Master's thesis
Master's program in materials research
Computational materials physics

Vacancy diffusion in high-entropy alloys by a kinetic Monte Carlo methodology based on collective variable-driven hyperdynamics and machine learning

Mikko Koskenniemi

November 7, 2023

Supervisor(s): Dr. Jyri Kimari
Prof. Flyura Djurabekova

Examiner(s): Dr. Jyri Kimari
Prof. Flyura Djurabekova

University of Helsinki
Faculty of science

PL 64 (Gustaf Hällströmin katu 2a)
00014 Helsingin yliopisto

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Koulutusohjelma — Utbildningsprogram — Degree programme
Faculty of science	Master's program in materials research Computational materials physics

Tekijä — Författare — Author

Mikko Koskenniemi

Työn nimi — Arbetets titel — Title

Vacancy diffusion in high-entropy alloys by a kinetic Monte Carlo methodology based on collective variable-driven hyperdynamics and machine learning

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidantal — Number of pages
Master's thesis	November 7, 2023	72

Tiivistelmä — Referat — Abstract

High-entropy alloys (HEAs), esteemed for their exceptional resistance to radiation damage, carry considerable potential for deployment within fusion reactors. Nonetheless, due to their compositional complexity, comprehending the diffusion behaviour in these multifaceted alloys continues to be a daunting task. This thesis proposes a novel approach to modelling vacancy diffusion in body-centred cubic (BCC) HEAs, particularly Mo-Nb-Ta-V-W.

The methodology involves the tactical application of collective variable-driven hyperdynamics (CVHD) to procure data for training a Gaussian process regression (GPR) and feed-forward neural network (FNN) model. The trained FNN model is subsequently employed within kinetic Monte Carlo (KMC) simulations for accurately predicting jump rates, whereas the GPR model is used to elucidate experimental findings related to the behaviour of vacancies in Mo-Nb-Ta-V-W. The robustness of the FNN model is manifested by its capacity to generalise to unseen data, whilst the efficacy of the overall method is corroborated by Monte Carlo molecular dynamics (MCMD) simulations.

The CVHD methodology, uniquely capable of functioning at finite temperatures, can capture the entropic contribution to the free energy and model kinematics explicitly. This singular ability facilitates a more comprehensive understanding of the system's behaviour under authentic conditions.

The findings presented in this thesis signify a considerable stride forward in the study of HEAs, providing a robust framework for the design of advanced materials. These results underscore the potential of the CVHD-trained FNN-KMC methodology in exploring complex environments, thereby establishing a firm foundation for future investigations and emphasising the need for its continued refinement and augmentation.

Avainsanat — Nyckelord — Keywords

Molecular dynamics; Collective variable-driven hyperdynamics; Vacancy diffusion; High-entropy alloys

Säilytyspaikka — Förvaringsställe — Where deposited

Muita tietoja — Övriga uppgifter — Additional information

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Koulutusohjelma — Utbildningsprogram — Degree programme
Matemaattis-luonnontieteellinen tiedekunta	Materiaalitutkimuksen maisteriohjelma Laskennallinen materiaalifysiikka

Tekijä — Författare — Author

Mikko Koskenniemi

Työn nimi — Arbetets titel — Title

Vacancy diffusion in high-entropy alloys by a kinetic Monte Carlo methodology based on collective variable-driven hyperdynamics and machine learning

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidantal — Number of pages
Maisterin tutkielma	November 7, 2023	72

Tiivistelmä — Referat — Abstract

Korkean entropian seokset (*high-entropy alloy*, HEA), jotka tunnetaan niiden erinomaisen säteilyvaurion sietokyvyn vuoksi, ovat lupaavia materiaaleja mahdolliseen fuusioreaktorikäyttöön. Koostumuksellisen monimutkaisuudensa vuoksi näiden seosten diffuusiokäyttäytymisen ymmärtäminen kuitenkin on edelleen haastavaa. Tässä tutkielmassa esitetään uutta lähestymistapaa mallintaa vakanssien diffuusiota tilakeskisissä kuutiollisissa (*body-centred cubic*, BCC) HEAssa, nimittäin Mo-Nb-Ta-V-W:ssä.

Tämän tutkielman menetelmä käyttää niinsanotun *collective variable-driven hyperdynamics* menetelmän (CVHD) avulla kerättyä dataa Gaussinen prosessi regressio (GPR) - ja eteenpäin kytketty neuroverkko -mallin (*feed-forward neural network*, FNN) kouluttamiseen. Koulutettu FNN-malli otetaan myöhemmin käyttöön kineettisissä Monte Carlo (KMC) -simulaatioissa vakanssien diffuusio-todennäköisyyksien ennustamiseen, kun taas GPR-mallia käytetään selittämään kokeellisia havaintoja vakanssien käyttäytymisestä Mo-Nb-Ta-V-W:ssä. FNN-mallin luotettavuus ilmenee sen kyvyssä yleistää havaitsemattomaan dataan, kun taas koko menetelmän paikkaansapitävyys vahvistetaan Monte Carlo molekyylidynamiikka (MCMD) -simulaatioiden avulla.

CVHD-menetelmässä, joka ainutlaatuisesti kykenee mallintamaan äärellisiä lämpötiloja, kinematiikka ja entropian kontribuutio vapaaseen energiaan ovat eksplisiittisesti mallinnettavissa. Tämä ainutlaatuinen kyky mahdollistaa järjestelmän käyttäytymisen kattavamman ymmärtämisen todellisissa olosuhteissa.

Tässä tutkielmassa esitetyt tulokset merkitsevät huomattavaa askelta eteenpäin HEA-tutkimuksessa ja tarjoavat vankan kehyksen edistyneiden materiaalien tutkimukselle. Nämä tulokset korostavat CVHD-koulutetun FNN-KMC-menetelmän potentiaalia tutkia monimutkaisia ympäristöjä, perustaen tukevan perustan tuleville tutkimuksille ja korostaen sen jatkuvan parantamisen ja laajentamisen tarvetta.

Avainsanat — Nyckelord — Keywords

Molecular dynamics; Collective variable-driven hyperdynamics; Vacancy diffusion; High-entropy alloys

Säilytyspaikka — Förvaringsställe — Where deposited

Muita tietoja — Övriga uppgifter — Additional information

Contents

1. Introduction

High-entropy alloys (HEAs) constitute a revolutionary category of materials, offering a unique amalgamation of properties that render them highly attractive for high-performance applications. Notably, their exceptional resilience to radiation damage [1] has kindled interest in their potential employment within fusion reactors, a challenging environment where material durability is of utmost importance. Nonetheless, the intrinsic complexity of these alloys, marked by their diverse composition and the vast range of potential atomic configurations, poses a significant hurdle in understanding their diffusion behaviour. This understanding is pivotal as it underlies the long-term evolution of radiation damage, a critical consideration for materials utilised within fusion reactors.

The, as of writing, unpublished experimental findings of Dr. Eryang Lu in W-based HEAs hinted at something profound; in contrast to pure W, initially sparsely-distributed vacancies in an equimolar Mo-Nb-Ta-V-W HEA were found to form clusters when the material was annealed. Due to the long distances involved, one would expect that the binding between vacancies would have a non-existent effect on this clustering, ergo something else must be behind this strange phenomenon. Recently, Ebina et al. [2] formulated a way to apply collective variable-driven hyperdynamics (CVHD) [3]—a flexible and explorative way to accelerate the timescale of molecular dynamics—to accelerate the diffusion of vacancies. In an effort to understand how adatoms behave on copper surfaces when subjected to electric fields, Kimari and collaborators utilised CVHD to model the diffusion of adatoms on various copper surfaces ($\{1\,1\,1\}$, $\{1\,1\,0\}$, and $\{1\,0\,0\}$) [4], which is of interest in vacuum arcing research for linear particle accelerators. Conventional molecular dynamics approaches would have been rendered ineffective for modelling the $\{1\,0\,0\}$ surface at ambient temperatures due to its formidable activation energy of 0.5 eV. However, the enhanced capabilities of CVHD overcame this limitation. Through their rigorous research, Kimari and his team furnished evidence that corroborates the hypothesis suggesting the emergence of surface protrusions on copper surfaces when subjected to an electric field.

The curious findings of Dr. Lu set me out to elucidate the cause. In this thesis, I put forward an innovative approach to modelling vacancy diffusion in body-centred cubic (BCC) HEA systems. These methodologies combine the CVHD method with feed-forward neural network (FNN) and Gaussian process (GPR) regression models, providing potent tools for exploring the complex dynamics of these intricate systems. The GPR model is used to provide qualitative understanding to the diffusivity in HEAs, whilst the FNN model is meant for more quantitative modelling. The effectiveness of the novel CVHD-FNN technique is further corroborated by Monte Carlo molecular dynamics (MCMD) simulations, underscoring its potential as a reliable tool for diffusion studies in HEAs.

The unique ability of CVHD to function at finite temperatures distinguishes it from prevalent methods, such as the nudged elastic band (NEB) [5, 6]. This trait enables the examination of pressure effects elicited by trapped gases within lattice vacancies and captures the entropic contribution to the free energy, providing vital insights into the behaviour of systems under authentic conditions. These insights are especially relevant in the context of fusion reactors, where material embrittlement due to trapped gases [7, 8], namely hydrogen and helium, can potentially precipitate disastrous failures.

However, the path towards a comprehensive comprehension of these complex systems is far from being fully trodden. The current implementation of CVHD supports the modelling of only single vacancies, and overcoming this limitation to enable the modelling of vacancy clusters is a critical forthcoming step. This thesis, therefore, lays the foundation for this endeavour, demonstrating the potential of my CVHD-FNN methodology in investigating these elaborate environments and emphasising the requirement for its continued refinement and expansion.

In essence, this thesis represents a significant stride forward in the study of HEAs. It not only debuts innovative methodologies for modelling and understanding diffusion processes in complex materials but also offers a robust framework that can be adapted and extrapolated to similar systems. In doing so, it signals new possibilities for advanced materials design and paves the way for a more profound understanding of these fascinating materials.

In Sec. 5, I elucidate the CVHD implementation and the parameters utilised. This is succeeded by an in-depth discourse on the transformation of CVHD results

into a training set for an FNN model devised for application in kinetic Monte Carlo (KMC) simulations and a GPR model. I further clarify the methodology applied to validate the FNN's effectiveness in forecasting jump rates for KMC. A comprehensive explanation of my approach to verifying the transferability of the FNN model, initially trained on non-rigid lattices from CVHD simulations, to rigid lattices used in KMC, is also presented. In Sec. 6, I outline my discoveries. Initially, I depict the predictions of the GPR model on different cases, followed by the innovative solutions crafted to tackle the obstacles inherent in applying CVHD in high-temperature settings and intricate lattice configurations. Subsequently, I disclose the outcomes of validating the transferability of the FNN model and assess the performance of the integrated FNN-KMC model. Lastly, in Sec. 7, I collect my insights and distil my findings. This section encapsulates the culmination of my research journey, where I summarise the significant discoveries, evaluate the impact of my work, and contemplate the potential implications of my research. Furthermore, I provide a reflective commentary on the challenges encountered during my investigation and consider opportunities for future research in this domain.

2. Materials for nuclear fusion

Thermonuclear fusion reactors, such as the tokamak, are hostile environments for materials. The immensely large temperature of the fusion fuel means reactor materials have to retain desirable properties at elevated temperatures. Furthermore, the particles emanating from fusion reactions have extraordinarily high kinetic energy, which imposes further problems for the plasma-facing wall materials, because the heat produced when these particles collide with reactor materials is what can be converted to electricity, making them an imperative part of the whole process. However, the aforementioned issues are only the most immediate ones; promising materials, that can seemingly put up with them, might have pitfalls in the longer-term effects, as is the case with the material discussed next, in Sec. 2.1.

2.1 Tungsten

As of now, ITER uses tungsten (W) as a plasma-facing material, namely the divertor is made up of W [9]. The divertor is tasked with absorbing about 15 % of the reactor power output and extracting impurities from the fuel [10]. These impurities include ejecta from high-energy collisions as well as fusion products, primarily helium. Having the highest known melting point of all pure metals [11], and relatively high thermal conductivity, tungsten (W) is well-suited for heat absorption. This enables the use of elevated temperatures, thereby enhancing the fusion rate and, consequently, operational efficiency. Moreover, the mechanical strength of W at high temperatures contributes to its durability and reliability, making it an excellent choice for the rigorous conditions in the divertor. However, certain properties of W do impose limitations on its economical viability and usability.

The immense energies associated with the particles originating from fusion reactions wreak havoc in the plasma-facing materials, introducing copious amounts of defects. Areas containing defects expand or contract in order to minimise the local potential energy. The relaxation volumes associated with defects tend to be quite large [12]; in metals, areas with interstitials expand whereas vacancies contract. In W,

the relaxation volume of a self-interstitial can be upwards for 4.5 times larger in magnitude than of a vacancy (1.68 and -0.37 atomic volumes, respectively) [13]. Abundant defects can therefore induce severe distortions and residual strain in plasma-facing materials, which is where one of W's inherent weaknesses becomes apparent: its brittleness.

W is a hard and brittle material. The maximum strain (measure of deformation) a brittle material can experience before breaking is small compared to ductile materials, often making their fracture toughness, which is related to the energy required to cause a fracture to grow, small. However, some metals can undergo a transition from brittle to ductile above a certain temperature, which is called the brittle-ductile transition [14] temperature. Without any special treatment, W's brittle-ductile transition temperature tends to be more than 300 °C [15, 16], restricting its operating temperature-range.

All in all, radiation-induced strains make W more likely to crack or fracture than its more ductile counterparts, making its operational lifespan, and therefore economical viability, limited. In addition, W has been observed to experience blistering when irradiated by deuterium or helium [7, 8] (deuterium being one of the isotopes used in ITER's fuel, and helium being a product of the deuterium-tritium fusion), which imparts increasing strain to the W lattice.

2.2 Tungsten-based high-entropy alloys

It has been long-known that the properties of metals can be improved by introducing different elements to the metal. For instance, the incorporation of chromium into steel—an alloy— enhances the resistance of steel to corrosion and rust [17, 18].

High-entropy alloys (HEA) are alloys comprising multiple different elements, usually five or more [19], in roughly equal concentrations. If one wishes to improve the properties of W, alloying it with other elements is the next logical step. For example, Mo-Nb-Ta-W can possess greater ductility than W, whilst retaining its mechanical properties under extended periods of elevated temperature better than W (3 days at 1 100 °C) [20]. Moreover, Mo-Nb-Ta-V-W has been shown to better retain its yield-strength at elevated temperatures than austenitic super-alloys [21]. Though, if produced incorrectly, refractory HEAs can be very brittle at room temperature [21, 22].

For fusion applications, it is important to understand how HEAs behave under irradiation; computational methods such as classical molecular dynamics (MD) [23] can be deployed to study the underlying mechanisms of radiation damage in detail

unobtainable by experimental means. The results obtained from classical MD, however, are only as good as the potential used, that is the method of computing interatomic potential energies. Fortunately, a variety of potentials exist, some of which yield results comparable to those obtained from quantum-mechanical simulations. One such potential will be the focus of the subsequent section.

3. Tabulated Gaussian approximation potential: Achieving exceptional accuracy with computational efficiency

In classical mechanics, atomic behaviour is governed by the spatial dependence of their potential energy. The electric potential between atoms is known to decay quickly with increasing interatomic distance. Thus, a popular way to make MD simulations more efficient is to use interatomic potentials, which only take into account nearest neighbouring atoms within a specified distance when evaluating atomic energies. The traditional way of computing atomic potential energy in MD is to use a fixed mathematical form with few parameters, a so-called *analytical* potential. The ability of analytical potentials to predict a wide array of phenomena is, however, limited. Fortunately, there exists a relatively new way to produce accuracy that is on par with quantum-mechanical simulations to the faster, classical MD: machine-learned potentials.

Machine-learned potentials use machine learning to predict atomic energies and have been shown to mimic quantum-mechanical simulation-results to profound success [24, 25], outperforming their traditional, analytical counterparts [26, 27, 28]. The Gaussian approximation potential (GAP) [28] is a widely used approach for developing machine-learned potentials and offers high accuracy when properly constructed. Though faster than quantum-mechanical methods, GAP is tremendously slow compared to most analytical potentials, such as the popular embedded atom method (EAM) potentials. Tabulated GAP (tabGAP) bridges the differences in performance between GAP and analytical potentials, being as much as two orders of magnitude faster than GAP [29]. Evidently, tabGAP is less of a encumbrance to use than GAP, but how it manages the impressive feat of speeding up GAP is what is truly fascinating, introducing its own set of accuracy-related disadvantages—and

benefits—over the most commonly-used form of GAP.

3.1 The computational advantage of tabGAP

For any given interatomic potential, there exists a wide array of different ways to evaluate the potential contribution of neighbouring atoms. The most commonly used terms include the two- and three-body terms (discussed in more detail in Ref. [30]), which compute the bonds between atom pairs and triplets, respectively (two-body introduces distance-dependency, three-body additionally implements angular dependency), and with machine-learning potentials, the *smooth overlap of atomic positions* (SOAP) [31]. GAP is often used with a two-body and SOAP term. Multi-body terms like SOAP and three-body terms are vectors with multiple dimensions, making them computationally demanding. For instance, depending on the material, the SOAP term can have more than 100 dimensions.

The computation of atomic energies in GAP is very intensive, but tabGAP circumvents this by using low-dimensional terms for their analytical form and fewer degrees of freedom; this way tabGAP can pre-evaluate values for these terms using GAP's slow but accurate methods, creating a feasibly sized grid of data points, in a process called tabulation. These data are then interpolated during simulations [32], usually resulting in a noticeable speed-up in performance. In addition to being faster, few-body terms require less fitting data from quantum-mechanical methods than the popular SOAP term, particularly when developing the potential for multiple elements [32]. Hence, low-dimensional terms are easier to make accurate for multi-element materials like HEAs.

However, two- and three-body terms have innate restrictions, namely they do not account for atomic coordination, that is the number of atoms surrounding a given atom, when evaluating energies of atomic bonds. Coordination influences the energy per atomic bond, as a higher coordination number signifies that the electron cloud of the atom interacts with a greater number of atoms. Terms that account for coordination often have many dimensions, making them incompatible with tabGAP, due to its interpolation-focused technique. Serendipitously, there exists a low-dimensional solution to the problem, and for that, an approach similar to EAM can be employed.

