UNIVERSITY OF HELSINKI REPORT SERIES IN PHYSICS

HU-P-D283

Atomistic Simulations of Self-diffusion on Copper Surfaces with and without an Applied Electric Field

Jyri Kimari

Helsinki Institute of Physics
Helsinki Accelerator Laboratory
Division of Materials Physics
Department of Physics
Faculty of Science
University of Helsinki
Helsinki, Finland

DOCTORAL DISSERTATION

To be presented for public discussion with the permission of the Faculty of Science of University of Helsinki, in auditorium U3032, main building of the University of Helsinki, on the 20th of January, 2023, at 1 p.m.

HELSINKI 2023

ISBN 978-951-51-8885-4 (printed version)
ISSN 0356-0961
Helsinki 2023
Unigrafia

ISBN 978-951-51-8886-1 (PDF version)
https://ethesis.helsinki.fi
Helsinki 2023

Jyri Kimari, **Atomistic Simulations of Self-diffusion on Copper Surfaces with and without an Applied Electric Field**, University of Helsinki, 2023, 73 p., University of Helsinki Report Series in Physics, HU-P-D283, ISSN 0356-0961, ISBN 978-951-51-8885-4 (printed version), ISBN 978-951-51-8886-1 (PDF version)

God made the bulk; surfaces
were invented by the devil.
Wolfgang Pauli

Abstract

Surface diffusion is the major process that can drive metal surface modification in temperatures significantly below the melting point. Diffusion is random movement with the tendency toward minimization of the total free energy. In the absence of external forces, the equilibrium shape of the metal surface is determined by the anisotropy of the surface energies in different orientations. At finite temperatures, anisotropy also in the temperature dependence of the surface free energies, together with the temperature and orientation dependent kinetic rates of surface diffusion events, make metal surface morphology a multifaceted phenomenon involving a range of competing effects.

Computational modeling of surface diffusion process is thus challenging in itself, but the task is even further complicated if the surface is subject to external factors. One such factor can be an applied electric field, the electrostatic contribution to the free energy of which alters the equilibrium shape. This is hypothesized to be one of the contributors to the development of surface roughening that further develops into field emitting tips. The latter, in turn, are thought to evolve into vacuum arc breakdowns, occurring in many devices operating in high electric fields.

In this work, surface diffusion in Cu with and without electric fields was studied by kinetic Monte Carlo (KMC) and molecular dynamics (MD). A multitude of encountered obstacles are described, from calculating KMC parameters in the highly heterogeneous surface environments, to the challenge of reaching sufficiently long time scales in MD.

The pre-parameterized on-lattice KMC model, chosen for efficiency, was improved by employing machine learning for migration barrier prediction. Better accuracy regarding the stability of small nanostructures on differently oriented surfaces was reached this way, compared to the earlier KMC model. The new model also correctly predicts the relative surface energies in Cu and the fragmentation of crossing nanowires at the junction points. The Cu $\{1\,1\,0\}$ surface also exhibits a self-roughening instability at elevated temperatures, commonly known to occur in this system both in experiments and MD simulations.

For simulations in electric fields, we utilized the coupling of MD to a continuum, finite elements method solver of electric fields and surface charges that was earlier developed in our group. This method enables the imposition of a realistic electric field gradient on the MD system, while keeping the number of dynamically simulated atoms very low for improved efficiency. Furthermore, the MD simulations on the surface orientations with the highest migration energy barriers were accelerated by a metadynamics-based method. As a result, we were able to capture the qualitative trend of biased diffusion toward higher electric fields, with the possibility of extracting surface polarization parameters from the MD results to be directly compared to *ab initio* calculations.

Tiivistelmä

Pintadiffuusio on merkittävin metallipintojen muotoon vaikuttava mekanismi huomattavasti sulamispistettä alhaisemmissa lämpötiloissa. Diffuusio on satunnaisliikettä, joka ajautuu kohti vapaan energian minimoimista. Ulkoisten voimien poissaollessa metallien pinnan muotojen tasapainotilan määrittää erisuuntaisten pintojen pintaenergioiden anisotropia. Äärellisissä lämpötiloissa myös pintojen vapaan energian sekä pintadiffuusion kinetiikan anisotropia tekevät metallipintojen morfologiasta monitahoisen, kilpailevien vaikutusten muodostaman ilmiön.

Pintadiffuusion laskennallinen mallintaminen on täten jo itsessään haastavaa, mutta tehtävä on vieläkin monimutkaisempi, jos pintaan vaikuttaa jokin ulkoinen tekijä. Yksi tällainen tekijä voi olla sähkökenttä, jonka vaikutus pintojen vapaaseen energiaan muuttaa tasapainotilan muotoa. Tämän muodon muuttumisen arvellaan olevan yksi ilmiöistä, jotka voivat johtaa metallipintojen karhenemiseen, ja edelleen kenttäemittoivien nanopilarien kasvamiseen. Pilarien puolestaan arvellaan voivan kehittyä tyhjiövalokaariläpilyönneiksi, joita tapahtuu monissa voimakasta sähkökenttää hyödyntävissä laitteissa.

Tässä työssä tutkittiin diffuusiota kuparipinnalla ilman sähkökenttää ja sen kanssa, käyttäen kineettistä Monte Carloa (KMC) ja molekyylidynamiikkaa (MD). Väitöskirjassa kuvailen monia kohdattuja haasteita, mukaanlukien KMC-simulaatioiden parametrien laskemisen erittäin heterogeenisissä atomiympäristöissä, ja riittävän pitken simulaatioaikojen saavuttamisen MD-simulaatioissa.

Ennakkoparametrisoitua hilapohjaista KMC-mallia, joka valittiin sen tehokkuuden vuoksi, paranneltiin käyttämällä koneoppimista migraatioenergiavallien ennustamiseen. Tällä tavalla nanorakenteiden vakautta erisuuntaisilla pinnoilla pystyttiin mallintamaan tarkemmin verrattuna aiempaan malliin. Tämä paranneltu malli myös ennustaa oikein erisuuntaisten pintojen suhteelliset pintaenergiat, sekä risteävien nanojohtojen murtumisen niiden risteyskohdissa. Tässä mallissa kupari-$\{1\,1\,0\}$-pinta myös karhenee korkeissa lämpötiloissa, minkä on havaittu tapahtuvan sekä kokeissa että aiemmissa MD-simulaatioissa.

Simulaatioihin sähkökentässä käytettiin ryhmässämme aiemmin kehitettyä MD-mallin kytkentää elementtimenetelmään, joka ratkaisee sähkökentän ja pintavaraukset jatkumossa. Tämä menetelmä mahdollistaa realistisen sähkökenttägradientin luomisen MD-systeemiin, samalla pitäen tarvittavien atomien määrän vähäisenä ja simulaation siten tehokkaana. Elementtimenetelmän lisäksi käytimme myös metadynamiikkapohjaista kiihdytystä pinnoilla, joilla migraatioenergiavallit ovat korkeimpia. Täten pystyimme kvalitatiivisesti mallintamaan diffuusion taipumusta suuntautua kohti voimakkaampia sähkökenttiä, sekä laskemaan pintapolarisaatiota kuvaavia suureita, joita voidaan suoraan vertailla *ab initio* -tulosten kanssa.

Contents

List of abbreviations

1nn	first-nearest neighbor
2nn	second-nearest neighbor
AI	artificial intelligence
ANN	artificial neural network
bcc	body-centered cubic
CLIC	Compact Linear Collider
CV	collective variable
CVHD	collective variable -driven hyperdynamics
DFT	density functional theory
fcc	face-centered cubic
FEM	finite elements method
hcp	hexagonal close-packed
KMC	kinetic Monte Carlo
LAE	local atomic environment
MD	molecular dynamics
MEP	minimum energy path
ML	machine learning
MLP	multilayer perceptron
NEB	nudged elastic band
RBFN	radial basis function network

Chapter 1

Introduction

Vacuum arc breakdowns—strong currents through nothing—have puzzled physicists since the dawn of electronic devices. The earliest observations of arcing in vacuum are from, not only more than one, but more than two hundred years ago [1]. In 1802, the vacuum was probably not exactly ultrahigh, but the sparks were not only due to contamination: despite the developments in vacuum technology since the Industrial Revolution, current spikes at sufficiently high electric fields have not stopped occurring.

From what experimental data can be collected under the conditions of vacuum arcs, it is evident that the current is carried by plasma. However, starting from the material origin of the plasma, all the way back to the onset of increased field emission current prior to breakdown, the details of the phenomenon are still largely shrouded in mystery.

Breakdowns plague a myriad of devices operating in high electric fields, ranging from particle accelerators [2] to free-electron lasers [3] and fusion reactors [4].

Experimental mysteries are opportunities for us in the computational modeling community to prove our worth. Having had relatively little time to develop, only some 70 years, the field of computer simulations in physics has already grown to encompass scales from singularly accurate description of the smallest molecules, to continuum models of the evolution of the universe [5].

The hypothesized sequence of events of vacuum arcs are, nevertheless, not trivial to model. Different phenomena playing a role include the response of bulk material to the strain by the electric field; the morphology of the interface between matter and vacuum, making for the minimum of free energy; the current of electrons removed from the surface; the thermal effects in the material conducting the current, eventually melting and evaporating; and, finally, the dance of particles thrown into the vacuum, developing into a conductive plasma. In order to study these processes, it is necessary to employ simulation techniques down from the quantum mechanical level, to intermediate, semi-empirical descriptions of atomic interactions, up to modeling the bulk, the surfaces, and the vacuum as continuous objects. The thesis at hand considers the simulation of surface effects

in the intermediate scales, utilizing both kinetic Monte Carlo (KMC) and molecular dynamics (MD), as well as coupling continuous finite elements method (FEM) calculations to the former.

A key mechanism of surface modification in solid materials is the self-diffusion of atoms. Known to be biased toward stronger electric fields, diffusive mass transport could contribute to formation of surface asperities, which would explain the observation of enhanced field emission currents.

The purpose of this doctoral dissertation work is to develop a more sophisticated tool for modeling diffusion on an arbitrarily oriented, arbitrarily rough surface in high electric fields. The included Publications **I** and **II** investigate the challenges and solutions to parametrizing KMC simulations of surface diffusion. In Publications **III**–**V**, artificial intelligence (AI) techniques to augment surface KMC simulations are developed. Finally, in Publication **VI**, the different approach of MD modeling is taken to surface diffusion under electric field.

Chapter 2 outlines the content of the publications and the author's contributions to them. Chapter 3 covers our understanding of vacuum arcs in greater detail, and takes the first look at the basics of the interactions between the electric field and metallic surfaces, and chapter 4 treats surface diffusion on a more extended theoretical level. Chapter 5 introduces the range of methods and techniques used in this work, and chapter 6 summarizes the results.

Chapter 2

Purpose and structure

2.1 Summaries of the original publications

Publication I: Migration barriers for surface diffusion on a rigid lattice: challenges and solutions

Baibuz, E., Vigonski, S., **Lahtinen, J.**, Zhao, J., Jansson, V., Zadin, V., and Djurabekova, F., *Computational Materials Science* **146**, 287–302 (2018).

In this publication, we discussed various challenges related to calculating migration barriers for atomic jumps in lattice kinetic Monte Carlo: concerted movement, unstable lattice sites and migration paths, and spontaneous processes. We proposed many solutions for these challenges, including an extended parametrization scheme.

Publication II: Data sets of migration barriers for atomistic Kinetic Monte Carlo simulations of Cu self-diffusion via first nearest neighbour atomic jumps

Baibuz, E., Vigonski, S., **Lahtinen, J.**, Zhao, J., Jansson, V., Zadin, V., and Djurabekova, F., *Data in Brief* **17**, 739–743 (2018).

This publication is parallel to Publication **I**, describing the Cu migration barrier sets calculated in the work. The purpose of this brief publication format is to make it easier for others to efficiently use the open data in their own research.

Publication III: Application of artificial neural networks for rigid lattice kinetic Monte Carlo studies of Cu surface diffusion

Kimari, J., Jansson, V., Vigonski, S., Baibuz, E., Domingos, R., Zadin, V., and Djurabekova, F., *Computational Materials Science* **183**, 109789 (2020).

Here we applied artificial neural networks—a machine learning method—to estimate Cu surface migration barriers in kinetic Monte Carlo simulations. The performance of the machine learning parametrization scheme was tested in simulations of small nanotip flattening, nanoparticle relaxation, and crossing nanowires, with reasonable accuracy achieved in each case.

Publication IV: Data sets and trained neural networks for Cu migration barriers

Kimari, J., Jansson, V., Vigonski, S., Baibuz, E., Domingos, R., Zadin, V., and Djurabekova, F., *Data in Brief* **32**, 106094 (2020).

This brief publication is parallel to Publication **III**, describing the training data sets and the trained neural networks and instructing how to use them.

Publication V: Gradient-based training and pruning of radial basis function networks with an application in materials physics

Määttä, J., Bazaliy, V., **Kimari, J.**, Djurabekova, F., Nordlund, K., and Roos, T., *Neural Networks* **133**, 123-131 (2021).

In this work, we proposed a novel way of training and pruning radial basis function networks, achieving a good approximation of the full model with a lower execution time. We applied this method to the Cu migration training data sets of Publications **III**–**IV**, reaching good accuracy. Furthermore, the radial basis function model found physically reasonable patterns in the training data, that could guide further development in describing local atomic environments for machine learning.

Publication VI: Biased self-diffusion on Cu surface due to electric field gradients

Kimari, J., Wang, y., Kyritsakis, A., Zadin, V. and Djurabekova, F. *Journal of Physics D: Applied Physics* **55**.46, 465302 (2022).

We studied Cu adatom migration under electric field gradient with molecular dynamics, and observed diffusion to be biased toward stronger electric fields. Notably, the classical simulation with electrostatic modification conformed surprisingly well to the quantum mechanical model of the Cu surface. The mechanism of biased diffusion can contribute to field enhancer sharpening, leading to higher field emission and possibly to runaway effects all the way to vacuum arc breakdowns.

2.2 Author's contribution

The author conducted the simulation and analysis work, except for the quantum mechanical calculations, for Publications **III**, **IV** and **VI**, and also prepared the first versions of the manuscripts for those publications. For Publications **I** and **II**, the author calculated the extended parametrization data set and conducted the analysis to compare it to the standard parametrization, and wrote the corresponding sections of the manuscript. In Publication **V**, the author acted as the materials science expert in the otherwise computer science oriented team, and wrote all materials science related sections of the manuscript.

Chapter 3

Current understanding of the vacuum arcing phenomenon

Vacuum arcs, or vacuum arc breakdowns, are bursts of current through vacuum in very high electric fields, carried by plasma. The phenomenon is not well understood despite more than a century of research. Open questions include the origin of the particles required for plasma, the mechanism causing enhanced field emission prior to breakdowns, and generally the evolution of the surface and bulk matter in the high electric field.

In this chapter, I will outline the fundamentals of vacuum arcs, including the devices where they occur, the known facts about the phenomenon, and some hypotheses regarding their formation. In the last section of the chapter, I will focus on the interaction between the high electric field and the metallic surface, the condition preceding breakdowns.

3.1 Application of high electric fields in scientific and industrial installations

High electric fields appear in many research and industrial applications. A classic example are linear particle accelerators, that, for sufficient particle energies, need to operate at higher accelerating voltages than their circular counterparts. For instance, in the proposed Compact Linear Collider (CLIC) [2], the nominal voltage gradient in the accelerating electromagnetic waves is of $100\,\mathrm{MVm}^{-1}$, corresponding to fields larger than $200\,\mathrm{MVm}^{-1}$ at the surfaces encasing the particle beam. Much research has been dedicated to the study of vacuum arcs in relation to the CLIC device, with the objective of reducing the occurrence of breakdowns below $3 \times 10^{-7}\,\mathrm{m}^{-1}$ per pulse. The bulk of the investigations have been carried out for Cu, the material of choice for CLIC due to its high thermal and electrical conductivity and relatively low cost.

Also the linear accelerators in medical applications (cancer treatment) suffer

from fairly frequent vacuum arcs [6]. Other devices prone to breakdowns include free-electron lasers [3] and fusion reactors [4].

Power loss caused by frequent vacuum arcs is not the only undesirable effect of this phenomenon. Often, metal surfaces that are exposed to electric fields become damaged because of plasma surface interaction. Post mortem surfaces are littered with craters. Additionally, depending on the application, the breakdowns can have instantaneous detrimental effects, such as the diversion of accelerated particle beams [7] or contamination of the fusion reactor plasma [8].

Vacuum arcs are not always unwanted. They are utilized for a number of applications such as vacuum interrupters [9], ion sources [10], and surface deposition [11]. The design of such techniques stands to gain from a deeper understanding of the vacuum arc phenomenon for better control over them, further motivating breakdown research.

3.2 State-of-the-art understanding of the vacuum arcing onset

Some of the direct observations regarding vacuum arc breakdowns are:

- Increased current prior to breakdown due to enhanced field emission [12].
- Voltage drop and current spike, as plasma shortcuts the vacuum gap.
- An increase in pressure in the vacuum chamber [13, 14].
- Glow emanating first from the cathode side, then also from the anode [15].
- Post mortem, the surface is covered in craters.

Related indirect experimental data include the distribution of the breakdown spots, the distribution of the breakdown events in time, the *conditioning curves* of the electrodes subjected to breakdowns, and the varying breakdown voltages in different materials. Metals ranging from Al to stainless steel, C, Cu alloys [16] and Cu of different hardness levels, plasma-cleaned Cu, have been investigated. Conditioning of the electrodes refers to pulsing gradually higher voltages between them, while keeping the breakdown rate roughly constant. The procedure yields a conditioning curve, i.e., the breakdown voltage as a function of the number of pulses.

Generally, the trend of breakdown voltages suggests metals with the hexagonal close-packed (hcp) structure to be the most resistant to breaking, and face-centered cubic (fcc) structures the least. Hard Cu (with smaller polycrystalline grain size) performs better than soft Cu [12, 17]. These observations suggest that dislocation activity plays role in the phenomenon, since they are trapped by grain boundaries, and generally less mobile in hcp materials than fcc [16]. Models around the dislocation activity and a hypothesized catastrophic dislocation proliferation have been developed [18].

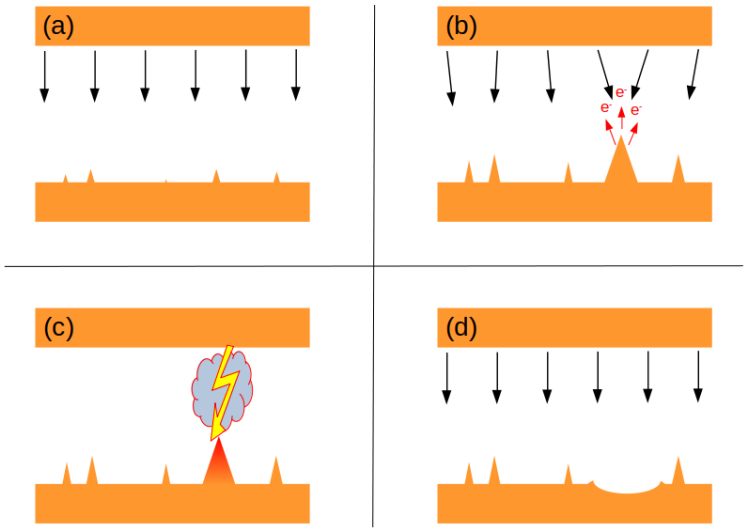


Figure 3.1: Hypothesized evolution of a breakdown. (a): The surface under electric field has nuclei of field emitters present. (b): A dominant field emitter grows taller and enhances the field and field emission locally. (c) From the particles produced in field-assisted evaporation and ionized by electron current, a plasma develops and conducts current from anode to cathode. (d) After the plasma is extinguished, a crater remains where the field emitter was.

Concluding from this wealth of experimental evidence, the following hypothesis for the sequence of events in a breakdown has been proposed. The initial material in the vacuum originates in the cathode, which is heated up and eventually evaporated due to the increased field emission current—an idea first put forth by Ahearn almost 90 years ago [19, 20]. The reason for enhanced field emission is not known, but one possible explanation are geometrical field emitters, i.e., protruding structures that grow under the field [21, 22]. After evaporation, the particles in the vacuum are ionized by the field emitted electrons, developing into plasma that allows current to conduct [23]. The full process is illustrated in Fig. 3.1.

The issue of directly testing this hypothesis experimentally lies in the short time and length scales. Not much can be seen before macroscopic plasma illuminates the scene. If tall field emitting protrusions are there, they grow and wither too fast to be seen under microscope after the vacuum is opened; atomistic simulations suggest that the time scales associated with the flattening of nanoscale features on Cu surfaces range from microseconds to hours [24, 25].

The study of these mechanisms are thus largely in the domain of computational modeling. Each hypothesized step can be modeled by at least a few different techniques, some methods able to span multiple steps. Dislocation activity underneath the surface is accessible, e.g., by discrete dislocation dynamics or rate equation based models of dislocation density [18]; modification of the surface under the field can be described by a range of continuum and atomistic techniques [26–28]; thermal runaway was quite recently captured in MD by Kyrit-

sakis et al. [29]; and, finally, the interplay of electrons and ions in the vacuum can be simulated with particle-in-cell codes [23].

The scope of this thesis with respect to the vacuum arc breakdown phenomenon is limited to the surface effects, not extending down into the bulk region nor much to the vacuum above. This does not make the task too easy, since, as Wolfgang Pauli knew, surfaces were invented by the devil.

3.3 Interaction of metal surface with an electric field

In this section I will introduce some necessary theory regarding macroscopic metal surfaces under electric fields, before moving onto the atomic level in the next chapter.

Metallic systems are characterized by freely moving electrons. When a metal object is in an electric field, it will polarize perfectly: negative charges accumulate on one side, leaving the other side net positive, with the field between these charges exactly canceling the external field inside the object. Because the charges always equilibrate, there can be no voltage gradient along the surface of the metal—if there were, charges would move across the voltage to again balance it. The metallic surface is an electrostatic equipotential surface, with the local electric field everywhere directed perpendicular to it.

