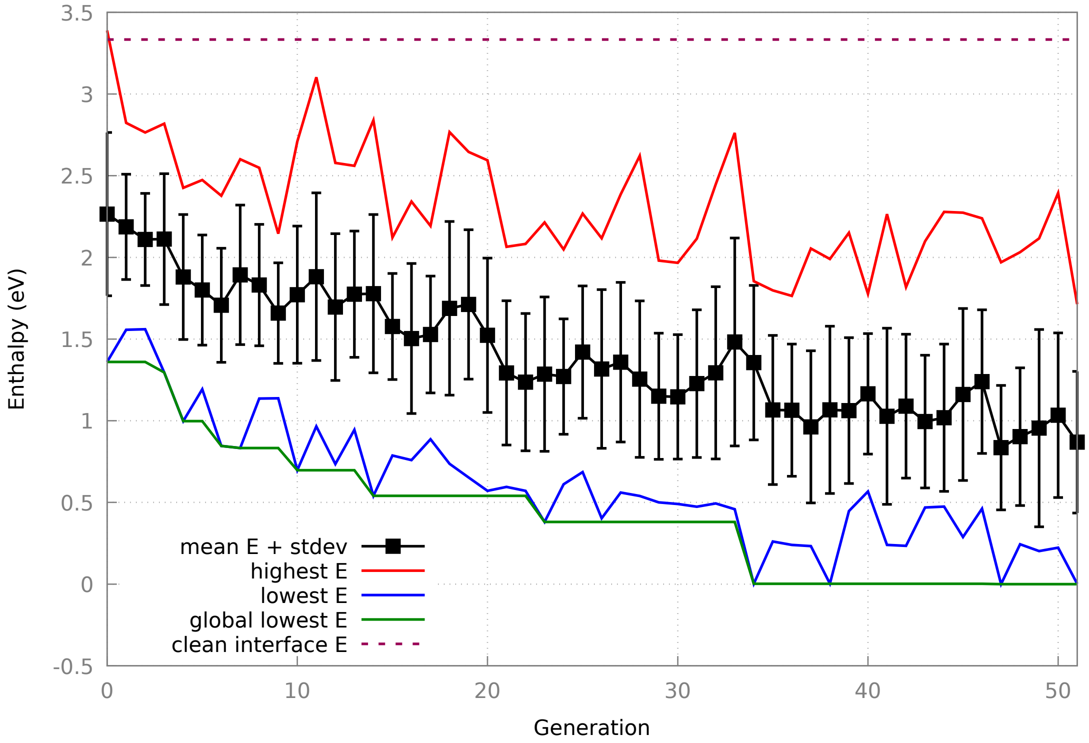

Figure 7.6: Highest (red), mean (black) and lowest (blue) enthalpies of the new children found against generation number for NiO/MgO. The dashed lined represents the enthalpy of the clean interface and the green line represents the lowest overall structure at each given generation.

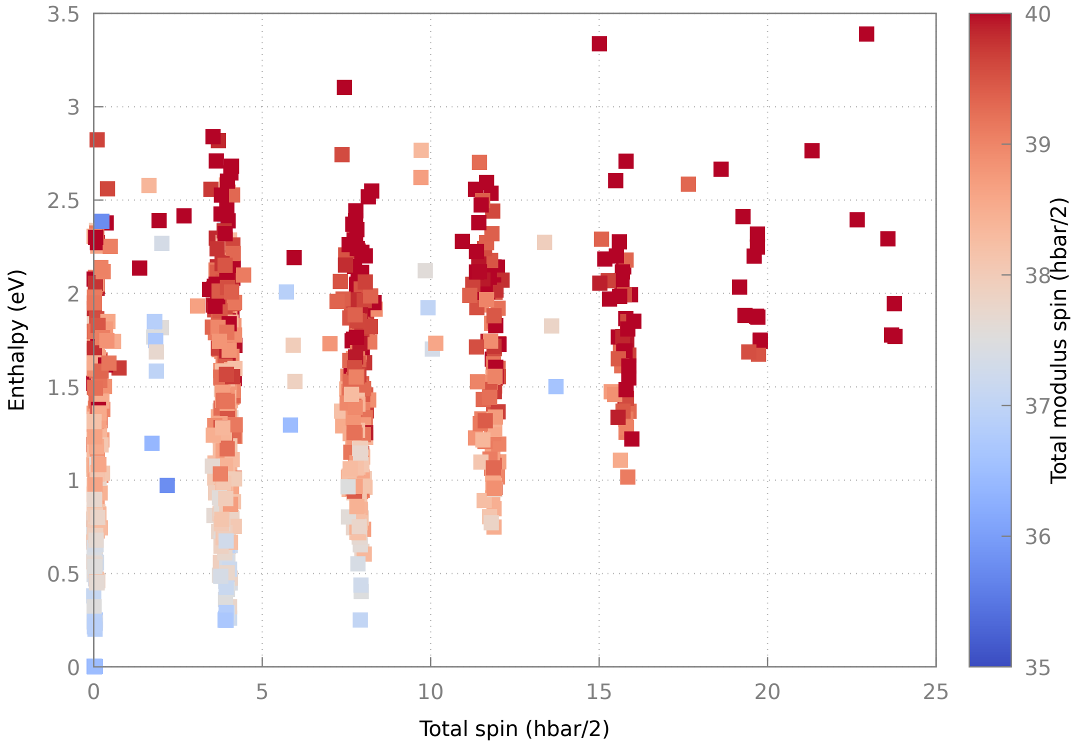

Figure 7.7: Enthalpy of each structure plotted against the total spin of that structure. The colour represents the sum of the modulus of the spins.

since these typically related to lower enthalpy structures. However, the GA also sampled regions of higher spin, demonstrating its ability to explore a wide variety of magnetic structures.

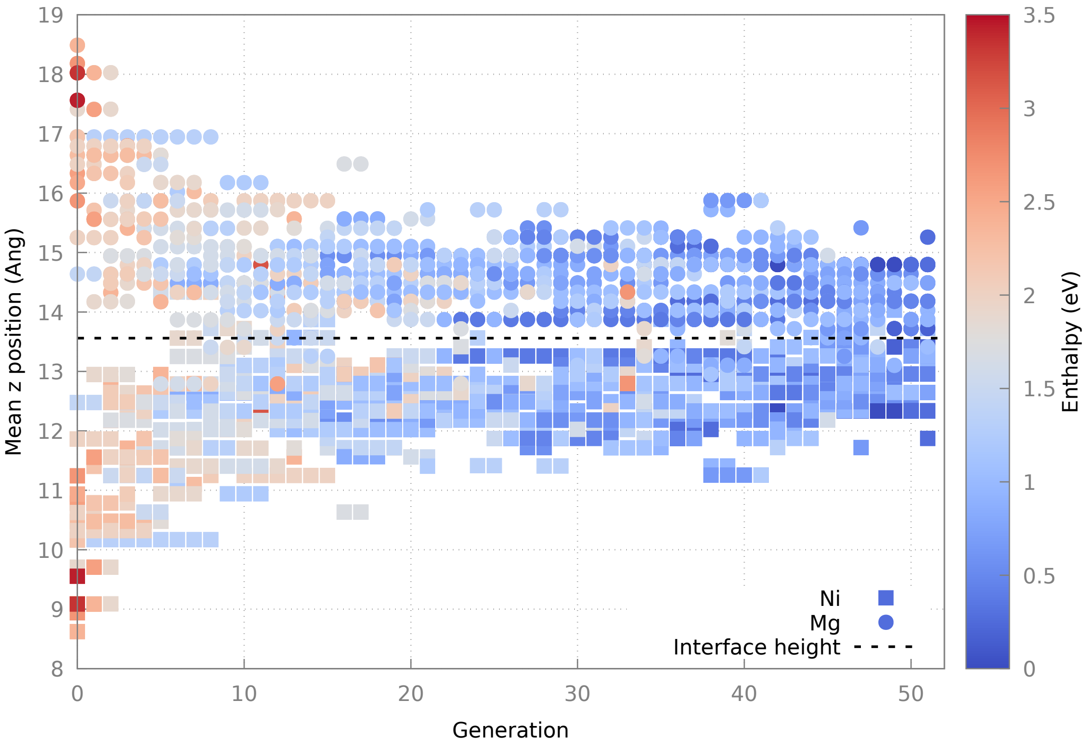

Figure 7.8: Mean z position of the nickel atoms (square) and the magnesium atoms (circle) for each structure against generation number for NiO/MgO. The colour represents the enthalpy of the structure.

Figure 7.8 demonstrates the amount of mixing between NiO and MgO layers by displaying the mean position of nickel and magnesium along the z direction. Initial mixing occurs up to around generation 7 due to the tendency for random swaps to disorder the structure. However, after around generation 14, the structures start to separate slightly, with the lowest enthalpy structures having a separation of mean atomic positions between the nickel and magnesium atoms of around 2.3 Å. This is consistent with the observation experimentally that the materials prefer to gradually mix, rather than either fully mixing or forming clean interfaces.

Figure 7.9 shows just the mean z distribution of just the nickel atoms colored by the total spin of the structure. There is no clear correlation between the spin polarisation of the structure and the amount of mixing of nickel atoms into the MgO after around generation 15. The apparent trend for more FM-like structures closer to the

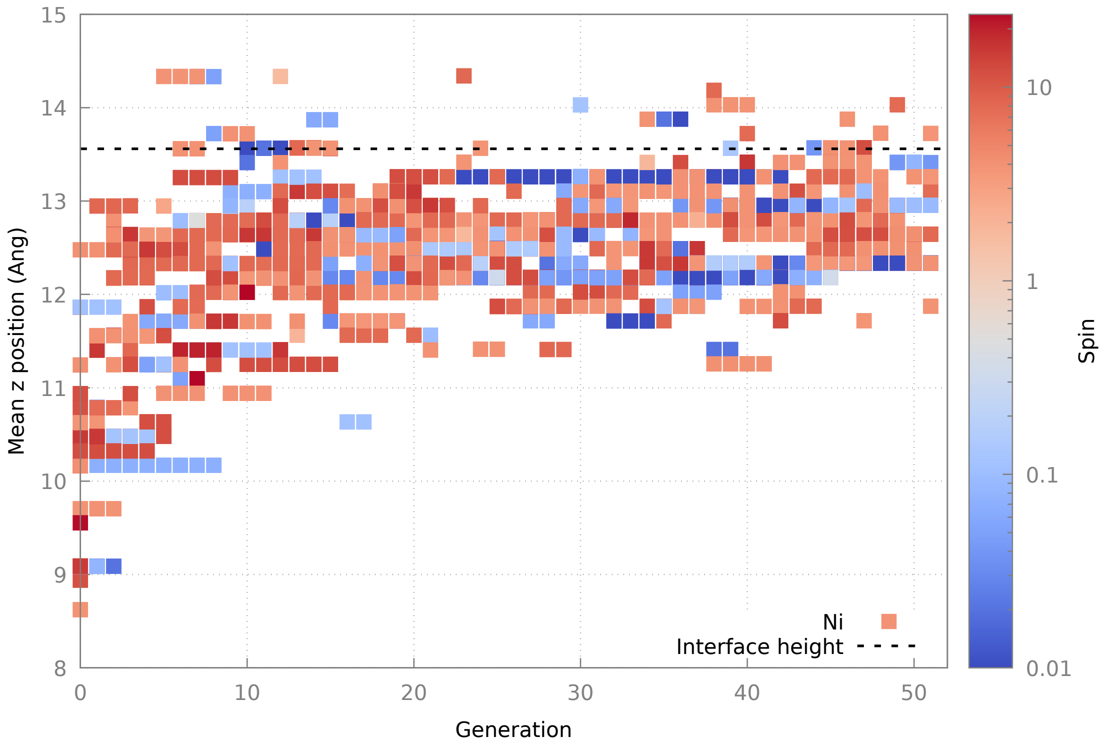

Figure 7.9: Mean z position of the nickel atoms for each structure against generation number. The colour represents the total spin of that structure.

clean interface (lower mean z position) in early generations is most likely due to the structures randomly generated in generation 0 rather than that being the more favourable configuration.