3.2 EAM-augmented tabGAP

As one will learn in this section, the distinctive feature of EAM, the embedding term, is rather suitable for use in tabGAP. The energy in EAM, E_{EAM}, is divided into two terms [33, 34], as follows:

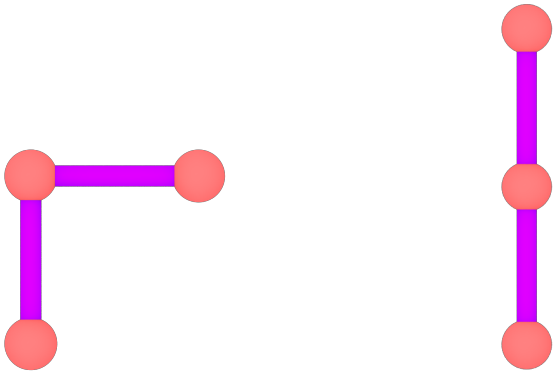
$$E_{\mathrm{EAM}} = E_{2\mathrm{b}} + \sum_j F_j\left(\rho_j\right) , \tag{3.1}$$

in other words, the energy consists of a two-body term, $E_{2\mathrm{b}}$, and the embedding term, $F = \sum_j F_j$. The j-th embedding term is solely a function of the local electron density ρ_j at the position of the j-th atom, on which the embedding energy of the same atom also depends. The local electron density, in turn, is a sum of spherically symmetric functions centred at neighbouring atoms (within the cut-off radius of the potential). The embedding function is unique because it provides coordination-dependence to the energies of atomic bonds, but has a slight drawback: the assumption of near-uniform atom configurations.

The embedding term works well when the atoms surrounding any given atom j are (roughly) uniformly arranged [33]. The root cause of this also manifests itself in EAM's choice of energy terms, shown in (3.1); much like the 2-body term, the embedding energy has no angular dependence (electron density is purely a function of distance), leaving EAM deficient in some instances.

An example of a problematic situation for the embedding function is illustrated in Fig. 3.1: the coordination numbers of the centre atom (the one with two bonds) in Fig.s 3.1a and 3.1b are the same, but the angles between bonds are different. When the embedding function is only paired with the 2-body term, the energies in both cases are the same for the middle atom, while in fact, they should not. The embedding term is unique because it is only 1-dimensional and implementable in tabGAP [35], whilst also accounting for atomic coordination.

Fortunately, the lack of angular dependence of the embedding term is remedied by the inclusion of the 3-body term in tabGAP. In fact, the inclusion of the embedding term, in addition to the usual 2- and 3-body terms, has been shown [35] to complement the strengths and weaknesses of tabGAP, leading to better agreement between tabGAP results and those obtained from density-functional theory (an accurate quantum-mechanical simulation method).



(a) (b)

Figure 3.1: Example of a shortcoming of EAM embedding term: the lack of angular dependence. Red spheres represent atoms; purple lines constitute bonds. The bonds in both cases are of the same length.

4. Modelling of diffusion processes using MD

As excellent of a tool as MD is, the time elapsed in *traditional* MD simulations is limited to nanoseconds, at most microseconds, making it unsuitable for observing rare phenomena. For instance over time, defects move around in metals, and this evolution is of interest when one wishes to understand the long-term effects of irradiation in materials. However, defect diffusion is very slow. As an example, consider vacancy diffusion in W. A rough estimation for the rate of vacancy diffusion (how many times the vacancy changes lattice sites per unit time), denote it by ν, can be obtained from the Arrhenius equation [36]:

$$\nu = \nu_0 \exp\left(-\beta E_{\mathrm{mig}}\right) ,$$

where $\beta \equiv 1/\left(k_{\mathrm{B}} T\right)$ (thermodynamic beta), k_{B} is the Boltzmann constant, E_{mig} is the energy needed for a vacancy to migrate to a new lattice site, and the attempt frequency, ν_0, is roughly equal to the Debye frequency of W (which is related to the frequency of thermal vibrations), hence $\nu_0 \approx 10^{13}\,\mathrm{s}^{-1}$. The experimental value of E_{mig} for W is between 1.7 eV and 2.02 eV [37, 38]. Therefore at room temperature, $T = 300\,\mathrm{K}$, the average time for a vacancy to migrate once is between 10^8 and $8 \cdot 10^{20}$ years; raise the temperature to 1000 K, and the time falls between 40 µs and 2 ms, which is close to MD time-scale, but would seldom be observed. There exist a myriad of methods to accelerate MD, allowing one to peer far beyond the nanosecond limit, the limitations of traditional MD notwithstanding.

4.1 Accelerating rare events

Rare events can be accelerated if one understands the fundamental mechanisms behind them. For instance, for a defect to migrate in solid matter, an atom usually has to cross a region of higher potential energy in order to leave its lattice site (this transition will be referred to as a site-leap hereafter). Doing this the classical way requires an atom to possess enough kinetic energy to momentarily supply the potential

boost necessary to perform the site-leap. Diffusion processes are slow if the barrier is significantly larger than the average atomic kinetic energy.

One way to speed up defect diffusion is to only consider the occurrence probability of the leap from one configuration to another, essentially avoiding computing everything in between the leaps, hence accelerating the simulation. This is the basis for the kinetic Monte Carlo (KMC) technique [39]. When a defect site-leaps, the state of the system, or configuration, changes. KMC-like techniques work only if one is able to predict the occurrence probabilities for all possible transitions that may transpire during the course of a simulation. For instance, the presence of multiple elements in HEAs can make the number of possible elemental configurations so large that manually sampling them in simulations becomes unfeasible. There do, however, exist methods that need nothing but the initial conditions, in addition to few parameters.

4.1.1 Hyperdynamics

Where KMC shortened computation time by only using transition probabilities, hyperdynamics methods effectively lowers the site-leap potential barriers for atoms; this is done by introducing a *bias potential* surface in addition to the potential [40], increasing the total potential energy of the system. By increasing the potential energy, atoms with the lower threshold to cross are the most likely to do so, as often is the case in reality. Lowering the potential threshold accelerates diffusion, decreasing computational load.

The beauty of hyperdynamics lies in the simplicity of its implementation: unlike in KMC, one does not need to know all the possible leap paths, rather the bias potential encourages atoms to pick the correct path [41]. This means that only the initial state needs to be known. The benefit of this method is that it allows the system to evolve as it naturally would (dictated by the choice of potential), without the behaviour being constrained by what is already known (e.g. the paths in KMC). In other words, this makes it possible to discover phenomena that were not known in advance, which are usually of interest when studying new, exotic materials.

One trade-off for not having to know all possible states and leap paths is the following: during a site-leap, the bias potential must be zero. In other words, the bias potential needs to have such a mathematical solution that it can differentiate between states, or at least their dividing boundaries, without knowing explicitly what those states are [40, 41]. Furthermore, these solutions to the bias potential can be exclusive

to a specific subset of materials [3]. This puts most hyperdynamics methods in a situation comparable to that of KMC where, instead of having to come up with leap paths and configurations, one needs to come up with the proper solution to the bias potential which depends on the type of system being studied. Although hyperdynamics can be used to explore the unknown aspects of materials, locked behind the typical time limit of MD, it can be challenging to adapt from one material to another. A solution to this challenge can be found in another method, metadynamics, which also uses the bias potential method, but has a more flexible way to distinguish between states.

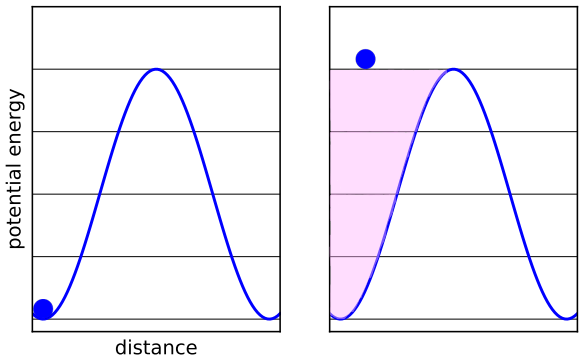
4.1.2 Metadynamics

To distinguish between different states, metadynamics [42, 43] utilises so-called collective variables (CV). CVs are functions of the atomic coordinates that describe the progress along a reaction coordinate or the evolution of a particular physical or chemical process. In metadynamics, they are used to enhance the sampling of the system's potential energy landscape by adding a history-dependent bias potential along the chosen CVs. In contrast to hyperdynamics-based methods, where the bias potential is greatly dependent on the system of interest, metadynamics is more readily adjustable to different systems, since only the CVs need to be changed [3].

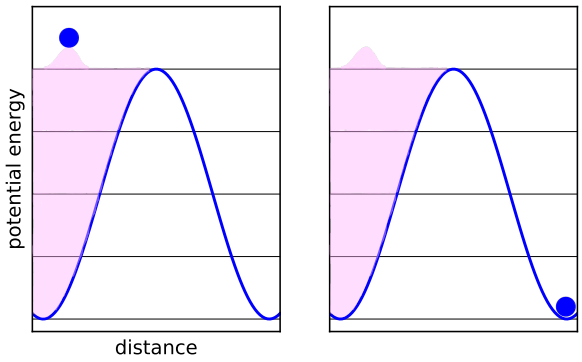
A greatly simplified example of applying the bias potential is illustrated in Fig. 4.1, where an atom sits in a potential well. In this example, the position of the atom could be the CV. The bias potential is added over time in packets whose distribution is Gaussian,* The addition of bias potential promotes state-leaps. This can be seen in Fig.4.1b where on the left-hand-side graph, a bias potential packet centred at the atom's position is introduced, resulting in the atom migrating over the potential barrier to another potential minimum on the right-hand-side graph. The atom could as well have moved left, but would ultimately move right as more bias potential is introduced to the left-hand side. During the transition, no bias potential is introduced. After the site leap, the adding of the bias potential would continue to the new potential minimum. In metadynamics, the bias potential in the old site is left there, which discourages the system from returning to its old states [42].

To better understand the possible limitations of metadynamics, consider some

*The bias potential needs not be applied at every time-step. This way, the atom is allowed to move along this potential surface, bounded only by the potential well (the potential pertaining to the system itself).



(a) left: $t = 0$, right: $t = 1$



(b) left: $t = 2$, right: $t = 3$

Figure 4.1: A simplified example of the bias potential addition. The blue circle represents an atom. The sine-like blue line is the base potential energy of the atom (potential without biasing). The shaded region is the bias potential. t is time, given in arbitrary units. The atom is free to move on the bias-potential (and system's potential) surface in the confines of its kinetic energy.

arbitrary system. Denote the amplitude (the maximum amount of potential energy introduced at once to any point in the system) of the Gaussian packets by A , and the standard deviation of each CV by σ_i. Moreover, let s_i denote the CVs of the system. The total bias potential, V_b , at time-step* $M = k\,\tau$, where k is an integer, and τ defines the number of steps between the depositions of the bias packets, is simply the Gaussian distribution over the CVs [44]:

$$V_\mathrm{b}\left(s_i\left(\mathbf{r}\right),\ M = n\,\tau\right) = A \sum_{k<n} \exp\left\{-\Sigma_i \frac{\left[s_i\left(\mathbf{r},\ M = n\,\tau\right) - s_i\left(k\,\tau\right)\right]^2}{2\,\sigma^2}\right\}, \tag{4.1}$$

*To pre-empt potential misunderstandings borne out of the temporal scale disparity between MD and sampling/accelerative methods discussed in this section, a decision has been made to represent some typically time-associated quantities in the context of time-steps. This approach guarantees an invariant frame of reference, as the count of time-steps remains steadfast between these simulation methods, notwithstanding the disparate timescales they inhabit.

where $s_i(\mathbf{r},\, M)$ are variables and $s_i(k\,\tau)$ are, at step M, considered bygone values of s_i.

The lower limit of the sum over all k in Eq. (4.1) is always the first instance of adding bias potential, which implies that the bias potential is always added on top of the previous bias and is not removed after a state leap. This makes the bias potential time-dependent, which allows metadynamics to encourage state-leaps. Note that the CVs depend on position, $\mathbf{r}$, making the bias potential dependent on $\mathbf{r}$. This will be important in addressing the limitations of metadynamics, later on, because the CVs need to be able to differentiate between different, *known* states, making it possible to only study transitions between these known states, much like in KMC.

The force related to a potential energy can be obtained by evaluating the potential's spatial dependence. For the sake of clarity, let the bias potential comprise only one CV, s_i. Along any given spatial direction, denoted by x, the force is given by:

$$\begin{aligned} F_x(M) &= -\frac{\partial}{\partial x} V_\mathrm{b} \\ &= -\frac{\partial V_\mathrm{b}}{\partial s_i(M)} \frac{\partial s_i(M)}{\partial x}, \end{aligned} \tag{4.2}$$

where in Eq. (4.2), the CV is assumed to have explicit dependence on x and the second line expands the derivative according to the chain rule:

$$\frac{d}{dx} = \frac{d}{du}\frac{du}{dx}$$

The main takeaway from Eq. (4.2) is that the CV must have a dependence on the positions of real atoms, so that when the CV changes, the particles in the system are affected by a change in the force they experience. In accordance with the bias potential's definition in Eq. (4.1), the term $\partial V_\mathrm{b} \,/\, (\partial s_i)$ (in Eq. (4.2)) has the same mathematical form for any CV, s_i, thereby pointing to the second term as the force-driving term; for the bias force to push a system to a new state, the CV(s) must change along the paths through which state-leaps can take place. Otherwise, the term $\partial s_i \,/\, (\partial x)$ will result in the bias force being too weak to drive the system's evolution, or nonexistent. In other words, the value(s) of the CV(s) must differ noticeably between states. Hence, one needs to know in advance the states they wish to study.

To summarise, in metadynamics, the states of a system that one wishes to study must be reflected in the values of the CVs, and hence known in advance. This restricts evolution to occur between known states, unlike hyperdynamics, which allows the system to evolve in a natural, unrestricted manner insofar as the initial state is known.

Metadynamics is simpler to adapt to different systems, because only the CVs need to be changed, as opposed to the functional form of the bias potential of hyperdynamics. For the goals of this thesis, the perfect method is one that combines the ease of adaptation of CVs with the natural evolution that hyperdynamics permits. *collective variable-driven hyperdynamics* is one such method.

4.1.3 Collective variable-driven hyperdynamics

In order to remedy the disadvantages of the techniques discussed above, collective variable -driven hyperdynamics (CVHD) reduces the number of CVs to one, denoted by η, resulting in a bias potential, $V_{\mathrm{b}}(\eta)$, that depends only on the CV, not position. Furthermore, the CV in CVHD is *unable* to discern the possible states that the system can assume, instead being capable of only telling the extent to which the system is in its current state [3], as will be exemplified later. The construction of the CV, for which only the initial state needs to be known, can be broken down to two variables: local distortion, χ_i, whose task is to establish a connection between select local properties and their involvement in state-leaps; global distortion, χ, whose task is to interpret, from the global distortions, how close the entire system is to a state-leap, so as to direct the bias potential to properties that contribute to the leap the most.

Local distortion needs to be able to perceive if a state-leap has occurred, even in unknown states. One such mathematical definition relies on atomic bonds being broken during transitions. This is realised by first building for some set of atoms a list of neighbouring atoms (atoms that lie within a spherical region of radius $r_{\mathrm{cut}}^{\max}$). In the original formulation, two parameters need to be defined: the equilibrium length, that is the minimum-potential distance between bound atoms, r_{ref}, and the maximum bond length, $r_{\max}$, beyond which a bond can be considered to be broken. Then, the degree to which the bond of the ith atomic pair is distorted, that is the local distortion, is given by [3]:

$$\chi_i = \begin{cases} 0\,, & (\text{if } r_i \leq r_{\mathrm{ref}}) \\ \dfrac{r_i - r_{\mathrm{ref}}}{r_{\max} - r_{\mathrm{ref}}}, & (\text{if} r_{\mathrm{ref}} < r_i < r_{\max}) \\ 1\,, & (\text{if } r_i \geq r_{\max}) \end{cases} \tag{4.3}$$

where r_i is the bond length. The global distortion [3] is:

$$\chi = \left(\sum_i \chi_i^p \right)^{1/p},$$

where $p > 1$. The purpose of p is to provide larger distortions more weight in the value of the global distortion. This way, larger distortions are assigned more bias potential

in order to encourage state-leaps.

If one were to use the notation in Eq. (4.3), then the CV would be [3]:

$$\eta = \begin{cases} \frac{1}{2}\,[\,1-\cos(\pi\,\chi^2)\,]\,, & (\text{if } X \le 1) \\ 1\,, & (\text{if } X > 1)\,, \end{cases} \tag{4.4}$$

However, an alternative notation has been used in recent literature [45, 2]. Since local distortion is always non-negative, it is apparent that $\chi \ge \chi_i$ for any χ_i. Instead of defining $r_{\max}$, such that when $r_i \ge r_{\max}$, $\chi_i = 1$ (and subsequently $\chi \ge 1$, leading to $\eta = 1$), one could indicate a broken bond by defining a cut-off value for global distortion, χ_{cut}, such that $\eta = 1$ if $\chi \ge \chi_{\text{cut}}$, i.e.:

$$\eta = \begin{cases} \frac{1}{2}\left\{1-\cos\left[\pi\,\left(\frac{\chi}{\chi_{\text{cut}}}\right)^2\right]\right\}\,, & (\text{if } \chi \le \chi_{\text{cut}}) \\ 1\,, & (\text{if } \chi > \chi_{\text{cut}})\,, \end{cases} \tag{4.5}$$

Now, one only needs to define the equilibrium bond length for local distortion:

$$\chi_i = \begin{cases} 0\,, & (\text{if } r_i < r_{\text{ref}}) \\ \dfrac{r_i - r_{\text{ref}}}{r_{\text{ref}}}\,, & (\text{if } r_i > r_{\text{ref}}) \end{cases} \tag{4.6}$$

To understand why the above-given definitions are equivalent, suppose that a bond breaks at $r_{\max} = a\,r_{\text{ref}}$, where a is a constant. If $r_i \ge r_{\max}$, then $\chi_i \ge (a-1)$, hence $\chi \ge (a-1)$, and one can therefore do away with $r_{\max}$ by defining $\chi_{\text{cut}} = a-1$. This is equivalent to the alternative notations in equations 4.3 and 4.4 because η will be equal to 1 when $r_i \ge r_{\max}$ because then $\chi \ge \chi_{\text{cut}}$, just as η was in the old notation, in spite of the newer notation's χ_{i} not capping off at 1.* In this thesis, the newer notations in equations (4.5) and (4.6) is preferred.