Among the direct effects taking place on metal surfaces under electric field is the Maxwell stress [30]. The field applies a stress onto the charged surfaces equal to

$$\sigma = \frac{\varepsilon_0}{2} F^2, \tag{3.1}$$

where ε_0 is the vacuum permittivity and F is the magnitude of the electric field. This stress is always tensile, i.e., away from the surface, since the charge was brought onto the surface by the field in the first place; the sign of the charge is such that the interaction with the field is attractive. The cathode surface is the negatively charged one, and the anode surface is positive.

On the cathode side, the field increases the probability of the accumulated electrons to tunnel out of the surface. This effect is known as electron field emission. The current density of the emitted electrons can be modeled with the Fowler-Nordheim equation [31]

$$I = \kappa_1 \left(\beta F_0\right)^2 \exp\left(-\frac{\kappa_2}{\beta F_0}\right), \tag{3.2}$$

where β is the local field enhancement factor and F_0 is the applied electric field. The material dependent functions κ_1 and κ_2 are not dependent on the applied field nor the enhancement factor, and are thus not necessary to delve into here. The essential dependency of current density is upon the field enhancement factor: by measuring field emission current, this factor can be estimated. For Cu, β associated with breakdowns ranges from 30 to as high as 140 [13, 17, 32].

As mentioned earlier, one possible explanation for the field enhancement factor is geometrical: a protrusion on the surface can enhance the field by a factor of β approximately equal to its aspect ratio, h/r [20, 33]. Connecting the dots to the measured field enhancement values conjures an image of monstrously tall and thin nanotips on the surface. In this thesis, I focus on surface diffusion as a possible mechanism of growth of such protrusions.

Chapter 4

Surface diffusion in metals

This chapter introduces the theoretical background of metallic systems that were used in this thesis. Starting from the crystal structures of pure metals and the properties of differently oriented surfaces, I will continue to describe surface diffusion as a path to surface energy minimization. Finally, I will discuss the theory of biased surface diffusion under electric field, a core topic of the work.

4.1 Metallic crystal structure and surfaces

Crystalline systems are defined by a periodic lattice structure and an atomic basis (locations of atoms) within the unit cells of the lattice. Unit cell is the repeating structure that fills space, not uniquely defined: a given crystal can be equivalently described with any integer multiple of a smaller unit cell, for example. The unit cell that has the smallest possible volume is called the *primitive* unit cell.

There exist a finite number of distinct lattice types: five in two dimensions, and 14 in three dimensions [34]. The lattices are defined by lattice vectors, equal in number to the dimensionality of the space. Only two of the three-dimensional ones, appearing commonly in metal systems, are treated here: the cubic and the hexagonal lattice.

4.1.1 Lattice vectors and directions

In cubic lattice, the three lattice vectors can be defined as

$$\mathbf{a}_1 = a\hat{\mathbf{x}} \tag{4.1}$$

$$\mathbf{a}_2 = a\hat{\mathbf{y}} \tag{4.2}$$

$$\mathbf{a}_3 = a\hat{\mathbf{z}}, \tag{4.3}$$

where a is called the lattice constant, the side length of the chosen unit cell. The simplest atomic basis in the cubic lattice is one atom located in the vertex of the cell (or, equivalently, anywhere else). See Fig. 4.1a for illustration. This is called

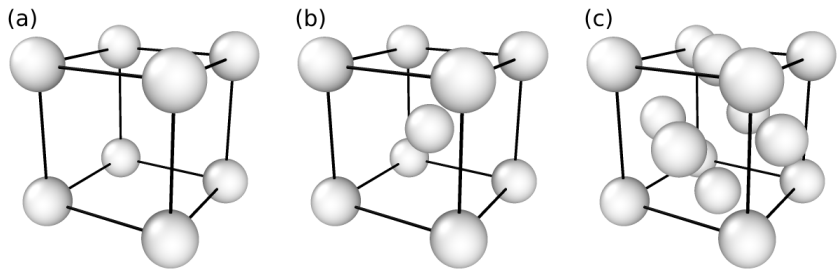


Figure 4.1: Conventional unit cells of the three cubic crystal structures. From left to right: simple cubic, bcc, and fcc.

the *simple cubic* lattice, common in ionic compounds such as NaCl—although, Na^+ and Cl^- alternate between primitive unit cells, and hence the structure would be more rigorously defined with a twice larger tetragonal unit cell with two atom basis.

The simple cubic system is rather unstable, and hence rare in monoelemental metal crystals, due to its low packing fraction. If space is filled with touching spheres in simple cubic manner, the spheres will occupy a fraction of $\frac{4\pi}{24} \approx 0.52$ of the total volume. The two cubic crystals reaching a higher packing fraction are the *body-centered cubic* (bcc) and the *face-centered cubic* (fcc), with the volume of touching spheres equal to 0.68 and 0.74 of the total volume, respectively. The packing fraction of fcc is the largest possible in three dimensions. The structures are illustrated in Fig. 4.1b–c.

The third common crystal structure for metals is the hexagonal close-packed (hcp). The unit cell can be defined as

$$\mathbf{a}_1 = a\left(\frac{1}{2}\hat{\mathbf{x}} + \frac{\sqrt{3}}{2}\hat{\mathbf{y}}\right) \tag{4.4}$$

$$\mathbf{a}_2 = a\left(\frac{1}{2}\hat{\mathbf{x}} - \frac{\sqrt{3}}{2}\hat{\mathbf{y}}\right) \tag{4.5}$$

$$\mathbf{a}_3 = c\hat{\mathbf{z}}, \tag{4.6}$$

with a two atom basis at $\mathbf{0}$ and $\frac{2}{3}\mathbf{a}_1 + \frac{1}{3}\mathbf{a}_2 + \frac{1}{2}\mathbf{a}_3$. For an ideal hcp crystal, $c = \frac{2\sqrt{2}}{3}a$, giving the same packing fraction as the fcc structure.

The hcp and the fcc structures share other similarities. They both consist of two-dimensional hexagonal close-packed atomic layers, differing only in the stacking order: in hcp, the atomic positions alternate every second layer, and in fcc, every third.

The directions and planes in crystals are often denoted in shorthand based on the vector components in the unit cell basis using Miller indices [34]. The components are scaled into integers: for instance, the direction along the line from the origin of the cubic lattice to the center of the opposing face perpendicular to the x-direction, i.e., to the point $\mathbf{a}_1 + \frac{1}{2}\mathbf{a}_2 + \frac{1}{2}\mathbf{a}_3$, is denoted as [2 1 1]. The

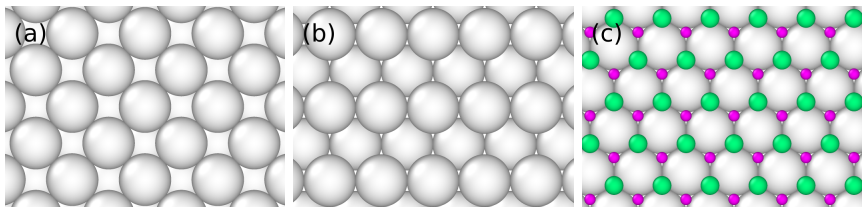


Figure 4.2: Lowest-index surfaces in the fcc crystal. From left to right: {1 0 0}, {1 1 0}, and {1 1 1}. On the {1 1 1} surface, the triangular positions for the next fcc layer are shown in small green spheres, and the positions for an hcp layer (stacking fault) are shown in smaller, purple spheres.

plane perpendicular to this direction is denoted as (2 1 1). Families of symmetrically equivalent directions, i.e. [2 1 1], [1 2 1], [1 1 2], etc., are marked in angled brackets: ⟨2 1 1⟩, and correspondingly the equivalent planes in braces: {2 1 1}. A negative component is marked with a line above the corresponding index, e.g., $[2\,\bar{1}\,1]$.

4.1.2 Surface properties of face-centered cubic structures

The major part of this work considers Cu, which exhibits the fcc crystal structure. Surface diffusion studies here concentrate on the three *lowest-index* surfaces, i.e., the {1 0 0}, the {1 1 0}, and the {1 1 1} surface (see Fig. 4.2). Looking at the fcc unit cell in Fig. 4.1c, these surfaces are sliced along a face of the cube, along the plane between two opposing edges, and along the plane between two opposing vertices, respectively. In Fig. 4.2c, the fcc sites and the off-lattice "hcp sites" in the next atomic layer are marked on top of the surface.

Surface energy is defined as

$$\sigma = \frac{E_A - E_{\text{bulk}}}{A}, \tag{4.7}$$

where E_A is the total energy of a system with an open surface of area A, and E_{bulk} is the total energy of a system of a corresponding size but without open surfaces. In crystalline matter, the surface energies can be highly anisotropic, i.e., the energies between surface orientations differ. As a general rule, the relative surface energies between different surfaces in a given material can be estimated based on the number of missing neighbors each atom has. In fcc bulk, the maximum number of first nearest neighbors (1nn) around each atom is 12. On the {1 1 1} surface, each surface atom has three missing neighbors, on the {1 0 0}, four, and on the {1 1 0}, five in the top layer and one in the layer below. Generally, the order of surface energies is $\sigma_{\{1\,1\,1\}} < \sigma_{\{1\,0\,0\}} < \sigma_{\{1\,1\,0\}}$. The {3 3 1} orientation is another low-energy surface, often even lower than the {1 1 0} [35]. The {3 1 1} surface also has surface energy almost equal to {1 1 0}.

When the metal bulk is cut to expose surface, the surface structure distorts compared to the bulk equilibrium structure to reduce energy. The top layer generally contracts inward, so that the top atoms can compensate for the reduced

electronic density due to the lost neighbors. Correspondingly, the second layer expands to compensate for the increased density due to the top layer contraction [35].

Aside from contraction and expansion of surface layers, more significant reconstruction is sometimes observed. In fcc metals, the most common one is the (1×2) reconstruction of the $\{1\,1\,0\}$ surface, i.e., breaking of the surface symmetry so that the periodically repeating unit cell is twice the size of that in the unreconstructed surface. In Cu $\{1\,1\,0\}$, the most probable reconstruction mode is the missing-row reconstruction, where every second atomic row is removed, exposing thin facets of $\{1\,1\,1\}$ surface between [36].

4.2 Diffusion on metal surfaces

Physical systems have the tendency to evolve toward lower total energies. In crystalline surfaces, different surface facets have different energies, as described in Sec. 4.1.2. Thus, low energy surfaces tend to grow at the expense of higher energy ones. The dominant mechanism for surface modification in low temperatures is diffusion, briefly outlined in Sec. 4.2.1. Common phenomena emerging from surface energy minimization are discussed in Sec. 4.2.2.

4.2.1 Surface diffusion

Diffusion in general can mean many different phenomena where the ordering of a system changes via random movement of its constituent parts. A common example is a droplet of ink in a glass of water: the originally segregated system of water and ink becomes mixed when the ink diffuses throughout the glass.

In the case of solid materials, as opposed to liquid and gas, the atoms are mostly constrained in their positions. Even in this case, where no significant flow can occur, the material can diffuse via individual atoms either exchanging their positions, or, in crystalline solids, by the migration of defects, such as additional atoms between lattice sites (interstitials) or empty lattice sites (vacancies).

On the surfaces of crystalline solids the picture is still somewhat different. Extra atoms on top of the surface layer—adatoms—are highly mobile compared to atoms inside the bulk of the material. These adatoms can appear by adsorption or by detaching from the edges of atomic layers, for instance.

Adatoms experience a bumpy 2-dimensional potential energy surface, with stable and metastable binding site valleys separated by ridges and hills (see Fig. 4.3). While hopping between stable sites, the most likely trajectory for the adatom is the minimum energy path (MEP), from the initial state to the final state via the saddle point.

The jump rates can be estimated using the Arrhenius equation:

$$\Gamma = \nu \exp\left(-\frac{E_\mathrm{m}}{k_\mathrm{B}T}\right), \tag{4.8}$$

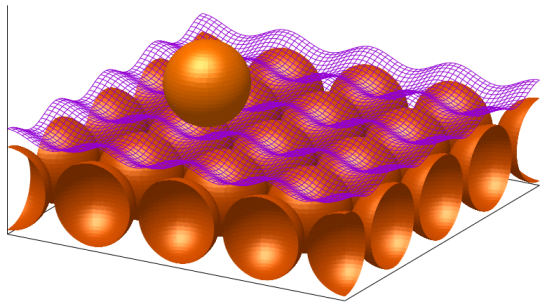


Figure 4.3: Schematic of the 2-dimensional potential energy surface overlaid on the fcc {1 0 0} crystal surface with one adatom. The tetragonal hollow sites above the empty spaces in the surface are stable binding locations for the adatom. The MEP for the adatom goes via the saddle points in the bridge sites.

where ν is called the prefactor or the attempt frequency, E_m is the migration energy barrier, T is the temperature, and k_B is the Boltzmann constant. Eq. (4.8) can be understood to describe the rate of an event that is *attempted* ν times per a unit of time, with the probability of each attempt to succeed depending exponentially on the migration barrier E_m.

The barrier is defined as difference between the energy in the saddle point of the migration event and the energy in its initial state, illustrated in Fig. 4.4. The height of the barrier depends on the material, but also on local atomic environment (LAE) around the jumping atom. If the path of the atom to the target site is partially obstructed by other atoms, the barrier is higher. Likewise, if the adatom is attached to an edge of the next atomic layer, it will have a higher barrier to jump away from the edge. Surface orientation matters as well: adatoms on the fcc {1 0 0} are sunk rather deep in the layer, and have much higher barrier than the adatoms on the {1 1 1} surface that can almost glide on the near-flat potential energy landscape.

4.2.2 Surface energy minimization

The tendency of surface evolution is the minimization of the total surface energy. In the absence of external driving forces, this means that, e.g., any protruding structures on the surface tend to flatten. Small macroscopic particles generally find lower energy in more spherical shapes, minimizing the surface area.

In microscopic scales, the anisotropy of crystalline surface energies becomes relevant. E.g., an fcc system may reach a lower energy by making its {1 1 1} facets larger and {1 1 0} facets smaller, even if the total surface *area* grows. Thus, small nanoparticles tend to facet, favoring low-energy surfaces at the expense of deviating from the minimum area spherical shape. The equilibrium shape minimizes

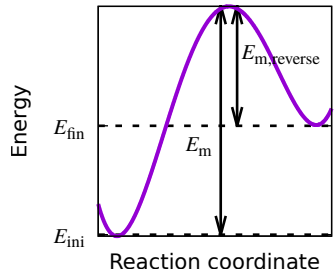


Figure 4.4: Potential energy surface of an atomic migration event between the initial and the final potential energy minimum. E_{ini} is the potential energy in the initial position of the atom, and E_{fin} in the final position of the atom. The migration energy barrier E_{m} is the difference between the saddle point and the initial energy. $E_{\mathrm{m,reverse}}$ is the barrier for the atom returning to the original position. [Publication **III**]

the total surface energy of the N facets of the particle,

$$E_{\mathrm{s}} = \sum_{i}^{N} \sigma_i A_i, \tag{4.9}$$

where σ_i is the surface energy density of facet i, as defined in Eq. 4.7, and A_i is the area of that facet. The equilibrium shape can be found by the Wulff construction, illustrated in Fig. 4.5.

Note that, precisely speaking, the surface energy to be minimized is not only the potential energy U, but the free energy

$$F = U - TS, \tag{4.10}$$

where S is the entropy of the system. In other words, surface energy, and especially the relative energies of different surfaces, are functions of temperature [37, 38]. This manifests in, e.g., the equilibrium shape deforming at finite temperatures [39].

Finally, it should be understood that, while metallic systems generally tend to minimize their total surface energy, the question of time scale to reach the energy minimum is not trivial. In temperatures well below melting point, the main mechanism of surface energy minimization is self-diffusion. As can be seen from Eq. (4.8), the migration rates of diffusion also depend on temperature: surface minimization is a kinetic process, with time scales approaching infinity at very low temperatures. Further, the temperature dependency is not linear, meaning that at low temperatures, the different diffusion rates on different surface orientations become accentuated, while at very high temperatures the rates become similar. This adds more complexity to the evolution of surfaces, with kinetic diffusion rates competing with each other in tandem with the anisotropic surface energies.

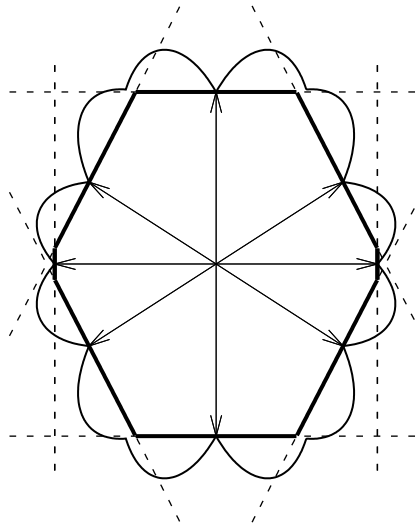


Figure 4.5: Schematic of a two-dimensional Wulff construction, with the surface energy as a function of polar angle in thin solid line. Planes (lines in two dimensions) normal to the radial vectors, intersecting the surface energy curve, are called the Wulff planes. The equilibrium volume (area; in thick solid line) is the one that is enclosed by all Wulff planes. The planes corresponding to the lowest energies, that end up defining the shape in this example, are drawn in dashed lines, along with the associated radial vectors.

4.3 Theory of modified diffusion under electric field

Surface energy is only one component of the total (free) energy of a system. Other contributing energies compete with surface energy minimization, which can result in equilibrium shapes that do not minimize surface free energy.

In the context of this work, the major other energy contributor to be considered is the energy stored in the electric field,

$$E_{\mathrm{C}} = \frac{1}{2} C V^2, \tag{4.11}$$

where C is the capacitance and V is the voltage over the capacitor [30]. By introducing this energy term in the total free energy of a protrusion on a capacitor surface, Fujita and Shimoyama [40] derived an expression for the change in the effective total surface energy:

$$\Delta\sigma = \frac{\varepsilon_0}{2} V F, \tag{4.12}$$

where ε_0 is the vacuum permittivity and F is the total free energy of the capacitor system in the absence of electric field. In other words, the free energy F of the system is reduced by a factor only dependent on the voltage. Crucially, this factor is *iso*tropic with respect to the surface orientations; it imposes a constant reduction of surface energy on the original equilibrium shape of the protrusion

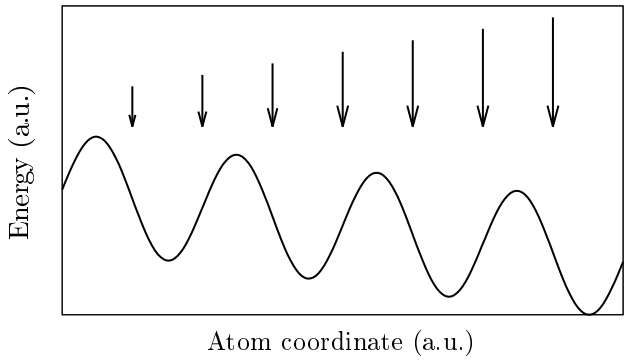


Figure 4.6: Effective potential energy surface experienced by an adatom in inhomogeneous electric field. Arrows indicate the electric field.

given by the Wulff construction, effectively increasing its relative *an*isotropy. According to theory, this will manifest as faceting of the protrusion, giving rise to enhanced field emission. Faceting of nanotips under electric field in W has also been observed in experiments [41].

Ref. [40] considered the protrusion as a continuous object. If the electric field drives a faceting of protrusions, there still needs to be an atomistic mechanism to provide a kinetic path to the free energy minimum. As explained earlier, the major surface minimization mechanism in low temperatures is self-diffusion. Thus, the electric field effect must be somehow coupled to the migration energy barriers that determine diffusion behavior on the microscale.

In 1975, Tsong and Kellogg [42] introduced a theory of biased diffusion under non-uniform electric field. Due to the polarization of the surface atoms, adatoms will experience a tilted potential energy surface with lower barriers toward the higher field (see Fig. 4.6). As charge in conductors accumulates in regions of high curvature, the electric field is strongest there, and thus atomic mass transfer would also be directed toward those regions. As material moves from flat areas to edges and vertices, facets will grow. This can be regarded essentially as the microscopic view of the same phenomenon Fujita and Shimoyama treated in meso/macroscale.

Kyritsakis et al. [43] refined Tsong and Kellogg's theory by considering the polarization of the metal system as a whole, rather than ambiguous dipoles of metal adatoms on the surface. They express the bias of migration barrier in electric field as:

$$\delta E_{\mathrm{m}} = -\mathcal{M}_{\mathrm{sl}} F_0 - \frac{1}{2}\mathcal{A}_{\mathrm{sl}} F_0^2 - \mathcal{M}_{\mathrm{sr}} d\gamma - \mathcal{A}_{\mathrm{sr}} F_0 d\gamma, \tag{4.13}$$

where F_0 is the value of the electric field at the initial lattice site of the jumping adatom, d is the distance between the lattice site and the saddle point, and γ is the electric field gradient, Quantities $\mathcal{M}$ and $\mathcal{A}$ are the permanent dipole moment of the system, and its polarizability, i.e., polarization response to an external

field. Subscripts sl refer to the difference between the saddle point and the lattice site, and sr to the difference between the saddle point and a *reference system* of a flat surface without the adatom: $\mathcal{M}_{\mathrm{sr}} \equiv \mathcal{M}_{\mathrm{s}} - \mathcal{M}_{\mathrm{r}}$, and analogously for the other cases. Kyritsakis et al. calculated the polarization characteristics for the W surface by density functional theory (DFT), and found good agreement between the prediction of Eq. (4.13) and the barriers calculated directly by DFT.

The non-zero permanent dipole moment of an adatom can be attributed to the so-called "Smoluchowski effect" [44], a qualitative explanation involving a smoothing of the electronic density in the vicinity of sharp surface features. More rigorously, the accumulation of positive net charge near the adatom is due to the valence electrons of the adatom shifting closer to the surface when it forms bonds with the substrate atoms [45].

