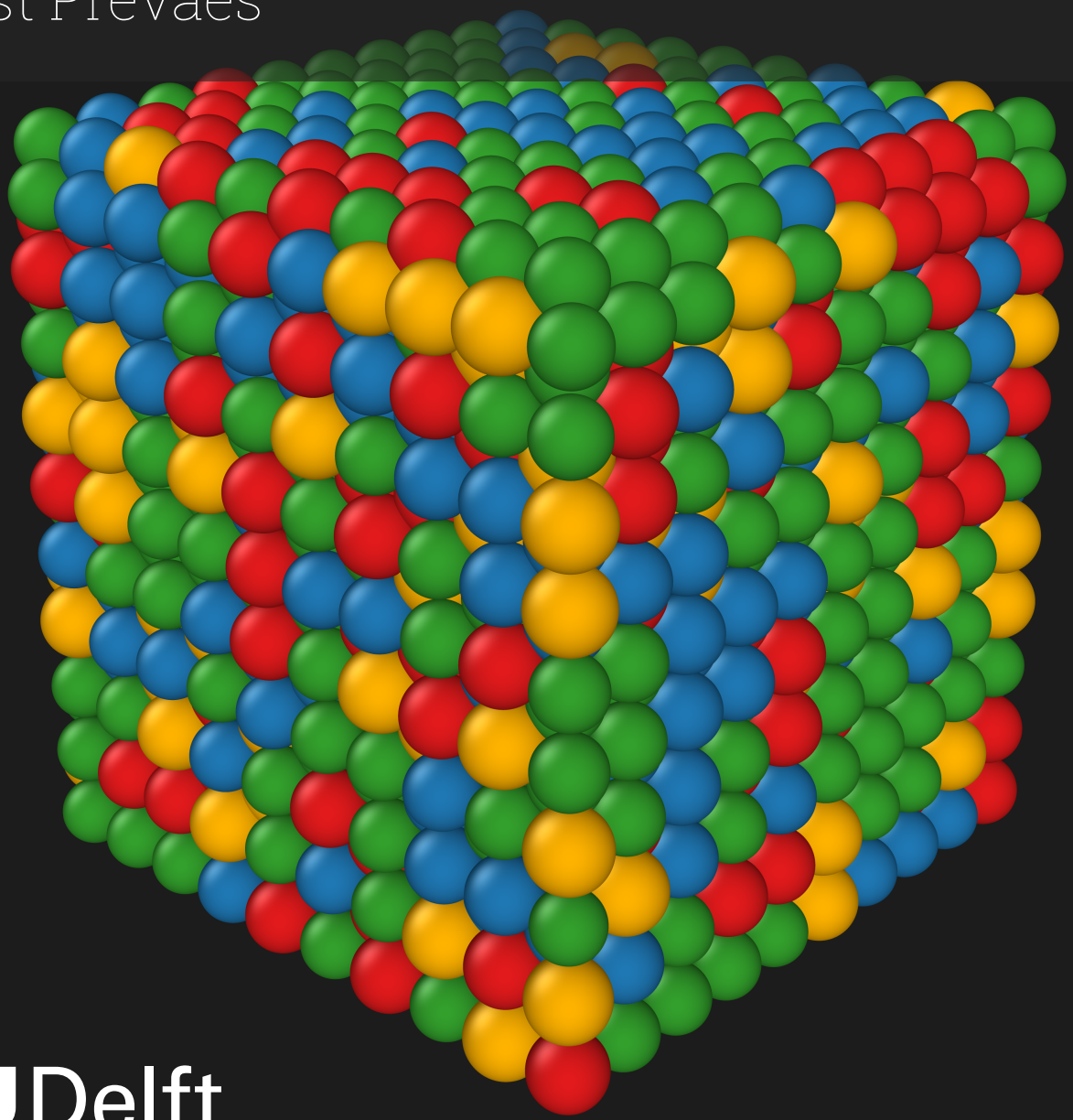
Hydrogen Diffusion in Multi-Principal Element Alloys

A Kinetic Monte Carlo and Machine Learning Framework for Hydrogen Diffusion in Chemically Complex BCC Alloys

Joost Prevaes



Delft University of Technology

TU Delft

Hydrogen Diffusion in Multi-Principal Element Alloys

A Kinetic Monte Carlo and Machine Learning Framework for Hydrogen Diffusion in Chemically Complex BCC Alloys

by

Joost Prevaes

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on Tuesday September 9, 2025 at 10:00.