The bias potential at any given time-step, M, is similar to that of metadynamics, shown in Eq. (4.1):

$$V_{\text{b}}\,(\,\eta\,(M = n\,\tau)\,) = A \sum_{k<n} \exp\left\{-\frac{[\eta\,(M = n\,\tau) - \eta\,(k\,\tau)]^2}{2\,\sigma^2}\right\}\,, \tag{4.7}$$

when the bias potential is added at τ time-step intervals. Now, the force acting on any atomic pair i (in the direction of its bond) is given by:

$$F_i\,(M) = \frac{\partial\,V_{\text{b}}}{\partial\eta\,(M)}\,\frac{\partial\,\eta\,(M)}{\partial\chi\,(M)}\,\frac{\partial\,\chi\,(M)}{\partial\chi_i\,(M)}\,\frac{\partial\,\chi_i\,(M)}{\partial r_i} \tag{4.8}$$

*The newer (equivalent) notation simplifies the parameters, because instead of assigning $r_{\max}$ a numerical value, say $1.5\,r_{\text{ref}}$, whose value depends on lattice properties (hence varies between materials), one can assign $1.5-1 = 0.5$ as the cut-off for χ, and have to change only r_{ref} to suit the material of choice.

With the force expanded in Eq. (4.8), the local distortion, in Eq. (4.6), is easier to interpret. If the ith bond's length is $r_i > r_{\text{ref}}$, and thus $\chi_i > 0$, then the right-most term in Eq. (4.8) implies that the bond experiences a force that aims to stretch it. However, when $r_i < r_i^{\text{ref}}$, the right-most term is zero, and the bond is deemed too short to apply a stretching force to.

When $r_i > r_i^{\text{ref}}$ such that $\chi \geq \chi_{\text{cut}}$, and hence $\eta = 1$, it alludes the existence of a broken bond, implying that a state-leap has taken place. Such an event is illustrated in Fig. 4.2. In the aforementioned figure, there exists 1 vacancy in a 2-dimensional lattice. The initial state is in Fig. 4.2a, from which the system transitions to the final state, Fig. 4.2b, as a result of atom number (#) 4 leaping to the vacant site. Consequently, atom #4 ends up outside of atom #0's neighbour shell, indicated by the dashed circle. In the final state, the local distortion pertaining to atoms #0 and #4 becomes great enough that $\chi > \chi_{\text{cut}}$,* terminating all bias forces. This is where CVHD implements a time-step-based reset, which is why V_{b} depends not on position but only on η, for it allows the CV to not need to able to differentiate between states.

After the site leap in Fig. 4.2, η stays equal to 1. If it were left that way, no further evolution would occur. In metadynamics, the evolution would continue if one had chosen CVs that can identify the new state, which implies that one would need to know which states they wanted to study in advance. CVHD gets around this by checking if its CV stays equal to 1 for some number of consecutive time-steps of τ_{reset}, which alludes that a state-leap has taken place, after which the neighbours of atoms are re-evaluated, the CV has its value consequently reset, and, in contrast to metadynamics, the old bias potential is thereby removed [3]. Because the CV is reset, it can take on the same values it did in the previous state, hence eliminating dependence on position. For example, after the reset in Fig. 4.2b, atom #0's neighbours would include atoms 1-3, excluding #4 that was initially a part of the list. This way, atom #0 would have no broken bonds, resulting in its local distortions being less than 1.

Hypertime

Due to the accelerative nature of CVHD, it is useful to define a time-frame which corresponds to in-simulation time that *would* have, on average, elapsed during an event had no accelerative measures like biasing been applied. This accelerated time estimate is known as hypertime, and formulating a link between MD time and this estimate is not trivial.

*Because $r_i > (\chi_{\text{cut}} + 1)\, r_{\text{ref}}$.

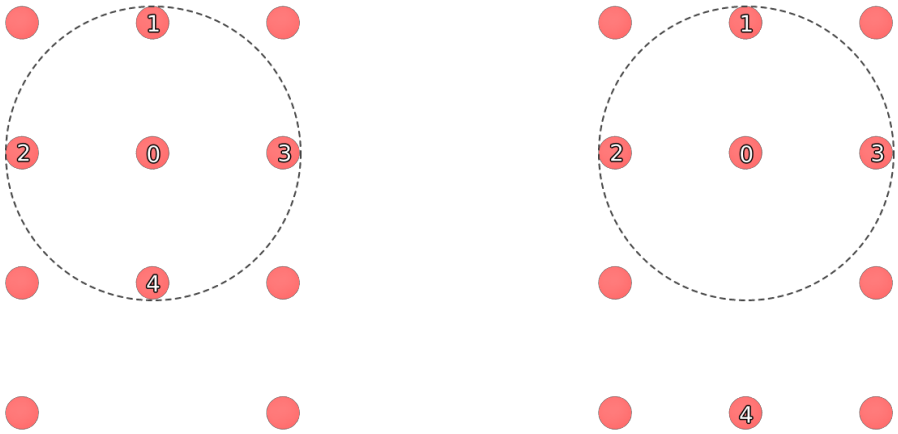


Figure 4.2: Example of vacancy diffusion in CVHD. The red circles are atoms. Fig. 4.2a takes place before Fig. 4.2b. The dashed line (circle of radius $r_{\text{cut}}^{\max}$) is the shell within which atom 0's neighbours lie. Distances between adjacent atoms (atoms that have no vacancy separating them) are very close to r_{ref} , such that the local distortions pertaining to these atoms are very small.

In CVHD, hypertime can be estimated from the amount of introduced bias potential. CVHD [3] uses the same method for evaluating hypertime as hyperdynamics, which is described in Ref. [41], but will be done here in more detail, addressing how it applies to CVHD.

The unbiased system is not observed in CVHD simulations, and therefore a relation between the properties of the biased and unbiased systems must be established, in order to evaluate hypertime. To do so, consider an unbiased system made up of N particles, which is connected to a heat bath, making it a canonical ensemble (N, volume V, and temperature T are constant). Such an NVT ensemble is obtainable over time in MD by applying a good thermostatting algorithm.

To describe the state of a system, one can combine the positions and momenta of all particles into single vectors, $\mathbf{r}$ and $\mathbf{p}$ respectively, making them $3N$-dimensional vectors, so-called generalised coordinates. Different states (configurations) can then be identified as $3N$-dimensional regions in the $\mathbf{r}$ space, with the boundaries between them having $3N-1$ dimensions. * For some unbiased state α , the escape frequency

*For example, consider the two-particle system of a single hydrogen molecule, H_2. If the distance between the hydrogen atoms exceeds some value $r_{\max}$, then the molecular bond breaks. By fixing the origin to one of the atoms, the generalised coordinate $\mathbf{r}$ would simply be the position of the other atom. This way, the unbound state occupies the volume of a sphere of radius $r_{\max}$ in the 3-dimensional $\mathbf{r}$ space, and the unbound state occupies the remaining space, where the interatomic distance is greater

(how often an escape is attempted) from this state is given by [3, 46]:

$$k_{\alpha\rightarrow} = \langle \mid v_\alpha \mid \delta_\alpha(\mathbf{r}) \rangle_\alpha , \tag{4.9}$$

where $\langle\rangle_\alpha$ denotes the average taken within the phase space, $(\mathbf{r}, \mathbf{p})$, of state α, v_α is the velocity (in terms of $\mathbf{r}$) perpendicular to the dividing surface of α and δ_α is the Dirac delta-function, centred at the boundary of α. By defining a step function, $\Theta_\alpha(\mathbf{r})$, whose value is 1 inside α and 0 outside, Eq. (4.9) can be written as:

$$k_{\alpha\rightarrow} = \langle \mid v_\alpha \mid \delta_\alpha(\mathbf{r}) \, \Theta_\alpha(\mathbf{r}) \rangle$$

In the canonical ensemble, the ensemble average of some quantity S is given by:

$$\langle S \rangle = \frac{1}{Z_N} \int S \, e^{-\beta H} \, \mathrm{d}\Gamma , \tag{4.10}$$

where $\beta = 1/(k_\mathrm{B} T)$, $Z_N \equiv \int \exp(-\beta H) \, \mathrm{d}\Gamma$ is the canonical partition function, Γ lies in the $6N$-dimensional phase space $(\mathbf{r}, \mathbf{p})$, and the Hamiltonian H is the sum of the kinetic energy $K(\mathbf{p})$ and the (unbiased) potential energy $V(\mathbf{r})$. Eq. (4.10) may now be written as:

$$k_{\alpha\rightarrow} = \frac{\int [\mid v_\alpha \mid \delta_\alpha(\mathbf{r}) \, \Theta_\alpha(\mathbf{r})] \, e^{-\beta[K(\mathbf{p})+V(\mathbf{r})]} \, \mathrm{d}\Gamma}{Z_N}$$

Note that Γ spans the phase space containing all possible states, but the step function limits the contributions of the integrand to the space of α. Next, one can introduce to the numerator the bias potential, by noting that $\exp(\beta V_\mathrm{b}) \exp(-\beta V_\mathrm{b}) = 1$, thereby:

$$k_{\alpha\rightarrow} = \frac{\int \left[\mid v_\alpha \mid \delta_\alpha(\mathbf{r}) \, \Theta_\alpha(\mathbf{r}) \, e^{\beta V_\mathrm{b}(\eta)} \right] e^{-\beta[K(\mathbf{p})+V(\mathbf{r})+V_\mathrm{b}(\eta)]} \, \mathrm{d}\Gamma}{Z_N} , \tag{4.11}$$

where the Hamiltonian of the system whose potential has been boosted by the bias potential, call it the boosted Hamiltonian, can be identified as: $H_\mathrm{b} \equiv K(\mathbf{p}) + V(\mathbf{r}) + V_\mathrm{b}(\eta)$. As previously, by multiplying by 1, Eq. (4.11) can be written as:

$$\begin{aligned} k_{\alpha\rightarrow} &= \frac{\int e^{-\beta H_\mathrm{b}} \, \mathrm{d}\Gamma}{\int e^{-\beta H_\mathrm{b}} \, \mathrm{d}\Gamma} \, k_{\alpha\rightarrow} \\ &= \frac{\int \left[\mid v_\alpha \mid \delta_\alpha(\mathbf{r}) \, \Theta_\alpha(\mathbf{r}) \, e^{\beta V_\mathrm{b}(\eta)} \right] e^{-\beta H_\mathrm{b}} \, d\Gamma}{\int e^{-\beta H_\mathrm{b}} \, \mathrm{d}\Gamma} \, \frac{\int e^{-\beta H_\mathrm{b}} \mathrm{d}\Gamma}{Z_N} \end{aligned} \tag{4.12}$$

Let the state α with the boosted Hamiltonian be denoted by α_b, and denote the partition function of the boosted system by $Z_{b,N} \equiv \int \exp(-\beta H_\mathrm{b}) \, \mathrm{d}\Gamma$. Now, the terms in Eq. (4.12) can be written as expectation values in the boosted system. The left-hand term becomes:

$$\frac{\int \left[\mid v_\alpha \mid \delta_\alpha(\mathbf{r}) \, \Theta_\alpha(\mathbf{r}) \, e^{\beta V_\mathrm{b}(\eta)} \right] e^{-\beta H_\mathrm{b}} \, \mathrm{d}\Gamma}{Z_{b,N}} \equiv \langle \mid v_\alpha \mid \delta_\alpha(\mathbf{r}) \, e^{\beta V_\mathrm{b}(\eta)} \rangle_{\alpha_\mathrm{b}} \tag{4.13}$$

than r_max

and the right-hand term, after introducing $H = H_b - V_\mathrm{b}(\eta)$ the same way as in Eq. (4.11):

$$\begin{aligned}\frac{Z_{\mathrm{b},N}}{Z_N} &\equiv \left(\frac{\int e^{-\beta H}\mathrm{d}\Gamma}{Z_{\mathrm{b},N}}\right)^{-1} \\ &= \left(\frac{\int e^{\beta V_\mathrm{b}(\eta)}\, e^{-\beta H_\mathrm{b}}\mathrm{d}\Gamma}{Z_{\mathrm{b},N}}\right)^{-1} \\ &\equiv \frac{1}{\langle\, e^{\beta V_\mathrm{b}}\,\rangle_{\alpha_\mathrm{b}}}\end{aligned}$$

The term in Eq. (4.13) can be further simplified by noting that $\delta_\alpha = 0$ everywhere but at the dividing boundary of α; in the case of CVHD, the definition of the CV in Eq. (4.5) guarantees that the term $\partial\eta\,/\,(\partial X)$ vanishes at the dividing boundary, $X = X_\mathrm{cut}$, thereby causing the bias force to be zero at the boundary. Consequently, the bias potential should be very flat (no spatial dependence) at the boundary and thus likely close to zero,* hence giving the escape frequency the final form:

$$\begin{aligned}k_{\alpha\rightarrow} &= \langle\, |\, v_\alpha\,|\, \delta_\alpha(\mathbf{r})\, e^{\beta V_\mathrm{b}(\eta)}\,\rangle_{\alpha_\mathrm{b}}\, \frac{1}{\langle\, e^{\beta V_\mathrm{b}}\,\rangle_{\alpha_\mathrm{b}}} \\ &= \frac{\langle\, |\, v_\alpha\,|\, \delta_\alpha(\mathbf{r})\, e^{0}\,\rangle_{\alpha_\mathrm{b}}}{\langle\, e^{\beta V_\mathrm{b}}\,\rangle_{\alpha_\mathrm{b}}} \\ &= \frac{\langle\, |\, v_\alpha\,|\, \delta_\alpha(\mathbf{r})\,\rangle_{\alpha_\mathrm{b}}}{\langle\, e^{\beta V_\mathrm{b}}\,\rangle_{\alpha_\mathrm{b}}}\end{aligned} \tag{4.14}$$

The above-given equation establishes that the escape frequency in the unbiased state α can be expressed in terms of ensemble averages of quantities in the biased state α_b. This is of significance because in CVHD simulations, the time that one has access to belongs to the biased state. In other words, Eq. (4.14) gives access to the "real" time that cannot be directly observed in the biased system due to it being accelerated.

After a state-leap, i, has transpired, the hypertime, or the total time it would have on average taken for the leap to occur (leap-time) without the bias potential, can

*The Gaussian distribution is flat far from its peak (where it is almost zero), but also at the peak. Unless the bias potential is added precisely at the dividing boundary, the former case has to be satisfied, and the bias potential should be close to zero at the boundary.

be obtained as follows:

$$\begin{aligned} t_{\mathrm{h},\,i} &= \frac{1}{k_{\alpha\rightarrow}} \\ &= \frac{\langle\, e^{\beta\, V_\mathrm{b}} \,\rangle_{\alpha_\mathrm{b}}}{\langle\, |\, v_\alpha \,|\, \delta_\alpha\,(\mathbf{r}\,) \,\rangle_{\alpha_\mathrm{b}}} \\ &\equiv \frac{\langle\, e^{\beta\, V_\mathrm{b}} \,\rangle_{\alpha_\mathrm{b}}}{k_{\alpha_\mathrm{b}\rightarrow}}, \end{aligned} \tag{4.15}$$

where $k_{\alpha_\mathrm{b}\rightarrow}$ is the escape frequency in the biased state. Since the leap-time, t_leap, inside the biased system is merely the elapsed MD time, the aforementioned frequency is simply $k_{\alpha_\mathrm{b}\rightarrow} = t_\mathrm{leap}^{-1}$. Bearing in mind that time in MD is quantised, the average in the numerator of Eq. (4.15) is taken over all n MD time-steps, $\Delta\, t_\mathrm{MD}$, leading up to $t_\mathrm{leap} = n\,\Delta\, t_\mathrm{MD}$:

$$\langle\, e^{\beta\, V_\mathrm{b}} \,\rangle_{\alpha_\mathrm{b}} = \sum_k^n \frac{e^{\beta\, V_\mathrm{b}[\,\eta(k\,\Delta\, t_\mathrm{MD})\,]}}{n}$$

Hence, the hypertime for one leap is:

$$t_{\mathrm{h},\,i} = \sum_k^n \Delta\, t_\mathrm{MD}\, e^{\beta\, V_\mathrm{b}[\,\eta(k\,\Delta\, t_\mathrm{MD})\,]} \tag{4.16}$$

Note that k marks the time-step when the system first enters state α, and n when it leaves it. If multiple leaps occur, the total hypertime is obtained by substituting $k = 0$ and n with the final step of the simulation. The evolution of hypertime in a simulation is illustrated in Fig. 4.3.

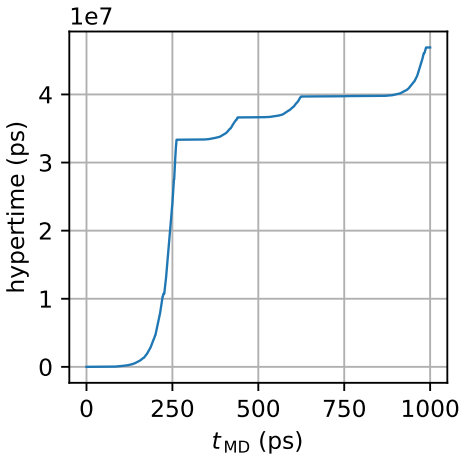


Figure 4.3: Hypertime in a vanadium-tungsten cell at 1000 K. MD time, t_MD, is the time elapsed in the biased system, hypertime is the equivalent accelerated time. The abrupt change in hypertime are caused by a vacancy site-leaping, causing the state of the system to change, after which bias potential is removed, and the acceleration momentarily becomes zero.

Ensuring the physicality of CVHD through well-tempered metadynamics

The freedom of not having to know all the possible states a system can evolve to introduces its own set of perils. One such pitfall is over-biasing. In cases where the CV is sluggish to change, the applied bias potential can easily exceed the site-leap energy barrier; this can lead to unrealistic dynamics, and can inflate the hypertime estimation, which leads to incorrect measurements. The latter happens because more bias is added than necessary, and on the account of Eq. (4.16), the hypertime is over-estimated. A measure for counteracting over-biasing can be formulated if one knows at least roughly the magnitude of the energy barriers that can be encountered in the simulations. When considering a canonical ensemble, which an NVT ensemble at equilibrium for a large time in MD approximately is, one can formulate from the deposition rate an analytical solution for the value of bias potential at any given value of the CV at $t \to \infty$. Well-tempered metadynamics [47] (WTMetaD) is a method that utilises such an approach, and can be utilised in CVHD to allow one to estimate the limit of this bias potential, i.e., the maximum value it can assume at any given CV value.