One can derive an expression for the mean "bias rate", i.e., the expected displacement in a unit of time, analytically in the case of one-dimensional migration of an atom under an electric field. With the field gradient toward the left being $-\gamma$, the mean drift rate toward the right is the sum of the rates to the right, Γ_+, and to the left, Γ_-, equal to

$$\Gamma_{\mathrm{b}} = \Gamma_+ - \Gamma_- = \nu \exp\left(-\frac{E_{\mathrm{m}} - \mathcal{M}_{\mathrm{sl}}F_0 - \frac{1}{2}\mathcal{A}_{\mathrm{sl}}F_0^2 - \mathcal{M}_{\mathrm{sr}}d\gamma - \mathcal{A}_{\mathrm{sr}}F_0 d\gamma}{k_{\mathrm{B}}T}\right) \tag{4.14}$$

$$- \nu \exp\left(-\frac{E_{\mathrm{m}} - \mathcal{M}_{\mathrm{sl}}F_0 - \frac{1}{2}\mathcal{A}_{\mathrm{sl}}F_0^2 + \mathcal{M}_{\mathrm{sr}}d\gamma + \mathcal{A}_{\mathrm{sr}}F_0 d\gamma}{k_{\mathrm{B}}T}\right) \tag{4.15}$$

$$= \underbrace{\nu \exp\left(-\frac{E_{\mathrm{m}} - \mathcal{M}_{\mathrm{sl}}F_0 - \frac{1}{2}\mathcal{A}_{\mathrm{sl}}F_0^2}{k_{\mathrm{B}}T}\right)}_{\equiv \Gamma'} \tag{4.16}$$

$$\times \left[\exp\left(\frac{\mathcal{M}_{\mathrm{sr}}d\gamma + \mathcal{A}_{\mathrm{sr}}F_0 d\gamma}{k_{\mathrm{B}}T}\right) - \exp\left(-\frac{\mathcal{M}_{\mathrm{sr}}d\gamma + \mathcal{A}_{\mathrm{sr}}F_0 d\gamma}{k_{\mathrm{B}}T}\right)\right] \tag{4.17}$$

$$= 2\Gamma' \sinh\left(d\gamma \frac{\mathcal{M}_{\mathrm{sr}} + \mathcal{A}_{\mathrm{sr}}F_0}{k_{\mathrm{B}}T}\right), \tag{4.18}$$

where Γ' denotes the migration rate of atoms in homogeneous electric field F_0. Migrating by steps of length l for a time τ, the expected displacement of the atom is equal to

$$\langle x \rangle_{\mathrm{b}} = \tau l \Gamma_{\mathrm{b}} \tag{4.19}$$

Using the knowledge that the mean square displacement of an atom diffusing at rate Γ' is

$$\langle x^2 \rangle = \tau l^2 \Gamma', \tag{4.20}$$

the "normalized" expected displacement is equal to

$$\frac{\langle x \rangle_{\mathrm{b}}}{\langle x^2 \rangle} = \frac{2}{l} \sinh\left(l\gamma \frac{\mathcal{M}_{\mathrm{sr}} + \mathcal{A}_{\mathrm{sr}}F}{2k_{\mathrm{B}}T}\right), \tag{4.21}$$

where the length of the step is assumed to be $l = 2d$. This expression is conveniently independent of the diffusion time τ and the height of the barrier Γ, making it an easily comparable quantity between different experimental settings.

A central objective of this thesis was to develop atomistic surface diffusion models that would be able to incorporate the electric field bias effect predicted by Eq. (4.13). In the following chapters, I will discuss this development work, covering a breadth of simulation methods and techniques.

Chapter 5

Computer simulation models of surface diffusion

Atomistic simulations have been an essential tool to complement experimental results and theoretical predictions since late 1950s, when Alder and Wainwright simulated a system of elastically colliding hard spheres [46]. Along with the developments in computational power and maturation of theory, the range of different techniques suitable for different time and length scales has vastly expanded. In this section, I will introduce the models used in this work to study surface diffusion.

5.1 Kinetic Monte Carlo methods in materials physics

KMC is a well-established simulation method for studying processes that take place over timescales of several microseconds to hours or even years [47]. Its efficiency is based on describing the simulated system evolution as a chain of transition events—all faster motion between events is disregarded, saving computational effort. The time the system resides in the intermediate states is estimated based on the total rate of events.

In the context of materials physics, the events can be any thermally activated processes, such as atomic jumps, adsorption or desorption, chemical reactions, generation or recombination of defects, etc. Temperature only appears in the simulation as a parameter to calculate event rates, while thermal vibrations are not taken into account. In some cases, not even every atom needs to be considered in the simulation: for example, a crystal may be modeled in the terms of only the individual defects that diffuse through the system and interact with each other [48].

In the rejection-free KMC algorithm [49], at every step, a list of possible transition events is compiled. For every atom (or a more generalized object) in

the system, all processes available to it are added to the list. Then, one event is chosen at random, with the probability P of each event directly proportional to the rate Γ of that event type. The probabilities can be found from the rates by normalizing each rate by the sum of all rates. For event i,

$$P_i = \frac{\Gamma_i}{\sum_j \Gamma_j} \tag{5.1}$$

After the event is chosen and the system state updated accordingly, the simulation time τ is advanced by

$$\Delta\tau = -\frac{\ln u}{\sum_j \Gamma_j}, \tag{5.2}$$

where u is a random number in the interval (0, 1). $\Delta\tau$ is the estimate for the time the system *resides* in the current state, before the next event; Eq. (5.2) is also called the residence-time algorithm [49].

The justification for the randomized time increment lies in the unknown probability that no event happens before time $\Delta\tau$ has elapsed since the previous event; see Ref. [49] for details. The residence-time algorithm gives an estimate for the time evolution *proportional* to the real elapsed time—thus, it may be necessary to scale the simulation duration by a constant factor.

The rates Γ are often calculated by an expression similar to Eq. (4.8). Before discussing the estimation of the parameters required for rate calculations in Sec. 5.2, I will introduce the two primary types of atomistic KMC models used in materials physics: off-lattice and on-lattice.

5.1.1 Off-lattice kinetic Monte Carlo

KMC simulation proceeds as a series of jumps between adjacent potential energy minima. In the most general case, the minima can be any relaxed configurations, not constrained to lattice. To simulate such system, the adjacent minima to any given configuration, and, crucially, the saddle points leading to them, must be found at every step.

Starting from an arbitrary initial configuration, saddle points can be found by a number of algorithms, briefly covered in Sec. 5.3. After finding the saddle points, the corresponding final states of the event can be found with methods such as the activation-relaxation technique (ART) [50, 51]. Alternatively, the final states can be identified directly by searching for available free volume in the lattice by a heuristic approach, after which the saddle points can be found by techniques discussed later in Sec. 5.3. Such algorithm was suggested for a lattice-free KMC model in Ref. [52].

Off-lattice KMC is implemented in numerous codes, including k-ART (which relies on the aforementioned ART technique) [53], the SEAK-MC [54], the LE-KMC [55], and the SKMC [56]. Lattice-free KMC models are reviewed more extensively by Trochet et al. in Ref. [57].

The drawback of off-lattice KMC models is that they cannot easily be pre-parameterized. Barrier calculations can be, depending on the level of the utilized theory, prohibitively expensive to carry out on-the-fly.

5.1.2 Lattice kinetic Monte Carlo

For modeling crystalline matter, a reasonable approximation is that of a *rigid lattice*. In other words, the atoms are assumed to stay in their regular lattice positions, only momentarily jumping into the interstitial space to cross over to neighboring lattice sites. The rigid lattice approximation is more justified in temperatures well below the melting point of the material, where the lattice is not very distorted and the atoms do not have enough energy to easily jump into interstitial sites.

The usefulness of the rigid lattice for KMC simulations lies in the ease of parametrizing atomic jump processes beforehand, as well as compiling the lists of events during the simulation. As will be explained later in Sec. 5.2, the jump rates Γ depend on the LAE—the immediate surroundings–of the jumping atoms. With all atoms constrained to well-defined lattice sites, the environment of each distinct jump can be described by the states of its surrounding lattice sites, each either occupied by an atom of a given species, or vacant. The number of distinct jump event is thus finite, allowing pre-calculation of their parameters. During the simulation, since atoms can only jump into vacant lattice sites, the algorithm can simply find the vacancies in the lattice and look up the required parameters tabulated in terms of the state of the LAE.

The state can be expressed as a list of atomic species in the neighborhood, or, as in Ref. [25], as the list of the number of neighbors in different neighbor shells—1nns of the initial position, second-nearest neighbors (2nn) of the initial position, 1nns of the final position, and 2nns of the final position. This labeling scheme by four numbers is also called the 4D parametrization. Since the mapping of precise LAEs to 4D parameters is many-to-one, an additional method for choosing the environment for given 4-number combination is needed. Both random selection and a systematic rule of choosing environments with the lowest initial and final energy [58] have been successful.

Simulation suites exploiting the rigid lattice include the LAKIMOCA package [59, 60], the kmos code [61], the KMCLib library [62], and the Kimocs program [25].

With the target system in the scope of this thesis being Cu at room temperature, we utilized the rigid lattice approximation in the KMC simulations for greater efficiency.

5.1.3 Application of kinetic Monte Carlo to surface diffusion

As explained in Sec. 4.2, diffusion on crystalline surfaces can be described in terms of individual jumps between specific stable or metastable sites in the surface lattice. These jumps can be defined and parameterized for KMC simulations analogously to migration events within crystalline bulk.

A crucial difference to simulating the bulk case is the much larger variance in the possible LAEs, jump lengths, and diffusion mechanisms. Atoms near the surface have plenty of free volume around them, and thus more nearby energy minima to jump to. Effectively long-distance jumps can happen via exchange processes, where an adatom becomes a surface atom, pushing another atom out from the surface [63–65]. Off-lattice structures can form much easier than within bulk, e.g., when atoms start occupying the metastable "hcp sites" on the fcc {1 1 1} surface. The activation energies of jumps can range from tens of meV, as in the case of the almost gliding motion on the Cu {1 1 1} surface [66], to almost 1 eV for atoms detaching from well-bound positions on, e.g., adatom island edges [67].

All this complexity poses great challenges to simulating surface diffusion by KMC. The models are often restricted to predefined surface orientations, and/or require on-the-fly calculations [55, 68–71]. The Kimocs code, short for kinetic Monte Carlo on crystal surfaces, supports a fully three-dimensional pre-parameterized diffusion model, but not without challenges of its own. Finding solutions to these challenges constitute a major portion of this thesis work.

5.2 Parametrization of kinetic Monte Carlo models for diffusion processes

Estimation of the event rates Γ in a KMC simulation was already in earlier sections hinted to involve the calculation of the parameters of the Arrhenius equation (4.8). The Arrhenius equation itself arose from experimental work in the latter part of the 19th century [72], when the exponential behavior of reaction rates was observed empirically. The average attempt frequency and the average migration barrier can be inferred from the so-called Arrhenius plot, where the measured diffusion rate atoms is shown as a function of temperature (see Fig. 5.1 for illustration). First approximations to KMC modeling of diffusion could be based on experimental estimates of the parameters.

This section will give an overview on how to estimate of the attempt frequency and the migration barrier on the microscopic level. With the barrier being the dominant parameter regarding KMC simulation accuracy, I will discuss methods to calculate them in greater detail in Sec. 5.3.

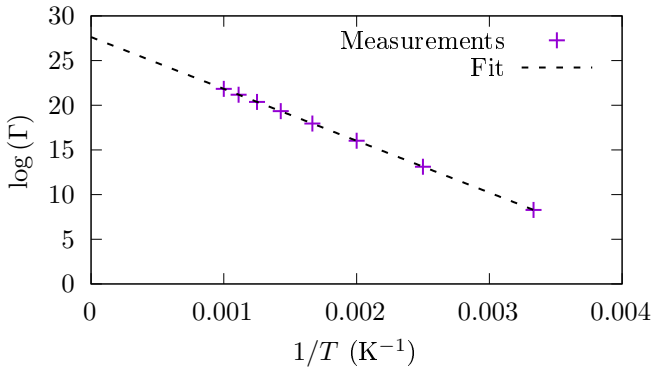


Figure 5.1: Hypothetical Arrhenius plot for a reaction with barrier equal to 0.5 eV and attempt frequency equal to 10^{12} Hz, measured at temperatures from 300 K to 1000 K. The slope of the fitted curve indicates the barrier height and the intercept at $1/T = 0\,\mathrm{K}^{-1}$ indicates the logarithm of the attempt frequency.

5.2.1 Attempt frequencies for calculation of event rates

The prefactor can be rigorously derived from the the transition state theory [73]. Using the harmonic approximation for the potential energy well in the initial state and in the saddle point, the attempt frequency can be expressed as a ratio of the products of the normal mode vibration frequencies [74]:

$$\nu = \frac{\prod_j^N \nu_j^{\mathrm{init}}}{\prod_j^{N-1} \nu_j^{\mathrm{saddle}}} \tag{5.3}$$

N is the number of degrees of freedom in the system, equal to three times the number of atoms in the three-dimensional system. Normal mode frequencies ν^{init} are for the initial state of the event, and ν^{saddle} are for the saddle point. The saddle-point system is an $(N-1)$-dimensional slice of the N-dimensional configuration space, since the dimension along the migration path has an imaginary normal mode. The products of the normal modes are, in fact, the partition functions of the initial position and the saddle point.

Moving from the N-dimensional TST to a simpler approximation, where each jump is considered a 1-dimensional event, independent of other jumps in the system, yields [75]

$$\nu = \frac{k_{\mathrm{B}}T}{h}\frac{nl^2}{2\alpha}\exp\left(\frac{\Delta S}{k_{\mathrm{B}}}\right), \tag{5.4}$$

where h is the Planck constant, n is the number of equivalent final positions available for the jump, l is the jump distance, α is the dimensionality of the motion (on the surface $\alpha = 2$), and ΔS is the difference in vibrational entropy between the saddle point and the initial site of the jump. Note that temperature and entropic effects are plainly explicit here, comparing to Eq. (5.3), where they are included implicitly in the partition functions [76].

In practice, the attempt frequencies vary between different events quite little, compared to the activation energies E_m—especially in cases where temperature stays the same throughout the simulation. Thus, it is often sufficient to use a single attempt frequency for all events to save computational resources.

For estimating a single attempt frequency ν for all events, another approach exists. Since ν effectively only modifies the time evolution of the simulation through equation (5.2), it can be obtained by fitting the time scale of a suitable process simulated in KMC to, e.g., a MD simulation of the same process. In other words, one can first run the KMC simulation with an arbitrary ν' and observe the time scale of the chosen process to be t_KMC. The "exact" time scale t_MD for the same process is obtained from MD, and the final ν for subsequent KMC simulations of this system is set to

$$\nu = \nu' \frac{t_\mathrm{KMC}}{t_\mathrm{MD}} \tag{5.5}$$

This approach was used in Refs. [25, 58, 65], where the reference process was chosen to be the flattening of cuboid nanotips on the surface. In a sense, by obtaining ν this way we are returning to the experimental roots of diffusion research, where the single prefactor of Eq. (4.8) "absorbed" all thermal and entropic effects of the attempt frequency.

5.2.2 Energy barriers for event rate calculations

As stated earlier, the migration energy barrier is a function of the LAE of the jumping atom. One straightforward way to include LAE-dependent variation in the barrier is to modify an estimated mean barrier E_0 by the total energies at the initial and the final states of the jump [60]:

$$E_\mathrm{m} = E_0 + \frac{E_\mathrm{fin} - E_\mathrm{ini}}{2} \qquad \text{(model 1)} \tag{5.6}$$

$$E_\mathrm{m} = E_0 + \max\left(0,\, E_\mathrm{fin} - E_\mathrm{ini}\right) \qquad \text{(model 2)} \tag{5.7}$$

The energies E_fin and E_ini were calculated on KMC runtime in an unrelaxed lattice in Ref. [60].

Another, very simple approach is to "drag" the jumping atom along a straight path from the initial to the final position, and calculate the energy in a number of points along the way [77]. The barrier is taken as the initial energy subtracted from the highest energy along the path. Avoiding iterative relaxation steps, this method can be implemented even by using an *ab initio* technique for the calculating energies.

Barriers can also be estimated without any runtime energy calculations, parameterizing the events in terms of the bonds broken by the jumping atoms when it leaves the initial lattice position, and the new ones formed at the saddle position and/or the final position. This method is commonly known as the bond-counting scheme [78], or the cut bond model [79]. The migration barrier for an atom of

species A is calculated as [80]

$$E_{\mathrm{m}}^{A} = E_{\mathrm{b,\,saddle}}^{A} - \sum_{i \in \mathrm{nn(ini)}} \epsilon_{AB_i} - \sum_{i \in \mathrm{nn(fin)}} \epsilon_{\mathrm{V}B_i}, \tag{5.8}$$

where $E_{\mathrm{b,\,saddle}}^{A}$ is the binding energy of the atom at the saddle position, B_i is the species of atom i, ϵ_{AB_i} is the interaction energy between species A and B_i (which can also be of the vacancy "species"), and $\epsilon_{\mathrm{V}B_i}$ is the interaction energy between the final, vacant position and an atom of species B_i. Summations are over the 1nn positions of the initial and the final state. The interaction energies can be precalculated from an interatomic potential or *ab initio*, or they can be treated as fitting parameters to produce the correct barriers for a number of known jumps. Modifications of the basic model include having $E_{\mathrm{b,\,saddle}}^{A}$ also depend on the LAE by calculating it as another sum of interaction energies, and extending the sums to the 2nn positions as well [79].

Migration barriers are very quick to estimate by these rather simple techniques, all of which avoid the computationally expensive task of finding the exact saddle points in relaxed lattices. However, they rely on the variations in the barriers to be quite small, finding more application inside bulk or being limited to single surface orientations. For a more general model of surface diffusion, more accurate methods are often required.

5.3 Computational approaches for estimation of relaxed energy barriers

The migration energy barrier is defined as the difference between the energy of the saddle point and the initial configuration. Thus, for accurate estimation, the energy of the relaxed saddle point must be calculated. This is not possible by placing an atom near the expected saddle point in the middle and performing a regular energy minimization, since the saddle point is by definition the energy *maximum* in one of the degrees of freedom.

The need for minimization altogether can be circumvented by an approach called metadynamics [81, 82]. It involves running a regular MD simulation, with an additional bias potential ΔV as a function of the selected collective variables (CV) that describe the reaction coordinates of the system as it exits the initial energy minimum. The bias is grown gradually over the course of the simulation, resulting in the end a complete energy profile $-\Delta V$ as a function of the CVs. The choice of CVs require prior knowledge of the system, but otherwise no initial guess for the reaction path is required. Another advantage of metadynamics is that it, in fact, calculates the *free* energy profile of the system at the finite temperature in which the MD simulation is run—instead of the 0 K potential energy barrier, the free energy barrier automatically includes entropic effects.

A more straightforward algorithm for finding saddle points from an initial energy minimum, not requiring any complicated parameters such as collective

coordinates, is the dimer method [83]. There, two replicas of the initial system are created, with the coordinates perturbed in a specific manner: a direction $\hat{\mathbf{N}}$ in the (for a 3-dimensional system of N atoms) $3N$-dimensional space is chosen randomly, and one replica is moved by a set amount ΔR along that direction, and the other by ΔR in the opposite direction. Then, the dimer is rotated to find the minimum curvature mode in the potential energy surface, and translated to maximize energy in the orientation of the dimer and minimizing it otherwise. Continuing this iteration, the dimer travels to the nearest saddle point.

The metadynamics and the dimer method are well suited for finding barriers when the final configuration of the event is not known beforehand, or when it is not even desired to be fixed, such as in certain styles of lattice-free KMC. In fact, the k-ART algorithm mentioned in Sec. 5.1.1 uses the dimer method for saddle point search. In lattice KMC, the target positions for jumps are defined as vacant lattice points, and the task is instead solving the correct barriers between given initial and final configurations. In this section, I will describe the algorithm used to calculate migration barriers for the KMC simulations in this work: the nudged elastic band (NEB) method [84, 85].

5.3.1 Nudged elastic band method

The NEB method calculates the migration barrier by relaxing the system at various points between the initial and the final configuration, but with the additional constraint of pushing the intermediate points away from the end points; this way, they cannot relax into the initial and the final minima. Instead of only the saddle point, NEB solves the entire MEP that goes through the saddle point. The algorithm is as follows:

1. Create n copies (images) of the atomic system, with the jumping atom in a different position in each image. The initial guess for the atom path can be, e.g., a linear interpolation between the initial and the final position (equal to the path used in the drag method).

2. Relax the images so that the path twists to follow the actual MEP. In addition to the regular interatomic forces within each image, the images interact with each other in a specific manner: the jumping atom interacts via a spring force with the jumping atom in the neighboring images in the series. In other words, the jumping atom images form a springy, *elastic* band.

3. Finally, certain forces in the system of images are ignored. The jump's *reaction coordinate* is defined as the 1-dimensional distance along the MEP. When the total forces acting on the jumping atom are calculated for the next relaxation step, the components of the real interatomic force within each image that are *parallel to the reaction coordinate* are ignored; and the components of the cross-image spring force that are *perpendicular to the reaction coordinate* are ignored. This way, the spring force does not affect

the "shape" of the path, but only keeps the images roughly equidistant. Correspondingly, the real interatomic forces between the images do not move the jumping atom along the band, but only *nudge* it perpendicular to the jump path. Hence the name, "nudged elastic band".

There are a number of modifications that can be made to this basic algorithm. The cross-image spring force can be applied to other atoms in the system as well, useful mostly in other applications of NEB where the process directly involves more than one atom. Another modification is the "climbing image" [85]: after a certain number of regular NEB steps, the image with the highest energy (closest to the saddle point) is cut loose from the band to be able to find the saddle point coordinate without restrictions, and the real interatomic forces acting to it parallel to the reaction coordinate are re-enabled, but, similarly to the dimer method, *reversed in direction*. This way, the image climbs up the potential energy surface, instead of relaxing down, and finds the true saddle point (maximum along the reaction coordinate, minimum perpendicular to it) more accurately. The climbing image algorithm was used in the NEB calculations throughout Publications **I**–**V**.