7.6 Resultant structures

The atomic structure of the lowest enthalpy structure is shown in figure 7.10. As observed in the experimental data, the number of nickel atoms per cation layer varies gradually, with (3,3,2,2,1,1,2,2) nickel atoms per layer. Based on this ordering, it is possible that a (3,3,2,2,1,1) ordering would be more stable, however constraints on the stoichiometry of the cell make this unreachable by the GA.

In addition to the concentration of the nickel atoms in the cation layers, it is clear that the AFM magnetic structure is retained from the bulk.

The next two lowest enthalpy distinct structure is shown in figure 7.11. The second lowest structure is equivalent to the lowest enthalpy structure with the spin of a

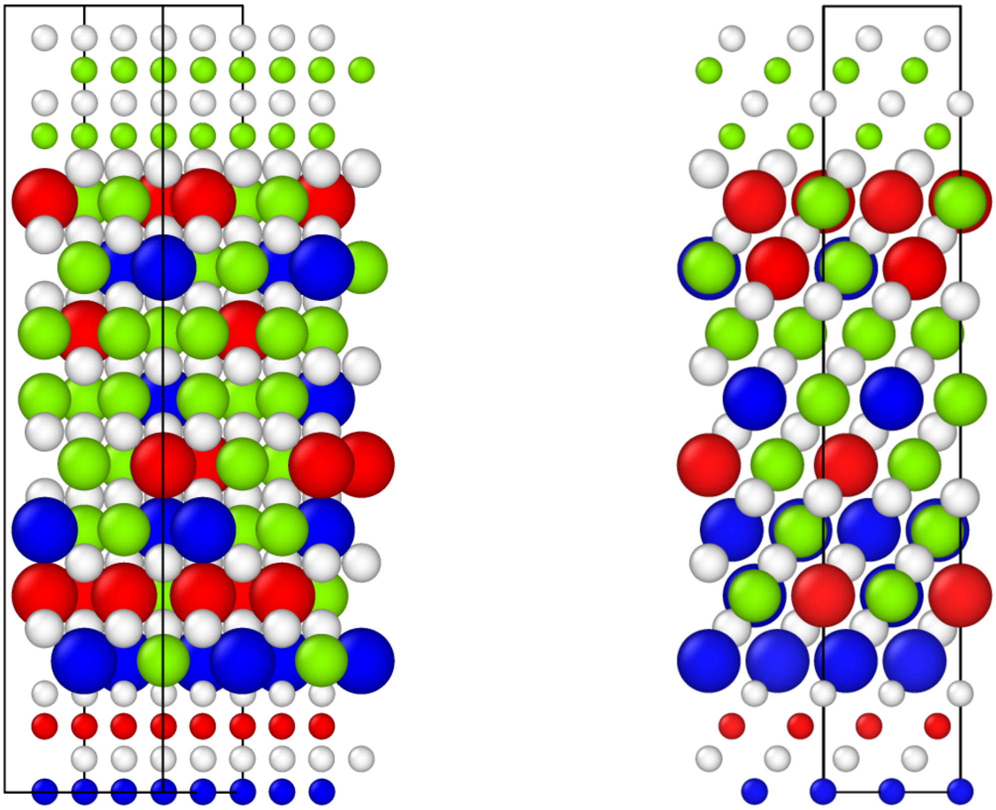

Figure 7.10: 2x1x1 unit cells of the lowest enthalpy MgO/NiO interface structure found by the GA from the front (left) and from the side (right). The gradual variation of nickel atoms (red=spin up, blue=spin down) and the AFM ordered layers can be seen.

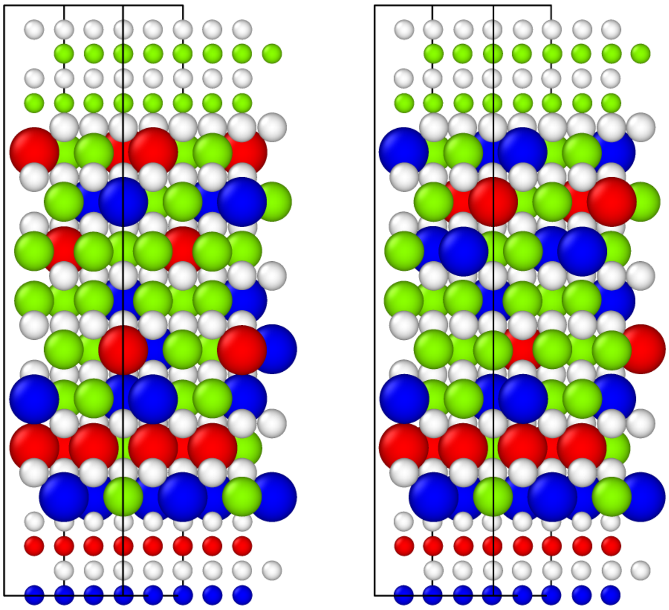

Figure 7.11: Structure of the second lowest enthalpy structure with a Ni spin flipped (left) and third lowest enthalpy structure with reordering of the cation layers.(right)

single nickel atom flipped in the 4th unconstrained cation layer up. This structure is 0.250 eV above the lowest energy structure. Similarly, the third lowest enthalpy distinct structure is 0.251 eV above the lowest enthalpy structure. This structure is similar to the lowest enthalpy structure, but with:

a) Reordering of the atomic layers (3,3,2,1,1,2,2,2)
b) The spins of the 6th - 8th cation layers reversed.

Both of these structures may be thermally accessible at the annealing temperatures of around 375 K, since the enthalpy/atom for them is around 0.005 to 0.0025 eV / atom, depending on which atoms are considered.

7.7 Summary

This chapter explored the mixing of nickel and magnesium in the NiO(111)/MgO(111) interface. The GA was able to explore a range of atomic and magnetic structures, sampling mixed and unmixed interfaces, along with a range of magnetic orderings. The lowest enthalpy structures appear to gradually mix nickel and magnesium along the interface, preferring single steps in mixing concentration rather than retaining entire planes of similar atoms. This is supported by experimental results showing similar incremental mixing at the interface, although over a length scale beyond what is currently feasible with the GA. The magnetic ordering of the nickel atoms appears to favour AFM ordering as is observed in the bulk, with each atomic plane anti-aligned to the previous plane. Since mixed AFM energies emerge as the lowest enthalpy structures, this is likely the driving mechanism for the emergence of mixed regions

Chapter 8

Heusler alloy interfaces

8.1 Introduction

In this chapter, the GA is used to investigate an unknown phase which forms when the $Co_2FeAl_{0.5}Si_{0.5}$ (CFAS)/n-Ge interface is annealed. A Half Heusler structure is proposed as a candidate for this phase. Unlike CFAS, this new phase is not half metallic, making it undesirable for spintronic devices where CFAS may be used.

8.2 Heusler Alloys

One class of materials which can exhibit half-metallicity are the Heusler alloys. Heusler alloys are ternary alloys with the stoichiometry X_2YZ, where X and Y are transition metals and Z is a p block element. The structure of Heusler alloys can be described in terms of 4 interlocking FCC sub-lattices with positions $X_1=(0,0,0)$, $X_2=(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ $Y=(\frac{1}{4},\frac{1}{4},\frac{1}{4})$, $Z=(\frac{3}{4},\frac{3}{4},\frac{3}{4})$ [102, 103]. The Heusler structure is shown in figure 8.1. In addition to full Heusler alloys, there also exists a class of material known as the half Heusler alloy. This is the same structure as the full Heusler but with vacancies on the X_2 sites.

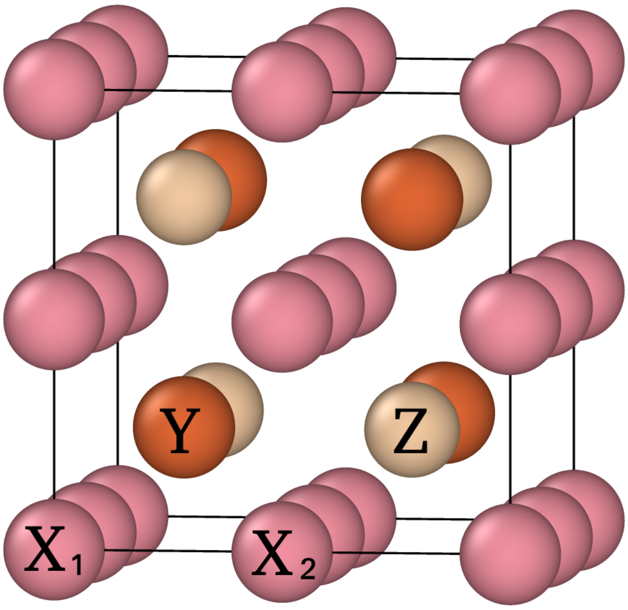

Figure 8.1: Heusler alloy structure

8.2.1 Half metallicity

A number of Heusler alloys have been predicted to exhibit half metallic [103]. In particular, Co_2MnSi (CMS) and Co_2FeSi (CFS) have been studied as potential device materials for a number of spintronic applications, such as spin valves and magnetic tunnel junctions. However while these materials are predicted to be half metallic as pure crystals, the Fermi energy sits very close to the conduction band for CFS and very close to the valance band for CMS, as illustrated in figure 8.2. As a result, the half metallicity is very sensitive to effects such as disorder in the lattice [104].

