Early Stages of Precipitation In Aluminum Alloys by First-Principles and Machine-Learning Atomistic Simulations

THÈSE N^O 8766 (2018)

PRÉSENTÉE LE 14 SEPTEMBRE 2018
À LA FACULTÉ DES SCIENCES ET TECHNIQUES DE L'INGÉNIEUR
LABORATOIRE DE SCIENCE COMPUTATIONNELLE ET MODÉLISATION
PROGRAMME DOCTORAL EN SCIENCE ET GÉNIE DES MATÉRIAUX

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Daniele GIOFRÉ

acceptée sur proposition du jury:

Prof. P. Muralt, président du jury
Prof. M. Ceriotti, directeur de thèse
Prof. G. C. Sosso, rapporteur
Prof. G. Tribello, rapporteur
Prof. W. Curtin, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2018

Dedicated to *Silvia*

I suppose the best way to find out
where you've come from
is to find out where you're going,
and then work backwards.

— Doctor Who, "City of Death"

ABSTRACT

Age hardening induced by the formation of (semi)-coherent pre-
cipitate phases is crucial for the processing and final properties
of the widely used Al-6000 alloys. The early stages of precipitation
are particularly important from the fundamental and technological
side, but are still far from being fully understood. In this thesis, this
crucial step in the technology of Al-based alloys is studied by means
of multi-scale simulations that include first-principles atomistic mod-
elling, surrogate models based on statistical learning, as well as ki-
netic Monte Carlo and continuum elasticity models. All with the
purpose of bridging time and length scales. We begin with an anal-
ysis of the energetics of nanometric precipitates of the meta-stable
β'' phases (that are understood to play a crucial role in this system).
This analysis allows us to identify the bulk, elastic strain and inter-
face energies that contribute to the stability of a nucleating cluster.
The results of this analysis show that needle-shape precipitates are
unstable to growth even at the smallest size β'' formula unit and
that there is no energy barrier to growth. This study made it possi-
ble to develop a semi-quantitative classical nucleation theory model,
including also elastic strain energy, that captures the trends in pre-
cipitate energy versus size and composition. This validates the use
of mesoscale models to assess stability and interactions of β'' precipi-
tates. Studies of smaller 3D clusters also show stability relative to the
solid solution state, indicating that the early stages of precipitation
may be diffusion-limited.

Our results thus point toward the need for a systematic study of
the energetics of aggregates in the Guinier-Preston zone regime. In
particular, a study of the interactions between aggregates and vacan-
cies and/or trace elements is required in the alloy in order to un-
derstand and fine-tune the behaviour of Al-6000 alloys in the early
stages of precipitation. To enable full atomistic-level simulations of
the whole precipitation sequence for this important alloy system, two
Neural Network (NN) potentials have been created by representing
just 2-body interactions and by also including 3-body terms. In order
to simplify the construction of NN potentials, we developed an au-
tomatic scheme to determine the most appropriate representation of
the structural features of this ternary alloy.

Training of the NN uses an extensive database of energies and
forces computed using first-principles Density Functional Theory, in-
cluding complex precipitate phases. The Neural Network potentials
accurately reproduce most of the properties of pure Al which are

relevant to the mechanical behaviour as well as the heat of solution, solute-solute and solute-vacancy interaction energies, and formation energies for small solute clusters and precipitates. An accurate model that reproduces all these properties is required for modelling the early stage of precipitation and mechanical strengthening. This success does not only enable future detailed studies of Al-Mg-Si but also highlights the ability of machine learning methods to generate useful potentials in complex alloy systems.

Finally, we used this NN potential to implement a kinetic Monte Carlo scheme to study the formation of pre-precipitation clusters. While quantitative accuracy will probably require further refinement of the training set for the NN, to achieve a more complete description of the interactions between solute atoms and vacancies, we could already observe some of the key mechanisms that determine the ultra-fast formation of aggregates of a handful of solute atoms. This work lays the foundations for a thorough investigation of the behaviour of Al-6000 alloys over time and size scales that are technologically relevant, and demonstrates a combination of atomistic modelling techniques that could be adapted and used to model a large number of similar metallic alloys.

KEYWORDS   Aluminium alloy, precipitation, natural ageing, artificial ageing, first principles simulations, Al-Mg-Si system, density functional theory, neural network potential, deep learning, kinetic Monte Carlo, unsupervised machine learning, long-range interactions, nucleation theory, molecular dynamics

SOMMARIO

L′ indurimento per invecchiamento indotto dalla formazione di fasi di precipitato (semi-)coerenti è cruciale per la lavorazione e le proprietà finali delle ampiamente utilizzate leghe Al-6000. Le prime fasi delle precipitazioni sono particolarmente importanti dal punto di vista fondamentale e tecnologico, ma sono ancora ben lungi dall'essere pienamente comprese. In questa tesi, questo passo cruciale nella tecnologia delle leghe a base di Al è studiato attraverso simulazioni a multiscala che includono la modellazione atomistica da primi principi, modelli surrogati basati sull'apprendimento statistico, nonché modelli elastici del continuum e modelli cinetici Monte Carlo per superare le scale di tempo e dimensione. Iniziamo con un'analisi energetica dei precipitati nanometrici delle fasi meta-stabili β'' (che si ritiene svolgano un ruolo cruciale in questo sistema) identificando le energie di deformazione elastica, di interfaccia e di formazione che contribuiscono alla stabilità di un agglomerato di nucleazione.

I risultati mostrano che i precipitati β'' a forma di ago sono precari per la crescita anche nella più piccola dimensione, un'unità di formula, cioè non vi è nessuna barriera energetica alla crescita. Le piccole differenze tra le diverse composizioni indicano la necessità di studiare una possibile ricostruzione dell'interfaccia precipitato/matrice. Un approccio classico semi-quantitativo basato sulla teoria della nucleazione, che include l'energia di deformazione elastica, riesce a catturare le tendenze dell'energia di precipitazione in funzione delle dimensioni e della composizione. Questo convalida l'uso di modelli in scala mesoscopica per valutare la stabilità e le interazioni dei precipitati β''. Gli studi sui più piccoli agglomerati 3D mostrano anche maggiore stabilità rispetto allo stato della soluzione solida, indicando che le fasi iniziali delle precipitazioni possono essere limitate alla diffusione.

I nostri risultati indicano quindi la necessità di uno studio sistematico dell'energia degli aggregati nel regime delle zona di Guinier-Preston, e delle interazioni tra questi aggregati e le vacanze e/o oligoelementi presenti nella lega per capire e perfezionare il comportamento delle Al-6000 leghe nelle prime fasi di precipitazione. Per consentire una completa simulazione a livello atomistico dell'intera sequenza di precipitazione di questo importante sistema ternario, sono stati creati due potenziali con un'infrastuttura basata su una rete neurale, rappresentando solo interazioni a 2 corpi e includendo anche le interazioni a 3 corpi. Per quest'ultimo, abbiamo sviluppato uno schema automatico per determinare la rappresentazione più ap-

propriata delle caratteristiche strutturali di questa lega ternaria. L'addestramento della rete neurale utilizza un ampio catalogo di energie e forze calcolate utilizzando la teoria a principi primi del funzionale della densita comprendendo le fasi complesse del precipitato. I potenziali ottenuti tramite rete neurale riproducono accuratamente la maggior parte delle proprietà del puro Al che sono rilevanti per il comportamento meccanico così come il calore di soluzione, le energie di interazione soluto-soluto e soluto-vacanza e le energie di formazione di piccoli gruppi e precipitati solitari, che sono necessari per modellare la fase iniziale di precipitazione e d'indurimento meccanico. Questo successo non solo consente futuri studi piú dettagliati sul sistema Al-Mg-Si, ma mette anche in evidenza la capacità dei metodi di apprendimento automatico per generare potenziali utili in sistemi di leghe complesse.

Infine, abbiamo usato questo potenziale basato su reti nuerali per implementare uno schema Monte Carlo cinetico per studiare la formazione di aggregati di pre-precipitazione. Mentre la precisione quantitativa probabilmente richiederà un ulteriore affinamento del catalogo di apprendimento, per ottenere una descrizione più completa delle interazioni tra atomi di soluto e vacanze, abbiamo giá potuto osservare alcuni dei meccanismi chiave che determinano l'ultra-veloce formazione di aggregati per un gruppetto di atomi solubili. Questo lavoro getta le basi per un'indagine approfondita del comportamento delle leghe Al-6000 nel tempo e su una scala dimensionale rilevanti dal punto di vista tecnologico, e dimostra una combinazione di tecniche di modellizzazione atomistica che potrebbero essere adattate per un gran numero di simili leghe metalliche.

PAROLE CHIAVI   Leghe di alluminio, precipitazione, invecchiamento naturale, invecchiamento artificiale, simulazioni da principi primi, sistema Al-Mg-Si, teoria del funzionale della densità, potenziale con rete neurale, apprendimento profondo, Monte Carlo cinetico, apprendimento macchina non supervisionato, interazioni a lungo raggio, teoria della nucleazione, dinamica molecolare.

CONTENTS

ACRONYMS

2B   2-Body, page 78

3B   3-Body, page 78

AI-KMC *ab-initio* Kinetic Monte Carlo, page 89

AIMD   *ab-initio* Molecular Dynamics, page 8

AMD   Accelerated Molecular Dynamics, page 33

BCP   Bond Critical Point, page 81

BM   Bond Midpoint, page 82

BO   Born-Oppenheimer , page 10

BZ   Brillouin Zone, page 17

CC   Coupled Cluster, page 9

CG   Conjugate Gradient , page 26

CI   Configuration Interaction , page 9

CI-NEB Climbing Image Nudged Elastic Band, page 34

CP   Car-Parinello, page 10

CPF   Conjugate Peak Refinement, page 34

CPU   Central Processing Unit, page 7

DB   DataBase, page 55

DFT   Density Functional Theory, page 7

DLM   Deep-Learning Method, page 21

EAM   Embedded Atom Method, page 53

EH   Ergodic Hypothesis, page 9

FC-NN   Fully Connected Neural Network, page 22

FCC   Face-Centered Cubic, page 90

FF   Force Field, page 9

FPS   Farthest-Point Sampling, page 73

FU   Formula Unit , page 4

GAP   Gaussian Approximation Potential, page 54

GD   Gradient Descent, page 26

GEA   Gradient Expansion Approximation , page 16

GGA   Generalized Gradient Approximation, page 9

GS   Ground State , page 14

GSF   Generalized Stacking Fault, page 56

HF   Hartree Fock, page 9

HK   Hohenberg and Kohn, page 13

KF   Kalman Filter, page 26

KMC   Kinetic Monte Carlo , page 4

KS   Kohn and Sham , page 13

L-BFGS Limited memory Broyden Fletcher Goldfarb Shanno, page 90

LAMMPS Large-scale Atomic/Molecular Massively Parallel Simulator, page 55

LDA   Local Density Approximation, page 9

LEP   Linear Energy Path, page 36

LMS   Local Minimum States, page 35

LSDA   Local Spin Density Approximation , page 15

MD   Molecular Dynamics, page 8

MEAM   Modified Embedded Atom Method, page 53

MEP   Minimal Energy Path, page 36

ML   Machine-Learning, page 8

MP   Møller-Plesset, page 9

N-N   Nearest-Neighbours, page 109

NEB   Nudged Elastic Band, page 34

NN   Neural Network , page 4

NPA   Natural Population Analysis, page 81

PBC   Periodic Boundary Conditions, page 17

PBE   Perdew-Burke-Ernzerhof , page 16

PES   Potential Energy Surface, page 9

QE   Quantum Espresso, page 17

ReLu   Rectified linear units, page 25

RL   Reinforcement Learning, page 20

RMS   Root Mean Square Error, page 57

SF   Symmetry Functions, page 30

SGD   Stochastic Gradient Descent , page 26

SL   Supervised Learning, page 20

SSL   Semi-Supervised Learning, page 20

SSS   Supersaturated Solid Solution, page 3

TST   Transition State Theory, page 8

USL   UnSupervised Learning, page 20

USPP   UltraSoft PseudoPotentials, page 18

VP   Voronoi Population, page 81

WC   Wulff Construction, page 108

ALUMINIUM ALLOYS TECHNOLOGY

Contents

1.1 OVERVIEW AND APPLICATIONS

Aluminium represents the second most extruded metal in the world, exceeded only by iron, despite the fact that its production became common only in the late 19th century owing to fundamental breakthroughs in the technologies for its extraction. Aluminium is essential because of its light weight, its weldability, and because of exceptional increase in strength that can be obtained by precipitation hardening. These factors combined with a comparatively low price make it commercially very attractive and widely used. As a pure metal, Aluminium is soft and has little strength or resistance to plastic deformation. However, as already mentioned, it can provide the strength of steel at only half of the weight, if alloyed with small amounts of other elements and with appropriate pre-aging treatments.[1–3] In fact, because of the mechanical properties of these alloys can be modulated by heat treatment, these alloys can be easily shaped into a given form and then rapidly strengthened. Aluminium alloys have many applications that include fuel-efficient transportation (e.g., it is increasingly used in automotive and aerospace industries for critical structure applications due to their excellent castability and corrosion resistance and in environmental protection due to their good strength/weight ratio), building construction and food packaging.

1.2 AL-MG-SI ALLOYS

A widely used family of Al alloys is the 6000 series. Although it is not among the strongest aluminium alloys, it constitutes a high share of the aluminium products in the world. Since 1989, about 90% of the tonnage extruded in Western Europe, was $\langle AlMgSi \rangle$ alloys.[4, 5] The 6xxx Al series is characterized by including Si and Mg atoms, in the range of 0.5–1 wt%, usually with a Si/Mg ratio larger than one. The deliberate addition of these solute atoms induces an age hardening

through nanometer-sized precipitates (e.g stable phases, metastable phases, or Guinier-Preston zones), which act as obstacles to dislocation movement in the bulk matrix, and which thus strengthen the aluminium by a factor of 2 or more.

Such alloys have an excess of silicon relative to stable phases (Mg_2Si). In fact, studies in the 90s by a three-dimensional atom probe (3DAP) and transmission electron microscopy (TEM) revealed that the atomic ratio in the co-clusters, Guinier-Preston (GP) zones and β″ is close to that of the alloy composition, namely close to 1 rather than 2.[1, 2, 6, 7] In Fig. 1, we can appreciate the conformation of these precipitates seen through high-resolution TEM analysis.
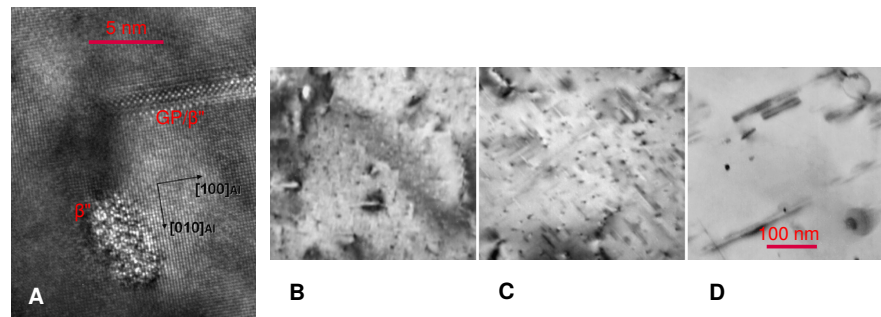

Figure 1: High-resolution TEM observation of (needle) precipitates in 6061-Al alloy after artificial ageing. **A:** Lattice image at high magnification shows a clear semi-coherent β″ phase and a not very clear and almost coherent phase. **B, C, and D:** Comparison of three precipitation states: no further treatment (Fig. B), after heating up to 300 °C at 15 Ks^{-1} (Fig. C), and after heating up to 400 °C at 15 Ks^{-1} (Fig. D). We can note in Fig. B as numerous β″ precipitates are present in the shape of thin needles. These are even smaller after heating up to 300 °C up to evolve into post-β″ phase with larger sizes after heating up to 400 °C. Reproduced from Ref. [8].

1.2.1 *Precipitation sequence*

In commercial 6000-series Al alloys precipitation starts at room temperature shortly after quenching. This is a "natural aging" effect that is usually undesirable. "Artificial aging" at elevated temperature is then used to achieve the desired precipitate type(s) and sizes. At the present time, the accepted precipitation sequence in the Al–Mg–Si alloy system[9] can be schematized to:

SSS → Mg- and Si- atomic clusters → GP-zones → β″ → β′ , U1, U2, B′ → β (stable),

where SSS refers to the initial Supersaturated Solid Solution (SSS). Some authors call the GP zones as GP-1 zones and the β'' phases as GP-2 zones [10]. Recently, one possible structure for the GP-zone and for β'' phase was proposed by high-resolution electron microscopy and quantitative electron diffraction. [4, 7]. These studies showed that such GP-zones have a unit cell, which is only slightly different from the one of the β'' phase, fully coherent with the matrix and with atomic positions very close to those of the Al-matrix. The study of β'' phase showed that its ideal composition has a Mg/Si ratio close to 1 proposing Mg_5Si_6 as a possible form. It has a needle shaped habit aligned along $\langle 100 \rangle$ Al that produces streaks parallel to it. [8] Subsequent studies suggested two other forms for the β'' phase, where Mg/Si ratio is still around 1 (e.g $Mg_5Al_2Si_4$ and $Mg_4Al_3Si_4$)[11], proving that the most stable form is $Mg_5Al_2Si_4$ and assuming such composition as that of β''-phase. In the transformation sequence, the next phases after the GP-zones and the β'' phase are the β', U1, U2, and B' phase. These have hexagonal rods structure and a lower Mg/Si ratio than the equilibrium β phase. [12] Moreover, their structure is related to the ones of previous stages through a similar arrangement of silicon atoms.[6, 13–15]
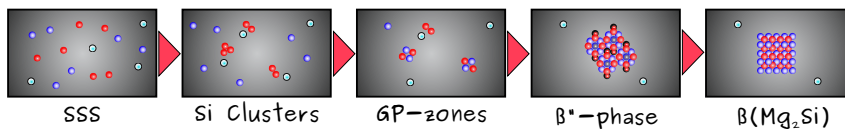

Figure 2: Sketch of precipitation sequence in the Al-6xxx alloys.

These other phases are still important in the process of precipitation, in the commercial Al alloys 6xxx-series the most effective hardening phase under aged conditions is β'' (or a combination of fully coherent GP-zones and semi-coherent β'').[16] Indeed, such β'' phase existing as nano-needles at high number densities provides optimal material strength by hindering the movement of dislocations.

These finite-size nano-needles have sizes that range in size from 200 to 1000 Å in length, and 60 Å in diameter.[4, 17] So, a model that allows us to understand the nucleation process of the β'' phase, based on a prediction for small clusters (e.g. few formula units) and that is extensible to clusters of these sizes, is of fundamental importance.

Challenges

While considerable progress has been made in understanding the structure of the β'' phase and the behavior of the SSSS [18], little is known on the early stages of the aging mechanism, and in particular on the thermodynamics of the initial clustering of solutes to form the precipitate [1, 2, 6, 7]. Little quantitative data has been, indeed, published to explain the details of changes in hardness versus annealing time. The dependence of hardness and nanostructure on the storage time at room temperature (RT), where the GP-zone and the β'' phase

are stable, is still an open problem.[7, 19] Because of the different time scales of the phenomena involved and because of the fundamental role of the environment variables, the nucleation mechanisms that direct the various steps are still difficult to characterize, and a considerable amount of confusion remains concerning the "clustering" processes. Such knowledge is crucial to gain better control over the balance between natural and artificial aging. The aim of this thesis work was to investigate the nature of the nucleation mechanisms of GP-zones and β'' phase by an atomistic study of the clustering barriers (i.e. due to the growth of a semi-coherent surface), and the associated thermodynamic forces.

Because of the complexity of the entire precipitation sequence, that involves both the computational cost and the treatment of the different time scales, the need for more than one computational method to achieve a meaningful description is unavoidable. For this reason we have used different frameworks to understand the nucleation mechanisms, as explained in Chap. 2. In particular, our approach to the problem has been to subdivide the entire clustering sequence into two sub-sequences, where the subdivision element is one formula unit (FU) of β'' phase. In this way we have been able to study the clustering post-β'' phase by a classical semi-quantitative nucleation theory approach, as explained in Chap. 3, and pre-β'' phase by a semi-*ab-initio* Kinetic Monte Carlo (KMC) driven by a Neural Network Potential (NN Potential) , as explained in the Chap. 5.
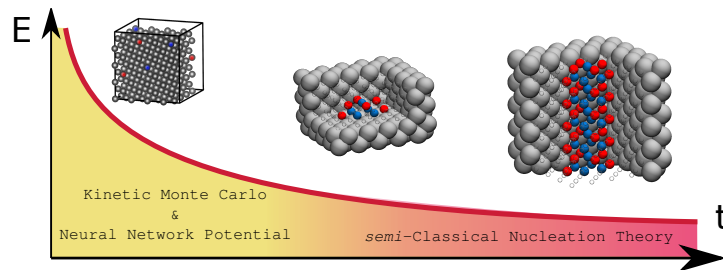

Figure 3: How to handle the entire precipitation sequence.

OUTLINE

The rest of the thesis is organized as follows. After a brief introduction about *ab-initio* atomistic simulations (e.g molecular dynamics and kinetic Monte Carlo) and machine-learning techniques in Chapter 2, Chapter 3 will be discussed, which contains a detailed analysis of the clustering post-β'' phase. It will expose the important interplay among composition-dependent bulk, interface, and elastic strain energies in determining nanoscale precipitate stability and growth, and the limitations of such model (adapted from ref. [20]). In Chapter 4

we discuss how it is possible to use ML to overcome the challenges proposed by this system and how the achievement of a NN potential based on these concepts could be the new frontier for the treatment of this ternary system and for other (even non-)aluminium-kind alloys (adapted from refs. [21, 22]). In Chapter 5, we adopt the NN potential obtained to drive a KMC dynamics. This analysis can give us preliminary and promising results on the first sub-sequence of precipitation. Finally, in Chapter 6 we draw our conclusions.

METHOD

Contents

Deductive scientific reasoning underlies experiments, theory, and, with the advent of the computer, simulations. Simulations and modelling arise as a way of overcoming the intrinsic limitations of mathematical calculations, associated with the complexity of natural phenomena, and they make room for another scientific discipline based on numerical calculations.

Numerical models have been able to overcome the limits of over simplified/approximate analytical theories. In the last decades, in materials research as in other fields, simulations have seen a boost of computational resources (e.g intrinsic CPU power) and algorithms performance. That has allowed us to adopt (increasingly) sophisticated and complete theories in the description of physical phenomena.

The emergence of CPU demanding theories based purely on first principles applied to a wide branch of complex systems and extensive problems, such as the resolution of Schrödinger equation for a many-body system, was in turn one of main driver of this process. Obviously, this boost in computational science was not sufficient to enable solving this equation in its most explicit form. As a consequence, physical deductions and approximations that simplify the treatment of the problem are still highly employed. The most famous, in the atomistic study of metals and alloys, is certainly the one proposed by Kohn-Sham and it is based on the density of the ground state, the so-called Density Functional Theory (DFT). In the Sec. 2.1.1, it will be presented with more attention.

As approximated method to solve Schrödinger's equation, DFT has proved to be remarkably accurate in describing structural and electronic properties of molecular and condensed matter systems, despite its computational simplicity. For these reasons, it is also used as a tool to predict not only static but even dynamical properties. It interfaces well with two big approaches based on different principles according to statistical mechanics: Molecular Dynamics (MD) (also called *ab-initio* Molecular Dynamics (AIMD)) and kinetic Monte Carlo. They are treated respectively in the Secs. 2.1 and 2.3.

MD simulations solve the equations of motion based on the Hamiltonian among atoms in an initial configuration in order to find out the next configuration. [23] Vice versa, the kinetic- Monte Carlo method is a simple and straightforward technique, based on stochastic approaches that relies on Transition State Theory (TST) [24]. Unlike the simple Monte Carlo, it is applicable to evaluate even out-of-equilibrium systems and it is possible to extrapolate dynamic properties.

An issue that stands out for both approaches is the amount of configurations generated. The large amount of easily produced information is hard to process manually, which triggered widespread adoption of Machine Learning to process the outcome of simulations to extract essential information. The branch of computer science that deals with this procedure is commonly called Machine-Learning (ML).

In sec. 2.2, we will give a more detailed explanation on how this new discipline can be applied to the study of materials, and in Chap. 4 we will show the particular case of 6xxx Al alloys.

2.1   MOLECULAR DYNAMICS

Molecular dynamics (MD) computes the movement of atoms by integrating the evolution of positions, velocities, and orientations with respect to time. It describes the atomistic behaviour of matter as a classical system composed of N interacting particles, thus MD given a Hamiltonian H, which is in most cases time-independent, produces a series of configurations based on the initial configuration and velocities. Their evolution over time is performed by solving Hamilton's equations:

$$
\begin{aligned}
\dot{\mathbf{r}}_i &= \frac{\partial H(\mathbf{p}, \mathbf{r})}{\partial \mathbf{p}_i} = \frac{\mathbf{p}_i}{m_i} \\
\dot{\mathbf{p}}_i &= -\frac{\partial H(\mathbf{p}, \mathbf{r})}{\partial \mathbf{r}_i} = -\frac{\partial V(\mathbf{r})}{\partial \mathbf{r}_i} = \mathbf{F}_i \,,
\end{aligned}
\tag{1}
$$

where $\mathbf{r}_i$, $\mathbf{p}_i$ and $\mathbf{F}_i$ are respectively the position, momentum and force associated with the i-th particle, V is the interatomic potential and H is the classical N-particle Hamiltonian.

The dynamics thus produced in turn is studied through a thermodynamic description. A thermodynamic description is characterized by a set of parameters, the so-called thermodynamic observables, which constitute the macroscopic observables of the system. In case we wish to calculate a macroscopic observable $\mathscr{G}_{obs}$ of our system, we need to properly conceive an average of the motion (position or velocity) of the particles, which is an usual way of trying to extract a significant and unique value from an infinite number of irregularly moving particles. Now, the question is how can MD meets the needs of the real-world experiments, which normally occur in a finite time τ. In a mathematical formalism, it translates that the averaged features observed must be evaluated during that interval as follows:

$$\mathscr{G}_{obs} = \bar{A}(\tau) = \frac{1}{\tau} \int_0^\tau dt\, A(\mathbf{p}(t), \mathbf{r}(t)), \tag{2}$$

where $A(\mathbf{p}(t), \mathbf{r}(t))$ is a physical observable. The main purpose of the MD is to compare itself with the experiments both to understand and justify, and even to predict them, and on this note the Ergodic Hypothesis (EH) comes to the rescue. If the EH is justified then $\bar{A}(\tau) \to \langle A \rangle_{NVE}$ as $\tau \to M$, where M is a large number and $\langle A \rangle_{NVE}$ is the micro-canonical ensemble average of A.

Depending on the level of detail of the Potential Energy Surface (PES) , the Hamiltonian in Eq. 1, takes on different forms. It is mainly divided into two macro-groups that are accepted in the scientific world under the names of Force Field MD (FF-MD) and *ab-initio* MD (AIMD). The first macro-group includes all those methods where the interactions are defined by several parameters normally tuned from experiments, or more accurate simulations, to reproduce defined systems. In this group, we can include methods like classical FF, polarizable FF, reactive FF, and coarse-grained FF. The major disadvantage of most empirical FFs is the fact that the predictive power of a FF is guaranteed only for those systems and thermodynamic conditions for which they were designed. All the other methods, where the forces are computed on-the-fly by accurate electronic structure calculations and without relying on any adjustable parameter are included in the second group, such as Density Functional Theory (Local Density Approximation (LDA) [25–27] , Generalised Gradient Approximation (GGA) [27–29] , Hybrid Exchange-Correlation functionals [26, 30–32]), many-bodies perturbation theory (Hartree Fock (HF) [33, 34] , Møller-Plesset (MP) [35, 36] , and Coupled Cluster (CC) [37–39]), Configuration Interaction (CI) (Standard, Quadratic, and Full) [40] and Quantum chemistry composite methods[41, 42].

In order to obtain a realistic time evolution of these alloys, it is mandatory to describe properly the interatomic interactions. The theory that best relates between computational cost and accuracy of PES, in this immense theoretical scheme, is the DFT. For these reasons, this

is method that will be adopted in the Chaps. IV, V and VI to sample the PES.

2.1.1   *Ab-initio Molecular Dynamics*

This approach unifies Hamilton's and Schrödinger's equations, allowing to deal with complex simulations without worrying about the development of empirical potential models. There are two types of approaches: Born-Oppenheimer (BO) MD and Car-Parinello (CP) MD [43]. We discuss the former in the following section.

Born Oppenheimer Approximation

In a many-body system, the Schrödinger equation is the one which regulates the behaviour of all the particles involved (nuclei and electrons)

$$\hat{H}\,\Psi(\mathbf{r}_1,\cdots,\mathbf{r}_N,\mathbf{R}_1,\cdots,\mathbf{R}_M,t) = i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r}_1,\cdots,\mathbf{r}_N,\mathbf{R}_1,\cdots,\mathbf{R}_M,t),$$

(3)

in which $\hat{H}$ is the Hamiltonian, composed by a kinetic energy operator and a potential operator. $\Psi(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N,\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_M,t)$ is the wave function which describes the quantum-mechanical system, $\mathbf{r}_i$ are the space coordinates of the *i*-th electron (for a system with N electrons), and $\mathbf{R}_j$ are the space coordinates of the *j*-th nucleus (for a system with M nuclei).
The many-body Hamiltonian describing a system of interacting electrons and nuclei is

$$\begin{aligned}
\hat{H} &= \hat{T}_e + \hat{T}_n + \hat{V}_{ee} + \hat{V}_{nn} + \hat{V}_{en} \\
&= -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 + \\
&\quad +\frac{1}{2}\sum_{i\neq j}\frac{e^2}{|\mathbf{r}_i-\mathbf{r}_j|} + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J e^2}{|\mathbf{R}_I-\mathbf{R}_J|} - \sum_{i,I}\frac{Z_I e^2}{|\mathbf{r}_i-\mathbf{R}_I|},
\end{aligned}$$

(4)

with T_e and T_n the sums of the kinetic energy terms of both electrons (with mass m_e and charge $-e$) and nuclei (with mass M_I and charge $Z_I e$) and V_{ee}, V_{nn}, and V_{en} the sums of the Coulomb interaction terms among two electrons, two nuclei, and one electron and one nucleus. If the Hamiltonian is time-independent, the Schrödinger equation can be solved by the method of the separation of variables and the properties of the interacting system in principle may be determined from the secular equation

$$\hat{H}\,\Psi(\mathbf{r},\mathbf{R}) = E\,\Psi(\mathbf{r},\mathbf{R}),$$

(5)

where E is the total energy of the $(N + M)$-particle system, and $\mathbf{r} = \{\mathbf{r}_i\}$ and $\mathbf{R} = \{\mathbf{R}_i\}$ are the full set of electronic and nuclear coordinates. The solution for the time-dependent function is simply written by adding a phase factor. The expectation values of any operator can be therefore obtained without the knowledge of this phase factor since it cancels out. However, for many-particle systems the solution of Eq. 5 has not yet a direct resolution. A powerful and accepted way to simplify Eq. 5 is to adopt the so-called Born Oppenheimer approximation.

The Born Oppenheimer approximation considers the motion of the electrons and nuclei decoupled. Since the nuclei are much heavier than the electrons their movement is much slower than the electrons one,

$$\frac{t_I}{t_e} \sim \sqrt{\frac{M_I}{m_e}} \approx 10^2 - 10^3, \tag{6}$$

with t_I the typical nuclear motion time-scale and t_e the electronic one. Then the nuclear coordinates are described as external terms which vary slowly. This means that for each instantaneous configuration of the nuclei the electrons are approximately in a stationary state. From a mathematical point of view, it is possible to separate the electronic motion from the ionic one, factorizing the wave function in an ionic part $\Phi(\mathbf{R})$ and an electronic one $\psi(\mathbf{r}, \mathbf{R})$:

$$\Psi(\mathbf{r}, \mathbf{R}) = \Phi(\mathbf{R})\psi(\mathbf{r}, \mathbf{R}). \tag{7}$$

Then $\psi(\mathbf{r}, \mathbf{R})$ solves the following electronic equation:

$$\hat{H}_e \psi(\mathbf{r}, \mathbf{R}) = [\hat{T}_e + \hat{V}_{ee} + \hat{V}_{en}(\mathbf{R})]\psi(\mathbf{r}, \mathbf{R}) = \epsilon(\mathbf{R})\psi(\mathbf{r}, \mathbf{R}), \tag{8}$$

in which $\hat{V}_{en,ext}(\mathbf{R})$ assumes the role of an external potential, since the coordinates of nuclei result as fixed parameters. Eq. 8 can be derived by substituting the wavefunction equation 7 in 5. The details of the formal derivation are described in Refs. [44, 45]. After having solved the electronic problem, the adiabatic motion of the nuclei is governed by the following expression

$$\hat{H}_n \Phi(\mathbf{R}) = [\hat{T}_n + \hat{V}_{nn} + \epsilon(\mathbf{R})]\Phi(\mathbf{R}) = E(\mathbf{R})\Phi(\mathbf{R}). \tag{9}$$

We get an effective Hamiltonian which depends on the position of nuclei only. This method is based on the principles of Mean Field Theory (MTF), in which the effect of the electrons on the nuclear motion is described introducing a mean field potential acting on nuclei $[\hat{V}_{mf} = \hat{V}_{nn} + \epsilon(\mathbf{R})]$. Despite this first simplification Eq. 8 cannot be solved exactly, except in some cases[1], at least with the means available

---

1 We know analytically the exact wavefunction for a single particle in a vacuum or immersed in potentials which show particular symmetries and for the hydrogen atom.

nowadays. For this reason we have to introduce DFT which allows to study the many-body systems, through sensible approximations and rigorous theorems. The main idea of DFT is not to calculate the wave function of the whole system, but to calculate handier quantities which can describe the system properties in the ground state. In the next section we will discuss the main aspects of DFT, which is nowadays the most used first principles approach.

2.1.2 *Density Functional Theory*

DFT was formulated by Hohenberg, Kohn and Sham to obtain the ground state (GS) properties of many-electron systems without making any explicit reference to the wave function.[47, 48] This method has the electron density as fundamental quantity. The advantages of such a theory are clear. Whereas the complicated many-body wave function depends on three spatial variables for each of the N electrons, the electron density only depends on three spatial variables, making the electron density a quantity that is easier to deal with.

Subsequently, in 1965 Kohn and Sham showed that the electron density of a system of interacting particles under the influence of external potential can be reproduced by a system of non-interacting particles under the influence of an effective potential. According to Hohenberg-Kohn theorem this effective potential is also a functional of the electron density. Besides, the Kohn-Sham formulation allows one to write the Schrödinger equation (Eq. 8) as a set of single-particle equations (i.e. non-interacting particles) that are solvable in a self-consistent way.

The effective potential is composed by the external potential, the Hartree potential (which is the mean Coulomb potential generated by all the electrons) and by the so-called exchange-correlation potential (which contains all the remaining small contributions). Moreover, the latter potential is the functional derivative of the exchange-correlation energy with respect to the electron density. Since we do not know the exact form of the exchange-correlation energy and of the equivalent potential we have to use some approximations.

In this section we will discuss the two basic theorems of DFT, the Kohn-Sham equations, the main approximations of exchange-correlation term, and how DFT is implemented in a plane-wave code such as Quantum Espresso[49].

While with numerical software we can calculate the wave-function for systems which have a reduced number of particles, indeed when studying more complex systems we meet the one which Kohn defines an "*Exponential Wall*" [46] which reduces the actual possibilities to systems containing about 10 electrons.

The Hohemberg-Kohn Theorems

Hohenberg and Kohn (HK) showed in 1964 through two theorems as follows:

> *The total energy of a N-electron system is a unique functional of GS electron density $\rho(\mathbf{r})$.*

> *The GS energy can be obtained variationally: the electron density that minimizes the total energy is the exact GS density.*

that the internal potential is a functional of the electron density. As a result every ground-state property, in particular the GS energy, is a functional of the electron density.[47] Indeed, from the first theorem we can then write again the time independent Schrödinger equation (8), as:

$$E[V_{ext}]|\Psi[V_{ext}]\rangle = (\hat{T} + \hat{V}_{ext} + \hat{V}_{ee})|\Psi[V_{ext}]\rangle \equiv (\hat{T} + \hat{V}_{ext} + \hat{V}_{ee})|\Psi[\rho]\rangle, \tag{10}$$

in which $\hat{T}$, $\hat{V}_{ext}$ and $\hat{V}_{ee}$ are the operators corresponding to the kinetic energy, the external potential and to the interaction among electrons, respectively. While we can write the total energy as

$$E[\rho(\mathbf{r})] = V_{ext}[\rho(\mathbf{r})] + T[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r})] = \int \rho(\mathbf{r})V_{ext}(\mathbf{r})d\mathbf{r} + F_{HK}[\rho(\mathbf{r})], \tag{11}$$

where F_{HK} is defined the universal HK functional. Furthermore, from the second theorem we can obtain the GS energy E_0 by the minimization of the functional $E_{V_{ext}}[\rho(\mathbf{r})]$, i.e.,

$$E_0 = \inf_{\rho \in \Pi} E_{V_{ext}}[\rho] = \inf_{\rho \in \Pi} \left[\int \rho(\mathbf{r})V_{ext}(\mathbf{r})d\mathbf{r} + F_{HK}[\rho] \right] \tag{12}$$

with Π the set of the V-representable electron density². Now the necessary step is to determine the density and the external potential. But in order to do this, we have to know the form of the functional $E[\rho(\mathbf{r})]$, which unluckily is unknown, since we do not know the exact form of $F_{HK}[\rho]$. Mostly it is not clear how we can define the kinetic functional for many particles system. This is the reason why Kohn and Sham (KS) reformulated the kinetic term and opened way to practical applications of DFT [50].

---

2 An electron density is called *V-representable* if it is the density of a GS of the Hamiltonian in Eq. 10 with N and $\hat{V}_{ee}$ specified.

The Kohn-Sham Approach

Hohemberg and Kohn (HK) have provided us with an elegant evidence to obtain the GS density and energy through the minimum principle, Eq. 12. However, we do not know practically any applicable forms of $F_{HK}[\rho]$. Kohn and Sham proposed a scheme to be able to find a good expression for $F_{HK}[\rho]$, defining the main terms of that functional and reducing the approximation only to a particular addend which was called exchange-correlation energy $E_{xc}[\rho]$. Their idea was that:

> *The GS density $\rho(\mathbf{r})$ of a system of interacting particles with external potential $V_{ext}(\mathbf{r})$ could be reproduced in a system of noninteracting particles whose $\rho(\mathbf{r})$ is built by the eigenstate of an single particle effective Schrödinger equation through a suitable local potential $V_S(\mathbf{r})$.*

Based on KS postulate, we get a set of three equations, that can be solved in a self-consistent way, after having chosen the form of the functional $E_{xc}[\rho]$ We have an exact closed expression of the kinetic energy function for single electron for a system of non-interacting electrons

$$1. \quad \left(-\frac{\hbar^2}{2m}\nabla^2 + V_S(\mathbf{r})\right)\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}) \tag{13}$$

where the charge density can be represented as a sum over one-electron orbitals $\psi_i(\mathbf{r})$:

$$2. \quad \rho(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2. \tag{14}$$

and the effective potential as follows

$$3. \quad V_S(\mathbf{r}) = V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) + V_{ext}(\mathbf{r}), \tag{15}$$

where $V_{xc} = \delta E_{xc}[\rho]/\rho$ and V_H is Hartree potential. Indeed, the iterative method starts examining a proof density with which we calculate the effective potential $V_S(\mathbf{r})$. Afterwards we resolve Schrödinger equation (Eq. 13), obtaining a series of orbitals $\psi_i(\mathbf{r})$, the so-called KS orbitals, which approximate better the electronic density (Eq. 14) and then the cycle starts again using the new density. This process goes on up to get some values of the electronic density within the fixed threshold. Then we can find the GS quantities of the system with the density obtained in this way.

We want to clarify that the Kohn-Sham wave function constructed from the N lowest KS orbitals is not intended to be an approximation to the true GS wave function and that the KS eigenvalues are not generally related to the excitation energies of the system. However,

the eigenvalue corresponding to the highest occupied KS orbital has a clear physical meaning [51, 52]. Furthermore, it can be shown that the eigenvalues of the other occupied KS orbitals are a good estimate of the true energies of ionization [53]. In the following section we will show the most used approximations for $E_{xc}[\rho]$, in particular we will focus on GGA which is the most accurate in the study of elastic properties in alloys compared to its rival LDA[54].

Functional Approximations for the Exchange-Correlation Energy

In order to apply the theory outlined in the previous sections we need an approximation for the exchange-correlation functional $E_{xc}[\rho]$.

The first approximation has been introduced by Kohn and Sham in 1965, the so-called Local Density Approximation (LDA). In this approximation $E_{xc}[\rho]$ is given by

Local Density Approximation

$$E_{xc}^{LDA}[\rho] = \int d\mathbf{r}\, \epsilon_{xc}(\rho(\mathbf{r}))\rho(\mathbf{r}), \tag{16}$$

in which $\epsilon_{xc}(\rho(\mathbf{r}))$ is the exchange-correlation energy per unit volume of a homogeneous electron gas of density $\rho(\mathbf{r})$. The corresponding exchange-correlation potential is given by

$$V_{xc}^{LDA}[\rho] = \epsilon_{xc}(\rho(\mathbf{r})) + \rho(\mathbf{r})\frac{\partial \epsilon_{xc}(\rho)}{\partial \rho}\bigg|_{\rho=\rho(\mathbf{r})}. \tag{17}$$

From the above mentioned expressions, it is clear that LDA approximates locally the true non homogeneous system with a homogeneous electrons gas. It is a very rough approximation which works well for systems that have an electronic density varying very slowly (i.e. metals). Vice versa that approximation is not suitable for the atoms and the molecules (where the electrons are localized), because of the sudden variations of density.