The NEB method is well suited for pre-calculating barriers for a lattice KMC model, since the events are defined in terms of the initial and final configurations that the algorithm needs as inputs. However, depending on the size and variation of the LAE and the number of atomic species present in the simulation, the sheer number of required barriers can be overwhelming. This motivates the final method used for barrier estimation in this work: machine learning (ML).

5.3.2 Machine learning algorithms for energy barrier predictions

ML is a subfield of AI research. Using ML to produce a function mapping means letting it "learn" its behavior by fitting into existing examples of correct responses to given inputs, rather than by handcrafted rules (predefined "if–then" statements). A classic example is image recognition: the inputs are the pixels of the images, and the expected responses are the patterns to recognize, such as the types of objects present in the images. The AI is given a functional form with a set of parameters, and the parameters are tuned by minimizing the classification error in a labeled training set of images.

In the context of computational physics, ML is most often used to approximate complicated functions. Here the input is a representation the physical system, such as the atomic configuration or the chemical composition, while the output can be any function the physicist is interested in, ranging from molecular energies to macroscopic parameters such as melting points and structural properties. One of the better known examples of ML applications in atomistic modeling are machine learned interatomic potential energy functions for MD simulations. Calculating the potential energy in an atomic system is typically *either* fast *or* accurate, with semiempirical formulas and DFT occupying the extreme

ends, respectively. By generating a training set of potential energy values that correspond to atomic configurations with an accurate method such as DFT, one can, in the best case scenario, train a function approximator that can produce quite fast, quite accurate energy values even for configurations it has never seen during training, offering a new tradeoff between speed and accuracy. Bartók et al. [86] and Behler [87] did some of the earliest work in this field.

In this work, ML was not used to directly estimate energies, but rather the energy *differences* between migration event saddle points and initial configurations—i.e., the migration barriers E_m of Eq. (4.8).

Earlier work solving the same task with ML exists. Djurabekova et al. used artificial neural networks (ANN) to estimate migration barriers in bulk Fe-Cu alloys [88]. Castin et al. have continued applying similar methodology to numerous different bulk materials [52, 89–99]. Simulating *surface diffusion*, however, is a greatly more difficult challenge, as explained in Sec. 5.1.3. Migration barrier prediction by ML on the surface specifically has been studied at least by Sastry et al. with genetic programming [100], and Verma et al. by cluster expansion [101]. Both of these works were limited to the fcc $\{1\,0\,0\}$ surface orientation. This thesis represents early work in constructing a ML assisted KMC model for diffusion on arbitrarily oriented, coexisting fcc surfaces.

The ML models used in this work were multilayer perceptrons (MLP) [102] and radial basis function networks (RBFN) [103], both subclasses of ANNs. This section outlines the mathematics behind the training processes of these methods.

Multilayer perceptron

The MLP is a mathematical model for non-linear classification and function approximation. It is constructed of nodes, also called neurons, organized in layers. The nodes are connected by "synapses" with adjustable weight parameters, representing the strength of the connections. See Fig. 5.2 for an illustration.

The purpose of the hidden layers is to produce a mapping of the input to the output. When output is requested for a given input, the nodes in the first hidden layer each calculate weighted sums of input vector $\mathbf{x}$ elements:

$$a_j = \sum_i w_{ji} x_i \tag{5.9}$$

The weight w_{ji} corresponds to the connection between input node i and hidden node j. The sum a_j is passed to the next layer through a non-linear *activation function* $g_j(\cdot)$:

$$z_j = g_j(a_j) \tag{5.10}$$

The nodes in the next layer calculate new weighted sums of the previous layer output and pass them through activation functions, until the signal arrives in the last layer that produces the "predicted" output vector $\hat{\mathbf{y}}$ with elements

$$\hat{y}_k \equiv z_k = g_k\left(\sum_j w_{kj} z_j\right) \tag{5.11}$$

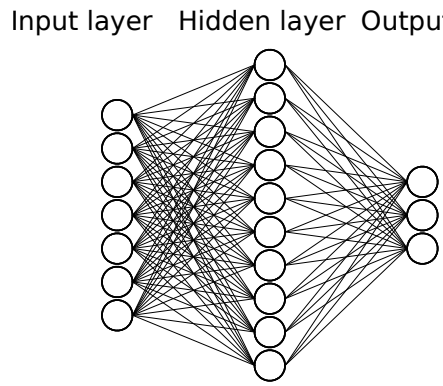


Figure 5.2: Schematic of an MLP. Circles are the nodes of the network, and lines are connections between nodes, each corresponding to one weight parameter that is optimized during training. [Publication **III**]

Here, the sum runs over all nodes i that are connected to output node k.

The network is "taught" to approximate a desired function by showing it examples of known input-output pairs $\{\mathbf{x}, \mathbf{y}\}$. Learning happens via optimization of the weight parameters between all nodes. The standard learning procedure of the MLP is the backpropagation algorithm [102]. It is based on iteratively predicting output vectors $\{\hat{\mathbf{y}}\}$ for input values $\{\mathbf{x}\}$ and calculating an error function E, assumed here to be of the form

$$E = \sum_n E^{(n)}, \tag{5.12}$$

i.e., the total error is a sum over the individual errors for each input-output pair (pattern) n. Common choices for the error function $E^{(n)}$ are the sum-of-squares $\sum_k \left(\hat{y}_k^{(n)} - y_k^{(n)}\right)^2$ and the log-loss $-\sum_k \hat{y}_k^{(n)} \log y_k^{(n)} + \left(1 - \hat{y}_k^{(n)}\right) \log \left(1 - y_k^{(n)}\right)$.

At every step, the network weights are updated to correct for the error by descending the gradient of the error function with respect to each weight. Using the chain rule, the gradient with respect to the weight w_{ji} for the pattern n is

$$\frac{\partial E^{(n)}}{\partial w_{ji}} = \frac{\partial E^{(n)}}{\partial z_j^{(n)}} \frac{\partial z_j^{(n)}}{\partial a_j^{(n)}} \frac{\partial a_j^{(n)}}{\partial w_{ji}} \tag{5.13}$$

Evaluating $\partial E^{(n)}/\partial z_j^{(n)}$ expands into calculating the corresponding derivatives in the next layer of the network, and so on, until in the output layer it can be directly evaluated knowing the form of the error function. Thus, the evaluation must be started from the output and propagated backward; hence the name of the algorithm. The interested reader is referred to, e.g., Ref. [104] for further details.

After all derivatives have been calculated, the weights are updated according to gradient descent. This can be done either after showing each pattern (on-line training):

$$\Delta w_{ji} = -\eta \frac{\partial E^{(n)}}{\partial w_{ji}}, \tag{5.14}$$

or after showing every pattern in the training set (batch training):

$$\Delta w_{ji} = -\eta \sum_n \frac{\partial E^{(n)}}{\partial w_{ji}} \tag{5.15}$$

Here, η is the learning rate, a user-defined hyperparameter of training.

This is the basic form of the algorithm. Improvements exist, such as the resilient backpropagation (RPROP) [105], and the improved RPROP [106]. The latter was used to train the networks in Publication **III**.

Also the network architecture can be optimized during training. One example is the cascade-correlation algorithm [107], which constructs the network one node at a time: each node is placed in its own layer, and every preceding node is connected directly to it by *shortcuts*. The new node is connected to the output node, along with every preceding node. All newly formed weights are trained until no improvement is observed. Then, the weights are *frozen*, and another new node is added, as long as the desired error or the desired maximum number of nodes is reached. A variation of this algorithm was used in Publication **III**.

Choosing activation functions $g(\cdot)$ is an important step of designing the network. In the original perceptron the MLP is based on, the activation function was the Heaviside step function

$$H(x) = \begin{cases} 1, & \text{if } x > 1 \\ 0, & \text{if } x \leq 0, \end{cases} \tag{5.16}$$

or its antisymmetric version, $2H(x) - 1$. This function would be sufficient for providing the non-linearity in the MLP as well, but it is incompatible with the backpropagation algorithm, which requires the derivative of the activation function when updating the weights. The activation functions are often "smoothened" step functions, such as $\tanh(x)$ or the logistic sigmoid:

$$s(x) = \frac{1}{1 + \exp(-x)} \tag{5.17}$$

Radial basis function network

RBFNs are neural networks that use radial basis functions for the activation of hidden nodes. In other words, the activation of each node does not depend on the weighted sum over the elements of the input vector $\mathbf{x}^{(n)}$, such as in the case of MLPs, but on its distance to the node's centroid vector $\mathbf{x}'$. Typically, RBFNs have only one layer of hidden nodes. The non-linear hidden node outputs are

combined linearly into output with elements

$$\hat{y}_k^{(n)} = \sum_j w_{kj}\phi_j\left(||\mathbf{x}^{(n)} - \mathbf{x'}_j||\right) \tag{5.18}$$

The shape of activation functions $\phi(\cdot)$ is to be decided by the user. Common choices include the multiquadric function (used in Publication **III**) and the Gaussian function (used in Publication **V**). The multiquadric function is

$$\phi_j(||\mathbf{x}^{(n)} - \mathbf{x'}_j||) = \sqrt{1 + \left(\varepsilon_j||\mathbf{x}^{(n)} - \mathbf{x'}_j||\right)^2} \tag{5.19}$$

Here, ε_j is a width parameter of the function. Included in all types of RBFN basis functions, it controls how much the basis functions overlap, i.e., how slowly the contribution of the basis function $\phi_j(\cdot)$ vanishes for input vectors that are further away from its centroid $\mathbf{x'}_j$.

The RBFN training process can be carried out in two steps. In the first step of this training approach, the basis function widths and centroid vectors are chosen. This procedure is usually is *unsupervised*, i.e., controlled only by the input values $\{\mathbf{x}\}$ in the training set. The centroids can be, e.g., sampled from among the input vectors, or they can be formed by a clustering algorithm. The width parameters $\{\varepsilon\}$ can be chosen, e.g., based on the density of the centroids, so that they smoothen the regression result appropriately. After setting the basis function parameters in this unsupervised manner, the weights $\{w\}$ are learned with supervised training, producing an optimal mapping of $\{\mathbf{x}\}$ to $\{\hat{\mathbf{y}}\}$. In this special case of one layer of linear weights, the least squares solution can be found with linear algebra. Eq. (5.18) can be expressed as a matrix equation: for N instances of training data with d-dimensional output vectors $\{\mathbf{y}\}$, and M basis functions in the RBFN,

$$\underbrace{\begin{bmatrix} \phi_1\left(||\mathbf{x}^{(1)} - \mathbf{x'}_1||\right) & \cdots & \phi_M\left(||\mathbf{x}^{(1)} - \mathbf{x'}_M||\right) \\ \phi_1\left(||\mathbf{x}^{(2)} - \mathbf{x'}_1||\right) & \cdots & \phi_M\left(||\mathbf{x}^{(2)} - \mathbf{x'}_M||\right) \\ \vdots & \ddots & \vdots \\ \phi_1\left(||\mathbf{x}^{(N)} - \mathbf{x'}_1||\right) & \cdots & \phi_M\left(||\mathbf{x}^{(N)} - \mathbf{x'}_M||\right) \end{bmatrix}}_{\mathbf{\Phi}} \underbrace{\begin{bmatrix} w_{11} & \cdots & w_{d1} \\ w_{12} & \cdots & w_{d2} \\ \vdots & \ddots & \vdots \\ w_{1M} & \cdots & w_{dM} \end{bmatrix}}_{\mathbf{W}} \tag{5.20}$$

$$= \underbrace{\begin{bmatrix} y_1^{(1)} & \cdots & y_d^{(1)} \\ y_1^{(2)} & \cdots & y_d^{(2)} \\ \vdots & \ddots & \vdots \\ y_1^{(N)} & \cdots & y_d^{(N)} \end{bmatrix}}_{\mathbf{Y}} \tag{5.21}$$

$$\mathbf{\Phi W} = \mathbf{Y} \tag{5.22}$$

The optimal weights are obtained by calculating the pseudoinverse of matrix $\mathbf{\Phi}$

$$\mathbf{W} = \left(\mathbf{\Phi}^\intercal\mathbf{\Phi}\right)^{-1}\mathbf{\Phi}^\intercal\mathbf{Y} \tag{5.23}$$

In the case of square matrix $\mathbf{\Phi}$, i.e., $M = N$, the pseudoinverse reduces to the regular inverse $\mathbf{\Phi}^{-1}$.

Alternatively, the basis function parameters and the linear weights can be learned simultaneously, with, e.g., gradient descent techniques. The advantage here is greater adaptability of the basis functions; on the other hand, the two-step approach can be faster, if more memory-intensive due to matrix inversion calculations. For more technical details on RBFN, see, e.g., [104].

5.4 Simulating surface diffusion dynamics in high electric fields

As explained in earlier chapters, the purpose of this work is to develop methods and apply them to simulate Cu surface diffusion specifically under electric fields, in effort to contribute toward improved understanding the vacuum arcing phenomenon. The previous section focused on simulating surface diffusion with KMC. In KMC, the interatomic forces are not considered explicitly, but rather they are hidden in the migration energy barrier parameters. Thus, introducing external forces, such as those caused by electric fields, is not straightforward. To a degree, the effect of the field can be introduced as a perturbation to the potential energy barriers E_{m}, as was done, e.g., in Ref. [108]. In that paper, KMC simulation was coupled with an electric field solver to estimate the local field and its gradient experienced by the migrating surface atoms. The permanent dipole moment and polarizability of Eq. (4.13) were assumed to be independent of the LAE of the migration event, allowing for a simple modification of the energy barriers.

According to a more recent study [109], dipole moment and polarizability have a significant and complicated dependency on the LAE. The authors suggest that full parametrization of surface diffusion in long-term, large scale KMC simulations require extensive density functional theory calculations.

As an alternative to the field-modified KMC, in Publication **VI** we explored MD for simulating surface diffusion under electric field. In MD, the motion of atoms is determined by the forces acting on them, with migration energy barriers appearing as an emergent property. The forces due to the electric field can be added directly in the simulation, instead of including them via barrier perturbations predicted by theory. The effect of the LAE of migration events is automatically included in the force calculations as accurately as the local field solver follows the atomic surface geometry. This flexibility comes at the cost of significantly limited simulation time and length scale, compared to KMC.

This section will cover the methodology necessary for MD surface diffusion simulations in electric field. To alleviate the time scale limitation for systems where atomic migration is slow, I will give an example of an acceleration technique also utilized in Publication **VI**: collective variable -driven hyperdynamics (CVHD) [110]. Finally, I will very briefly describe the theoretical background

of the finite elements method (FEM) for solving the differential equations of electric fields, and coupling of FEM to MD [27].

5.4.1 Classical molecular dynamics

MD is a method for simulating atomic systems according to given set of rules for interactions between atoms. The atoms are given initial positions $\mathbf{R_0}$, and initial velocities with magnitudes $\{v\}$ sampled, e.g., from the Maxwell-Boltzmann distribution at the desired temperature T

$$f(v) = \left(\frac{m}{2\pi k_\mathrm{B} T}\right)^{3/2} 4\pi v^2 \exp\left(-\frac{mv^2}{2k_\mathrm{B}T}\right), \tag{5.24}$$

where m is the atom's mass. After initialization, the total force on each atom i is found as the gradient of the given potential energy function $V(\mathbf{R})$

$$\mathbf{F}_i = -\nabla_i V(\mathbf{R}) \tag{5.25}$$

The atomic accelerations are found from the Newton's second law

$$\mathbf{a}_i = \frac{1}{m_i}\mathbf{F}_i \tag{5.26}$$

The simplest possible algorithm for integrating acceleration is the Euler method, where velocities and positions are updated for the next *timestep* as

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\Delta t + \frac{1}{2}\mathbf{a}_i(t)\Delta t^2 \tag{5.27}$$

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \mathbf{a}_i(t)\Delta t, \tag{5.28}$$

where Δt is the length of the timestep, usually in the order of femtoseconds. Eqs. (5.25)–(5.28) are iterated until the desired number of timesteps or some other end condition has been met.

In practice, the acceleration is integrated with higher-order methods, such as the velocity Verlet [111]. This algorithm was also used in Publication **VI**.

Without further modifications, the thermodynamical *ensemble* realized in an MD simulation is the *microcanonical* ensemble: the number of particles, the volume, and the total energy of the system remain constant. This corresponds to a completely isolated system that exchanges neither particles nor heat with its environment. While this is the default mode of MD, in the real world, no experimental setting can be perfectly isolated—such state can be reached only approximately. In an uncontrolled "laboratory" environment, or the macroscopic world in general, the properties usually observed to stay constant are the temperature and pressure. It is often desirable to model such conditions in MD as well. Instantaneous temperature T in MD relates to the average kinetic energy of the particles; in a three-dimensional system without internal degrees of freedom (rotation or vibration of single particles),

$$T = \frac{1}{3Nk_\mathrm{B}} \sum_i m_i v_i^2, \tag{5.29}$$

where N is the total number of particles in the system. Pressure P in MD is calculated based on the interatomic forces:

$$P = \frac{Nk_\mathrm{B}T}{V} + \frac{1}{3V}\sum_i \mathbf{r}_i \cdot \mathbf{F}_i, \tag{5.30}$$

where V is the volume of the system. To emulate thermal conduction to an external heat bath at temperature T_0, the rate of temperature change should be proportional to the temperature difference:

$$\frac{\mathrm{d}T}{\mathrm{d}t} \approx \frac{T(t+\Delta t) - T(t)}{\Delta t} = \frac{T_0 - T(t)}{\tau_T} \tag{5.31}$$

τ_T is a proportionality constant in the units of time. The temperature at the next timestep is thus

$$T(t+\Delta t) = \frac{\Delta t}{\tau_T}\left[T_0 - T(t)\right] + T(t) = T(t)\left[\frac{\Delta t}{\tau_T}\left(\frac{T_0}{T(t)} - 1\right) + 1\right] \tag{5.32}$$

With velocities proportional to the square root of temperature, this gives the scaling constant to velocities:

$$\mathbf{v}_{i,\,\mathrm{scaled}} = \mathbf{v}_i\sqrt{\frac{\Delta t}{\tau_T}\left(\frac{T_0}{T(t)} - 1\right) + 1} \tag{5.33}$$

The scaling constant for the system volume to control pressure can be derived analogously. Scaling like this is called the Berendsen thermostat [112] (and barostat). The choice of the time constant τ_T (and the analogous τ_P) between Δt and infinity tunes the system between a rather unrealistic behavior without any fluctuations, and the uncontrolled microcanonical ensemble. For more sophisticated control that enforces the proper thermodynamic *isothermal-isobaric* ensemble, more advanced methods exist—in Publication **VI**, we used the Nosé-Hoover thermostat [113] and the analogous barostat.

Interatomic potentials

A crucial part of carrying out MD simulations is choosing a suitable set of "rules" to describe the atomic interactions. Within classical MD, the rules typically have the form of potential energy function directly dependent on atomic positions. In *ab initio* MD, the potential energy can be a functional of the electronic density, or, for very small systems, it can be calculated even on the wavefunction level.

The simplest potential energy function only has 2-body (pair) terms, i.e., interatomic distances. A well-known example is the Lennard-Jones potential [114], parameterized by the binding energy ϵ and the length scale σ:

$$V_\mathrm{LJ}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^6 - \left(\frac{\sigma}{r}\right)^{12}\right] \tag{5.34}$$

The length scale determines the distance at which the energy between two particles is zero; the equilibrium distance, where the force between them is zero, is

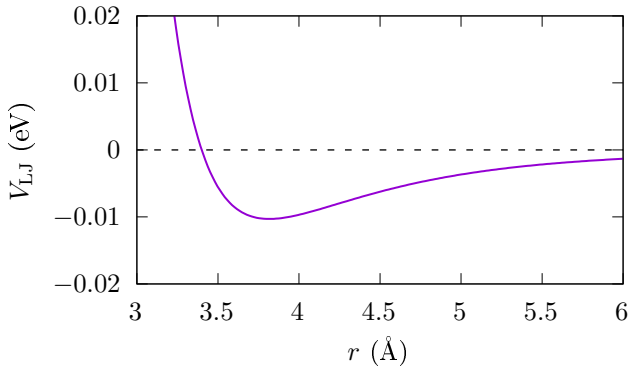


Figure 5.3: Lennard-Jones potential for Ar [115].

equal to $2^{1/6}\sigma$. The shape of the potential is illustrated in Fig. 5.3. The Lennard-Jones potential applicable primarily for weakly interacting systems, such as noble gases; it cannot describe metallic, ionic or covalent bonds realistically.

The interatomic potential for Cu interactions in this work follows the embedded atom model (EAM) [116] formalism. EAM is a commonly used manybody potential for metals, combining pair interaction ϕ and an *embedding energy* F. The embedding energy can be understood as the interaction between the metal atom and the electronic density where it is "embedded". The electronic density in turn is calculated as the sum of the spherically symmetric "electron clouds" ρ of all the other atoms in the system. The energy of atom i is calculated as

$$E_i = F\left(\sum_{j\neq i}\rho(||\mathbf{r}_i - \mathbf{r}_j||)\right) + \frac{1}{2}\sum_{j\neq i}\phi(||\mathbf{r}_i - \mathbf{r}_j||) \tag{5.35}$$

The functions F, ρ, and ϕ can be defined semi-empirically to reproduce the desired physical properties of the solid, such as the lattice structure, lattice parameter, and elastic moduli. There are also ways to determine them from first principles, e.g., by calculating the embedding energy of atoms in electron gas (jellium) in DFT.