In order to reduce these effects, the material can be doped in order to move the Fermi energy closer to the middle of the band gap. One option is to try a hybrid of CFS and CMS, i.e. $Co_2Fe_xMn_{1-x}Si$ (CFMS). While this does successfully move the band gap, it also disrupts the up-spin bands, reducing their metallic nature. One other option is to p-dope the silicon site with aluminium in order to reduce the Fermi energy, resulting in $Co_2FeAl_{0.5}Si_{0.5}$ (CFAS). This turns out to be far less disruptive to the up spin conduction states.

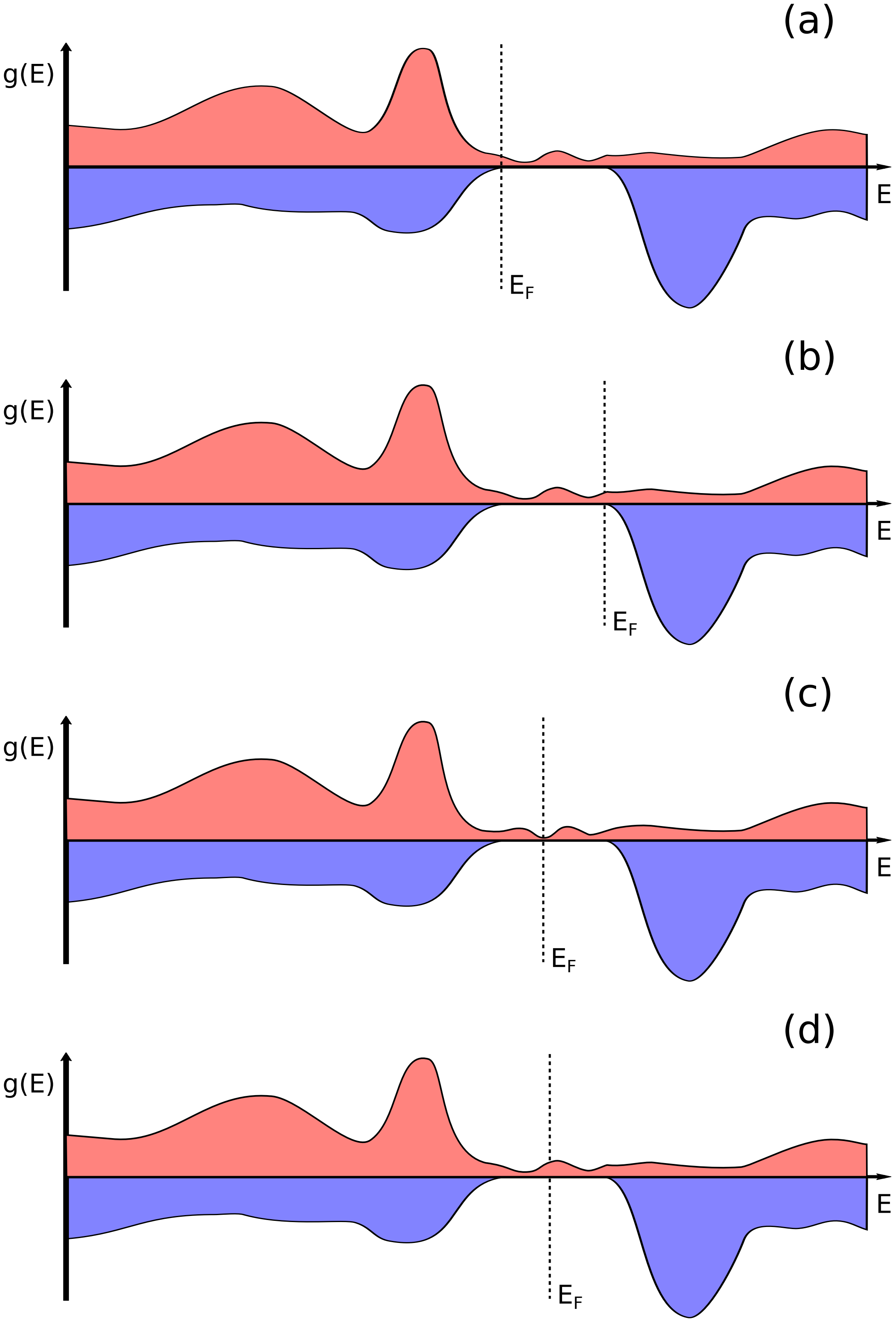

Figure 8.2: Illustration of the spin density of states for Heusler alloys. Represented are situations typical to (a) CMS, (b) CFS, (c) CFMS and (d) CFAS. The effect of the composition on the Fermi energy E_F is demonstrated.

8.2.2 Heusler/contact interfaces

In order to use these materials for device applications, Heusler alloys need to be attached to electrical contacts. For a number of spintronic applications of which Heusler alloys could be used for, both metallic and semiconducting contacts are required. When considering a material for the contacts, a number of considerations need to be made, including:

- Lattice match between Heusler and contact material,
- Heusler/contact interface roughness or mixing,
- Electronic effects at the interface.

It is known that the half metallic nature of Heusler alloys at the interface is sensitive to the exact configuration of the interface atoms, and can affect the device's effectiveness. For example, CMS grown on a silver surface changes significantly depending on whether it terminates with a Co layer or a Mn/Si layer [105].

8.3 The CFAS/Ge interface

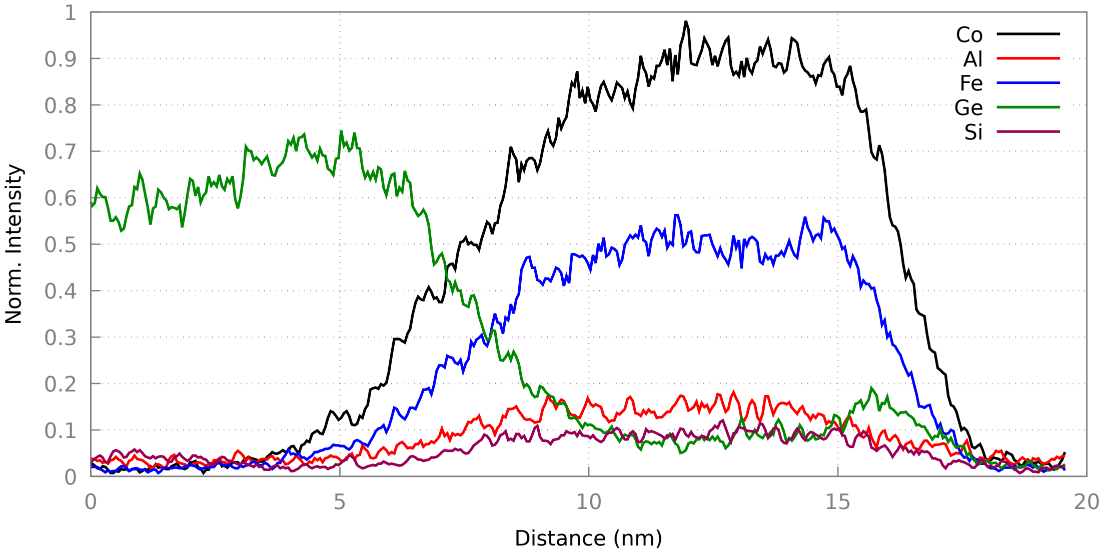

Figure 8.3: Energy Dispersion Spectroscopy (EDS) results across the un-annealed CFAS/n-Ge interface, showing significant mixing of the atomic species. Data from [4]

For the CFAS/semiconductor interface, germanium provides an extremely good

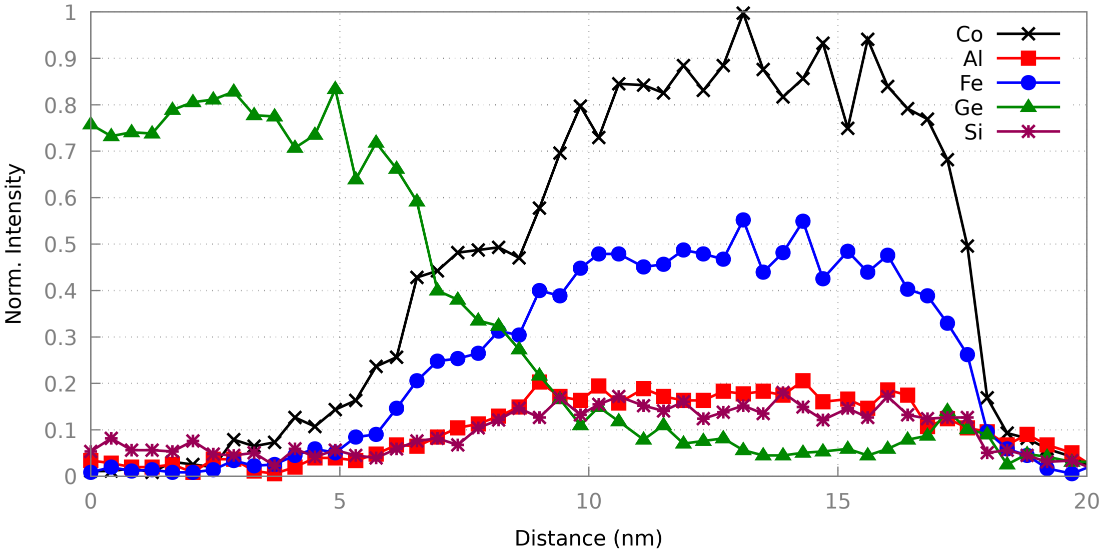

Figure 8.4: EDS results across the annealed CFAS/n-Ge interface, showing significant mixing of the atomic species. Plateaus can be seen around 6.5-9.0 nm in cobalt, iron and germanium. Data from [4]

lattice match with only a 0.2% mismatch, compared to around 4.5% for CFAS/Si. However, when CFAS is grown on n-type germanium, significant mixing at the interface can occur, as is shown in figure 8.3 [4].

In addition, annealing the sample at 575 K forms a plateau in the cobalt intensities and, to a lesser extent, the iron and germanium intensities. This suggests that a stable phase is forming at this point. Since this is in the centre of the interface region and all the EDS energies are about half of their respective bulk values, it is proposed that this structure contains half the number of atoms in the bulk CFAS and half the number of atoms in a germanium unit cell. This results in a $Ge_4Co_4Fe_2AlSi$ structure, containing 12 atoms, although the crystal structure of this interface phase is unknown. Since the electronic properties of this phase will affect the half metallicity of the CFAS/n-Ge interface, it would be useful to understand what it is. For this reason, the phase was investigated using the GA.