So far the KS equations have been considered independent of spin, but the DFT can be also extended to calculate GS properties of spin-polarized systems. In this case the electronic density is decomposed into two spin terms

Local Spin Density Approximation

$$\rho(\mathbf{r}) = \rho^{\uparrow}(\mathbf{r}) + \rho^{\downarrow}(\mathbf{r}), \tag{18}$$

and the energy is a functional of both these components, $E = E[\rho^{\uparrow}, \rho^{\downarrow}]$. So in the Local Spin Density Approximation (LSDA) the exchange-correlation functional can be written as

$$E_{xc}^{LSDA}[\rho] = \int d\mathbf{r}\, \epsilon_{xc}(\rho^{\uparrow}(\mathbf{r}), \rho^{\downarrow}(\mathbf{r}))\rho(\mathbf{r}), \tag{19}$$

where ϵ_{xc} is the exchange-correlation energy per unit volume of a homogeneous electron gas at densities $\rho^{\uparrow}(\mathbf{r})$ and $\rho^{\downarrow}(\mathbf{r})$. An important quantity is the local polarization

$$\xi = \frac{\rho^{\uparrow}(\mathbf{r}) - \rho^{\downarrow}(\mathbf{r})}{\rho(\mathbf{r})}. \tag{20}$$

If it is equal to one, all the spins are oriented along the same direction and the system is completely polarized.[55]

A well-known failure of LDA and LSDA is the systematic underestimation of the band gap of semiconductors and insulators. More precisely, LDA and LSDA do not show derivative terms which enable to describe the correct jump[3], caused by the supplement or the removal of an infinitesimal fraction of the integer number of electrons [56].

Gradient Expansion Approximation

The obvious way to go beyond the LDA and the LSDA is to extend the exchange-correlation functional with terms containing gradients of the density. This so-called Gradient Expansion Approximation (GEA) [47, 57] has the following form

$$
\begin{aligned}
E_{xc}^{GEA}[\rho] \;=\; & E_{xc}[\rho]^{LDA} + \int d\mathbf{r}\, f_1(\rho(\mathbf{r}))(\nabla\rho(\mathbf{r}))^2 + \\
& + \int d\mathbf{r}\, f_2(\rho(\mathbf{r}))(\nabla^2\rho(\mathbf{r}))^2 + \cdots,
\end{aligned}
\tag{21}
$$

where the functions $f_i(\rho(\mathbf{r}))$ are uniquely determined by the density response functions of the homogeneous electron gas. However, the GEA is often found to give results that are less accurate than the one obtained with the LSDA, since this type of approximation does not satisfy the charge neutrality and Slater's theorem [58].

Generalized Gradient Approximation

This problem of the GEA has been overcome with the so-called Generalized Gradient Approximation (GGA) [59]. In comparison with LDA, GGA tends to improve total energies, atomization energies, energy barriers and structural energy difference [60]. In GGA, the exchange-correlation functional depends both on the homogeneous electron gas density and on its gradient:

$$
E_{xc}^{GGA}[\rho] = \int d\mathbf{r}\, F(\rho(\mathbf{r}), \nabla\rho(\mathbf{r})),
\tag{22}
$$

where F is a parametrized analytic function. To obtain reasonable results the function F must be chosen with care among the conditions we have the charge neutrality and Slater's theorem.

With the idea to account both the advantages of the two approximations GGA and LSDA, we have used another exchange-correlation functional $E_{xc}[\rho]$, the so-called Perdew-Burke-Ernzerhof (PBE) functional [60], due to good adaptability to metal alloys.[61, 62] Indeed, PBE functional does not have any fitting parameters because it retains the correct features of LSDA, but it combines them with the most energetically important features of nonlocal GGA.

---

3 They do not estimate accurately the electronic density discontinuity.

2.1.3   *Numerical Methods*

The use of DFT in numerical calculation requests further approximations in addition to those regarding the exchange-correlation functional. In carrying out a computational algorithm based on DFT we must apply and know a lot of technical details, which are also peculiar to informatics.

For our aims we will deal with only the main aspects of DFT software used. If a reader is interested in a more detailed description, he can find it in refs. [63–65]. In this dissertation, the electronic properties of systems have been calculated with Quantum Espresso code [49] (and VASP code[66]).

In order to solve the differential equations (like KS equations) it is necessary to introduce appropriate boundary conditions. In QE as so as in VASP and in other DFT codes, the Periodic Boundary Conditions (PBC) are applied. They correspond to the periodical reproduction of a supercell in three dimensions. This is useful when we consider infinite systems (like crystals), but it is also extensible to the management of finite systems. Indeed, we can consider a supercell big enough to separate the objects of two near supercells. The key of the supercell approach is Bloch's theorem, which states:

Periodic Boundary Conditions

> *For a periodic system the electronic wave-function can be written as a product of wavelike part and a cell-periodic part*

$$\psi_{j,\mathbf{k}}(\mathbf{r}) = u_{j,\mathbf{k}}(\mathbf{r})\, e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{23}$$

> *where* j *is a discrete band index and* $\mathbf{k}$ *a reciprocal lattice vector belonging to the first Brillouin zone (BZ) corresponding to the supercell.*

The theorem allows for mapping the KS eigenvalue problem into the reciprocal space, where one can separately obtain a discrete set of eigenstates of the Hamiltonian for each *k*-point. Then the expectation value of a generic one-body operator $\hat{O}$ can be written like this:

k-points Grid

$$\langle \hat{O} \rangle = \frac{1}{\Omega_{BZ}} \int_{BZ} d\mathbf{k}\, O(\mathbf{k}) \approx \sum_{\mathbf{k}\in BZ} \omega_{\mathbf{k}} O(\mathbf{k}) \tag{24}$$

with

$$\sum_{\mathbf{k}\in BZ} \omega_{\mathbf{k}} = 1, \tag{25}$$

where the integral over the first Brillouin Zone (BZ) , with volume Ω_{BZ}, for practical purposes is approximated by a sum over *k*-points with weight factors $\omega_{\mathbf{k}}$. In QE as in VASP, this discrete BZ sampling is based on a so-called Monkhorst-Pack grid [67]. We can observe also

that larger supercells have a smaller BZ. For large enough cells the number of *k*-points required can be reduced considerably. In fact for these supercells the use of the only Γ-point ($\mathbf{k} = 0\hat{i} + 0\hat{j} + 0\hat{k}$) already enables a good description.

Pseudopotential   Most DFT implementations use pseudopotentials to get rid of the treatment of core electrons in order to simplify the calculation. The idea is to replace the actual atomic potential, which describes also core electrons that are chemically inert, with an effective potential (the so-called pseudopotential) which provides the same description for valence electrons. In QE we prefer the use of UltraSoft PseudoPotentials (USPP) decomposed on spherical Bessel-functions[68] because they are widely used and tested for these systems.

Plane-wave Cutoffs   Furthermore, to get some numerical results from the KS equations we must calculate wave-functions on a basis. In QE as in VASP, we use the plane-wave decomposition to define $\psi_j, k(\mathbf{r})$ that guarantees that the Hamiltonian and overlap matrix are sparse so that the variational principle is computationally easy to evaluate. In this way, Eq. 23 becomes

$$\psi_j, k(\mathbf{r}) = \sum_{\mathbf{G}} c_{\mathbf{G},j} e^{i(\mathbf{k}+\mathbf{G}\cdot\mathbf{r})}, \tag{26}$$

where $\mathbf{G}$ are the reciprocal primitive vectors. Considering this last expression, we see that the solution for each $\mathbf{k}$ need to be evaluated over an infinite number of possible values of $\mathbf{G}$.

In order to make computational possible the $\psi_j, k(\mathbf{r})$ definition is introduced a so-called energy cut-off. Since solutions at very high energies are negligible. This energy cut-off is so defined

$$E_{cut}^{WF} = \frac{\hbar^2}{2m} G_{cut}^2. \tag{27}$$

The sum in Eq. 26, then, is finite. Using the same procedure, an energy cut-off is also defined for charge density and potential description with wave-function decomposition. This value for USPP is normally from 8 to 12 times the value of the wave-function energy cut-off.

Smearing   Another important DFT computational parameter for metals is the smearing. This is fundamental for the integration over Fermi-surface at zero temperature because it shows a discontinuity there (e.g we have a step function). So a super-dense grid of *k*-points would be required to reach the right converged calculation, especially in the case of metals (See Fig. 4), making the simulation unacceptable from the computational point of view. A well accepted remedy in the DFT field is the adoption of a function that smears out such discontinuity by transforming the integrand so that it becomes continuous across the Fermi level. This method is equivalent to having a fictitious finite temperature or to introduce a fictitious entropy term in our system.
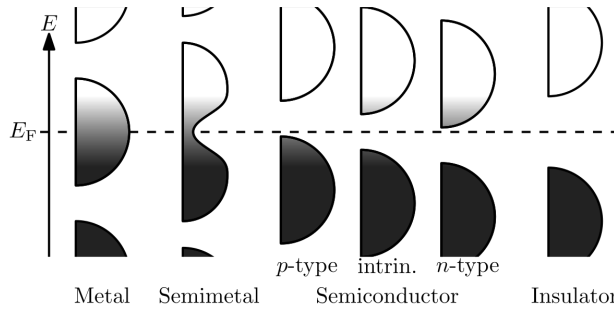
Figure 4: Different occupation of the electronic bands. this behaviour also defines the membership in the three 4 different macro-classes.

This is the reason why the smearing is also associated to the electronic temperature. So we have resolved the k-point convergence problem, introducing a fictitious entropy term that diverges the value of energy from the real one. For this reason, the use of smearing must be dosed carefully always keeping in mind a good combination between it and k-point sampling. However, given its importance in the treatment of metals, different sophisticated smearing functions were born with the aim of minimizing the entropic term, as shown in Fig. 5. In our case we have adopted a Gaussian smearing.
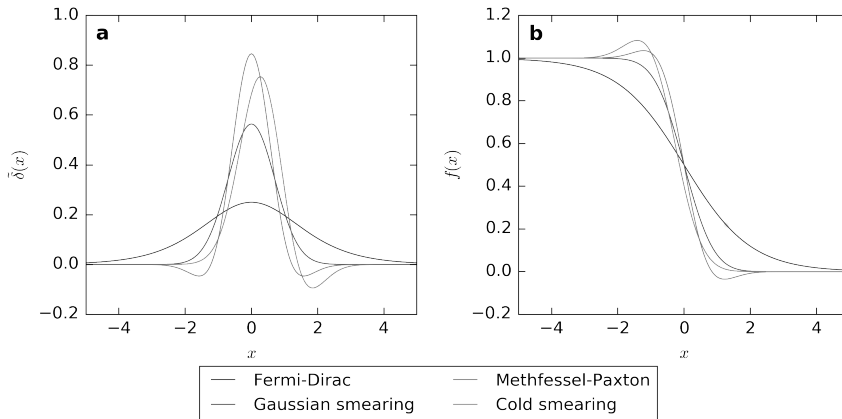

Figure 5: **a:** Different occupation functions[69, 70]. **b:** We show how the different occupation functions damp the discontinuous Fermi step-function at zero temperature.

2.2 MACHINE-LEARNING

Machine-learning (ML) is a term that encompasses a set of methods developed in various scientific communities with different names such as: computational statistics, pattern recognition, artificial neural networks, adaptive filtering, theory of dynamic systems, image pro-

cessing, data mining. These methods give to the computer systems the ability to "learn" from data, without being explicitly programmed to execute it.[71–78]

The goal of ML is to improve a system without the need of human intervention. Two components are required in order to learn:

- a set of data relating to the application domain;

- a learning algorithm capable of extracting knowledge from the data set.

It is important to underline that the quality of the data on which the algorithm works determines the outcome of the whole learning process. A clever and efficient algorithm will never be able to extract useful knowledge from a set of data that does not contain information about the application domain.

The ML algorithms can use four learning methodologies depending on the type of "knowledge" we expect: Supervised Learning (SL), UnSupervised Learning (USL), Semi-Supervised Learning (SSL), Reinforcement Learning (RL).

Supervised Learning   SL is the automatic learning methodology by which examples are passed to the machine as a pair of data $\mathcal{X}, \mathcal{Y}$ containing the input data and the expected result. The task of the machine is to find the rule, $\mathcal{K}$ (function or model), that establishes a relationship between the two so that it can obtain a correct result even when presented with a new sample

$$\mathcal{Y} = \mathcal{K}[\mathcal{X}]. \tag{28}$$

Possible applications of SL can be classification, regression, and image, speech, and pattern recognition.

UnSupervised Learning   As can be deduced from its name, USL, unlike the previous one, does not use data classified and labelled a priori. We do not know, therefore, to which category they belong. The machine is required to extract a rule that groups the cases presented according to characteristics that are derived from the data itself. For this reason it is also called "feature learning". Example of USL methods include clustering, dimensionality reduction, anomaly detection, and latent variable analysis.[79]

Semi-Supervised Learning   Located halfway between the two previous ones, SSL is based on mixed data in which a small part is already labelled and a large majority consists of unlabelled data. This approach is used to improve the machine's predictions of unlabelled data. Its possible applications are generally an extension of those of SL, where the cost related to the labeling evaluation might make a fully labeled training set infeasible.[80]

Reinforcement Learning   RL is an automatic learning technique that aims to develop sys-

tems capable of learning and adapting to changes of the environment in which they are immersed, through the distribution of a "reward" called "reinforcement", given by the correct evaluation of performance. It has potential applications in the fields of: game theory, control theory, operations research, information theory, simulation-based optimization, multi-agent systems, and swarm intelligence.[81]

ML algorithms make it possible to disentangle all the regularities and correlations hidden in simulation data. Applications of ML algorithms to computational solid state physics have already been successfully proposed in literature[82–89]

2.2.1   *Neural Networks*

A particularly promising class of models for ML of materials properties consists in the so-called deep-learning methods (DLMs) [86, 90, 91], where the term "deep" comes from the fact that there are multiple layers (at least two hidden layers, i.e. intermediate) (non-)linearly connected with simple nodes/objects. This kind of multi-layered networks organized hierarchically are capable of learning very complex functions. For these reasons, they can be applied as fitting algorithms [92]. Unlike other alternative estimation methods of the high-dimensional functions[84, 93, 94], DLMs are able to fit a real-valued function of any dimension making the evaluation statistically and computationally more efficient. The curve-fitting of a high-dimensional function (as a DFT-PES can be[4]) requires an enormous amount of input data, that is often limited for computational reasons. In fact, when the dimensionality of a function increases, the volume of the space increases so fast that the available data become sparse. In order to overcome this lack of information and to get a statistically reliable result we must increase exponentially the data set. This problem is known as *curse of dimensionality* and DLMs are not liable to it.

Moreover, hierarchical organization allows to catalog information (similarly to object-oriented programming). Along the hierarchy we can select specific features and discard unnecessary details in order to maximize the invariance.[95]

DLMs take their cue from the functioning of the human brain to create models capable of solving complex problems related to the perceptual sphere. This is why they are also commonly called Neural Network methods. Dated studies suggest, in fact, that the sensory cortex of mammals is approximately organized in a hierarchical way. The lower layers of this complex hierarchy extract elementary information from the sensory input: in the visual system, for example, neurons of the primary visual cortex encode some characteristics of

---

4 For example an atomic structure of 100 Al-atoms, its PES lives in a 300D space.

the stimulus as a set of simple spatially oriented bars. [96] As the hierarchy is moved up, sensory input is represented in an increasingly abstract way, thanks to the progressive extraction of invariances and statistics of a higher order. Think about how the brain organizes ideas and concepts, how it learns new knowledge, starting from simple basic concepts that it composes and extends in order to obtain more advanced concepts (See Fig. 6).
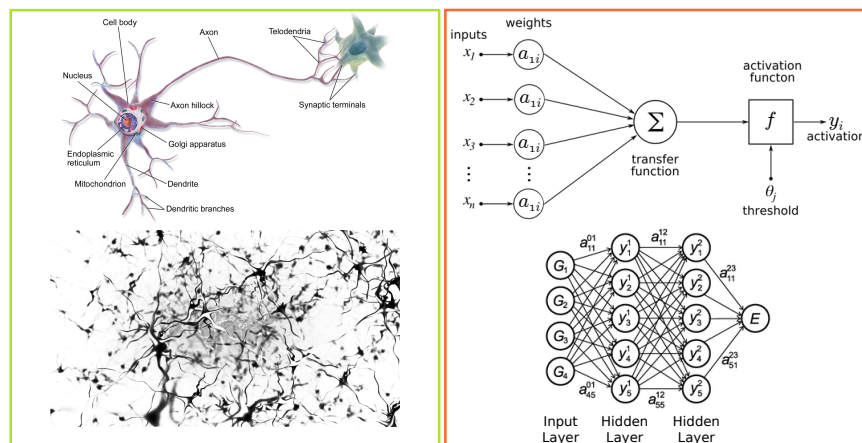

Figure 6: We show the similarities between an artificial and biological neural network. Artificial NN can manage the modelling of complex data through the use of probabilistic graphic theories, capable of learning the representation of data in input at different levels of abstraction.

NN methods can be subdivided according to the type of ML task they perform. In fact, we can have "discriminatory" feed-forward models for classification (or regression) by supervised or semi-supervised training (e.g Fully Connected Neural Network [97], Convolutional Neural Network[98], Hierarchical Temporal Memory[99]), generative models trained to reconstruct input and to produce salient features by unsupervised training (e.g. Stacked Auto-Encoders[100], Restricted Boltzmann Machine[101] and Deep Belief Networks[102]), and recurring model by reinforcement training in order to learn behaviour (e.g. Recurrent Neural Network[103], Long Short-Term Memory[104], and Deep Q-Learning[105]).

For the purposes of this thesis, we will deal with the feed-forward models and, in particular, only the Fully Connected Neural Networks (FC-NN), also called in literature in certain cases "Multilayer Perceptron" and often only Neural Networks (NN). Owing to a simpler structure, they are easy to implement. Furthermore, it is not obvious how the local connectivity of a convolutional NN might be helpful in an AIMD context. By always referring to AIMD simulations no memory or recurrency is involved: the energies and forces estimation at

given time do not require any knowledge of the earlier states of the system.

The feed-forward FC-NN was the first and simplest type of NN devised. In this network, the information moves in only one direction: forward through the layers. An example feed-forward FC-NN is shown in Figure 6(bottom right), consisting of an input layer, two hidden layers and an output layer. Nodes are represented by circles, while the arrows display the connections between the nodes, including the direction of information flow. Each connection is associated to an appropriate weight $a_{m,i}^{j-1,j}$ that represents the connection correlation between the two nodes (however, not all weights are shown here). We also observe that each node in a layer is connected to all nodes in the subsequent layer. For simplicity of notation and to avoid confusion, hereafter we will always name FC-NNs as Neural Networks (NN) that is their simplest and more common name.

In a NN, each artificial neuron (node) is sketched as in Figure 6(above right). It mimics a single biological neuron and we can express, for example, the i-th node of the j-th layer by a mathematical formulation as follows

$$y_i^j = f(X_i^j) = f\left(\sum_{m=1}^{M} a_{m,i}^{j-1,j} x_m + b_{m,i}^{j-1,j}\right) , \tag{29}$$

where x_m are the outputs of the nodes in the $(j-1)$-th layer. The sum runs over all M nodes(/initial inputs) in the previous layer since NN is fully connected. Instead, $a_{m,i}^{j-1,j}$ and $b_{m,i}^{j-1,j}$ are the weight connecting the m-th node in the $(j-1)$-th layer with the i-th node of the j-th layer, and its bias value, respectively.

In this framework, each complex multivariate function has a one-to-one correspondence with the associated NN diagram. In other words, any topology of NN diagram can be mapped as a specific function as long as there is no feedback loop in the NN diagram, namely, the information must travel in one direction through the layers. For the sake of clarity, we report the full expression of NN diagram shown in the Fig. 6(bottom right):

$$E = f\left(\sum_{k=1}^{5} a_{k,1}^{2,3} f\left(\sum_{r=1}^{5} a_{r,k}^{1,2} f\left(\sum_{m=1}^{4} a_{m,r}^{0,1} x_m + b_{m,r}^{0,1}\right) + b_{r,k}^{1,2}\right) + b_{k,1}^{2,3}\right) . \tag{30}$$

Of course, if we add a node or, even better a layer, we increase the complexity/flexibility of NN. This increased flexibility allows to fit more complex functions thanks to the universality of NN diagram[106]. But to do it, we are including more parameters on which both the NN and the estimated function depend, leading to side effects that fall under the terminology of overfitting. Overfitting occurs

when a set of points in a zone of configuration space is fit very accurately at the cost of a poor fit for the points in another zone. In Fig. 7, we show a typical example of overfitting.
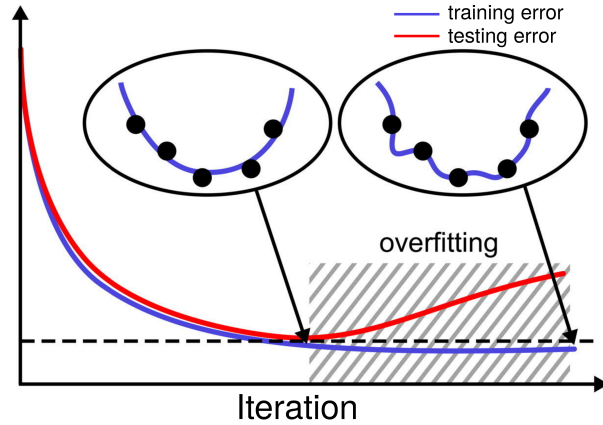

Figure 7:   The *early-stopping* method.   We see that at the beginning both errors (testing and training) decrease until the testing error reaches a minimum and, then, starts increasing. After that optimization step, we improve the fit of a set of points at the expense of others.[107]

In the NN framework, we can solve this problem by the *early-stopping* method: We train the NN on a sub-part of the whole input database (also called training set) and we use the remaining sub-part to test the optimized NN (also called testing set). The comparison of error obtained from training and testing sets is a good criterion of the quality of our NN. For more information see Fig. 7.

Activation Functions

As we can see in Eq. 30, the activation functions are the core of artificial NN, they calculate the output response according to the weighed sum that reaches the input. In order to learn non-linear complex functional mappings and to remove the linearity relationship between input and output, the activation function cannot be a simple linear function.

The activation functions are nothing more than non-linear mathematical functions that respect the following characteristics to fulfill the NN universality (the so-called universal approximation theorem [106]):

- $f \in C^1$ where C^1 is the the space of continuously differentiable functions;

- bounded;

- $f' \geqslant 0$.

These properties are also essential to use backpropagation as an optimization strategy, discussed in the next section.

Examples of commonly used activation functions:

- **Sigmoid function**

$$f(x) = \frac{1}{1 + e^{-x}} \; ; \tag{31}$$

- **Hyperbolic tangent function**

$$f(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \; ; \tag{32}$$

- **Rectified linear units** (ReLu)

$$f(x) = max(0, x) \; . \tag{33}$$

In NNs the most (historically) used activation function is the sigmoid. In very deep networks, sigmoid use is problematic for gradient back-propagation (vanishing gradient problem). The sigmoid derivative is typically lower than 1 and the application of the chain rule at the derivative leads to a multiplication of many terms lower than 1, with the result of reducing quite the gradient values in levels far from the output. In addition, its output is not zero centered and it makes optimization harder. Finally, sigmoids have slow convergence which facilitates error fluctuations during the training stage.

The hyperbolic tangent is often preferred to other functions since calculating the gradient is inexpensive.

The rectified linear units has, instead, become very popular recently, after it proved to be 6 times better in convergence than *tanh* function. It overcomes the vanishing gradient problem. But it does not have an upper bound and has an isolated singularity, and therefore does not satisfy the universal approximation requirements. Because of this it can only be used in the hidden layers. Also, it can occur that ReLu manifests the problem of dead neurons[5]. [108]

2.2.2 *Training and Optimization*

In the case of NN, the learning is achieved by interactively feeding the network with input data. With the help of certain learning algorithms, the parameters of network will be then optimized in order to extract salient information from the inputs. In material science we make use of NNs for regression: the goal is to interpolate inputs of atomic coordinates with first-principles energies and forces to evaluate interatomic potentials with comparable accuracy. Interpolation through

---

5 A dead neuron has a associated gradient with 0 value due to this fact it will never be adjusted by the optimizer

NNs is essentially an *optimization* problem, where the NN parameters (e.g. weights and biases) are refined iteratively to minimize the error in reproducing a set of expected energies values. This procedure is defined as *training*. Like any optimization algorithm we need to define a cost function to be minimized during the procedure. In our case, we can write this such as the mean squares error

$$\Gamma = \frac{1}{2N} \sum_{i=1}^{N} (E_{i,NN} - E_{i,DFT})^2 \, , \tag{34}$$

where N is the number of data points in the train set, $E_{i,NN}$ is the energy calculated by NN and $E_{i,DFT}$ is the reference energy calculated by DFT. This cost function, if it is differentiable[6], can be minimized by a wide class of algorithms. Among these we can find the Gradient Descent (GD), the Conjugate Gradient (CG), BFGS method (such as Simple and Limited-memory BFGS), the Stochastic Gradient Descent (SGD) and its variants (such as Momentum, Adadelta, Adagrad, AMSGrad, Adagrad, Nadam, Nestov accelerated gradient, and RMSprop) and the Kalman Filter (KF) category (such as Simple, Extended KF, and Decoupled Extended KF).

The GD applied to NN normally does not guarantee the convergence to a minimum if we do not tune correctly its step size during search procedure. It mainly aims at finding local minimums at the expense of the global one. CG, as also GD, are often redundant in operations and BFGS is inefficient from the point of view of memory consumption. Limited-memory BFGS can be a good alternative to solve the memory management, but a clear convex dynamics is still needed.

The intrinsically non-convex nature of NNs makes hard to optimize their cost function. In order to put a remedy to this, we have to use SGD and its variants. The SGD category, indeed, works well for non-linear dynamics, as in NN context, and is often used where the dynamics are not evident too. On the other hand, like all gradient-based algorithms, it tends to be slow (e.g. it is model free and relies on noisy measurements of the gradient).

The KF category is among the main ones used in the deep learning framework, despite the adaptive approach of SGD (e.g Adagrad, Adadelta and Adam) are widely used for the NN training, aslo thanks to their robustness and their excellent management of the sparse data,. In our case we used a KF methods since they are less likely to get trapped in shallow local minima. For example, from a recent study it was discovered that the training by extended KF does lead to a better positioning performance and it requires a lower effort than SGD algorithms.[109]

---

6 The cost function is differentiable only if the activation function is differentiable.

Having discussed possible optimizers, now we need to explain how to calculate the gradients of the cost function. In the feed-forward NN field, this customary procedure is referred to as back-propagation. In back-propagation, an input sample is propagated forward through the NN to produce an output. This output is compared to the expected output (reference value), and the error is then propagated backwards through the layers to obtain the estimation according to which each parameter must be adjusted, hence the name. The method can be seen as an implementation of the chain rule, and it allows us to calculate the partial derivatives of the cost function with respect to all the parameters, thus obtaining the gradient of the network. We define the derivatives of the cost function with respect to all the parameters (weights and biases) as follows

$$\mathcal{A}_{m,i}^{j-1,j} \;=\; \frac{\partial \Gamma}{\partial a_{m,i}^{j-1,j}} \, , \tag{35}$$

$$\mathcal{B}_{m,i}^{j-1,j} \;=\; \frac{\partial \Gamma}{\partial b_{m,i}^{j-1,j}} . \tag{36}$$

Now, we get the following equations if we apply the chain rule and we use Eq. 29

$$\mathcal{A}_{m,i}^{j-1,j} \;=\; \frac{\partial \Gamma}{\partial X_i^j} \frac{\partial X_i^j}{\partial a_{m,i}^{j-1,j}} = \sum_r^{N_{node}^{j+1}} \underbrace{\frac{\partial \Gamma}{\partial X_r^{j+1}}}_{1} \underbrace{\frac{\partial X_r^{j+1}}{\partial y_i^j}}_{2} \underbrace{\frac{\partial y_i^j}{\partial X_i^j}}_{3} \underbrace{\frac{\partial X_i^j}{a_{m,i}^{j-1,j}}}_{4} , \tag{37}$$

$$\mathcal{B}_{m,i}^{j-1,j} \;=\; \frac{\partial \Gamma}{\partial X_i^j} \frac{\partial X_i^j}{\partial b_{m,i}^{j-1,j}} = \sum_r^{N_{node}^{j+1}} \underbrace{\frac{\partial \Gamma}{\partial X_r^{j+1}}}_{1} \underbrace{\frac{\partial X_r^{j+1}}{\partial y_i^j}}_{2} \underbrace{\frac{\partial y_i^j}{\partial X_i^j}}_{3} \underbrace{\frac{\partial X_i^j}{\partial b_{m,i}^{j-1,j}}}_{4} . \tag{38}$$

By carrying out the partial derivatives of the previous equation and going deeper into the network, we obtain the full gradient of the NN for each node in any layer

$$\mathcal{A}_{m,i}^{j-1,j} \;=\; \underbrace{y_m^{j-1}}_{4} \underbrace{y_i'^j}_{3} \sum_r^{N_{node}^{j+1}} \underbrace{a_{i,r}^{j+1,j}}_{2} \underbrace{\frac{\partial \Gamma}{\partial X_r^{j+1}}}_{1} , \tag{39}$$

$$\mathcal{B}_{m,i}^{j-1,j} \;=\; \underbrace{1}_{4} \underbrace{y_i'^j}_{3} \sum_r^{N_{node}^{j+1}} \underbrace{a_{i,r}^{j+1,j}}_{2} \underbrace{\frac{\partial \Gamma}{\partial X_r^{j+1}}}_{1} . \tag{40}$$

2.2.3 *NN Potentials*

NN potentials built to predict atomic-scale properties are revolutionizing molecular modelling, making it possible to get first-principles

accuracy, at a fraction of the costs.[110–114] But before we attain a level of quality that makes it highly competitive in the MD field, we need to adopt additional precautions. The first NNs applied to the world of atomistic simulations found many limitations associated mostly with the low extensibility of the method.[115, 116] They used a single NN to estimate energy by feeding directly on the atomic coordinates of the system. This type of single NN, once fitted, is dependent on the number of atoms, in other words, it can map just PES at fixed dimensionality. If you wish to apply a single NN to systems with thousands of atoms, it would require the adoption of thousands of input nodes, an excessive number of hidden layers and, consequently, hidden nodes, which destabilize and give even to the most advanced optimization algorithms a hard time.

It is clear that their direct implementation is not feasible in the regime to which we wish to apply it. Typically, kinetic-Monte Carlo simulations used to deal with the precipitation issue involve samplings with variable compositions for different (large) system-size. Creating a new NN each time, based on the number of atoms and chemical species included, makes the procedure costly and hard to control.

The first NN scheme, which was really practicable to deal with extensive high-dimensional PES, has been suggested by Behler and Parinello[117], concurrent with the equally promising Bholoa's one[118]. In the next section, we will focus on Behler's algorithm because it is one of the two used in this thesis, togheter with the alternative implementation of Kobayashi (See Sec. 2.2.3).

Behler-Parinello NN

Behler and Parinello propose to replace the single NN by a set of atomic NNs, where each NN provides the energetic contribution E_i of the i-th atom to the total energy E of the system. In this way, we have again that the NN architecture is fixed but for each element (for example, in the case of Al-Mg-Si alloy we have 3 fixed NN). So we get a set of NN potentials generated ad-hoc for each atomic species (that is an atom-centered approach) in a similar way to when we achieve a standard class of DFT pseudo-potential. In addition, as usual, to achieve a FF potential, they assumed that the main contribution to E_i and, therefore, to E came from the atoms closest to the i-th atom.

Then, we can express the total energy as

$$E = \sum_{X}^{N_{as}} \sum_{i}^{N_X} E_i[\mathbf{r}_{ij}], \tag{41}$$

where N_{as} is the number of atomic species and N_X is the number of atoms for the X-element; while r_{ij} are the atomic distances between

the atoms i and j, in which the j-th atoms are placed within a sphere centered on the i-th atom with a cutoff radius r_c of the order of 10Å (in order to have a tractable chemical environment), and the positions are written in the Cartesian coordinate system.

Another problem which arises when Cartesian coordinates (that are typical format in the atomistic simulations) are adopted as NN inputs, is the need to preserve the invariance of the total energy with respect to symmetry operations: e.g. exchanges of equal atoms, rotations and translations of the entire system. This is viable applying a function which turns all the Cartesian atomic distances r_{ij} in a vector of values $\{G^i\}$, which respects these invariances/symmetries. This new set of vectors $\{G^i_X\}$ are then the input nodes of the i-th X-atom for an atomic NN that computes the atomic energy E_i (See Fig. 8).

Eq. 41 can be re-written as follows

$$E = \sum_{X}^{N_{as}} \sum_{i}^{N_X} E_i[G^i_X] \tag{42}$$

Using these "symmetry functions" as inputs decouples the number of input nodes from the number of neighbouring atoms, by storing angular and radial information of the various local environments.

Through this process, each atom has a specific atomic NN with a fixed set of symmetry function as input but with different $\{G^i_X\}$ values. All atoms of the same atomic species have the same atomic NN diagram and the symmetry function set. For example, in the case of Al-Mg-Si alloys, we have only three exclusive NNs that need to be trained. In this manner, the total energy E is invariant compared to exchanges of the positions of equal atomic species.
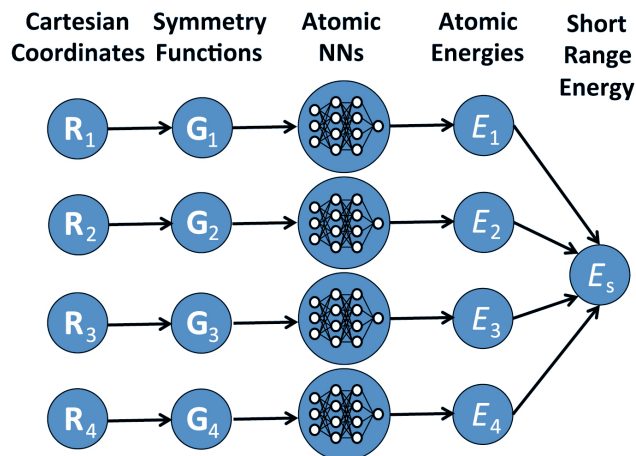

Figure 8: The NN diagram for mapping a potential energy surface in atomistic simulations. [86] It is implemented in RuNNer package [119].

Now, we just summarize briefly the form and parameterization

Symmetry Functions

of two important families of Symmetry Functions (SF), proposed by Behler and Parrinello, which will be used in our results.

All the symmetry functions describe the correlations between atoms in a neighbourhood of a central atom with index i. The first functional form, called G_2 following the convention used in previous works [112, 120, 121], provides information about pair correlations between the atoms

$$G_2^i = \sum_{j \nleqslant i} e^{-\eta(R_{ij} - R_s)^2} \cdot f_c(R_{ij}), \qquad (43)$$

where the parameters η and R_s control the width and the position of the Gaussian with respect to the central atom, and $f_c(R_{ij})$ is a cutoff function that ensures that the symmetry function smoothly decreases to 0 in value and slope at a fixed cut-off r_c. The sum is over all neighbouring atoms which are closer than r_c. The second type of symmetry functions, called G_3, provides information about angular correlations, and has the form

$$G_3^i = 2^{1-\zeta} \sum_j \sum_{k \neq j} (1 + \lambda \cdot \cos \theta_{ijk})^\zeta \cdot$$
$$e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} \cdot f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk}), \quad (44)$$

where ζ, η, and λ are the three parameters that determine the shape of this type of symmetry function. The indices j and k run over all the atoms in the neighbourhood of the tagged atom i. The cut-off function that we have used has the form

$$f_c(R_{ij}) = \begin{cases} \tanh^3 \left[1 - \frac{R_{ij}}{r_c} \right] & \text{for} \quad R_{ij} \leqslant r_c \\ 0.0 & \text{for} \quad R_{ij} > r_c \end{cases}. \qquad (45)$$

Forces and Stress   From what we can see from Eqs. 43 and 44, SFs have analytic derivatives. This is indispensable in order to get properties such as forces and stress. Indeed, we can write the force along the k-direction acting on atom m with atomic positions $\mathbf{R}^m$ by applying the chain rule

$$\mathbf{F}_k^m = -\frac{\partial E}{\partial \mathbf{R}_k^m} = -\sum_X^{N_{as}} \sum_i^{N_X} \frac{\partial E_i}{\partial \mathbf{R}_k^m} = -\sum_X^{N_{as}} \sum_i^{N_X} \sum_s^{N_G} \frac{\partial E_i}{\partial G_s^i} \frac{\partial G_s^i}{\partial \mathbf{R}_k^m}, \quad (46)$$

where N_X and N_{as} are respectively the number of atoms and the atomic species within local environment of the atom m included itself; instead G_s^i and N_G are the SFs of all N_X atoms and their corresponding number. The first term of Eq. 46 is easily given by the NN scheme while the second one is given by the SF definition.

With regards to the static contribution to the stress tensor, we can get its expression along α and β Cartesian coordinates using the Virial theorem[7] [122] and Eq. 46:

$$\sigma_{\alpha,\beta} = \sum_n \sum_m (\mathbf{R}^m - \mathbf{R}^j)_\alpha \, \mathbf{F}_\beta^m; \tag{47}$$

n and m are atomic indexes.

Kobayashi NN

The NN potential developed by Kobayashi is basically the same as the one originally developed by Behler and Parrinello, shown in previous section.[21, 113] It is implemented in the NAP (Nagoya Atomistic-simulation) package.[123] However, since there are some differences, here we describe the structure of the NN potential.

The energy of an atom-i in a structure-s is defined as

$$E_i^s = \sum_m w_{1m}^2 y_{i,m}^1, \tag{48}$$

$$y_{i,m}^1 = f_a\left(\sum_n w_{mn}^1 G_2^{i,n}\right). \tag{49}$$

Here w_{mn}^l is the weight of a line from a node-n in the $(l-1)$-th layer to a node-m in l-th layer of the NN, and $y_{i,m}^l$ is the value of node-m in l-th layer, Also, $G_2^{i,n}$, the so-called *symmetry function*, is the n-th input which depends on the interatomic bond distances R_{ij} from an atom-i to surrounding atom-j. The activation function $f_a(x)$ is defined using the sigmoid function as

$$f_a(x) = \frac{1}{1+e^{-x}} - \frac{1}{2} \tag{50}$$

where $1/2$ is subtracted from the sigmoid function so that $y_{i,m}^1$ becomes zero when all the inputs are zero, resulting in the energy being zero.

There are many choices of symmetry functions. Here we employ the Gaussian function

$$G_2^{i,n} = \sum_{j\neq i} e^{-\eta_n(R_{ij}-R_n)^2} f_c(R_{ij}), \tag{51}$$

where η_n and R_n are the parameters that are determined heuristically before training the network. The cutoff function $f_c(R)$ is defined as

$$f_c(R) = \begin{cases} 1, & \text{for } R \leqslant R^*, \\ \frac{1}{2}\left[\cos\frac{\pi(R-R^*)}{(R_c-R^*)} + 1\right], & \text{for } R^* < R \leqslant R_c, \\ 0, & \text{for } R > R_c. \end{cases} \tag{52}$$

---

7 The Virial theorem is valid since the total energy is a function of inter-atomic distances.

where R_c is the cutoff radius and $R^* = 0.9R_c$.

As in Behler-based algorithm, the objective function to be minimized in training of the NN potential is

$$\mathcal{L}\left(\{w\}\right) = \frac{1}{2M} \sum_s^M \left[\left(\frac{\Delta E^s}{\varepsilon_e^s}\right)^2 + \sum_i^{N_a^s} \sum_\alpha^{xyz} \frac{1}{3N_a^s} \left(\frac{\Delta F_{i,\alpha}^s}{\varepsilon_f^s}\right)^2 \right] \quad (53)$$

where $\Delta E^s = \left(E^{s,NN} - E^{s,ref}\right)/N_a^s$ and $\Delta F_{i,\alpha}^s = F_{i,\alpha}^{s,NN} - F_{i,\alpha}^{s,ref}$ are the differences in energies and force components, respectively, as obtained by the NN potential and the DFT calculations. The superscript s indicates the index of a sample, N_a^s is the number of atoms in the sample-s, and M is the number of samples. The quantities ε_e^s and ε_f^s are the convergence criteria for the energy per atom and the force component for a sample-s, respectively; the objective function $\mathcal{L}$ becomes less than one when ΔE^s and $\Delta F_{i,\alpha}^s$ are smaller, on average, than these convergence values.

A parameter-rich model like the NN can often show overfitting, in which the model is well optimized to the training set data but reproduces very poorly data not included in the training set. This is also known as low transferability of the interatomic potential. There are various ways to avoid overfitting, such as the *early stopping* approach[107, 124], which stops the optimization before the overfitting begins, and the *weight decay* approach[124] which adds a penalty term to $\mathcal{L}$ to suppress the parameters having large values that usually causes the overfitting. Because it is not easy to determine when to stop the optimization in the *early stopping*, we adopted the *weight decay* approach by adding a penalty term, the so-called *ridge penalty*, to Eq. (53) as

$$\mathcal{L}^*(\{w\}) = \mathcal{L}(\{w\}) + \lambda \sum_i w_i^2 \quad (54)$$

where λ is the penalty parameter to improve the generalization or transferability of the NN potential. The hyperparameter λ is determined heuristically to be as large as possible so that the penalty term is small enough not to exceed the $\mathcal{L}$ value. The minimization of the function $\mathcal{L}^*(\{w\})$ is carried out using a quasi Newton method, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm.[125]

Parameters which are not trained from data, the so-called hyperparameters, must be specified before optimization. In this NN model, the hyperparameters are the number of hidden layers, the number of nodes in a hidden layer and the set of symmetry functions. In general, these hyperparameters are determined by trial and error, which validate the obtained NN potential using several physical values. The resulting hyperparameters of the NN potential proposed were eventually chosen as follows: the number of hidden layers is 1 and the

number of nodes in a hidden layer is 30; while for what concerns the SFs selection see Sec 4.2.