The interatomic potential model used in this work is based on the developments around the effective medium (EM) theory [117], originally used to describe the energy to embed atoms in inhomogeneous electron densities by expressing it in terms of the embedding energy in a homogeneous *effective* density. Note that the embedding energy function F also treats the density as homogeneous, since it does not depend on any derivatives of ρ. Kress and DePristo introduced the corrected effective medium (CEM) theory by adding Coulombic interaction and accounting for the difference in kinetic, exchange, and correlation energies between the effective homogeneous system and the actual, inhomogeneous one [118]. Finally, Stave et al. developed a version of CEM optimized for large MD and Monte Carlo simulations by simplifying the computationally expensive CEM corrections to be absorbed back into the embedding term F [119]—now, the full model can

be expressed in the form of Eq. (5.35). This is called the MD/MC-CEM theory, used in the barrier calculations of Publications **I**–**IV** and the MD simulations of Publication **VI**.

5.4.2 Accelerated molecular dynamics simulations

As was described in Sec. 5.1, what makes KMC an attractive method for simulating diffusion, is the ability to disregard short timescale thermal motion to gain efficiency in modeling slower processes. The slower the diffusion, the greater advantage KMC has over MD.

There exist, however, techniques to accelerate MD. Among the simplest of them is the parallel replica method [120], which runs N copies of the system simultaneously, monitored to detect transition events. Once an event is detected in one of the copies, this replica is taken as the starting point for the regeneration of all the other replicas. Whenever new replicas are generated, the atomic velocities are reinitialized randomly. Simulation time advances at N times higher pace compared to unaccelerated MD—the acceleration factor is equal to the factor of added computational effort required to run the replicas.

The acceleration method used also in Publication **VI** is CVHD [110]. Unlike in the parallel replica method, in CVHD the additional computational cost scales with the number of accelerated atoms, while the acceleration factor depends on the *total* rate of the processes of interest. Depending on the process rates (temperature and migration barriers), factors of up to 10^9 can be reached.

The CV is a number that describes the state of the whole system with respect to the vicinity of a transition state. CVs were mentioned in Sec. 5.3 in the context of metadynamics, indeed a close relative of hyperdynamics. In CVHD, this reaction coordinate is normalized such that when the system is near equilibrium, the CV close to zero, and as it approaches a transition, CV approaches one. The suitable definition of the CV depends on the type of process to be accelerated: in diffusion simulations, a good choice is based on the interatomic "bond lengths" r, i.e., the distances between nearest neighbors; the higher the distances, the larger the CV.

The acceleration is enforced with a bias potential ΔV as a function of CV, illustrated in Fig. 5.4a. The total potential energy of the system is

$$V^*(\mathbf{R}) = V(\mathbf{R}) + \Delta V(\eta), \tag{5.36}$$

Where η is the CV. The bias potential biases the interatomic forces through η, itself also dependent on the atomic positions $\mathbf{R}$. The bias acting on atom i is

$$\Delta\mathbf{F}_i = -\frac{\partial \Delta V}{\partial \eta} \sum_{j \in \mathrm{nn}(i)} \frac{\partial \eta}{\partial r_{ij}} \left(\frac{\partial r_{ij}}{\partial x}\hat{\mathbf{x}} + \frac{\partial r_{ij}}{\partial y}\hat{\mathbf{y}} + \frac{\partial r_{ij}}{\partial z}\hat{\mathbf{z}} \right), \tag{5.37}$$

where r_{ij} is the distance between atoms i and j, and the summation goes over i's nearest neighbors.

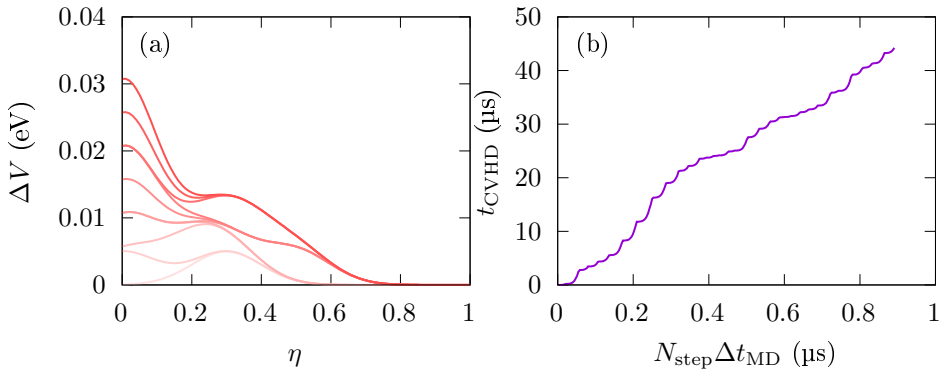


Figure 5.4: (a) Bias potential in CVHD as a function of the CV η. The bias drives the system closer to the transition state at $\eta = 1$. ΔV is grown gradually during the simulation, allowing for flexibility. (b) Accelerated CVHD time as a function of the regular MD simulation time. [adapted from Publication **VI**]

The acceleration factor at each MD timestep depends on the bias potential:

$$\Delta t_{\mathrm{CVHD}} = \Delta t_{\mathrm{MD}} \left\langle \exp\left(\frac{\Delta V}{k_{\mathrm{B}}T}\right) \right\rangle, \tag{5.38}$$

where Δt_{MD} is the regular MD timestep length and t_{CVHD} is the accelerated time. As the bias potential grows over time, the acceleration factor grows. This is illustrated in Fig. 5.4b. When the system crosses the transition state, detected by monitoring of η's saturation to one, all bias potential is removed and the interatomic bonds are reformed to reset the equilibrium state. This reduces the applicability of the method to large systems where events happen frequently (the total rate of events is high).

The precise functional forms of ΔV and η, along with further details and parameters of the CVHD method, are covered in Publication **VI**.

5.4.3 Electric field in molecular dynamics

In classical MD, atoms are modeled as pointlike objects that interact via the defined interatomic potential. Using simple interaction models, such as Lennard-Jones or EAM type functions, electrons of the system are "hidden" in the potential parameters. Thus, there is no meaningful definition of an electric field in bare MD, since it would not interact with the atoms in any manner. To simulate diffusion in electric field, further modifications are needed.

As a first approximation, when the main effect desired to be included in the simulation is the Maxwell stress, it can be added directly by adding a corresponding force to act on surface atoms in the simulation [121–123].

A rather simple way to introduce atomic polarizabilities in MD is to separate the atoms into positively charged cores and negatively charged shells, connected to each other by a harmonic spring [124, 125]. This way, the particles can interact

electrically with each other and also with an external electric field, but in general the method is not applicable to metals in the absence of free electrons.

A step closer to the goal of simulating larger, freely polarizable objects, is the charge equilibration technique [126]. Atoms can exchange charge freely, and the total charge distribution is solved self-consistently within the constraint of total charge conservation. Originally applied to the polarization of chemical compounds, this method was recently employed in simulations of metallic surfaces by Bagchi and Perez [28].

Another approach is coupling the MD system to a continuum electric field solver. Djurabekova et al. developed an electrodynamics-MD method, called HELMOD [26], where the fields and surface charges are obtained from the solution of the Laplace equation

$$\nabla^2 \Phi = 0 \tag{5.39}$$

in a finite-difference mesh. Φ here is the electrostatic potential. The electric field is obtained as the gradient of Φ, and the charge within a given (surface) volume is calculated by Gauss's law. Based on the charges and the field, Lorentz and Coulomb forces are calculated and added to the interatomic forces in MD.

In modeling of diffusion by MD, maintaining the system size small is critical to reach sufficiently long time scales. However, applying an electric field *gradient* is most straightforward by geometrical means, where the field is enhanced toward the edges and the vertices in the system. To allow large variation in the effective field and its gradient, larger systems are preferable.

A solution to the conflicting requirements is offered by moving from the uniformly dense finite-differences method to the more adaptive FEM. Such tool, named Femocs [27], was developed as the continuation of the HELMOD code by Veske et al. It takes the atomic coordinates from MD as inputs to generate a surface mesh, where it solves either the Laplace or the Poisson equation

$$\nabla \left(\varepsilon \nabla \Phi \right) = -\rho, \tag{5.40}$$

depending on whether the space charge density ρ is included in the simulation or not. ε is the permittivity function (different for inside and outside the surface). Boundary conditions can be set as needed, e.g., in the case of an infinitely wide capacitor plate with its counterpart placed infinitely far, the potential is constant on the surface, its gradient is constant at the opposite side of the system, and no electric flux is allowed through the periodic boundaries.

The advantage of FEM over finite-differences is that the elements can be made larger far away from the regions of interest to save computational effort. They will also be naturally larger where the surface is relatively flat. The size of the elements can guide the optimization of surface charge calculations: in flat areas, the charge is calculated for the finite element faces and distributed to atoms within the corresponding elements, while in rougher regions the charges are calculated individually for surface atom Voronoi cells (limited by additional support points outside the surface) [127].

A caveat of FEM-MD is that the permanent dipole moment of adatoms, mentioned in Sec. 4.3, is not included in this model. The "electron density" function ϕ of the EAM potential does not contribute to the actual electrostatic calculation by the FEM, while the Laplace equation solver treats the surface mesh as a macroscopic object and does not account for valence electron shifts responsible for adatom polarization.

What makes the inclusion of electrostatics via FEM coupling particularly attractive for the purpose of this work, is that the surface mesh can be freely expanded beyond the atomistically defined system at a relatively low computational cost. The electric field can be solved in an arbitrary geometry to provide precisely the desired field and gradient on the surface atoms.

Chapter 6

Atomistic simulations of Cu surface diffusion

In this chapter I will summarize the results published in Publications **I**–**VI**. The first section describes the challenges and developments in the KMC model based on a rigid lattice approximation for Cu surface diffusion. The second section covers the results of the MD simulation of Cu diffusion in electric field of Publication **VI**.

6.1 Energy barriers for self-diffusion on Cu surfaces

The KMC simulations in Publications **I**–**V** were carried out with the Kimocs software [25]. Kimocs is specifically developed for simulating surface diffusion, used successfully in numerous studies of, e.g., Cu [128], Au [58], and W [65, 108] surfaces. Kimocs, as it was published in 2016, uses the rigid lattice approximation with the 4D encoding scheme of the LAE, briefly described in Sec. 5.1.2. Migration energy barrier parametrizations have been produced using the NEB method, described in Sec. 5.3.

Calculating migration energy barriers for surface jump processes in rigid lattice has some challenges not present in barrier calculations inside bulk material. Within crystalline bulk, the atoms are well-supported by the surrounding atoms, even when there are vacant lattice sites or other defects nearby. Each lattice site constitutes a stable potential energy basin, to and from which atoms can jump along well-defined MEPs and associated energy barriers. On the surface, however, the atoms near the surface are, by definition, significantly under-coordinated. In fcc crystals, the site for an individual adatom on the $\{1\,0\,0\}$, the $\{1\,1\,0\}$, and the $\{1\,1\,1\}$ surfaces is surrounded by 4, 5, and only 3 nearest neighbor atoms, respectively. On an arbitrarily rough surface, surface-adjacent sites can have any coordination ranging from 1 almost to the full 12. When calculating migration barriers of other atoms jumping into these available sites, a number of issues can

arise:

1. The target site is *unstable*, i.e., when a relaxation calculation is performed for an atom jumping there, the configuration collapses into something completely different. The migration barrier for the original jump cannot be found.

2. The initial site is unstable, i.e., the jumping atom changes its position during the initial relaxation simulation. It either spontaneously falls into the final site of the transition of interest or to a completely new unrelated to the given transition neighboring site. An event like this could, in principle, have a zero energy barrier associated to it. However, this violates the assumption of equilibrium in the transition state theory.

3. The MEP is not well-defined, i.e., it has intermediate minima that can even have lower energy than the initial or the final configuration. This happens especially frequently on the fcc {1 1 1} surface, where the off-lattice hcp sites are almost as stable as the on-lattice fcc ones.

A straightforward way to avoid these problems would be simply marking unstable sites and jumps to them forbidden in the KMC simulation, instead of attempting to define a barrier for them. However, if we consider 1nn jumps to unstable positions as *intermediate states* of longer-distance jumps, we can limit the parametrization process significantly, while still being able to include long-distance events in our model. For instance, atoms jumping down from single monolayer steps (crossing the Ehrlich-Schwoebel barrier [129, 130]) must jump further than their 1nn positions (either by hopping or by exchange process). The position at the step edge is not stable, and thus it is impossible to assign barrier for a 1nn hopping jump ending there in regular NEB. If these jumps are forbidden, atoms finding themselves on top of steps would be indefinitely stuck there. It would be desirable to let the atoms temporarily jump onto the unstable edge position instead.

All the challenges listed above make the systematic calculations of the barriers either to be used directly in KMC or for training ML algorithms impossible. To resolve this issue, we have taken two approaches to enable systematic parametrization by 1nn jump barriers. The first is the tethering force approach, briefly described in Sec. 6.1.1, and the second is the use of AI to estimate barriers for arbitrary events based on a training data set. The latter technique was explored in Publications **III**–**V**, summarized in Sec. 6.1.2.

6.1.1 Tethering force approach in migration barrier calculations

The solution to the problem of atoms falling out of place during NEB calculations, proposed in Publication **I** and Ref. [58], is to apply an additional *tethering force*. It is a 3-dimensional spring force that loosely binds each atom to its original

lattice site. Effectively, the NEB calculation is carried out in a *semi-rigid* lattice, instead of a fully relaxed one. The estimated distorting effect of tethering to *stable processes* is within 0.1 eV. At this cost, the method will allow the estimation of unstable but necessary process barriers for the KMC simulation.

The tethering force approach has been successfully employed already for producing 4D parametrizations for lattice KMC models. The barrier calculation process can be fully automatized, requiring as input only the potential energy function that describes the interatomic interactions and a method to choose the LAE corresponding to each 4-number combination. With the LAE restricted to the 2nn neighborhood around the event, the number of different events is only some thousands, and thus all barriers can be calculated with reasonable computational cost.

What if the surface of interest consists of two or more elements? The scheme of counting of neighbors could be generalized to counting neighbors of different species separately, but a fundamental issue lies in the explosion of the number of different events. This annihilates the attractive property of the 4D encoding space being compact enough to be fully explored within reasonable time. The same problem is encountered even within a single element, if it turns out that neighbors beyond the 2nn shell make a significant contribution to the migration energy barriers.

The next section of this work presents an approach that can handle the expansion of dimensions, either in the number of elements or the number of neighbor shells, or both. Instead of an exhaustive search of the LAE space, we only need to sample a slice of it with NEB to produce a training set for an ML algorithm.

6.1.2 Migration barrier prediction by machine learning

The basic principles of the ML methods used in Publications **III–V** were outlined in Sec. 5.3.2. In this section, I will describe their specific application to estimate migration barriers on the Cu surface.

Local atomic environment encoding

In order to train an ML algorithm to interpolate and extrapolate a function, a mapping of input to output, the function's input and output must be in a suitable format. In our case, the migration barrier output does not need any special encoding. The encoding of the LAE input, on the other hand, is much less straightforward. The input encoding scheme in ML is also called the "descriptor".

The LAE means the configuration of atoms around the jumping atom—thus, a first idea might be to encode the environment simply with the Cartesian coordinates of each atom in the neighborhood. However, a critical shortcoming becomes apparent when investigating the *invariances* of the migration energy barrier: the barrier does not change if the atomic configuration undergoes a translation, a rotation or a reflection. The Cartesian presentation is not invariant in any of these transformations, meaning that infinitely many different presentations map

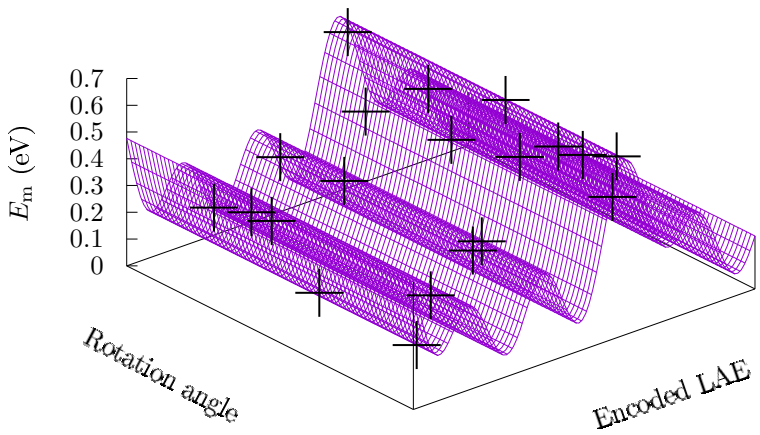


Figure 6.1: Schematic of a migration barrier function with a LAE descriptor that does not respect rotational invariance. The barrier varies as a function of the LAE, but it is invariant with respect to its rotation angle. Now, if the ML training set is formed by sampling LAEs at arbitrary angles (marked in the plot with crosses), learning the correct E_m surface will be more difficult than in the case where the descriptor has the proper invariances.

to each physically unique configuration. In other words, the migration barrier function will have degeneracies that the AI has to learn *in addition* to being able to interpolate accurately (see Fig. 6.1).

The second idea for the descriptor could be the 4D scheme, introduced in Sec. 5.1.2. Since the 4D descriptor only counts the numbers of 1nn and 2nn atoms, disregarding their coordinates, it is certainly invariant with respect to translations, rotations, and reflections. However, here we encounter the opposite problem to that of Cartesian descriptors. The same number of 1nn and 2nn atoms can be arranged around the jumping atom in physically different ways, each leading to the same 4D descriptor. From ML perspective, the input space becomes "too compact", with each input point hiding complexity of the exact atomic configurations and their effects on the migration barrier, with the AI struggling to find the systematic behavior of the function (see Fig. 6.2).

Concluding from these two example encodings unsuitable for ML purposes, a good encoding scheme needs to 1) respect the invariances of the LAE itself, and 2) map, ideally, descriptors and LAEs one-to-one (bijectively). There are multiple such descriptors developed for ML in materials science to describe atomic arrangements. Some are based on the interatomic distances [131], nearest-neighbor distances, angles and/or torsions [132] present in the structure, while others consider the local atomic density and describe it in terms of, e.g., spherical harmonics expansion coefficients [133–135]. In addition to the desirable invariances and bi-

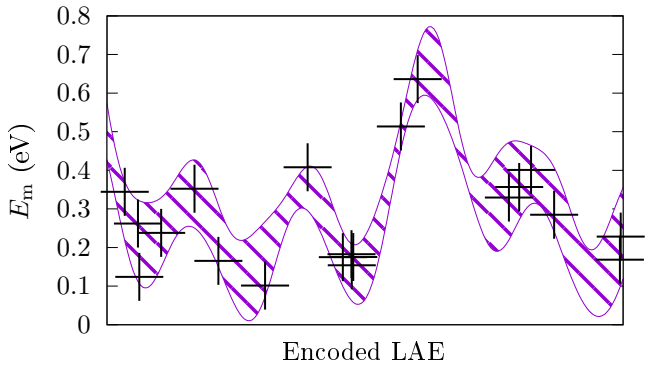


Figure 6.2: Schematic of a LAE descriptor with hidden complexity. Each input corresponds to several different LAEs with different migration barriers E_{m}. It will be more difficult for the ML algorithm to converge at a good approximation for the underlying function based on the sampled training set (marked with crosses), than in the case where the descriptor was one-to-one with physically distinct LAEs.

jectivity, these presentations strive for the property of *smoothness*, i.e., small changes in the atomic configuration result in proportionally small changes in the descriptor.

For describing environments in the rigid lattice, there exists a special case of descriptors. Since the LAE is only defined by the states (either empty or occupied by one atom) of the surrounding lattice sites, the configuration can be described simply as a list of those states. The LAE including N lattice sites is an N-dimensional integer vector in this encoding. See Fig. 6.3 for an illustration of this descriptor in mono-elemental fcc system with lattice sites up to the 2nn distance included in the LAE. Since the encoding includes 26 neighboring lattice positions in total, it is called the 26-dimensional (26D) descriptor. This type of input encoding was used for training AIs to predict migration barriers in various atomic systems in Refs. [68, 70, 88–92, 94, 95, 97, 98], as well as in Publications **I**–**V**.

As the integer vector descriptor does not hold any information on the three-dimensional locations of atoms, it is definitely invariant in translations, as well as *most* rotations and reflections. Owing to the symmetry of the elongated octahedral LAE itself, the arrangement of the lattice sites maps to itself under two-fold rotations about three perpendicular axes, as well as in inversion about the center point (in Schoenflies notation [136, 137], this symmetry group is D_{2h}). When the jumping direction is taken into account, it breaks the symmetries perpendicular to the jump, reducing the symmetry group to C_{2v}. Now, when some lattice sites are occupied by atoms, any configuration that *also has the* C_{2v} *symmetry*, is invariant in two-fold rotations and reflections, as desired. However, those configurations that break the C_{2v} symmetry have either two or four alternative descriptors that correspond to the same physical shape. In Publications **III**–**V**, this ambiguity was resolved by systematically choosing the encoding with the lowest numerical

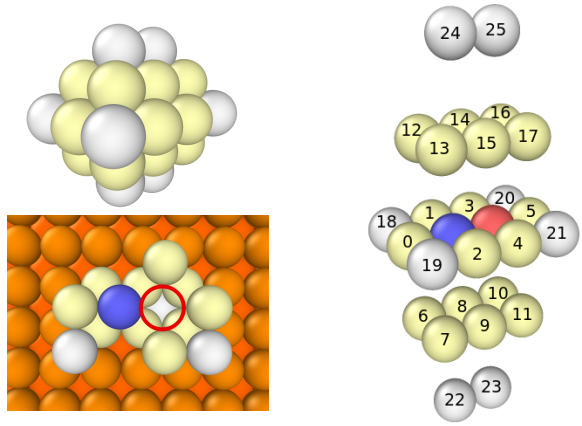


Figure 6.3: LAE descriptor as an integer vector listing the occupation states of every lattice site. The vector indices of the sites, 26 in total, are shown in the right-hand side panel. The initial site of the jumping atom itself is marked in blue, and its final site in red. With the LAE extending to the 2nn distance of the initial *and* final position of the jumping atom, its shape is an elongated octahedron. With vacant sites marked as 0's and occupied sites as 1's, the LAE in the bottom left panel would be encoded as "10011111111100000001011100". [Publication **III**]

value, when the 26D vector was written out as a binary number.