In WTMetaD, the height of the deposited Gaussian packets, A, is not treated as a constant. Rather, the amplitude of deposited bias packets is attenuated as the total deposited value at some CV approaches a threshold value. The parameter that specifies the magnitude of the aforementioned threshold, called the biasing factor, γ, is conveniently defined as [48]:

$$\gamma = \frac{T + \Delta T}{T},$$

where T is the physical temperature of the system, and ΔT is the so-called bias temperature, which is *generally* defined such that $k_\mathrm{B}\,(T + \Delta T)$ is in the order of the threshold one wishes to limit the bias energy to. Now, one can write the height of the Gaussian packet deposited at time-step $M = k\,\tau$ as:

$$A\,(k\,\tau) = A_0\,\exp\left(-\frac{V_\mathrm{b}\,(\,\eta\,(M = k\,\tau)\,)}{(\gamma - 1)\;k_\mathrm{B}\,T}\right), \tag{4.17}$$

where $V_\mathrm{b}\,(\,\eta\,(M = k\,\tau)\,)$ is the *total* deposited bias potential at time-step $k\,\tau$, and A_0 is the largest possible height a Gaussian packet can assume.

WTMetaD can be applied to CVHD, as well. Then, the total bias energy is obtained by inserting the $A\,(k\,\tau)$ of Eq. (4.17) in place of the A in Eq. (4.7), such that $A\,(k\,\tau)$ is included in the sum over all $k < n$.

5. Methods

5.1 Software

The LAMMPS code [49] was used for all MD simulations. Additionally, for the CVHD simulations, the PLUMED2 code [50], which is a modification of the PLUMED code [51], was utilised. The KMC simulations were run using the Kimocs code [52], specifically my custom branch thereof, called **D**iffusion **I**nference - **S**oftware **W**ith **O**ptimised **R**igid-lattice **D**ynamics (Di-Sword), which is tailored for my specific set-up of machine learning and descriptors, and the material of interest. As the interaction model in all MD simulations, the tabGAP developed in Ref. [53] was adopted. Moreover, the time integration in the MD simulations was done using the velocity Verlet algorithm [54]. The OVITO code [55] was used for data visualisation. Figures were generated using the Matplotlib [56] Python library. The Python library Keras [57] was used for its feed-forward neural network implementation. The Python library, DScribe [58], was employed for its implementation of the atom-centred symmetry function descriptor necessary for my machine-learning predictions. For GPR [59, 60], the Python library scikit-learn was used [61].

5.2 Collective variable-driven hyperdynamics

In the research underpinning this thesis, I have employed the advanced technique of CVHD [3] with the intention of providing an accurate representation of the intricate process of vacancy diffusion in a BCC HEA system. The principal focus of my efforts was on the quinary system Mo-Nb-Ta-V-W. From this, I derived a total of 95 unique compositions for in-depth scrutiny. My exploration considered a broad range of potential subsystems, including unary, binary, ternary, and quaternary combinations, in addition to the quinary system. The compositions selected for scrutiny were conscientiously chosen, aided by GPR [59, 60]. This facilitated a preliminary evaluation of correlations between diffusivity and composition. The uncertainty in compositional prediction, ascertained through GPR, provided direction as to which compositions re-

quired analysis. Beyond the study of five unary lattices, I generated five distinctively seeded random cells for each non-unary composition that was under investigation. This resulted in a structure comprising of $5 \times 5 \times 5$ unit cells, with each cell incorporating 250 atoms. Subsequently, a vacancy was introduced at a random position within each cell. The small size of the cells helped us counteract the probability of encountering unwanted vacancy formation, elaborated in Sec. 5.4.1, and to enhance the computational efficiency of the CVHD simulations.

5.2.1 CVHD deployment

The execution of CVHD necessitated the amalgamation of the bond-breaking CV [3]. To safeguard against the assumption of unrealistic dynamics by the system due to over-biasing, I adopted well-tempered metadynamics [47], which proved instrumental in circumventing the challenges related to over-biasing while accurately modelling my HEA system.

5.2.2 Implementation of the virtual atom

A crucial step in my simulation procedure involved monitoring the vacancy's location. Since a vacancy is essentially defined as an absence, and given the thermal noise inherent in my system's lattice framework, the task of pinpointing a vacancy proves challenging. This problem was addressed by employing the method proposed by Ebina et al. [2], which entails inserting a virtual atom at the site of the vacancy, the bonds of which are elongated relative to the first-nearest neighbours (1NNs).

In particular, I leveraged the atoms' altered coordination number, denoted as c_i, thereby permitting non-integer values for each atom i [2]:

$$c_i = \sum_{j \neq i} S_{ij} \tag{5.1}$$

where

$$S_{ij} = \begin{cases} \dfrac{1 - \left(\frac{r_{ij} - d_0}{r_0}\right)^6}{1 - \left(\frac{r_{ij} - d_0}{r_0}\right)^{12}}, & \text{for } r_{ij} > d_0 \\ 1, & \text{for } r_{ij} \leq d_0 \end{cases}$$

where r_{ij} is the separation between particles i and j, d_0 signifies the maximum distance for a full coordination contribution, and r_0 delineates how swiftly this contribution decays for distances exceeding d_0. Analogous to Ebina et al., I utilised this coordination number to detect atoms with inadequate coordination, indicative of a nearby vacancy.

These coordination numbers were transformed into a range from 0 to 1 using a switching function, denoted as f_i [2]:

$$f_i = \begin{cases} \dfrac{1 - \left(\frac{c_i - D_0}{R_0}\right)^6}{1 - \left(\frac{c_i - D_0}{R_0}\right)^{12}}, & \text{for } c_i > D_0 \\ 1\,, & \text{for } c_i \leq D_0 \end{cases} \tag{5.2}$$

where D_0 and R_0 are the threshold and tolerance, respectively, for an atom to be deemed under-coordinated. The switching function for various sets of D_0 and R_0 are shown in Fig. 5.1.

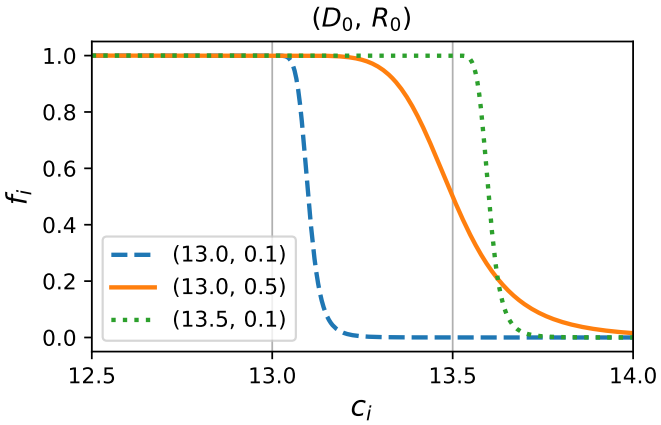


Figure 5.1: Illustration of how D_0 and R_0 affect the switching function's outcome.

Finally, the virtual atom's position is determined as the weighted, geometric centre of the under-coordinated atoms. Denote the component of the position of atom i as $x_{i,\alpha}$, where $\alpha \in \{x, y, z\}$. The components of the position of the virtual atom, g_α, are [2]:

$$g_\alpha = \frac{1}{2\,\pi} \arctan\left(\frac{\sum_i f_i \sin\left(2\,\pi\, x_{i,\alpha}\right)}{\sum_i f_i \cos\left(2\,\pi\, x_{i,\alpha}\right)}\right) \tag{5.3}$$

This methodology facilitated the tracking of vacancies in a convenient manner, and enabled the modelling of the bond-breaking CV, denoted as η, as a function of both local and global distortions, χ_i and χ, respectively.

5.2.3 Selection of parameters for CVHD

During the execution of CVHD in my system, I encountered several inadequacies in the conventional selection of parameters, which were insufficient to address the

intricacies presented by my multi-elemental, high-temperature lattice systems. My research, until the present time, remains one of the most intricate systems for the application of CVHD, comprising a multitude of diverse elements, with a significant portion presenting high transition barriers for vacancy movements in unary forms. Consequently, numerous systems exhibited remarkable resistance against alterations in response to the biasing potential, necessitating an elevated temperature of 800 K to facilitate noticeable vacancy movements.

A unique difficulty lay in the system's marked unwillingness to evolve under biasing, thereby heightening the risk of excessive biasing, potentially disrupting the physicality of the simulations. This circumstance demanded a significant divergence from the traditionally recommended biasing factor, γ [48]. Further details can be found in Sec. 6.1.

In addition, my parameters R_0 and D_0 required significant revisions from the selections made by Ebina and colleagues in their exploration of Fe-Cr BCC structures. R_0 experienced a change of an order of magnitude to accommodate the complexity of my system, whereas D_0 underwent a slight yet vital adjustment.

These intentional deviations from the conventional practices were necessitated by the unique composition of my studied lattices, the high temperature required for noticeable vacancy movements, and the systems' inherent resistance to adapt to biasing. The specifics of these parameter choices are given in Table 5.1.

For an in-depth discussion of the challenges and my devised solutions during the implementation of CVHD on this intricate system, the reader is referred to Sec. 6.1.

5.2.4 Configuration for CVHD-based MD simulations

Non-monoelemental cells were predominantly subjected to simulations lasting 2.5 ns, although certain outliers existed. For instance, monoelemental and non-monoelemental cells exhibiting lower diffusivity were simulated for an extended 6 ns period, aiming to document more than a singular event.

The CVHD MD simulations operated with a time-step of 1 fs. Before the commencement of these simulations, the cells were permitted to equilibrate at a temperature of 800 K under an NPT ensemble deploying Nosé-Hoover thermostatic

Table 5.1: The CVHD parameters used in this study. $a_{\mathrm{E}} = 2.87\,\text{Å}$ is the lattice constant of the α-Fe in Ref. [2]. Also, $d_{0,\mathrm{E}} = 3.2\,\text{Å}$ and $r_{0,\mathrm{E}} = 0.2\,\text{Å}$ are the d_0 and r_0 that Ebina and colleagues used in Ref. [2]. For reference, the numeric values for the scaling factors are $\frac{d_{0,\mathrm{E}}}{a_{\mathrm{E}}} \approx 1.115$ and $\frac{r_{0,\mathrm{E}}}{a_{\mathrm{E}}} \approx 0.070$. a is the material-specific lattice constant for each studied lattice, determined for the thermally equilibrated lattice at the temperature of interest before the CVHD simulation. $R_{1\mathrm{NN}} = \frac{\sqrt{3}}{2}a$ is the distance between the first nearest neighbours (1NN) in the lattice, and $R_{2\mathrm{NN}} = a$ is the distance between the second nearest neighbours (2NN) in the lattice. τ is the pace, or the number of time-steps between the additions of bias packets, and τ_{reset} is the number of time-steps that the collective variable needs to remain equal to 1, before it is reset (diffusion event is registered). Empty brackets after a parameter denote a dimensionless quantity.

Parameter	Value
d_0 [Å]	$\frac{d_{0,\mathrm{E}}}{a_{\mathrm{E}}}\ a$
r_0 [Å]	$\frac{r_{0,\mathrm{E}}}{a_{\mathrm{E}}}\ a$
D_0 [ ]	13.0
R_0 [ ]	0.01
r_{ref} [Å]	$R_{1\mathrm{NN}}$
p [ ]	8
χ_{cut} [ ]	0.6
A_0 [eV]	0.01
σ [ ]	0.01
γ [ ]	6
$r_{\mathrm{cut}}^{\mathrm{max}}$ [Å]	$\frac{R_{1\mathrm{NN}}+R_{2\mathrm{NN}}}{2}$
τ [ ]	500
τ_{reset} [ ]	2000

and barostatic control [62, 63]. In all the diffusivity simulations, the cells were maintained at a steady temperature of 800 K through the application of a Nosé-Hoover thermostat. This configuration essentially fostered the creation of environments closely mirroring canonical ensembles—a requirement for the specific variant of accelerated

time utilised, namely, hypertime [41].

5.3 Gaussian process regression for qualitative diffusivity predictions

I employed a GPR model, utilising a White kernel to account for noise in the data and a radial basis function (RBF) kernel to ensure smoothness in the data. The respective hyperparameters for these kernels are 'noise level' and 'length scale'. To optimise the hyperparameters, I computed the average root-mean-square error (RMSE) over 300 distinct 5-fold cross-validation runs for various pairs of hyperparameters, each initialised with a unique random seed. Additionally, I calculated the mean correct fraction, defined as the proportion of predictions that lie within the 95% credible interval of the true value, i.e., $\pm 1.96\,\sigma$, where σ denotes the standard deviation or uncertainty returned by the GPR for each prediction. The optimal hyperparameters were identified as those that minimised the mean RMSE and yielded a mean correct fraction close to 95%.

To see how diffusivity would differ locally inside an ordered HEA, I also generated a 39 366-atom ordered, equimolar Mo-Nb-Ta-V-W cell at 300 K using MCMD, on which GPR would be used to predict these diffusivities. The MCMD simulation ran for roughly 600 000 time-steps in total, where every 10 time-steps, 60 swaps between each unique pair of distinct elements (600 swaps in total) were attempted.

5.3.1 Evaluation of diffusivity

For every lattice composition under investigation, I computed the mean-square-displacements at all hypertime t values of the virtual atom symbolising the vacancy using*:

$$\langle \mathbf{r}^2 \rangle(t) = |\mathbf{r}(t) - \mathbf{r}(0)|^2 \tag{5.4}$$

where $\mathbf{r}(t)$ represents the position of the virtual atom at hypertime t, $\mathbf{r}(0)$ is its position at hypertime equals zero.

Should spontaneous vacancy creation transpire within a simulation, as elaborated in Sec. 5.4.1, the pertinent simulation was excluded from my analysis.

*Usually, mean-square-displacement is computed over multiple particles. In this case, it is the mean for a sample size (number of particles) of 1.

For each simulation, a linear fit was executed as:

$$\langle \mathbf{r}^2 \rangle(t) = a\,t + b \tag{5.5}$$

In which, setting $b = 0$ and $a = \mathrm{d}/\mathrm{d}t\ \langle \mathbf{r}^2 \rangle$. According to the Einstein relation of Browinan motion, the average diffusivity, D, of N particles in three dimensions is (see Eq. (13.3.34) in Ref. [64]):

$$D = \frac{1}{6} \lim_{t\to\infty} \frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{N} \sum_{i=1}^{N} \langle |\mathbf{r}_i(t) - \mathbf{r}_i(0)|^2 \rangle \tag{5.6}$$

Here, the limit inferior is used to capture the long-term behaviour of the system, ensuring that transient, short-term fluctuations do not dominate the measure of diffusivity. The hypertime in the accelerated MD simulations is likely long enough to satisfy this limit. In my simulations, only a single vacancy is present, so $N = 1$. Therefore, the diffusivity of the vacancy, D_v, is:

$$\begin{aligned} D_\mathrm{v} &= \frac{1}{6} \frac{\mathrm{d}}{\mathrm{d}t} \langle \mathbf{r}^2 \rangle \\ &\equiv \frac{a}{6} \end{aligned} \tag{5.7}$$

Given that multiple simulations were conducted for each composition, I averaged the diffusion coefficients derived from each simulation. This average, in conjunction with the composition vector (a quintet vector detailing the fractions of each chemical element), served as the input for the GPR.

5.3.2 Rationale for GPR selection

The decision to employ GPR for the diffusivity-composition fitting stemmed from the nature of these simulations, which produce a single datum per composition. Notably, GPR exhibits superior performance on limited training sets compared to neural networks [65]. The outcomes derived from the GPR predictions are delineated in Sec. 6.3.

5.3.3 Transition to kinetic Monte Carlo simulations

For a more nuanced quantification of vacancy diffusion, the elementary character of the composition descriptor proves inadequate for parameterising extended-timescale methodologies, specifically kinetic Monte Carlo simulations. The ensuing sections delve deeper into this aspect.

5.4 Feed-forward neural network modelling

This section outlines the training of a Feed-Forward Neural Network (FNN) model utilising my CVHD results. This model primarily seeks to ascertain whether CVHD

is capable of producing valid outcomes for diffusion dynamics in intricate systems, notwithstanding the inherent obstacles. The affirmation of my FNN model's robustness is further expounded in Sec. 5.5.

The selection of the FNN model was informed by its robustness, computational efficiency, and the ability to manage high-dimensional data whilst learning sophisticated non-linear relationships amidst noise [66, 67]. Of particular note is the FNN model's capacity for swift processing of large datasets, made possible by its parallel computing capabilities.

Selection of descriptor

In order to facilitate the training of an FNN model employing my CVHD results, it became necessary to establish a connection between the atomic environments of jumping atoms and the corresponding observed leap-(hyper)time.

In the pursuit of a quantifiable representation of atomic environments, my approach favoured the application of the atom-centred symmetry function (ACSF) descriptor [68]. The choice of ACSF as my preferred descriptor stemmed from its capacity to encapsulate both radial and angular properties through the adoption of two- and three-body terms. This allowed for an exhaustive representation of the environment encircling a prospective leap-candidate.

At the inception of this endeavour, consideration was given to the deployment of the SOAP descriptor [31], which likewise encompasses two- and three-body terms. Nevertheless, the employment of SOAP descriptors turned out to demand a greater computational effort and did not deliver a noticeable advantage over ACSF. Even when settling for a modest selection of two radial basis and two spherical harmonics functions—regarded generally as insufficient even for mono-elemental systems as delineated in Fig. 22 of Ref. [69]—SOAP engendered a substantially more extensive feature set than ACSF. This effect can be attributed to the manner in which the feature count of SOAP escalates with the count of distinct chemical elements. In my case, a single two-body and three-body term produced merely 25 features for ACSF, an order of magnitude fewer than the aforementioned rudimentary SOAP.

Contrary to initial suppositions, the robust feature set proffered by SOAP did not result in a better performing machine learning model. Indeed, the performance of this model was significantly inferior to that of the ACSF model. The complexity

of SOAP, brought about by the multitude of elements in my lattice, emerged as the chief cause of this divergence. It merits attention that this complexity, contingent on the chemical elements, is the very raison d'être of tabGAP, the potential I deployed, which is a GAP potential that circumvents SOAP for lower-dimensional terms.

Consequently, my selection favoured ACSF over SOAP owing to its lesser computational demand and diminished complexity, which collectively resulted in a more efficacious model.