8.4 Interface phase investigation with the GA

The calculation was performed on the 12 atom unit cell: 4 germanium , 4 cobalt, 2 iron, 1 aluminium and 1 silicon. The GA was allowed to fully optimise the atomic positions, however the lattice was constrained to that of the full Heusler alloy, since this closely matches the germanium lattice, and it is not expected that this would change much over such a short region. The GA was run with a population size of 20.

A mutation amplitude of 3 Å was used for the ions, at a rate of 0.03 mutations per atom. This was chosen to allow on average one atom every 2 members to mutate, but significantly enough to hop lattice sites. A slightly higher permutation rate of 0.04 was chosen since, for the bulk Heusler, atomic swaps on the lattice sites can be comparable in energy to the ground state. If a similar atomic structure is found, it is likely that exploring this kind of disorder would be a good idea. The spin mutation rate was 0.06 per atom. This was chosen to be higher than the atomic mutation rate because not all atoms would be expected to magnetise, meaning that some mutations would get negated by the local optimisation.

Each member was optimised using the BFGS algorithm to a tolerance of 1 meV. The energies were computed using CASTEP, with a cutoff energy of 650 eV and k-point spacing of 0.04 $Å^{-1}$. The PBE functional was used in conjunction with Hubbard U values of 2.1 eV on the d orbitals of iron and cobalt, keeping with previous Heusler work [105]. Figure 8.5 shows the calculated density of states of bulk CFAS using these parameters, demonstrating that they reproduce the expected half metallicity.

8.4.1 Convergence of the GA

The GA was initially run for 50 generations. However, due to an insufficient niching distance, the population stagnated. The effects of this can be seen in figure 8.6, many of the children generated after generation 5 represent the same structure. While in this case the GA was able to find a lower energy structure, this behavior is characteristic of stagnation and often results in the calculation being stuck in a local minimum.

The GA was then re-run with an increased niching distance for 70 generations. This resulted in a broader range of structures throughout the calculation, as can be seen in figure 8.7. In addition, each generation had a broader range of structures, providing

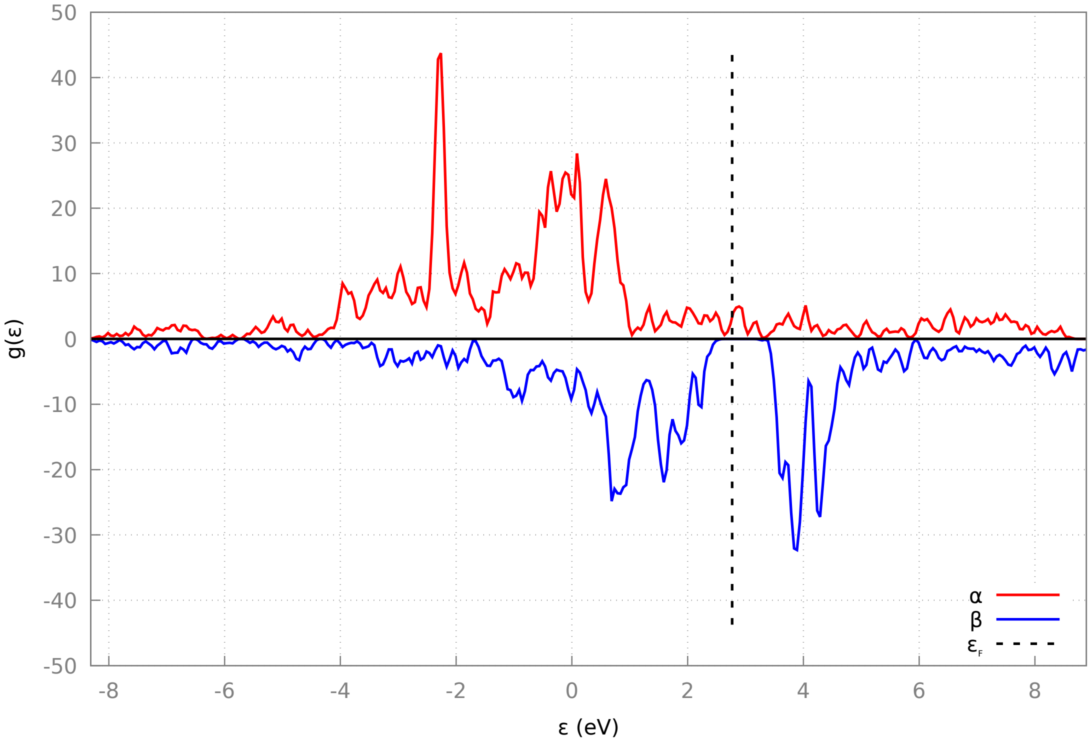

Figure 8.5: Calculated density of states for bulk CFAS using the simulation parameters.

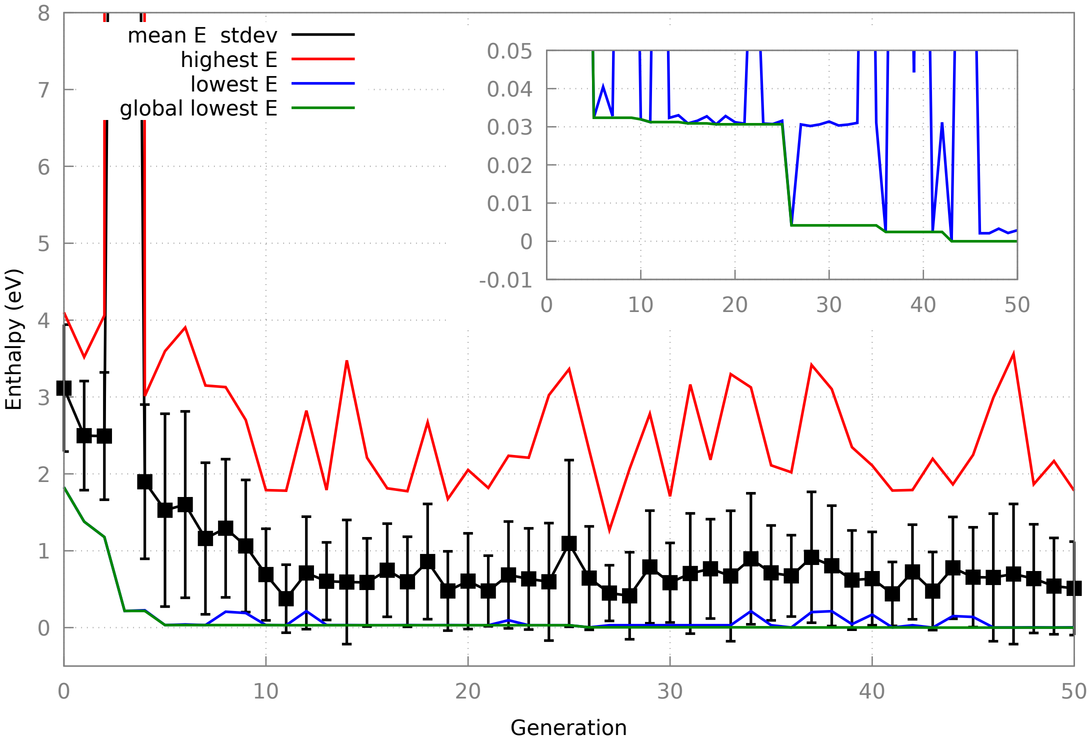

Figure 8.6: Convergence of the 1st GA run for CFAS/Ge interface structure. Stagnation of the population can be seen in later generations.

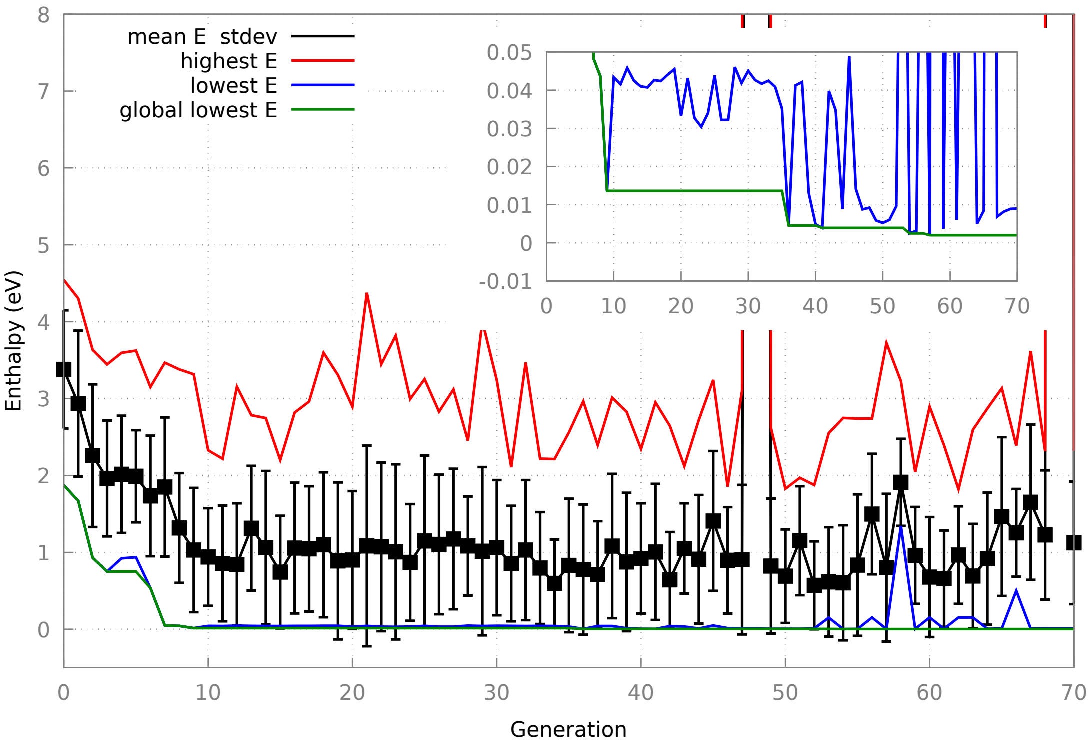

Figure 8.7: Convergence of the 2nd GA run for CFAS/Ge interface structure. Niching forces a wider range of structures to be explored.

more options for crossover.

Both GA runs found the same lowest enthalpy structure, although due to noise in geometry convergence, the first GA run found it with an enthalpy around 0.16 meV lower than the second run.

Figure 8.8 shows the distribution of magnetic structures found, plotted against their energy. Two clusters of members are found in the energy range which would be thermally accessible during annealing. The first of these are ferromagnetic, with a spin of about 8 $\hbar/2$ and a modulus spin of about 8.8 $\hbar/2$. The second cluster is antiferromagnetic, with a spin of about 0 $\hbar/2$ and a modulus spin of about 6.5 $\hbar/2$.