2.3 KINETIC-MONTE CARLO

As we have seen in Sec 2.1, MD integrating the equations of motion, is not so feasible due to the timescale. Indeed, the time step needs to be in the order of fs for a meaningful integration of the equations of motion limiting the simulations in the best cases to some μs and in our case (that is AIMD) to some ns (See the fig. 9). The challenge in MD simulations is to find reliable ways to capture infrequent processes and extend them to longer time scales with reasonable computational resources. This known issue of time scale of
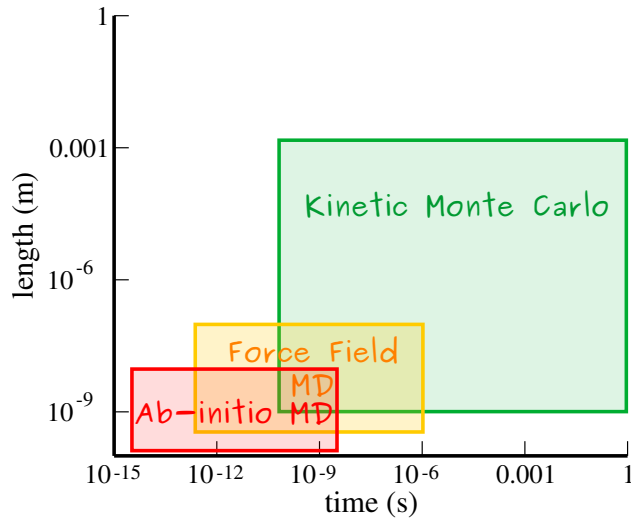
Figure 9: The computational limits of three theories.

AIMD can be overcome by other sophisticate techniques such as Accelerated Molecular Dynamics (AMD) or by the Kinetic Monte Carlo (KMC) technique. KMC simulations appear very promising because they provide the ability to perform atomic-level simulations of the defect kinetics and microstructure evolution over relevant length and time scales. KMC manages to be more efficient than MD to study state-to-state kinetics without significant loss in accuracy by using an accurate list of process rates.[126, 127] On the other hand, the main disadvantage of KMC is the need to know a priori the list of all possible events/processes and their respective rates. The determination of these rates is feasible through Transition State Theory in the context of validity of the harmonic approximation. In this framework, the rates of transition for a given i-th event, k_i, in the list can then be calculated using

$$k_i = \nu_i^0 e^{(-E_i/k_B T)} \tag{55}$$

where ν_i^0 is the prefactor and E_i is the activation energy for the i-th process and k_B is the Boltzmann constant and T is the temperature. The challenge of this model, which is more manageable from the mathematical point of view w.r.t. the equations of motion, is to know how to represent the vector of the rates in the best way. It is precisely that vector that characterizes our system and reproduces its dynamics. From Eq. 55, we can understand that the description of an event depends on two parameters:

1. The activation energies

2. The prefactors

Afterwards we are able to evaluate these reaction rates, they are introduced into a discrete-state KMC model to explore the diffusion and aggregation phenomena in the contest of alloys.

2.3.1 *Transition State Theory*

TST provides the basis for the calculation of transition process rates, assuming only knowledge of the equilibrium and saddle point states by using the Arrhenius expression, previously seen in the Eq. 55. It also defines the overall rate of all possible process of the system as

$$R = \sum_i R_i = \sum_i n_i k_i \tag{56}$$

where R_i is the global rate and r_i is the local rate (as reported in the Eq. 55 associated with the *i*-th process, while n_i is the number of particles in the system that are candidates for this kind of process. As discussed in the previous section, a complete knowledge of the rates vector, eqs. 55 and 56, is a prerequisite for all KMC methods. In order to know the rates vector, it is necessary to have a complete awareness of the energy landscape. An energy landscape is a $d \cdot N$ dimensional PES with the N number of bodies (e. g. atoms) in a d-dimensional system. In a typical state-to-state motion, the system crosses the energy space between two adjacent minima via a saddle point, which is a critical point where the Hessian has both positive and negative eigenvalues. It is generally assumed that only the transition states (e.g. PES points wherein only one eigenvalue of Hessian is negative) describe the relevant kinetics of the system.

There are different approaches to determine the transition state energies[128] (e.g. Drag method, Nudged Elastic Band (NEB), Climbing Image-NEB (CI-NEB), Conjugate Peak Refinement (CPF) , Ridge method, and Dimer method) and for the purpose we want to achieve, we follow the Ci-NEB approach, which it will be presented more clearly in the Sec. 2.3.1.

Since the early stage precipitate structures (as Si-clusters and/or GP zones) are coherent phases within the Al-matrix, which is periodic, the full energy landscape can be defined by identifying all the possible permutations of nearest-neighbor atomic arrangements.

An on-lattice based approach (that is an Al-sites grid), then, would be enough to study the dynamics of Al-Mg-Si alloy precipitation. However the precipitation sequence is not so trivial and KMC must know how to represent also semi-coherent phases such as β'', that do not fit within the on-lattice picture.

The value of the rates is much more influenced by a small variation in the energy barriers than a small one in ω, and so in many cases the prefactor has been assumed to be a constant value in the range between 1 and 10 Thz.[129, 130]

Nevertheless, in order to describe all accessible defect diffusion mechanisms, it is necessary to have a careful calculation of the prefactor; as shown in recent studies, where the prefactors can space in a more extensive range (1 Thz-10^7 Thz).[131, 132] The standard procedure to determine the prefactor imposes the validity of the harmonic regime of the TST and the use of the Vineyard equation. This reduces the determination of the prefactor to the calculation of the normal modes as follows

$$\nu_i^0 = \frac{\prod_{j=1}^{M} \omega_j^{LMS(i)}}{\prod_{j=1}^{M-1} \omega_j^{TS(i)}}, \tag{57}$$

here $\omega_j^{LMS(i)}$ and $\omega_j^{TS(i)}$ are the vibrational normal modes at the Local Minimum States (LMS) and Transition States (TS) for each of the possible events, respectively.[133] The normal frequencies used in this equation at the LMS and TS are derived from the eigenvalues of the Hessian $\mathcal{H}$ using a finite-difference construction.

$$\mathcal{H}_{ij} = \frac{1}{2\sqrt{m_i m_j}} \left(\frac{F_i(+\delta_j) - F_i(-\delta_j)}{2\delta} + \frac{F_j(+\delta_i) - F_j(-\delta_i)}{2\delta} \right), \tag{58}$$

in which $F_i(+\delta_j)$ is the force acting on i-th component due to the positive position displacement in the j-th component, δ is the displacement, and m_i is atomic mass of the i-th atom. The value of δ has to be chosen small enough to get correct finite-difference derivatives.

NEB method with Climbing Image

A model of defect stability and migration barriers is accomplished better using first-principles evaluation of the electronic structure of the material. A first-principles theory, which also considers the electronic structure of the material, is suited to predict the plausible defect energies, the stability and migration barriers. This was confirmed by the fact that the use of theories based on Embedded Atom Model

has often led to predictions in disaggreement with the results from *ab-initio* theories and experiments.[134, 135]

Since the whole process is strongly vacancy diffusion governed[20], we have limited ourselves to evaluate the events involving the displacement of a vacancy. In a lattice, the displacement of a vacancy takes place through an exchange with one of the sites adjacent to the one of vacancy. In the case of ternary metal alloy these possible exchanges/events are reduced to three if one neglects the effect of the atoms that are not directly involved in the exchange. These events have a known final and initial state that has prompted us to choose a NEB approach with climbing image in order to search for these three saddle points and their energy barriers.

In our case, we will use the CI-NEB implemented in the QE [49] package to evaluate all Minimal Energy Paths due to vacancies migrations.

The general scheme works by linearly interpolating a set of images between the known initial and final states, also called Linear Energy Path (See Fig. 10), and then minimizes the energy of this string of images, following this force definition for i-th image

$$\mathbf{F}_i^0 = -\nabla V(\mathbf{R}_i)|_\perp + \mathbf{F}_i^s \cdot \hat{n}_\parallel \, \hat{n}_\parallel \,, \tag{59}$$

where $\hat{n}_\parallel$ is the tangent versor to the path, $-\nabla V(\mathbf{R}_i)|_\perp$ is the perpendicular component of the gradient of the potential with respect to $\hat{n}_\parallel$, and $\mathbf{F}_i^s$ is the spring force between two adjacent images.
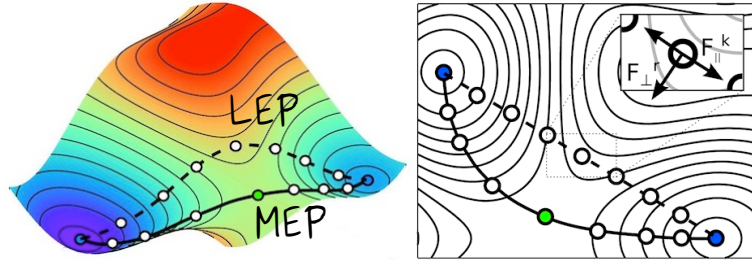

Figure 10: MEP and LEP between reagent and product. Each image finds the lowest energy possible minimizing the force due to the potential perpendicular to this elastic band as in Eq.59.

In this way, the choice of the spring forces is arbitrary since it only influences the division of the images along the path.

However, if the energy of system changes suddenly along the path and $-\nabla V(\mathbf{R}_i)|_\perp$ is small, the MEP searching could have convergence problems. Indeed, the images perceive a strong parallel force that tries to slice them down while the NEB maintains constant the image spacing. In this regime, the NEB chain could reach the convergence only by a strong stretching and, then, "breaking" the chain. So, the

consequential evaluation of new tangents may become problematic. A possible solution to this problem introduces an additional term to the equation 59 that smoothly turns on the perpendicular component..

$$\mathbf{F}_i^{NEB} = \mathbf{F}_i^0 + f(\theta_i) \left(\mathbf{F}_i^s - \mathbf{F}_i^s \cdot \hat{n}_\parallel \, \hat{n}_\parallel \right) , \tag{60}$$

with $f(\theta_i)$ s a switching function which goes from 0 for a straight path to 1 if two subsequent segments along the path form an angle of 90°. A small amount of the perpendicular spring force is often enough to get out of pathological curvatures of PES.[136]

The CI-NEB, which does not differ much from the NEB method, leads to a better estimate of the exact saddle point for the highest energy image. This image does not feel the spring force along the band and its tangent force is inverted. In this manner, the image tries to maximize its energy along the band, and minimize in all other directions. [137]

We will discuss our results for these barriers, that are consistent with previous literature, in Sec. 4.3.3.

2.3.2 *Markov Chain and Poisson Process*

KMC evolves a system through numerical sampling of (Markovian) stochastic processes which generate a series of kinetic states, and determines the physical time passed during this evolution by calculating the probabilities of each event. Since KMC deals with stochastic processes that jump from one to other state in a continuous time, it can be viewed as a method to solve the Master Equation[138]. According to the Markov state assumptions, Master Equation has to preserve two important conditions: *steady state* and *detailed balance*.[138]

> **The steady state** implies the sum of all the transition into a particular state n equals the sum of all the transition out of a particular state n′
>
> $$\sum_m W_{mn} p_n(t) = \sum_m W_{mn} p_n'(t) \tag{61}$$

where $p_n(t)$ is the probability to be at n-state at time t and $W_{n'n}$ is the matrix of transition probability per unit time, namely the process rate of going from configuration n to that m.

> **The detailed balance** implies the ratio of the transition probabilities at equilibrium for a move $n \rightarrow m$, whilst the inverse move $m \rightarrow n$ depends only on the energy change.
>
> $$\frac{W_{mn}}{W_{nm}} = \frac{p_{eq,n}}{p_{eq,m}} \tag{62}$$

The detailed balance condition is fundamental because it ensures the equilibrium distribution of occupied states is consistent with Boltzmann distribution.[139]

However, the mere imposition of having a Markovian stochastic process is not sufficient to establish a connection between Monte Carlo time-step and real physical time. For this reason, KMC predicts that the system evolves as if it were a one-step process[8], and, in particular, assumes a Poisson distribution of residence times. A Poisson process, in addition to supporting the conditions of one-step process, describes only independent events where the transition probabilities depend only on the starting state. So, the probability that n events happen depends only on the time interval. In this way, the Master Equation assumes the following simple form

$$\dot{p}_n(t) = R(p_{n-1} - p_n) \tag{63}$$

where R is the sum of rates as shown in Eq. 56. The solution to this equation, which is given by

$$p_n(t) = \frac{Rt}{n!} e^{-Rt} \tag{64}$$

clearly shows a stationary series of independent, random events occuring with an average transition rates R in the framework of a Poisson process.[140, 141]

It is important not forget that the purpose of all this discussion is to find the real time associated with the occurrence of a process. Now this can be determined by integrating the probability density of Eq. 64 in a time τ so we attain the probability for an event to happen within a time τ

$$p_\tau = \int_0^\tau R\, e^{-Rt'} dt' = 1 - e^{-R\tau}. \tag{65}$$

Therefore, the probability for an event not to happen within τ is $\bar{p}_\tau = e^{-R\tau}$. Since $\bar{p}_\tau$ could be any random number in the interval [0,1], the real time of two successive events is estimated as

$$\tau = -\frac{\ln \bar{p}_\tau}{R}. \tag{66}$$

Based on the above theory, the KMC method is very well suited to simulate the real-time evolution of slow process such as nucleation phenomena and precipitation.[142–146] Indeed, the fact that KMC has a kinetic character makes it possible to study non-equilibrium dynamic processes, in particular those where energy barriers govern the transition between subsequent states.

---

8 An one-step process allows to have one jump at a time and only jumps between adjacent states.

POST-β''-PHASE PRECIPITATION

Contents

1

I n this chapter we study the energetics of nanoscale β''-phase precipitates using *ab initio* electronic structure methods so as to identify the different contributions to the thermodynamic in-situ precipitation energetics. We compute the energy contributions due to the precipitate formation energy, the precipitate/matrix interface energies, and the elastic energy due to lattice and elastic mismatch between precipitate and matrix. We show that these contributions semi-quantitatively capture the total energy of in-situ precipitates as a function of precipitate size. Our results demonstrate that – down to the size of a single formula unit of the β'' phase, fully encapsulated in the Al matrix – the precipitate growth process can proceed without energetic barriers. Since the nucleation process of the β'' phase has nearly zero barrier, control of precipitation kinetics should focus on aggregates of atoms of even smaller size.

The remainder of this chapter is organized as follows. In Section 3.1 we describe the details of our *ab initio* simulations. In Section 3.2 we report a few benchmarks on the bulk properties of the different stoichiometries proposed for the β'' phases. In Section 3.3 we discuss a classical-nucleation-theory (CNT) model of precipitate stability, including surface energies and the continuum elasticity model of lattice mismatch relaxation, and compare with DFT results for needle-like precipitates. In Section 3.4 we present *ab initio* simulations of fully-encapsulated clusters. We finally draw conclusions.

---

1 This Chapter is adapted from Ref. [20].

3.1   COMPUTATIONAL DETAILS

Density functional theory (DFT) has been shown to provide reliable energetics for aluminum and its alloys [11, 14, 18, 147]. We have used self-consistent DFT as implemented in the Quantum ESPRESSO (QE) package[49]. We used a gradient corrected exchange and correlation energy functional (PBE)[148], together with a plane-waves expansion of Kohn-Sham orbitals and electronic density, using ultra-soft pseudopotentials for all the elements involved [149–151]. All calculations were performed with a k-point sampling of the Brillouin zone using a grid density of $\approx 5 \cdot 10^{-6}$ Å^{-3} and a Mokhorst-Pack mesh[152]. The plane-wave cut-off energy was chosen to be 35 (280) Ry for the wavefunction (the charge density) when evaluating the energetics of defects (i.e. for computing formation, surface, and precipitation energies). Test calculations performed at larger cutoffs showed that these parameters are sufficient to converge the atomization energy of Al at a level of 0.3 meV/atom. Cutoffs were increased to 50 (400) Ry so as to converge the value of the elastic constants to an error below 1 GPa. Comparison with previous literature results, where available, will be presented below.

3.2   BULK PROPERTIES OF MATRIX AND PRECIPITATE PHASES

Bulk properties (lattice structure, lattice constants, elastic constants) of Al and the various β″-precipitates studied here have been previously computed in the literature. Here, we present our results as a means of benchmarking our methods, verifying literature results, and most importantly obtaining reference values that are fully consistent with our computational details – which is crucial to evaluate the energy differences that determine surface and defect energies.

For bulk *fcc* Al, we computed the lattice parameter to be 4.057 Å, in excellent agreement with the experimental value and with previous modelling using the same functional [153, 154]. These lattice parameters are used throughout our study to build supercells representing the Al matrix. All of the β″ phases we consider can be described by a monoclinic cell containing two formula units (f.u.). We consider three compositions, Mg_5Si_6, $Mg_5Al_2Si_4$ and $Mg_4Al_3Si_4$, as shown in Figure 11. We computed the crystal structures of these β″-precipitates starting from the geometries proposed in previous works [4]. The equilibrium lattice parameters and monoclinic angles are shown in Table 1, and agree well with existing literature [155]. Inside the Al matrix, the main crystallographic directions (lattice vectors) of the precipitate are aligned with those in the *fcc* lattice of aluminum as follows:

$$[100]_{β″} \parallel [203]_{Al} \qquad [010]_{β″} \parallel [010]_{Al} \qquad [001]_{β″} \parallel [\bar{3}01]_{Al}. \quad (67)$$
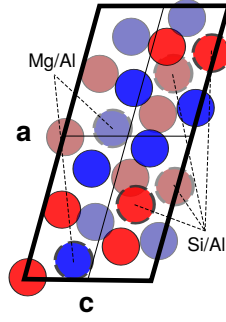
Figure 11: A view along the **b** lattice vector of the monoclinic unit cell of β″ phases. The red and blue circles represent Si and Mg atoms, respectively, while the different shading indicates the position of the atoms at a height of zero and $|\mathbf{b}|/2$ along the **b** vector. Circles with dashed outline indicate the atoms that can be substituted to obtain the three β″ stoichiometries (that is, Mg_5Si_6, $Mg_5Al_2Si_4$ and $Mg_4Al_3Si_4$)[11].

The ideal monoclinic unit cell can be deformed, relative to the fully relaxed structures, to substitute for 22 Al atoms. The corresponding lattice vectors and lattice constants of the 22-atom Al are shown Table 1. The difference between the ideal monoclinic unit cell and the 22-atom Al unit cell uniquely determines the misfit strain tensor of the precipitate in the Al lattice, which will be used below to determine the corresponding elastic energy of precipitates in the matrix.

Composition	a [Å]	b [Å]	c [Å]	β [°]
Mg_5Si_6 (*DFT* [11])	15.14 (*15.13*)	4.08 (*4.07*)	6.93 (*6.91*)	109.9 (*110.1*)
$Mg_5Al_2Si_4$ (*DFT* [11])	15.33 (*15.32*)	4.05 (*4.07*)	6.84 (*6.77*)	106.0 (*105.8*)
$Mg_4Al_3Si_4$ (*DFT* [11])	15.13 (*15.11*)	4.12 (*4.13*)	6.65 (*6.60*)	106.6 (*106.4*)
Al matrix (*Exp.* [4])	14.63 (*15.16*)	4.06 (*4.05*)	6.41 (*6.74*)	105.3 (*105.3*)

Table 1: The fully-relaxed bulk lattice parameters of the β″ phases, compared with those that correspond to an ideal embedding within the Al matrix and, in brackets, with those already presented in literature (i.e. they come from a previous DFT study and from HR-TEM analysis .

We computed the elastic constants of all bulk phases by evaluating the stresses generated by small displacements of the unit cell around the equilibrium structure. A suitable set of displacements was used[2], and the stresses were then modelled as a linear function of the displacements to obtain the elastic constants [156]. The elastic constants for bulk Al and for the three β″ phases studied here are shown in Table 2, and were computed according to a reference system consistent

---

2 We used a set of 12 displacements of which six are not equivalent. These correspond to the three axial strains and the three shear strains.

[GPa]	C_{11}	C_{22}	C_{33}	C_{44}	C_{55}	C_{66}
Al	106.1 (114.3)			31.9 (31.6)		
Mg_5Si_6	98.4	84.6	88.0	21.9	29.1	51.2
$Mg_5Al_2Si_4$	107.1	94.7	99.1	26.9	36.3	49.4
$Mg_4Al_3Si_4$	106.7	96.5	97.1	25.9	35.6	46.3
C_{12}	C_{13}	C_{23}	C_{15}	C_{25}	C_{35}	C_{46}
55.9 (61.9)			0.			
50.0	47.7	45.7	8.2	5.8	5.4	-10.1
40.3	45.6	43.0	-13.1	4.3	11.9	5.4
46.5	48.0	48.8	9.3	5.7	9.3	6.3

Table 2: Elastic constants obtained by a linear fit of ab-initio stress tensors for small cell deformations. The values in parentheses are the experimental ones, extrapolated to 0K.[157]

with the Al matrix, as shown in Fig. 13. Our values are in good agreement with available experimental values [157] and previous computations [11, 158, 159].

In order to define a reference state for the thermodynamics of the precipitates we define the solid solution energies as

$$E_{Al}^{ss} = E_{Al_M}^{tot}/M \tag{68}$$

$$E_{x}^{ss} = E_{Al_{M-1}(x)}^{tot} - (M-1)E_{Al}^{ss}, \tag{69}$$

for $x = Si, Mg$. Here, $E_{Al_M}^{tot}$ and $E_{Al_{M-1}(x)}^{tot}$ are the total energies of a bulk-Al supercell containing M Al atoms and $(M-1)$ Al atoms and 1 atom of $x = Si, Mg$, respectively. The energy $E_{Al_{M-1}(x)}^{tot}$ is computed using a single solute in a 4x4x4 unit periodic cell with the cell volume held fixed. The cell develops a small pressure due to the misfit volume of the solute, but this contribution to the energy is negligible for the large cell size used.

The formation energy for a precipitate can then be defined as the total energy of a precipitate formula unit relative to that of the total energies of the precipitate atoms in the solid solution state. Thus, the formation energy is

$$E_{form} = \frac{1}{2}E_{\beta''}^{tot} - \sum_{x=Al,Si,Mg} n_x \cdot E_x^{ss}, \tag{70}$$

where $E_{\beta''}^{tot}$ is the (DFT) total energy of a fully-relaxed unit cell of the β'' phase containing 22 atoms (2 formula units), n_x is the number of atoms of element x in one formula unit, and E_x^{ss} is the energy of solute x in the (dilute) solid solution state. Knowing all the terms in eq. 70, we can compute the formation energies of the three proposed β''-phase compositions as shown in Table 3). The precipitates

	E_{form} [eV/f.u.]	γ [meV/Å^2] [mJ/m^2]			E_{strain} [meV/f.u.] (size: f.u./l.u.)			
					dilute	N : 1	4	16
					1 × 1	1 × 1	2 × 2	4 × 4
		A	B	C	96×96	5×5	7×7	12×12
Mg_5Si_6	-2.607	8.36	21.1	2.69	140	171	203	223
		134	*338*	*43.1*				
$Mg_5Al_2Si_4$	-2.769	11.8	23.5	9.11	128	161	198	223
		189	*376*	*146*				
$Mg_4Al_3Si_4$	-2.366	10.1	20.4	8.24	74	89	106	117
		162	*327*	*132*				

Table 3: Bulk, strain, and surface-energy terms computed for the three stoichiometries of the β'' phase which we considered in this study. The elastic strain energy E_{strain} has been computed for the dilute case and for the three periodic cases with varying numbers of formula units (N). The precipitate size is reported in formula units (f.u.) and the matrix size in fcc lattice unit cells (l.u.).

are strongly favorable, with negative formation energies in excess of -2eV/f.u., or greater than -0.2 eV/atom on average. Precipitate formation is thus thermodynamically highly preferable relative to the solid solution state.

3.3   IN-SITU PRECIPITATES

Bulk properties provide important information on the thermodynamic driving forces for precipitation, but are incomplete for understanding in-situ precipitation nucleation and growth. The system of precipitate plus matrix has additional energetic contributions from the precipitate/matrix interfaces, precipitate/matrix lattice and elastic constant mismatches that give rise to elastic energies when the precipitate is coherent, and precipitate/matrix edge and corner energies. All of these additional contributions determine the total thermodynamic driving force for precipitate growth as a function of precipitate size, shape, and density. While not addressed here, the elastic interactions between precipitates at finite densities also influences their spatial arrangement and orientation [160–162].

We thus need to predict the size, shape, and energy of a critical precipitate nucleus. At some critical precipitate size, the precipitate becomes thermodynamically unstable to further growth, i.e. increasing size leads to decreasing total energy. Below the critical precipitate size, the precipitate is unstable and should re-dissolve in the solid solution. Here, we take a model based on classical nucleation theory

(CNT) to assess the precipitate stability as a function of size, shape and density (which influences the elastic energy). In this analysis, we ignore edge and corner energies. Also assuming, for the moment, a low density of precipitates, the total energy of a precipitate containing N formula units, relative to the SSSS, can be written as

$$E_{prec}(N) = NE_{form} + NE_{strain} + E_{surf}(N). \tag{71}$$

There are two new terms in Eq. 71. First, there is the elastic strain energy E_{strain} due to the lattice and elastic mismatch between the precipitate and the Al-matrix per β″ formula unit for a single precipitate in an infinite matrix (the dilute limit). Second, there is the surface (interface) energy E_{surf} of the precipitate, which will depend on both the size and the shape of the nucleus. In order to evaluate the precipitation energy, we first obtain quantitative values for the strain and interface energies. Then, we will make predictions for the thermodynamics in the dilute limit. Finally, we will perform DFT studies of in-situ precipitates and compare the DFT energies versus the CNT model, adapted to the geometry of the DFT supercells.

3.3.1 Interface Energies for β″ Precipitates

Based on TEM analyses [7, 9, 163], and the correspondence between β″ structure and the closely-related 22-atom Al unit that accommodates one precipitate unit cell, we study three interface orientations as shown in Figure 12. The orientations are denoted $A \equiv (103)_{Al} \equiv (100)_{β''}$, $B \equiv (010)_{Al} \equiv (010)_{β''}$, and $C \equiv (\bar{3}02)_{Al} \equiv (001)_{β''}$. Given the relatively complex structure of the β″ phase, there are many possible ways to terminate the precipitate. Previous computational studies of the β″-Mg_5Si_6/α-Al interface have found that the associated surface energies can change significantly between different choices [164]. To compare with previous studies of finite-size precipitates, we chose the interfaces used in Ref. [11]. Figure 12 shows only one monoclinic unit cell of the precipitate and one for the matrix for $Mg_5Al_2Si_4$ but all three compositions were studied, and simulations were performed with much larger supercells of sizes 4β″ + 4Al unit cells for the A orientation, 6β″ + 6Al unit cells for the B orientation, and 6β″ + 6Al unit cells for the C orientation.

Since the precipitate and matrix have a structural mismatch, the total energy computed in a given simulation cell includes an elastic deformation energy. This energy must be computed independently and subtracted from the total energy obtained in the interface simulation to estimate the specific interface energy $\gamma_{\Lambda = A,B,C}$. First, we compute the energy per formula unit of the partially-relaxed β″ phase. For each interface orientation, we define $E_{\Lambda}^{\beta''}$ as the energy per formula unit of a β″ cell that is fully coherent with the Al matrix in the Λ

$[100]_{\beta''}$

$[001]_{\beta''}$

$[010]_{\beta''}$

$[001]_{\beta''}$

$[100]_{\beta''}$

$[203]_{Al}$

$[\bar{3}01]_{Al}$

$[010]_{Al}$

$[\bar{3}01]_{Al}$

$[100]_{Al}$

$[203]_{Al}$

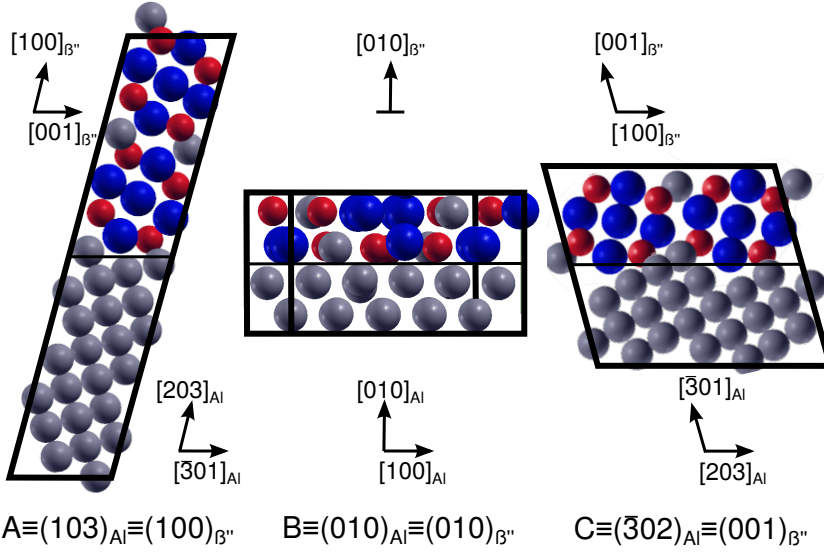$A \equiv (103)_{Al} \equiv (100)_{\beta''}$ $B \equiv (010)_{Al} \equiv (010)_{\beta''}$ $C \equiv (\bar{3}02)_{Al} \equiv (001)_{\beta''}$

Figure 12: We considered three orientations for the interfaces the between β'' precipitate and the Al matrix: the A orientation (left), the B orientation (center), and the C orientation (right). While we chose to represent only the $Mg_5Al_2Si_4$ phase, for simplicity, the other two compositions can be obtained by performing the substitutions indicated in Tab. 11.

plane, and relaxed in the orthogonal direction. We then prepared an interface between the Al matrix and the β'' phase, once again fixing the dimensions parallel to the interface to be fully coherent with the matrix, and relaxing it in the orthogonal direction. The interface energy can then be obtained from the total energy of this supercell E_Λ^{sc} as

$$\gamma_{\Lambda=A,B,C} = \frac{\left(E_\Lambda^{sc} - n_{Al}E_{Al}^{SS} - n_{\beta''}E_\Lambda^{\beta''}\right)}{2S_{supercell}^\Lambda}, \tag{72}$$

where $S_{supercell}^\Lambda$ is the cross-section of the simulation supercell corresponding to the orientation of the interface, n_{Al} is the number of Al atoms in the matrix, and $n_{\beta''}$ is the number of β'' formula units inside the supercell. The computed surface energies for each orientation are shown in Table 3.

As previously noted [164], the B surface energy is relatively large but the anisotropy is not sufficient to fully explain the observed needle-shaped habit of the precipitates. Given the large range of values observed for different terminations [164], a change in composition or some degree of interface reconstruction may significantly lower the energies of the A and C interfaces, leading to larger anisotropy. For instance, we obtain a considerably lower surface energy for the C interface in Mg_5Si_6 than any of the values reported in Ref. [164]. For

this specific case – that is associated with a relatively large mismatch in the unit cells between the β″ phase and the matrix – we observe significant relaxation of atoms at the interface, extending for several layers in the bulk, that was probably not captured fully in the smaller supercells[3] used in Ref. [164]. The issue of interface energies of β″ phases in Al thus merits further study.

3.3.2 *Elastic Strain Energies of Needle-like β″ Precipitates*

During the aging process, β″ precipitates show a strongly anisotropic habit, extending along the b ≡ [010] direction forming needle-like semi-coherent particles. The lattice mismatch between Al and β″ along the crystallographic b direction is also quite small. For this reason, two-dimensional slices along the a, c axes of the precipitate capture the main contributions to the energetics of large precipitates, and have already been studied to characterize both the energetics and elastic deformation of the matrix in this regime [164, 165]. To compute the magnitude of the elastic strain energy contribution for such a two-dimensional slice, we will use anisotropic continuum elasticity. The boundary value problem is formulated to correspond to the direct DFT studies below. We study a periodic two-dimensional plane-strain problem with a fully three-dimensional eigenstrain within the precipitate due to the misfit between the precipitate and the matrix. Figure 13 shows a schematic of the geometry with the relevant coordinate axes.

The Al matrix Ω_{matrix} is modeled as linearly elastic,

$$\sigma = C_{\mathrm{matrix}}\epsilon \quad \text{in } \Omega_{\mathrm{matrix}}, \tag{73}$$

where σ and ϵ are the Cauchy stress and strain tensors and C_{matrix} is the anisotropic fourth-order stiffness tensor of the matrix expressed in the global frame of reference $\hat{\bar{e}}_x$-$\hat{\bar{e}}_y$-$\hat{\bar{e}}_z$ aligned with the cubic lattice vectors of the pure aluminum matrix. The precipitate Ω_{prec} is also linearly elastic, but with an additional eigenstrain $\bar{\epsilon}$ relative to the reference Al lattice that accounts for the size and shape misfit of the precipitate,

$$\sigma = C_{\mathrm{prec}}(\epsilon - \bar{\epsilon}) \quad \text{in } \Omega_{\mathrm{prec}}. \tag{74}$$

Determination of the eigenstrain $\bar{\epsilon}$ and the rotation of the stiffness tensor C_{prec} into the global frame of reference are described in the 7.1.

---

3 Calculations in Ref. [164] used 44+44 atoms supercells, while our calculations for the C interface contained 132+132 atoms. We verified that when using a supercell with 66+66 atoms the surface energy for Mg_5Si_6(C) increased to 63 mJ/m², getting closer to previous results.
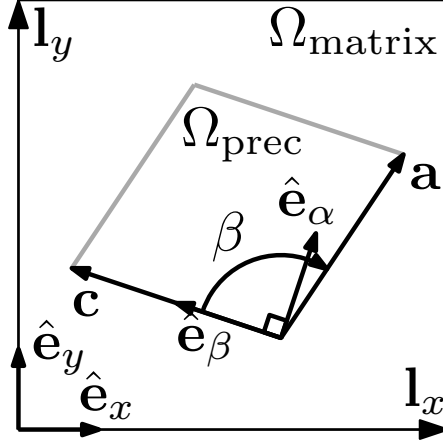
Figure 13: Schematic of the computational domain and definition of frames of reference. The directions of the vectors $\vec{c}$ and $\vec{a}$ are drawn as defined by (67), vectors $\vec{b}$ and $\hat{\vec{e}}_z$ point out of the sheet (not depicted). The global frame of reference is $\hat{\vec{e}}_x$-$\hat{\vec{e}}_y$-$\hat{\vec{e}}_z$ while the elastic constants listed in Table 2 are measured in the material frame $\hat{\vec{e}}_\alpha$-$\hat{\vec{e}}_\beta$-$\hat{\vec{e}}_z$.

As a plane-strain problem, there is zero out-of-plane displacement $u_z = 0$. Therefore the total strain tensor has $\epsilon_{xz} = \epsilon_{yz} = \epsilon_{zz} = 0$. The eigenstrain $\bar{\epsilon}$ retains these components, however, so that the effects of the mismatch in the z direction are included. We impose periodic Dirichlet boundary conditions on the displacement $\vec{u}$ in the horizontal and vertical directions

$$\vec{u}(\vec{x}) = \vec{u}(\vec{x} + n\vec{l}_x + m\vec{l}_y), \quad n, m \in \mathbb{Z}, \quad \forall \vec{x} \in \partial\Omega, \tag{75}$$

where $\vec{l}_x$ and $\vec{l}_y$ are the vectors linking the bottom left corner to the bottom right and the top left, respectively. We fix an arbitrary point $u(\vec{x}_p) = 0$ to exclude solid body motion. The static equilibrium stress and strain fields throughout the body are then determined by solving the standard equilibrium equation $\nabla \cdot \sigma = \vec{0}$. With the computed stess field, the strain fields are obtained from the constitutive models above and the elastic strain energy (per unit length in the out-of-plane direction) E_{strain} is then computed as

$$E_{strain} = \frac{1}{2}\left(\int_{\Omega_{matrix}} \epsilon C_{matrix}\epsilon \, d\Omega + \int_{\Omega_{prec}} \epsilon - \bar{\epsilon})C_{prec}(\epsilon - \bar{\epsilon}) \, d\Omega\right) \tag{76}$$

Note that the energy per unit length is independent of absolute model size and so the energy only depends on the size of the precipitate relative to the size of the computational cell, or equivalently on the area fraction (equal to the volume fraction) of the precipitate.

The boundary value problem is solved using the finite-element method (see 7.2). Note that, although the problem is nominally two-dimensional (plane-strain), the evaluation of the elastic strain energy remains fully three-dimensional due to the eigenstrain $\bar{\epsilon}$. Using the above implementation, we first computed the elastic strain energy per formula unit in the dilute limit where interactions among precipitates are negligible. This is done by using one formula unit in a cell of 96 x 96 fcc unit cells, and the results are shown as the "dilute" limit in Table 3. The elastic energies are small compared to the chemical energies, but are not small compared to differences in energies among precipitate compositions.

3.3.3   *In-situ Energetics of Dilute Needle-like β″ Precipitates*

Having evaluated separately the bulk, surface, and elastic relaxation energies for a needle-like precipitate of the β'' phases, we can then proceed to estimate the overall energetics of a nucleus. Assuming for simplicity the surface area of the interfaces to be that of the matrix-coherent unit cell (that is 26.02 Å^2 for each formula unit along the A facets, and 29.7 Å^2 for each formula unit along the C facets) we find that a needle-like precipitate with a cross-section of a single formula has already a negative formation energy. Considering the elastic energy associated with the infinite-dilution limit, one obtains $E_{prec} =$ -1.872 eV for 1 f.u. of Mg_5Si_6, E_{prec} = -1.486 eV/f.u. for $Mg_5Al_2Si_4$, and E_{prec} = -1.278 eV/f.u. for $Mg_4Al_3Si_4$. The formation of the β'' phases starting from the SSS is so exoenergetic that needle-like precipitates can form without overcoming a free energy barrier. Due to the much lower surface energy for the C interface, in the small-precipitate limit Mg_5Si_6 forms the most stable precipitate. In the limit of macroscopic precipitates, the energy per f.u. tends to the precipitation energy plus the dilute-limit elastic contribution, given as $E_\infty =$ -2.467 eV/f.u. for Mg_5Si_6, E_∞ = -2.651 eV/f.u. for $Mg_5Al_2Si_4$, and E_∞ = -2.292 eV/f.u. for $Mg_4Al_3Si_4$. Thus, $Mg_5Al_2Si_4$ is predicted to be the most stable form in the large-precipitate limit. The elastic strain energy does not change the order of stability but does narrow the energy difference between the most and least stable down to 0.35 eV/f.u. or 0.032 eV/atom

3.3.4   *DFT of Needle-shaped Precipitates and Comparison to CNT Model*

The CNT model of precipitate energetics we have introduced in Eq. 71, including self-consistent elasticity terms, could be very useful to examine the interaction between growing precipitates. In order to assess its accuracy, we use the same needle-like geometry to evaluate the energetics of precipitates using DFT, and perform a comparison

with the results of the model. To be consistent with the definition of formation energies used above, we define the precipitation energy using the SSSS as reference, i.e.

$$E_{prec}(N) = E_{sys}^{tot}(N) - M\,E_{Al}^{ss} - N \sum_{x=SiMgAl} n_x \cdot E_x^{ss}, \qquad (77)$$

where M is the number of Al atoms in the matrix for a give simulation supercell, and n_x and E_x^{ss} indicate the β'' composition and the solid-solution energy for Al, Si and Mg, as in Eq. (70).

To benchmark the model across different precipitate sizes, we study three systems whose cross-section contains 1, 4, and 16 formula units of precipitate in an equiaxed geometry. These precipitates are embedded in an Al matrix supercells of sizes ($a \times b \times c$) $5 \times 1 \times 5$, $7 \times 1 \times 7$, and $12 \times 1 \times 12$ *fcc* unit cells, respectively, as shown in Fig. 15(a) for the supercell containing 16 f.u. of the β'' phase).

As noted above, the elastic energy depends on the precipitate density or cell geometry. The DFT cells are not in the dilute limit. Therefore, for comparison to the DFT energies, the CNT model is modified to account for the elastic energy changes in the non-dilute limit as

$$E_{prec}(N) = N E_{form} + N E_{strain}(N, V) + E_{surf}(N), \qquad (78)$$

where $E_{strain}(N, V)$ is the elastic strain energy per formula unit in a supercell of volume V containing a precipitate of size N formula units. Elasticity calculations have been performed using the method described earlier for precisely the geometries studied in DFT, and the strain energies $E_{strain}(N, V)$ are shown in Table 3. These values are generally larger than the dilute limit, and increase with increasing N due the larger fraction of β'' precipitate included in the supercell.

Figure 14 compares the DFT precipitate energies, per formula unit, versus precipitate size with predictions obtained using (i) surface energy terms only (CNT(γ)) and (ii) surface energies terms plus elastic strain energy in the DFT simulation cell (CNT($\gamma + \epsilon$)). The results generally follow the expected trend, in that larger precipitates are thermodynamically more stable due to the reduction in relative importance of the interface, edge, and corner energies with increasing size, and the energies approach the (size-independent) formation energies plus dilute-limit elastic energies for each of the three stoichiometries (Table 3). A CNT model that uses only the surface energies captures qualitatively the asymptotic behavior for different β'' compositions, as well as relative ordering. However, it under-estimates the energy of the precipitates in the large-precipitate-size limit, due to the absence of the positive contribution of the elastic energy.