5.4.1 Data analysis and challenges in CVHD simulations

In certain instances of CVHD simulation, I noted an unanticipated advent of surplus vacancies, identified by a rapid rise in the number of under-coordinated atoms (NUCA), which accumulates all f_i (refer to Eq. (5.2) for more detail), prompting the positioning of the virtual atom between two vacancies and frequently distant from a vacant lattice site, as revealed in Fig. 5.2 for one of my initial test runs of CVHD in a Mo-Nb-Ta-V-W cell, larger than the 250-atom cells I ultimately chose to study with CVHD.

One can see in Fig. 5.2a, the 1NN under-coordinated atoms form a single shell, inside of which the virtual particle resides. At a later instance, two such shells, which alludes to there being two vacancies, as can be seen in Fig. 5.2b, and the virtual atom now resides between these shells. This incited an incorrect detection

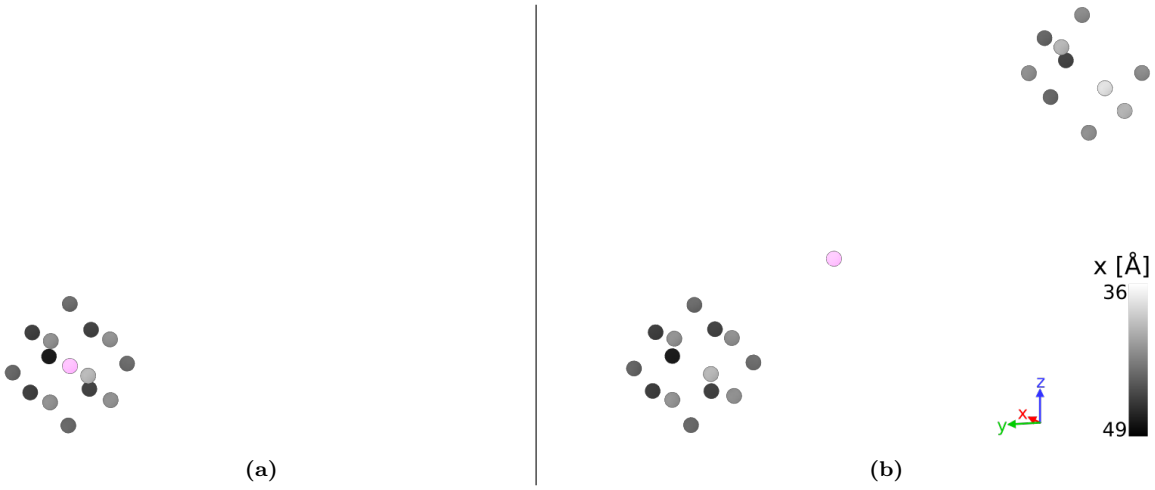


Figure 5.2: Under-coordinated atoms (greyscale) and virtual particle (pink) in Mo-Nb-Ta-V-W. Fig. 5.2a takes place before Fig. 5.2b. The greyscale mapping is used to convey a sense of depth to the under-coordinated particles, and it indicates the x-coordinate of these particles. The coordination analysis is performed within the 2NN shell.

of a site-leap until the end of the simulation or the recombination of the extra

defects. This phenomenon was particularly pronounced in lattices characterised by sluggish vacancy diffusivity, propelling me to speculate that over-biasing might be the root cause in the rare instances where this event transpired. For example, if a bias hill enlarges excessively, as illustrated in Fig. 6.3, a particle sliding down this hill could feasibly amass enough energy to produce a defect, considering the high thermal energy and bias-induced distortion existent in the lattice. In my simulations, I observed instances where the number of under-coordinated atoms (NUCA) exceeded 16. I selected this value as a conservative threshold based on empirical observations; values slightly above 16, such as 17, might have been similarly effective. Exceeding this threshold led me to question the reliability of the particular simulation frame. As a precautionary measure, if the NUCA remained above 16 for two consecutive frames—equivalent to 1000 time-steps—I discarded the subsequent frames in the related simulation. Conversely, I encountered situations where some systems' NUCA values consistently remained under 10, despite the virtual atom continuing to trail the vacancy as expected. Hence, I empirically determined a lower bound for NUCA at 5.

Next, I pinpointed the time-frames in which diffusion events were occurring. These time-frames were extracted from the LAMMPS dump file, each 500 time-steps apart. The atom with the greatest displacement between the event time-frame and the one before it was identified as the leaping atom.

After identifying the leaping atom, my subsequent objective was to depict its environmental context before the site-leap, establishing a correlation between this context and the observed leap-time. This depiction was vital for comprehending the influences on the initial transition barrier. I targeted frames prior to the event, in which the system's global distortion (GD) χ was maintained within the 0.2 threshold, henceforth designated as *acceptable frames*.

The GD threshold was necessary to maintain the accuracy of my analysis. In the context of CVHD, the GD is an indicator of how close the system is to a state transition or site-leap. An extreme rise in GD suggests that the leaping atom is transitioning between lattice sites. The environment of the atom during this state may not be a reliable indicator, as the atom is not tied to a specific lattice site.

Therefore, by confining my analysis to frames with a GD of 0.2 or lower, my aim is to accurately portray the environment prior to the diffusion event, thereby avoiding possible inaccuracies due to observing the transition too closely.

The 0.2 GD threshold represents a considered compromise between achieving an accurate environmental depiction and retaining sufficient training data for model development. A lower threshold may yield a more accurate atomic environment representation, but runs the risk of excluding too many events from the training data, undermining the accuracy of the FNN model.

To strike a balance between computational expenditure and descriptor precision, I restricted my environmental consideration up to the third nearest neighbours (3NNs) within the acceptable frame window. This constraint lessened the computational load during the correlation process and facilitated ACSF evaluations in the following KMC simulations.

If more than 15 acceptable frames preceded an event, I analysed only the first 15. For less than 15 frames, I analysed all. This method took into account that escalating bias energy gradually distorts the lattice, making initial frames more favourable. However, events with no frames demonstrating a GD of 0.2 or lower were omitted, indicating swift transitions.

The selected ACSFs were accumulated, with features calculated as the mean across the individual ACSFs. This tactic aimed to create an ACSF representing atoms near their equilibrium site, enhancing the relevance of the FNN model for rigid-lattice KMC simulations, especially considering the influence of thermal noise at 800 K.

The limitation of 15 frames balanced computational feasibility with the diminishing returns from evaluating additional frames, particularly as multiple ACSF parameters required cross-validation. This compromise aided model development by maintaining data sufficiency while limiting computational cost.

Following this process, I was left with 16 228 leap-times and corresponding ACSFs. Acknowledging the inherent randomness and noise in diffusion data, I employed K-means clustering [70] to group ACSFs based on their similarity, inherently causing the corresponding leap-times to be grouped as well. The optimal cluster number was chosen based on the silhouette score [71], with 1860 clusters yielding the highest score. Each cluster's ACSFs and leap-times were averaged, managing noise and ensuring accurate leap-time representation for KMC simulations. These processed data were utilised to train the FNN model.

Before training, further processing was conducted on the averaged leap-times.

The leap-times were initially converted to leap-frequencies, by inverting them (one divided by leap-time). Subsequently, a logarithmic transformation was applied, benefitting the back-transformation of FNN-predicted leap-frequencies. This adjustment was necessary due to the considerable variation in leap-time (and thus leap-frequency) magnitudes, from picoseconds to tens of seconds. Moreover, the logarithmic transformation ensures that even potentially negative numbers become non-negative upon exponentiation, upholding the physical impossibility of a negative leap-frequency. Following the logarithmic transformation, standardisation was applied, adjusting the leap-frequencies to a mean of 0 and a standard deviation of 1.

The combined application of logarithmic transformation and standardisation enhanced data comparability, facilitated efficient FNN learning, and ensured non-negative predictions. Crucially, these steps preserved the physical interpretation of predicted leap-times when returned to the original scale post-standardisation.

5.4.2 Refinement of parameters

During the stage of model construction, an array of techniques was implemented to bolster performance whilst safeguarding against potential complications of overfitting and underfitting [72, 73]. Among these was the L2-regularisation technique [74], crafted to forestall the model from allocating excessive weightage to a single feature, thus minimising the likelihood of overfitting. Concurrently, dropout [75] was utilised, a strategy which randomly deactivates a predetermined percentage of neurons during the training process to instil resilience within the model. This strategy promotes the network's capability to identify useful features within diverse contexts, thereby diminishing the model's reliance on a singular neuron. These approaches were supplemented with the Rectified Linear Unit (ReLU) activation function [76], renowned for its efficacy in addressing non-linearity and its contribution towards computational efficiency in the training procedure. Alongside the application of the R^2 metric* to evaluate model performance, these methods collectively augment the model's capacity to successfully extrapolate from the training data to unobserved data.

Employing the R^2 metric is of critical significance when the model is put into practice for KMC simulations. Its foremost merit is its emphasis on the overall trend between atomic environment context and leap-time rather than focusing on the exact

*The R^2 metric, also known as the coefficient of determination, is the square of the Pearson correlation coefficient r. It quantifies the proportion of variance in the dependent variable explained by the independent variables in a linear model. In essence, $R^2 = r^2$, where r is the correlation between the model's predictions and the actual data.

value of the leap-time. Indeed, within the realm of KMC simulations, the precise leap-time value bears no intrinsic physical relevance except as a measure of the elapsed time. The model's adeptness at accurately distinguishing relative energy barrier (leap-time) differences among leap-candidates with unique environments greatly influences the system's evolution and, hence, the outcomes of the simulation. The R^2 score adequately addresses this requirement by highlighting accuracy in these relative differences, thus ensuring a reliable and meaningful model.

The data, once pre-processed, logarithmically transformed, and standardised, was subjected to a k-fold cross-validation [77] with 10 folds. Optimal hyperparameters and ACSF parameters were ascertained based on their respective R^2 scores across a selection of training sets. The ultimate optimal values that were chosen for the application in the FNN model, which elevated performance, are delineated in Tables 5.2 and 5.3.

Table 5.2: Optimal hyperparameters for the FNN model.

Hyperparameters

Dropout rate	Learning rate	Neurons	Weight decay[a]
10^{-3}	0.02	70	10^{-5}

[a]In regularisation, the modified loss function L' can be expressed as: $L' = L + \lambda \sum_i w_i^2$. Here, L is the original loss function, w_i are the weights of the neural network, and λ is the weight decay coefficient. The term $\lambda \sum_i w_i^2$ is the L2 regularisation term, and λ controls the extent of regularisation. A higher value of λ will result in more regularisation, hence pulling the weights towards zero and simplifying the model.

Table 5.3: Optimal ACSF parameters. Here, η_r and R_s are η and R_s for the radial symmetry function (two-body term) in [68], and η_a , ζ , and λ are η , ζ , and λ for the angular symmetry function (three-body term) therein.

ACSF parameters

η_r	R_s	η_a	ζ	λ
0.03	0	0	2	-1

5.4.3 Assessing the applicability of FNN model to rigid lattices

In order to substantiate the viability of the FNN model, educated with CVHD, for the representation of rigid lattices that frequently occur in KMC simulations where temperature is only indirectly considered, an evaluation procedure was embarked

upon, the details of which are elaborated in this section.

The intention was to verify that the FNN model, in spite of its training using atomic environment information sourced from non-rigid lattices within CVHD simulations that incorporate thermal oscillation, could still provide accurate predictions of relative differences in leap times. This was based on atomic environment data obtained from the fixed lattices within KMC simulations where there is no presence of thermal oscillation.

The process was initiated by producing $12 \times 12 \times 12$ BCC cells, each containing a single vacancy, for a randomly generated Mo-Nb-Ta-V-W alloy, alongside several binary solutions. These originated as fixed lattices, their lattice constants mirroring those of the thermally equilibrated MD cells with equivalent compositions. These cells were subsequently thermalised with the use of a 1-fs time step, within an NPT ensemble at 800 K, for a duration of 10^4 steps.

After thermalisation, the first nearest neighbours (1NNs) of the vacancy, the candidates for leaping, were identified in both the rigid and non-rigid lattices. The FNN model was utilised to estimate the leap frequencies of these candidates in the rigid structure, based on their ACSF descriptors. Furthermore, I determined the equivalent ACSF descriptors, averaged over the concluding 10 frames (each separated by 100 time steps) of the thermalisation trajectory, to estimate the leap times of the matching 1NNs in the thermalised, non-rigid structure.

The findings are introduced in Sec. 6.4.

5.5 KMC simulations and MCMD verification

The validity of my FNN model was tested utilising Di-Sword (Di-Sword is described in Sec. 5.1). Di-Sword is capable of modelling systems and subsystems of Mo-Nb-Ta-V-W with the aid of ACSF-fed FNN models. Di-Sword incorporates a mechanism that restricts the system from reverting to prior states, capitalising on the efficient framework and integer-oriented calculations of the Kimocs code.

Within the KMC simulations, I prohibited systems from returning to the immediate preceding state to prevent stagnation in energy basins. This restriction is executed in Di-Sword by augmenting the transition barriers back to recent states to 1 keV. Although this adjustment may seem to significantly modify the simulation dynamics, its influence on the results was marginal in reality. To validate this, I

performed simulations both with and without the mechanism to inhibit return to recent states. The predicted concluding SRO parameters were remarkably similar in both situations. However, simulations incorporating the prevention mechanism required fewer time-steps to converge.

In this research, I created random equimolar lattices of a quinary Mo-Nb-Ta-V-W alloy, along with various binary solutions distinguished by differing mixing energies, specifically Mo-W (approximately zero), Nb-V (positive), Ta-V (positive), and Ta-W (negative) [32]. The lattice parameter for each cell, relevant to the ACSF descriptors, was calculated from the CVHD cells with corresponding compositions. For each composition, I generated five separate random $12 \times 12 \times 12$ lattices, each with a distinct seed, and inserted a vacancy at a random location. Each of these lattices underwent a Di-Sword KMC simulation, encompassing approximately 10^6 steps.

To confirm that my FNN-KMC model precisely represented the physical phenomena of interest, I utilised MCMD simulations, a blend of the Metropolis-Hastings Monte Carlo algorithm [78] (employed for atomic swapping) and Molecular Dynamics approaches. MCMD strives to reach the configurational energy minimum of a system via the execution of atomic swaps, following a stochastic process. In a multi-element system, such atomic exchanges often result in a degree of short-range ordering (SRO) of the elements.

MCMD simulations, functioning within an NPT ensemble at 800 K, encompassed $4 \cdot 10^6$ steps, adopting a 1-fs time-step. For every unique pair of different elements, five swaps were executed every 500 steps. The lattices were initially thermalised to 800 K for 10^5 steps within the NPT ensemble, before commencing MCMD swapping. Given the reliable convergence of the MCMD algorithm, only a single simulation was run for each composition.

SRO parameters were extracted from the *steady state interval* at the conclusion of each simulation. For the MCMD simulations, both SRO parameters and potential energy were taken into account when ascertaining if the system had reached a steady state. In the KMC simulations, I relied solely on the SRO parameters. An illustration of the steady-state region in a V-W KMC simulation is depicted in Fig. 5.3.

Upon pinpointing the frames which exhibit steady-state behaviour in each simulation, I averaged the SRO parameters within these frames, thereby generating a single set of average SRO parameters for each separate simulation.

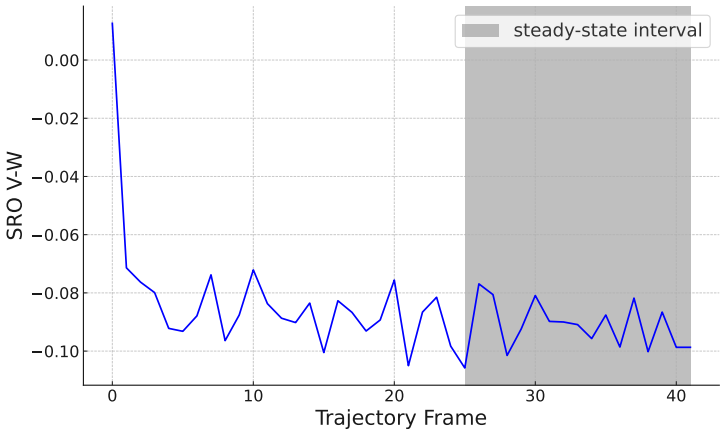


Figure 5.3: Illustration of what I define as a steady-state region (shaded region) in a V-W KMC simulation. The SROs are evaluated within the 1NN shell.

The Cowley SRO parameters are defined as [79]:

$$\alpha_{\mathrm{A,B}} = 1 - \frac{p_{\mathrm{A,B}}}{n_{\mathrm{B}}} \tag{5.8}$$

where $p_{\mathrm{A,B}}$ denotes the *observed* proportion of element B atoms located around an atom of element A, and n_{B} signifies the fraction of B atoms within the entire system. Thus, the Cowley SRO parameter encompasses the preference level for A and B to be together (negative value), separate (positive value), or show indifference (zero value), relative to the overall composition of B atoms in the system.

This metric permitted me to directly compare the predictions of the FNN-KMC model with the equivalent outcomes from the MCMD model, thereby offering a gauge of the model's capacity to capture the inherent environment-dependent, configurational energetics.

The results are exhibited in Sec. 6.5.

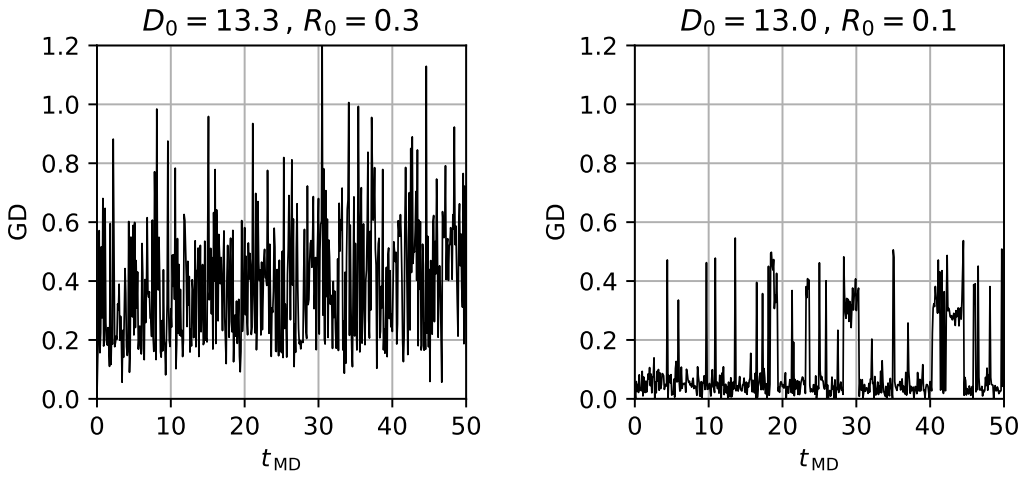
6. Results

6.1 Addressing challenges in implementing CVHD

The implementation of CVHD raised several distinct difficulties. Primarily, the system, described by a BCC structure, exhibits very small separation between the first and second nearest neighbours (1NN and 2NN, respectively). The complexity of the system induced substantial distortion to the BCC lattice, making it far from an ideal lattice. Additionally, specific pure elements present in my system, notably W and Mo, displayed considerable energy barriers for vacancy migration in their elemental form. As a result, the HEA introduced significant energy barriers for vacancy migration, necessitating high temperatures (specifically, 800 K) to record a reasonable count of vacancy leaps. The increased temperature added to the system noise, which further obscured the differentiation between 1NNs and 2NNs.