Calculating migration barriers and training neural networks

With the formalism for input encoding and the general algorithm for barrier calculations—NEB—decided, there still remain non-trivial issues regarding the production of training data for the ML algorithm. While the 26D descriptor exhibits the excellent property of easy conversion from integer vector to 3D structure and back, the neighborhood *beyond* the descriptor volume still needs to be defined somehow for the NEB calculation. At the extremes, the LAE could be completely embedded inside bulk material, or, alternatively, the LAE "cluster" of atoms could float free in the vacuum. Intermediate solutions were investigated in Publications **I**–**III**.

The second question regarding training data generation is the sampling scheme of the 26-dimensional space. In monoelemental lattice with a large number of vacancies (such as nearby a surface), or, equivalently, inside the bulk of a binary lattice, the 26D descriptor can have $2^{26} \approx 67$ million different values. Due to the symmetry of the LAE, roughly an eighth of this number of barrier calculations are required for a complete table of parameters (assuming the forward and the backward barriers can be resolved in a single calculation for each process). Regardless, depending on the cost of the barrier calculations, this space may be unfeasible to cover entirely, and choosing a sampling method becomes relevant.

The first barrier set using the 26D encoding was calculated in Publications **I**–**II**. This set was not used to train an AI, but rather to enable comparison of 4D described barriers to a more accurate description (see Fig. 6.4). 334 725 different 26D configurations were selected randomly for this set, called *Cu set 3* in Publi-

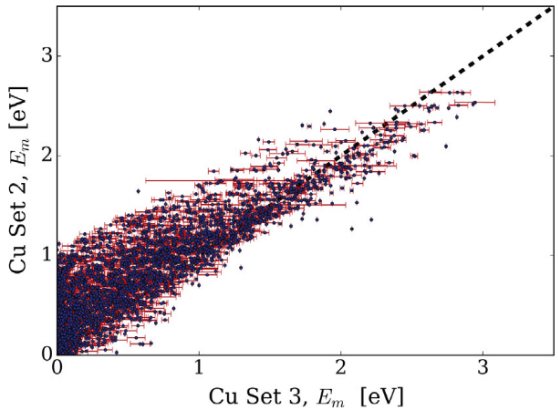


Figure 6.4: Comparison of 4D described barriers (Cu set 2) to 26D described barriers (Cu set 3). Mapping of 26D to 4D is many-to-one, with the dark blue points indicating the mean and the red bars indicating the standard deviation of every 26D barrier corresponding to the given 4D barrier. [Publication **I**]

cations **I**–**II**. The tethering force constant was set to a sufficiently high level to enable successful calculation of all selected barriers: $2.0\,\mathrm{eV/\AA^2}$.

At first, the method used to define the atomic environment beyond the 2nn atoms was to embed the LAE completely within crystalline Cu bulk. However, this resulted in abnormally high barriers, in the order of tens of eV for some configurations. The reason was found to be the negative stress of the LAEs with very low number of atoms: due to attractive interactions, forces became very high at some points of the MEP. Thus, all barriers initially higher than 1.5 eV were recalculated with the LAE embedded on the Cu surface instead.

The first barrier sets used for ML purposes were calculated in Publications **III**–**IV**. Teaching MLPs and RBFNs was initially tested on small barrier sets representing "flat" surfaces, i.e., perfect {1 0 0}, {1 1 0}, or {1 1 1} surfaces with the LAE varying only in the adatom layer. As these sets are disjoint, the resulting AIs are difficult to apply to arbitrarily rough surfaces where different facets can appear. Furthermore, jumps into unstable positions, enabling more flexible kinetics, are completely absent. Thus, a more complete training set was calculated. The 26D configurations were selected by thorough sampling with one stability criterion: when the LAE was embedded in the surface, *if every atom in the LAE had at least three 1nn atoms around it*, the configuration was accepted. For the details of the embedding scheme, see Publication **III**. Three neighbors is an approximately the lower bound for stabilizing an atom, although it is not always sufficient. Complemented with the tethering force, it was possible to calculate the barriers for 11 652 085 configurations. These calculations were a serious computational effort, taking many months to complete.

As the embedding was always done in one of the low-index surfaces, this barrier set can also be considered to consist of three disjoint sets. We were able

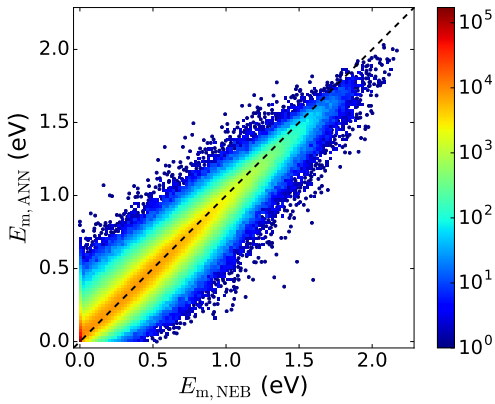


Figure 6.5: Correlation between ML predictions ($E_{\mathrm{m,\,ANN}}$) and the original barriers used in training ($E_{\mathrm{m,\,NEB}}$). Color coding is according to the density of points, with the majority of points lying on the identity curve (dashed line). [Publication **III**]

to construct a more accurate barrier prediction model by training separate neural networks to each of these three sets, and combining them with a classifier that decides which networks should be called for each 26D environment encountered during KMC simulation; see Fig. 6.5.

Programmatic considerations

The ML library that was decided to be integrated into the Kimocs KMC code was the Fast Artificial Neural Networks (FANN) [138]. It is written in the same language as Kimocs, namely C, and implements all necessary features for the project. Training of the neural networks was carried out in a separate program, producing files containing the network architecture (number of layers and their widths and connections between nodes) and the weights of the connections. Routines to read these network files and to call the networks for output for given encoded LAE input were implemented in Kimocs.

6.1.3 Kinetic Monte Carlo simulations with machine learned barriers

In addition to calculating sets of migration barriers and training neural networks to perform barrier prediction in Publication **III**, KMC simulations using these networks were run to validate the predictive model.

First, the attempt frequency ν of Eq. (4.8) was fitted by Eq. (5.5) according to the scheme described in Sec. 5.2.1. We used the same process for the fitting of ν as in Ref. [25], namely the flattening of a nanocuboid on the $\{1\,1\,0\}$ surface. After calibration of ν, the surface energies predicted by the model were validated by running simulations for nanoparticle shape equilibration. The nanoparticle shapes can be compared to the theoretically optimal shape given by Wulff construction.

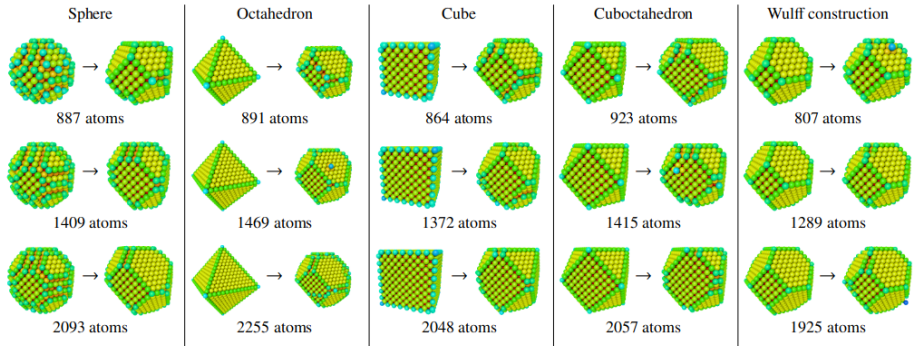


Figure 6.6: Relaxation of nanoparticles from various initial configurations, using machine learned barriers in KMC simulation. The Wulff construction in the rightmost column is the theoretical prediction for the optimal shape of particles of similar size according to the surface energies given by the interatomic potential energy function used in the training set barrier calculations. [Publication **III**]

Table 6.1: Flattening times of cuboid nanotips at 1000 K on Cu surfaces of different orientations. [Publication **III**]

Surface	KMC+ML (ns)	KMC [25] (ns)	MD [25] (ns)
{1 0 0}	5.2 ± 0.6	31.0 ± 6.61	1.62 ± 0.60
{1 1 0}	7.5 ± 0.7	9.25 ± 1.10	9.29 ± 1.44
{1 1 1}	4.8 ± 0.9	18.8 ± 0.96	6.01 ± 1.48

The kinetics were validated by simulating cuboid nanotips on the {1 0 0} and the {1 1 1} surfaces, to be compared to corresponding simulations in Ref. [25]. Finally, systems of nanowires, both individually and crossing at 90° angle, were simulated, and compared to simulations of the similar fcc metal Au in Ref. [58], and some existing experiments of Cu nanowires [139, 140]. All simulations were carried out at 1000 K.

The model predicts the surface energies in Cu well, as evidenced by the relaxation of nanoparticles close to the equilibrium shape given by Wulff construction according to the the underlying potential energy function (see Fig. 6.6). The stability of nanotips on the {1 0 0} and the {1 1 1} surfaces were observed to agree better with MD simulations than the corresponding experiments using the 4D parametrization. The comparison is shown in Tab. 6.1. See Fig. 6.7 for examples of the nanotip simulations.

Thin Cu nanowires exhibited similar behavior as corresponding systems of Au nanowires in KMC simulations and experiments of Ref. [58]. The wires are much more stable individually, taking 250±130 ns to fragment into pieces, than when they are crossed close to one another: the crossing wires fragment in the junction in already 8±4 ns. Snapshots of the simulations are shown in Fig. 6.8. Fragmentation at Cu nanowire junctions has also been observed experimentally [139, 140].

Finally, it is worth mentioning that during the simulations on the {1 1 0}

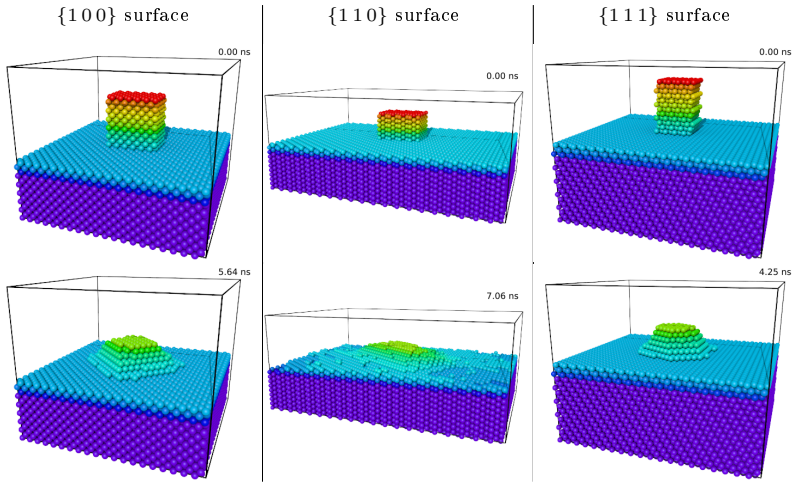


Figure 6.7: Examples of the initial and final frames of the cuboid nanotip flattening simulations. Top row is the initial configuration, and bottom row is the final. [Publication **III**]

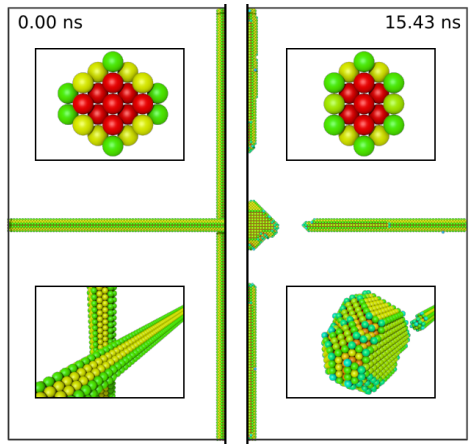


Figure 6.8: Fragmentation of Cu nanowires at the junction. Left panel is the initial configuration, and the right panel is the final. [Publication **III**]

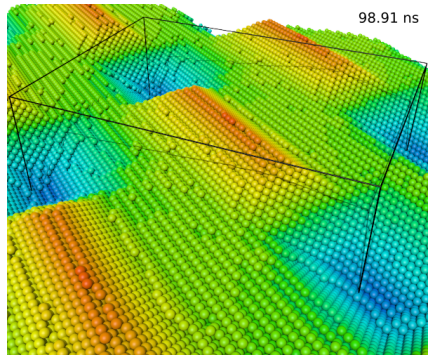


Figure 6.9: Thermal instability of the Cu {1 1 0} surface at 1000 K. [Publication **III**]

surface, a phenomenon not observed in the earlier KMC models emerged: self-roughening at elevated temperatures (see Fig. 6.9). This is known to occur in experiments roughly at temperatures of 80 % of the Cu melting point, or 900–1070 K [141, 142]. It has also been observed in MD simulations [143]. In Publication **III**, the Cu {1 1 0} surface destabilized at temperatures above 950 K. Further research is needed to verify whether the mechanism of roughening in the model is realistic, given that the validity of the rigid lattice approximation is already pushed at temperatures this close to the melting point of Cu.

6.1.4 Increasing the neural network descriptor resolution

As described earlier in Sec. 6.1.2, the 4D encoding of the LAE *hides* some information of the geometry of the environment, namely the precise locations of the 1nn and 2nn atoms around the jump. Including this information can be considered an *increase in resolution* of the descriptor, with each 4D encoded LAE splitting into multiple different 26D ones.

However, the same statement of hiding information can be made of the 26D description as well, regarding the neighbors outside the 2nn shell. In the calculation of the barriers the definition of the further environment is left to the arbitration of the user, with different choices producing slightly different barriers.

The effect of the atoms further in the environment—another increase in the resolution of the descriptor—was investigated in the Master's Thesis by Jonna Romppainen [144]. In that work, a 72-dimensional integer descriptor, up to the fourth nearest neighbors of the jumping atom, was constructed with the same principle as the 26D descriptor. The 72D binary space has impressive 4.7×10^{21} possible values, compared to the mere 6.7×10^{7} of the 26D space. The extreme size of the space makes the choice of sampling scheme of utmost importance. Combining different heuristic sampling methods, a total of 4.5 million configurations were chosen for barrier calculations.

Comparison between the 26D and the 72D barrier sets revealed an effect within ± 0.5 eV to most barriers, with rare exceptions of up to 2.0 eV difference

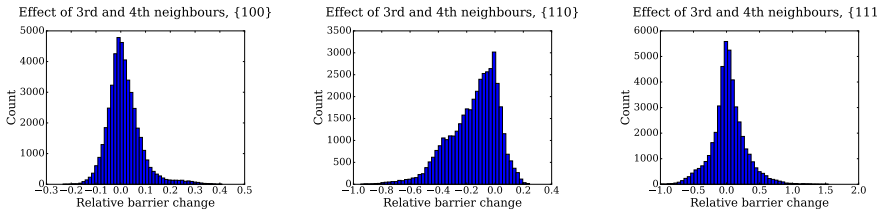


Figure 6.10: Relative change in 72D barriers (including third and fourth neighbors) compared to the corresponding 26D barriers. From Ref. [144].

of some 72D barriers to the corresponding 26D barrier. Histograms depicting the relative barrier changes are shown in Fig. 6.10.

Such large differences depending on the descriptor resolution indicate the need to expand the LAE beyond the 2nn atoms. However, when neural networks trained on the 72D parameterized barriers were used in KMC simulations, the physical behavior of the system was not strictly better than with the earlier 26D parametrization. The self-roughening instability of the {1 1 0} surface, observed in Publication **III**, disappeared, but in turn the stability of the {1 1 1} surface was overestimated. In the nanotip flattening simulation on the {1 0 0} surface, the cuboid tip relaxed to an indefinitely stable pyramidal shape with {1 1 1} facets. This is contrary to MD simulations in the corresponding system, as well as the earlier parametrizations of the Kimocs code. The lesser accuracy in some simulation settings, despite the higher resolution of the LAE encoding, highlights the importance of the sampling scheme in the 72D space. With only a negligible slice of the configurations—a fraction of $\sim 10^{-16}$—within the reach of computational resources, the LAEs must be chosen very carefully to provide good accuracy for the trained ML models *in the regions most important for processes simulated by the KMC model.*.

When the model can describe simple processes, such as nanotip flattening, in a qualitatively correct manner, more qualitative tests can be devised. In addition to estimating the thermodynamic stability of different surface facets, as was done in Publication **III**, the diffusion constants in various geometries could be compared to the corresponding results in MD simulations using the same potential energy function that was used in the training set barrier calculations. This may require the acceleration of MD, for which some techniques were introduced in Sec. 5.4.2.

6.1.5 Interpretable radial basis function networks

A widely encountered issue in ML is the lack of transparency of the trained structures as well as the predictions they make. A neural network, for instance, is a collection of weight parameters that have no physical interpretation. It is a truly "black box", that produces output that we hope approximates the desired function sufficiently well.

In addition to the lack of straightforward interpretation, efforts to visualize

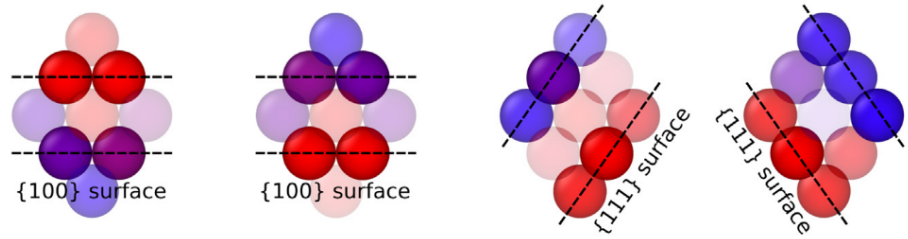
the AI are hindered by the vast amount of parameters they contain. E.g., even the smallest networks with one hidden layer trained in Publication **III** have almost 1000 weight parameters, while the largest cascade networks with 70 shortcut-connected nodes have more than 4000.

RBFNs have a slight advantage over MLPs and especially the more complicated deep neural networks: as was explained in Sec. 5.3.2, the RBFN parameters consist of *centroids* and a single layer of weights between the radial basis functions and the output node(s). The centroid vectors have dimensionality equal to the input data, allowing direct physical interpretation of their coordinates. Even so, in the case of hundreds or thousands of centroids, interpretation of the network as a whole is challenging.

In Publication **V**, we explored a new method to make RBF networks more easily interpretable. Using the barriers calculated in Publications **III**–**IV** as training data sets, we first trained conventional RBF networks using gradient descent to simultaneously optimize both the centroids and the weight parameters. These initial RBFNs still elude interpretation, containing thousands of centroids. We subsequently *pruned* these networks down to only sixteen centroids, by minimizing the expected deviation between the large, original network and the small, pruned network, *over the full distribution of the input data*. The expectation can be calculated analytically by assuming the distribution of input variables. This way, the pruned network approximates the large network globally for input data distributed like this, not only within the actualized distribution of the training set. In the publication, we considered distributions with the Gaussian mixture probability density, the uniform density, and the Bernoulli density.

The sixteen centroid vectors of the pruned networks can provide a compact visualization the relevant patterns discovered during training. The centroids lie in the same 26-dimensional space as the encoded LAE input—by plotting the components of the centroids as "atoms" with opacity proportional to their magnitudes, different patterns become visible. This is illustrated in Fig. 6.11. Panels (a) and (b) are example centroids from the network taught with {1 0 0} barrier set, and panels (c) and (d) are from the one taught with {1 1 1} barriers. The corresponding surfaces are shown on top of the images: the centroids have anti-symmetry with respect to the surface orientations.

To be clear, the purpose of the pruned RBFNs is not to provide migration barrier predictions to be used in physical simulations—even though they approximate the original networks everywhere for a given input data distribution, their predictive accuracy is reduced compared to the original. However, the centroid vector patterns can guide the design of new descriptors for LAEs of this type.



(a) {1 0 0} (10) (b) {1 0 0} (14) (c) {1 1 1} (1) (d) {1 1 1} (6)

Figure 6.11: Examples of pruned RBFN centroid vectors visualized as atoms, colored by the sign of the vector component and opacities proportional to the magnitudes of the components. [Publication **V**]

6.2 Biased diffusion observed in molecular dynamics in electric field gradients

KMC, while quite ideal for simulating atomistic diffusion in crystalline matter, faces numerous challenges when brought from the bulk onto the surface. The rigid lattice approximation becomes less accurate, leading to issues with defining the necessary migration events. These challenges and their solutions, including ML techniques, were explored in length in Publications **I**–**V**.

The ultimate goal of this thesis is to simulate diffusion in *electric field*. Electric fields can be incorporated in KMC by modification of the migration energy barriers by surface polarization, discussed in Sec. 4.3. However, the estimation of the polarization characteristics in different LAEs is a complicated and computationally expensive problem, with solutions not yet sufficiently mature for simulations beyond simple approximations [108]. Hence, in the last publication in the scope of this work, we applied MD coupled with FEM electrostatic solver to circumvent the problem of direct estimation of polarization. As explained in Sec. 5.4.3, the solution of the electric field in continuous mesh in the given boundary conditions produces the correct overall polarizability, and any local effects—except for the permanent dipole moment—are taken into account by refining the mesh around surface atoms of interest.

The simulations in Publication **VI** were carried out with the LAMMPS molecular dynamics program [145]. The FEM solver Femocs [27] is readily interfaced with LAMMPS, earlier applied to simulate, e.g., catastrophic thermal runaway processes [29] without emphasis on slower evolution by diffusion.