Student number:	5871182	
Project duration:	February – September, 2025	
Thesis committee:	Dr. Poulumi Dey,	TU Delft, MSE, supervisor
	Dr. Fei Shuang,	TU Delft, MSE, supervisor
	Prof. Dr. Maria Santofimia Navarro,	TU Delft, MSE
	Dr. Luca Laurenti,	TU Delft, DCSC

Cover: $5 \times 5 \times 5$ Atom Supercell of a Body Centred-Cubic High-Entropy Alloy visualised using OVITO.

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

Acknowledgements

First and foremost, I would like to express my sincere gratitude towards Dr. Poulumi Dey and Dr. Fei Shuang for guiding me into the fascinating field of computational materials science. Their openness to questions, even outside of scheduled meetings, and guidance were invaluable and will continue to inspire me. I am also grateful to Prof. Dr. Maria Santofimia Navarro and Dr. Luca Laurenti for taking the time to participate in the assessment committee.

My thanks go to all my friends, who have made my student life at TU Delft far more enjoyable, including, but not limited to, the countless, much-needed coffee breaks during the past months of thesis work and the evenings and weekends spent together outside of campus. Last, but not least, I would like to thank my mother, father, and brother for their unconditional love and support, without which this work would not have been possible.

Joost Prevaes
Delft, September 2025

Abstract

Hydrogen is a promising energy carrier for sustainable energy systems, but its interaction with metallic structures poses significant challenges, particularly hydrogen embrittlement. Multi-principal element alloys (MPEAs), including high- and medium-entropy alloys, offer resistance to hydrogen embrittlement and potential for hydrogen storage due to their disordered atomic lattices, which create effective trapping sites for hydrogen. However, the vast compositional space of MPEAs limits experimental exploration, and conventional simulation approaches are often too computationally intensive for high-throughput screening. This thesis introduces an efficient and comprehensive computational framework for predicting hydrogen diffusivity in body-centred cubic (BCC) MPEAs. The diffusion energy landscape is characterised by statistical parameters that describe the distribution of saddle point and well-energies. Through kinetic Monte Carlo (KMC) simulations, a large dataset of hydrogen diffusivity was generated using synthetic energy landscapes defined by these statistical parameters. Machine learning symbolic regression (MLSR) was then employed to derive analytical expressions that relate the statistical descriptors to macroscopic diffusivity. To apply the model to real alloys, hydrogen diffusivity is obtained through the MLSR expressions based on energy landscape statistics that were determined using climbing-image nudged elastic band (CI-NEB) calculations with universal machine learning interatomic potentials (uMLIPs). The predictions were validated against molecular dynamics (MD) simulations, showing reasonable agreement. This framework enables fast, scalable prediction of hydrogen diffusion in complex alloys, supporting accelerated materials discovery for hydrogen-related applications.

Contents

Nomenclature

Abbreviations

Abbreviation	Definition
BCC	Body-Centred cubic
CCA	Chemically Complex Alloy
(CI)-NEB	(Climbing-image) Nudged Elastic Band
DB	Diffusion Barrier
DFT	Density Functional Theory
EAM	Embedded Atom Method
FCC	Face-Centred Cubic
GB	Grain Boundary
HE	Hydrogen Embrittlement
HEA	High-Entropy Alloy
KMC	Kinetic Monte Carlo
LAMMPS	Large-scale Atomic/Molecular Massively Parallel Simulator
MD	Molecular Dynamics
MEA	Medium-Entropy Alloy
ML	Machine Learning
MLIP	Machine Learning Interatomic Potential
MLSR	Machine Learning Symbolic Regression
MPEA	Multi-Principal Element Alloy
NN	Neural Network
PES	Potential Energy Surface
RMPEA	Refractory Multi-Principal Element Alloy
RT	Random Trap
SE	Solution Energy
SRO	Short-Range Order
sMLIP	Specialised Machine Learning Interatomic Potential
uMLIP	Universal Machine Learning Interatomic Potential