In order to overcome these challenges, I substantially modified the parameters used in Ebina et al. [2] for determining the position of the virtual atom. This modification yielded a more stable position for the virtual atom and a more consistent list of 1NNs. As stated in Sec. 5.4.1, I noted that the biasing generated by CVHD sometimes produces additional vacancies in systems with slow vacancy diffusion, an issue that was resolved by significantly altering the biasing factor, γ (refer to Sec. 6.1).

In order to address the issues of lattice distortion and substantial thermal agitation, modifications were made to the parameters pertinent to the estimation of atom coordination: D_0 and R_0. These parameters impact the calculated position of the virtual atom (vacancy). For the validation of these parameters, the alloy V-W was modelled at 1000 K, a temperature notably above the 800 K employed in my actual CVHD simulations, to heighten the manifestation of the challenges encountered. The selected alloy, V-W, which consists of elements with substantially varying sizes, added complexity in distinguishing 1NNs and 2NNs by distance due to a highly distorted lattice structure.



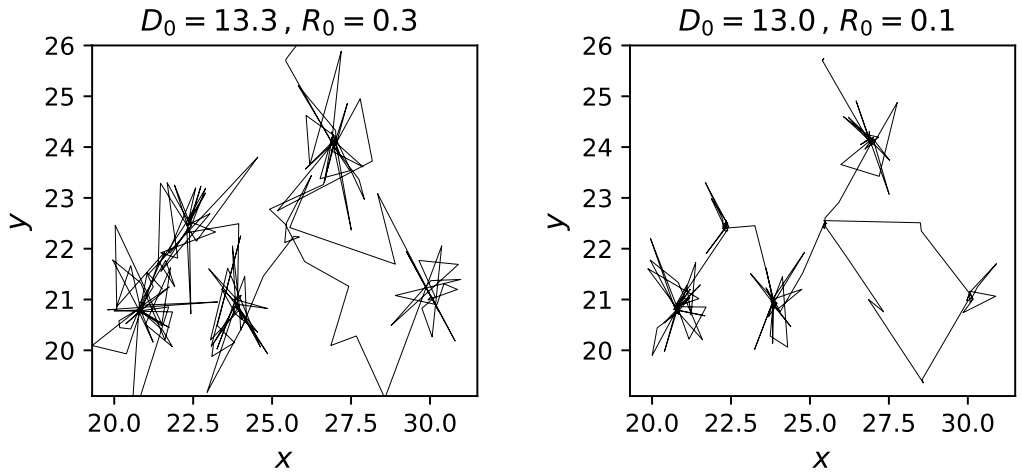
(a) **(b)**

Figure 6.1: Comparison of global distortion (GD) in a V-W alloy at 1000 K using different parameter sets for calculating atoms' coordination numbers. The x-axis represents time in picoseconds (non-accelerated, MD time), and the y-axis illustrates the global distortion. In Fig. 6.1a, I employed the parameters of Ebina et al. [2], and in Fig. 6.1b, I used parameters more closely aligned with my own. Please note that any GD exceeding my cut-off value of 0.6 halts the application of CVHD's bias potential.

Figures 6.1 and 6.2 illustrate the effect of these parameter modifications. A tidier, less distorted GD is noticeable with parameters near to those used in my research, as shown in Fig. 6.1, while Fig. 6.2 reveals a more stable trajectory of the virtual atom accompanied by reduced disturbance.

My investigation of one of the most formidable test cases—the W system at 800 K—unveils the complexities of choosing the suitable biasing factor, γ for CVHD applications. This system's inherent challenge is rooted in its large energy barriers. Efforts to impose the bias potential on the system were thwarted by the virtual atom's marked hesitance to relocate, resulting in an undesired concentration of the bias potential at a singular point. This led to the creation of an exaggerated 6-eV bias potential hill with a γ of 30, vastly exceeding the actual barrier. While conventional wisdom for the choice of γ proposes it should approximate the barriers one anticipates to encounter [48], my research questions this concept.

In relation to tungsten, the $\langle 111 \rangle$ direction vacancy diffusion barrier, measured at approximately 1.8 eV via the NEB method [5, 6], would typically recommend a γ around 30. However, my results suggest that allocating a γ value in this range led to an unwanted aggregation of the bias potential. This discovery bears two important



(a) **(b)**

Figure 6.2: Trajectory of the virtual atom projected onto the x-y plane for an equimolar V-W alloy at 1000 K, contrasting the use of different parameters. Fig. 6.2a employs Ebina et al.'s parameters [2], and Fig. 6.2b utilises parameters more similar to ours.

ramifications: firstly, the hypertime calculation for larger gamma values proved to be non-physical, over 10^{30} times the age of the universe, in sharp contrast to an expected timescale on the order of seconds (according to the Arrhenius equation). Secondly, the excessive biasing within the system had the capacity to warp the lattice, possibly triggering the formation of additional defects.

Remarkably, a significantly divergent γ of 6, not usually advocated by traditional guidelines (and not illustrated in Fig. 6.3) was chosen. This is due to unwanted piling of the bias potential for values of γ too high, as demonstrated in Fig. 6.3. My chosen value of 6 yielded a leap-time of approximately 20 seconds, which aligns closely with the timescale predicted by the Arrhenius equation, on the order of 10^0 seconds, assuming an attempt frequency of 10^{12} Hz, a temperature of 800 K, and a barrier of 1.8 eV.

This research emphasises the criticality of empirical exploration in determining the appropriate γ value, particularly in complex systems such as tungsten, where large barriers pose unique challenges. My findings also underscore the potential ramifications of excessive accumulation of bias potential, which may result in non-realistic hypertime predictions and potential lattice distortions.

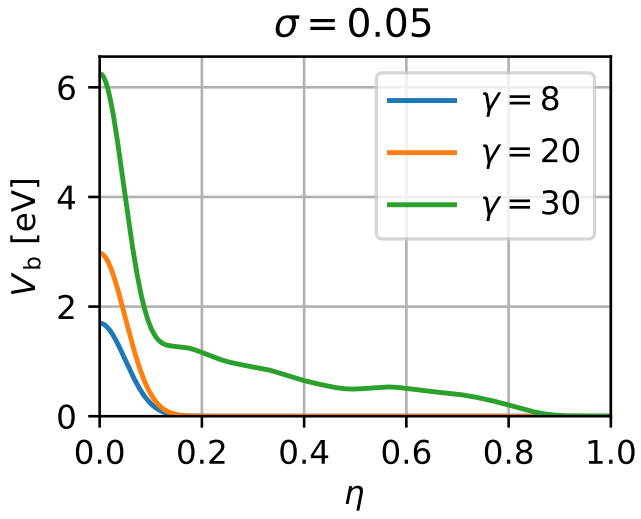


Figure 6.3: The impact of the biasing factor, γ, on the distribution of the total deposited bias potential (V_b) relative to the collective variable, η, in a W lattice at 800 K prior to the first site-leap. σ is the standard deviation of the deposited Gaussian bias packets.

6.2 SRO curiosity in Mo-Nb-Ta-V-W

A curious finding came about when I applied MCMD to an equimolar Mo-Nb-Ta-V-W alloy 39 366 atoms in size. In stark contrast to the elemental ordering observed in a smaller, 6750-atom cell studied in [32], I observed that slab-like elemental formations had formed in the final lattice, as shown in Fig. 6.5 . I confirmed that this did not occur in the smaller cell with similar settings, which implies that whether the formation of these slabs is energetically auspicious depends on the size of the cell.

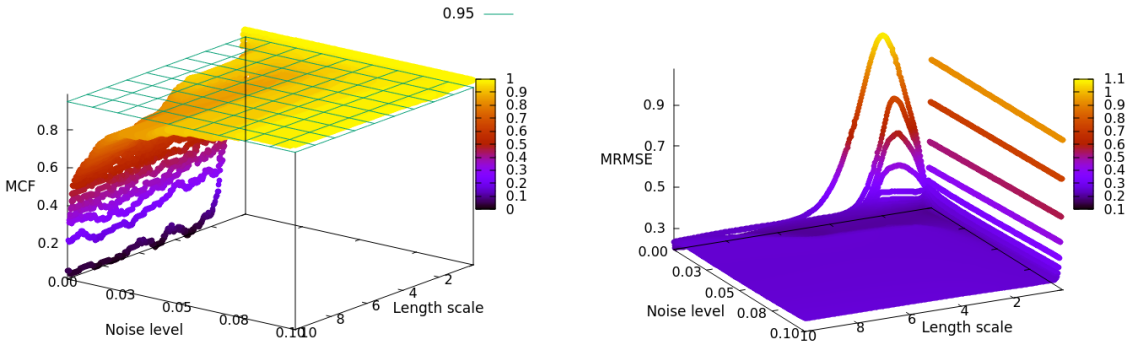
The *slabbing* could have to do with, e.g., interplay of the energy penalty arising from the interface between the V–W-rich phase and the Mo–Nb–Ta-rich phase (in terms of lattice parameter, V is the biggest outlier of the five elements, which induces strain when V is placed into these lattices), as well as the energy penalty from mixing these phases together. In a larger cell, the ratio between interface area and intraphase (inside-of-phase) bulk volume becomes small enough that the energy increase due to the formation of an interface is more advantageous than the trade-off of mixing these phases. It could be that, in reality, this sort of structure is behind energy barriers too high to surmount to be observed in real HEAs, as MCMD does not take kinematics into account, but rather finds the thermodynamically most advantageous configuration. To that end, due to the presence of periodic boundaries, these slabs are infinitely long; to form such a perfectly uninterrupted structure, many energy barriers

would probably need to be overcome in reality, making the formation unlikely to occur (slabs are a very specific structure, so there is the probabilistic challenge that this slabs would occur over the innumerable other configurations). On the contrary, Mo-Nb-Ta-V-W has been experimentally found to possess dendritic formations. Though not strictly dendrites, slabs might be more energetically propitious, which could imply that dendrites are the precursors to these slabs, but the energy barriers separating these could be too great to overcome.

The slab-like formations present an opportunity to try out the capabilities of my GPR model.

6.3 Hyperparameters and GPR Predictive Case Studies

As discussed in Sec. 5.3, having an optimal set of hyperparameters is imperative for a reliable GPR model. The results of the GPR hyperparameter optimisation are illustrated in Fig. 6.4. Based on these results, I decided to choose 1.95 for length scale, and 0.0753 for noise level.



(a) **(b)**

Figure 6.4: Results of optimising various hyperparameter pairs, 'Noise level' (White kernel) and 'Length scale' (RBF). MCF stands for 'mean correct fraction' and MRMSE for mean RMSE (both elaborated on in Sec. 5.3).

My findings indicate that vacancy diffusion, in the monoelemental cases, is slowest in W, and fastest in V. Though the V-rich regions contain W, W seems to be more spread about in the lattice, which could explain why the V-W slabs have the highest diffusivity, as is corroborated by Fig. 6.6, wherein one can see that a decreasing

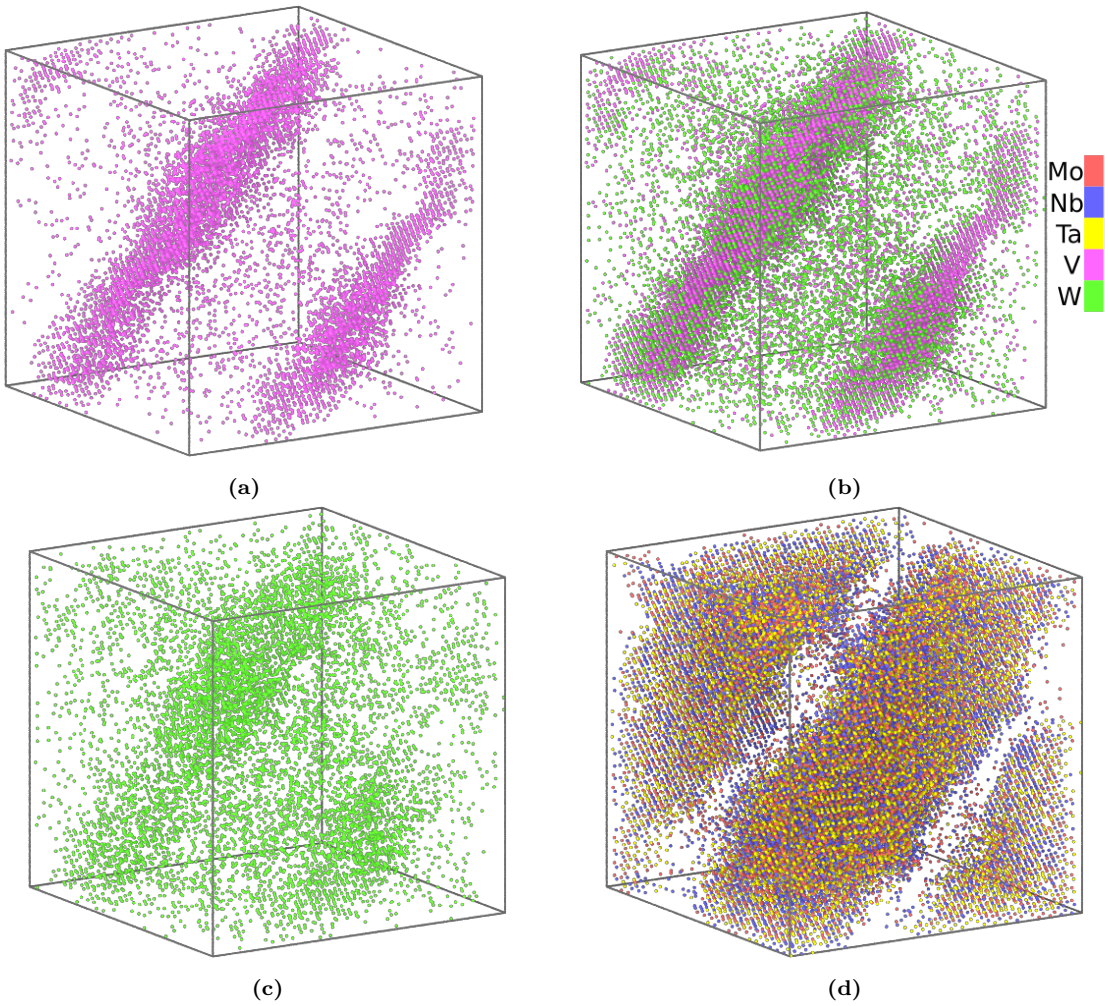


Figure 6.5: Elemental distribution in an equimolar Mo-Nb-Ta-V-W alloy after applying MCMD.

concentration of W in V-W results in an exponential increase in diffusivity. If the slabs present in my MCMD HEA are indicative of the chemical configurations of dendrites in a real Mo-Nb-Ta-V-W HEA, it could mean that vacancies in the regions of higher diffusivity would be restricted in how free they are to migrate around in the material, as they would get trapped in these regions of lower diffusivity. This could allow for vacancies to form clusters, even if they are initially sparsely distributed.

In Fig. 6.7, one can observe a colour map, represented by coloured atoms in a lattice, that illustrates the predicted diffusivity of a single vacancy should it replace any of the existing atoms in the slab-containing HEA, as forecast by GPR. It is worth noting that the GPR predictions are for 800 K, whilst the cell is 300 K, as the slabbing is present at low temperatures, whilst training the model is more challenging at lower temperatures, but the GPR predictions should provide a qualitative insight nonetheless. In this figure, one can see that vacancies in the V-rich regions are many

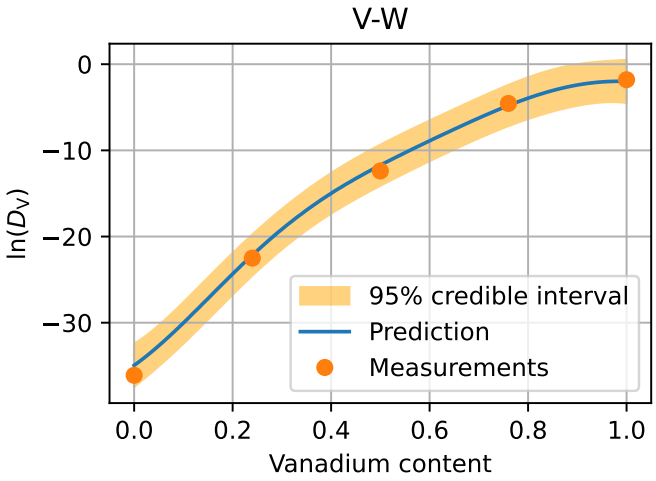


Figure 6.6: Figure of GPR-predicted logarithmic vacancy diffusivity in a V-W as a function of V-concentration. The markers represent the studied compositions of V-W, the connecting line is the GPR-predicted function, and the shaded region is the 95-% credible interval, as per the uncertainty (standard error) GPR gives for each prediction. Note that the uncertainty at the training data markers is non-zero due to the White kernel, which means that the fitting points are not assumed to be free of noise.

orders of magnitude more likely to diffuse than in the Mo–Nb–Ta-rich regions.

6.4 Assessing the transferability of the CVHD-FNN model: A comparative analysis of rigid and non-rigid lattice settings

The CVHD method, integrating system dynamics such as pressure and entropy explicitly, operates at decidedly non-zero temperatures. Consequently, this leads to distorted lattices where atomic distances and angles may take on a continuum of values. It is therefore of utmost importance that an FNN model, trained on these distorted lattices, is capable of generalising to the rigid lattices observed in my KMC simulations, characterised by quantised distances and angles. In this segment, I delve into the transferability of my FNN model from CVHD to KMC.