The CNT($\gamma + \epsilon$) model predicts quite accurately the energetics of the larger precipitates. However, it significantly overestimates the energy at the smaller sizes. The full DFT energies are up to 0.4eV/f.u
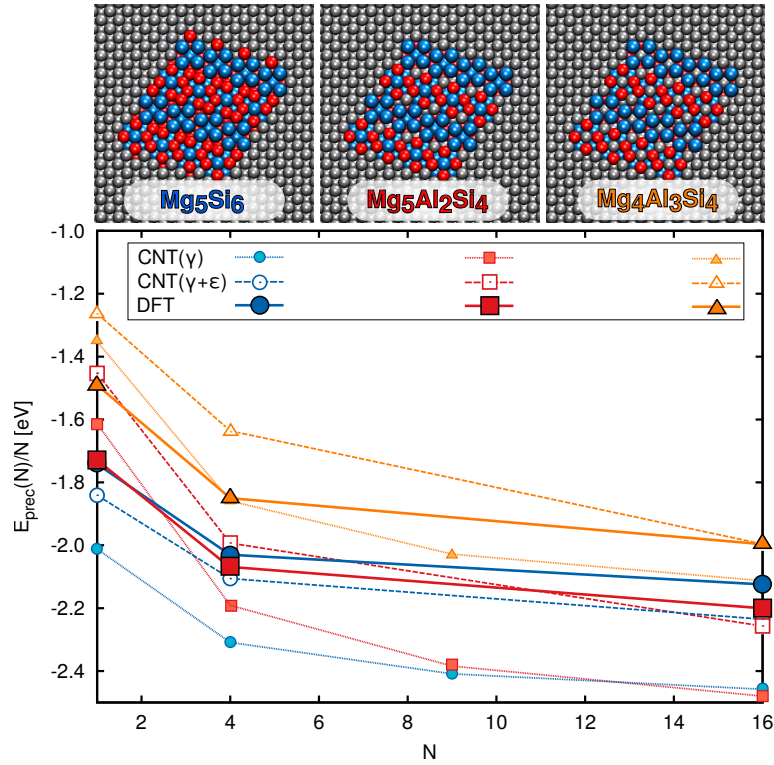
Figure 14: (Top) A view along the **b** direction of an infinite needle-shaped β″ precipitate with 16 (4x4) formula units cross-section, embedded in a $12 \times 1 \times 12$ Al supercell for the three precipitate compositions. (Bottom) Precipitation energies per formula unit, $E_{prec}(N)/N$, calculated from explicit DFT calculations using Eqn. 77 and estimated from the thermodynamic CNT model in Eq. 71.

lower than predicted by Eqn (78). One would normally expect that edge and corner terms would destabilize the nucleus (increase the energy) at the smaller sizes. Thus, the fact that the self-consistent energetics leads to stronger stabilization suggests that the surface energies computed assuming ideal interfaces provides only an upper-bound to the actual $\gamma^{A,B,C}$. Further relaxation (which is hindered for the larger precipitates, and for periodic surface calculations) could significantly lower the interface energy. Searching for reconstructions of the β″ ‖Al interfaces with a top-down approach and using electronic structure calculations constitutes a formidable challenge. We expect that the development of machine-learning models[21] for classical inter-atomic potentials, together with Monte Carlo sampling techniques, might help elucidate this important contribution to the stability and morphology of precipitates in the Al-6000 series.

Comparison between the calculations we report here and those presented in Ref. [11] underscores the importance of accounting for elastic relaxation in this kind of simulations. While part of the discrep-
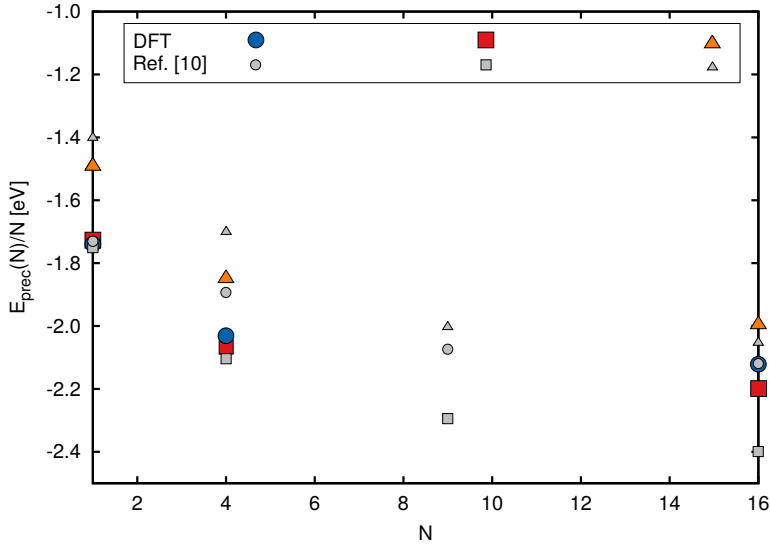
Figure 15: Precipitation energies per formula unit, $E_{prec}(N)/N$, calculated from explicit DFT calculations using Eqn. 77. Results from literature calculations that apply the same equation but include relaxation of the supercell are shown for comparison [165]. Symbols and colors are consistent with Fig. 14.

ancy could be attributed to minor differences in the computational details, we note a general trend where the energies for the 4×4 precipitates reported by Ref. [11] are considerably lower than those for the smaller 2×2 precipitates, values – and in all cases but for Mg_5Si_6 – lower than our values. As shown in the Appendix, this trend can be understood in terms of the boundary conditions chosen for DFT calculations. Simulations in Ref. [11] allowed the supercell dimensions to relax, which underestimates the energy of the encapsulated precipitate relative to the dilute limit. In our calculations, instead, we fixed the cell parameters to match the Al bulk lattice parameter, which, conversely, overestimates the energy. Use of a fixed supercell simplifies the comparison between calculations, and the definition of consistent surface energies. However, only a multi-scale analysis that includes a FE model makes it possible to compute the elastic corrections to the "dilute" limit and to interpret quantitatively DFT results in terms of the physical contributions to the precipitate energy.

3.4 NUCLEATION OF A PRECIPITATE IN 3D

The analyses in the previous section show that there is no barrier for the growth of needle-like precipitates starting at the smallest size $N = 1$ for the in-plane precipitate structure. The inclusion of interface and elastic energies was essential in this analysis to verify that

nanoscopic precipitates are stable despite the high interface and elastic energy contributions. We note that possible lower-energy interfaces will only enhance the stabilization of the smallest precipitates. Therefore, nucleation of all three β'' phases studied here occurs at the in-plane unit cell level or below. However, the in-plane analysis neglects the additional energy cost of the high-energy B $[010]_{\beta''}$ interface. We thus investigate here the formation energy of 3D precipitates, to better understand the precipitate nucleation process and possible nucleation barriers.

We simulated 3D precipitates composed of a single formula unit fully-embedded in the Al matrix. As shown in Table 4, the fully-relaxed DFT energy is negative for all compositions. This confirms that precipitation is barrierless down to a single 3D formula unit even when considering the high-γ B interfaces. At this scale, the CNT(γ) model is very inaccurate, predicting positive formation energy for all the stoichiometries except Mg_5Si_6. The elastic strain energy computation requires a full 3d analysis, and is not performed here since the elastic term would increase the energy relative to the CNT(γ) model.

It is not surprising that a mesoscopic model cannot capture the energetics of a precipitate that consists of just eleven atoms. It is however interesting that – just as for the needle-like geometry – the mesoscale model *overestimates* the energy cost associated with the precipitate-matrix interfaces, indicating that local relaxations can significantly lower the interface excess energy as compared to the ideal unreconstructed interfaces.
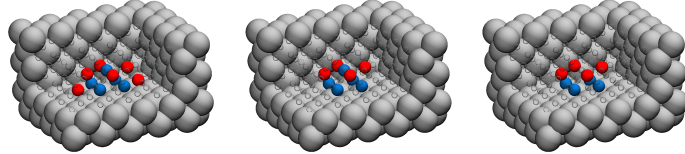
E_{prec} [meV]	Mg_5Si_6	$Mg_5Al_2Si_4$	$Mg_4Al_3Si_4$
DFT	-658	-558	-351
CNT(γ)	-102	513	497

Table 4: (Top) Snapshots of the simulation cells used to model single-formula-unit precipitates fully-encapsulated in the Al-matrix. (Table) Precipitation energies of one formula unit precipitates computed from DFT calculations (Eqn. (77)) and the CNT model (71) with and without the finite element strain term. The area of the different interfaces has been assumed to correspond to those of half of the monoclinic unit cell.

A NN POTENTIAL FOR AL-MG-SI ALLOYS

1

Atomistic simulations such as AI-MD, and *ab-initio* KMC are powerful tools for studying the early stage of clustering of solutes and the interaction between dislocations and solutes, clusters and precipitates. *Ab initio* methods, as we have seen, provide chemical accuracy for arbitrary atomic arrangements but are computationally prohibitive, with respect to both sizes and times, for addressing problems related to clustering, and strength. Thus, FF-like MD or even FF-like KMC methods must be used, but these methods require the existence of interatomic potentials for all the interactions among alloying elements, and must be accurate enough to provide realistic predictions. The development of accurate multi-component interatomic potentials has proven to be a serious challenge, and thus a significant impediment to the application of FF-like MD and FF-like KMC methods to alloys.

In spite of these challenges, many efforts to create interatomic potentials have been made. For pure metals and solid solution alloys, the Embedded Atom Method (EAM) potentials[166] have been widely used because of their reasonable accuracy and simple form. The Modified EAM (MEAM) [167] approach enables additional directional

---

1 This Chapter is adapted from Refs. [21] and [22].

bonding and so has been pursued for binary alloys that form intermetallic compounds[168, 169]. However, it is difficult for both EAM and MEAM methods to reproduce a wide variety of compound phases for alloys with more than two components. For instance, Jelinek *et al.* developed an MEAM potential with pair and triplet interaction parameters for the five elements Al, Mg, Si, Cu and Fe that reproduces several properties of binary compounds. However, generally these potentials do not provide "chemical accuracy" at the level needed ($\sim k_B T$) for realistic simulations at characteristic temperatures ($T \sim 300-600K$). For the particular Al-Mg-Si ternary system, the lattice constants and heats of formation of important precipitate phases are not well-predicted. Thus, there continues to be a great need for accurate multi-component interatomic potentials.

In this chapter, we develop a Neural-Network (NN) interatomic potential for the ternary Al-Mg-Si system that is useful for the study of precipitation processes and strengthening of the Al 6xxx alloys. In fact, as already mentioned in Sec. 2.3, it will be the engine of our KMC algorithm allowing a deeper understanding of the dynamics of the first precipitation sub-sequence, ie the one that leads to the formation of a single formula unit β''-phase. As shown in the Chap. 3, it also represents the smallest precipitates and the lower description limit of our CNT-type model.

The NN potential approach, developed by Behler and Parrinello[170], introduced in Sec. 2.2, is one of several classes of machine-learning potentials, such as the Gaussian Approximation Potential (GAP) [84, 171] and the linear regression potential[172]. These potentials introduce a large number of functions and parameters with no direct physical interpretation, rather than a few functions based on physical concepts, but the high flexibility then allows for the fitting of complex potential energy landscapes that govern the observed structures and the evolution of a material system.

In order to be efficient in the interpolation between reference structures, and to achieve some degree of transferability, these functions (that are the input of ML algorithms) should encode the physical features, and the mandatory symmetries of the problem, such as invariance with respect to rotations, translations, and permutations of identical atoms [117, 173–175]. To satisfy this requirement, many descriptors (e.g Symmetry Functions)[2] have been introduced that are able to characterize, to a various degree, atomic-scale systems, identify their similarities and differences, and form the basis for effective statistical learning schemes of energies and other properties [176–178].

---

2 We want to point out that in this dissertation we will use interchangeably the term descriptor, which represents a more general concept, with the term "symmetry functions", which indicates a sub-class of the environment descriptors. The methods, we will present, can also be applied to other types of descriptors, although we will present just the particular case of the SFs.

Therefore, the accuracy, speed and reliability of NN potentials, depends strongly on the way atomic configurations are represented, i.e. the choice of these descriptors/SFs. Choosing the set of parameters that characterizes the possible configurations that is simultaneously economic and thorough is one of the crucial steps in the optimization of ML schemes.

In Sec. 4.3, we will discuss automatic protocols to select a reduced number of descriptors out of a large pool of candidates, based on the correlations that are intrinsic to the training data. We will demonstrate this recipe in the particular case of the achievement of an accurate, fast and reliable NN potential for Al-Mg-Si alloy, although this procedure can be applied to any multi-component interatomic potentials. In the same section, we will discuss and compare the results obtained with those of Sec. 4.2, where the SF selection and the consequent achievement of NN potential are based on a combination of physical intuition and trial-and-error. In simple terms, we want to compare a protocol that is completely human supervised and time consuming with an automatic machine-driven one.

Finally, in Sec. 4.4 we show some open problems in building a NN potential associated with its strong dependence on the training DataBase (DB) and with the lack of description of long-range charge transfer. We will reveal these limitations by analysing the real effect of charge fluctuations for high symmetry systems containing point defects (for example, solute atoms in the Al-matrix). Since the NN technology for inter-atomic potential does not currently provide a unbiased description of electrically charged systems, we have adopted some solutions that act directly on the DFT training DB. These are based on having a coherent database of structures, where the effects of the charge fluctuations are reduced in order to minimize spurious terms in the interpolation stage.

4.1 SIMULATION METHODS AND GEOMETRIES

Energies and forces of all the structures used for training the network are calculated by DFT using the same computational framework shown in the Sec. 3.1. To estimate the uncertainty within the DFT scheme, some computations were repeated using the so-called PBEsol exchange-correlation functional[179][3].

Calculations using the NN potential and optimization of parameters in the NN potential are performed using the MD program, NAP[123]. Calculations using the MEAM potentials of both Jelinek's and Kim's[180] are performed using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS).[181]

---

3 It is based on a simple PBE with a final fit of the exchange-correlation energy to that of surface jellium

In calculations of pure bulk structures and ordered compounds, the minimum conventional cells are used with PBCs in all the directions. In finite temperature calculations of pure Al to compute the thermal expansion coefficient, we use $3 \times 3 \times 3$ cubic FCC cells and equilibrate the system to the target temperature using the Langevin thermostat[182] and the target pressure (0 GPa) using the Berendsen barostat[183] over a total time of 10 ps. Atomic volume at a given temperature is obtained as the average volume during 10 ps after the initial 10 ps equilibration time.

In calculations of single solutes/vacancies and solute-solute or solute-vacancy interactions, we use $4 \times 4 \times 4$ cubic FCC cells with one or two Al atom replaced by solutes as appropriate. In calculations of "random" distribution of solutes, $2 \times 2 \times 2$ cubic FCC cells are used and a half of the atoms (16) are chosen randomly and replaced with two vacancies and 14 solute atoms, Mg and Si. In both cases, PBCs are applied to all the directions and atom positions are relaxed with the lattice constants being fixed to that of pure Al. These are reference calculations of specified structures for input into the NN algorithm, and so there is no need to relax these structures fully.

In calculations of stacking faults, the simulation cell is oriented with axes along [1$\bar{1}$0], [11$\bar{2}$] and [111] as x, y and z directions, respectively. PBCs are used along x and y and the surfaces along z are free with a vacuum region wider than 11 Å. There are 12 atomic layers along z and the stacking fault is introduced by the rigid shift of upper half atoms in z followed by relaxation of all atoms only along the z direction.

In calculations of the properties for the (111), (001) and (110) surfaces, the energy versus separation is calculated by rigidly separating upper and lower halves of atoms. The slab was made thicker than 8 Å to reduce spurious interactions between surfaces. The cell size in the plane of the surface is fixed to the bulk equilibrium value.

In calculations of precipitates, the geometry of the system follows that employed in the Chap. 3 (also shown in Fig. 21).[20]

Reference data   Sample configurations used to train the network should contain structures that well represent the environments which are relevant for the phenomena of interest. To create the potential used for the study of precipitation strengthening of Al-Mg-Si alloys, we used the following base structures: (i) FCC, BCC and HCP structures of Al and Mg, diamond structure of Si; (ii) the Generalized Stacking Fault (GSF) structure along the {111} plane of Al; (iii) several surfaces of FCC Al; (iv) FCC Al structures containing one, two, and more vacancies or solute atoms; and (v) binary compound phases among Al, Mg and Si. From each of those base structures, we created deformed structures by changing the cell vectors of the system and displaced structures by random displacements of atoms or from MD snapshots. These

displaced structures are necessary additions to the equilibrium struc-
tures because otherwise the NN could give unphysical lower energies
for deformed structures relative to the true low-energy structures. In
total, we generated 10,237 structures and divided them randomly into
a training set used for the training and a test set used for monitoring
the convergence.

4.2   2-BODY NN POTENTIAL

In this section, we use a NN potential based on the implementation
by R. Kobayashi. This potential has a simpler form for the SF than
that of Behler and collaborators. Indeed, as discussed in Sec. 2.2.3,
the implementation in NAP represents the environment of the co-
ordinates with a family of functions that only consider correlations
between two objects (i.e. atoms). The use of a two-body schemes,
that does not consider correlations to three atoms, makes it possible
to reduce the number of free parameters and generate a reasonable
set of SF just based on physical intuition and trial-&-error validation.

The SF parameters of the NN potential were chosen as follows: The
number of symmetry functions for each pair is 20, so the total number
of symmetry function is 120 because there are six pairs among three
elements. For all the symmetry functions, $\eta_n = 10.0$ Å^{-2}. There are
20 R_n for each pair at regular intervals from 1.5 to 5.7 Å. We set the
cutoff radius $R_c = 5.8$ Å to cover 3rd neighbors interactions, which is
important to distinguish FCC and HCP structures and to reproduce
the generalized stacking fault (GSF) energy curve.

4.2.1   *Optimization of the Potential*

Figure 16 compares the energies of sample structures predicted by the
optimized NN potential against the reference values DFT. The Root
Mean Square Error (RMSE) for the training data is 0.5 meV/atom.
More importantly, the RMSE for the test data is 2.0 meV/atom. This
indicates that the present NN potential is able to accurately reproduce
all the structures in the entire sample data set, which is designed to
include many structures and configurations that are relevant for mod-
eling of evolution and mechanical performance in the Al-Mg-Si sys-
tem. Because the present NN potential has 120 symmetry functions,
3600 1st-layer weights and 30 2nd-layer weights per atom, the speed
of calculating forces of the NN potential implemented in the code,
NAP. is about 8x slower than that of the Al-Mg-Si MEAM potential
implemented in LAMMPS.[4]

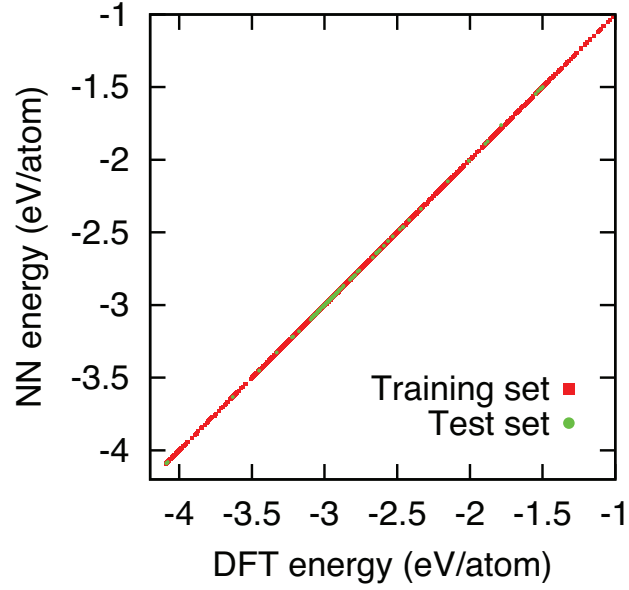4 The NN potential is implemented also in LAMMPS (it is available on request).

Figure 16: Energies per atom of sample structures obtained by the NN potential and the DFT calculation. The $y = x$ line indicates ideal matching between NN and DFT values.

4.2.2 *Properties of Aluminum*

There are a number of interatomic potentials for bulk Aluminum that reproduce most important properties with good accuracy. It is important that the NN potential give similar, or better, properties for the matrix Al material of the Al-Mg-Si alloy. Here, we thus present validation of the NN for pure Al.

Table 5 shows the bulk properties of FCC Al as obtained by DFT (the reference data), the Jelinek *et al.* MEAM potential (as an example), and the present NN potential. Although the present DFT value of cohesive energy is lower than that of experimental data, the lattice constant, bulk modulus, and other elastic moduli are in good agreement with experimental and previous *ab initio* values. The NN potential reproduces the DFT values of elastic properties with good accuracy.

The generalized stacking fault (GSF) energy is an important property relevant to the dissociation of a perfect dislocation into partial dislocations and to dislocation emission at a crack tip. The GSF energy of a shift vector (x, y) on {111} is defined as

$$\gamma^{\text{SF}}(x, y) = \frac{E[N](x, y) - N\varepsilon_{\text{Al}}}{A} \tag{79}$$

where N is the number of atoms, A the area of xy plane of the calculation cell, and ε_X the chemical potential of species-X, which is the cohesive energy of the most stable structure of the species. Table 5 shows

Table 5: Pure Al bulk properties from experiments or *ab initio* calculations and as computed using the present NN potential and two MEAM potentials (Jelinek and Kim). The structures relevant to these properties are in the traning data.

Al	Exp/*ab initio*	NN	Jelinek	Kim
a (Å)	4.05 [44], *4.06*	4.05	4.05	4.05
E_c (eV)	3.39 [44], *3.056*	3.057	3.353	3.360
B (GPa)	79.1 [44], *72.2*	76.9	78.4	79.0
C_{11} (GPa)	114.3 [44], *106.1* [[20]	109.9	111.1	113.8
C_{12} (GPa)	61.9 [44], *55.9* [20]	55.5	60.9	61.4
C_{44} (GPa)	31.6 [44], *31.9* [20]	31.6	28.6	31.4
γ_{sf} (mJ/m²)	135-166 [44], *122-164* [29,34-38], *125.8*	116.4	141.9	147.1
γ_{us} (mJ/m²)	*224* [184, 185], *166.2*	156.7	280.9	236.2
$\gamma_{(111)}$ (mJ/m²)	*710* [185],*720.2*,	742.9	716.1	516.3
$\gamma_{(001)}$ (mJ/m²)	*873.8*	878.0	1071.6	743.9
$\gamma_{(110)}$ (mJ/m²)	*927.7*	945.1	1104.8	820.9
α (10^{-6}/K)	23.6-25.4 [186]	23.3	14.4	

the stable and unstable stacking fault energies and Fig. 17 shows the GSF curve along the $\langle 112 \rangle$ direction. The NN potential slightly underestimates the *ab initio* results but reproduces the GSF curve well, especially in the range from 0.5 to 1.0 that includes the stable and unstable stacking fault energies.

The surface energy is defined as

$$\gamma^{surf}(d) = \frac{E[N](d) - N\varepsilon_{Al}}{2A} \tag{80}$$

where d is the separation of two surfaces. Figure 18 shows the energy versus rigid separation of two blocks of Al as computed by DFT, by two MEAM potentials (Jelinek and Kim), and by the NN potential. The fully-relaxed surface energies $\gamma_{(111)}$, $\gamma_{(001)}$ and $\gamma_{(110)}$ obtained from the largest d are shown in Table 5. Although the two MEAM potentials reproduce the order of stability of these surfaces, $\gamma_{(111)} < \gamma_{(001)} < \gamma_{110}$, the Jelinek potential has large unphysical barriers for separation in all the directions and, although they are much smaller, the Kim potential also has unphysical barriers. The NN potential reproduces the DFT curves along all directions and all distances. The NN potential thus accurately reproduces the surface energies but also forces during separation which are the cohesive tractions relevant in fracture processes.

We have also performed MD simulation at finite temperatures up to 500 K, which is higher than the normal paint bake temperature of ~ 450 K used during aging of Al 6xxx alloys, to confirm that the
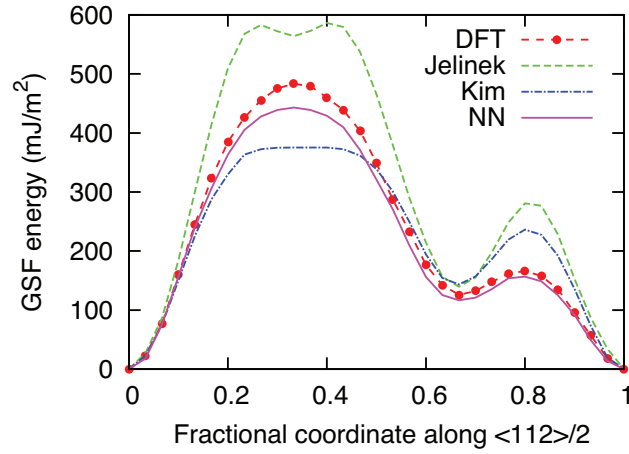
Figure 17: Generalized stacking fault curve for bulk Al along the [112] direction, as computed using DFT, two MEAM potentials, and the present NN potential. These structures are included in the training data.
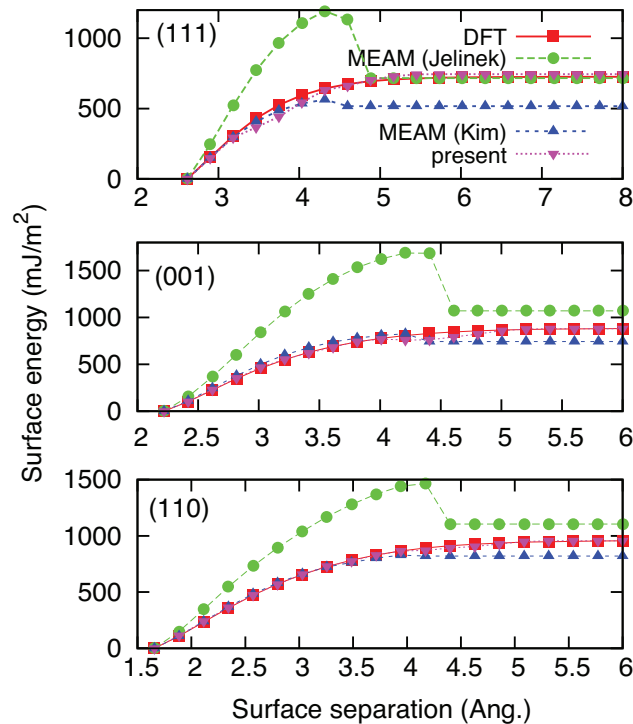

Figure 18: Energy versus separation for rigid block separation across the (111), (110), and (100) surfaces of bulk Al, as computed via DFT, two MEAM potentials, and the present NN potential. These structures are included in the training data.

present potential is suitable over the necessary temperature range and also to measure the thermal expansion coefficient. The thermal expansion coefficient is extracted from the slope of the volume-temperature relation as

$$\alpha = \frac{1}{a_0} \left[\frac{da}{dT} \right] \tag{81}$$

where a_0 is the equilibrium lattice constant at 0 K. The α value obtained using the NN potential, shown in Table 5, is in good agreement with experimental values. This indicates that the NN potential captures the anharmonicity around the equilibrium lattice constant as well as the harmonic region (related to the bulk modulus).

4.2.3 Binary and Ternary Al-Mg-Si Systems

We now turn to comparison of the predictions of the NN potential versus DFT for many properties of the Al-Mg-Si ternary system. Note that, although a lot of important structures are included in the training data set, it is impossible to include all the structures of interest because they require large number of atoms or configurations that are not easily calculated by DFT. We will mention in each subsection, figure and table if the structures are not in the training data.

The alloy phase with the lowest heat of formation ΔH is the one most likely to form at zero temperature. For accurate modeling of the evolution of the system toward precipitation of the thermodynamically favorable phases, the potential must reproduce the heats of formation of many possible alloy phases. The heat of formation of a compound is defined as

Ordered binary and ternary compounds

$$\Delta H^{comp} = \frac{E[\{N_X\}] - \sum_X N_X \varepsilon_X}{\sum_X N_X} \tag{82}$$

where N_X is the number of atoms of species X. Table 6 shows the lattice constants and heats of formation of binary compound phases, and Table 7 shows the equilibrium cell parameters, heats of formation, bulk moduli and elastic moduli of precipitate phases, as computed by DFT and as predicted by the Jelinek *et al.* MEAM potential and the present NN potential. Even for the complex MEAM potential that is calibrated to various binary phases, the predictions of other binary and ternary phases is challenging. In particular, the lattice constants and heats of formation of important Mg-Si phases such as Mg_2Si and Mg_5Si_6 are significantly different from the DFT values. The NN potential, on the other hand, reproduces well the DFT values of the equilibrium cell parameters, bulk moduli and heats of formation for the important precipitate phases such as Mg_5Si_6, $Al_2Mg_5Si_4$ and $Al_3Mg_4Si_4$ as shown in Table 7. Table 6 and 7 include some *ab-initio* values of

Table 6: Lattice constant a (Å) and heat of solutions ΔH^{comp} (meV/atom) for binary and ternary compounds calculated using DFT-PBE, the Jelinek *et al.* MEAM potential, and the present NN potential. *Ab-initio* values computed using DFT-PBEsol are indicated in parentheses. These structures are in the traning data.

Composition	Structure	*ab initio*		NN		Jelinek	
		a	ΔH^{comp}	a	ΔH^{comp}	a	ΔH^{comp}
AlMg	B1	5.745	424	5.760	425	5.714	236
AlMg	B2	3.388	66	3.396	66	3.440	-31
Al_3Mg	L1$_2$	4.138	4	4.145	0	4.238	-39
$AlMg_3$	L1$_2$	4.375	5	4.146	4	4.448	-46
$Al_{12}Mg_{17}$		10.506	-18	10.539	-18	10.315	389
AlSi	B1	5.217	261	5.229	261	5.241	280
AlSi	B2	3.160	239	3.164	239	3.188	142
Al_3Si	L1$_2$	3.994	96	4.000	96	4.059	113
$AlSi_3$	L1$_2$	3.899	333	3.907	333	4.194	532
MgSi	B1	5.505	384	5.532	383	5.507	192
MgSi	B2	3.308	144	3.313	144	3.384	64
Mg_3Si	L1$_2$	4.263	-7 (-4)	4.267	-8	4.367	24
$MgSi_3$	L1$_2$	3.988	269	3.988	269	4.176	322
Mg_2Si	C1	6.365	-136 (-105)	6.362	-136	6.530	44

heat of formation obtained using the PBEsol functional, from which we see that the difference between DFT with PBE and NN potential is smaller than the difference between DFT with PBE and PBEsol; i.e. the NN potential is within the accuracy of the DFT itself. The difference in lattice constants for Mg_5Si_6 and FCC Al for the NN potential is less than 1%, indicating that the precipitates of Mg_5Si_6 in Al matrix will be stable and remain coherent when using the NN potential. Some of the shear elastic constants of NN potential differ from those of DFT and could contribute the difference in formation energies of precipitates in Al matrix, which are evaluated in Sec. 4.2.4.

Single Solute and Vacancy Properties   Heats of solution of isolated solutes in FCC Al are important because they are relevant to the solid solution phase and thus to the stability of ordered phases relative to the solid solution state. The heat of formation of a vacancy is related to equilibrium and non-equilibrium vacancy concentrations and thus to vacancy-mediated diffusion of solutes and matrix atoms. The heat of solute or formation energy is computed as

$$\Delta H^{sol} = E[Al_{(N-1)}X_1] - (N-1)\varepsilon_{Al} - \varepsilon_X. \tag{83}$$

where X denotes the solute or vacancy, and we hereafter treat a vacancy as a "solute" with $\varepsilon_{Vac} = 0$. Table 8 shows the calculated heats of solution via DFT, the two MEAM potentials, and the present NN potential. The Kim MEAM potential gives quite good values for the heats of formation for Mg and vacancy, whereas the Jelinek MEAM potential is not good for these solutes. This shows the difficulty in

Table 7: Equilibrium cell parameters, heat of solutions ΔH^{comp} (meV/atom), bulk modulus B (GPa) and elastic constants C_{ij} (GPa) of precipitate structures obtained using the DFT-PBE, the present NN potential and Jelinek *et al.* MEAM potential. The geometry of the precipitate systems follow that employed by Ninive *et al.*[11] (also shown in Fig. 21). Heat of solutions computed using DFT-PBEsol is indicated in parentheses. Elastic constants C_{ij}'s by the *ab initio* calculation are taken from D. Giofré *et al.*[20] The structures relevant to these properties are included in the training data.

| | Mg_5Si_6 | | | $Al_2Mg_5Si_4$ | | $Al_3Mg_4Si_4$ | |
	ab initio	NN	Jelinek	*ab initio*	NN	*ab initio*	NN
a (Å)	15.138	15.173	17.012	15.299	15.343	15.095	15.109
b (Å)	4.081	4.074	4.322	4.054	4.052	4.122	4.131
c (Å)	6.982	6.940	7.142	6.818	6.846	6.633	6.651
α	90.0	90.0	89.2	90.0	90.0	90.0	90.0
β	110.4	109.9	89.6	105.9	106.0	106.5	106.6
γ	90.0	90.0	89.9	90.0	90.0	90.0	90.0
volume (Å^3)	400.4	403.2	525.0	405.7	408.9	395.0	397.7
ΔH^{comp} (meV/atom)	13 (24)	17	311	-71	-67	-44	-39
B (GPa)	62.1	58.5	125.0	61.2	62.5	63.8	64.4
C_{11} (GPa)	98.4	109.4		107.1	118.3	106.7	110.1
C_{22} (GPa)	84.6	94.0		94.7	98.6	96.5	102.6
C_{33} (GPa)	88.0	103.6		99.1	112.4	97.1	108.5
C_{44} (GPa)	21.9	29.8		26.9	27.4	25.9	30.3
C_{55} (GPa)	29.1	38.6		36.3	45.3	35.6	48.0
C_{66} (GPa)	51.2	68.1		49.4	60.8	46.3	52.7
C_{12} (GPa)	50.0	26.1		40.3	33.0	46.5	36.6
C_{13} (GPa)	47.7	46.2		45.6	60.0	48.0	55.6
C_{23} (GPa)	45.7	46.7		43.0	44.0	48.8	42.3
C_{15} (GPa)	8.2	-0.2		-13.1	3.6	9.3	5.9
C_{25} (GPa)	5.8	-2.5		4.3	9.5	5.7	7.0
C_{35} (GPa)	5.4	-1.6		11.9	-3.8	9.3	-6.2
C_{46} (GPa)	-10.1	0.8		5.4	2.6	6.3	4.8

constructing MEAM potentials for more than two elements. The NN potential predictions for the heats of solution for Mg and Si agree well with those computed by DFT.

Table 8 also shows the misfit volumes of Mg and Si calculated by DFT and the NN potential. The DFT misfit volumes are taken from Leyson *et al.*[135] and the NN potential misfit volumes are calculated using the same method. The NN potential underestimates the misfit volume of Mg by about 10 %. This is acceptable for reasonable estimates of the interaction energy of Mg with the Al dislocation or with precipitate-induced pressure fields. The misfit volume of Si is in very good agreement with the DFT result.

4.2.4 *Precipitates*

The interactions among solutes and vacancies is crucial to the early-stage formation of solute clusters during aging. Trapping of vacancies

Solute-Solute Binding and Energies of Solute Clusters

Table 8: Heats of formation ΔH^{sol} (eV) and misfit volumes ($\mathrm{\AA}^3$) (in parentheses) of a substitutional solute and vacancy in FCC Al. These values are obtained from $4 \times 4 \times 4$ cubic FCC cells that are not in the training data, but the relevant structures with $2 \times 2 \times 2$ cubic FCC cells are included in the training data.

Substitute	*ab initio*	NN	Jelinek	Kim
Mg	*0.090 (5.71)*	0.100 (5.15)	-0.200	0.098
Si	*0.375 (-2.65)*	0.376 (-2.63)	0.500	
Vacancy	*0.654*	0.647 (-3.39)	0.670	0.708

by clusters (so-called "vacancy prisons"[187]) has been suggested as important in Al-6xxx aging in particular. A useful potential must therefore reproduce these binding energies within approximately an energy of $k_B T$ to provide accurate metastable energetics during evolution of the system.

The binding energy between two solutes X and Y is defined as

$$
\begin{aligned}
-E_{bind}^{X-Y} &= E[Al_{(N-2)}X_1 Y_1] + N\varepsilon_{Al} \\
&\quad -E[Al_{(N-1)}X_1] - E[Al_{(N-1)}Y_1]
\end{aligned}
\tag{84}
$$

With this definition, a positive value indicates that X and Y tend to bind to each other. Figure 19(a) shows the binding energies of pairs among Mg and Si versus pair separation distance, and Figure 19(b) the binding energies that involve a vacancy, as computed by the DFT and as predicted by the NN potential. These values are obtained using $4 \times 4 \times 4$ cubic FCC cells and these structures are not in the training data, but the relevant structures with $2 \times 2 \times 2$ cells are in the training data. The figures show that the differences between the NN prediction and the DFT computation are less than 20 meV in all cases, which is slightly below $k_B T = 25.4$ meV at $T = 293$ K. The sign of the nearest-neighbor binding of Si-Si is the opposite of the DFT results, but the absolute value is small so that this difference is not of great consequence at the temperatures of interest. Of more importance are the trends of strong binding between Si-Vac and Mg-Si, and the repulsive interaction between vacancies, all of which are predicted well with the NN potential. These results indicates that the NN potential can be useful for meaningful simulations of clustering of solutes.

To move beyond simple pair interactions toward larger clusters such as those that might emerge during aging, we have also computed the heats of formation of rather complex random Al-Mg-Si-vacancy configurations via both DFT and the NN potential. The many specific configurations studied here (consisting of 16 Al atoms, 2 vacancies, and a total of 14 Mg and Si atoms) are not important on their
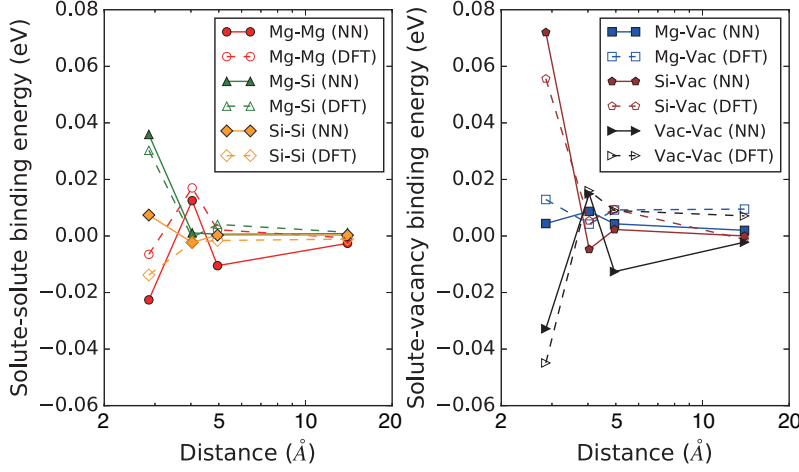
Figure 19: (a) Solute-solute binding energies as a function of distance between solutes, and (b) solute-vacancy binding energies, obtained by the NN potential (filled markers and solid lines) and DFT (open markers and broken lines). These values are obtained using $4 \times 4 \times 4$ cubic FCC cells which are not in the training data, but the relevant structures with $2 \times 2 \times 2$ cells are in the training data.

own, and hence we do not show all the structures but rather concentrate on the spectrum of energies. Figure 20 shows the difference in formation energies of all configurations as computed by the NN potential and by DFT. The formation energy is defined as

$$E_f = E[N_{Al}, N_{Mg}, N_{Si}] - N_{Al}\varepsilon_{Al} - N_{Mg}\varepsilon_{Mg}^{SS} - N_{Si}\varepsilon_{Si}^{SS} \tag{85}$$

where $\varepsilon_X^{SS} \equiv E[Al_{255}X_1] - 255\varepsilon_{Al}$ is the solid solution energy of solute X. Since these configurations are not included in the training data set, these results provide a measure of transferability, or conversely error prediction, for the present NN potential for solute clusters. The mean difference is +6 meV/atom with a standard deviation of +/- 8 meV/atom. Figure 20 also shows the differences between PBE and PBEsol DFT results, with a mean difference of +8 meV/atom and standard deviation of +/- 4 meV/atom. The NN potential predictions are thus somewhat larger than the differences among different DFT methods, but are almost within the statistical scatter. In contrast, the Jelinek MEAM potential predicts differences with DFT of mean +256 meV/atom and standard deviation +/- 108 meV/atom, which are not only 10-20 times larger, but far too large to be suitable for studies of aging.

In the early stages of precipitation in Al-Mg-Si alloys, the Mg/Si ratio in a needle-shape precipitate is smaller than that of the larger equilibrium Mg$_2$Si precipitate, which has the C1 structure. The structure of the early-stage precipitate is thought to be Mg$_5$Si$_6$[188] or the
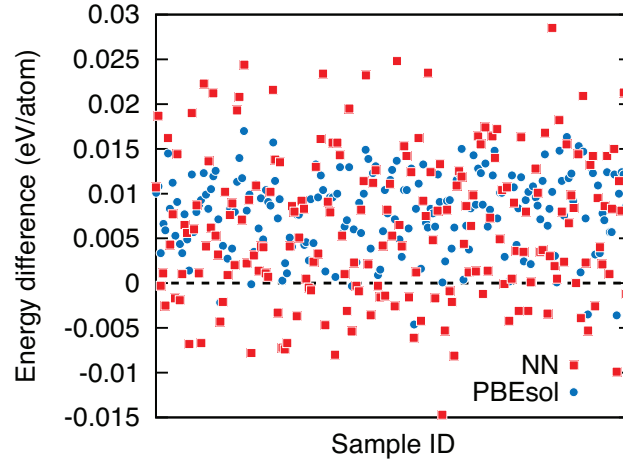
β″-phase

Figure 20: Formation energies per atom of the system described in Sec. 4.1 obtained by the NN potential (red squares) and the DFT-PBEsol (blue circles), as differences from the DFT-PBE results. Since these structures are not in the training data set, this can be treated as the prediction error of solution/precipitation energies by the NN potential.

same crystalline structure but with some Mg or Si atoms replaced by Al as shown in Fig. 21(a).[11, 20] We have thus computed the formation energies of these precipitates as a function of size when embedded in the Al matrix using the NN potential and have compared the predictions with DFT results of the previous Chapter. The simulation cells here for both DFT and NN are identical, making direct comparison possible. These structures are not in the training data set and thus the comparison shows prediction errors of the NN potential for precipitates.