Symbols

Symbol	Definition	Unit
Latin symbols		
$A(t)$	Activity of a KMC step, sum of all rates	[s^{-1}]
D	Diffusion coefficient	[m^2/s]
D_0	Pre-exponential factor, measure of the diffusion coefficient at infinite temperature	[m^2/s]
D_u	Diffusion coefficient with uniform barrier distribution ($\sigma_s = 0$ & $\sigma_w = 0$)	[m^2/s]
E	Diffusion energy barrier	[eV]
k_b	Boltzmann constant	[eV/K]
Q_{VFT}	Effective activation energy of the VFT model	[eV]
R	Random number between 0 and 1	[-]
r_k	Rate at which transition k occurs	[s^{-1}]
s	Transition-state energy or saddle point	[eV]
T	Temperature	[K]
T_0	Vogel Temperature	[K]
w	Site-energy or well-energy	[eV]
Greek symbols		
μ	Mean value of transition state distribution	[eV]
ν_0	Attempt frequency	[s^{-1}]
σ_s	Standard deviation of transition-state distribution energy	[eV]
σ_w	Standard deviation of well-energy distribution	[eV]

1 Introduction

The release of CO_2 from fossil fuels and the impact it has on the environment is one of the biggest challenges humanity faces today. Levels of CO_2 have been rising since the first continuous measurements, this increase is mainly attributed to the use of fossil fuels such as coal, oil, and gas. Figure 1.1 shows the worldwide CO_2 emission by fuel or industry type since the year 1900, showing the rapid increase of global CO_2 emission due to fossil fuels. Increased levels of CO_2 in the atmosphere lead to an increase in temperature and global climate change [1]. A promising alternative to fossil fuels, that has the ability to be carbon neutral if made from renewable electricity, is hydrogen. The most sustainable way to produce hydrogen is through solar and wind power, excess electricity from these sources can be stored in the form of hydrogen gas and metallic hydrides to level out fluctuations in energy supply and demand [2]. This green hydrogen can then be used in fuel cells to produce electricity or in combustion engines to generate power, with water and heat as the only by-products [3]. Moreover, hydrogen can directly be used in industrial processes, for example through the direct reduction of iron oxides in iron- and steelmaking [4].