Fig. 6.8 demonstrates the predicted jump frequencies for the eight leap-candidates around a vacancy in two diverse lattices: the equimolar HEA, Mo-Nb-Ta-V-W, and the equimolar binary, Ta-V. Ta-V has been selected to be highlighted as it exhibits the most blurred boundary between 1NN and 2NN distances, indicative of a highly distorted lattice. This implies that if the FNN model can generalise to Ta-V, given

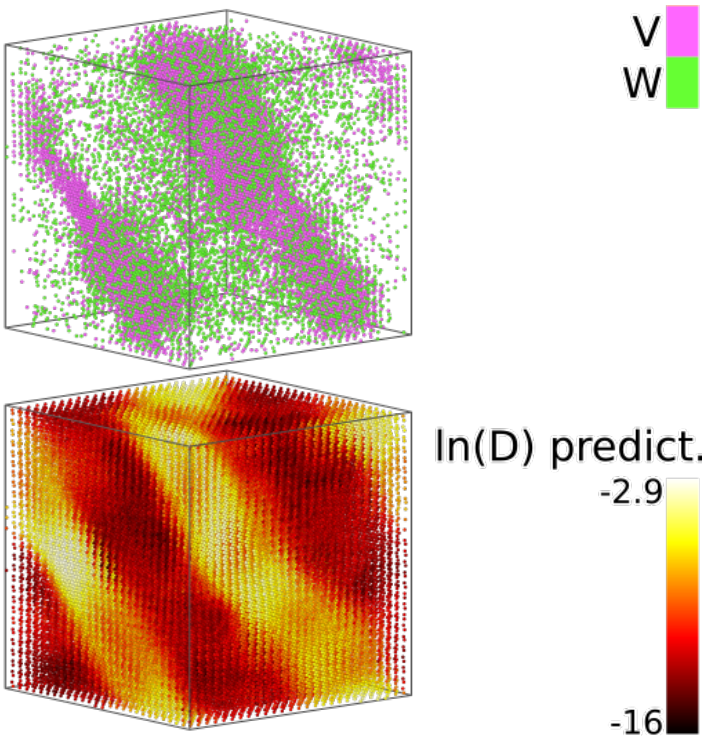
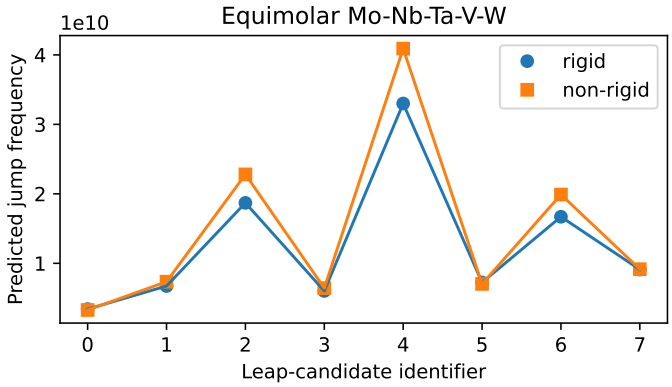


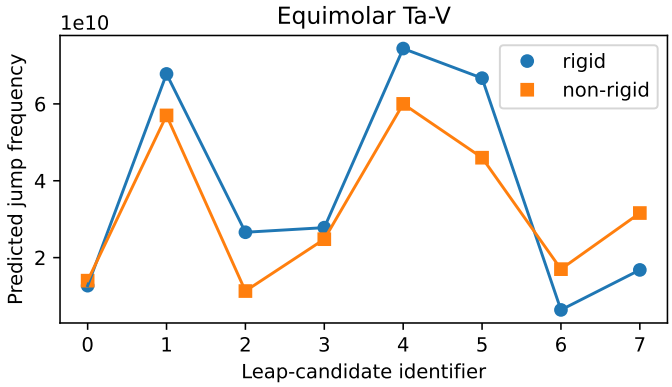
Figure 6.7: At the top, V and W in the 300-K, MCMD Mo-Nb-Ta-V-W lattice are given to guide the eye. The bottom is a colour map of the logarithmic diffusivity a vacancy would be expected to have depending on where in the system it were placed. The predictions at any given atom's position are determined from the chemical composition of atoms within a 10-Å cutoff radius.

the continuum of angles and distances in non-rigid lattices, it ought to generalise to all other lattices with adequate training. This suggests that the differences between the quantised properties of a rigid lattice and the continuum properties of a non-rigid lattice do not significantly influence the robustness of an ACSF-equipped FNN model.

Despite Fig. 6.8 revealing some variations in the forecasted jump frequencies between the rigid and non-rigid lattices, the overall pattern (i.e., relative variations between the predictions) stays consistent. This is true even for Ta-V, arguably the most difficult lattice. Given enough training data, these findings propose that an FNN model trained on non-rigid lattices can generalise to rigid lattices without a considerable loss in precision. This endorses the use of CVHD as a feasible method for training FNN models for KMC simulations.



(a)



(b)

Figure 6.8: Comparison of forecasted jump frequencies for eight leap-candidates around a vacancy in two lattices of identical elemental occupation. The rigid lattice, indicative of my KMC simulations, is an ideal BCC structure without thermal motion. The non-rigid lattice, indicative of my CVHD simulations, is a thermally distorted BCC structure at 800 K. The y-axis denotes the FNN-predicted jump frequency, in units s^{-1}, of the leap-candidates on the x-axis. Each lattice is $12 \times 12 \times 12$ unit cells in dimensions.

6.5 Evaluating the FNN-KMC model: An analysis of elemental ordering in MCMD and FNN-KMC

The MCMD technique is an efficacious method for detecting the global energy minimum of a system through swapping the positions of distinct chemical elements, which results in elemental ordering. In view of the variations in configurational energies in different chemical contexts, one could anticipate that a vacancy moving within a lattice would prefer paths that sequentially lower the system's energy.

On successful implementation, the CVHD-FNN model ought to predict these energy-favourable paths accurately, yielding an elemental sequence akin to that observed in MCMD simulations. Hence, a comparison of the elemental ordering anticipated by the CVHD-FNN model and MCMD can offer not only a measure of validation but also a glimpse into the predictive capabilities of the model.

In this section, I elaborate on the outcomes of this comparative study and delve into the implications of my discoveries.

Fig. 6.9 displays the Cowley SRO parameters as reported by the KMC and MCMD simulations for the Mo-Nb-Ta-V-W HEA, the Mo-W binary (zero mixing energy [32]), the Ta-V and Nb-V binaries (highly positive mixing energy [32]), and the Ta-W binary (highly negative mixing energy [32]).

As discernible in Fig. 6.9, the KMC and MCMD predictions align well overall. The KMC simulations generally capture the sign of the SRO parameters accurately, signalling the configurational affinity among specific elements. For example, in Fig. 6.9a, the KMC simulations predict most of the SRO signs correctly, notwithstanding the intricacy of the lattice. Additionally, in Fig. 6.9b, despite the signs being different, both the KMC and MCMD simulations predict near-zero affinity, as anticipated for the Mo-W binary with zero mixing energy.

The most common divergence between the KMC and MCMD predictions usually lies in the magnitude of the SRO parameters. The KMC simulations struggle to achieve the global energy minimum compared to MCMD, which depends on the route a vacancy takes and the time I can afford for vacancy diffusion. Energy-optimal configurations may also be located beyond significant barriers, which are challenging to overcome within the time limit of my KMC simulations. This divergence is particularly noticeable for the Ta-V and Nb-V binary alloys, possibly due to their significant lattice distortion.

In the future, I recommend investigating the use of descriptors encapsulating higher-order terms or expanding the ACSF descriptor with supplementary two- and three-body terms. However, such modifications could result in increased descriptor complexity and computational expenditure, likely requiring larger quantities of training data to ensure reliable performance of the FNN model.

Another consideration would be to implement a binning or bucketing process during the training stage. In this strategy, continuous values such as distances or

angles could be categorised into distinct ranges or 'bins', and then input into the machine learning model as a single representative value. This might permit better management of complexities introduced by the non-rigid lattice and its wide range of distances and angles.

My work primarily aimed at demonstrating the efficiency and advantages of training an FNN model on CVHD data for use in KMC simulations. Moreover, I endeavoured to demonstrate that the CVHD methodology could be successfully applied to study vacancy diffusion in complex BCC high-entropy alloys. I am confident that my methods and discoveries lay the foundation for future refinements and investigations of these techniques.

The results imply that the FNN-KMC model can generalise to unseen atomic environments and lattices larger than the $5 \times 5 \times 5$ unit cells used in the CVHD simulations. Although the model encountered some obstacles with the most distorted lattices, it generally captured the correct sign in the SRO affinities. This implies that an FNN model trained on CVHD data can serve as a reliable tool for future KMC simulations.

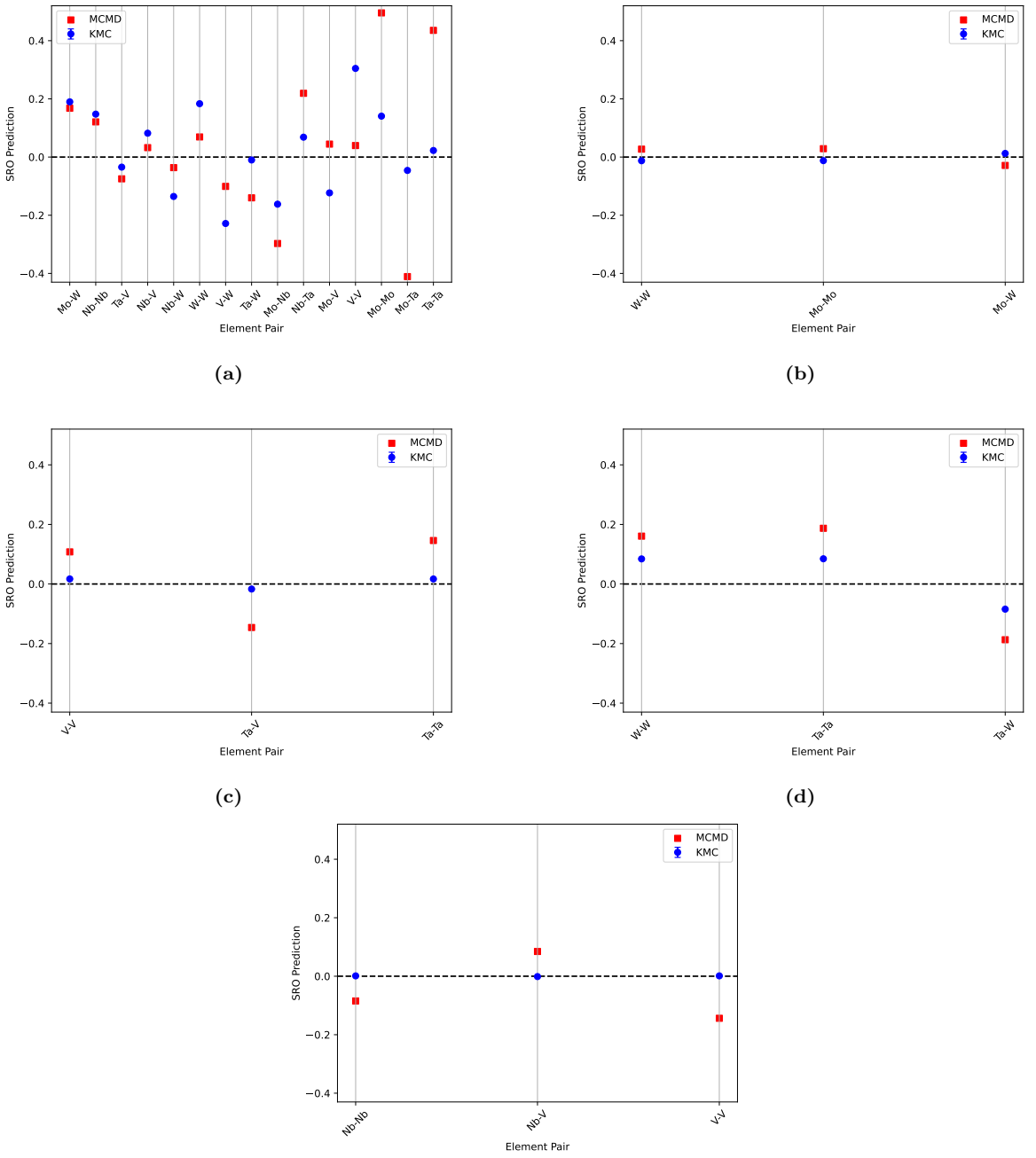


Figure 6.9: Comparison of Cowley SRO parameters as reported by the KMC and MCMD simulations. The barely discernible error bars denote the population standard errors of the mean values from each of the five KMC simulations. The SRO parameters are determined for the atoms inside the 1NN shell. Each cell measures $12 \times 12 \times 12$ unit cells.

7. Summary of conclusions

This investigation has presented a comprehensive methodology for modelling vacancy diffusion within BCC HEA systems. The integration of the CVHD method with FNN models, subsequently deployed within the KMC framework, has been a significant step in this direction. The FNN model has showcased remarkable robustness in its ability to generalise to unobserved data.

In a separate endeavour, a GPR model was trained on CVHD data to study vacancy diffusivity in relation to chemical composition. The GPR model, despite its known limitations, provided valuable qualitative insights when applied to findings from MCMD simulations. These MCMD simulations revealed a slab-like separation of V–W-rich and Mo–Nb–Ta-rich phases. Though these slabs might be more energetically favourable than the dendritic structures confirmed by experimental efforts to exist in equimolar Mo-Nb-Ta-V-W [80], their formation could be locked behind energy barriers too large to be reached in reality. Notably, GPR predicted that these phases exhibit distinct vacancy diffusivities. Though slabs are not dendrites, both are elongated in shape, and could thus offer paths for vacancies to traverse through, which could restrict the degrees of freedom associated with this motion. This could explain why vacancies that are initially separated by a great distance could form clusters if heated up, which is what the experimental findings that Dr. Eryang Lu has discussed with me show happens in Mo-Nb-Ta-V-W. Vacancies could move faster along the higher-diffusivity dendrites in a restricted region of space, like a tunnel, whilst vacancies that end up in regions of lower diffusivity would be static in comparison to the faster vacancies, which would increase the likelihood of clustering, which is driven by the short-range binding of vacancies.

The potential of high-entropy alloys, especially in the context of future fusion reactors, is vast. In these reactors, trace elements, such as minor gases like hydrogen and helium, can significantly influence the material's attributes and longevity. The CVHD method, with its unique capability to function at non-zero temperatures, stands out, especially when compared to techniques like NEB. This functionality not

only allows for the study of pressure effects by confined gases in lattice vacancies but also incorporates the entropic component of free energy. The potential embrittlement caused by trapped gases remains a focal point of interest, given its implications for fusion reactor safety. The CVHD method, enriched by the insights from MCMD simulations, could promise a deeper understanding of these systems, potentially paving the way for the design of safer and more efficient fusion reactors. Furthermore, the CVHD method's versatility is evident when contrasted with other finite temperature accelerative/sampling methods, such as metadynamics. Its adaptable collective variable, which focuses on proximity to a state-transition rather than the state itself, is particularly beneficial for studying complex alloys, due to the limited knowledge required of the future evolution of the system.

However, it is worth noting that the current CVHD implementation predominantly supports single vacancy modelling.* Future work will necessitate modifications to this framework to model vacancy clusters, vital for studying gas embrittlement. This research establishes a robust foundation for addressing such complexities, highlighting the potential of the combined FNN-KMC methodology in probing these intricate systems.

In conclusion, this investigation underscores the immense potential of the evolved methodology for modelling diffusion processes in complex materials. It represents a significant advancement in the study of high-entropy alloys and offers a robust framework adaptable to analogous systems, heralding new avenues for advanced material design.

*While CVHD can technically model multiple vacancies, its representation becomes less meaningful as they come close enough to each other to form a cluster.

Bibliography

[1] O. El-Atwani et al. “Outstanding Radiation Resistance of Tungsten-Based High-Entropy Alloys”. In: *Science Advances* 5.3 (2019), eaav2002. DOI: 10.1126/sciadv.aav2002. eprint: https://www.science.org/doi/pdf/10.1126/sciadv.aav2002.

[2] Hiroyasu Ebina, Satoru Fukuhara, and Yasushi Shibuta. “Accelerated Molecular Dynamics Simulation of Vacancy Diffusion in Substitutional Alloy with Collective Variable-Driven Hyperdynamics”. In: *Computational Materials Science* 196 (Aug. 2021), p. 110577. ISSN: 09270256. DOI: 10.1016/j.commatsci.2021.110577.

[3] Kristof M. Bal and Erik 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. ISSN: 1549-9618, 1549-9626. DOI: 10.1021/acs.jctc.5b00597.

[4] Jyri Kimari et al. “Biased Self-Diffusion on Cu Surface Due to Electric Field Gradients”. In: *Journal of Physics D: Applied Physics* 55.46 (Sept. 2022), p. 465302. ISSN: 0022-3727. DOI: 10.1088/1361-6463/ac91dd.

[5] Blas Uberuaga and Hannes Jonsson. “A Climbing Image Nudged Elastic Band Method for Finding Saddle Points and Minimum Energy Paths”. In: *J. Chem. Phys.* 113 (Dec. 2000), pp. 9901–9904. DOI: 10.1063/1.1329672.

[6] Hannes Jonsson. “Improved Tangent Estimate in the Nudged Elastic Band Method for Finding Minimum Energy Paths and Saddle Points”. In: *J. Chem. Phys.* 113 (Dec. 2000), pp. 9978–9985. DOI: 10.1063/1.1323224.

[7] W. M. Shu, G. -N. Luo, and T. Yamanishi. “Mechanisms of Retention and Blistering in Near-Surface Region of Tungsten Exposed to High Flux Deuterium Plasmas of Tens of eV”. In: *Journal of Nuclear Materials*. Proceedings of the Twelfth International Conference on Fusion Reactor Materials (ICFRM-12) 367–370 (Aug. 2007), pp. 1463–1467. ISSN: 0022-3115. DOI: 10.1016/j.jnucmat.2007.04.005.

[8] S. Nagata et al. “Helium and Hydrogen Trapping in W and Mo Single-Crystals Irradiated by He Ions”. In: *Journal of Nuclear Materials* 307–311 (Dec. 2002), pp. 1513–1516. ISSN: 0022-3115. DOI: 10.1016/S0022-3115(02)01269-2.