The precipitate energy is defined as

$$E_{prec} = E_f/N_{FU} \tag{86}$$

where E_f is defined in Eq. 85 and N_{FU} the number of precipitate formula units in the precipitate, as shown in Fig. 21(a). Figure 21(c) shows the precipitate energies obtained by DFT and by the NN potential as a function of N_{FU}. The DFT and NN results are in good agreement for larger precipitates, but the NN potential underestimates the precipitate stability for the smallest precipitates (one formula unit), with a difference of ∼0.2-0.3 eV/unit. These energy differences are thus rather larger than those found for individual solute-solute interactions on a per-atom basis as the one precipitate unit contains 11 atoms. If we assume that not only precipitate atoms but also surrounding Al atoms, for example Al atoms interacting with the precipitate within the cut-off range of the NN potential, are contributing
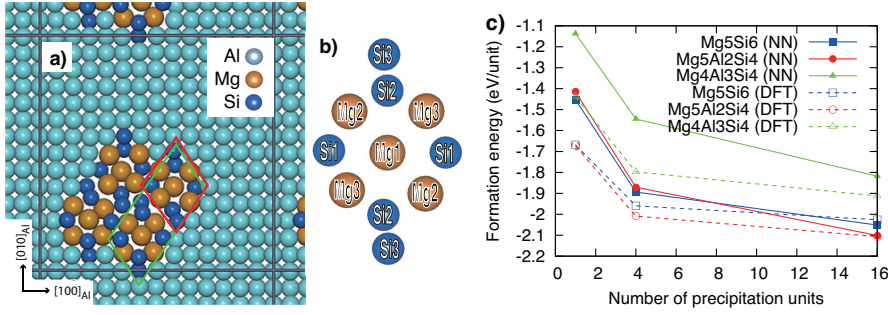
Figure 21: a) Typical simulation cell of Mg_5Si_6 precipitate of $N_{FU} = 4$ in Al matrix, which corresponds to $7 \times 7 \times 1$ FCC Al cell. The system is periodic in z-direction, $[001]_{Al}$. The green rectangle indicates a monoclinic cell of Mg_5Si_6. b) Formation unit of Mg_5Si_6 indicated in the red rectangle in a). In case of $Mg_5Al_2Si_4$, Si_3 sites in Mg_5Si_6 are replaced by Al, while Mg_1 site is also replaced by Al in $Mg_4Al_4Si_4$. c) Size dependency of the precipitation energy of Mg_5Si_6, $Al_2Mg_5Si_4$ and $Al_3Mg_4Si_4$ obtained by the NN potential and the DFT [20]) These structures are not in the training data set.

to the energy difference, it would decrease close to the uncertainty of the DFT method found for the 32-atom Al-Mg-Si-vacancy clusters.

Proceeding further, the contributions to the precipitation energy can be divided into four parts: the bulk formation energy, the strain energy, and contributions from interfaces and edges. The E_{prec}, which is the precipitation energy per N_{FU}, of bulk and strain parts are independent of N_{FU}, while the interface and edge energies should scale as $N_{FU}^{-1/2}$ and N_{FU}^{-1}, respectively. Therefore, as N_{FU} increases, the bulk and strain energies should dominate E_{prec} while the edge contribution could be significant at small sizes, e.g. $N_{FU} = 1$. The good accuracy of the bulk formation energy and elastic constants as shown in Table 5 and 7 the underestimation of the NN potential for the small precipitates with $N_{FU} = 1$ imply that the NN potential overestimates the formation energies of the precipitate edges and surfaces. Of course, one precipitate unit (11 atoms) is nearly entirely "interface" and "edges", and hence deviations per atom are comparable to the deviations for other solute-solute interaction energies. Nonetheless, the cumulative effect of these per-atom energy differences (reaching 0.2-0.3 eV/unit) can have an effect on overall aging behavior since the total difference in precipitate energy is not negligible. Improving the error of these energy differences is still challenging even for NN machine-learning potential and may not be necessary because the uncertainty of the DFT method could also result in non-negligible differences.

4.3   3-BODY NN AND AUTOMATIC FINGERPRINTS

In this section, we explore the fitting of a NN potential within Behler-Parinello framework for the same ternary system including also symmetry functions that describe 3-body interactions.[112, 121, 189, 190] The structure of neural network involves 2 hidden layers and a number of nodes for each hidden layer which ranges from 15 to 25. As evidenced from related studies, this NN structure is able to capture well the manifold features of DFT PES. [114, 117, 191–196] Given the presence of an enormous number of possible SF, choosing the set of descriptors that characterizes the possible configurations in an equally economic and thorough way is one of the crucial steps in the optimization of ML schemes. The idea of doing this descriptors selection by a non-automated method is, therefore, prohibitively time consuming and does not guarantee an optimal result.

The reason for embarking in an alternative training that includes 3-body SF are essentially two:

1. showing an explicit comparison with a purely 2-body potential and assessing the importance of 3-body terms

2. demonstrating an automatic protocol to select SFs for a complex ternary NN potential where the number of possible parameters is too high to proceed by trial&error validation

SF Database Generation

The SFs that are used in Behler-Parrinello NN potentials have been already introduced in detail in the Sec. 2.2.3. [112] The magnitude and variability of these symmetry functions depends on the relative distribution of one (for G_2, see Eq. 43) or two (for G_3, see Eq. 44) elements around the selected reference atom. The order of magnitude of a SF depends trivially on its spatial extent and on the concentration of the species involved. To treat them on more equal grounds, we decided to normalize the value of each symmetry functions based on the value it would take if it were computed for a uniform ideal gas,

$$
\begin{aligned}
I_2 &= 4\pi\rho_A \int d\, r_A\, r_A^2\, G_2(r_A) \\
I_3 &= 8\pi^2\rho_A\rho_B \int d\, r_A\, d\, r_B\, d\, \theta r_A^2\, r_B^2 \sin\theta\; G_3(r_A, r_B, \theta),
\end{aligned}
\tag{87}
$$

where ρ_A and ρ_B are the average densities for the element corresponding to the first and second neighbor considered in the evaluation of G_2 and G_3. We scale the symmetry functions by the square root of the integrals (87) so as to guarantee that the variance of the values would be constant in a uniform gas.

Given that the method we will show in Sec. 4.3.1 is based on the sparsification of a large set of these SF fingerprints, a first preparatory step involves the determination of a thorough yet manageable pool of candidate SFs. The generation is done spanning over all of the meaningful sets of parameters, using simple heuristic rules to represent most of the possible correlations within the cutoff distance. We generate two separate sets of radial symmetry functions, G_2. The first group $R^{(1)}$ is centered on the reference atom (i.e. $R_s = 0$) and the width varies as

$$\eta_m = \left(\frac{n^{m/n}}{r_c} \right)^2,$$ (88)

where n is the number of intervals in which we have chosen to divide the space and $m = \{0, 1, ..., n\}$. The second group $R^{(2)}$ is centered along the path between the central atom and its neighbours, at increasing distances following

$$R_{s,m} = \frac{r_c}{n^{m/n}},$$ (89)

while the Gaussian widths are chosen as

$$\eta_{s,m} = \frac{1}{(R_{s,n-m} - R_{s,n-m-1})^2}$$ (90)

in order to have narrow Gaussians close to the central atom and wider ones as the distances increases. This effectively creates a finer grid closer to the central atom, where small variations in the position have a larger effect on the potential (see Fig. 22).
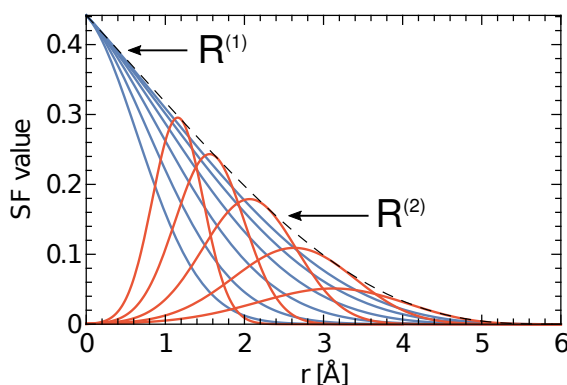

Figure 22: Examples of radial symmetry functions generated using $N = 5$ and $r_c = 6$ Å. The blue curves are the symmetry functions centered in the origin ($R_s = 0$) and η varying as in Eq. 88, while the red ones have their center shifted using R_s as described in Eq. 89 and η is described by Eq. 90. The black dashed curve is the cutoff function for $r_c = 6$ Å.

The G_3 symmetry functions were generated with a similar proce-
dure, choosing values for η according to Eq. 88, setting λ to both val-
ues $\{-1, 1\}$ that were originally proposed and choosing a few values
of ζ on a logarithmic scale, e.g. $\{1, 4, 16\}$.

By increasing the cutoff radius and the number N of symmetry
functions that are generated, one can make the description of the
environment more and more complete. This comes however at the
expense of greater computational costs, since a large number of SFs
would then have to be generated at each potential evaluation. Less
obviously, using too many, strongly correlated symmetry functions
could lead to overfitting and difficulties in the regression process. We
will discuss in Section 4.3.1 how to identify a small subset that con-
veys the essential structural information.

4.3.1   *SF Selection*

The SF DB discussed in the previous Section can easily lead to an
arbitrarily high-dimensional feature space. We can now turn to the
discussion of how one can select a N'-dimensional subset of the initial
descriptors, that captures the essential features of the atomic environ-
ments and can be used with little or no performance loss as the basis
of a statistical regression scheme to predict the properties of a given
set of materials. Symmetry functions can be selected, in a more or
less automatic fashion, by evaluating empirically the accuracy of a
ML model based on a trial set of SF. For instance, genetic algorithms
have been recently proposed as a method to generate an optimal se-
lection [197], similar to what has been done in Sec. 4.3.2 to select an
optimal set of reference structures. [198] Here, we focus on an unsu-
pervised approach that relies only on knowledge of the geometries of
the reference structures (e.g. SF description), without using informa-
tion on energy and forces, nor on the performance of the ML model
that results from a given choice of input features.

The approach we will discuss here is based on a relatively simple
idea: given a set of M structures $\{A_i\}$ that are representative of the
system one wants to study, and a large number N of fingerprints
$\{\Phi_j\}$, one can build the $M \times N$ matrix $\mathbf{X}$ such that $X_{ij} = \Phi_j(A_i)$.
The most effective fingerprints can then be chosen by using standard
linear algebra techniques to approximate $\mathbf{X}$. The ML schemes we
discuss in this work are based on a decomposition of the energy of
the system in local contributions, each of which is associated with
an atom-centered spherical environment with cutoff radius r_c. Un-
less otherwise specified, we will consider these environments, rather
than the entire structure, as the core of our discussion. The elements
of $\mathbf{X}$ refer to the fingerprints defining these environments, which we
consider, in order to simplify the notation, without explicit reference

to the structure they are part of. A given set of features can be used in a variety of regression schemes to predict atomic-scale properties, ranging from linear fits to Gaussian Process regression [84] (GPR) and neural networks [117, 199] (NN). It has been shown that in many cases the quality of the input representation plays a much more important role than the regression algorithms in determining the accuracy of predictions [200]. Here we will consider as example the NN scheme for an Al-Mg-Si alloy, using the two paradigmatic families of SFs described in Sec. 2.2.3.

We aim to find the optimal $M \times N'$ feature matrix $\mathbf{X}'$, where $N' \ll N$, that still provides a satisfactory representation of the space while reducing the computational load of the ML scheme. This is essentially a dimensionality reduction problem, that can be interpreted in terms of the construction of a low-rank approximation $\tilde{\mathbf{X}}$ of the feature matrix. Most of the dimensionality techniques available for this task, such as Singular Value Decomposition (SVD), generate new features that are a linear combination of the initial set and cannot be used for our current purpose, as they would still require the evaluation of all the N features and, only as a second step, project them onto a lower-dimensional space. We have therefore considered methods that strive to obtain a low-rank approximation of the feature matrix or its associated covariance using only rows and columns of $\mathbf{X}$. We discuss in particular the CUR decomposition approach.

CUR Decomposition

CUR decomposition [201] is a feature selection method that has been developed to deal with data where the information provided by the singular vectors cannot be properly interpreted, such as gene expression data. In analogy with the low-rank approximation obtained with a singular value decomposition, one writes

$$\mathbf{X} \approx \tilde{\mathbf{X}} = \mathbf{C}\,\mathbf{U}\,\mathbf{R} \tag{91}$$

where $\mathbf{C}$ and $\mathbf{R}$ are actual rows and columns of the original matrix and $\mathbf{U}$ is just a small dense matrix which satisfies the equation. The objective is still to find the best low-rank approximation to $\mathbf{X}$, but in this case only actual elements of the matrix are used, which implies that $\tilde{\mathbf{X}}$ can be obtained without having to compute all N fingerprints.

We discuss in particular the procedure for selecting a reduced number of columns (i.e. fingerprints), but the method can also be used to reduce the number of rows (i.e. reference structures) [171]. Each column c of the initial feature matrix is given an "importance score" calculated as

$$\pi_c = \sum_{j=1}^{k} (v_c^{(j)})^2, \tag{92}$$

where $v_c^{(j)}$ is the c-th coordinate of the j-th right singular vector, and k is the number of features that have yet to be selected and runs from N′ to 1. We also found that a very effective selection can be obtained by using a fixed number of singular vectors k = 1 at each iteration in the procedure (CUR(k = 1)). Not only this makes the method numerically more stable and significantly faster, but it makes the selection independent on the target number of symmetry functions, so that one can effectively perform a single selection with a large N′, obtaining a list of SF that is sorted from the most important to the least important. The importance score can also be weighted by a factor if one wants to prioritize the selection of a certain type of features, e.g. if the cost of evaluating different fingerprints varies greatly, and one would rather take several "cheap" fingerprints than a single "expensive" one.

Most CUR schemes employ a probabilistic criterion for feature selection, to guarantee e.g. that if several nearly-identical features are present, any of them will have approximately the same probability of being selected. To obtain a deterministic selection, we pick at each step the column with the highest score, and avoid selecting multiple nearly-identical features with an orthogonalization procedure. After having selected the l-th column with the highest importance score, every remaining column in $\mathbf{X}$ is orthogonalized relative to it

$$X_j \leftarrow X_j - X_l \, (X_l \cdot X_j) / |X_l|^2 \,. \tag{93}$$

The SVD is then re-computed based on the orthogonalized matrix, and the column weights are re-evaluated. The procedure is iterated until all N′ features have been chosen to build the $\mathbf{C}$ matrix, that corresponds to the reduced feature matrix $\mathbf{X}'$. Since in this application we are only interested in reducing the number of fingerprints, $\mathbf{R} = \mathbf{X}$, and we can compute $\mathbf{U} = \mathbf{C}^+ \mathbf{X} \mathbf{X}^+$, where $\mathbf{A}^+$ indicates the pseudoinverse. One can then compute the accuracy of the approximation as

$$\epsilon = \|\mathbf{X} - \mathbf{CUR}\|_F \, / \, \|\mathbf{X}\|_F \tag{94}$$

The total number of features to be selected, can either be fixed a priori, or increased until ϵ becomes smaller than a prescribed threshold.

4.3.2 *Farthest Point Sampling of Train Structures*

In this section, we discuss a way to determine the most representative selection of train structures from a DB of initial structures. This can be useful to reduce the cost of evaluating a ridge regression model, or to minimize the number of property evaluations that need to be performed in order to train the model [171, 202, 203]. In order to do so, it is useful to construct a set of fingerprints associated with the

whole structure, rather than with individual atomic environments. A straightforward definition of a "global" fingerprint associated with a structure A, $\bar{\Phi}(A)$ is the average of all the local fingerprints for the environments that compose the structure A, i.e.

$$\bar{\Phi}_j(A) = \sum_{\mathcal{X}_k \in A} \Phi_j(\mathcal{X}_k)/N_{at}(A). \tag{95}$$

In the case of Behler-Parrinello symmetry functions, that are defined separately for each chemical species, we consider that the global fingerprint is composed by concatenating sections corresponding to each element. In other terms, one can see this as a sparse representation for a larger fingerprint vector that is padded with zeros in all sections but the relevant one, even though in a practical NN implementation one only computes symmetry functions associated with the identity of the central atom. The fingerprint vector for the entire structure can then be built according to (95), summing these zero-padded vectors over all atoms in the structure.

In this context, we select the structures that are most diverse within an initial DB using a Farthest-Point Sampling (FPS) approach which seems to be the most promising in line with other systems.[22] This is analogous to the strategy with which one can select uniformly-spaced reference points (see e.g. [204]), but here we apply it to the set of vectors $\{\bar{\Phi}_j(A)\}$, so as to select the train structures that are as diverse as possible for the DB being investigated. In a FPS scheme, successive points are chosen so as to maximize the Euclidean distance between them. After arbitrarily selecting the first structure, each subsequent one is chosen as

$$k = \mathrm{argmax}(\min_j |\Phi_j(A) - \Phi_k(A)|), \tag{96}$$

where j runs over all of the structures that have already been selected. The procedure is repeated until the desired number of structures has been chosen.

So, from the above diagram, one can use the FPS method (or even the CUR method) to sparsify the train reference DB. The reference DB we used as a starting point is composed of the 10551 structures shown in Sec. 4.2, supplemented by 609 structures of β''-phase precipitates and interfaces shown in Chapt. 3. [20, 21] Given that many of the resulting 11160 structures are taken from short MD runs and are highly correlated, we selected 2000 structures with FPS, that have been used both for the selection of SF and the training/testing procedure. This sparser selection leads to a larger absolute magnitude of the fit error, but does not affect the quality of the fit, while making the optimization procedure faster and more stable.

4.3.3 *Optimization of the Potential*

A NN potential for Al-Si-Mg alloys including only 2-body functions has been already shown in Sec. 4.2 [21], where all of the interactions among the different species and defects must be accounted for to obtain accurate predictions across the full range of relevant compositions. The presence of multiple interactions at different length scales makes the manual selection of SF a particularly cumbersome task. In the Sec. 4.2, the problem was circumvented by restricting the SF pool to the 2-body G_2 components, making it possible to obtain a systematic - if not optimal - selection. [21] The automatic selection procedure we introduce in this section makes it much easier to automatically determine an efficient feature set that includes both G_2 and G_3 SFs, which makes it possible to take into account the angular dependence of the atomic interactions explicitly.

We follow the protocol discussed in Sec. 4.3 for the initial generation of SF with random initial weights and a 3:1 random split of train:test points. Six sets of G_2 SF have been generated using $N = 4, 12$ and $r_c = 8, 16, 20$ bohr, and two sets of G_3 SF have been generated using $N = 8$ - one with $r_c = 8$ bohr and $\zeta = 1, 2, 4, 8, 16$ and the other with $r_c = 12$ bohr and $\zeta = 1, 2, 4$. Final results are not sensitive to these choices, that we only made to have intermediate files of manageable size. Duplicate SF have been eliminated, together with those that had a width smaller than 1.06 Å for the radial ones and smaller than 1.32 Å for the angular ones. We weighted the importance scores (92) by a factor proportional to $\rho_A \rho_B r_c^3$ for G_2 functions between atoms A and B and $\rho_A \rho_B \rho_C r_c^6$ for G_3 functions between atoms A, B and C, to reflect the cost associated with evaluating them. We note that these importance scores do not enter the functional form of the SFs finally used in the fit.

The details of the fingerprints can be inferred from the input files provided in the supporting information of Ref. [22], and the performance of the resulting NN potentials can be seen in Table 9. The test set RMSE decreases systematically as the number of selected SF increases, up to 64 SFs per species. We also compared the results with those obtained with the SF selection from Sec. 4.2. To ensure a fair comparison we re-optimized and tested the potential using the RuNNer [119] software and the same FPS selection we discuss above. We verified that the accuracy of the re-trained NN for the properties we test here is comparable of better than that of the original potential. [21] Already at $N' = 96$ (32 SFs per species) the automatic selection that includes 3-body SFs leads to a better test set error than the systematic selection of 120 G_2 SFs.

While the test set RMSE is a good measure of the quality of a potential, it is important to also verify the stability of the NN when com-

$N'_{Al}, N'_{Mg}, N'_{Si}$	$\epsilon_{Al}, \epsilon_{Mg}, \epsilon_{Si}$ $\times 10^4$	RMSE(E) [meV/at.]	RMSE(f) [eV/Å]
CUR selection			
16,16,16	79,99,101	16.22	0.084
32,32,32	7.9,14,10	4.08	0.052
64,64,64	0.9,1.3,0.8	2.47	0.022
SFs of Ref. [21]			
40,40,40	-	9.2	0.069

Table 9: The table reports, for different numbers of SF, the error in the approximation of the feature matrix, and the RMSE for energies and forces from a test set. Results from the SF used in Sec. 4.2.1 and in Ref. [21] are also shown for comparison.

puting a property for which configurations had not been explicitly included in the train DB. As an example of the behavior of the different potentials, that is very relevant for the potential application of this NN in the description of the early stages of precipitation in Al-6xxx alloys [20], we computed the configuration energy along the minimum energy pathways for the vacancy-assisted migration of Al, Si, Mg atoms in a matrix of 256 Al atoms. Atomic configurations along the pathway between the minimum energy states were obtained by linear interpolation, and by local optimization using the NEB method [205] with the climbing image algorithm [206] as implemented in Quantum ESPRESSO [49]. The details of the DFT calculations were the same as described in Refs. [20, 21]. 7 images have been used for Mg and 13 have been used for Al and Si, and lead to relaxed vacancy migration barriers that are consistent with previous DFT calculations [207]. Keeping the configurations fixed, we computed the energy along the migration barrier for both the linear transition path between the initial and final configurations and the corresponding relaxed positions.

As shown in Figure 23, there is a considerable improvement in the quality of the fit when going from 16 to 32 SFs per species, whereas the improvement is less dramatic when using a larger number of SFs, and actually in the case of the vacancy-assisted diffusion of Si the 64-SF NN performs worse than the 32-SF NN. This observation underscores the fact that refining the SF selection does systematically improve the accuracy in the interpolative regime, as probed by cross-validation, but not necessarily to a systematic improvement in the extrapolative regime. For all of the vacancy-assisted diffusion processes we considered, however, NN potentials reproduce the correct qualitative behavior. Excluding the case with 16 SF per element, which is clearly insufficient for this system, the error in the relaxed barrier
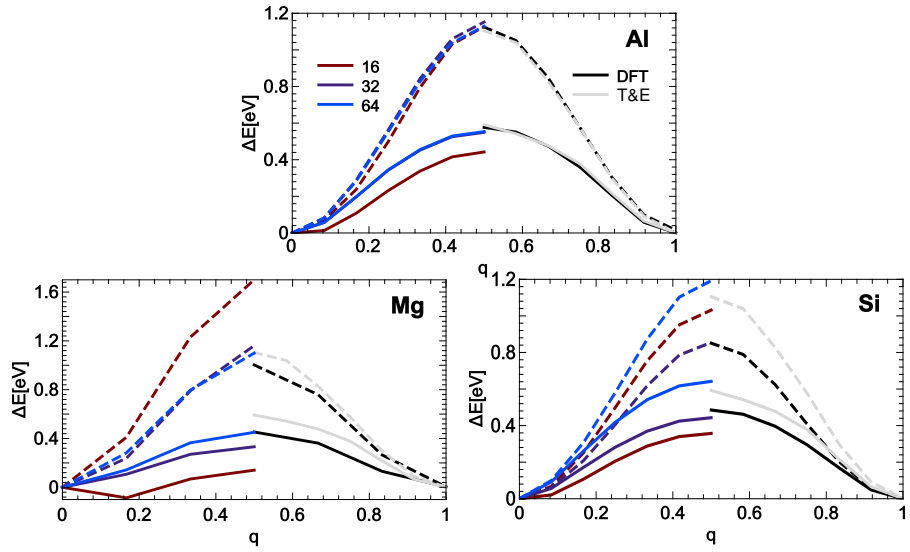
Figure 23: The energy barrier for the vacancy-assisted migration of Al, Mg, and Si using an increasing number of symmetry functions are presented on the left, compared to DFT and the choice of SF for the Trail&Error protocol in Sec. 4.2.1, presented on the right. Dashed lines correspond to the unrelaxed configurations, solid lines to the minimum energy pathway. The energies are shown as a difference from the minimum energy structure.

is below 0.1 eV, which is comparable to the typical DFT error. Automatic SF selections that include 3-body terms as reported in Fig. 23 with a blue, brown and violet color, perform better in the case of 32 SF than the 2-body NN discussed in Sec 4.2 (shown with a grey color) in case of Silicon, and slightly better in the case of 64 SF in the case of Aluminium and Magnesium than the same 2-body NN. Nevertheless, the 2-body NN predicts diffusion barriers with a remarkably small error.

4.3.4 *Precipitates*

In this section, we will benchmark the NN for the properties linked to the phenomenon of precipitation that is central to this dissertation. For an exhaustive list of comparisons of other properties obtained from NN potentials generated through the two protocols see App. 7.3. We can however summarize App. 7.3 by stating that the potential obtained through an automatic protocol has a slightly better accuracy in the prediction mode and a worse accuracy on the training set. Suggesting that the 2-body potential suffers from a degree of over-fitting. Here, all the quantities determined use a 3-body NN potential with 64 SFs for each atoms (e.g. 192 SFs as a total).
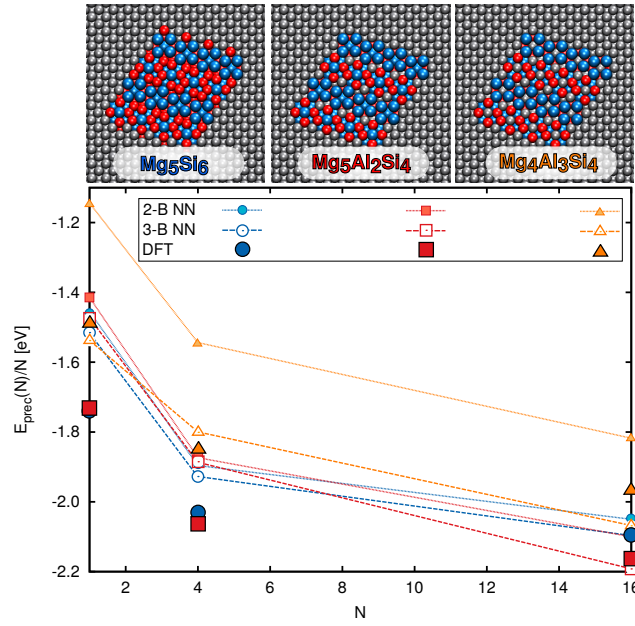
Figure 24: Top) A view along the **b** direction of an infinite needle-shaped β'' precipitate with 16 (4x4) formula units cross-section, embedded in a $12 \times 1 \times 12$ Al supercell for the three precipitate compositions. (Bottom) Precipitation energies per formula unit, $E_{prec}(N)/N$, calculated from explicit DFT calculations (as in the previous Chapter) and estimated by the 2-body and 3-body NN Potentials.

As in Sec. 4.2.4, we have computed the precipitation energy of β''-phase as a function of size. We can see in Fig. 24 that both protocols underestimate the precipitation energies for the small precipitates with $N_{FU} = 1$. This implies that even the 3-body NN potential overestimates the edges energies of precipitates. The edge contribution scales as N_{FU}^{-1} and it could be prominent at 1 FU. On the other hand, the 3-body NN potential seems to describe the interface energies much better. This is apparent in the results for 4 FU. Indeed, the surface terms scale as $N_{FU}^{-1/2}$ so that they predominate over the edge energies at intermediate precipitates. This better accuracy of interface terms is not only due to the fitting protocol; in fact in this later fit we included in the train set structures that contain $\beta''//$Al interfaces (i.e. the ones shown in Fig. 12). This shows how the extensibility of NN potentials in the introduction of new knowledge is the real strength of this method compared to other fitting methods.

As reported in Sec. 4.2.4, we calculate the solute-solute binding energies, and we use the standard definition of binding energies[208]. Therefore, the values of the binding energies in Fig. 25(Right) are of opposite sign w.r.t. those in Fig. 19. From what we have understood, the most evident errors between DFT and NN Potential appear to be

β″-phase

Solute-Solute Binding and Energies of Solute Clusters

related to interactions between solutes, which is troubling given the importance of these terms during the early stages of precipitation.As already discussed, an appropriate potential should reproduce these binding energies within approximately an energy of k_BT to provide accurate metastable energetics during KMC evolution of the system. Both the 2B and 3B potentials, while being much better than EAMs, barely meet this target. This is surprising given that the structures are part of the training set.

An important observation, which we will discuss in detail in Sec. 4.4, is that the interaction energy between solutes depends strongly on the size of the supercell, more than would be expected based on misfit volume and elastic terms. We traced the source of this discrepancy to long-range charge oscillations induced by the solute atom. These kinds of long-range effects cannot be captured by a short-range NN potential, which partially explains the difficulty in accurately learning solute interaction energies.

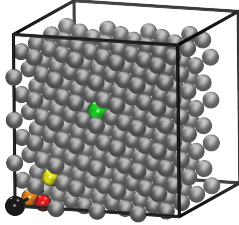## 4.4   PARTIAL CHARGES AND SHORTCOMINGS OF THE NN

Long-range phenomena are problematic in Machine Learning algorithms since they are strongly dependent on system sizes and exceed the range of the environments described by the symmetry functions. [209–212] A preliminary study on electrically charged systems and long-range interactions conducted through a Behler-Parrinello NN Potential proved to be strongly charge definition dependent. [213, 214] It allowed one train the long-range term of NN Potential by combing the numerical accuracy of NN with an electrostatic term based on Hirshfeld charges.[215]

Both versions of the potentials we built do not include any information about the charge density during the fitting mode which is reasonable given that metallic screening should eliminate any long range effect.

In the next sections, we will demonstrate the existence of long-distance charge fluctuations in aluminium alloys, and discuss possible remedies to reduce the impact of these long range electronic effects. In short, our aim will be to avoid the problem of electrostatic long-range terms by smoothing out these charge fluctuations until they have disappeared.

4.4.1   *System-size Dependence of Defect Interaction Energies*

In Secs. 4.2 and 4.3, we have seen that an Al-Mg-Si NN Potential built with 2B interactions and 3B interactions is able to predict several mechanical and thermodynamic properties of the aluminium ternary system (6xxx-series).

Substitute 1	Substitute 2	*ab initio*	2-B NN	3-B NN
Mg	—	*90*	100	92
Si	—	*375*	376	376
Vacancy	—	*654*	647	673
Mg	Mg	*8 (35)*	23	63 (49)
	Mg	*-16 (-12)*	-12	-3 (-9)
	Si	*-30 (-38)*	-38	-29 (-50)
	Si	*1 (-1)*	2	-1 (-14)
	Vacancy	*-12 (2)*	-2	12 (0)
	Vacancy	*-3 (-16)*	-13	-26 (-41)
Si	Si	*14 (16)*	-17	49
	Si	*3 (9)*	2	58
	Vacancy	*-55 (-54)*	-76	-74 (-63)
	Vacancy	*-5 (2)*	9	-26 (-27)
Vacancy	Vacancy	*46 (78)*	36	54
	Vacancy	*-15 (-22)*	-19	30

Figure 25: **Left:** We show a 4x4x4 Al-supercell (e.g 256 Al-atoms) where we substitute Al atoms with three impurities (Mg, Si, and Vacancy). The substitute 1 is indicated by a black sphere, while the substitute 2 is positioned based on the color in four different relative positions with respect to the substitute 1 and is indicated by a coloured sphere. The grey spheres are Al-atoms. **Right:** Heats of formation of a substitutional solute and the solute-solute binding energies both expressed in meV units. The text color for substitute 1 and 2 is correlated to their positions in the image on the left. The terms in brackets have been calculated considering an Al-matrix with 32 Al-atoms (e.g. Al supercell with 2x2x2 size). These structures are in the training data for the 3-B NN and are not in the training data for the 2-B NN.

Solute-solute interaction energies are the aspect that showed most pronounced inconsistencies in the prediction mode for both NN potentials. Since these are fundamental to simulate the dynamics of the early stage of precipitation, their potential divergences have to be understood. This section will be dedicated to an in-depth study of the reasons behind this inconsistency. The initial idea that these energetic discrepancies could be associated to the inaccurate training of NN was definitely excluded by analysing the *ab-initio* simulations of different-sized supercells, which have the same arrangements of solutes at different concentrations in Fig. 25(Right).

We can note that the solute-solute binding energies calculated using a cell of 256 atoms (e.g. 4x4x4 Al-matrix) and a cell of 32 atoms (e.g. 2x2x2 Al-matrix) differ significantly and their energetic discrepancies are too large to be attributed to a simple misfit volume effect.[216] This kind of discrepancies is worrying, and can affect negatively the accuracy of the Neural Network. In order to understand the origin of these discrepancies, we firstly checked carefully all possible DFT convergence parameters. In particular, we focused on the size dependent parameters such as the number of Kohn-Sham bands and the k-points sampling. We discovered that the number of KS bands defined as default by QE was not sufficient to fully represent unoccupied states due to considerable smearing (See Sec. 2.1.3). This affected the energy of the most excited states and, therefore, the calculation of the Fermi level. For these reasons, we have increased the number of KS bands equal to the total number of electrons.

The other computational parameter that needs to be reassessed according to the system size is the the *k*-points grid. The comparison between two systems having the same supercell should always be done using the same *k*-point set, so that we can make vanish plausible errors from a non-converged *k*-point grid. A similar strategy can also be applied when we compare structures with different unit cells (e.g. we adopt the equivalent k-point sampling).[217] In addition, we use 8 times denser sampling to ensure complete independence of the total energy from the choice of the *k*-point set. Adopting this last precaution to the whole train set has a positive effect on the stability of the optimizer during the interpolation and on the predictive accuracy of the potential NN too.[218] However, such discrepancies remained even taking these measures into account as shown in Tab. 10.

In the literature[219–222], similar results are traceable and strongly correlated with the presence of long-range electrostatic terms. When we incorporate the substitutes into the 2x2x2 Al-supercell and in the 4x4x4 Al-supercell, the interaction is so delocalized that, besides interacting with each others at long distances, they feel the effects of their periodic images.

This is visible in the first two columns of Tab. 10 in which the binding energies do not tend to zero for solutes relatively far apart and are very different for systems with different solute concentration but identical placing. Eventually, we traced the problem to long-range charge fluctuations, that we analyse in the following.

4.4.2 *Which Atomic Charge Definition?*

An unambiguous unbiased definition of the (first principles) atomic charges has not been established yet.[214, 223] Their quantification should be determined by many-atoms electronic density which, by definition, does not belong to any atom. For these reasons, different ways of dividing the space or the charge density lead to different numbers for the atomic charges and therefore possibly to different ways of interpreting the chemical bond.[214] Then the question is of course which region of space should be attributed to a certain atom and where we should use diffuse or discrete boundaries to define the integration regions (i.e. overlapping versus non-overlapping regions).[223]

The oldest and also the best known definition of atomic charges is based on the Mulliken population analysis. [224–226] The Mulliken population of (atom-centered) basis functions is obtained by assigning half of each total overlap population between two atoms at each atom. In this way, it does not take into account the difference in electronegativity between the two atoms. Furthermore, this approach does not, in principle, converge with increasing basis set size. Therefore, unrealistic atomic charges can result. [227–230] A possible improvement to this approach is the Natural Population Analysis (NPA)[228] Both methods are based on a representation of the molecular wave function using atomic basis functions.

A second class of methods, which is based directly on the electron density as a function of space, includes among those: Hirshfeld method [215], Bader approach[231] and Voronoi Population (VP) analysis[223, 232]. These methods rely on the direct spatial integration of the electron density function over an atomic region. They do not project the wave function on any form of basis functions, and are more representative of our problem due to electronic charge density fluctuations. We will evaluate the atomic charges using the electronic charge density and two very different ways of approaching the problem of space partitioning: Bader approach and VP analysis. Both approaches can be used for periodic or molecular systems.

The Bader approach uses the topology of the electron density to define an atomic domain. It starts with locating the Bond Critical Points (BCPs), which are the points of minimum charge density (zero gradient) along a bond path and a maximum in the normal direction

to the bond. The BCPs are used to determine the zero-flux surfaces, which then divide the atomic system into atomic regions. Integrating the electronic density ρ_e within an atomic region we obtain the charge assigned to that atom as follows

$$Q_i^{Bader} = \int_{\Lambda} \rho_e(\mathbf{r}) d\mathbf{r}. \tag{97}$$

In Eq. 97, Λ is the region bounded by BCPs for the i-th atom. It is also clear that atomic volumes obtained in this way are not spherical and do not overlap.

For the VP analysis, we use a generalization of the Politzer method based on Voronoi polyhedra.[232–234] The polyhedra cells are defined following Voronoi process: we construct a plane at the bond midpoint perpendicular to the line connecting two atoms. We repeat this process for all atoms and we get a set of planes assigned to each atom that defines a polyhedron around it, the so-called Voronoi polyhedron. In the standard process the perpendicular planes are put at the bond midpoint. This overestimates the electronic charge transfer between different kinds of atoms.[214] For this reason, we use a modified Voronoi Population analysis[223] which is more reliable. It consists in moving the perpendicular bisecting plane away from the Bond Midpoint (BM) by a distance related to the relative sizes of the atoms (e.g. "van der Waals radii") as in Eq. 98.

$$\mathbf{R}_{\perp}^{ij} = \frac{\chi_i}{\chi_i + \chi_j}\mathbf{r}_{ij} \, , \tag{98}$$

where χ_i and χ_j are the "van der Waals radii" of atom i and j and $\mathbf{r}_{ij}$ is their internuclear separation. We can, then, write the VP charges in the following way

$$Q_i^{VP} = \int_{\Omega} \rho_e(\mathbf{r}) d\mathbf{r} \tag{99}$$

in which Ω is the volume of i-th Voronoi cell.

Both Eq. 97 and Eq. 99 can be normalized subtracting the corresponding valence atomic charge Z_i.

4.4.3   Bader Versus Voronoi

In this section we compare the two techniques presented in the previous section. For the calculation of the electronic density $\rho(\mathbf{r})$ we will use the same DFT computational parameters expressed in Sec. 3.1. In particular, we apply an energy cutoff of 280 Ry to electronic charge density which implies a mesh density of $5 \cdot 10^{-11} \text{Å}^{-3}$. This value makes the calculation of atomic charges for both methods converge within a tolerance of ± 0.01 electron per atom[5]. We use the "Bader

---

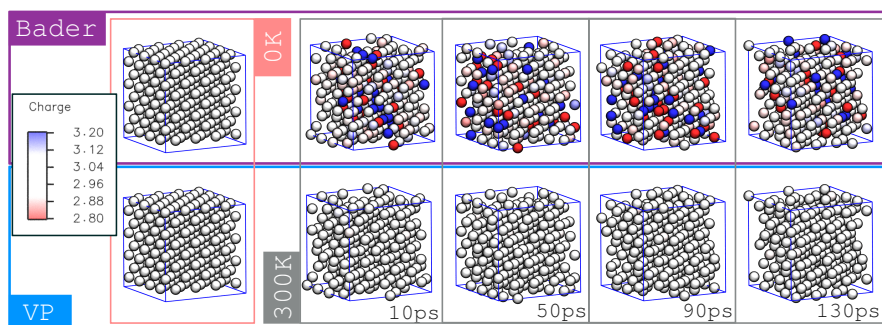5 This value has been confirmed in the case of Al-bulk at 300 K.

Figure 26: In 4x4x4 Al-bulk, we show atomic charges calcolated by Bader approach and by VP analysis. They have been analysed at zero temperature and for 4 frames at 300K after a time of 10ps, 50ps, 90ps and 130ps. The dynamics were driven by a velocity rescaling thermostat [238] and our NN potential, while the electronic charge density was evaluated by a DFT approach.

code" in order to make the charge analysis consistent.[235–237] We consider as test for both methods the electronic charge density associated to a fully relaxed 4x4x4 Al-bulk at 0 K and a 4 configurations of Al-bulk dynamics at 300 K (See Fig. 26).

It is clear from Fig. 26 that the Bader construction is very sensitive to even moderate levels of disorder in Al bulk, assigning large formal charges to different atoms that, given the metallic nature of the bond in bulk Al, would be expected to be almost neutral. For this reason, in the following we use VP charges that give a more physical description for this system.

Through a similar process we have calculated the charge densities and, then, the atomic charges of the configurations utilised to determine the solute-solute binding energies and the heats of formation of a substitutional solute in Fig. 25. The atomic charge analysis by (both Bader and) VP approach confirmed a strong charge oscillation due to a charge "point"-defects into a high-symmetry perturbed system.