Figure 8.9 shows these regions in more detail. It can be seen that each cluster contains a number of structures. These are be atomic swaps of the lowest enthalpy structure in each magnetic state. Candidate structures are labelled A1-A3 for the AFM structures and F1-F4 for the FM structures.

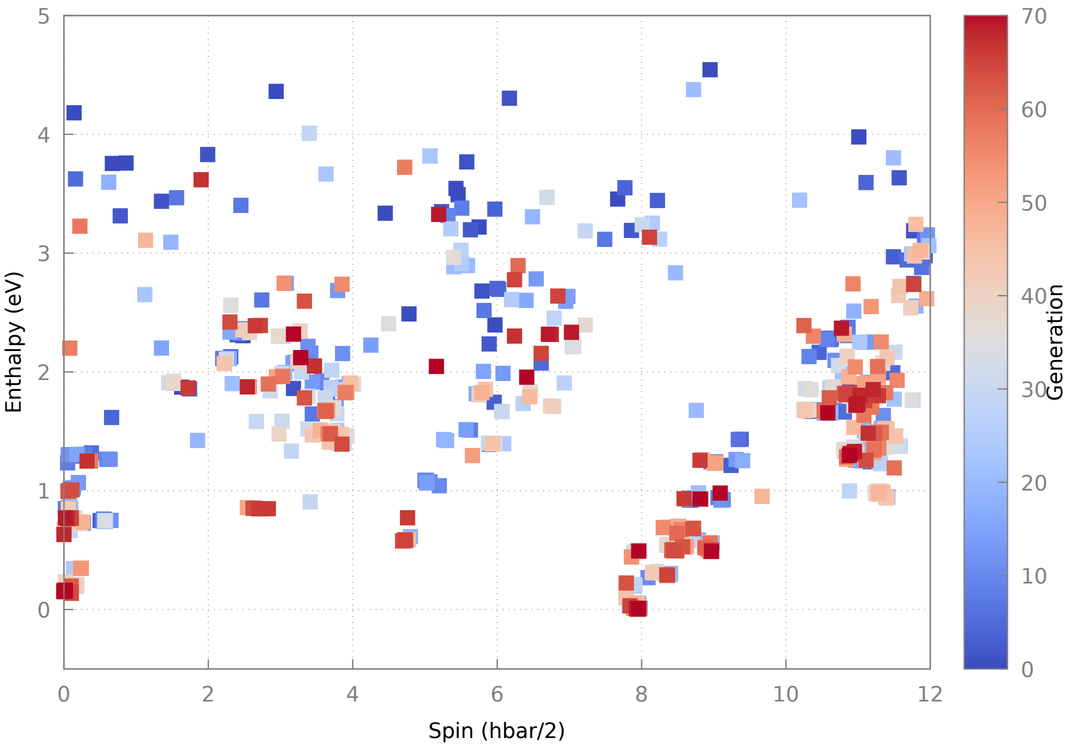

Figure 8.8: Total spin of the GA structures plotted against enthalpy.

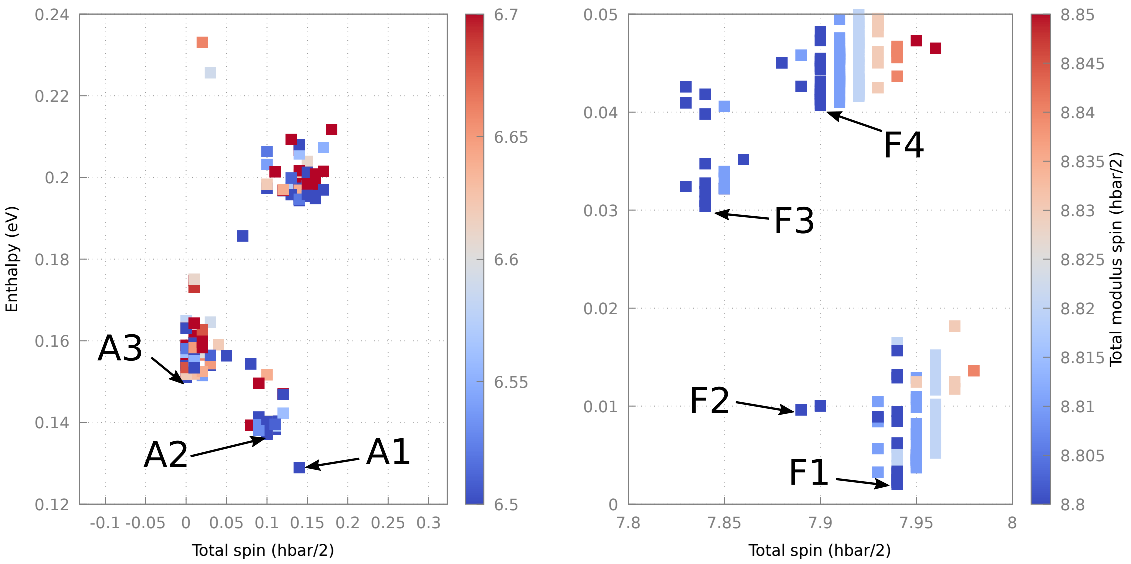

Figure 8.9: Clustering of structures around the AFM and FM configurations of the interface phase. The colors represent the total modulus spin.

8.4.2 Candidate structures

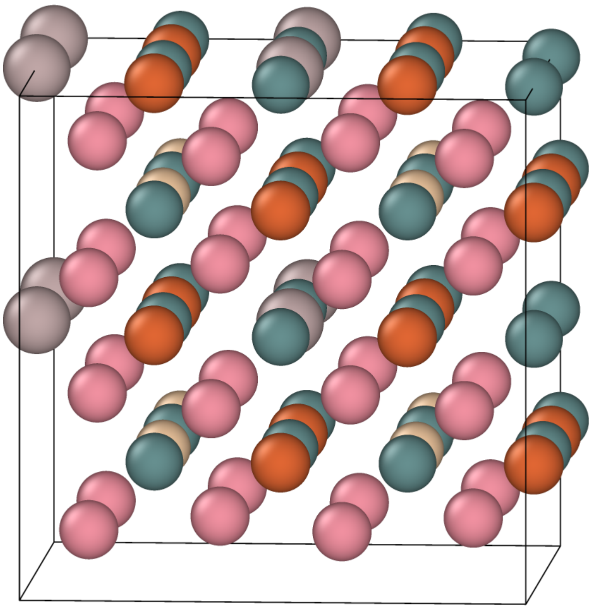

Figure 8.10: 2x2x2 unit cells of the lowest enthalpy structure found by the GA, showing the positions of the cobalt (pink), aluminium (grey), germanium (turquoise), silicon (gold) and iron (orange) atoms.

The lowest enthalpy structure F1, shown in figure 8.10 turns out to be a half Heusler alloy, where the X_1 site contains cobalt, the X_2 site contains a vacancy, the Y site contains the iron, silicon and aluminium, and the Z site contains germanium. The F2 structure is the same structure as F1. The F3 and F4 structures are the same structures with one germanium atom swapped with the silicon atom from the Y site, shown in table 8.1. According to the Hirshfeld analysis, in each of these structures the spin exists on the iron and cobalt, with a spin of around 3 $\hbar/2$ on the iron atoms and 0.5 $\hbar/2$ on the cobalt atoms.

Structure	**Enthalpy (eV)**	**Total spin ($\hbar/2$)**	**Disorder**
F1	0.0	7.94	None
F2	0.01	7.89	None
F3	0.03	7.84	Ge <-> Si
F4	0.04	7.89	Ge <-> Si
A1	0.13	0.14	2Ge <-> Si,Fe
A2	0.14	0.10	Ge <-> Fe
A3	0.15	0.00	None

Table 8.1: Table showing the energies and spins of the best structures, together with the lattice permutations in relation to F1.

Figure 8.11 shows the density of states for the 4 FM structures. Unlike the full CFAS structure, none of these structures are half metallic, with a significant number of states around the Fermi energy. It does not appear that the atomic disorder separating F1-4 has a significant effect on the electronic structure of the structures.

The 3 AFM structures A1, A2 and A3 sit around 0.15 eV above F1, making them thermally accessible at the annealing temperature of 575 K, with an enthalpy/atom of around 0.0125 eV. The A3 structure has the same crystal structure as F1 and the A1 and A2 structures are the same structures with some disorder. In the case of the AFM structures, this disorder lowers the enthalpy of the AFM structures. According to Hirshfeld analysis, all the spin exists on the iron, with about 3.1 $\hbar/2$ spin on one iron atom and about -3.1 $\hbar/2$ spin on the other.

Figure 8.12 shows the density of states of the 3 AFM structures. Again, none of the AFM structures are half metallic.

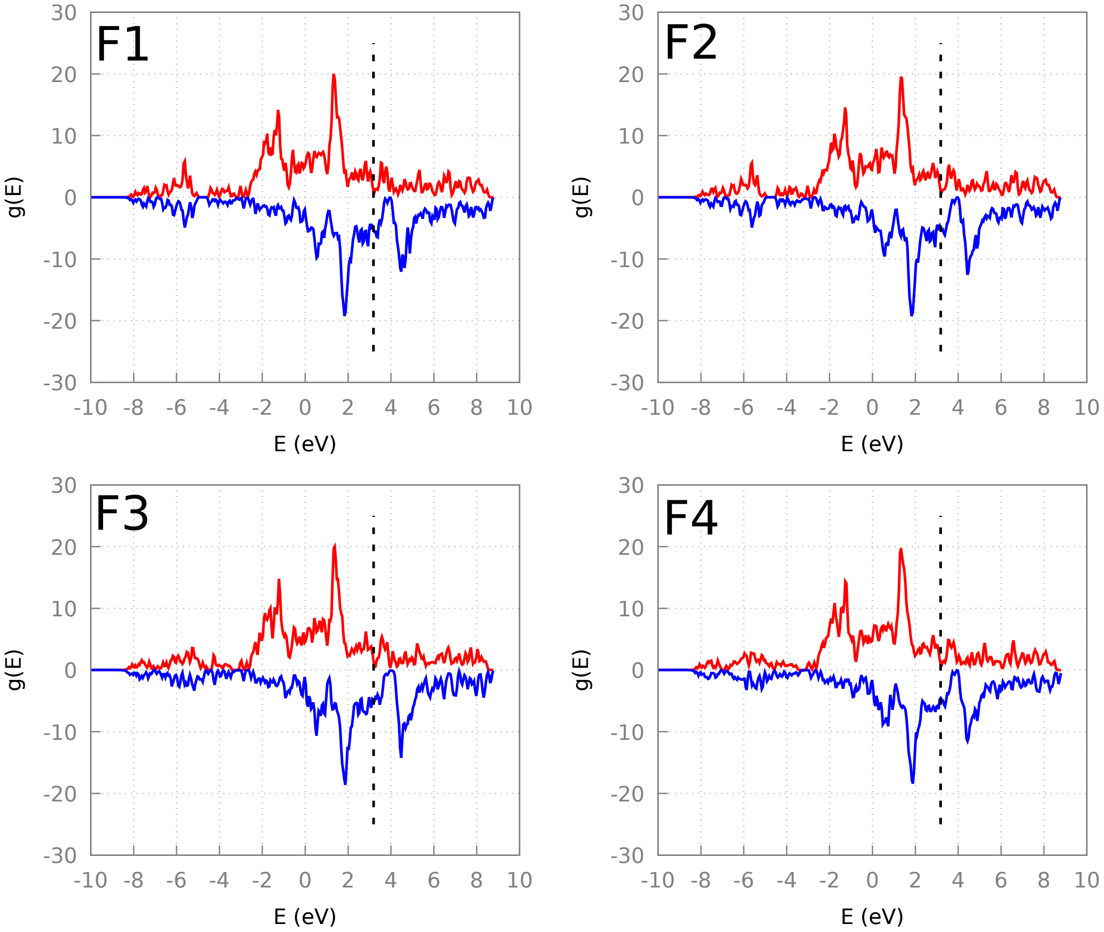

Figure 8.11: Density of states of the 4 low enthalpy FM structures. None of these structures are half metallic.