[9] R. Tivey et al. “ITER R&D: Vacuum Vessel and In-Vessel Components: Divertor Cassette”. In: *Fusion Engineering and Design* 55.2 (July 2001), pp. 219–229. ISSN: 0920-3796. DOI: 10.1016/S0920-3796(01)00212-5.

[10] M. Rieth et al. “Recent Progress in Research on Tungsten Materials for Nuclear Fusion Applications in Europe”. In: *Journal of Nuclear Materials* 432.1 (Jan. 2013), pp. 482–500. ISSN: 0022-3115. DOI: 10.1016/j.jnucmat.2012.08.018.

[11] *Handbook of Chemistry and Physics 102nd Edition*.

[12] S. L. Dudarev and Pui-Wai Ma. “Elastic Fields, Dipole Tensors, and Interaction between Self-Interstitial Atom Defects in Bcc Transition Metals”. In: *Physical Review Materials* 2.3 (Mar. 2018), p. 033602. ISSN: 2475-9953. DOI: 10.1103/PhysRevMaterials.2.033602.

[13] F. Hofmann et al. “Lattice Swelling and Modulus Change in a Helium-Implanted Tungsten Alloy: X-ray Micro-Diffraction, Surface Acoustic Wave Measurements, and Multiscale Modelling”. In: *Acta Materialia* 89 (May 2015), pp. 352–363. ISSN: 1359-6454. DOI: 10.1016/j.actamat.2015.01.055.

[14] “19 - Fracture Toughness Properties of Aerospace Materials”. In: *Introduction to Aerospace Materials*. Ed. by Adrian P. Mouritz. Woodhead Publishing, Jan. 2012, pp. 454–468. ISBN: 978-1-85573-946-8. DOI: 10.1533/9780857095152.454.

[15] Chao Yin et al. “Ductile to Brittle Transition in ITER Specification Tungsten Assessed by Combined Fracture Toughness and Bending Tests Analysis”. In: *Materials Science and Engineering: A* 750 (Mar. 2019), pp. 20–30. ISSN: 0921-5093. DOI: 10.1016/j.msea.2019.02.028.

[16] D. H. Lassila, F. Magness, and D. Freeman. *Ductile-Brittle Transition Temperature Testing of Tungsten Using the Three-Point Bend Test*. Tech. rep. UCRL-ID-108258. Lawrence Livermore National Lab., CA (United States), Mar. 1991. DOI: 10.2172/5273161.

[17] Yang Yu et al. “Effect of Chromium Content on the Corrosion Resistance of Ferritic Stainless Steels in Sulfuric Acid Solution”. In: *Heliyon* 4.11 (2018), e00958. ISSN: 2405-8440. DOI: 10.1016/j.heliyon.2018.e00958.

[18] Ming Liu et al. “Effect of Cr on the Passive Film Formation Mechanism of Steel Rebar in Saturated Calcium Hydroxide Solution”. In: *Applied Surface Science* 389 (2016), pp. 1182–1191. ISSN: 0169-4332. DOI: 10.1016/j.apsusc.2016.08.074.

[19] B. S. Murty, Jien-Wei Yeh, and S. Ranganathan. *High-Entropy Alloys*. Oxford, UNITED STATES: Elsevier Science & Technology, 2014. ISBN: 978-0-12-800526-2.

[20] Yu Zou, Huan Ma, and Ralph Spolenak. "Ultrastrong Ductile and Stable High-Entropy Alloys at Small Scales". In: *Nature Communications* 6.1 (Nov. 2015), p. 7748. ISSN: 2041-1723. DOI: 10.1038/ncomms8748.

[21] O.N. Senkov et al. "Mechanical Properties of Nb25Mo25Ta25W25 and V20Nb20Mo20Ta20W20 Refractory High Entropy Alloys". In: *Intermetallics* 19.5 (May 2011), pp. 698–706. ISSN: 09669795. DOI: 10.1016/j.intermet.2011.01.004.

[22] Yu Zou et al. "Size-Dependent Plasticity in an Nb25Mo25Ta25W25 Refractory High-Entropy Alloy". In: *Acta Materialia* 65 (Feb. 2014), pp. 85–97. ISSN: 1359-6454. DOI: 10.1016/j.actamat.2013.11.049.

[23] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Oxford, England: Oxford University Press, 1989.

[24] Ganesh Sivaraman et al. "Machine-Learned Interatomic Potentials by Active Learning: Amorphous and Liquid Hafnium Dioxide". In: *npj Computational Materials* 6.1 (July 2020), p. 104. ISSN: 2057-3960. DOI: 10.1038/s41524-020-00367-7.

[25] Samuel Tovey et al. "DFT Accurate Interatomic Potential for Molten NaCl from Machine Learning". In: *The Journal of Physical Chemistry C* 124.47 (2020), pp. 25760–25768. DOI: 10.1021/acs.jpcc.0c08870.

[26] Jörg Behler and Michele Parrinello. "Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces". In: *Physical Review Letters* 98.14 (Apr. 2007), p. 146401. DOI: 10.1103/PhysRevLett.98.146401.

[27] Jesper Byggmästar. *Analytical and Machine-Learning Interatomic Potentials for Radiation Damage in Fusion Reactor Materials*. Helsinki, 2020.

[28] Albert P. Bartók et al. "Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons". In: *Physical Review Letters* 104.13 (Apr. 2010), p. 136403. DOI: 10.1103/PhysRevLett.104.136403.

[29] Mikko Koskenniemi et al. "Efficient Atomistic Simulations of Radiation Damage in W and W–Mo Using Machine-Learning Potentials". In: *Journal of Nuclear Materials* 577 (Apr. 2023), p. 154325. ISSN: 0022-3115. DOI: 10.1016/j.jnucmat.2023.154325.

[30] Albert P. Bartók and Gábor Csányi. "Gaussian Approximation Potentials: A Brief Tutorial Introduction". In: *International Journal of Quantum Chemistry* 115.16 (2015), pp. 1051–1057. DOI: 10.1002/qua.24927. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qua.24927.

[31] Albert P. Bartók, Risi Kondor, and Gábor Csányi. "On Representing Chemical Environments". In: *Physical Review B* 87.18 (May 2013), p. 184115. DOI: 10.1103/PhysRevB.87.184115.

[32] J. Byggmästar, K. Nordlund, and F. Djurabekova. "Modeling Refractory High-Entropy Alloys with Efficient Machine-Learned Interatomic Potentials: Defects and Segregation". In: *Physical Review B* 104.10 (Sept. 2021), p. 104101. DOI: 10.1103/PhysRevB.104.104101.

[33] Murray 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. ISSN: 0163-1829. DOI: 10.1103/PhysRevB.29.6443.

[34] M.I. Mendelev et al. "Development of New Interatomic Potentials Appropriate for Crystalline and Liquid Iron". In: *Philosophical Magazine* 83 (Dec. 2003), pp. 3977–3994. DOI: 10.1080/14786430310001613264.

[35] J Byggmästar et al. "Multiscale Machine-Learning Interatomic Potentials for Ferromagnetic and Liquid Iron". In: *Journal of Physics: Condensed Matter* 34.30 (July 2022), p. 305402. ISSN: 0953-8984, 1361-648X. DOI: 10.1088/1361-648X/ac6f39.

[36] Jean Philibert. "III Diffusion Mechanisms and Correlation Effects". In: *Atom Movements: Diffusion and Mass Transport in Solids*. Les Ulis, FRANCE: EDP Sciences, 2020. ISBN: 978-2-7598-0172-5.

[37] R.W. Balluffi. "Vacancy Defect Mobilities and Binding Energies Obtained from Annealing Studies". In: *Journal of Nuclear Materials* 69–70 (Feb. 1978), pp. 240–263. ISSN: 00223115. DOI: 10.1016/0022-3115(78)90247-7.

[38] J. N. Mundy, S. T. Ockers, and L. C. Smedskjaer. "Vacancy Migration Enthalpy in Tungsten at High Temperatures". In: *Philosophical Magazine A* 56.6 (Dec. 1987), pp. 851–860. ISSN: 0141-8610, 1460-6992. DOI: 10.1080/01418618708204493.

[39] Arthur F. Voter. "Introduction To The Kinetic Monte Carlo Method". In: *Radiation Effects in Solids*. Ed. by Kurt E. Sickafus, Eugene A. Kotomin, and Blas P. Uberuaga. NATO Science Series. Dordrecht: Springer Netherlands, 2007, pp. 1–23. ISBN: 978-1-4020-5295-8. DOI: 10.1007/978-1-4020-5295-8_1.

[40] Arthur F. Voter. "Hyperdynamics: Accelerated Molecular Dynamics of Infrequent Events". In: *Physical Review Letters* 78.20 (May 1997), pp. 3908–3911. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.78.3908.

[41] Arthur F. Voter. "A Method for Accelerating the Molecular Dynamics Simulation of Infrequent Events". In: *The Journal of Chemical Physics* 106.11 (Mar. 1997), pp. 4665–4677. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.473503.

[42] Alessandro Laio and Michele Parrinello. "Escaping Free-Energy Minima". In: *Proceedings of the National Academy of Sciences* 99.20 (Oct. 2002), pp. 12562–12566. DOI: 10.1073/pnas.202427399.

[43] Alessandro Barducci, Massimiliano Bonomi, and Michele Parrinello. "Metadynamics". In: *WIREs Computational Molecular Science* 1.5 (2011), pp. 826–843. ISSN: 1759-0884. DOI: 10.1002/wcms.31.

[44] Alessandro Laio et al. "Assessing the Accuracy of Metadynamics". In: *The Journal of Physical Chemistry B* 109.14 (Apr. 2005), pp. 6714–6721. ISSN: 1520-6106, 1520-5207. DOI: 10.1021/jp045424k.

[45] Satoru Fukuhara et al. "Accelerated Molecular Dynamics Simulation of Large Systems with Parallel Collective Variable-Driven Hyperdynamics". In: *Computational Materials Science* 177 (May 2020), p. 109581. ISSN: 09270256. DOI: 10.1016/j.commatsci.2020.109581.

[46] J. D. Doll. "A Unified Theory of Dissociation". In: *The Journal of Chemical Physics* 73.6 (Sept. 1980), pp. 2760–2762. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.440496.

[47] Alessandro Barducci, Giovanni Bussi, and Michele Parrinello. "Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method". In: *Physical Review Letters* 100.2 (Jan. 2008), p. 020603. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.100.020603.

[48] Pratyush Tiwary and Michele Parrinello. "A Time-Independent Free Energy Estimator for Metadynamics". In: *The Journal of Physical Chemistry B* 119.3 (Jan. 2015), pp. 736–742. ISSN: 1520-6106. DOI: 10.1021/jp504920s.

[49] Steve Plimpton. "Fast Parallel Algorithms for Short-Range Molecular Dynamics". In: *Journal of Computational Physics* 117.1 (1995), pp. 1–19. ISSN: 0021-9991. DOI: 10.1006/jcph.1995.1039.

[50] Gareth A. Tribello et al. "PLUMED 2: New Feathers for an Old Bird". In: *Computer Physics Communications* 185.2 (Feb. 2014), pp. 604–613. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2013.09.018.

[51] Massimiliano Bonomi et al. “PLUMED: A Portable Plugin for Free-Energy Calculations with Molecular Dynamics”. In: *Computer Physics Communications* 180.10 (Oct. 2009), pp. 1961–1972. ISSN: 00104655. DOI: 10.1016/j.cpc.2009.05.011.

[52] Ville Jansson, Ekaterina Baibuz, and Flyura Djurabekova. “Long-Term Stability of Cu Surface Nanotips”. In: *Nanotechnology* 27.26 (July 2016). Comment: This is an author-created, un-copyedited version of an article accepted for publication/published in Nanotechnology. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it. The Version of Record is available online at http://dx.doi.org/10.1088/0957-4484/27/26/265708, p. 265708. ISSN: 0957-4484, 1361-6528. DOI: 10.1088/0957-4484/27/26/265708. arXiv: 1508.06870 [cond-mat].

[53] J. Byggmästar, K. Nordlund, and F. Djurabekova. “Simple Machine-Learned Interatomic Potentials for Complex Alloys”. In: *Physical Review Materials* 6.8 (Aug. 2022), p. 083801. DOI: 10.1103/PhysRevMaterials.6.083801.

[54] Loup Verlet. “Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules”. In: *Physical Review* 159.1 (July 1967), pp. 98–103. DOI: 10.1103/PhysRev.159.98.

[55] Alexander Stukowski. “Visualization and Analysis of Atomistic Simulation Data with OVITO-the Open Visualization Tool”. In: *MODELLING AND SIMULATION IN MATERIALS SCIENCE AND ENGINEERING* 18.1 (Jan. 2010). ISSN: 0965-0393. DOI: 10.1088/0965-0393/18/1/015012.

[56] J. D. Hunter. “Matplotlib: A 2D Graphics Environment”. In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95. DOI: 10.1109/MCSE.2007.55.

[57] François Chollet et al. *Keras*. 2015.

[58] Lauri Himanen et al. “DScribe: Library of Descriptors for Machine Learning in Materials Science”. In: *Computer Physics Communications* 247 (Feb. 2020), p. 106949. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2019.106949.

[59] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. Cambridge, Mass: MIT Press, 2006. ISBN: 978-0-262-18253-9.

[60] Jie Wang. *An Intuitive Tutorial to Gaussian Processes Regression*. Comment: 17 pages, 12 figures. Apr. 2022. arXiv: 2009.10862 [cs, stat].

[61] F. Pedregosa et al. “Scikit-Learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[62] William G. Hoover. "Canonical Dynamics: Equilibrium Phase-Space Distributions". In: *Physical Review A: Atomic, Molecular, and Optical Physics* 31.3 (Mar. 1985), pp. 1695–1697. DOI: 10.1103/PhysRevA.31.1695.

[63] Shuichi Nosé. "A Molecular Dynamics Method for Simulations in the Canonical Ensemble". In: *Molecular Physics* 100.1 (Jan. 2002), pp. 191–198. ISSN: 0026-8976. DOI: 10.1080/00268970110089108.

[64] Mark Tuckerman. *Statistical Mechanics: Theory and Molecular Simulation: Theory and Molecular Simulation*. Oxford, UNITED KINGDOM: Oxford University Press, Incorporated, 2010. ISBN: 978-0-19-152346-5.

[65] Aditya Kamath et al. "Neural Networks vs Gaussian Process Regression for Representing Potential Energy Surfaces: A Comparative Study of Fit Quality and Vibrational Spectrum Accuracy". In: *The Journal of Chemical Physics* 148.24 (June 2018), p. 241702. ISSN: 1089-7690. DOI: 10.1063/1.5003074.

[66] José Mira and José R. Álvarez, eds. *Artificial Neural Nets Problem Solving Methods: 7th International Work-Conference on Artificial and Natural Neural Networks, IWANN2003 Maó, Menorca, Spain, June 3–6, 2003 Proceedings, Part II*. Vol. 2687. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2003. ISBN: 978-3-540-40211-4 978-3-540-44869-3. DOI: 10.1007/3-540-44869-1.

[67] Jakob Macke, Philipp Berens, and Matthias Bethge. "Statistical Analysis of Multi-Cell Recordings: Linking Population Coding Models to Experimental Data". In: *Frontiers in Computational Neuroscience* 5 (2011). ISSN: 1662-5188.

[68] Jörg Behler. "Four Generations of High-Dimensional Neural Network Potentials". In: *Chemical Reviews* 121.16 (Aug. 2021), pp. 10037–10072. ISSN: 0009-2665. DOI: 10.1021/acs.chemrev.0c00868.

[69] Volker L. Deringer et al. "Gaussian Process Regression for Materials and Molecules". In: *Chemical Reviews* 121.16 (Aug. 2021), pp. 10073–10141. ISSN: 0009-2665. DOI: 10.1021/acs.chemrev.1c00022.

[70] J. Macqueen. "Some Methods For Classification and Analysis of Multivariate Observations". In: *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* 1.14 (1967), pp. 281–297.

[71] Peter J. Rousseeuw. "Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis". In: *Journal of Computational and Applied Mathematics* 20 (Nov. 1987), pp. 53–65. ISSN: 0377-0427. DOI: 10.1016/0377-0427(87)90125-7.

[72] Yikui Zhai et al. "Robust SAR Automatic Target Recognition Based on Transferred MS-CNN with L2-Regularization". In: *Computational Intelligence and Neuroscience* 2019 (Nov. 2019), e9140167. ISSN: 1687-5265. DOI: 10.1155/2019/9140167.

[73] Xiaoyun Xie et al. "A Hybrid Improved Neural Networks Algorithm Based on L2 and Dropout Regularization". In: *Mathematical Problems in Engineering* 2022 (Nov. 2022), e8220453. ISSN: 1024-123X. DOI: 10.1155/2022/8220453.

[74] Brett Byram et al. "A Model and Regularization Scheme for Ultrasonic Beamforming Clutter Reduction". In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 62.11 (Nov. 2015), pp. 1913–1927. ISSN: 1525-8955. DOI: 10.1109/TUFFC.2015.007004.

[75] Alex Labach, Hojjat Salehinejad, and Shahrokh Valaee. *Survey of Dropout Methods for Deep Neural Networks*. Oct. 2019. DOI: 10.48550/arXiv.1904.13310. arXiv: 1904.13310 [cs].

[76] Johannes Schmidt-Hieber. "Nonparametric Regression Using Deep Neural Networks with ReLU Activation Function". In: *The Annals of Statistics* 48.4 (Aug. 2020), pp. 1875–1897. ISSN: 0090-5364, 2168-8966. DOI: 10.1214/19-AOS1875.

[77] Roozbeh Valavi et al. *blockCV: An R Package for Generating Spatially or Environmentally Separated Folds for k-Fold Cross-Validation of Species Distribution Models*. June 2018. DOI: 10.1101/357798.

[78] W. K. Hastings. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications". In: *Biometrika* 57.1 (1970), pp. 97–109. ISSN: 0006-3444. DOI: 10.2307/2334940.

[79] J. M. Cowley. "An Approximate Theory of Order in Alloys". In: *Physical Review* 77.5 (Mar. 1950), pp. 669–675. ISSN: 0031-899X. DOI: 10.1103/PhysRev.77.669.

[80] O. N. Senkov et al. "Refractory High-Entropy Alloys". In: *Intermetallics* 18.9 (Sept. 2010), pp. 1758–1765. ISSN: 0966-9795. DOI: 10.1016/j.intermet.2010.05.014.