Having established a reliable method for defining atom charges in this system, and leaving aside the fact that in different systems the opposite conclusion has been reached that Bader charges are more "physically meaningful" [239–243], we can now move on to discuss the effect of lattice defects. As shown in Fig. 27, vacancy-solute clusters generate large atomic charges in their vicinity (that are even more pronounced when studied in terms of Bader charges), that differ markedly depending on the size of the simulation cell. These charge fluctuations are very long-range, can be seen as analogue of Friedel that are introduced by the presence of a surface in a homogeneous electron gas. [244, 245] We believe that these effects are at the
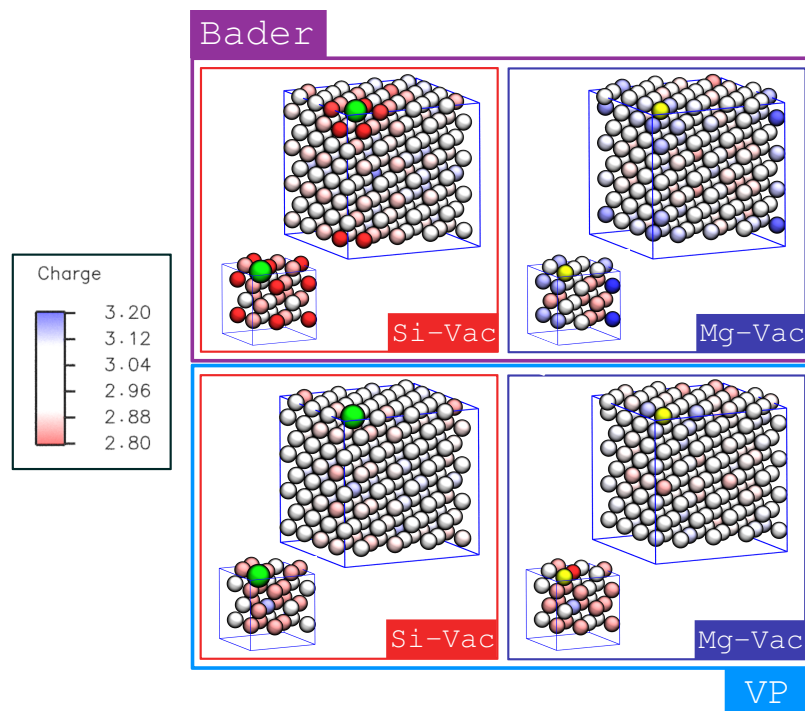
Figure 27: Atomic charges were calculated by Bader approach and by VP analysis in the cases of a **Mg**-**Vac** system and the **Si**-**Vac** one in 2x2x2 and 4x4x4 Al-supercell. In the previous sentence the substitutes were carefully marked and coloured so that it is possible to appreciate their relative positions in Fig. 25(Left).

origin of the large errors in the description of solute-solute interactions, and of the discrepancies between cells of different size. While this kind of electronic effects cannot included in a binding short-range NN and in a straightforward manner, one to reduce their impact on the stability and reliability of the fit which will be the subject of the following section.

4.4.4   *Solutions*

In this section, we will discuss possible solutions in order to reduce these charge fluctuations. We will use for our analysis the more reliable VP analysis, but similar conclusions could be reached based on a Bader analysis (see App. 7.4).

Consistently with our interpretation of these charge fluctuations as coherent phenomena related to Friedel oscillations, these effects are very sensitive to disorder. As shown in Figure 28, small random displacements of the atoms, or finite-temperature fluctuations reduce or eliminate almost completely the charge inhomogeneity. This suggests that the charge fluctuations, although physically real, might be less relevant in finite-temperature conditions. However, highly-symmetric minimum-energy geometries are important for building and testing the potential, and so it is not possible to just ignore the problem.

Another possibility to reduce the impact of the charge fluctuations, that is also effective in the case of ordered minimum-energy configurations, involves increasing the electronic temperature. Just as Friedel oscillations, that are largely a consequence of the sharp Fermi edge at T=0 K, get smeared for finite electronic temperature, the fluctuations of the charge around the solutes are considerably reduced as the electronic smearing is increased (see Fig. 29). As shown in Table 10, interactions energies between defects are better behaved when using an electronic temperature of 0.2Ry. Interactions decay to zero for far-away defects, and the discrepancy between small and large simulation cell are sizeably reduced, except for the case of solute-vacancy interactions that still show a discrepancy larger than 10meV. More tests are needed to determine if the error introduced by such a large electronic temperature is preferable to the inconsistencies due to charge fluctuations.
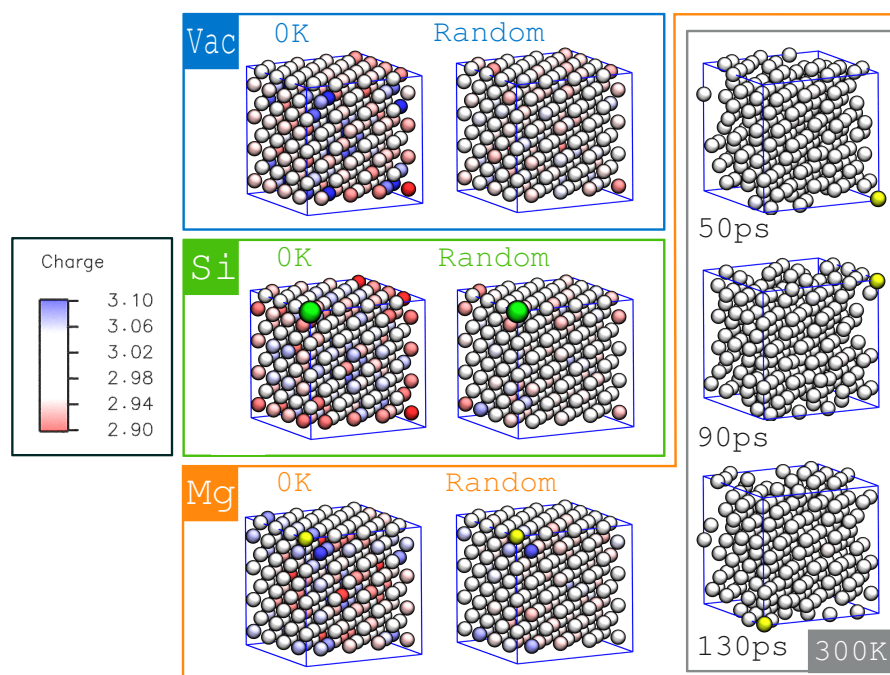
Figure 28: We show atomic charges determined by VP analysis for the solute atom systems (e.g one Al atom into 4x4x4 Al-matrix is substituted with a Si, Vac and Mg atom). As the case of Al-bulk, the atomic charges have been analysed at zero temperature and, only, in the case of Mg-solute system at 300K. Moreover, for all 3 solute atom systems the label "Random" indicates that the atomic positions are randomly displaced from the equilibrium positions for a factor ±0.01Å. We observe that the charge fluctuations decrease when we introduce an entropic component into the atomic positions.
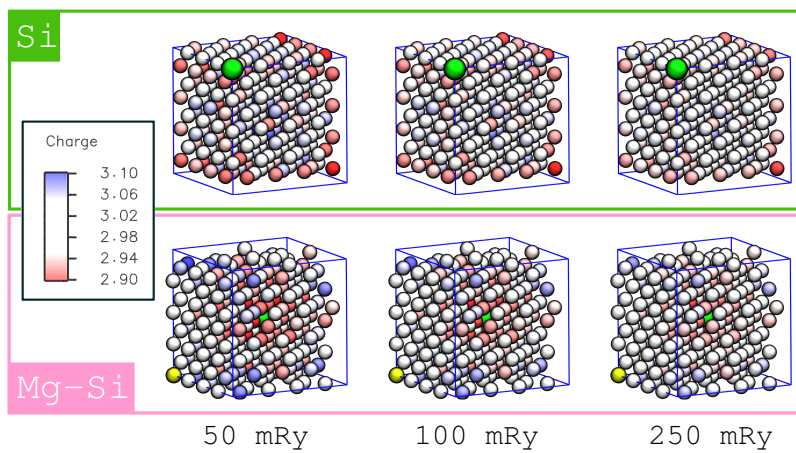
Figure 29: We show the atomic charges determined by VP analysis for the Si-solute atom system and **Mg**-**Si** system into 4x4x4 Al-matrix varying the smearing from 0.05 Ry to 0.25 Ry. In the previous sentence the substitutes were carefully marked and coloured so that it is possible to appreciate their relative positions in Fig. 25(Left). It is clear that the charge fluctuations decrease when we raise the electronic temperature.

Table 10: First-principles solute-solute binding energies (in meV) calculated for the various combinations of substitues and in their various relative positions represented by four different colours as shown in the Fig. 25(Left): in 4x4x4 Al-matrix with 256 Al-atoms/in 2x2x2 Al-matrix with 32 atoms. The values in the "standard" column were obtained using the computational framework discussed in Sec. 3.1 where atoms have been required to relax in a super-cell with the geometry of the Al-matrix.In spite of this framework, the energy effects of the misfit volume are not sufficient to explain these discrepancies. Meanwhile, the values in the "converged" column were obtained by adopting other two further precautions discussed in Sec. 4.4.1, i.e. paying attention to the convergence of the KS bands and of the *k*-point sampling. Finally, the values indicated in the "0.2 Ry" column were obtained by simulations with the previous measures and using an electronic smearing of 0.2 Ry.

Substitute 1	Substitute 2	Standard	Converged	0.2 Ry
Mg	Mg	8 / 35	7 / 34	3 / 18
	Mg	-16 / -12	-17 / -13	-15 / -13
	Mg	-1 / –	-1 / –	-1 / –
	Mg	-2 / –	-2 / –	0 / –
Mg	Si	-30 / -38	-30 / -40	-26 / -30
	Si	1 / -1	1 / -2	2 / 2
	Si	-4 / –	-4 / –	-3 / –
	Si	-1 / –	-1 / –	0 / –
Mg	Vacancy	-12 / 2	-13 / 1	-16 / -22
	Vacancy	-3 / -16	-4 / -17	-8 / -13
	Vacancy	-8 / –	-7 / –	-6 / –
	Vacancy	-8 / –	-8 / –	-6 / –
Si	Si	14 / 16	14 / 15	17 / 14
	Si	3 / 9	2 / 7	2 / 6
	Si	3 / –	2 / –	0 / –
	Si	3 / –	2 / –	0 / –
Si	Vacancy	-55 / -54	-56 / -56	-35 / -45
	Vacancy	-5 / 2	-5 / 1	-9 / 0
	Vacancy	-9 / –	-8 / –	-5 / –
	Vacancy	2 / –	-8 / –	-4 / –
Vacancy	Vacancy	46 / 78	45 / 77	31 / 49
	Vacancy	-15 / -22	-16 / -23	-4 / -17
	Vacancy	-4 / –	-5 / –	-5 / –
	Vacancy	-6 / –	-7 / –	-3 / –

5

PRE-β''-PHASE PRECIPITATION AND KINETIC MONTE CARLO

In the previous chapter we dealt in detail with all the construction of an accurate and robust Al-Mg-Si NN potential. The results have confirmed its adaptability and accuracy suggesting it as an effective strategy to avoid expensive first principles techniques. In this Chapter we exploit its accuracy combined with its low cost to describe diffusion-controlled phenomena. In fact, the use of KMC methods driven by the NN potential make it possible to achieve this at the atomic level.[246–248] The KMC methods, as discussed in chapter 2, are suitable to study a wide variety of materials up to experimentally relevant length and time scales [249–252], shedding light on the resulting (micro and) nano-structural and (micro- and) nano-chemical evolution during operational conditions, e.g ageing.[253–256]

Clustering during natural ageing in Al–Mg–Si alloys is a low temperature reaction that is driven by diffusion and we use *ab-initio* KMC (AI-KMC) simulation to understand its kinetics. The simulation is based on the diffusion of a single vacancy, characterised by both diffusion activation and atomic interaction energies.[257]

Even if KMC approaches are in principle well suited for the investigation of the underlying atomic-scale mechanisms, the task, in this case, is complicated by the chemical complexity of the alloy. Specifically, in AI-KMC, the evolution of the alloy proceeds through migration events of single defects[257–259], which are stochastically selected at each step based on their transition rates.

In the specific case of Al-Mg-Si alloys, KMC model has to take into account that all proposed crystal structure of the β'' phase or post β'' phase cannot be simply projected on the sites of Al lattice (namely we still have an issue of commensurability). The solution we adopt consists in determining possible states and transition based on a lattice model. However, we recalculate on the fly the minimum local energy associated with each lattice state so as to update the vector of the rates preserving the detailed balance principle (see Sec. 2.3.2).

A similar vacancy-diffusion model has previously been developed for the same ternary system. [257] However, it relies on a rigid lattice, and considers only DFT pair interactions between first nearest-neighbours (calculated in the case of isolated solute atoms) in order to estimate the rates. Our previous analysis of the importance of elastic relaxation in the early stages of precipitation, and the analysis of pair

interactions that extend well beyond the first neighbour, both suggest the need to go beyond this over-simplified model.

The cost of this minimization procedure is only affordable thanks to the use of a NN potential. At the same time, this raises the question of whether the NN is transferable enough to describe the multitude of pre-nucleation clusters observed in a typical KMC run. For this reason, after having discussed the details of our KMC implementation in Sec. 5.1 and presented some qualitative preliminary results in Sec. 5.2, we will re-assess in Sec. 5.3 the quality of the NN by testing directly on selected structures from these preliminary runs. This check demonstrates the need to extend the training dataset to include representative configuration from the KMC run, suggesting an iterative strategy to refine the simulations.

5.1   ON-LATTICE ALGORITHM FOR ALMOST OFF-LATTICE SYSTEMS

The general theory behind a KMC simulation has been presented in Sec. 2.3. Here, we show how this can be applied specifically to the study of aluminium alloys. Our KMC simulation algorithm carries out elementary swaps on a virtual 3D grid with PBC applied. This grid is on-lattice and follows the Al-sites. It can represent an idealized picture of vacancy migrations and precipitation sequence for Al-alloy with Mg- and Si-trace elements. The defects as Mg- and Si-solute atoms and Vacancies, are artificially and randomly introduced into the system to represent an initial state that corresponds to a Supersaturated Solid Solution (SSS). Since we are looking to the vacancy mediated diffusion in a Face-Centered Cubic (FCC) crystal in a fixed volume, which evolve according to a finite set of physical swaps, the number of configurations available at any KMC step is finite and enumerable. The so-called configuration space is discrete. Furthermore, at each KMC step, we can determine all of the potential events that the system can possibly undergo. For example, in the Fig. 30 we show how we classify all meaningful events at a possible KMC step.

Next, the energy of local minima is calculated for all these tracked-down process. The energy minimization cycle is performed by a fixed number of steps of the Limited memory BFGS (L-BFGS) optimization algorithm. It is in the family of quasi-Newton methods that approximates the Broyden Fletcher Goldfarb Shanno (BFGS) algorithm using a limited amount of computer memory.

Implementation of this scheme was achieved by incorporating the KMC algoritm into i-Pi, which is a universal force engine interface code written in Python, designed to be used together with an *ab-initio*, force-field, or NN potential evaluation of the interactions between the atoms.[260] Our choice to adopt the NN potential implemented ac-
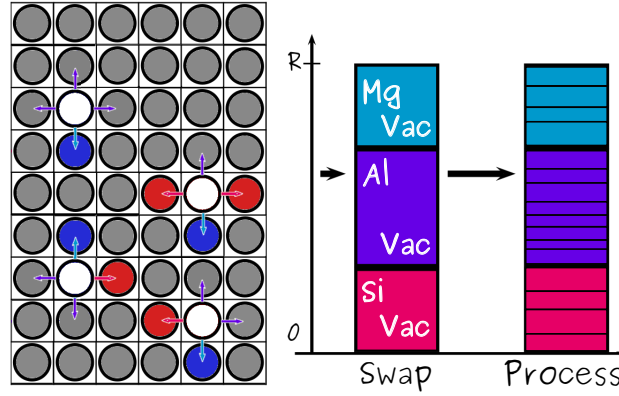
Figure 30: The criterion for assigning events. The arrows represent the possible events which are weighed according to the Eq. 55 so as to construct the vector of rates which adds up to R.

cording to the procedure presented in Ch. 4, instead of an *ab-initio* potential, is only because of the computational cost and the real time we wish to simulate. The rates are, in turn, estimated based on activation barriers, which are a correction to those calculated with the NEB method for the vacancy-assisted jump of an isolated solute atom. The new activation barrier for a swap involving a solute X is given by

$$E_{X,i} = \frac{(E_{i,LMS}^{NN}(n+1) - E_{LMS}^{NN}(n))}{2} + E_{X}^{DFT}, \tag{100}$$

where $E_{LMS}^{NN}(n)$ is the local minimized NN energy at KMC step n and $E_{i,LMS}^{NN}(n+1)$ is the minimized NN energies for the i-th candidate at step $n+1$. E_{X}^{DFT} is the DFT energy barrier for a swap of an isolate atom (e.g Vacancy/Al-atom, Vacancy/Mg-atom, or Vacancy/Si-atom). Rates computed based on these barriers are consistent with detailed balance.

Clearly this model that assumes constant prefactors for each solute, and an activation barrier that only depends on relaxed initial and final states and the dilute limit diffusion barriers is an approximation, but one that would affect kinetics and not the relative stability of clusters.

After this intermediate stage of revaluation of the rates, our KMC continues following the normal KMC steps:

- extraction of two normalized random numbers ρ_1 and ρ_2,

- evolution of the system according to the process k chosen through the condition:

$$\sum_{i=0}^{k} n_i r_i \geqslant \rho_1 R > \sum_{i=0}^{k-1} n_i r_i, \tag{101}$$

- updating of the clock according to Eq. 66 with $\rho_2 = \bar{p}_\tau$.

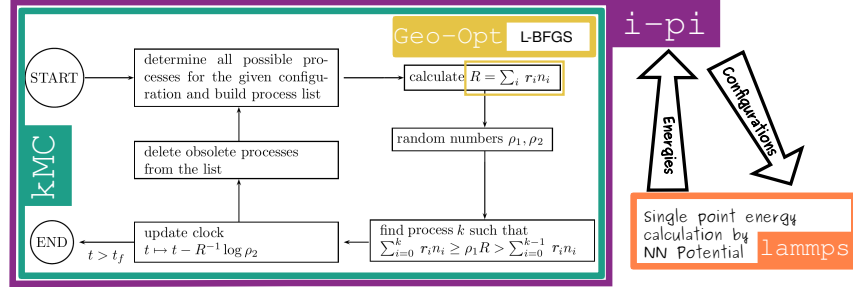In Fig. 31 , we show the complete scheme of our KMC algorithm.

Figure 31: Scheme of KMC algorithm interfaced with LAMMPS package[261] by means of i-pi code. The geometry optimization is internally managed by i-pi that communicates with LAMMPS only the new atomic positions receiving back from it both the energies and the forces (and even stress tensor if necessary).

In calculating the rates, we used as prefactors, ν_i^0, and DFT activation barriers, E_X^{DFT}, the values quoted by Mantina et al. [262] since the barriers are very similar to our results as reported in Fig. 23.

5.2 RESULTS

Using this NN driven vacancy-diffusion model, we run KMC simulations of the annealing of supersaturated Al–Mg–Si solid solutions. In order to avoid unphysical interactions between periodic images, we consider two Al-supercells: one with {6,6,6} primitive vectors and one vacancy, and the other with {8,8,8} primitive vectors and two vacancies. Alloys corresponding to a supersaturated solid solution containing 2.8 at.% Mg and 2.8 at.% Si (case of {6,6,6} Al-supercell) and 2.3 at.% Mg and 2.3 at.% Si (case of {8,8,8} Al-supercell) are prepared.

This concentration of vacancies (i.e. around $4{\cdot}10^{-3}$) is almost an order of magnitude greater than the estimated vacancy concentration at equilibrium (i.e. $7.2{\cdot}10^{-4}$)[1] in a pure aluminium at a heating temperature of 873 K. However, in a SSS the total concentration of vacancies at thermal equilibrium due to substitutional impurity atoms is given by

$$n_{vac} \sim \left((1 - 13\, n_{sol}) + 12\, n_{sol} e^{\frac{-S_{vac}}{k_B}} e^{\frac{-\bar{\varepsilon}_1}{k_B}} \right) e^{\frac{-S_{vac}}{k_B}} e^{\frac{-E_{vac}}{k_B T}} , \quad (102)$$

---

1 At equilibrium, the vacancy concentration at a given temperature is given approximately by the equation $n_{vac} = \exp(-S_{vac}/k_B)\exp(-E_{vac}/k_B T)$, where E_{vac} is the heat of formation of a vacancy, S_{vac} is the entropy of formation of a vacancy (e.g. $S_{vac} = 1.5\, k_B$ for closed packed metal [263], T is the temperature and k_B is the Boltzmann constant.

where n_{sol} is the concentration of solutes and $\bar{\epsilon}_1$ is the vacancy-solute binding average energy at first nearest neighbours. [263] Based on this expression, we can estimate n_{vac} at equilibrium to be $3.4 \cdot 10^{-3}$ for a supersaturated solid solution containing 4.6 at.% solute atoms at 873 K.

Starting from these SSS conditions, we ran KMC simulations at room temperature and observed the formation of some agglomerates containing Al-atoms in the core after ten milliseconds. As it was deduced from the previous study [257], we can assert that the clustering is a fast process where the most significant agglomerations take place within few hundredths of a second.

In the case of {6,6,6} Al-supercell and one vacancy (see Fig. 32, above: "1Vacancy"), we observe that the precipitate with composition $Mg_6Al_2Si_6Vac_1$ is the most stable during the first hundredth of a second of ageing, although it later dissolves. In the case of {8,8,8} Al-supercell with two vacancies (see Fig. 32, below: "2Vacancies"), we could also study the phenomenon of vacancies trapping by solute atoms. In particular, in the area marked in blue in the figure we have observed as a Si-Si dimer can strongly trap a vacancy for more than one millisecond (seed 2) and a small cluster of two Si atom and one Mg atom can trap both vacancies (seed 1). In both cases it has been noticed that this phenomenon happens in the first 10 milliseconds of simulation that is when clusters of two/three atoms move to aggregate. After ten milliseconds, the two simulations lead to two different nucleation mechanisms. In the first case (seed 1), the solute atoms cluster forming a single precipitate with a needle-like shape which is coherent with the Al-matrix. This looks similar to what D. Maisonnette et al. [8] have observed and it was reported in Fig. 1. These preliminary observations must be validated by a self-consistent procedure of refinement of the NN Potential but until then we can identify it provisionally as a possible GP zone. Given the finite concentration of solutes, the object is not stable throughout the simulation and eventually, dissolves. In the second case (seed 2), we have the formation of two precipitates instead of a big one. We observe that these pair of precipitates around 20 ms have the same energetics of the big needle-shaped precipitate, and $Al_2Mg_5Si_4Vac$ as chemical composition. These precipitates have the same size of 1 FU of β"-phase and the most stable ones, around 25 ms, have one less Al atom in the core (i.e. $AlMg_5Si_4Vac$). In both case the core of these precipitates contain Al-atoms, this is comparable to what has been published in a KMC study [259] regarding ternary aluminium-zinc-scandium alloys, where $L1_2$ ordered stable precipitates of Al_3Sn and Al_3Zn (typical of their binary alloys [264]) are encapsulated within a more complex agglomeration.
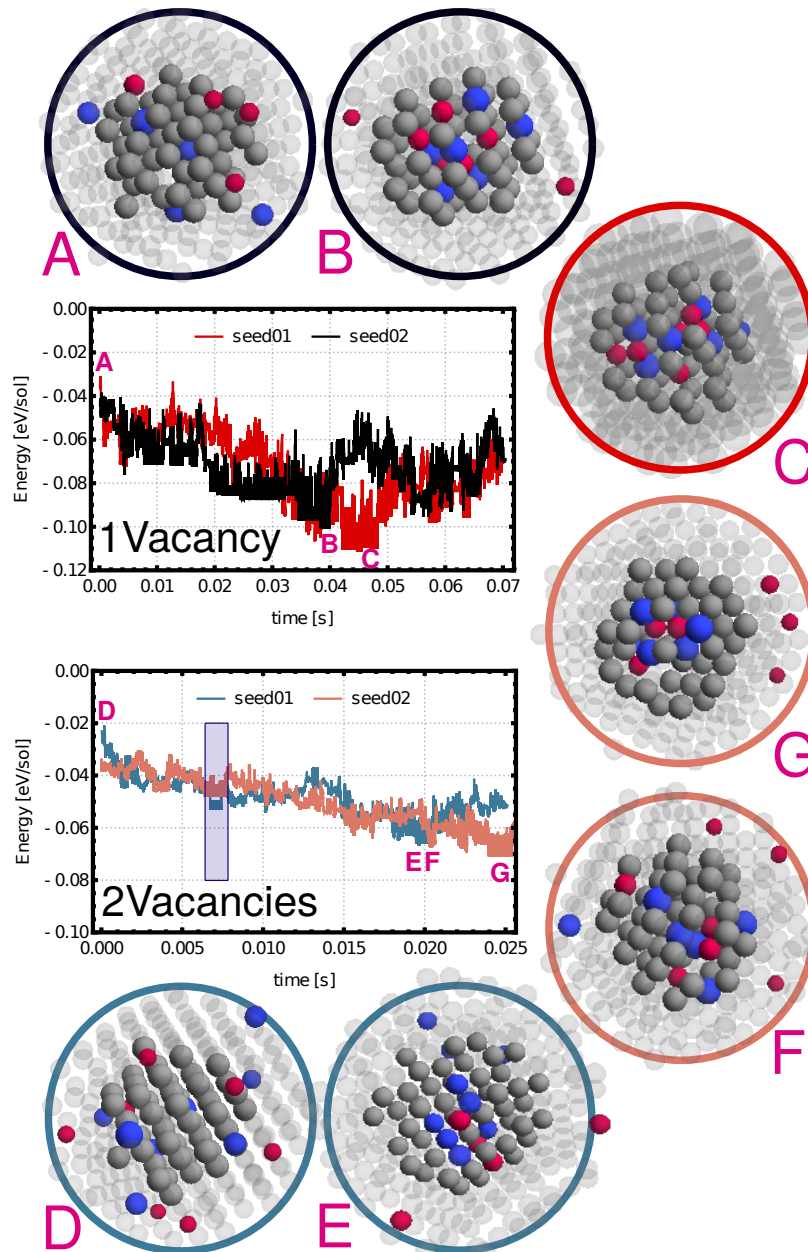
Figure 32: The evolution with time of the total energy per solute for a
           KMC simulation with two different SSS state (e.g different
           seed) in the case of {6,6,6} Al-supercell (above) and of {8,8,8}
           Al-supercell (below). We show some frame of the precip-
           itate morphology during the evolution. They are plotted
           by sitting on the centre of mass of the solutes making up
           the precipitate and using two cut-off radii to indicate prox-
           imity of atoms to that centre of mass. The first cut-off ra-
           dius of 6Å shows all atoms as solid spheres (e.g Al atoms
           as grey, Mg atoms as blue, and Si atoms as red spheres).
           While the second cut-off radius of 10Å shows all Al atoms
           within the spherical shell of these two spheric regions as
           grey circle and all solute atoms as spheres with reduced
           radius according to distance from the centre of mass.

5.3 OPTIMIZATION OF THE NN POTENTIAL

These qualitative observations obviously require longer simulations, larger cell sizes, and more sampling of different starting conditions. Before doing so, however, it is important to verify that the NN potential is transferable enough to determine precisely the energies of different configurations encountered during the KMC procedure, because in our simplified diffusion model they determine entirely the height of the barrier as well as the relative stability of different clusters.
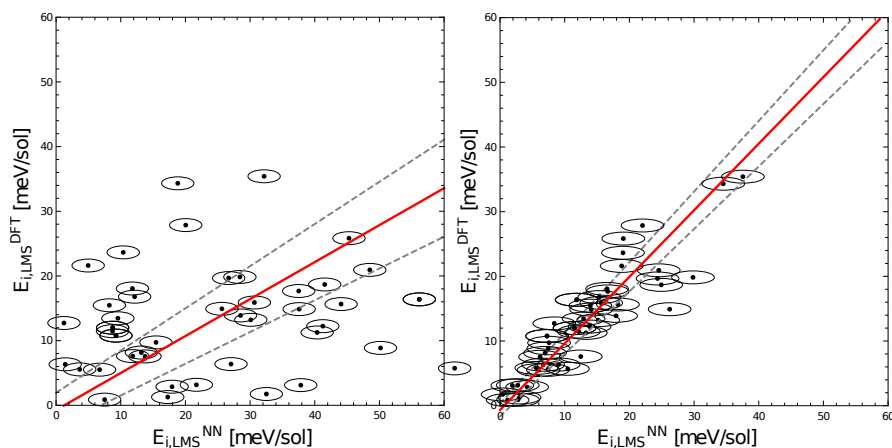
Figure 33: Energies per solute atom of 50 KMC structures evaluated by the 3B NN potential and the DFT calculations (left image), and by the extended 3B NN potential and the DFT calculations (right image). Each point has an oval that circumscribes its tolerance associated with the DFT error (± 1mev/at) and with 3B-NN error ($\pm$ 2.47 meV/at) (right image) or extended 3B-NN error (3.35 me/at) (left image). The $y = x$ indicates ideal matching between NN and DFT values and the red line is the orthogonal regression between these values, while the dashed lines represent its validity range.

To achieve such validation we extracted 50 random frames from two KMC simulations with 10'000 steps each, in the same temperature and concentration conditions. For each of these configurations we performed a DFT calculation with the same computational details used to generate the reference database for the NN, and compared the energy predicted by the NN and by DFT. As shown in the left panel of Fig. 33, there is little to no correlation between the two sets of energies (the R^2 coefficient is just 0.21). Even though the energy errors in terms of energy per solute atoms are only marginally higher than those seen in the NN validation set, it is clear that this lack of correlation makes KMC results untrustwortyh. To remedy this, we

proceeded to generate new configurations to supplement the previous reference dataset and fit an improved NN potential without re-doing the CUR selection of SF. To obtain a more thorough catalogue of possible pre-precipitation clusters, we ran a further set of simulations with 2% and 3% Mg and Si concentration, and temperatures of 300, 500, 1000 and 2000K. From these KMC trajectories we extracted 72'000 configurations, out of which we further selected by FPS 350 diverse configurations. For each, we ran new DFT calculations. We then added these structures to the reference database, and performed a NN fit using the same selection of 64 symmetry functions we used for the 3B NN discussed in section 4.3, with a 3:1 train:validation split. We could achieve a RMSE on the validation set of 3.35 meV per atom. We then re-computed the energetics of the same 50 random configurations extracted from the 72000 database using this extended NN; as shown in the right panel of Fig. 33, the correlation with DFT reference values is dramatically improved, with a R^2 of 0.82. Compared to the previous train data set, the new one also contains 350 structures/precipitates that are in an intermediate stage between the simple dimers and the complex β″-phases. These typology of structures was not present in the first database.

This extended NN can provide the basis for a more extensive set of KMC calculations, that will be the subject of future work. Additional validation with selected reference DFT calculations shall determine whether an additional iteration of re-training, including additional reference configurations, is needed to achieve the level of accuracy that is needed to determine precipitation energetics and kinetics with DFT accuracy.

6

CONCLUSIONS

I n this thesis we have engaged in a thorough analysis of the early
stages of precipitation in the ternary Al-6000 family of alloys,
with a multi-scale study involving *ab-initio* calculations, continuum
elasticity corrections, the optimization of a neural network interatomic
potential and the implementation of a quasi-on-lattice kinetic Monte
Carlo model.

By clearly identifying the chemical, surface, and elastic strain en-
ergies that contribute to the total precipitation energy versus size
and composition, and demonstrating that the overall trends are con-
sistent with a thermodynamic classical-nucleation-theory-like model,
we have provided new insights into the early stages of the formation
of β″ precipitates in Al-6000 alloys.

The in-situ needle-like β" precipitates are found to be stable rela-
tive to the solid solution down to the smallest in-plane formula unit,
indicating barrier-less growth at and above this size. The composition
dependence of the total energies is subtle, with two compositions be-
ing quite close in energy. Thus, the inclusion of surface energies and
elastic energies due to the different precipitate structures and com-
positions is essential for interpreting the DFT results and for then
determining the energetics in the more-dilute limit of real materials.
The benchmarking of the CNT-type model also provides a validation
for the use of such mesoscopic models in other systems.

The largest discrepancy between the thermodynamic CNT model
and DFT calculations is seen for the smallest precipitates, with the
ab initio energies being consistently much *lower* than those predicted
based on surface energies computed for a coherent interface between
the precipitate and the matrix. Together with the fact that the anisotropy
of γ is not sufficient to justify the aspect ratio of needle-like β″ pre-
cipitates, this observation hints strongly at the need for consideration
of more complex models of the interfaces of the precipitates – includ-
ing variable composition and a significant degree of reconstruction –
that may help reduce the interface and elastic energies and further
stabilize the small precipitates.

We further show that, down to a single formula unit that is fully
encapsulated in the Al matrix, the DFT energy of a nanoscale precip-
itate is lower than the reference supersaturated solid solution. This
underscores the fact that precipitation kinetics is likely to be diffusion-
limited. Aggregates of a few solute atoms that can act as vacancy
traps [187] would thus slow vacancy-mediated solute diffusion that

is necessary to form larger precipitates, greatly affecting the aging times. This conclusion of dominance of diffusion-controlled aging is also consistent with recent findings that the addition of 100 ppm of Sn to Al-6061 can significantly delay aging, attributed to trapping of the quenched-in vacancies by the Sn atoms [18, 265]. Our results thus point toward the need for a systematic study of the energetics of aggregates in the GP-zone regime, and the interactions between those aggregates and vacancies and/or trace elements in the alloy to understand and fine-tune the behavior of Al-6000 alloys in the early stages of precipitation.

In order to follow this lead we have developed a NN potential for studying mainly the precipitation strengthening in the pre-β'' stages. The NN potential is trained to reproduce DFT energies and forces for a diverse set of structures including several bulk phases, surfaces, vacancies and solid solutions. The NN potential is accurate in its prediction of (i) the lattice constant, bulk modulus, elastic moduli and surface energies of pure Al, (ii) the solute/solute and solute/vacancy binding energies to within 20 meV/pair, relative to the DFT results. Errors for properties included in the training DB are smaller than the uncertainties in the DFT method, as assessed by comparisons of DFT using two exchange-correlation functionals for a subset of properties. Errors for properties not included in the training data set are slightly larger but still generally comparable to the uncertainty of the DFT method.

In order to determine an appropriate NN parametrization to take into account 3 body interactions, we have introduced an unsupervised protocol that can be used to select the most important features out of a large initial pool of SFs candidates. This protocol is based on preserving the features that retain the most information, and are identified using relatively standard linear algebra methods, which makes the approach fully automatic and transferable to different systems and families of descriptors. Given that our strategy determines a low-dimensional description that requires the evaluation of only a small number of features, it can also reduce dramatically the cost of the property prediction. This is really apparent when considering a ternary system such as Al, Si and Mg alloy. The 3B-NN potentials based on AF-Protocol outperform a previous 2B-NN potential, both in terms of test RMSE and in terms of the accuracy of predicting new features (e.g. vacancy-assisted atom migration barriers).

Only the solute-solute binding energies show differences for both NN potentials that could be important for quantitative predictions of strengthening. The energies of small Al-Mg-Si clusters are also less accurate than other quantities, although not significantly on a per-atom basis, but trends with size and composition are followed quite well especially in the case of the 3B-NN potential. Since the machine

learning-type potentials are basically interpolating a potential energy landscape, these inaccurate properties could be improved by including training data such as dislocation core structures, interfaces and small precipitate structures. An unsolved problem remains in the modelling of solute atom interactions, due to the presence of long-range Friedel-like charge oscillations that lead to large finite-size effects that are difficult to incorporate with a short-range NN potential. We proposed a few remedies to this problem, including disrupting the coherent oscillations by randomization of the atomic positions or an artificial increase of the electronic temperature.

The overall success of the NN potential indicates that it can drive a KMC model for the studies of precipitate strengthening of Al-Mg-Si alloys involving the evolution of solute clusters, early formation of nano precipitates, and their interactions with dislocations. More broadly, these results show that machine learning-type potentials such based on the Neural-Networks can be quantitatively successful, and thus powerful tools, for modelling complex alloys that have proven to be extremely challenging cases for the other physics-based inter-atomic potential formulations.

Preliminary results on the first sub precipitation sequence (e.g. pre-β" phase) obtained by an almost off-lattice KMC model shed light on the mechanisms of vacancies trapping by solute atoms/clusters. In particular, their binding energies determine both the morphology and the size of the precipitates over time. Aggregates, which have similar composition to the more stable β"-phase proposed, are grossly observed in the first hundredths of a second.

While it was clear that a NN potential that is not fit to comparable precipitate structures as those encountered in a KMC run is not accurate enough to reproduce the correct energy trends between different configurations, we showed that a relatively straightforward extension of the data-set to include a few of such structures was sufficient to recover predictive energetics. Future work will focus on a more quantitative analysis of the pre-nucleation kinetics, and on a further cycle of validation and refinement of the NN potential.

The multi-scale analysis that we have presented here, leveraging state-of-the-art machine-learning potentials, ab initio energetics and coupling to continuum elasticity, open up to atomistic investigation the understanding of the microscopic phenomena that are crucial to the technology of precipitation-hardened alloys.

APPENDIX

7.1 CALCULATION OF EIGENSTRAIN AND STIFFNESS TENSORS

T he eigenstrain $\bar{\epsilon}$ is the strain required to compensate for the misfit between the matrix and precipitate lattices, i.e., the strain that deforms a formula unit of precipitate into the shape of a formula unit of undeformed matrix. Subsequently, we show how to compute $\bar{\epsilon}$ in the global frame of reference $\hat{\vec{e}}_x$-$\hat{\vec{e}}_y$-$\hat{\vec{e}}_z$ described in Figure 13. The formula unit geometries of the matrix and the precipitates are monoclinic cells for which the directions of $\vec{c}$ and $\vec{b}$ coincide but differ in the angle β and the edge lengths a, b, c. We start by determining the material frame of reference $\hat{\vec{e}}_\alpha$-$\hat{\vec{e}}_\beta$-$\hat{\vec{e}}_z$ as it simplifies both the expression of the edge vectors $\vec{a}, \vec{b}, \vec{c}$ and, since the elastic constants reported in Table 2 are computed in that frame, is required to compute stiffness tensors in the global frame.

The basis vectors $\hat{\vec{e}}_\beta$ and $\hat{\vec{e}}_z$ are collinear with the formula unit cell edge vectors defined in (67), $\vec{c}$ and $\vec{b}$, respectively, and the third basis vector $\hat{\vec{e}}_\alpha$ is chosen to complete a right-handed orthonormal basis

$$
\begin{aligned}
\hat{\vec{e}}_\beta &= \frac{\vec{c}}{c} = \tfrac{1}{\sqrt{10}} \left(-3, 1, 0\right)^{\mathrm{T}}, \\
\hat{\vec{e}}_z &= \frac{\vec{b}}{b} = \left(0, 0, 1\right)^{\mathrm{T}}, \\
\hat{\vec{e}}_\alpha &= \hat{\vec{e}}_\beta \times \hat{\vec{e}}_z = \tfrac{1}{\sqrt{10}} \left(1, 3, 0\right)^{\mathrm{T}}.
\end{aligned}
\tag{103}
$$

We use the basis vectors to express the edge vectors in the global frame of reference using Table 1

$$
\vec{c} = c\,\hat{\vec{e}}_\beta, \quad \vec{b} = b\,\hat{\vec{e}}_z, \quad \vec{a} = a\left(\sin\beta\,\hat{\vec{e}}_\alpha + \cos\beta\,\hat{\vec{e}}_\beta\right).
\tag{104}
$$

The eigenstrain $\bar{\epsilon}$ corresponds to a displacement gradient $\nabla\vec{u}$ that transforms the precipitate edge vectors into the matrix edge vectors, see Figure 34 (left). After defining matrices composed of the edge vectors for a precipitate $\boldsymbol{V}_{\mathrm{prec}} = \left(\vec{a}_{\mathrm{prec}}, \vec{b}_{\mathrm{prec}}, \vec{c}_{\mathrm{prec}}\right)$ and the matrix $\boldsymbol{V}_{\mathrm{matrix}} = \left(\vec{a}_{\mathrm{matrix}}, \vec{b}_{\mathrm{matrix}}, \vec{c}_{\mathrm{matrix}}\right)$, the displacement gradient $\nabla\vec{u}$ can be expressed as

$$
\boldsymbol{V}_{\mathrm{matrix}} = \nabla\vec{u}\,\boldsymbol{V}_{\mathrm{prec}} + \boldsymbol{V}_{\mathrm{prec}} \;\Rightarrow\; \nabla\vec{u} = \boldsymbol{V}_{\mathrm{matrix}}\,\boldsymbol{V}_{\mathrm{prec}}^{-1} - \boldsymbol{I},
\tag{105}
$$

where $\boldsymbol{I}$ is the identity matrix. The eigenstrain $\bar{\epsilon}$ is the symmetric part of $\nabla\vec{u}$

$$
\bar{\epsilon} = \tfrac{1}{2}\left(\nabla\vec{u} + \nabla\vec{u}^{\mathrm{T}}\right).
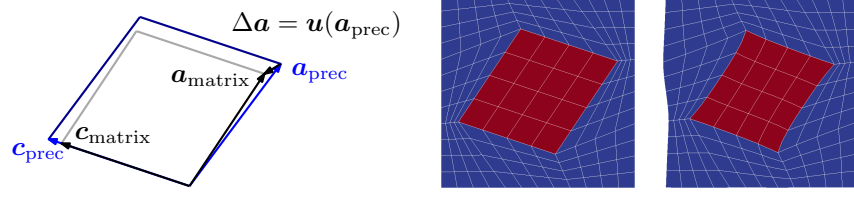\tag{106}
$$

Figure 34: Schematic illustration of eigendisplacement (left). Mesh for finite element analysis of the elastic problem (center). Note that the structured mesh follows the boundary of the precipitate (red parallelogram). Deformed elastic problem (right). The displacements have been magnified by 5 for better visibility. Note the periodic deformation.