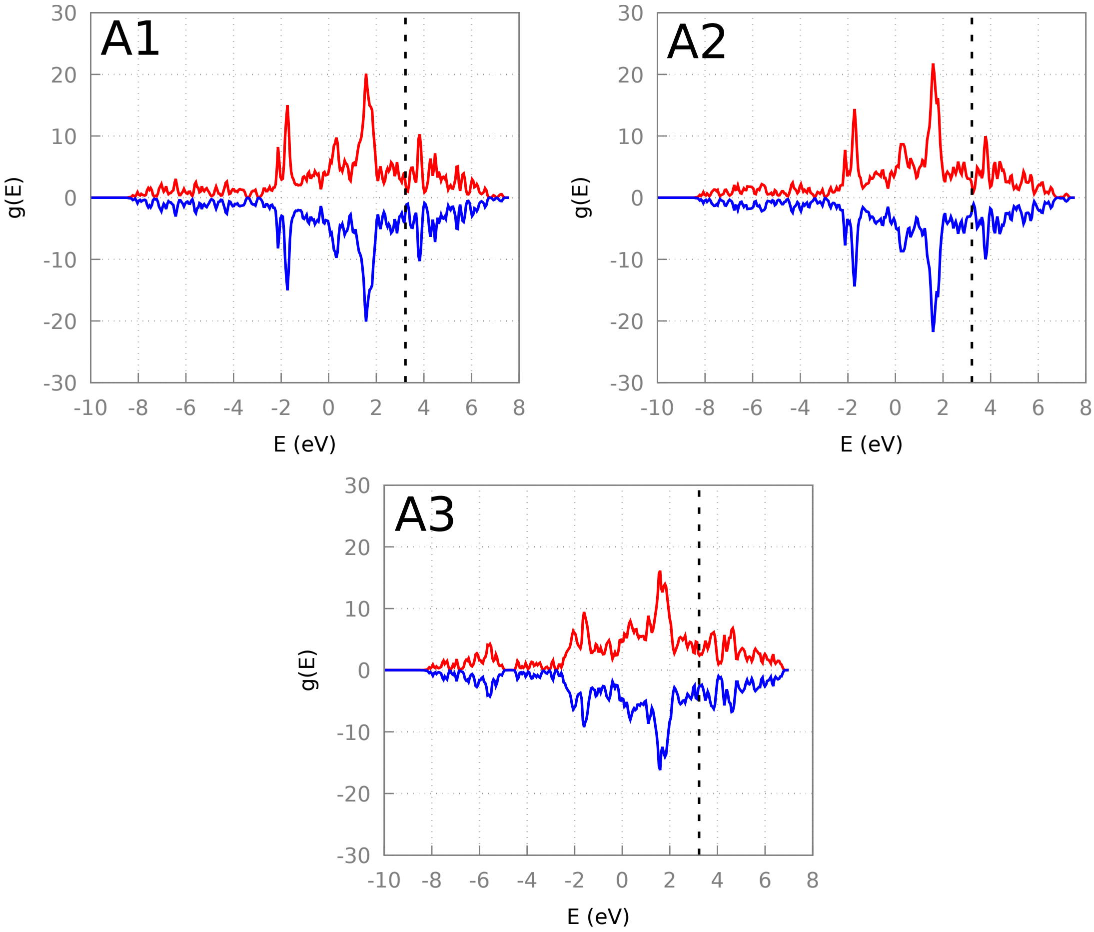

Figure 8.12: Density of states of the 4 low enthalpy AFM structures. None of these structures are half metallic.

8.5 Summary

The Heusler alloy CFAS is a good candidate structure for half metals. When grown on top of n-doped germanium, significant inter-layer mixing occurs. When annealed, the mixed region stabilises to form an unknown phase at the interface. The GA was used to predict what this phase could be, finding the Half Heusler alloy $\mathrm{CoFe_{0.5}Al_{0.25}Si_{0.25}Ge}$ to be a promising candidate. This candidate phase turns out to be fully metallic in both up and down spins, which is undesirable for device applications. The phase also turns out to be ferrimagnetic. Both structural and magnetic disorder are thermally accessible. Structural disorder does not appear to change the electronic structure significantly, although AFM ordering of the Fe atoms does.

Chapter 9

Conclusions and future work

9.1 Summary and conclusions

The aim of this work was to investigate the feasibility of, and develop the means for using the GA for structure prediction of magnetic materials. In chapter 5, issues with using the GA for predicting magnetic structures were raised and addressed by defining new operations such as crossover, mutation and permutation operations. In addition, the idea of structural fingerprinting using the crystallographic structure factor was extended, based on ideas from magnetic neutron scattering theory, and it was demonstrated that the new fingerprint could identify both similar and distinct crystal and magnetic structures.

To test the magnetic extensions to the GA, a new pair potential which included some magnetic-like effects was defined. While this potential does not accurately model any particular real life material, it provided a computationally efficient way of exploring magnetic materials without the cost of fully quantum calculations. Two parameterisations were given for this potential. The first parameterised off DFT simulations of iron dimers in both aligned and anti-aligned configurations. Since the iron dimers preferred to be aligned, it was expected that this parameterisation would yield a FM structure. This was observed as the lowest energy structure found by the GA was FM FCC, and there was a clear trend in the energies of all the structures searched towards FM structures. The second parameterisation mirrored the first, but with a reversed sign on the magnetic effects. This was expected to result in AFM structures, since pairs of aligned spins would raise the energy of the system. Indeed, the lowest energy structure discovered by the GA was an AFM BCC structure. In addition, there was a clear trend amongst the other members towards AFM alignment.

Next, the GA was applied to the problem of mixing at the NiO(111)/MgO(111) interface. In this case, the process behind experimentally observed mixing was unknown. The GA was allowed to evolve the Ni/Mg sites near a clean interface region, resulting in a number of lower energy structures. This suggested that the mixing was energetically driven rather than kinetically driven. In addition, the lowest energy structure found remained AFM ordered with the bulk layers, and displayed the gradual increase in Ni density observed in experimental interfaces. There were however both structural and magnetic disorders found within thermal energy of the lowest structure.

Finally, the interface between the Heusler alloy CFAS and n-doped germanium was investigated. Experimentally, the two materials showed significant mixing and, when annealed, a new phase formed at the interface. The structure of this phase was investigated with the GA, and a half-Heusler structure was predicted. This structure, and low energy disorders around it, show undesirable electronic properties for the devices in which they would want to be used.

To conclude, work set out to develop a GA for structure prediction capable of optimising both magnetic and crystal structures. The GA was able to achieve this, first demonstrating that it could correctly predict the magnetic structure of a simplified pair potential, and then being able to shed light on interface effects with magnetic materials, where it was not known before what was going on.

9.2 Future work

9.2.1 Multiobjective optimisation

One of the more interesting fields in evolutionary computation is that of multiobjective optimisation. By defining multiple objective functions, for example minimising energy *and* maximising some desirable property of a material such as half metallicity, a set of solutions can be found, whereby no structure can be improved in one objective without making it worse in another. This set is called a Pareto set. A number of algorithms exist to do this using the framework of genetic algorithms. Two popular multiobjective evolutionary algorithms are the NSGA-II algorithm [106] and the SPEA algorithm [107]. Some preliminary work has been done in this area, however it is not mature enough for inclusion in this thesis.

9.2.2 Surrogate models for selection

In the broader field of GAs, expensive optimisation problems sometimes use surrogate models to improve performance [108]. These are cheap models trained to the solutions the GA has already seen, and are used to approximate the landscape in order to guide which solutions are chosen in the future for full evaluation. This can be done using machine learning techniques such as neural networks or Gaussian processes [109], but could also involve fitting a known approximate function to previously seen solutions.

It is important to note that, unlike the machine learning methods used for molecular dynamics and structure optimisation like those mentioned in section 2.7, surrogate models do not need accurate forces for all atoms, but only an approximate energy for the entire system. This can then be used to bias the population at the selection stage away from structural motifs which produce high energy structures, such as unphysical bond lengths or angles.

Appendix A

Derivation of the Structure factor fingerprint

In the original paper by Abraham and Probert [84], the equation for the structure factor fingerprint, based on the x-ray scattering intensity Λ, was incorrectly reported as:

$$\Lambda(k) = N \sum_{I=1}^{N} \rho_I^2 + 2 \sum_{I=1}^{N} \sum_{J>I}^{N} \rho_I^2 \rho_J^2 \times J_0 \left(\sqrt{3} \pi k |\mathbf{R}_{IJ}| \right) \tag{A.1}$$

Where ρ_I is the charge density of the nucleus of atom I, defined to be the charge of the nucleus Z_I at the position of the nucleus $\mathbf{R}_I$, and zero everywhere else. J_0 is a Bessel function. The correct form of this equation is derived below.

The structure factor for distinguishable atoms at positions R_I in a crystal[110] is:

$$F(\mathbf{k}) = \sum_I \rho_I \times \mathrm{e}^{i\mathbf{k}\cdot\mathbf{R_I}}, \tag{A.2}$$

Where $\mathbf{k} = \mathbf{s} - \mathbf{s}_0$ is the change in wave vector between incident and scattered waves through the crystal.