To allow observation of surface diffusion over a long timescale in a stable environment, the effects of the field emission current were not considered. The dynamic region of the simulation setup was limited to a thin slice of a nanotip formation, extended with a static FEM mesh to provide a realistic electric field gradient along the tip. The goal of the study was to investigate whether the MD-FEM model can capture the biased diffusion in electric field gradient, predicted

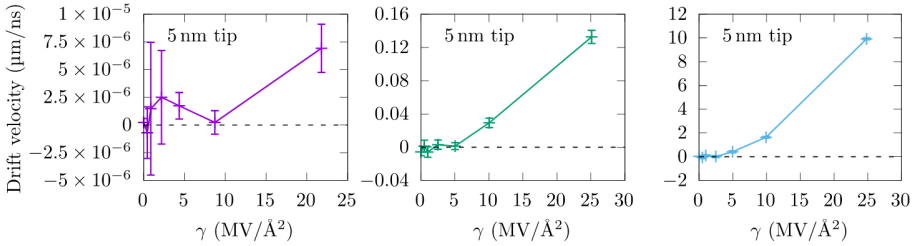


Figure 6.12: Drift velocity, indicating the bias of diffusion, as the function of electric field gradient γ on different surfaces. The simulated system was a 20 Å thick slice modeled in MD, set in a static 5 nm tall nanotip. [Publication **VI**]

Table 6.2: Permanent dipole moment difference $\mathcal{M}_{\mathrm{sr}}$ and the polarizability difference $\mathcal{A}_{\mathrm{sr}}$ between the saddle point and the reference system, as estimated in MD and in DFT, on different Cu surfaces. [Publication **VI**]

Surface	$\mathcal{M}_{\mathrm{sr}}$ (eÅ)		$\mathcal{A}_{\mathrm{sr}}$ (eÅ^2/V)	
	MD	DFT	MD	DFT
{1 0 0}	0.0 ± 0.2	0.106 ± 0.003	0.1 ± 0.1	0.27 ± 0.02
{1 1 0}	-0.008 ± 0.005	0.094 ± 0.006	0.034 ± 0.008	0.30 ± 0.04
{1 1 1}[a]	-0.016 ± 0.006	0.162 ± 0.003	0.02 ± 0.01	0.23 ± 0.02

[a] The saddle point value for {1 1 1} surface is approximated by the value at lattice position ($\mathcal{M}_{\mathrm{lr}}$, $\mathcal{A}_{\mathrm{lr}}$).

by Eq. (4.13).

On the Cu {1 0 0} surface, having the highest migration energy barrier of ~0.5 eV among the low-index surfaces, the dynamics were accelerated by CVHD (see Sec. 5.4.2).

Examples of the results are shown in Fig. 6.12. The growing tendency of the drift velocity, i.e., the average total displacement of migrating adatoms during the simulation divided by the simulation time, is plotted against the strength of the electric field gradient.

In addition to qualitative observation of biased diffusion, the realization of the theoretical prediction from Eq. (4.13) can be tested by fitting Eq. (4.21) to the simulation data for each surface orientation estimating the polarization characteristics of the surfaces can be estimated. In Publication **VI**, the same characteristics were also calculated by DFT, for comparison. The results are shown in Tab. 6.2.

A large discrepancy can be seen between the two simulation techniques in the estimates of the permanent dipole moment property $\mathcal{M}_{\mathrm{sr}}$. This can be explained by the missing Smoluchowski effect: the purely electrostatic system coupled with MD will not become polarized by adatoms, unlike the quantum mechanical system modeled in DFT. As for $\mathcal{A}_{\mathrm{sr}}$, the results differ by an order of magnitude. This is expected to be due to the different geometries used in the two methods: a nanotip for MD and a slab system for DFT. A direct comparison of the results calculated

in the same geometry is a subject of future studies.

The important conclusion of Publication **VI** is, that the MD-FEM model describes the electrostatic interaction of the atoms and the field to a sufficient degree to reproduce the effect of biased surface diffusion in electric field gradients. Thus, MD simulations can be considered for the future work of direct atomistic modeling of diffusive mass transport as a mechanism of surface protrusion growth and faceting under the field, as well as for validating KMC models developed for the same purpose.

Chapter 7

Summary and outlook

In this thesis, self-diffusion on Cu surfaces was investigated using lattice kinetic Monte Carlo (KMC) and molecular dynamics (MD), with and without applied external electric fields. The work represents an early effort to develop a comprehensive model of diffusion in complicated surface systems subject to external forces.

A major part of the studies revolved around the development of the parametrization of the surface KMC model. After identifying a number of challenges in the rigid lattice approach to surface diffusion modeling, machine learning was investigated as one possible solution. By employing artificial neural networks, the KMC model was improved to predict the stability of differently oriented nanostructures more accurately than the earlier parametrization. The self-roughening of the Cu {1 1 0} surface at temperatures near 1000 K was also observed.

Creating an artificial neural network requires training data, which is a useful object also for further research, both in machine learning model development and in the deeper investigation of the data itself. All such data sets produced within this work were also published in compact parallel publications for ease of future use. Within this thesis, we introduced a novel gradient descent -based method to train radial basis function networks, as well as to *prune* them into smaller, more easily interpretable structures. Studying the patterns in the data, found in the pruned network parameters, can guide the search for new, more efficient data encoding schemes for machine learning in materials science.

The thesis culminates in an MD study of biased self-diffusion on Cu in electric field gradient. Successful simulation of this phenomenon with feasible computational resources required metadynamical acceleration of the dynamics, as well as coupling the MD to an electrostatic continuum model. The theoretically predicted qualitative trend of diffusion bias toward stronger electric fields was captured correctly. This proves MD simulations a valid tool in atomistic modeling of the growth and faceting of nanotips under electric fields.

In terms of outlook, the first item that should be mentioned is the continued development of the machine learning KMC method. While the technique

presented here surpasses the earlier models in some aspects, it is possible that the elevated temperature behavior is partially artificial due to the rigid lattice approach. An accurate, fully three-dimensional KMC model of surface diffusion may be possible to realize with a lattice-free approach, where machine learning prediction of migration energy barriers can offer even more than on-lattice. Sophisticated atomic configuration descriptors designed for continuously defined coordinates in the context of machine learned interatomic potentials could be applied in an off-lattice KMC model as well. Generating the barrier training data by active learning during the KMC simulation could solve many of the issues related to sampling in extremely high dimensions.

For more efficient simulation of diffusion in electric field, further research on implementing the coupling between KMC and electrostatics is in order. The topic was mentioned in this dissertation only in passing, due to the new barrier prediction scheme did not reaching sufficient maturity for expecting the electrostatics coupling to give new insight. The earlier work in this area relies on simplifying assumptions regarding the surface polarization characteristics. Machine learning could be applicable here as well, predicting the dipole moments and polarizabilities in a given local atomic environment, based on either density functional theory or lower level charge equilibration calculations.

Finally, encouraged by the success of the performance of the MD simulations, I expect further studies in this direction to prove fruitful. The immediate task would be verifying a better quantitative agreement between the classical electrostatics model and *ab initio* calculations. It may be possible to apply a realistic electric field gradient by a simpler geometry directly to a slab-like MD system, for easier direct comparison to density functional theory.

These developments could power simulations able to seek the large scale mass transport mechanism that could contribute to the growth of field emitters, suspected to play a role in vacuum arc breakdowns.

Acknowledgements

I have been privileged to conduct my doctoral studies in an excellent group in an excellent university. Furthermore, there are many individuals who deserve special acknowedgement for the various ways they contributed toward the completion of this thesis.

First, I must thank my supervisor, Prof. Flyura Djurabekova, for guiding me all the way from my Master's Thesis work to the defence of the doctoral dissertation. You always stand up for your students and give them your full support and scientific expertise. You have kept me in track, while allowing me true academic freedom to explore and develop into a confident researcher of my own. I could not have wished for a better supervisor.

I gratefully acknowledge Profs. Adam Foster and Milica Todorević for accepting the invitations to pre-examine this book and for providing invaluable feedback to improve it. I am very thakful to Prof. Lorenzo Malerba for agreeing to serve as the opponent in the public examination.

There are a few giants on whose shoulders I have been able to stand tall and see far. Dr. Ville Jansson, you not only authored the code I used for most of the work presented here, but have also been a great academic role model for me. I have done my best to keep alive your tradition of delightfully bone-dry humor in the lab. Ekaterina Baibuz, you taught me many of the techniques vital for conducting this research. Drs. Simon Vigonski and Mihkel Veske, your skillful programming and expert advice has been immeasurably important for the success of this work. Prof. Andreas Kyritsakis, I thank you for sharing your knowledge of electrostatics, and also for our incredibly interesting discussions outside of academic topics. Viacheslav Bazaliy, Carlos Santana Vega, Dr. Jussi Määttä and Prof. Teemu Roos, you showed me a glimpse into the world of serious scholars of machine learning, and I tried and will try my best to make the most of this arcane practice. Doc. Antti Kuronen, you taught me most of what I know of computational methods in physics, and furthermore are the one person who never misses the lunch in Unicafe, even when others stray from the path.

I am thankful to all my friends in the lab. Henrique, I greatly enjoyed sharing the office with you while I did. Anton, Veronika, Jian, Alvaro, Milad, Junlei, Hilal, I thank you for all of our scientific collaboration and discussions. Ville Jantunen, as I said in your karonkka party, you are like my academic twin brother, giving me crucial pieces of information to navigate through the bureaucracy leading to

the defence. Morten, you were my Linux mentor when I arrived in the lab, barely knowing git from cat. Special thanks go to the bright students, Jonna and Mikko, whose Master's Theses I had the pleasure to co-supervise.

I thank my parents for supporting me during the very early stages of the process that has ultimately lead to the defence of this thesis.

Lastly, but most importantly, I thank my spouse, Aura. You have given me your unfaltering support throughout the ten years we have been together. You have made me the person, the scientist and the father I am today. I love you.

Bibliography

[1] I. I. Beilis. "Vacuum Arc Cathode Spot Theory: History and Evolution of the Mechanisms". In: *IEEE Transactions on Plasma Science* 47.8 (Aug. 2019), pp. 3412–3433.

[2] CLIC, M. J. Boland, U. Felzmann, P. J. Giansiracusa, T. G. Lucas, R. P. Rassool, C. Balazs, T. K. Charles, K. Afanaciev, I. Emeliantchik, et al. "Updated Baseline for a Staged Compact Linear Collider". In: *CERN Yellow Reports* 4.0 (2016).

[3] N. Shafqat, C. Serpico, and T. G. Lucas. "Design and High-Power Test of a Short Prototype of High Gradient S-band Accelerating Structure for the FERMI Free Electron Laser Linac Upgrade". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 979 (Nov. 2020), p. 164473.

[4] G. M. McCracken. "A Review of the Experimental Evidence for Arcing and Sputtering in Tokamaks". In: *Journal of Nuclear Materials* 93–94 (Oct. 1980), pp. 3–16.

[5] K. Heitmann, N. Frontiere, C. Sewell, S. Habib, A. Pope, H. Finkel, S. Rizzi, J. Insley, and S. Bhattacharya. "THE Q CONTINUUM SIMULATION: HARNESSING THE POWER OF GPU ACCELERATED SUPERCOMPUTERS". In: *The Astrophysical Journal Supplement Series* 219.2 (Aug. 2015), p. 34.

[6] A. W. Chao and W. Chou. *Reviews of Accelerator Science and Technology - Volume 2: Medical Applications of Accelerators*. World Scientific, Dec. 2009. ISBN: 978-981-4299-34-3.

[7] M. Aicheler and CERN, eds. *A Multi-TeV Linear Collider Based on CLIC Technology: CLIC Conceptual Design Report*. CERN 2012,7. Geneva: CERN, 2012. ISBN: 978-92-9083-379-6.

[8] J. Wesson. *Tokamaks*. Oxford University Press, 1997.

[9] P. G. Slade. *The Vacuum Interrupter: Theory, Design, and Application*. Boca Raton: CRC Press, Jan. 2017. ISBN: 978-1-315-22161-8.

[10] I. Brown and E. Oks. "Vacuum Arc Ion Sources: Recent Developments and Applications". In: *IEEE Transactions on Plasma Science* 33.6 (Dec. 2005), pp. 1931–1943.

[11] R. L. Boxman and V. N. Zhitomirsky. "Vacuum Arc Deposition Devices". In: *Review of Scientific Instruments* 77.2 (Feb. 2006), p. 021101.

[12] A. Saressalo, A. Kilpeläinen, K. Mizohata, I. Profatilova, A. Nolvi, I. Kassamakov, P. Tikkanen, S. Calatroni, W. Wuensch, and F. Djurabekova. "In-Situ Plasma Treatment of Cu Surfaces for Reducing the Generation of Vacuum Arc Breakdowns". In: *Journal of Applied Physics* 130.14 (Oct. 2021), p. 143302.

[13] A. Descoeudres, Y. Levinsen, S. Calatroni, M. Taborelli, and W. Wuensch. "Investigation of the Dc Vacuum Breakdown Mechanism". In: *Physical Review Special Topics – Accelerators and Beams* 12.9 (2009), p. 092001.

[14] I. Profatilova, X. Stragier, S. Calatroni, A. Kandratsyeu, E. Rodriguez Castro, and W. Wuensch. "Breakdown Localisation in a Pulsed DC Electrode System". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 953 (Feb. 2020), p. 163079.

[15] Z. Zhou, A. Kyritsakis, Z. Wang, Y. Li, Y. Geng, and F. Djurabekova. "Direct Observation of Vacuum Arc Evolution with Nanosecond Resolution". In: *Scientific Reports* 9.1 (May 2019), p. 7814.

[16] S. Calatroni. "Breakdown Studies for the CLIC Accelerating Structures". In: *Proceedings of Linear Accelerator Conference LINAC2010*. 2010, p. 4.

[17] A. Korsbäck, F. Djurabekova, L. M. Morales, I. Profatilova, E. R. Castro, W. Wuensch, S. Calatroni, and T. Ahlgren. "Vacuum Electrical Breakdown Conditioning Study in a Parallel Plate Electrode Pulsed Dc System". In: *Physical Review Accelerators and Beams* 23.3 (Mar. 2020), p. 033102.

[18] E. Z. Engelberg, A. B. Yashar, Y. Ashkenazy, M. Assaf, and I. Popov. "Theory of Electric Field Breakdown Nucleation Due to Mobile Dislocations". In: *Physical Review Accelerators and Beams* 22.8 (Aug. 2019), p. 083501.

[19] A. J. Ahearn. "The Effect of Temperature, Degree of Thoriation and Breakdown on Field Currents from Tungsten and Thoriated Tungsten". In: *Physical Review* 50.3 (Aug. 1936), pp. 238–253.

[20] G. E. Vibrans. "Vacuum Voltage Breakdown as a Thermal Instability of the Emitting Protrusion". In: *Journal of Applied Physics* 35.10 (Oct. 1964), pp. 2855–2857.

[21] W. Wuensch. "High-Gradient Breakdown in Normal-Conducting Rf Cavities". In: *Proceedings of the 8th European Particle Accelerator Conference, Paris*. June 2002, p. 6.

[22] M. Kildemo, S. Calatroni, and M. Taborelli. "Breakdown and Field Emission Conditioning of Cu, Mo, and W". In: *Physical Review Special Topics - Accelerators and Beams* 7.9 (Sept. 2004), p. 092003.

[23] H. Timko, K. Ness Sjobak, L. Mether, S. Calatroni, F. Djurabekova, K. Matyash, K. Nordlund, R. Schneider, and W. Wuensch. "From Field Emission to Vacuum Arc Ignition: A New Tool for Simulating Copper Vacuum Arcs". In: *Contributions to Plasma Physics* 55.4 (2015), pp. 299–314.

[24] J. Frantz, M. Rusanen, K. Nordlund, and I. T. Koponen. "Evolution of Cu Nanoclusters on Cu(100)". In: *Journal of Physics: Condensed Matter* 16.17 (Apr. 2004), p. 2995.

[25] V. Jansson, E. Baibuz, and F. Djurabekova. "Long-Term Stability of Cu Surface Nanotips". In: *Nanotechnology* 27.26 (July 2016), p. 265708.

[26] F. Djurabekova, S. Parviainen, A. Pohjonen, and K. Nordlund. "Atomistic Modeling of Metal Surfaces under Electric Fields: Direct Coupling of Electric Fields to a Molecular Dynamics Algorithm". In: *Physical Review E* 83.2 (Feb. 2011), p. 026704.

[27] M. Veske, A. Kyritsakis, K. Eimre, V. Zadin, A. Aabloo, and F. Djurabekova. "Dynamic Coupling of a Finite Element Solver to Large-Scale Atomistic Simulations". In: *Journal of Computational Physics* 367 (Aug. 2018), pp. 279–294.

[28] S. Bagchi and D. Perez. "Atomistic Modeling of the Coupling between Electric Field and Bulk Plastic Deformation in Fcc Metals". In: *Physical Review Accelerators and Beams* 25.3 (Mar. 2022), p. 033101.

[29] A. Kyritsakis, M. Veske, K. Eimre, V. Zadin, and F. Djurabekova. "Thermal Runaway of Metal Nano-Tips during Intense Electron Emission". In: *Journal of Physics D: Applied Physics* 51.22 (May 2018), p. 225203.

[30] D. J. Griffiths. *Introduction to Electrodynamics*. Fourth edition. Boston: Pearson, 2013. ISBN: 978-0-321-85656-2.

[31] R. H. Fowler and L. Nordheim. "Electron Emission in Intense Electric Fields". In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. Vol. 119. The Royal Society, 1928, pp. 173–181.

[32] A. Descoeudres, T. Ramsvik, S. Calatroni, M. Taborelli, and W. Wuensch. "Dc Breakdown Conditioning and Breakdown Rate of Metals and Metallic Alloys under Ultrahigh Vacuum". In: *Physical Review Special Topics - Accelerators and Beams* 12.3 (Mar. 2009), p. 032001.

[33] P. A. Chatterton. "A Theoretical Study of Field Emission Initiated Vacuum Breakdown". In: *Proceedings of the Physical Society* 88.1 (May 1966), pp. 231–245.

[34] C. Kittel. *Introduction to Solid State Physics*. 8th ed. Hoboken, NJ: Wiley, 2005. ISBN: 978-0-471-41526-8.

[35] S. B. Sinnott, M. S. Stave, T. J. Raeker, and A. E. DePristo. "Corrected Effective-Medium Study of Metal-Surface Relaxation". In: *Physical Review B* 44.16 (Oct. 1991), pp. 8927–8941.

[36] Z. P. Hu, B. C. Pan, W. C. Fan, and A. Ignatiev. "Structure Analysis of the Cu(110)-(1x2) Surface Reconstruction Induced by Alkali-Metal Adsorption". In: *Physical Review B* 41.14 (May 1990), pp. 9692–9696.

[37] E. E. Gruber and W. W. Mullins. "On the Theory of Anisotropy of Crystalline Surface Tension". In: *Journal of Physics and Chemistry of Solids* 28.5 (May 1967), pp. 875–887.

[38] M. McLean and H. Mykura. "The Temperature Dependence of the Surface Energy Anisotropy of Platinum". In: *Surface Science* 5.4 (Dec. 1966), pp. 466–481.

[39] C. Jayaprakash and W. F. Saam. "Thermal Evolution of Crystal Shapes: The Fcc Crystal". In: *Physical Review B* 30.7 (Oct. 1984), pp. 3916–3928.

[40] S. Fujita and H. Shimoyama. "Mechanism of Surface-Tension Reduction by Electric-Field Application: Shape Changes in Single-Crystal Field Emitters under Thermal-Field Treatment". In: *Physical Review B* 75.23 (June 2007), p. 235431.

[41] H. Yanagisawa, V. Zadin, K. Kunze, C. Hafner, A. Aabloo, D. E. Kim, M. F. Kling, F. Djurabekova, J. Osterwalder, and W. Wuensch. "Laser-Induced Asymmetric Faceting and Growth of a Nano-Protrusion on a Tungsten Tip". In: *APL Photonics* 1.9 (Dec. 2016), p. 091305.

[42] T. T. Tsong and G. Kellogg. "Direct Observation of the Directional Walk of Single Adatoms and the Adatom Polarizability". In: *Physical Review B* 12.4 (Aug. 1975), pp. 1343–1353.

[43] A. Kyritsakis, E. Baibuz, V. Jansson, and F. Djurabekova. "Atomistic Behavior of Metal Surfaces under High Electric Fields". In: *Physical Review B* 99.20 (May 2019), p. 205418.

[44] R. Smoluchowski. "Anisotropy of the Electronic Work Function of Metals". In: *Physical Review* 60.9 (Nov. 1941), pp. 661–674.

[45] H. Ibach. *Physics of Surfaces and Interfaces*. Berlin ; New York: Springer, 2006. ISBN: 978-3-540-34709-5.

[46] B. J. Alder and T. E. Wainwright. "Studies in Molecular Dynamics. I. General Method". In: *The Journal of Chemical Physics* 31.2 (Aug. 1959), pp. 459–466.

[47] V. Jansson and L. Malerba. "Simulation of the Nanostructure Evolution under Irradiation in Fe–C Alloys". In: *Journal of Nuclear Materials* 443.1 (Nov. 2013), pp. 274–285.

[48] C. Domain, C. S. Becquart, and L. Malerba. "Simulation of Radiation Damage in Fe Alloys: An Object Kinetic Monte Carlo Approach". In: *Journal of Nuclear Materials* 335.1 (Oct. 2004), pp. 121–145.

[49] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz. "A New Algorithm for Monte Carlo Simulation of Ising Spin Systems". In: *Journal of Computational Physics* 17.1 (Jan. 1975), pp. 10–18.

[50] G. T. Barkema and N. Mousseau. "Event-Based Relaxation of Continuous Disordered Systems". In: *Physical Review Letters* 77.21 (Nov. 1996), pp. 4358–4361.

[51] R. Malek and N. Mousseau. "Dynamics of Lennard-Jones Clusters: A Characterization of the Activation-Relaxation Technique". In: *Physical Review E* 62.6 (Dec. 2000), pp. 7723–7728.

[52] N. Castin, J. R. Fernández, and R. C. Pasianot. "Predicting Vacancy Migration Energies in Lattice-Free Environments Using Artificial Neural Networks". In: *Computational Materials Science* 84 (2014), pp. 217–225.