The elastic constants of the precipitates have been calculated in the material frame of reference $\hat{\bar{e}}_\alpha$-$\hat{\bar{e}}_\beta$-$\hat{\bar{e}}_z$ and the corresponding stiffness tensor has to be rotated into the global frame of reference for the finite-element analysis. The stress $\boldsymbol{\sigma}$ and strain $\boldsymbol{\epsilon}$ in the global frame of reference are related to the material frame stress $\boldsymbol{\sigma}'$ and strain $\boldsymbol{\epsilon}'$ by the rotation $\boldsymbol{R} = \left(\hat{\bar{e}}_\alpha, \hat{\bar{e}}_\beta, \hat{\bar{e}}_z \right)$

$$\boldsymbol{\epsilon}' = \boldsymbol{R}\boldsymbol{\epsilon}\boldsymbol{R}^{\mathrm{T}}, \quad \boldsymbol{\sigma}' = \boldsymbol{R}\boldsymbol{\sigma}\boldsymbol{R}^{\mathrm{T}}, \tag{107}$$

and the relationship between $\boldsymbol{\sigma}'$ and $\boldsymbol{\epsilon}'$ is governed by elasticity

$$\boldsymbol{\sigma}' = \boldsymbol{C}'\,\boldsymbol{\epsilon}', \tag{108}$$

where $\boldsymbol{C}'$ is the stiffness tensor in the material frame of reference. The stiffness tensor in the global frame of reference $\boldsymbol{C}$ can be obtained by combination (107) and (108) in index notation (Einstein summation applies to repeated indices)

$$R_{ij}\sigma_{jk}R_{lk} = C'_{ilmn}R_{mo}\epsilon_{op}R_{np},$$
$$\underbrace{R_{ia}R_{ij}}_{\delta_{aj}}\sigma_{jk}\underbrace{R_{lk}R_{lb}}_{\delta_{kb}} = R_{ia}C'_{ilmn}R_{mo}\epsilon_{op}R_{np}R_{lb},$$
$$\sigma_{ab} = R_{ia}R_{lb}C'_{ilmn}R_{mo}R_{np}\epsilon_{op},$$
$$C_{abop} = R_{ia}R_{lb}R_{mo}R_{np}C'_{ilmn}. \tag{109}$$

7.2 ELASTIC CALCULATIONS

The elastic calculations use the finite element method [266] and have been performed using a modified version of the open-source finite-element code Akantu [267]. This section explains the chosen procedure.

We modeled the elastic problem using a structured, quadrilateral, and periodic two-dimensional mesh of bi-quadratic serendipity elements with eight nodes [268]. The element type was chosen over

linear elements for its high accuracy in static problems. In order to enforce periodic boundary conditions, we define the boundary nodes i_s of the upper and right boundary as slave nodes to their counterparts on the bottom and left boundary (master nodes i_m). During the evaluation of nodal forces on master nodes $\vec{f}_{i_m}$, the forces acting their slave nodes are also assembled on the master $\vec{f}_{i_m}^{tot} = \vec{f}_{i_m} + \vec{f}_{i_s}$ and the slave node displacement is set to be equal to the displacement of their master $\vec{u}_{i_s} = \vec{u}_{i_m}$. In order to preclude solid body motion (and, thus, a singular stiffness matrix K), the center node in the precipitate is fully blocked $\vec{u}_c = 0$.

Figure 34 (center and right) shows such a mesh in its original and deformed state where the displacements have been amplified by a factor five for better visibility. The structured mesh follows the boundary of the precipitate, such that any element is either of matrix material (blue) or precipitate material(red). Note the periodic deformation of the simulation cell. The precipitate is preloaded with the eigenstrain $\bar{\epsilon}$ as described in Section 7.1 and the stiffness tensors for matrix C_{matrix} and precipitate and C_{prec} are assigned to the blue and red elements respectively. In absence of external loads, the assembled system of equations to solve is

$$K\vec{U} = \vec{0}, \tag{110}$$

where K is the assembled stiffness matrix and $\vec{U}$ is the vector of all displacement degrees of freedom. We solve this system using the direct solver Mumps [269]. The calculation of strain energy exploits the quadrature routines of Akantu using the shape functions of the elements to evaluate the integrals in (76). Figure 35 shows the distribution of strain energy density e_{strain} for the geometries considered using the example of $Mg_4Al_3Si_4$. A mesh that is eight times finer than
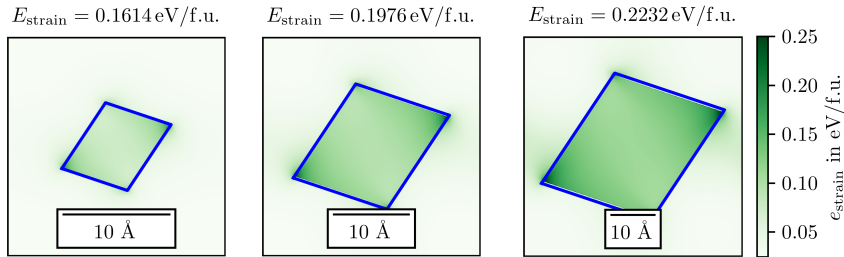
Figure 35: Distribution of strain energy density e_{strain} for different geometries in the example of $Mg_4Al_3Si_4$. The blue frame marks the boundaries of the precipitate.

the one represented in Figure 34 was used for smooth visualization.

7.2.1   *Relaxation of Boundary Conditions*

In order to compare our results more readily to those presented in [11], we have additionally performed elastic calculations with fully relaxed periodic boundary conditions, in which the simulation box was allowed to expand and tilt as needed to have no average stress. This was done by following the procedure described in Appendix 7.2, but with an additional uniform eigenstrain added to all elements. This additional eigenstrain was used as a degree of freedom in a minimization of the total strain energy.

Table 11 compares the strain energies per formula unit obtained with fixed periodic boundary conditions like the ones used in all DFT calculations in this work to the energies obtained using the relaxed boundary conditions used in [11]. One can see that the relaxed conditions lead to a consistent underestimation of the strain energy, while the fixed periodic conditions lead to overestimated energies.

Composition [meV per f.u.]	dilute	4×4 fixed	4×4 relaxed	2×2 fixed	2×2 relaxed	1×1 fixed	1×1 relaxed
$Mg_4Al_3Si_4$	74	117	53	106	59	89	66
$Mg_5Al_2Si_4$	128	223	89	198	98	161	113
Mg_5Si_6	140	223	114	203	122	171	132

Table 11: Comparison of elastic strain energies E_{strain} obtained for all considered geometries with periodic boundary conditions of fixed dimensions (as the DFT calculations in this work) or fully relaxed conditions for which there is no mean stress on the simulation box (as in [11]).

7.3 COMPARISON OF THE TWO NN POTENTIALS

Table 12: As in Sec. 4.2.2, pure Al bulk properties from *ab initio* calculations [20] and NN potentials with 2-body and 3-body interactions.

Al	Exp/*ab initio*	2-B NN	3B-NN
a (Å)	4.05, *4.06*	4.05	4.05
E_c (eV)	3.39 , *3.056*	3.057	3.055
B (GPa)	79.1, *72.2*	76.9	72.4
C_{11} (GPa)	114.3, *106.1*	109.9	105.4
C_{12} (GPa)	61.9, *55.9*	55.5	55.9
C_{44} (GPa)	31.6, *31.9*	31.6	31.4

Table 13: As in Sec. 4.2.3, lattice constant a (Å) and heat of solutions ΔH^{comp} (meV/atom) for binary and ternary compounds calculated using DFT-PBE and NN potential and NN potentials with 2-body and 3-body interactions.

Composition	Structure	*ab initio*		2-B NN		3B-NN	
		a	ΔH^{comp}	a	ΔH^{comp}	a	ΔH^{comp}
AlMg	B1	*5.760*	*424*	5.760	425	5.799	424
AlMg	B2	*3.389*	*66*	3.396	66	3.397	63
Al$_3$Mg	L1$_2$	*4.138*	*3*	4.145	0	4.182	5
AlMg$_3$	L1$_2$	*4.376*	*6*	4.146	4	4.379	3
Al$_{12}$Mg$_{17}$		*10.506*	*-18*	10.539	-18	10.520	-20
AlSi	B1	*5.217*	*260*	5.229	261	5.218	262
AlSi	B2	*3.160*	*239*	3.164	239	3.169	235
Al$_3$Si	L1$_2$	*3.994*	*95*	4.000	96	4.000	93
AlSi$_3$	L1$_2$	*3.899*	*333*	3.907	333	3.899	331
MgSi	B1	*5.507*	*384*	5.532	383	5.522	382
MgSi	B2	*3.308*	*144*	3.313	144	3.309	142
Mg$_3$Si	L1$_2$	*4.256*	*-7*	4.267	-8	4.285	-8
MgSi$_3$	L1$_2$	*3.988*	*269*	3.988	269	3.989	271
Mg$_2$Si	C1	*6.365*	*-136*	6.362	-136	6.364	-138
Mg$_5$Si$_6$		*4.081*	*13*	4.074	17	4.099	10
Al$_2$Mg$_5$Si$_4$		*4.054*	*-71*	4.053	-67	4.055	-72
Al$_3$Mg$_4$Si$_4$		*4.122*	*-44*	4.131	-39	4.112	-46

7.4   ATOMIC CHARGES BY BADER METHOD
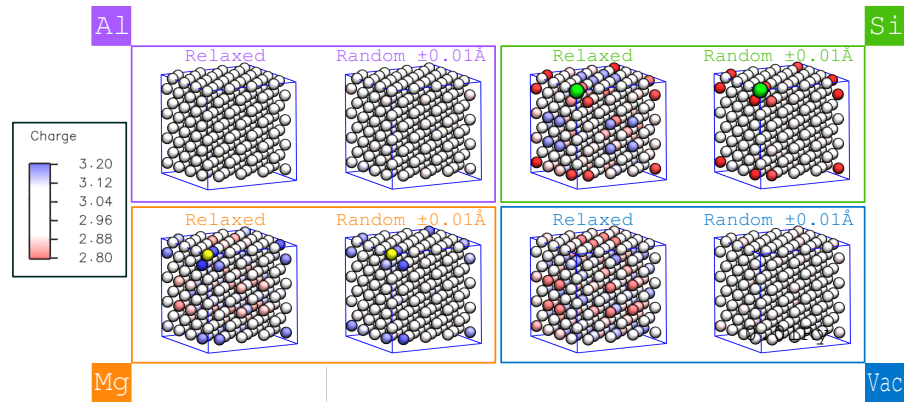
Figure 36: We show the atomic charges determined by Bader approach for one-solute atom systems (e.g one Al atom into 4x4x4 Al-matrix is substituted with an Al, Mg, Si, and Vac atom). In the "Relaxed" cases the atomic positions are fully-relaxed and the smearing is 0.05Ry, instead in the "Random ±0.01Å" cases the atomic positions are randomly displaced from the equilibrium positions for a factor ±0.01Å and the smearing is five times lower (e.g 0.01Ry). It is even clear that the charge fluctuations decrease when we introduce an entropic component into the atomic positions even if we decrease the electronic temperature.
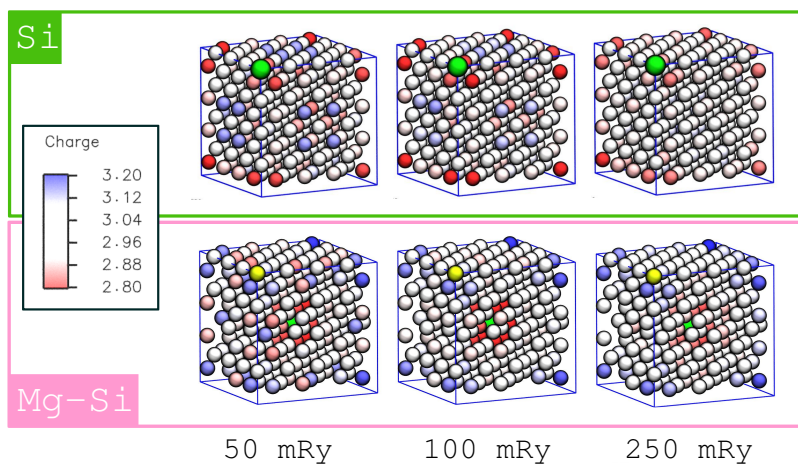
Figure 37: We show the atomic charges determined by Bader approach for the Si-solute atom system and **Mg**-**Si** system into 4x4x4 Al-matrix varying the smearing from 0.05 Ry to 0.25 Ry. In the previous sentence the substitutes were carefully marked and coloured so that it is possible to appreciate their relative positions in Fig. 25. It is clear that the charge fluctuations decrease when we raise the electronic temperature.

7.5 EQUILIBRIUM SHAPE OF β''-PHASE CLUSTERS

We can use the surface energies indicated in Table 3 in order to apply the Wulff Construction (WC) [270, 271] by presuming that these surfaces are associated to the cusp points in WC plot.

If we also assume that the real clusters are a simple stacking of single formula units of β''-precipitates, we can get their shape and express their 3D structure as a single size parameter which is trivially associable at the diagonal of their base. Now, we can estimate the structure of these finite-size precipitates imposing a diameter of 60 Å as reported in previous studies performed by means of a High-Resolution Transmission Electron Microscopy [4, 17]. We show the structures thus obtained in Fig. 38 and their 3D WC plot.
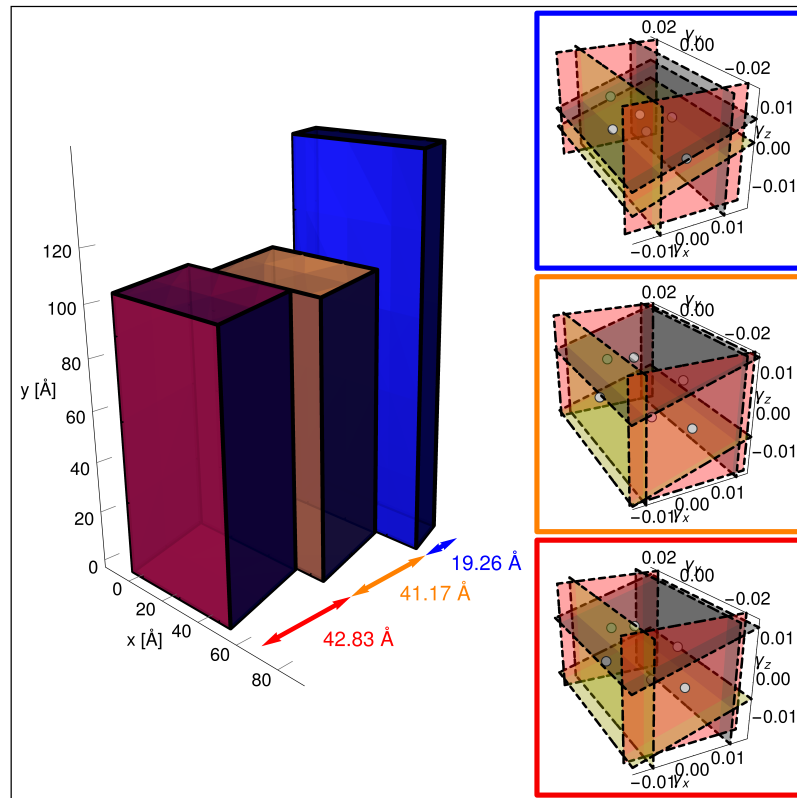
Figure 38: The equilibrium shape of three proposed precipitates, e.g. Mg_5Si_6, $Mg_4Al_3Si_4$, and $Mg_5Al_2Si_4$, in Al-matrix (on the left) and the associated Wulff Construction (on the right).

7.6 A PILOT TESTING OF KMC ALGORITHM

We show the "simple" case of one Si solute atom and one Vacancy in an Al-matrix as a criterion to test the potentialities of our KMC algo-

rithm. In this system we can map all possible interactions between these two defects based on their respective distance in the Al-matrix. We chose to consider a supercell of {4,4,4} primitive vectors containing 64 atoms because it is the smallest supercell where distances to the eighth Nearest-Neighbours (N-N) can exist. The NN Potential has a cut-off radius of 10.58 Å which, in Al-matrix, allows to consider the interactions between eighth N-N without being definitively damped to zero. its predictability and its major functionality

The simple case of the dumbbell (i.e. two defects) allows to exclude misunderstood cohesion terms from the discussion so as to illustrate more clearly how the non-locality of our model works. In Fig. 39, we
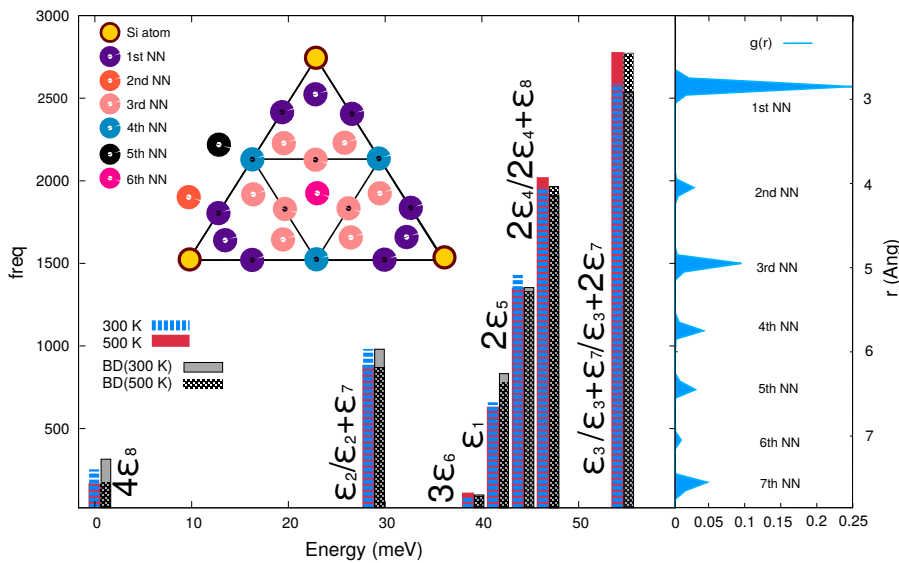
Figure 39: We plot the frequency counting of the various energy levels associated to different Si-Vac interactions (i.e. arrangements), ϵ_i (the i index specifies the membership of neighbours), for a KMC simulation running on 8'000 steps (i.e few tenths of a second of real time at room temperature). In the same picture, we indicate also the theoretical expectation by a Boltzmann distribution. On the top left side, a picture of i-N-N for Al-supercell with {4,4,4} primitive vectors in the plane that contains its shortest size is show for clarity purposes. The atomic sites on the bottom level have a black dot in the center of the circle. Vice versa, on the right side we have Si-Al radial distribution function averaged over frames and without applying the PBCs.

see all possible energy levels associated with this dumbbell. These are a consequence of the non-local effects introduced by NN potential through the minimisation cycles. This energy detail, which is not predictable by a basic model at first N-N, enriches the representation of the dynamics acting during the process. In this simple case, we can,

after an in-depth study, classify the possible configuration in which we can find the system. First of all, we notice that the interaction at seventh N-N is too damped to make significant contributions to energy and we can consider the energy of 4 eighth N-N (i.e. $4\,\epsilon_8$) as reference (i.e. it represents two isolated defects). This analysis clears up another important aspect that concerns the quarter, fifth and sixth N-N interactions for this supercell. They are associated by construction with configurations where an Si atom is approached by two fictitious vacancies (and vice versa), one the periodic image of the other. These two periodic images create a three-point system aligned in the case of fourth N-N arrangements, and desalinated in the other two cases. In the latter cases, this generates a transverse force. This combination of forces may be a possible explanation for why they deviate so much from the energy of isolated defects, $4\,\epsilon_8$.

In the same image we have also indicated the theoretical values of expectation by Boltzmann Distribution (BD). They were obtained counting the degeneracy of non-equivalent states. We note that both simulations at different temperature (i.e. 300K and 500K) respect this distribution. This ensures that the distribution of states is adequately simulated, essential requirement for properly developing a KMC algorithm.

[1] M Murayama and K Hono. "Pre-precipitate clusters and precipitation processes in Al–Mg–Si alloys." en. In: *Acta Materialia* 47.5 (Mar. 1999), pp. 1537–1548. ISSN: 13596454. DOI: `10.1016/S1359-6454(99)00033-6` (cit. on pp. 1–3).

[2] G.A. Edwards et al. "The precipitation sequence in Al–Mg–Si alloys." en. In: *Acta Materialia* 46.11 (July 1998), pp. 3893–3904. ISSN: 13596454. DOI: `10.1016/S1359-6454(98)00059-7` (cit. on pp. 1–3).

[3] S.P. Ringer and K. Hono. "Microstructural Evolution and Age Hardening in Aluminium Alloys." en. In: *Materials Characterization* 44.1-2 (Jan. 2000), pp. 101–131. ISSN: 10445803. DOI: `10.1016/S1044-5803(99)00051-0` (cit. on p. 1).

[4] S. J. Andersen et al. "The crystal structure of the β" phase in Al–Mg–Si alloys." In: *Acta Materialia* 46.9 (1998), pp. 3283–3298 (cit. on pp. 1, 3, 40, 41, 108).

[5] *Digital Activity Report*. 2017 (cit. on p. 1).

[6] C. D. Marioara et al. "Atomic model for GP-zones in a 6082 Al–Mg–Si system." In: *Acta materialia* 49.2 (2001), pp. 321–328 (cit. on pp. 2, 3).

[7] C.D Marioara et al. "The influence of temperature and storage time at RT on nucleation of the β" phase in a 6082 Al–Mg–Si alloy." en. In: *Acta Materialia* 51.3 (Feb. 2003), pp. 789–796. ISSN: 13596454. DOI: `10.1016/S1359-6454(02)00470-6` (cit. on pp. 2–4, 44).

[8] D. Maisonnette et al. "Effects of heat treatments on the microstructure and mechanical properties of a 6061 aluminium alloy." In: *Materials Science and Engineering: A* 528.6 (Mar. 2011), pp. 2718–2724. DOI: `10.1016/j.msea.2010.12.011` (cit. on pp. 2, 3, 93).

[9] C. D. Marioara et al. "The influence of alloy composition on precipitates of the Al-Mg-Si system." In: *Metallurgical and Materials Transactions A* 36.3 (2005), pp. 691–702 (cit. on pp. 2, 44).

[10] M. Murayama et al. "Atom probe studies on the early stages of precipitation in Al–Mg–Si alloys." In: *Materials Science and Engineering: A* 250.1 (1998), pp. 127–132 (cit. on p. 3).

[11]   Per Harald Ninive et al. "Detailed atomistic insight into the phase in Al–Mg–Si alloys." en. In: *Acta Materialia* 69 (May 2014), pp. 126–134. ISSN: 13596454. DOI: `10.1016/j.actamat.2014.01.052` (cit. on pp. 3, 40–42, 44, 50, 51, 63, 66, 104).

[12]   Dongdong Zhao et al. "Structure and thermodynamics of the key precipitated phases in the Al–Mg–Si alloys from first-principles calculations." en. In: *Journal of Materials Science* 46.24 (Dec. 2011), pp. 7839–7849. ISSN: 0022-2461, 1573-4803. DOI: `10.1007/s10853-011-5765-4` (cit. on p. 3).

[13]   S.J. Andersen et al. "The structural relation between precipitates in Al–Mg–Si alloys, the Al-matrix and diamond silicon, with emphasis on the trigonal phase U1-MgAl2Si2." en. In: *Materials Science and Engineering: A* 444.1-2 (Jan. 2007), pp. 157–169. ISSN: 09215093. DOI: `10.1016/j.msea.2006.08.084` (cit. on p. 3).

[14]   P. M. Derlet et al. "A first-principles study of the β"-phase in Al-Mg-Si alloys." In: *Journal of Physics: Condensed Matter* 14.15 (2002), p. 4011 (cit. on pp. 3, 40).

[15]   J. Gjønnes et al. "Electron crystallography of aluminum alloy phases." In: *Zeitschrift für Kristallographie - Crystalline Materials* 218.4 (Jan. 2003). ISSN: 2196-7105, 2194-4946. DOI: `10.1524/zkri.218.4.293.20745` (cit. on p. 3).

[16]   M. Takeda et al. "Stability of metastable phases and microstructures in the ageing process of Al–Mg–Si ternary alloys." In: *Journal of materials science* 33.9 (1998), pp. 2385–2390 (cit. on p. 3).

[17]   M. W. Zandbergen et al. "Data analysis and other considerations concerning the study of precipitation in Al–Mg–Si alloys by Atom Probe Tomography." In: *Data in Brief* 5 (Dec. 2015), pp. 626–641. ISSN: 2352-3409. DOI: `10.1016/j.dib.2015.09.045` (cit. on pp. 3, 108).

[18]   S. Pogatscher et al. "Diffusion on Demand to Control Precipitation Aging: Application to Al-Mg-Si Alloys." In: *Phys. Rev. Lett.* 112 (2014), p. 225701 (cit. on pp. 3, 40, 98).

[19]   Bin-Lung Ou and Chin-Hui Shen. "Effect of preaging on precipitation behavior in aluminum alloy 6063 during high-temperature aging." In: *Scandinavian journal of metallurgy* 33.2 (2004), pp. 105–112 (cit. on p. 4).

[20]   Daniele Giofré et al. "Ab initio modelling of the early stages of precipitation in Al-6000 alloys." In: *Acta Materialia* 140 (Nov. 2017), pp. 240–249. DOI: `10.1016/j.actamat.2017.08.017` (cit. on pp. 4, 36, 39, 56, 59, 63, 66, 67, 73, 75, 105).

[21] Ryo Kobayashi et al. "Neural network potential for Al-Mg-Si alloys." In: *Phys. Rev. Materials* 1 (2017), p. 053604 (cit. on pp. 5, 31, 50, 53, 73–75).

[22] Giulio Imbalzano et al. "Automatic selection of atomic fingerprints and reference configurations for machine-learning potentials." In: *The Journal of Chemical Physics* 148.24 (June 2018), p. 241730. DOI: 10.1063/1.5024611 (cit. on pp. 5, 53, 73, 74).

[23] M. P. Allen and D. J. Tildesley. *Computer simulation of liquids.* Oxford [England] : New York: Clarendon Press ; Oxford University Press, 1987. ISBN: 978-0-19-855375-5 (cit. on p. 8).

[24] A. P. J. Jansen. *An introduction to kinetic Monte Carlo simulations of surface reactions.* Lecture notes in physics 856. OCLC: 797324973. Heidelberg: Springer, 2012 (cit. on p. 8).

[25] John P Perdew and Alex Zunger. "Self-interaction correction to density-functional approximations for many-electron systems." In: *Physical Review B* 23.10 (1981), p. 5048 (cit. on p. 9).

[26] Seymour H Vosko, Leslie Wilk, and Marwan Nusair. "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis." In: *Canadian Journal of physics* 58.8 (1980), pp. 1200–1211 (cit. on p. 9).

[27] John P Perdew, Kieron Burke, and Matthias Ernzerhof. "Generalized gradient approximation made simple." In: *Physical review letters* 77.18 (1996), p. 3865 (cit. on p. 9).

[28] Chengteh Lee, Weitao Yang, and Robert G Parr. "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density." In: *Physical review B* 37.2 (1988), p. 785 (cit. on p. 9).

[29] John P Perdew. "Density-functional approximation for the correlation energy of the inhomogeneous electron gas." In: *Physical Review B* 33.12 (1986), p. 8822 (cit. on p. 9).

[30] Carlo Adamo and Vincenzo Barone. "Toward reliable density functional methods without adjustable parameters: The PBE0 model." In: *The Journal of chemical physics* 110.13 (1999), pp. 6158–6170 (cit. on p. 9).

[31] Axel D Becke. "Phys. Re V. A 1988, 38, 3098.(b) Becke." In: *J. Chem. Phys* 98 (1993), p. 5648 (cit. on p. 9).

[32] PJ Stephens et al. "Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields." In: *The Journal of Physical Chemistry* 98.45 (1994), pp. 11623–11627 (cit. on p. 9).

[33] John C Slater. "A simplification of the Hartree-Fock method." In: *Physical Review* 81.3 (1951), p. 385 (cit. on p. 9).

[34]   Charlotte Froese Fischer. "Hartree–Fock method for atoms. A numerical approach." In: (1977) (cit. on p. 9).

[35]   PJ Knowles and NC Handy. "A new determinant-based full configuration interaction method." In: *Chemical physics letters* 111.4-5 (1984), pp. 315–321 (cit. on p. 9).

[36]   Chr Møller and Milton S Plesset. "Note on an approximation treatment for many-electron systems." In: *Physical Review* 46.7 (1934), p. 618 (cit. on p. 9).

[37]   Hendrik J Monkhorst. "Calculation of properties with the coupled-cluster method." In: *International Journal of Quantum Chemistry* 12.S11 (1977), pp. 421–432 (cit. on p. 9).

[38]   Bogumil Jeziorski and Hendrik J Monkhorst. "Coupled-cluster method for multideterminantal reference states." In: *Physical Review A* 24.4 (1981), p. 1668 (cit. on p. 9).

[39]   John F Stanton and Rodney J Bartlett. "The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties." In: *The Journal of chemical physics* 98.9 (1993), pp. 7029–7039 (cit. on p. 9).

[40]   James B. Foresman and Æleen Frisch. *Exploring chemistry with electronic structure methods*. 2. ed. OCLC: 35319950. Pittsburgh, Pa: Gaussian, Inc, 1996. ISBN: 978-0-9636769-3-1 (cit. on p. 9).

[41]   Amir Karton. "A computational chemist's guide to accurate thermochemistry for organic molecules: A computational chemist's guide to accurate thermochemistry for organic molecules." In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 6.3 (May 2016), pp. 292–310. DOI: `10.1002/wcms.1249` (cit. on p. 9).

[42]   W. S. Ohlinger et al. "Efficient Calculation of Heats of Formation." In: *The Journal of Physical Chemistry A* 113.10 (Mar. 2009), pp. 2165–2175. DOI: `10.1021/jp810144q` (cit. on p. 9).

[43]   Kari Laasonen et al. ""Ab initio"liquid water." In: *The Journal of chemical physics* 99.11 (1993), pp. 9080–9089 (cit. on p. 10).

[44]   N.W. Ashcroft and N.D. Mermin. *Solid state physics*. Brooks Cole, 1976 (cit. on pp. 11, 59).

[45]   B.H. Bransden and C.J. Joachain. *Physics of Atoms and Molecules*. 2nd. Prentice Hall, 1994 (cit. on p. 11).

[46]   W. Kohn. "Electronic structure of matter - Wave functions and density functional." In: *Nobel Lecture* (1999) (cit. on p. 12).

[47]   P. Hohenberg and W. Kohn. "Inhomogeneous Electron Gas." In: *Phys. Rev.* 136.3B (Nov. 1964), B864–B871 (cit. on pp. 12, 13, 16).

[48] W. Kohn and L. J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects." In: *Phys. Rev.* 140.4A (Nov. 1965), A1133–A1138 (cit. on p. 12).

[49] Paolo Giannozzi et al. "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials." In: *J. Phys. Cond. Matt.* 21 (2009), pp. 395502–395519 (cit. on pp. 12, 17, 36, 40, 75).

[50] P. Giannozzi. *Density functional theory for electronic structure calculations*. Lecture Notes (cit. on p. 13).

[51] M. Levy, Perdew. J.P., and V. Sahni. In: *Phys. Rev. A* 30.2745 (1984) (cit. on p. 15).

[52] C.O. Almbladh and U. von Barth. In: *Phys. Rev. B* 31.6 (1985), p. 3231 (cit. on p. 15).

[53] O.V. Gritsenko, B. Braïda, and E.J. Baerends. In: *J. Chem. Phys.* 119.4 (2003), p. 1937 (cit. on p. 15).

[54] Levente Vitos, Pavel A. Korzhavyi, and Börje Johansson. "Modeling of alloy steels." In: *Materials Today* 5.10 (2002), pp. 14–23 (cit. on p. 15).

[55] C. Motta. "First-principles study of electronic transport in organic molecular junctions." PhD thesis. BOA, 2013 (cit. on p. 16).

[56] R.M. Dreizler and E.K.U. Gross. *Density Functional Theory: An Approach to the Quantum Many-Body Problem*. Berlin: Springer-Verlag, 1990 (cit. on p. 16).

[57] Y. Osaka. In: *J. Phys. Soc. Japan* 36.376 (1974) (cit. on p. 16).

[58] J.C. Slater. *Microwave Electronics*. New York: Van Nostrand, 1950 (cit. on p. 16).

[59] J.P. Perdew. "Density-functional approximation for the correlation energy of the inhomogeneous electron gas." In: *Phys. Rev. B* 33.12 (1986), p. 8822 (cit. on p. 16).

[60] J.P. Perdew, K. Burke, and Ernzerhof M. "Generalized gradient approximation made simple." In: *Phys. Rev. Lett.* 77 (1996), p. 3865 (cit. on p. 16).

[61] Joachim Paier et al. "The Perdew–Burke–Ernzerhof exchange-correlation functional applied to the G2-1 test set using a plane-wave basis set." In: *The Journal of Chemical Physics* 122.23 (June 2005), p. 234102. DOI: 10.1063/1.1926272 (cit. on p. 16).

[62] Nicola Marzari et al. *A Standard Solid State Pseudopotentials (SSSP) library optimized for accuracy and efficiency (Version 1.0, data download)*. 2018. DOI: 10.24435/materialscloud:2018.0001/v1 (cit. on p. 16).

[63] P. Ordejón, E. Artacho, and J.M. Soler. "Self-consistent order-N density-functional calculations for very large systems." In: *Phys. Rev. B* 53.10 (1996), p. 441 (cit. on p. 17).

[64] J.M. Soler et al. "The SIESTA method for ab initio order-N materials simulation." In: *J. Phys: Condens. Matter* 14.11 (2002), p. 2745 (cit. on p. 17).

[65] E. Artacho et al. "The SIESTA method: Developments and applicability." In: *J. Phys: Condens. Matter* 20.6 (2008), p. 4208 (cit. on p. 17).

[66] G. Kresse and J. Hafner. "*Ab initio* molecular dynamics for liquid metals." In: *Physical Review B* 47.1 (Jan. 1993), pp. 558–561. DOI: 10.1103/PhysRevB.47.558 (cit. on p. 17).

[67] H.J. Monkhorst and J.D. Pack. "Special points for Brillouin-zone integrations." In: *Phys. Rev. B* 13.12 (1976), p. 5188 (cit. on p. 17).

[68] Andrew M Rappe et al. "Optimized pseudopotentials." In: *Physical Review B* 41.2 (1990), p. 1227 (cit. on p. 18).

[69] M. Methfessel and A. T. Paxton. "High-precision sampling for Brillouin-zone integration in metals." In: *Physical Review B* 40.6 (Aug. 1989), pp. 3616–3621. DOI: 10.1103/PhysRevB.40.3616 (cit. on p. 19).

[70] Nicola Marzari et al. "Thermal Contraction and Disordering of the Al(110) Surface." In: *Physical Review Letters* 82.16 (Apr. 1999), pp. 3296–3299. DOI: 10.1103/PhysRevLett.82.3296 (cit. on p. 19).

[71] Christophe Andrieu et al. "An introduction to MCMC for machine learning." In: *Machine learning* 50.1-2 (2003), pp. 5–43 (cit. on p. 20).

[72] Nasser M Nasrabadi. "Pattern recognition and machine learning." In: *Journal of electronic imaging* 16.4 (2007), p. 049901 (cit. on p. 20).

[73] Simon S Haykin et al. *Neural networks and learning machines.* Vol. 3. Pearson Upper Saddle River, NJ, USA: 2009 (cit. on p. 20).

[74] Richard Fritz Lanman III et al. *Adaptive filtering with machine learning.* US Patent 9,769,553. Sept. 2017 (cit. on p. 20).

[75] Bart Kosko. "Neural networks and fuzzy systems: a dynamical systems approach to machine intelligence/book and disk." In: *Vol. 1Prentice hall* (1992) (cit. on p. 20).

[76] Milan Sonka, Vaclav Hlavac, and Roger Boyle. *Image processing, analysis, and machine vision.* Cengage Learning, 2014 (cit. on p. 20).

[77]   Ian H Witten et al. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016 (cit. on p. 20).

[78]   David J Hand, Heikki Mannila, and Padhraic Smyth. *Principles of data mining (adaptive computation and machine learning)*. MIT press Cambridge, MA, 2001 (cit. on p. 20).

[79]   Zoubin Ghahramani. "Unsupervised learning." In: *Advanced lectures on machine learning*. Springer, 2004, pp. 72–112 (cit. on p. 20).

[80]   Tao Li, Mitsunori Ogihara, and George Tzanetakis. *Music data mining*. CRC Press, 2011 (cit. on p. 20).

[81]   Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. Vol. 1. 1. MIT press Cambridge, 1998 (cit. on p. 21).

[82]   Lorenz C. Blum and Jean-Louis Reymond. "970 Million Drug-like Small Molecules for Virtual Screening in the Chemical Universe Database GDB-13." In: *Journal of the American Chemical Society* 131.25 (July 2009), pp. 8732–8733. DOI: 10.1021/ja902302h (cit. on p. 21).

[83]   Roman M. Balabin and Ekaterina I. Lomakina. "Neural network approach to quantum-chemistry data: Accurate prediction of density functional theory energies." In: *The Journal of Chemical Physics* 131.7 (Aug. 2009), p. 074104. DOI: 10.1063/1.3206326 (cit. on p. 21).

[84]   Albert P. Bartók et al. "Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons." In: *Physical Review Letters* 104.13 (Apr. 2010). DOI: 10.1103/PhysRevLett.104.136403 (cit. on pp. 21, 54, 71).

[85]   Pierre Baldi, Klaus-Robert Müller, and Gisbert Schneider. "Editorial: Charting Chemical Space: Challenges and Opportunities for Artificial Intelligence and Machine Learning." In: *Molecular Informatics* 30.9 (Sept. 2011), pp. 751–751. DOI: 10.1002/minf.201180003 (cit. on p. 21).

[86]   Jörg Behler. "Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations." In: *Physical Chemistry Chemical Physics* 13.40 (2011), p. 17930. DOI: 10.1039/c1cp21668f (cit. on pp. 21, 29).

[87]   Piero Gasparotto and Michele Ceriotti. "Recognizing molecular patterns by machine learning: An agnostic structural definition of the hydrogen bond." en. In: *The Journal of Chemical Physics* 141.17 (Nov. 2014), p. 174110. DOI: 10.1063/1.4900655 (cit. on p. 21).

[88] Albert P. Bartók et al. "Machine learning unifies the modeling of materials and molecules." In: *Science Advances* 3.12 (Dec. 2017), e1701816. DOI: `10.1126/sciadv.1701816` (cit. on p. 21).

[89] Aldo Glielmo, Peter Sollich, and Alessandro De Vita. "Accurate interatomic force fields via machine learning with covariant kernels." In: *Physical Review B* 95.21 (June 2017). DOI: `10.1103/PhysRevB.95.214302` (cit. on p. 21).

[90] Alberto Testolin and Marco Zorzi. "Probabilistic models and generative neural networks: towards an unified framework for modeling normal and impaired neurocognitive functions." In: *Frontiers in computational neuroscience* 10 (2016) (cit. on p. 21).

[91] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." In: *Nature* 521.7553 (2015), pp. 436–444 (cit. on p. 21).

[92] Ken-Ichi Funahashi. "On the approximate realization of continuous mappings by neural networks." In: *Neural Networks* 2.3 (Jan. 1989), pp. 183–192. DOI: `10.1016/0893-6080(89)90003-8` (cit. on p. 21).

[93] William H. Press, ed. *Numerical recipes in C: the art of scientific computing*. 2nd ed. Cambridge ; New York: Cambridge University Press, 1992. ISBN: 978-0-521-43108-8 978-0-521-43720-2 (cit. on p. 21).

[94] Josef Ischtwan and Michael A. Collins. "Molecular potential energy surfaces by interpolation." In: *The Journal of Chemical Physics* 100.11 (June 1994), pp. 8080–8088. DOI: `10.1063/1.466801` (cit. on p. 21).

[95] Danijar Hafner et al. "Learning Hierarchical Information Flow with Recurrent Neural Modules." In: *Advances in Neural Information Processing Systems*. 2017, pp. 6727–6736 (cit. on p. 21).

[96] David H. Hubel and Torsten N. Wiesel. "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex." In: *The Journal of physiology* 160.1 (1962), pp. 106–154 (cit. on p. 22).

[97] Daniel Svozil, Vladimir Kvasnicka, and Jiri Pospichal. "Introduction to multi-layer feed-forward neural networks." In: *Chemometrics and intelligent laboratory systems* 39.1 (1997), pp. 43–62 (cit. on p. 22).

[98] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "Imagenet classification with deep convolutional neural networks." In: *Advances in neural information processing systems*. 2012, pp. 1097–1105 (cit. on p. 22).

[99] Jeff Hawkins and Dileep George. *Hierarchical temporal memory: Concepts, theory and terminology*. Tech. rep. Technical report, Numenta, 2006 (cit. on p. 22).

[100] Jonas Gehring et al. "Extracting deep bottleneck features using stacked auto-encoders." In: *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE. 2013, pp. 3377–3381 (cit. on p. 22).

[101] George Dahl, Abdel-rahman Mohamed, Geoffrey E Hinton, et al. "Phone recognition with the mean-covariance restricted Boltzmann machine." In: *Advances in neural information processing systems*. 2010, pp. 469–477 (cit. on p. 22).

[102] Geoffrey E Hinton. "Deep belief networks." In: *Scholarpedia* 4.5 (2009), p. 5947 (cit. on p. 22).

[103] Tomáš Mikolov et al. "Recurrent neural network based language model." In: *Eleventh Annual Conference of the International Speech Communication Association*. 2010 (cit. on p. 22).

[104] Sepp Hochreiter and Jürgen Schmidhuber. "Long short-term memory." In: *Neural computation* 9.8 (1997), pp. 1735–1780 (cit. on p. 22).

[105] Volodymyr Mnih et al. "Human-level control through deep reinforcement learning." In: *Nature* 518.7540 (2015), p. 529 (cit. on p. 22).

[106] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. "Multilayer feedforward networks are universal approximators." In: *Neural Networks* 2.5 (Jan. 1989), pp. 359–366. DOI: 10.1016/0893-6080(89)90020-8 (cit. on pp. 23, 24).