From this, the scattering intensity $\Lambda(\mathbf{k})$ can be calculated:

$$
\begin{aligned}
\Lambda(\mathbf{k}) &= |F(\mathbf{k})|^2 \\
&= \sum_I \sum_J \rho_I \rho_J \times \mathrm{e}^{i\mathbf{k}\cdot(\mathbf{R}_I - \mathbf{R}_J)} \\
&= \sum_I \sum_J \rho_I \rho_J \times \mathrm{e}^{ik|\mathbf{R}_{IJ}|\cos\theta},
\end{aligned} \tag{A.3}
$$

Here we have defined $\mathbf{R}_{IJ} = \mathbf{R}_I - \mathbf{R}_J$ to be the vector between atoms I and J and θ to be the angle between $\mathbf{k}$ and $\mathbf{R}_{IJ}$.

The probability of $\mathbf{R}_{IJ}$ being angle θ with $\mathbf{k}$ is $\frac{\sin\theta}{2}\mathrm{d}\theta$. Therefore, the spherically averaged scattering intensity can be calculated by integrating over all θ:

$$
\begin{aligned}
\Lambda(k) &= \sum_I \sum_J \rho_I \rho_J \times \int_0^\pi \exp(ik|\mathbf{R}_{IJ}|\cos\theta)\frac{\sin\theta}{2}\mathrm{d}\theta \\
&= \sum_I \sum_J \rho_I \rho_J \times j_0(k|\mathbf{R}_{IJ}|)
\end{aligned} \tag{A.4}
$$

Separating out the diagonal term and making use of the symmetry in I and J, we get the final form:

$$
\Lambda(k) = \sum_I \rho_I^2 + 2\sum_I \sum_{J>I} \rho_I \rho_J \times j_0(k|\mathbf{R}_{IJ}|) \tag{A.5}
$$

Bibliography

[1] David J. Wales and Jonathan P. K. Doye. Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms. *The Journal of Physical Chemistry A*, 101(28):5111–5116, 1997.

[2] J. Kübler. Magnetic moments of ferromagnetic and antiferromagnetic bcc and fcc iron. *Physics Letters A*, 81(1):81 – 83, 1981.

[3] VK Lazarov, D Gilks, PJ Hsanip, K Matsuuxaki, H Hosono, and T Susaki. Polar oxide nio(111)/mgo(111) interface study by aberration corrected scanning transmission electron microscopy. *Unpublished*.

[4] B. Kuerbanjiang, Y. Fujita, M. Yamada, S. Yamada, A. M. Sanchez, P. J. Hasnip, A. Ghasemi, D. Kepaptsoglou, G. Bell, K. Sawano, K. Hamaya, and V. K. Lazarov. Correlation between spin transport signal and heusler/semiconductor interface quality in lateral spin-valve devices. *Phys. Rev. B*, 98:115304, Sep 2018.

[5] Jos Thijssen. *Computational Physics*. Cambridge University Press, 2 edition, 2007.

[6] J E Lennard-Jones. Cohesion. *Proceedings of the Physical Society*, 43(5):461, 1931.

[7] Murray S. Daw and M. I. Baskes. Semiempirical, quantum mechanical calculation of hydrogen embrittlement in metals. *Phys. Rev. Lett.*, 50:1285–1288, Apr 1983.

[8] Murray S. Daw, Stephen M. Foiles, and Michael I. Baskes. The embedded-atom method: a review of theory and applications. *Materials Science Reports*, 9(7):251 – 310, 1993.

[9] Benedict E. K. Snodin, Ferdinando Randisi, Majid Mosayebi, Petr Šulc, John S. Schreck, Flavio Romano, Thomas E. Ouldridge, Roman Tsukanov, Eyal Nir, Ard A. Louis, and Jonathan P. K. Doye. Introducing improved structural properties and salt dependence into a coarse-grained model of dna. *The Journal of Chemical Physics*, 142(23):234901, 2015.

[10] D. Dias da Costa, J. Alfaiate, L.J. Sluys, and E. Jùlio. A discrete strong discontinuity approach. *Engineering Fracture Mechanics*, 76(9):1176 – 1201, 2009.

[11] Vlado K. Lazarov, Zhuhua Cai, Kenta Yoshida, K. Honglian L. Zhang, M. Weinert, Katherine S. Ziemer, and Philip J. Hasnip. Dynamically stabilized growth of polar oxides: The case of mgo(111). *Phys. Rev. Lett.*, 107:056101, Jul 2011.

[12] E. Du Trémolet de Lacheisserie, University Joseph Fourier, D. Gignoux, and M. Schlenker. *Magnetism.* Number v. 1 in Collection Grenoble Sciences. Springer, 2005.

[13] J.M.D. Coey. Rare-earth magnets. *Endeavour*, 19(4):146 – 151, 1995.

[14] Elisa Alonso, Andrew M. Sherman, Timothy J. Wallington, Mark P. Everson, Frank R. Field, Richard Roth, and Randolph E. Kirchain. Evaluating rare earth element availability: A case with revolutionary demand from clean technologies. *Environmental Science & Technology*, 46(6):3406–3414, Mar 2012.

[15] R. L. Comstock. Review modern magnetic materials in data storage. *Journal of Materials Science: Materials in Electronics*, 13(9):509–523, Sep 2002.

[16] Melvin M. Vopson. Fundamentals of multiferroic materials and their possible applications. *Critical Reviews in Solid State and Materials Sciences*, 40(4):223–250, 2015.

[17] Dieter Weller, Gregory Parker, Oleksandr Mosendz, Andreas Lyberatos, Dmitriy Mitin, Nataliia Y. Safonova, and Manfred Albrecht. Review article: Fept heat assisted magnetic recording media. *Journal of Vacuum Science & Technology B*, 34(6):060801, 2016.

[18] Benjamin Balke, Sabine Wurmehl, Gerhard H Fecher, Claudia Felser, and Jürgen Kübler. Rational design of new materials for spintronics: Co 2 fe z (z =al, ga, si, ge). *Science and Technology of Advanced Materials*, 9(1):014102, 2008.

[19] Xingxing Li and Jinlong Yang. First-principles design of spintronics materials. *National Science Review*, 3(3):365–381, 2016.

[20] Lars Hedin. New method for calculating the one-particle green's function with application to the electron-gas problem. *Phys. Rev.*, 139:A796–A823, Aug 1965.

[21] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, Nov 1965.

[22] Richard M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2004.

[23] M. Born and R. Oppenheimer. Zur quantentheorie der molekeln. *Annalen der Physik*, 389(20):457–484.

[24] Martin Hilbert and Priscila López. The world's technological capacity to store, communicate, and compute information. *Science*, 332(6025):60–65, 2011.

[25] D. R. Hartree. The wave mechanics of an atom with a non-coulomb central field. parts i, ii and iii. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(1):89–110, 1928.

[26] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, Nov 1964.

[27] D. M. Ceperley and B. J. Alder. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45:566–569, Aug 1980.

[28] J. P. Perdew and Alex Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23:5048–5079, May 1981.

[29] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–3868, Oct 1996.

[30] Victor G. Ruiz, Wei Liu, Egbert Zojer, Matthias Scheffler, and Alexandre Tkatchenko. Density-functional theory with screened van der waals interactions for the modeling of hybrid inorganic-organic systems. *Phys. Rev. Lett.*, 108:146103, Apr 2012.

[31] Alberto Ambrosetti, Anthony M. Reilly, Robert A. DiStasio, and Alexandre Tkatchenko. Long-range correlation energy calculated from coupled atomic response functions. *The Journal of Chemical Physics*, 140(18):18A508, 2014.

[32] U von Barth and L Hedin. A local exchange-correlation potential for the spin polarized case. i. *Journal of Physics C: Solid State Physics*, 5(13):1629–1642, jul 1972.

[33] Vladimir I. Anisimov, Jan Zaanen, and Ole K. Andersen. Band theory and mott insulators: Hubbard u instead of stoner i. *Phys. Rev. B*, 44:943–954, Jul 1991.

[34] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton. Electron-energy-loss spectra and the structural stability of nickel oxide: An lsda+u study. *Phys. Rev. B*, 57:1505–1509, Jan 1998.

[35] Matteo Cococcioni and Stefano de Gironcoli. Linear response approach to the calculation of the effective interaction parameters in the LDA + U method. *Phys. Rev. B*, 71:035105, Jan 2005.

[36] Felix Bloch. Über die quantenmechanik der elektronen in kristallgittern. *Zeitschrift für Physik*, 52(7):555–600, Jul 1929.

[37] Hendrik J. Monkhorst and James D. Pack. Special points for brillouin-zone integrations. *Phys. Rev. B*, 13:5188–5192, Jun 1976.

[38] Chris-Kriton Skylaris, Peter D. Haynes, Arash A. Mostofi, and Mike C. Payne. Introducing onetep: Linear-scaling density functional simulations on parallel computers. *The Journal of Chemical Physics*, 122(8):084119, 2005.

[39] R. A. Buckingham, J. Corner, and Harrie Stewart Wilson Massey. Tables of second virial and low-pressure joule-thomson coefficients for intermolecular potentials with exponential repulsion. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 189(1016):118–129, 1947.

[40] Stephen L. Mayo, Barry D. Olafson, and William A. Goddard. Dreiding: a generic force field for molecular simulations. *The Journal of Physical Chemistry*, 94(26):8897–8909, 1990.

[41] B. W. H. van Beest, G. J. Kramer, and R. A. van Santen. Force fields for silicas and aluminophosphates based on ab initio calculations. *Phys. Rev. Lett.*, 64:1955–1958, Apr 1990.

[42] Frank H. Stillinger and Thomas A. Weber. Computer simulation of local order in condensed phases of silicon. *Phys. Rev. B*, 31:5262–5271, Apr 1985.

[43] P.C.L Stephenson, M.W Radny, and P.V Smith. A modified stillinger-weber potential for modelling silicon surfaces. *Surface Science*, 366(1):177 – 184, 1996.

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

[45] Albert P. Bartók, Mike C. Payne, Risi Kondor, and Gábor Csányi. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.*, 104:136403, Apr 2010.

[46] Jörg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.*, 98:146401, Apr 2007.

[47] J Behler. Representing potential energy surfaces by high-dimensional neural network potentials. *Journal of Physics: Condensed Matter*, 26(18):183001, 2014.

[48] Ralph J. Harrison and Genrich L. Krasko. Magnetic-state-dependent interatomic potential for iron (abstract). *Journal of Applied Physics*, 67(9):4585–4585, 1990.

[49] S L Dudarev and P M Derlet. A 'magnetic' interatomic potential for molecular dynamics simulations. *Journal of Physics: Condensed Matter*, 17(44):7097, 2005.

[50] Graeme J. Ackland. Two-band second moment model for transition metals and alloys. *Journal of Nuclear Materials*, 351(1):20 – 27, 2006. Proceedings of the Symposium on Microstructural Processes in Irradiated Materials.