[53] F. El-Mellouhi, N. Mousseau, and L. J. Lewis. "Kinetic Activation-Relaxation Technique: An off-Lattice Self-Learning Kinetic Monte Carlo Algorithm". In: *Physical Review B* 78.15 (Oct. 2008), p. 153202.

[54] H. Xu, Y. N. Osetsky, and R. E. Stoller. "Simulating Complex Atomistic Processes: On-the-fly Kinetic Monte Carlo Scheme with Selective Active Volumes". In: *Physical Review B* 84.13 (Oct. 2011), p. 132103.

[55] D. Konwar, V. J. Bhute, and A. Chatterjee. "An Off-Lattice, Self-Learning Kinetic Monte Carlo Method Using Local Environments". In: *The Journal of Chemical Physics* 135.17 (Nov. 2011), p. 174103.

[56] M. A. Lively. "A Hybrid Model of Ion-Induced Syrface Modification with Prompt Molecular Dynamics and Lattice-Free Kinetic Monte Carlo Diffusion". PhD thesis. University of Illinois, 2021.

[57] M. Trochet, N. Mousseau, L. K. Béland, and G. Henkelman. "Off-Lattice Kinetic Monte Carlo Methods". In: *Handbook of Materials Modeling*. Ed. by W. Andreoni and S. Yip. Cham: Springer International Publishing, 2018, pp. 1–29. ISBN: 978-3-319-42913-7.

[58] S. Vigonski, V. Jansson, S. Vlassov, B. Polyakov, E. Baibuz, S. Oras, A. Aabloo, F. Djurabekova, and V. Zadin. "Au Nanowire Junction Breakup through Surface Atom Diffusion". In: *Nanotechnology* 29.1 (Dec. 2017), p. 015704.

[59] C. Domain, C. S. Becquart, and J. C. Van Duysen. "Kinetic Monte Carlo Simulations of FeCu Alloys". In: *MRS Online Proceedings Library* 540.1 (Dec. 1998), pp. 643–648.

[60] C. Domain, C. S. Becquart, and J. C. van Duysen. "Kinetic Monte Carlo Simulations of Cascades in Fe Alloys". In: *MRS Online Proceedings Library (OPL)* 650 (2000), R3.25.

[61] M. J. Hoffmann, S. Matera, and K. Reuter. "Kmos: A Lattice Kinetic Monte Carlo Framework". In: *Computer Physics Communications* 185.7 (July 2014), pp. 2138–2150.

[62] M. Leetmaa and N. V. Skorodumova. "KMCLib: A General Framework for Lattice Kinetic Monte Carlo (KMC) Simulations". In: *Computer Physics Communications* 185.9 (Sept. 2014), pp. 2340–2349.

[63] L. S. Perkins and A. E. DePristo. "Self-Diffusion Mechanisms for Adatoms on Fcc(100) Surfaces". In: *Surface Science* 294.1 (Sept. 1993), pp. 67–77.

[64] L. S. Perkins and A. E. DePristo. "Self-Diffusion of Adatoms on Fcc(110) Surfaces". In: *Surface Science* 317.3 (Oct. 1994), pp. L1152–L1156.

[65] V. Jansson, A. Kyritsakis, S. Vigonski, E. Baibuz, V. Zadin, A. Aabloo, and F. Djurabekova. "Tungsten Migration Energy Barriers for Surface Diffusion: A Parameterization for KMC Simulations". In: *Modelling and Simulation in Materials Science and Engineering* 28.3 (Feb. 2020), p. 035011.

[66] C. L. Liu, J. M. Cohen, J. B. Adams, and A. F. Voter. "EAM Study of Surface Self-Diffusion of Single Adatoms of Fcc Metals Ni, Cu, Al, Ag, Au, Pd, and Pt". In: *Surface Science* 253.1 (Aug. 1991), pp. 334–344.

[67] M. Karimi, T. Tomkowski, G. Vidali, and O. Biham. "Diffusion of Cu on Cu Surfaces". In: *Physical Review B* 52.7 (Aug. 1995), pp. 5364–5374.

[68] O. Trushin, A. Karim, A. Kara, and T. S. Rahman. "Self-Learning Kinetic Monte Carlo Method: Application to Cu(111)". In: *Physical Review B* 72.11 (Sept. 2005), p. 115401.

[69] L. Li and G. Zhou. "3D KMC Simulations of Crater Growth during the Reduction of Oxide Nanoislands on Metal Surfaces". In: *Surface Science* 605.1 (Jan. 2011), pp. 54–61.

[70] A. Latz, L. Brendel, and D. E. Wolf. "A Three-Dimensional Self-Learning Kinetic Monte Carlo Model: Application to Ag(111)". In: *Journal of Physics: Condensed Matter* 24.48 (Oct. 2012), p. 485005.

[71] F. Nita, C. Mastail, and G. Abadias. "Three-Dimensional Kinetic Monte Carlo Simulations of Cubic Transition Metal Nitride Thin Film Growth". In: *Physical Review B* 93.6 (Feb. 2016), p. 064107.

[72] K. J. Laidler. "The Development of the Arrhenius Equation". In: *Journal of chemical Education* 61.6 (1984), p. 494.

[73] H. Eyring. "The Activated Complex in Chemical Reactions". In: *The Journal of Chemical Physics* 3.2 (1935), pp. 107–115.

[74] G. H. Vineyard. "Frequency Factors and Isotope Effects in Solid State Rate Processes". In: *Journal of Physics and Chemistry of Solids* 3.1-2 (1957), pp. 121–127.

[75] U. Kürpick, A. Kara, and T. S. Rahman. "Role of Lattice Vibrations in Adatom Diffusion". In: *Physical Review Letters* 78.6 (Feb. 1997), pp. 1086–1089.

[76] G. Henkelman and H. Jónsson. "Long Time Scale Kinetic Monte Carlo Simulations without Lattice Approximation and Predefined Event Table". In: *The Journal of Chemical Physics* 115.21 (Dec. 2001), pp. 9657–9666.

[77] H. Jonsson. "Methods for Finding Saddle Points and Minimum Energy Paths". In: *Progress on Theoretical Chemistry and Physics*. Kluwer Academic Publishers, Jan. 2000, pp. 269–300.

[78] M. Schroeder, P. Smilauer, and D. E. Wolf. "Bond Counting in a Simulation Model of Epitaxial Growth". In: *Physical Review B* 55.16 (Apr. 1997), pp. 10814–10818.

[79] E. Vincent, C. S. Becquart, C. Pareige, P. Pareige, and C. Domain. "Precipitation of the FeCu System: A Critical Review of Atomic Kinetic Monte Carlo Simulations". In: *Journal of Nuclear Materials* 373.1 (Feb. 2008), pp. 387–401.

[80] F. Soisson, A. Barbu, and G. Martin. "Monte Carlo Simulations of Copper Precipitation in Dilute Iron-Copper Alloys during Thermal Ageing and under Electron Irradiation". In: *Acta Materialia* 44.9 (1996), pp. 3789–3800.

[81] A. Laio and M. Parrinello. "Escaping Free-Energy Minima". In: *Proceedings of the National Academy of Sciences* 99.20 (2002), pp. 12562–12566.

[82] P. Tiwary and M. Parrinello. "From Metadynamics to Dynamics". In: *Physical review letters* 111.23 (2013), p. 230602.

[83] G. Henkelman and H. Jónsson. "A Dimer Method for Finding Saddle Points on High Dimensional Potential Surfaces Using Only First Derivatives". In: *The Journal of Chemical Physics* 111.15 (Oct. 1999), pp. 7010–7022.

[84] G. Henkelman and H. Jónsson. "Improved Tangent Estimate in the Nudged Elastic Band Method for Finding Minimum Energy Paths and Saddle Points". In: *The Journal of Chemical Physics* 113.22 (2000), pp. 9978–9985.

[85] G. Henkelman, B. P. Uberuaga, and H. 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.

[86] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi. "Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons". In: *Physical review letters* 104.13 (2010), p. 136403.

[87] J. Behler. "Neural Network Potential-Energy Surfaces in Chemistry: A Tool for Large-Scale Simulations". In: *Physical Chemistry Chemical Physics* 13.40 (2011), pp. 17930–17955.

[88] F. G. Djurabekova, R. Domingos, G. Cerchiara, N. Castin, E. Vincent, and L. Malerba. "Artificial Intelligence Applied to Atomistic Kinetic Monte Carlo Simulations in Fe–Cu Alloys". In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*. Computer Simulation of Radiation Effects in Solids 255.1 (Feb. 2007), pp. 8–12.

[89] N. Castin, R. Pinheiro Domingos, and L. Malerba. "Use of Computational Intelligence for the Prediction of Vacancy Migration Energies in Atomistic Kinetic Monte Carlo Simulations". In: *International Journal of Computational Intelligence Systems* 1.4 (Dec. 2008), pp. 340–352.

[90] N. Castin, L. Malerba, G. Bonny, M. I. Pascuet, and M. Hou. "Modelling Radiation-Induced Phase Changes in Binary FeCu and Ternary FeCuNi Alloys Using an Artificial Intelligence-Based Atomistic Kinetic Monte Carlo Approach". In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*. Proceedings of the Ninth International Conference on Computer Simulation of Radiation Effects in Solids 267.18 (Sept. 2009), pp. 3002–3008.

[91] N. Castin and L. Malerba. "Prediction of Point-Defect Migration Energy Barriers in Alloys Using Artificial Intelligence for Atomistic Kinetic Monte Carlo Applications". In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*. Proceedings of the Ninth International Conference on Computer Simulation of Radiation Effects in Solids 267.18 (Sept. 2009), pp. 3148–3151.

[92] N. Castin and L. Malerba. "Calculation of Proper Energy Barriers for Atomistic Kinetic Monte Carlo Simulations on Rigid Lattice with Chemical and Strain Field Long-Range Effects Using Artificial Neural Networks". In: *The Journal of Chemical Physics* 132.7 (Feb. 2010), p. 074507.

[93] F. Soisson, C. S. Becquart, N. Castin, C. Domain, L. Malerba, and E. Vincent. "Atomistic Kinetic Monte Carlo Studies of Microchemical Evolutions Driven by Diffusion Processes under Irradiation". In: *Journal of Nuclear Materials*. FP6 IP PERFECT Project: Prediction of Irradiation Damage Effects in Reactor Components 406.1 (Nov. 2010), pp. 55–67.

[94] M. I. Pascuet, N. Castin, C. S. Becquart, and L. Malerba. "Stability and Mobility of Cu–Vacancy Clusters in Fe–Cu Alloys: A Computational Study Based on the Use of Artificial Neural Networks for Energy Barrier Calculations". In: *Journal of Nuclear Materials* 412.1 (May 2011), pp. 106–115.

[95] 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.

[96] N. Castin, M. I. Pascuet, and L. Malerba. "Mobility and Stability of Large Vacancy and Vacancy–Copper Clusters in Iron: An Atomistic Kinetic Monte Carlo Study". In: *Journal of Nuclear Materials* 429.1 (Oct. 2012), pp. 315–324.

[97] L. Messina, N. Castin, C. Domain, and P. Olsson. "Introducing Ab Initio Based Neural Networks for Transition-Rate Prediction in Kinetic Monte Carlo Simulations". In: *Physical Review B* 95.6 (Feb. 2017), p. 064112.

[98] N. Castin, L. Messina, C. Domain, R. C. Pasianot, and P. Olsson. "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), p. 214117.

[99] L. Messina, A. Quaglino, A. Goryaeva, M.-C. Marinica, C. Domain, N. Castin, G. Bonny, and R. Krause. "A DFT-driven Multifidelity Framework for Constructing Efficient Energy Models for Atomic-Scale Simulations". In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 483 (Nov. 2020), pp. 15–21.

[100] K. Sastry, D. D. Johnson, D. E. Goldberg, and P. Bellon. "Genetic Programming for Multitimescale Modeling". In: *Physical Review B* 72.8 (2005), p. 085438.

[101] S. Verma, T. Rehman, and A. Chatterjee. "A Cluster Expansion Model for Rate Constants of Surface Diffusion Processes on Ag, Al, Cu, Ni, Pd and Pt(100) Surfaces". In: *Surface Science* 613 (July 2013), pp. 114–125.

[102] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. "Learning Representations by Back-Propagating Errors". In: *Nature* 323.6088 (Oct. 1986), pp. 533–536.

[103] D. S. Broomhead and D. Lowe. *Radial Basis Functions, Multi-Variable Functional Interpolation and Adaptive Networks.* Tech. rep. Royal Signals and Radar Establishment Malvern (United Kingdom), 1988.

[104] C. M. Bishop. *Neural Networks for Pattern Recognition.* Oxford university press, 1995.

[105] M. Riedmiller and H. Braun. "A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm". In: *IEEE International Conference on Neural Networks.* Mar. 1993, 586–591 vol.1.

[106] C. Igel and M. Hüsken. "Improving the Rprop Learning Algorithm". In: *Proceedings of the Second International ICSC Symposium on Neural Computation (NC 2000).* Vol. 2000. Citeseer, 2000, pp. 115–121.

[107] S. E. Fahlman and C. Lebiere. "The Cascade-Correlation Learning Architecture". In: *Advances in Neural Information Processing Systems.* 1990, pp. 524–532.

[108] V. Jansson, E. Baibuz, A. Kyritsakis, S. Vigonski, V. Zadin, S. Parviainen, A. Aabloo, and F. Djurabekova. "Growth Mechanism for Nanotips in High Electric Fields". In: *Nanotechnology* 31.35 (Aug. 2020), p. 355301.

[109] E. Baibuz, A. Kyritsakis, V. Jansson, and F. Djurabekova. *Polarization Characteristics of Adatoms Self-Diffusing on Metal Surfaces under High Electric Fields.* Jan. 2022.

[110] K. M. Bal and E. C. Neyts. "Merging Metadynamics into Hyperdynamics: Accelerated Molecular Simulations Reaching Time Scales from Microseconds to Seconds". In: *Journal of Chemical Theory and Computation* 11.10 (Oct. 2015), pp. 4545–4554.

[111] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. "A Computer Simulation Method for the Calculation of Equilibrium Constants for the Formation of Physical Clusters of Molecules: Application to Small Water Clusters". In: *The Journal of Chemical Physics* 76.1 (Jan. 1982), pp. 637–649.

[112] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. "Molecular Dynamics with Coupling to an External Bath". In: *The Journal of Chemical Physics* 81.8 (Oct. 1984), pp. 3684–3690.

[113] S. Nosé. "A Unified Formulation of the Constant Temperature Molecular Dynamics Methods". In: *The Journal of Chemical Physics* 81.1 (July 1984), pp. 511–519.

[114] J. E. Jones and S. Chapman. "On the Determination of Molecular Fields.—I. From the Variation of the Viscosity of a Gas with Temperature". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 106.738 (Oct. 1924), pp. 441–462.

[115] N. Bernardes. "Theory of Solid Ne, A, Kr, and Xe at 0° K". In: *Physical Review* 112.5 (Dec. 1958), pp. 1534–1539.

[116] M. S. Daw and M. I. Baskes. "Embedded-Atom Method: Derivation and Application to Impurities, Surfaces, and Other Defects in Metals". In: *Physical Review B* 29.12 (June 1984), pp. 6443–6453.

[117] J. K. Nørskov and N. D. Lang. "Effective-Medium Theory of Chemical Binding: Application to Chemisorption". In: *Physical Review B* 21.6 (Mar. 1980), pp. 2131–2136.

[118] J. D. Kress and A. E. DePristo. "Corrected Effective Medium Method. I. One-body Formulation with Applications to Atomic Chemisorption and Diatomic Molecular Potentials". In: *The Journal of Chemical Physics* 87.8 (Oct. 1987), pp. 4700–4715.

[119] M. S. Stave, D. E. Sanders, T. J. Raeker, and A. E. DePristo. "Corrected Effective Medium Method. V. Simplifications for Molecular Dynamics and Monte Carlo Simulations". In: *The Journal of Chemical Physics* 93.6 (Sept. 1990), pp. 4413–4426.

[120] A. F. Voter. "Parallel Replica Method for Dynamics of Infrequent Events". In: *Physical Review B* 57.22 (June 1998), R13985–R13988.

[121] A. S. Pohjonen, F. Djurabekova, K. Nordlund, A. Kuronen, and S. P. Fitzgerald. "Dislocation Nucleation from near Surface Void under Static Tensile Stress in Cu". In: *Journal of Applied Physics* 110.2 (July 2011), p. 023509.

[122] S. Vigonski, F. Djurabekova, M. Veske, A. Aabloo, and V. Zadin. "Molecular Dynamics Simulations of Near-Surface Fe Precipitates in Cu under High Electric Fields". In: *Modelling and Simulation in Materials Science and Engineering* 23.2 (Mar. 2015), p. 025009.

[123] K. Kuppart, S. Vigonski, A. Aabloo, Y. Wang, F. Djurabekova, A. Kyritsakis, and V. Zadin. "Mechanism of Spontaneous Surface Modifications on Polycrystalline Cu Due to Electric Fields". In: *Micromachines* 12.10 (Oct. 2021), p. 1178.

[124] M. Dixon and M. J. L. Sangster. "Effects of Polarization on Some Static and Dynamic Properties of Molten NaI". In: *Journal of Physics C: Solid State Physics* 9.6 (Mar. 1976), pp. 909–925.

[125] G. Jacucci, I. R. McDonald, and A. Rahman. "Effects of Polarization on Equilibrium and Dynamic Properties of Ionic Systems". In: *Physical Review A* 13.4 (Apr. 1976), pp. 1581–1592.

[126] A. K. Rappe and W. A. Goddard. "Charge Equilibration for Molecular Dynamics Simulations". In: *The Journal of Physical Chemistry* 95.8 (Apr. 1991), pp. 3358–3363.

[127] M. Veske, A. Kyritsakis, F. Djurabekova, K. N. Sjobak, A. Aabloo, and V. Zadin. "Dynamic Coupling between Particle-in-Cell and Atomistic Simulations". In: *Physical Review E* 101.5 (May 2020), p. 053307.

[128] T. Metspalu, V. Jansson, V. Zadin, K. Avchaciov, K. Nordlund, A. Aabloo, and F. Djurabekova. "Cu Self-Sputtering MD Simulations for 0.1–5 keV Ions at Elevated Temperatures". In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 415 (Jan. 2018), pp. 31–40.

[129] G. Ehrlich and F. G. Hudda. "Atomic View of Surface Self-Diffusion: Tungsten on Tungsten". In: *The Journal of Chemical Physics* 44.3 (Feb. 1966), pp. 1039–1049.

[130] R. L. Schwoebel and E. J. Shipsey. "Step Motion on Crystal Surfaces". In: *Journal of Applied Physics* 37.10 (Sept. 1966), pp. 3682–3686.

[131] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. Von Lilienfeld. "Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning". In: *Physical review letters* 108.5 (2012), p. 058301.

[132] B. Huang and O. A. von Lilienfeld. "Communication: Understanding Molecular Representations in Machine Learning: The Role of Uniqueness and Target Similarity". In: *The Journal of Chemical Physics* 145.16 (Oct. 2016), p. 161102.

[133] A. P. Bartók, R. Kondor, and G. Csányi. "On Representing Chemical Environments". In: *Physical Review B* 87.18 (May 2013), p. 184115.

[134] R. Drautz. "Atomic Cluster Expansion for Accurate and Transferable Interatomic Potentials". In: *Physical Review B* 99.1 (Jan. 2019), p. 014104.

[135] G. Dusson, M. Bachmayr, G. Csányi, R. Drautz, S. Etter, C. van der Oord, and C. Ortner. "Atomic Cluster Expansion: Completeness, Efficiency and Stability". In: *Journal of Computational Physics* 454 (Apr. 2022), p. 110946.

[136] C. Schutte, J. Bertie, P. Bunker, J. Hougen, I. Mills, J. Watson, and B. Winnewisser. "Notations and Conventions in Molecular Spectroscopy: Part 2. Symmetry Notation (IUPAC Recommendations 1997)". In: *Pure and applied chemistry* 69.8 (1997), pp. 1641–1650.

[137] N. F. M. Henry and K. Lonsdale. *International Tables for X-ray Crystallography. 1, Symmetry Groups*. Birmingham, England: The Kynoch Press, 1952.

[138] S. Nissen. *Implementation of a Fast Artificial Neural Network Library (Fann)*. Tech. rep. Department of Computer Science University of Copenhagen (DIKU), 2003.

[139] K. Mallikarjuna, H.-J. Hwang, W.-H. Chung, and H.-S. Kim. "Photonic Welding of Ultra-Long Copper Nanowire Network for Flexible Transparent Electrodes Using White Flash Light Sintering". In: *RSC advances* 6.6 (2016), pp. 4770–4779.

[140] J. S. Oh, J. S. Oh, T. H. Kim, and G. Y. Yeom. "Efficient Metallic Nanowire Welding Using the Eddy Current Method". In: *Nanotechnology* 30.6 (Dec. 2018), p. 065708.

[141] S. G. J. Mochrie. "Thermal Roughening of the Copper (110) Surface: An x-Ray Diffraction Experiment". In: *Physical Review Letters* 59.3 (July 1987), pp. 304–307.

[142] K. Kern. "Thermal Dynamics of (110) Fcc Metal Surfaces". In: *Surface X-Ray and Neutron Scattering*. Springer, 1992, pp. 69–71.

[143] H. Häkkinen, J. Merikoski, M. Manninen, J. Timonen, and K. Kaski. "Roughening of the Cu(110) Surface". In: *Physical Review Letters* 70.16 (Apr. 1993), pp. 2451–2454.

[144] J. Romppainen. "Artificial Neural Networks for Parametrisation of a Kinetic Monte Carlo Model of Surface Diffusion". MA thesis. Helsingin yliopisto, 2020.

[145] S. Plimpton. "Fast Parallel Algorithms for Short-Range Molecular Dynamics". In: *Journal of Computational Physics* 117.1 (Mar. 1995), pp. 1–19.