[107] Nongnuch Artrith and Alexander Urban. "An implementation of artificial neural-network potentials for atomistic materials simulations: Performance for TiO 2." In: *Computational Materials Science* 114 (Mar. 2016), pp. 135–150. DOI: 10.1016/j.commatsci.2015.11.047 (cit. on pp. 24, 32).

[108] Andrew L. Maas, Awni Y. Hannun, and Andrew Y. Ng. "Rectifier nonlinearities improve neural network acoustic models." In: *Proc. icml*. Vol. 30. 2013, p. 3 (cit. on p. 25).

[109] Claude Mbusa Takenga et al. "Comparison of gradient descent method, Kalman filtering and decoupled Kalman in training neural networks used for fingerprint-based positioning." In: *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*. Vol. 6. IEEE, 2004, pp. 4146–4150 (cit. on p. 26).

[110]   A. Bholoa, S.D. Kenny, and R. Smith. "A new approach to potential fitting using neural networks." en. In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 255.1 (Feb. 2007), pp. 1–7. DOI: 10.1016/j.nimb.2006.11.040 (cit. on p. 28).

[111]   Jörg Behler et al. "Metadynamics Simulations of the High-Pressure Phases of Silicon Employing a High-Dimensional Neural Network Potential." In: *Phys. Rev. Lett.* 100 (2008), p. 185501 (cit. on p. 28).

[112]   Jörg Behler. "Atom-centered symmetry functions for constructing high-dimensional neural network potentials." In: *J. Chem. Phys.* 134 (2011) (cit. on pp. 28, 30, 68).

[113]   Jörg Behler. "Perspective: Machine learning potentials for atomistic simulations." In: *The Journal of Chemical Physics* 145.17 (2016), p. 170901 (cit. on pp. 28, 31).

[114]   Gabriele C. Sosso et al. "Neural network interatomic potential for the phase change material GeTe." en. In: *Physical Review B* 85.17 (May 2012). ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.85.174103 (cit. on pp. 28, 68).

[115]   Kwang-Hwi Cho, Kyoung Tai No, and Harold A Scheraga. "A polarizable force field for water using an artificial neural network." In: *Journal of molecular structure* 641.1 (2002), pp. 77–91 (cit. on p. 28).

[116]   Helmut Gassner et al. "Representation of Intermolecular Potential Functions by Neural Networks." en. In: *The Journal of Physical Chemistry A* 102.24 (June 1998), pp. 4596–4605. ISSN: 1089-5639, 1520-5215. DOI: 10.1021/jp972209d (cit. on p. 28).

[117]   Jörg Behler and Michele Parrinello. "Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces." In: *Phys. Rev. Lett.* 98 (2007), p. 146401 (cit. on pp. 28, 54, 68, 71).

[118]   a. Bholoa, S.D. D Kenny, and R. Smith. "A new approach to potential fitting using neural networks." In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 255 (2007), pp. 1–7 (cit. on p. 28).

[119]   RuNNer - A Neural Network Code for High-Dimensional Potential-Energy Surfaces, Jörg Behler, Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, Germany. (Cit. on pp. 29, 74).

[120]   Jörg Behler. "Constructing high-dimensional neural network potentials: A tutorial review." In: *International Journal of Quantum Chemistry* 115.16 (2015), pp. 1032–1050 (cit. on p. 30).

[121]   Jörg Behler. "First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems." In: *Angewandte Chemie International Edition* 56.42 (Aug. 2017), pp. 12828–12840. DOI: 10.1002/anie.201703114 (cit. on pp. 30, 68).

[122]   Jonathan A Zimmerman et al. "Calculation of stress in atomistic simulation." In: *Modelling and simulation in materials science and engineering* 12.4 (2004), S319 (cit. on p. 31).

[123]   Ryo Kobayashi. *Nagoya Atomistic-simulation Package.* 2014 (cit. on pp. 31, 55).

[124]   Grégoire Montavon, ed. *Neural networks: tricks of the trade.* 2. ed. Lecture notes in computer science 7700. OCLC: 828098376. Heidelberg: Springer, 2012. ISBN: 978-3-642-35289-8 978-3-642-35288-1 (cit. on p. 32).

[125]   William H. Press, ed. *FORTRAN numerical recipes.* 2nd ed. Cambridge [England] ; New York: Cambridge University Press, 1996. ISBN: 978-0-521-43064-7 978-0-521-57439-6 (cit. on p. 32).

[126]   Xinghang Zhang et al. "Radiation damage in nanostructured materials." In: *Progress in Materials Science* (Mar. 2018) (cit. on p. 33).

[127]   M. Jaraiz et al. "Atomistic calculations of ion implantation in Si: Point defect and transient enhanced diffusion phenomena." In: *Applied Physics Letters* 68.3 (Jan. 1996), pp. 409–411 (cit. on p. 33).

[128]   Graeme Henkelman, Gísli Jóhannesson, and Hannes Jónsson. *Methods for finding saddle points and minimum energy paths.* Springer, 2002, pp. 269–302 (cit. on p. 34).

[129]   Kurt E Sickafus et al. *Radiation effects in solids: [proceedings of the NATO Advanced Study Institute on Radiation Effects in Solids : Erice, Sicily, Italy, 17-29 July 2004.* OCLC: 422734740. Dordrecht: Springer, 2007 (cit. on p. 35).

[130]   Fedwa El-Mellouhi, Normand Mousseau, and Laurent J. Lewis. "Kinetic activation-relaxation technique: An off-lattice self-learning kinetic Monte Carlo algorithm." en. In: *Physical Review B* 78.15 (Oct. 2008). DOI: 10.1103/PhysRevB.78.153202 (cit. on p. 35).

[131]   C Sasaoka, Y Kato, and A Usui. "Anomalous As desorption from InAs (100) 2× 4." In: *Applied physics letters* 62.19 (1993), pp. 2338–2340 (cit. on p. 35).

[132]   T Lazauskas, S D Kenny, and R Smith. "Influence of the pref-actor to defect motion in α-Iron during long time scale simula-tions." In: *Journal of Physics: Condensed Matter* 26.39 (Oct. 2014), p. 395007. DOI: 10.1088/0953-8984/26/39/395007 (cit. on p. 35).

[133]   Hisao Yamauchi. "The equilibrium solution of vineyard's equa-tion for the kinetics of long-range order." en. In: *Scripta Met-allurgica* 7.1 (Jan. 1973), pp. 109–113. ISSN: 00369748. DOI: 10.1016/0036-9748(73)90191-9 (cit. on p. 35).

[134]   Y. Mishin et al. "Structural stability and lattice defects in cop-per: *Ab initio* , tight-binding, and embedded-atom calculations." In: *Physical Review B* 63.22 (May 2001). DOI: 10.1103/PhysRevB.63.224106 (cit. on p. 36).

[135]   G.P.M. Leyson, L.G. Hector, and W.A. Curtin. "Solute strength-ening from first principles and application to aluminum al-loys." en. In: *Acta Materialia* 60.9 (May 2012), pp. 3873–3884. DOI: 10.1016/j.actamat.2012.03.037 (cit. on pp. 36, 63).

[136]   Hannes Jónsson, Greg Mills, and Karsten W. Jacobsen. "Nudged elastic band method for finding minimum energy paths of transitions." In: *Classical and quantum dynamics in condensed phase simulations*. World Scientific, 1998, pp. 385–404 (cit. on p. 37).

[137]   Graeme Henkelman, Blas P. Uberuaga, and Hannes Jónsson. "A climbing image nudged elastic band method for finding saddle points and minimum energy paths." In: *The Journal of Chemical Physics* 113.22 (Dec. 2000), pp. 9901–9904. DOI: 10.1063/1.1329672 (cit. on p. 37).

[138]   N. G. Van Kampen, TotalBoox, and TBX. *Stochastic Processes in Physics and Chemistry*. OCLC: 969056915. Elsevier Science, 2011 (cit. on p. 37).

[139]   Kerson Huang. *Statistical mechanics*. 2nd ed. New York: Wiley, 1987 (cit. on p. 38).

[140]   William Feller. *An introduction to probability theory and its appli-cations*. 3. ed., rev. print., [Nachdr.] Wiley series in probability and mathematical statistics. S.l.: Wiley, 2009 (cit. on p. 38).

[141]   E. Çınlar. *Introduction to stochastic processes*. Englewood Cliffs, N.J: Prentice-Hall, 1974 (cit. on p. 38).

[142]   E. Vincent et al. "Precipitation of the FeCu system: A critical review of atomic kinetic Monte Carlo simulations." In: *Journal of Nuclear Materials* 373.1-3 (Feb. 2008), pp. 387–401. DOI: 10.1016/j.jnucmat.2007.06.016 (cit. on p. 38).

[143] David Molnar et al. "Multiscale simulations on the coarsening of Cu-rich precipitates in α-Fe using kinetic Monte Carlo, molecular dynamics and phase-field simulations." In: *Acta Materialia* 60.20 (Dec. 2012), pp. 6961–6971. DOI: 10.1016/j.actamat.2012.08.051 (cit. on p. 38).

[144] W M Young and E W Elcock. "Monte Carlo studies of vacancy migration in binary ordered alloys: I." In: *Proceedings of the Physical Society* 89.3 (Nov. 1966), pp. 735–746. DOI: 10.1088/0370-1328/89/3/329 (cit. on p. 38).

[145] N. Castin, M. I. Pascuet, and L. Malerba. "Modeling the first stages of Cu precipitation in α-Fe using a hybrid atomistic kinetic Monte Carlo approach." In: *The Journal of Chemical Physics* 135.6 (Aug. 2011), p. 064502. DOI: 10.1063/1.3622045 (cit. on p. 38).

[146] G. Sha and A. Cerezo. "Kinetic Monte Carlo simulation of clustering in an Al–Zn–Mg–Cu alloy (7050)." In: *Acta Materialia* 53.4 (Feb. 2005), pp. 907–917. DOI: 10.1016/j.actamat.2004.10.048 (cit. on p. 38).

[147] Håkon S. Hasting et al. "Composition of β″ precipitates in Al–Mg–Si alloys by atom probe tomography and first principles calculations." en. In: *Journal of Applied Physics* 106.12 (2009), p. 123527. ISSN: 00218979. DOI: 10.1063/1.3269714 (cit. on p. 40).

[148] Jp P Perdew, K Burke, and M Ernzerhof. "Generalized Gradient Approximation made simple." In: *Phys. Rev. Lett.* Phys. Rev. Lett. (USA) 77 (1996), p. 3865 (cit. on p. 40).

[149] D Vanderbilt. "Soft self-consistent pseudopotentials in a generalized eigenvalue formalism." In: *Phys. Rev. B* 41 (1990), pp. 7892–7895 (cit. on p. 40).

[150] Georg Kresse and D. Joubert. "From ultrasoft pseudopotentials to the projector augmented-wave method." In: *Physical Review B* 59.3 (1999), p. 1758 (cit. on p. 40).

[151] Ivano Castelli and Nicola Marzari. *Standard Solid State Pseudopotentials*. 2015 (cit. on p. 40).

[152] H J Monkhorst and J D Pack. "Special points for Brillouin-zone integrations." In: *Phys. Rev. B* 13 (1976), pp. 5188–5192 (cit. on p. 40).

[153] Wheeler P. Davey. "Precision measurements of the lattice constants of twelve common metals." In: *Physical Review* 25.6 (1925), p. 753 (cit. on p. 40).

[154] Michael J. Tambe, Nicola Bonini, and Nicola Marzari. "Bulk aluminum at high pressure: A first-principles study." en. In: *Physical Review B* 77.17 (May 2008). ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.77.172102 (cit. on p. 40).

[155] C. Ravi and C. Wolverton. "First-principles study of crystal structure and stability of Al/Mg/Si/(Cu) precipitates." In: *Acta Materialia* 52 (2004), pp. 4213–4227 (cit. on p. 40).

[156] O. H. Nielsen and Richard M. Martin. "First-principles calculation of stress." In: *Physical Review Letters* 50.9 (1983), p. 697 (cit. on p. 41).

[157] G V Sin'ko and N A Smirnov. "Ab initio calculations of elastic constants and thermodynamic properties of bcc, fcc, and hcp Al crystals under pressure." In: *Journal of Physics: Condensed Matter* 14.29 (2002), p. 6989. URL: http://stacks.iop.org/0953-8984/14/i=29/a=301 (cit. on p. 42).

[158] C. Bercegeay and S. Bernard. "First-principles equations of state and elastic properties of seven metals." en. In: *Physical Review B* 72.21 (Dec. 2005). ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.72.214101 (cit. on p. 42).

[159] R. Yu, J. Zhu, and H.Q. Ye. "Calculations of single-crystal elastic constants made simple." en. In: *Computer Physics Communications* 181.3 (Mar. 2010), pp. 671–675. ISSN: 00104655. DOI: 10.1016/j.cpc.2009.11.017 (cit. on p. 42).

[160] D.Y. Li and L.Q. Chen. "Computer simulation of stress-oriented nucleation and growth of θ' precipitates inAl–Cu alloys." en. In: *Acta Materialia* 46.8 (May 1998), pp. 2573–2585. ISSN: 13596454. DOI: 10.1016/S1359-6454(97)00478-3 (cit. on p. 43).

[161] Kang Luo et al. "Stress/strain aging mechanisms in Al alloys from first principles." en. In: *Transactions of Nonferrous Metals Society of China* 24.7 (July 2014), pp. 2130–2137. ISSN: 10036326. DOI: 10.1016/S1003-6326(14)63323-9 (cit. on p. 43).

[162] Shang Fu et al. "Effects of external stress aging on morphology and precipitation behavior of θ'' phase in Al-Cu alloy." en. In: *Transactions of Nonferrous Metals Society of China* 24.7 (July 2014), pp. 2282–2288. ISSN: 10036326. DOI: 10.1016/S1003-6326(14)63345-8 (cit. on p. 43).

[163] Ji-Yong Yao et al. "A TEM study of precipitation in Al–Mg–Si alloys." In: *Micron* 32.8 (2001), pp. 865–870 (cit. on p. 44).

[164] Y. Wang et al. "First-principles calculations of β''-Mg5Si6/α-Al interfaces." en. In: *Acta Materialia* 55.17 (Oct. 2007), pp. 5934–5947. ISSN: 13596454. DOI: 10.1016/j.actamat.2007.06.045 (cit. on pp. 44–46).

[165]   Per Harald Ninive, Ole Martin Løvvik, and Are Strandlie. "Density Functional Study of the β" Phase in Al-Mg-Si Alloys." en. In: *Metallurgical and Materials Transactions A* 45.6 (June 2014), pp. 2916–2924. ISSN: 1073-5623, 1543-1940. DOI: 10.1007/s11661-014-2214-4 (cit. on pp. 46, 51).

[166]   Murray S. Daw and M. I. Baskes. "Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals." en. In: *Physical Review B* 29.12 (June 1984), pp. 6443–6453. DOI: 10.1103/PhysRevB.29.6443 (cit. on p. 53).

[167]   Byeong-Joo Lee and MI Baskes. "Second nearest-neighbor modified embedded-atom-method potential." In: *Physical Review B* 62.13 (2000), p. 8564. DOI: 10.1103/PhysRevB.62.8564 (cit. on p. 53).

[168]   B Jelinek et al. "Modified embedded-atom method interatomic potentials for the Mg-Al alloy system." In: *Physical Review B* 75.5 (2007), p. 054106. DOI: 10.1103/PhysRevB.75.054106 (cit. on p. 54).

[169]   Byeong-Joo Lee et al. "The modified embedded-atom method interatomic potentials and recent progress in atomistic simulations." In: *Calphad* 34.4 (Dec. 2010), pp. 510–522. DOI: 10.1016/j.calphad.2010.10.007 (cit. on p. 54).

[170]   Jörg Behler and Michele Parrinello. "Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces." en. In: *Physical Review Letters* 98.14 (Apr. 2007). ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.98.146401 (cit. on p. 54).

[171]   Albert P. Bartók and Gábor Csányi. "Gaussian approximation potentials: A brief tutorial introduction." en. In: *International Journal of Quantum Chemistry* 115.16 (Aug. 2015), pp. 1051–1057. ISSN: 00207608. DOI: 10.1002/qua.24927 (cit. on pp. 54, 71, 72).

[172]   Atsuto Seko, Akira Takahashi, and Isao Tanaka. "Sparse representation for a potential energy surface." en. In: *Physical Review B* 90.2 (July 2014). ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.90.024101 (cit. on p. 54).

[173]   Albert P. Bartók et al. "Machine-learning approach for one- and two-body corrections to density functional theory: Applications to molecular and condensed water." In: *Phys. Rev. B* 88 (2013), p. 054104 (cit. on p. 54).

[174]   Aldo Glielmo, Peter Sollich, and Alessandro De Vita. "Accurate interatomic force fields via machine learning with covariant kernels." In: *Physical Review B* 95.21 (2017), p. 214302 (cit. on p. 54).

[175]   Andrea Grisafi et al. "Symmetry-Adapted Machine Learning for Tensorial Properties of Atomistic Systems." In: *Physical Review Letters* 120.3 (2018), p. 036002 (cit. on p. 54).

[176]   Li Zhu et al. "A fingerprint based metric for measuring similarities of crystalline structures." In: *J. Chem. Phys.* 144 (2016), p. 034203 (cit. on p. 54).

[177]   Bing Huang and O Anatole von Lilienfeld. "Communication: Understanding molecular representations in machine learning: The role of uniqueness and target similarity." In: *J. Chem. Phys.* 145 (2016) (cit. on p. 54).

[178]   Felix A Faber et al. "Machine Learning Energies of 2 Million Elpasolite (A B C 2 D 6) Crystals." In: *Physical Review Letters* 117.13 (2016), p. 135502 (cit. on p. 54).

[179]   John P. Perdew et al. "Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces." en. In: *Physical Review Letters* 100.13 (Apr. 2008). ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.100.136406 (cit. on p. 55).

[180]   Young-Min Kim, Nack J. Kim, and Byeong-Joo Lee. "Atomistic Modeling of pure Mg and Mg–Al systems." en. In: *Calphad* 33.4 (Dec. 2009), pp. 650–657. ISSN: 03645916. DOI: 10.1016/j.calphad.2009.07.004 (cit. on p. 55).

[181]   Steve Plimpton. "Fast Parallel Algorithms for Short-Range Molecular Dynamics." en. In: *Journal of Computational Physics* 117.1 (Mar. 1995), pp. 1–19. ISSN: 00219991. DOI: 10.1006/jcph.1995.1039 (cit. on p. 55).

[182]   R. Biswas and D. R. Hamann. "Simulated annealing of silicon atom clusters in Langevin molecular dynamics." en. In: *Physical Review B* 34.2 (July 1986), pp. 895–901. ISSN: 0163-1829. DOI: 10.1103/PhysRevB.34.895 (cit. on p. 56).

[183]   H. J. C. Berendsen et al. "Molecular dynamics with coupling to an external bath." en. In: *The Journal of Chemical Physics* 81.8 (Oct. 1984), pp. 3684–3690. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.448118 (cit. on p. 56).

[184]   Gang Lu et al. "Generalized-stacking-fault energy surface and dislocation properties of aluminum." en. In: *Physical Review B* 62.5 (Aug. 2000), pp. 3099–3108. ISSN: 0163-1829, 1095-3795. DOI: 10.1103/PhysRevB.62.3099 (cit. on p. 59).

[185]   Xiao-Zhi Wu et al. "Ab initio calculations of generalized-stacking-fault energy surfaces and surface energies for FCC metals." en. In: *Applied Surface Science* 256.21 (Aug. 2010), pp. 6345–6349. ISSN: 01694332. DOI: 10.1016/j.apsusc.2010.04.014 (cit. on p. 59).

[186] A J C Wilson and H Lipson. "The calibration of Debye-Scherrer X-ray powder cameras." In: *Proceedings of the Physical Society* 53.3 (May 1941), pp. 245–250. ISSN: 0959-5309. DOI: 10.1088/0959-5309/53/3/306 (cit. on p. 59).

[187] S. Pogatscher et al. "Mechanisms controlling the artificial aging of Al-Mg-Si Alloys." In: *Acta Materialia* 59.9 (May 2011), pp. 3352–3363. DOI: 10.1016/j.actamat.2011.02.010 (cit. on pp. 64, 97).

[188] HW Zandbergen, SJ Andersen, and J Jansen. "Structure determination of Mg5Si6 particles in Al by dynamic electron diffraction studies." In: *Science* 277.5330 (1997), pp. 1221–1225 (cit. on p. 65).

[189] Jörg Behler. "Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations." In: *PCCP* 13 (2011), pp. 17930–55 (cit. on p. 68).

[190] K. V Jovan Jose, Nongnuch Artrith, and Jörg Behler. "Construction of high-dimensional neural network potentials using environment-dependent atom pairs." In: *J. Chem. Phys.* 136 (2012), p. 194111 (cit. on p. 68).

[191] Rustam Z Khaliullin et al. "Graphite-diamond phase coexistence study employing a neural-network mapping of the ab initio potential energy surface." In: *Phys. Rev. B* 81 (2010), p. 100103 (cit. on p. 68).

[192] Gabriele C. Sosso et al. "Neural network interatomic potential for the phase change material GeTe." In: *Phys. Rev. B* 85 (2012), p. 174103 (cit. on p. 68).

[193] Hagai Eshet et al. "Microscopic Origins of the Anomalous Melting Behavior of Sodium under High Pressure." In: *Phys. Rev. Lett.* 108 (2012), p. 115701 (cit. on p. 68).

[194] Nongnuch Artrith and Jörg Behler. "High-dimensional neural network potentials for metal surfaces: A prototype study for copper." In: *Phys. Rev. B* 85 (2012), p. 045439 (cit. on p. 68).

[195] Venkat Kapil, Jörg Behler, and Michele Ceriotti. "High order path integrals made easy." In: *J. Chem. Phys.* 145 (2016), p. 234103 (cit. on p. 68).

[196] Bingqing Cheng, Jörg Behler, and Michele Ceriotti. "Nuclear Quantum Effects in Water at the Triple Point: Using Theory as a Link Between Experiments." In: *J. Phys. Chem. Letters* 7 (2016), pp. 2210–2215 (cit. on p. 68).

[197]   Michael Gastegger et al. "wACSF—Weighted atom-centered symmetry functions as descriptors in machine learning potentials." In: *The Journal of Chemical Physics* 148.24 (2018), p. 241709 (cit. on p. 70).

[198]   Nicholas J Browning et al. "Genetic Optimization of Training Sets for Improved Machine Learning Models of Molecular Properties." In: *J. Phys. Chem. Letters* 8 (2017), pp. 1351–1359 (cit. on p. 70).

[199]   Justin S Smith, Olexandr Isayev, and Adrian E Roitberg. "ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost." In: *Chemical science* 8.4 (2017), pp. 3192–3203 (cit. on p. 71).

[200]   Felix A. Faber et al. "Prediction Errors of Molecular Machine Learning Models Lower than Hybrid DFT Error." In: *J. Chem. Theory Comput.* (2017), acs.jctc.7b00577 (cit. on p. 71).

[201]   Michael W Mahoney and Petros Drineas. "CUR matrix decompositions for improved data analysis." In: *Proc. Natl. Acad. Sci. USA* 106 (2009), pp. 697–702 (cit. on p. 71).

[202]   Sandip De et al. "Comparing molecules and solids across structural and alchemical space." In: *Phys. Chem. Chem. Phys.* 18 (2016), pp. 13754–13769 (cit. on p. 72).

[203]   Albert P Bartók et al. "Machine learning unifies the modeling of materials and molecules." In: *Science advances* 3.12 (2017), e1701816 (cit. on p. 72).

[204]   Michele Ceriotti, Gareth A. Tribello, and Michele Parrinello. "Demonstrating the Transferability and the Descriptive Power of Sketch-Map." In: *J. Chem. Theory Comput.* 9 (2013), pp. 1521–1532 (cit. on p. 73).

[205]   G Henkelman and H Jónsson. "A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives." In: *J. Chem. Phys.* 111 (1999), p. 7010 (cit. on p. 75).

[206]   G Henkelman, B P Uberuaga, and H Jónsson. "A climbing image nudged elastic band method for finding saddle points and minimum energy paths." In: *J. Chem. Phys.* 113 (2000), p. 9901 (cit. on p. 75).

[207]   M. Mantina et al. "First principles impurity diffusion coefficients." en. In: *Acta Materialia* 57.14 (Aug. 2009), pp. 4102–4108. ISSN: 13596454. DOI: 10.1016/j.actamat.2009.05.006 (cit. on p. 75).

[208]   A De Vita and M J Gillan. "The ab initio calculation of defect energetics in aluminium." In: *Journal of Physics: Condensed Matter* 3.33 (Aug. 1991), pp. 6225–6237. DOI: 10.1088/0953-8984/3/33/002 (cit. on p. 77).

[209]   Matthias Rupp et al. "Fast and accurate modeling of molecular atomization energies with machine learning." In: *Physical review letters* 108.5 (2012), p. 058301 (cit. on p. 78).

[210]   Grégoire Montavon et al. "Machine learning of molecular electronic properties in chemical compound space." In: *New Journal of Physics* 15.9 (2013), p. 095003 (cit. on p. 78).

[211]   Bryce Meredig et al. "Combinatorial screening for new materials in unconstrained composition space with machine learning." In: *Physical Review B* 89.9 (2014), p. 094104 (cit. on p. 78).

[212]   Patrick Bleiziffer, Kay Schaller, and Sereina Riniker. "Machine Learning of Partial Charges Derived from High-Quality Quantum-Mechanical Calculations." In: *Journal of chemical information and modeling* (2018) (cit. on p. 78).

[213]   Nongnuch Artrith, Tobias Morawietz, and Jörg Behler. "High-dimensional neural-network potentials for multicomponent systems: Applications to zinc oxide." In: *Physical Review B* 83.15 (Apr. 2011). DOI: 10.1103/PhysRevB.83.153101 (cit. on p. 78).

[214]   Celia Fonseca Guerra et al. "Voronoi deformation density (VDD) charges: Assessment of the Mulliken, Bader, Hirshfeld, Weinhold, and VDD methods for charge analysis." In: *Journal of Computational Chemistry* 25.2 (Jan. 2004), pp. 189–210. DOI: 10.1002/jcc.10351 (cit. on pp. 78, 81, 82).

[215]   Fred L Hirshfeld. "Bonded-atom fragments for describing molecular charge densities." In: *Theoretica chimica acta* 44.2 (1977), pp. 129–138 (cit. on pp. 78, 81).

[216]   Gerard Paul M Leyson et al. "Quantitative prediction of solute strengthening in aluminium alloys." In: *Nature materials* 9.9 (2010), p. 750 (cit. on p. 80).

[217]   MPAT Methfessel and AT Paxton. "High-precision sampling for Brillouin-zone integration in metals." In: *Physical Review B* 40.6 (1989), p. 3616 (cit. on p. 80).

[218]   Daniele Dragoni et al. "Achieving DFT accuracy with a machine-learning interatomic potential: Thermomechanics and defects in bcc ferromagnetic iron." In: *Physical Review Materials* 2.1 (2018), p. 013808 (cit. on p. 80).

[219]   N Chetty et al. "Vacancies and impurities in aluminum and magnesium." In: *Physical Review B* 52.9 (1995), p. 6313 (cit. on p. 80).

[220] PT Sprunger et al. "Giant Friedel oscillations on the beryllium (0001) surface." In: *Science* 275.5307 (1997), pp. 1764–1767 (cit. on p. 80).

[221] G Grüner and Csaba Hargitai. "Temperature Dependence of the Charge Oscillation Around Nearly Magnetic Impurities." In: *Physical Review Letters* 26.13 (1971), p. 772 (cit. on p. 80).

[222] G Grüner and M Minier. "Charge perturbations around impurities in metals." In: *Advances in Physics* 26.3 (1977), pp. 231–284 (cit. on p. 80).

[223] B. Rousseau, A. Peeters, and C. Van Alsenoy. "Atomic charges from modified Voronoi polyhedra." en. In: *Journal of Molecular Structure: THEOCHEM* 538.1-3 (Mar. 2001), pp. 235–238. DOI: `10.1016/S0166-1280(00)00692-8` (cit. on pp. 81, 82).

[224] Robert S Mulliken. "Electronic population analysis on LCAO–MO molecular wave functions. I." In: *The Journal of Chemical Physics* 23.10 (1955), pp. 1833–1840 (cit. on p. 81).

[225] Ernest R Davidson. "Electronic population analysis of molecular wavefunctions." In: *The Journal of Chemical Physics* 46.9 (1967), pp. 3320–3324 (cit. on p. 81).

[226] Pieter Ros and GCA Schuit. "Molecular orbital calculations on copper chloride complexes." In: *Theoretica chimica acta* 4.1 (1966), pp. 1–12 (cit. on p. 81).

[227] F Matthias Bickelhaupt et al. "The carbon- lithium electron pair bond in (CH3Li) n (n= 1, 2, 4)." In: *Organometallics* 15.13 (1996), pp. 2923–2931 (cit. on p. 81).

[228] Alan E Reed, Robert B Weinstock, and Frank Weinhold. "Natural population analysis." In: *The Journal of Chemical Physics* 83.2 (1985), pp. 735–746 (cit. on p. 81).

[229] GG Hall. "Atomic charges within molecules." In: *Advances in atomic and molecular physics*. Vol. 20. Elsevier, 1985, pp. 41–63 (cit. on p. 81).

[230] Steven M Bachrach. "Population analysis and electron densities from quantum mechanics." In: *Reviews in Computational Chemistry, Volume 5* (2007), pp. 171–228 (cit. on p. 81).

[231] RFW Bader. *A Quantum Theory, Clarendon.* 1990 (cit. on p. 81).

[232] Peter Politzer and Roger R Harris. "Properties of Atoms in Molecules. I. Proposed Definition of the Charge on an Atom in a Molecule." In: *Journal of the American Chemical Society* 92.22 (1970), pp. 6451–6454 (cit. on pp. 81, 82).

[233]   Georges Voronoi. "Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Deuxième mémoire. Recherches sur les parallélloèdres primitifs." In: *Journal für die reine und angewandte Mathematik* 134 (1908), pp. 198–287 (cit. on p. 82).

[234]   NN Medvedev. "The algorithm for three-dimensional Voronoi polyhedra." In: *Journal of computational physics* 67.1 (1986), pp. 223–229 (cit. on p. 82).

[235]   W Tang, E Sanville, and G Henkelman. "A grid-based Bader analysis algorithm without lattice bias." In: *Journal of Physics: Condensed Matter* 21.8 (2009), p. 084204 (cit. on p. 83).

[236]   Min Yu and Dallas R Trinkle. "Accurate and efficient algorithm for Bader charge integration." In: *The Journal of chemical physics* 134.6 (2011), p. 064111 (cit. on p. 83).

[237]   Graeme Henkelman, Andri Arnaldsson, and Hannes Jónsson. "A fast and robust algorithm for Bader decomposition of charge density." In: *Computational Materials Science* 36.3 (2006), pp. 354–360 (cit. on p. 83).

[238]   Giovanni Bussi, Davide Donadio, and Michele Parrinello. "Canonical sampling through velocity rescaling." In: *The Journal of chemical physics* 126.1 (2007), p. 014101 (cit. on p. 83).

[239]   Liwei Li et al. "On the Relationship between Mo K-Edge Energies and DFT Computed Partial Charges." In: *The Journal of Physical Chemistry C* 117.6 (2013), pp. 2769–2773 (cit. on p. 83).

[240]   Jeroen A Rombouts, Andreas W Ehlers, and Koop Lammertsma. "A quantitative analysis of light-driven charge transfer processes using voronoi partitioning of time dependent DFT-derived electron densities." In: *Journal of computational chemistry* 38.20 (2017), pp. 1811–1818 (cit. on p. 83).

[241]   Yaocen Wang et al. "Atomic packing and diffusion in Fe85Si2B9P4 amorphous alloy analyzed by ab initio molecular dynamics simulation." In: *Journal of Applied Physics* 117.17 (2015), 17B705 (cit. on p. 83).

[242]   Qingguo Feng et al. "The role of hydroxyl groups in interchain interactions in cellulose Iα and Iβ." In: *International Journal of Quantum Chemistry* 117.10 (2017) (cit. on p. 83).

[243]   Giacomo Giorgi et al. "On the dual deuterium/deuteron nature of D charge distribution in the Ti host matrix: A DFT analysis." In: *International Journal of Hydrogen Energy* 38.36 (2013), pp. 16477–16484 (cit. on p. 83).

[244] MV Petrik et al. "Ab initio based analysis of grain boundary segregation in Al-Mg and Al-Zn binary alloys." In: *arXiv preprint arXiv:1708.00211* (2017) (cit. on p. 83).

[245] Osamu Takai et al. "Vacancy-vacancy, vacancy-impurity, and impurity-impurity interactions in aluminum." In: *Physical Review B* 10.8 (Oct. 1974), pp. 3113–3125. DOI: `10.1103/PhysRevB.10.3113` (cit. on p. 83).

[246] N. Castin et al. "Advanced atomistic models for radiation damage in Fe-based alloys: Contributions and future perspectives from artificial neural networks." In: *Computational Materials Science* 148 (June 2018), pp. 116–130. DOI: `10.1016/j.commatsci.2018.02.025` (cit. on p. 89).

[247] N. Castin et al. "Improved atomistic Monte Carlo models based on *ab-initio* -trained neural networks: Application to FeCu and FeCr alloys." In: *Physical Review B* 95.21 (June 2017). DOI: `10.1103/PhysRevB.95.214117` (cit. on p. 89).

[248] Luca Messina et al. "Introducing *ab initio* based neural networks for transition-rate prediction in kinetic Monte Carlo simulations." In: *Physical Review B* 95.6 (Feb. 2017). DOI: `10.1103/PhysRevB.95.064112` (cit. on p. 89).

[249] E. Clouet et al. "Kinetic Monte Carlo Simulations of Precipitation." In: *Advanced Engineering Materials* 8.12 (Dec. 2006), pp. 1210–1214. DOI: `10.1002/adem.200600183` (cit. on p. 89).

[250] David Molnar et al. "Activation energies for nucleation and growth and critical cluster size dependence in JMAK analyses of kinetic Monte-Carlo simulations of precipitation." In: *Continuum Mechanics and Thermodynamics* 24.4-6 (Nov. 2012), pp. 607–617. DOI: `10.1007/s00161-012-0258-5` (cit. on p. 89).

[251] Frédéric Soisson. "Kinetic Monte Carlo simulations of radiation induced segregation and precipitation." In: *Journal of nuclear materials* 349.3 (2006), pp. 235–250 (cit. on p. 89).

[252] Tomoyuki Honma, David W Saxey, and Simon P Ringer. "Effect of trace addition of Sn in Al-Cu alloy." In: *Materials science forum*. Vol. 519. Trans Tech Publ. 2006, pp. 203–208 (cit. on p. 89).

[253] G Sha and A Cerezo. "Kinetic Monte Carlo simulation of clustering in an Al–Zn–Mg–Cu alloy (7050)." In: *Acta materialia* 53.4 (2005), pp. 907–917 (cit. on p. 89).

[254] E Vincent, CS Becquart, and C Domain. "Solute interaction with point defects in α Fe during thermal ageing: A combined ab initio and atomic kinetic Monte Carlo approach." In: *Journal of nuclear materials* 351.1-3 (2006), pp. 88–99 (cit. on p. 89).

[255] JL Bocquet. "Defects and Diffusion in Metals: An Annual Retrospective IV." In: *Trans Tech Publications Inc., Stafa-Zurich, Switzerland* 203 (2002), p. 2 (cit. on p. 89).

[256] H Gupta et al. "Microscopic computer simulations of directional coarsening in face-centered cubic alloys." In: *Acta materialia* 49.1 (2001), pp. 53–63 (cit. on p. 89).

[257] Zeqin Liang et al. "The kinetics of clustering in Al–Mg–Si alloys studied by Monte Carlo simulation." In: *International Journal of Materials Research* 103.8 (2012), pp. 980–986 (cit. on pp. 89, 93).

[258] E Vincent, CS Becquart, and C Domain. "Microstructural evolution under high flux irradiation of dilute Fe–CuNiMnSi alloys studied by an atomic kinetic Monte Carlo model accounting for both vacancies and self interstitials." In: *Journal of Nuclear Materials* 382.2-3 (2008), pp. 154–159 (cit. on p. 89).

[259] Emmanuel Clouet et al. "Complex precipitation pathways in multicomponent alloys." en. In: *Nature Materials* 5.6 (June 2006), pp. 482–488. DOI: 10.1038/nmat1652 (cit. on pp. 89, 93).

[260] Michele Ceriotti, Joshua More, and David E. Manolopoulos. "i-PI: A Python interface for ab initio path integral molecular dynamics simulations." en. In: *Computer Physics Communications* 185.3 (Mar. 2014), pp. 1019–1026. DOI: 10.1016/j.cpc.2013.10.027 (cit. on p. 90).

[261] Steve Plimpton, Paul Crozier, and Aidan Thompson. "LAMMPS-large-scale atomic/molecular massively parallel simulator." In: *Sandia National Laboratories* 18 (2007), pp. 43–43 (cit. on p. 92).

[262] M. Mantina et al. "First principles impurity diffusion coefficients." In: *Acta Materialia* 57.14 (Aug. 2009), pp. 4102–4108. DOI: 10.1016/j.actamat.2009.05.006 (cit. on p. 92).

[263] Masao Doyama. "Vacancy-solute interactions in metals." In: *Journal of Nuclear Materials* 69-70 (Feb. 1978), pp. 350–361. DOI: 10.1016/0022-3115(78)90253-2 (cit. on pp. 92, 93).

[264] Emmanuel Clouet, Maylise Nastar, and Christophe Sigli. "Nucleation of Al 3 Zr and Al 3 Sc in aluminum alloys: From kinetic Monte Carlo simulations to classical theory." In: *Physical Review B* 69.6 (Feb. 2004). DOI: 10.1103/PhysRevB.69.064109 (cit. on p. 93).

[265] M.F. Francis and W.A. Curtin. "Microalloying for the controllable delay of precipitate formation in metal alloys." In: *Acta Materialia* 106 (Mar. 2016), pp. 117–128. DOI: 10.1016/j.actamat.2016.01.014 (cit. on p. 98).

[266] O. C. Zienkiewicz. *The finite element method*. 3d expanded and rev. ed. London - New York: McGraw-Hill, 1977. ISBN: 978-0-07-084072-0 (cit. on p. 102).

[267] Nicolas Richart and Jean-François Molinari. "Implementation of a parallel finite-element library: test case on a non-local continuum damage model." In: *Finite Elements in Analysis and Design* 100 (2015), pp. 41–46 (cit. on p. 102).

[268] Joannis G Ergatoudis. "Isoparametric Finite Elements in two and Three Dimensional Stress Analysis." PhD thesis. University College of Swansea, 1968 (cit. on p. 102).

[269] P.R. Amestoy, I.S. Duff, and J.-Y. L'Excellent. *MUMPS MUltifrontal Massively Parallel Solver Version 2.0*. 1998 (cit. on p. 103).

[270] E Ringe, RP Van Duyne, and LD Marks. "Wulff construction for alloy nanoparticles." In: *Nano letters* 11.8 (2011), pp. 3399–3403 (cit. on p. 108).

[271] Thierry Bodineau. "The Wulff construction in three and more dimensions." In: *Communications in mathematical physics* 207.1 (1999), pp. 197–229 (cit. on p. 108).

CURRICULUM VITAE

Daniele Giofré was born in Martina Franca, Italy, on the 11th of August 1989. He achieved a Chemical diploma at institute of higher education J. Monnet (2008). After that he continued his studies in Physics at the University of Milano-Bicocca, obtaining his bachelor with a dissertation in "*Electron properties of aromatic cyclic molecules*" (2011) under the supervision of Prof. GianPaolo Brivio, and specializing in Solid-State Physics during the master. He got his master cum laude in 2013 discussing a thesis titled "*Ab-initio study of the electric field effects produced by nitrogen and boron dopants on the transport and electronic properties of the bilayer graphene*" supervised by Prof. Dr. Mario Italo Trioni. Since May 2014 he was hired as Ph.D at the department of Materials Science at the École Polytechnique Fédérale de Lausanne in the laboratory of Computational Science and Modelling under the direction of Prof. Michele Ceriotti. His field of research involves studying materials, e.g. alloys, based on atomistic simulations and machine-learning techniques.

LIST OF PUBLICATIONS

1. "Automatic selection of atomic fingerprints and reference configurations for machine-learning potentials", G. Imbalzano, A. Anelli, D. Giofré, S. Klees, J. Behler, M. Ceriotti, The Journal of Chemical Physics **148**, 241730 (2018)

2. "Ab initio modelling of the early stages of precipitation in Al-6000 alloys", D. Giofré, T. Junge, W.A. Curtin, M. Ceriotti, Acta Materialia **140**, 240 (2017)

3. "Neural network potential for Al-Mg-Si alloys", R. Kobayashi, D. Giofré, T. Junge, M. Ceriotti, W.A. Curtin, Physical Review Materials **1**, 053604 (2017)

4. "Electronic transport in BN substituted bilayer graphene nano-junctions", D. Giofré, D. Ceresoli, G. Fratesi, M.I. Trioni, Physical Review B **93**, 205420 (2016)

5. "Thermally-nucleated self-assembly of water and alcohol into stable structures at hydrophobic interfaces", (2016) K. Voïtchovsky, D. Giofré, J.J. Segura, F. Stellacci, and M. Ceriotti, Nature Communications **7** (2016)

ACKNOWLEDGEMENTS

momenti più bui, per avermi insegnato che le avversità si superano concentrandosi sulla soluzione e non sul problema e per rendermi semplicemente felice. A te, più di tutti, che hai corso mano nella mano con me questa gara posso dire che tale traguardo é tuo, quanto mio. Grazie!

Thank you very much.

Daniele Giofré
Lausanne, July 31, 2018