[51] Daniele Dragoni, Thomas D Daff, Gábor Csányi, and Nicola Marzari. Achieving DFT accuracy with a machine-learning interatomic potential: Thermomechanics and defects in bcc ferromagnetic iron. *Physical Review Materials*, 2(1):013808, 2018.

[52] R F L Evans, W J Fan, P Chureemart, T A Ostler, M O A Ellis, and R W Chantrell. Atomistic spin model simulations of magnetic nanomaterials. *Journal of Physics: Condensed Matter*, 26(10):103202, 2014.

[53] P.E. Gill, W. Murray, and M.H. Wright. *Practical optimization.* Academic Press, 1981.

[54] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos. Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. *Rev. Mod. Phys.*, 64:1045–1097, Oct 1992.

[55] Jonathan Barzilai and Jonathan M. Borwein. Two-point step size gradient methods. *IMA Journal of Numerical Analysis*, 8(1):141–148, 1988.

[56] C. G. Broyden. The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA Journal of Applied Mathematics*, 6(1):76–90, 1970.

[57] R. Fletcher. A new approach to variable metric algorithms. *The Computer Journal*, 13(3):317–322, 1970.

[58] Donald Goldfarb. A family of variable-metric methods derived by variational means. *Mathematics of Computation*, 24(109):23–23, jan 1970.

[59] D. F. Shanno and P. C. Kettler. Optimal conditioning of quasi-Newton methods. *Mathematics of Computation*, 24(111):657–657, sep 1970.

[60] Bernd G. Pfrommer, Michel Côtè, Steven G. Louie, and Marvin L. Cohen. Relaxation of crystals with the quasi-newton method. *Journal of Computational Physics*, 131(1):233 – 240, 1997.

[61] Michael P. Teter, Michael C. Payne, and Douglas C. Allan. Solution of schrödinger's equation for large systems. *Phys. Rev. B*, 40:12255–12263, Dec 1989.

[62] G. P. Kerker. Efficient iteration scheme for self-consistent pseudopotential calculations. *Phys. Rev. B*, 23:3082–3084, Mar 1981.

[63] M.I.J. Probert. Improved algorithm for geometry optimisation using damped molecular dynamics. *Journal of Computational Physics*, 191(1):130 – 146, 2003.

[64] Erik Bitzek, Pekka Koskinen, Franz Gähler, Michael Moseler, and Peter Gumbsch. Structural relaxation made simple. *Phys. Rev. Lett.*, 97:170201, Oct 2006.

[65] D. H. Wolpert and W. G. Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1):67–82, April 1997.

[66] David H Wolpert. What the no free lunch theorems really mean; how to improve search algorithms. 2012.

[67] R.E. Smallman and A.H.W. Ngan. Chapter 11 - work hardening and annealing. In R.E. Smallman and A.H.W. Ngan, editors, *Modern Physical Metallurgy (Eighth Edition)*, pages 443 – 471. Butterworth-Heinemann, Oxford, eighth edition edition, 2014.

[68] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[69] J. Pannetier, J. Bassas-Alsina, J. Rodriguez-Carvajal, and V. Caignaert. Prediction of crystal structures from crystal chemistry rules by simulated annealing. *Nature*, 346(6282):343–345, jul 1990.

[70] David J. Wales and Harold A. Scheraga. Global optimization of clusters, crystals, and biomolecules. *Science*, 285(5432):1368–1372, 1999.

[71] Stefan Goedecker. Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. *The Journal of Chemical Physics*, 120(21):9911–9917, 2004.

[72] Maximilian Amsler and Stefan Goedecker. Crystal structure prediction using the minima hopping method. *The Journal of chemical physics*, 133(22):224104, 2010.

[73] Chiara Panosetti, Konstantin Krautgasser, Dennis Palagin, Karsten Reuter, and Reinhard J Maurer. Global materials structure search with chemically motivated coordinates. *Nano letters*, 15(12):8044–8048, 2015.

[74] R. Martoňák, A. Laio, and M. Parrinello. Predicting crystal structures: The parrinello-rahman method revisited. *Phys. Rev. Lett.*, 90:075503, Feb 2003.

[75] Scott M. Woodley and Richard Catlow. Crystal structure prediction from first principles. *Nature Materials*, 7(12):937–946, dec 2008.

[76] Chris J Pickard and R J Needs. Ab initio random structure searching. *Journal of Physics: Condensed Matter*, 23(5):053201, 2011.

[77] Scott M. Woodley, Peter D. Battle, Julian D. Gale, and C Richard A. Catlow. The prediction of inorganic crystal structures using a genetic algorithm and energy minimisation. *Phys. Chem. Chem. Phys.*, 1:2535–2542, 1999.

[78] N. L. Abraham and M. I. J. Probert. A periodic genetic algorithm with real-space representation for crystal structure and polymorph prediction. *Phys. Rev. B*, 73:224104, Jun 2006.

[79] Colin W. Glass, Artem R. Oganov, and Nikolaus Hansen. Uspex—evolutionary crystal structure prediction. *Computer Physics Communications*, 175(11):713 – 720, 2006.

[80] David C Lonie and Eva Zurek. Xtalopt: An open-source evolutionary algorithm for crystal structure prediction. *Computer Physics Communications*, 182(2):372–387, 2011.

[81] Nicholas J Radcliffe and Patrick D Surry. Formal memetic algorithms. In *AISB Workshop on Evolutionary Computing*, pages 1–16. Springer, 1994.

[82] Tarek A El-Mihoub, Adrian A Hopgood, Lars Nolle, and Alan Battersby. Hybrid genetic algorithms: A review. *Engineering Letters*, 13(2):124–137, 2006.

[83] Fred Glover. Tabu search—part i. *ORSA Journal on computing*, 1(3):190–206, 1989.

[84] NL Abraham and MIJ Probert. Improved real-space genetic algorithm for crystal structure and polymorph prediction. *Physical Review B*, 77(13):134117, 2008.

[85] Stewart J Clark, Matthew D Segall, Chris J Pickard, Phil J Hasnip, Matt IJ Probert, Keith Refson, and Mike C Payne. First principles methods using castep. *Zeitschrift für Kristallographie-Crystalline Materials*, 220(5/6):567–570, 2005.

[86] Artem R Oganov and Mario Valle. How to quantify energy landscapes of solids. *The Journal of chemical physics*, 130(10):104504, 2009.

[87] Albert P. Bartók, Risi Kondor, and Gábor Csányi. On representing chemical environments. *Phys. Rev. B*, 87:184115, May 2013.

[88] N. L. Abraham and M. I. J. Probert. Erratum: Improved real-space genetic algorithm for crystal structure and polymorph prediction [phys. rev. b 77, 134117 (2008)]. *Phys. Rev. B*, 94:059904, Aug 2016.

[89] J B Pendry. Reliability factors for leed calculations. *Journal of Physics C: Solid State Physics*, 13(5):937, 1980.

[90] B. L. Miller and M. J. Shaw. Genetic algorithms with dynamic niche sharing for multimodal function optimization. In *Proceedings of IEEE International Conference on Evolutionary Computation*, pages 786–791, May 1996.

[91] D.E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Artificial Intelligence. Addison-Wesley Publishing Company, 1989.

[92] Fred L Hirshfeld. Bonded-atom fragments for describing molecular charge densities. *Theoretica chimica acta*, 44(2):129–138, 1977.

[93] GE Bacon. Neutron diffraction monographs on the physics and chemistry of materials, 1975.

[94] S. Blundell. *Magnetism in Condensed Matter*. Oxford Master Series in Condensed Matter Physics 4. OUP Oxford, 2001.

[95] Shinji Yuasa, Taro Nagahama, Akio Fukushima, Yoshishige Suzuki, and Koji Ando. Giant room-temperature magnetoresistance in single-crystal fe/mgo/fe magnetic tunnel junctions. *Nature Materials*, 3:868 EP –, Oct 2004.

[96] Susanne Stemmer and S. James Allen. Two-dimensional electron gases at complex oxide interfaces. *Annual Review of Materials Research*, 44(1):151–171, 2014.

[97] V. K. Lazarov, M. Weinert, S. A. Chambers, and M. Gajdardziska-Josifovska. Atomic and electronic structure of the fe3o4(111)/mgo(111)model polar oxide interface. *Phys. Rev. B*, 72:195401, Nov 2005.

[98] Yu Kumagai, Atsuto Seko, Fumiyasu Oba, and Isao Tanaka. Ground-state search in multicomponent magnetic systems. *Phys. Rev. B*, 85:012401, Jan 2012.

[99] C. Xu, W. S. Oh, and D. W. Goodman. Growth and chemical properties of layered, binary mixed oxides: No on mgo-nio and cao-nio. *The Journal of Physical Chemistry B*, 104(44):10310–10315, 2000.

[100] Phil J Hasnip. Calculations on nio/mgo interfaces. unpublished.

[101] Luiz G. Ferreira, Lara Teles, and Marcelo Marques. Band structure of nio revisited. 10 2009.

[102] Peter J. Webster. Heusler alloys. *Contemporary Physics*, 10(6):559–577, 1969.

[103] R. A. de Groot, F. M. Mueller, P. G. van Engen, and K. H. J. Buschow. New class of materials: Half-metallic ferromagnets. *Phys. Rev. Lett.*, 50:2024–2027, Jun 1983.

[104] P. J. Hasnip, J. H. Smith, and V. K. Lazarov. Ab initio studies of disorder in the full heusler alloy co2fexmn1-xsi. *Journal of Applied Physics*, 113(17):17B106, 2013.

[105] Zlatko Nedelkoski, Philip J Hasnip, Ana M Sanchez, Balati Kuerbanjiang, Edward Higgins, Mikihiko Oogane, Atsufumi Hirohata, Gavin R Bell, and Vlado K Lazarov. The effect of atomic structure on interface spin-polarization of half-metallic spin valves: Co2mnsi/ag epitaxial interfaces. *Applied Physics Letters*, 107(21):212404, 2015.

[106] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and TAMT Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2):182–197, 2002.

[107] E. Zitzler and L. Thiele. Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach. *IEEE Transactions on Evolutionary Computation*, 3(4):257–271, Nov 1999.

[108] Yew S Ong, Prasanth B Nair, and Andrew J Keane. Evolutionary optimization of computationally expensive problems via surrogate modeling. *AIAA journal*, 41(4):687–696, 2003.

[109] Volker L. Deringer, Chris J. Pickard, and Gábor Csányi. Data-driven learning of total and local energies in elemental boron. *Phys. Rev. Lett.*, 120:156001, Apr 2018.

[110] N.W. Ashcroft and N.D. Mermin. *Solid State Physics.* Saunders College, Philadelphia, 1976.