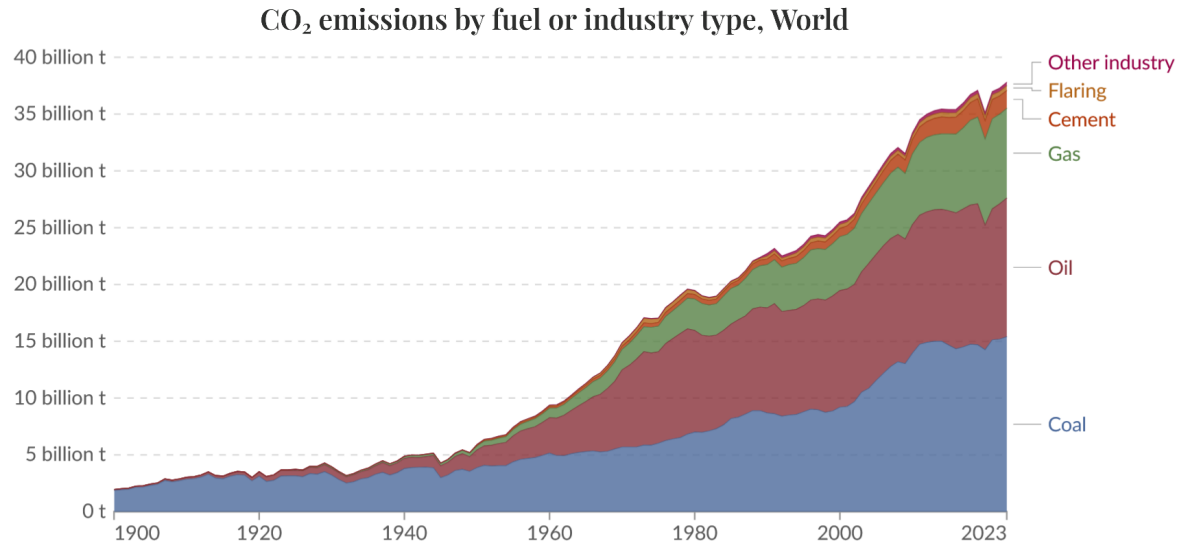


Figure 1.1: Worldwide CO_2 emission by fuel or industry type since the year 1900 [5].

Despite its advantages, hydrogen presents significant challenges in terms of material compati-

bility. Hydrogen is the smallest element. Therefore, it has the ability to rapidly move through metallic materials through interstitial diffusion. Mobile hydrogen atoms can accumulate at defects in the crystal structure, such as grain boundaries, interfaces, and micro-voids. This accumulation can induce decohesion or stress concentrations, eventually leading to crack initiation and propagation, and thus hydrogen embrittlement (HE) [6]. Stronger steels are, in general, more susceptible to HE. High-strength martensitic steels are highly susceptible to HE, while lower-strength ferritic steels are less susceptible [7]. HE can be mitigated through the introduction of hydrogen traps in the material, certain microstructural features such as precipitates and dislocations can act as traps, which reduces the amount of hydrogen that can reach critical sites. In general, face-centred cubic (FCC) materials have lower diffusivity than body-centred cubic (BCC) materials, leading to a higher resistance to HE for the former [8].

In addition, achieving efficient hydrogen storage remains a significant challenge. Hydrogen can be stored in metallic structures through the formation of metal hydrides, where the hydrogen is chemically bonded to the metal atoms. The hydrogen within these metal hydrides can be released through thermolysis when needed [9]. Among the most promising are Mg-based alloys, which have high gravimetric capacities, but suffer from disadvantageous thermodynamics. Specifically, high temperatures are needed to reversibly store H [10]. Intermetallic compounds such as TiFe and $LaNi_5$ also show promise for H storage due to their good reversibility under near ambient conditions. However, these intermetallics have low gravimetric capacities [9, 10].

To address the material compatibility issues associated with HE and H storage, new classes of alloys are being developed. Among the most promising are multi-principal element alloys (MPEAs), which include high-entropy alloys (HEAs) and medium-entropy alloys (MEAs), and are also referred to as chemically complex alloys (CCAs) [11]. Unlike conventional alloys based on a single base element, MPEAs consist of multiple principal elements in near-equimolar compositions. The resulting high configurational entropy can stabilise single-phase solid solutions and suppress the formation of brittle intermetallic compounds [12]. By convention, HEAs contain five or more principal elements, whereas MEAs contain three or four, yielding a vast compositional design space.

The nanostructure of MPEAs is characterised by a highly distorted lattice caused by atomic size mismatch, this can be seen in Figure 1.2 which shows a comparison of the lattices of a pure metal, a dilute alloy, and an HEA with five elements [13]. Each atom experiences a different local environment, resulting in lattice strain and a rough energy landscape. The complex chemical environment causes two main advantageous properties for hydrogen-related application. The first lies in its ability to mitigate HE, the lattice creates well dispersed hydrogen trapping sites on the nanoscale, which can inhibit hydrogen mobility and reduce susceptibility to embrittlement [14]. The good resistance to HE of MPEAs can be seen in Figure 1.3 which compares the stress strain relation of CrMnFeCoNi FCC HEA and two conventional alloys with and without hydrogen charging at room temperature [15]. The second advantage lies in hydrogen storage, MPEAs have demonstrated the ability to store a substantial amount of hydrogen owing to their ability to trap hydrogen due to their rough energy landscape [16]. MPEAs can reach metal-to-hydrogen ratios of 2.5 [17], surpassing that of pure metals [16]. In addition, the vast compositional space of MPEAs allows for the tuning of alloys to obtain H reversibility at advantageous temperatures and pressures by tailoring composition to balance hydride stability with fast absorption–desorption kinetics [10, 18].

In addition to the hydrogen related advantages, MPEAs can exhibit excellent mechanical properties. A subset of MPEAs called refractory MPEAs (RMPEAs), typically including elements such as Cr, Hf, Mo, Nb, Ta, Ti, V, W, and Zr, can maintain high yield strength far

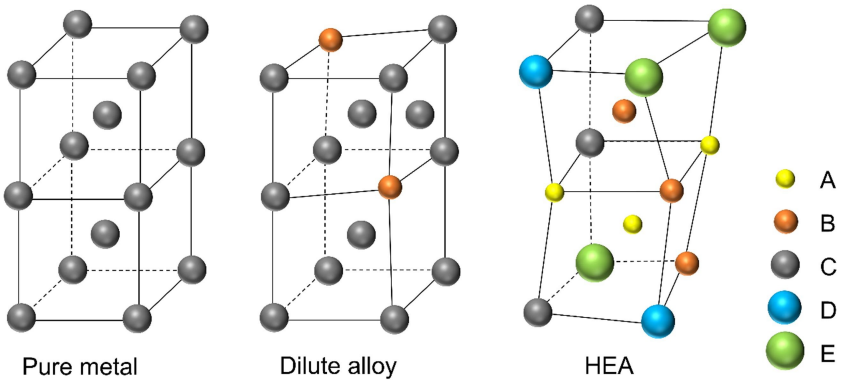


Figure 1.2: Comparison of lattices of a pure metal, a dilute alloy and an HEA [13].

surpassing 1000 °C [19], outperforming nickel-based superalloys [20]. The possible refractory elements offer wide range of properties, such as melting temperature (2128 - 3695 K), density (4.5 - 19.4 g/cm^3), and elastic moduli (68 - 411 GPa). Therefore, RMPEAs can be tailored to achieve specific properties. RMPEAs with low density favour Cr, Nb, V, and Zr, while the highest melting temperature can be achieved with Mo, Nb, Ta, and W [19]. The development of materials that remain strong at higher temperatures is crucial, as an increase in operation temperature leads to an increase in efficiency in power-generation and aerospace technology [21, 22]. Furthermore, good irradiation resistance of these RMPEAs makes them promising candidates for nuclear applications [23].

Despite the promising properties of MPEAs for hydrogen storage and resistance to HE, quantifying hydrogen transport in these chemically complex systems remains challenging. Detecting small atoms like hydrogen at low concentrations is inherently difficult [24], relying on techniques such as atom probe tomography [25] and thermal desorption spectroscopy [26], both of which are slow and limited in throughput. As a result, experimentally mapping the vast compositional space of MPEAs to identify alloys with low hydrogen diffusivity is impractical.

Computational modelling offers a practical route for screening candidate compositions for HE resistance and H storage. In simple metals and dilute alloys, bulk diffusion modelling, whether by kinetic Monte Carlo (KMC) or molecular dynamics (MD), is relatively straightforward, as the energy landscape can often be captured with empirical interatomic potentials and well-defined migration barriers. In MPEAs, however, local chemical fluctuations and lattice distortion create a rough energy landscape, producing a broad spectrum of diffusion barriers [27]. This distribution is an intrinsic material property. Capturing such complexity requires a statistical description of the diffusion landscape.

This thesis develops an efficient, comprehensive computational framework to predict hydrogen diffusivity in BCC MPEAs from statistical descriptors of their diffusion energy landscapes. The aim is to enable rapid, physics-based screening of alloy compositions with favourable hydrogen diffusion properties.

First, large-scale KMC simulations are performed on synthetic barrier landscapes that follow a Gaussian distribution to establish the link between statistical parameters and macroscopic diffusivity. Machine learning symbolic regression (MLSR) is then used to derive analytical expressions for this relationship. These models are applied to real complex BCC alloys by extracting diffusion barrier statistics from climbing-image nudged elastic band (CI-NEB) calculations using universal machine learning interatomic potentials (uMLIPs). Finally, predictions are validated against MD simulations of hydrogen diffusion.

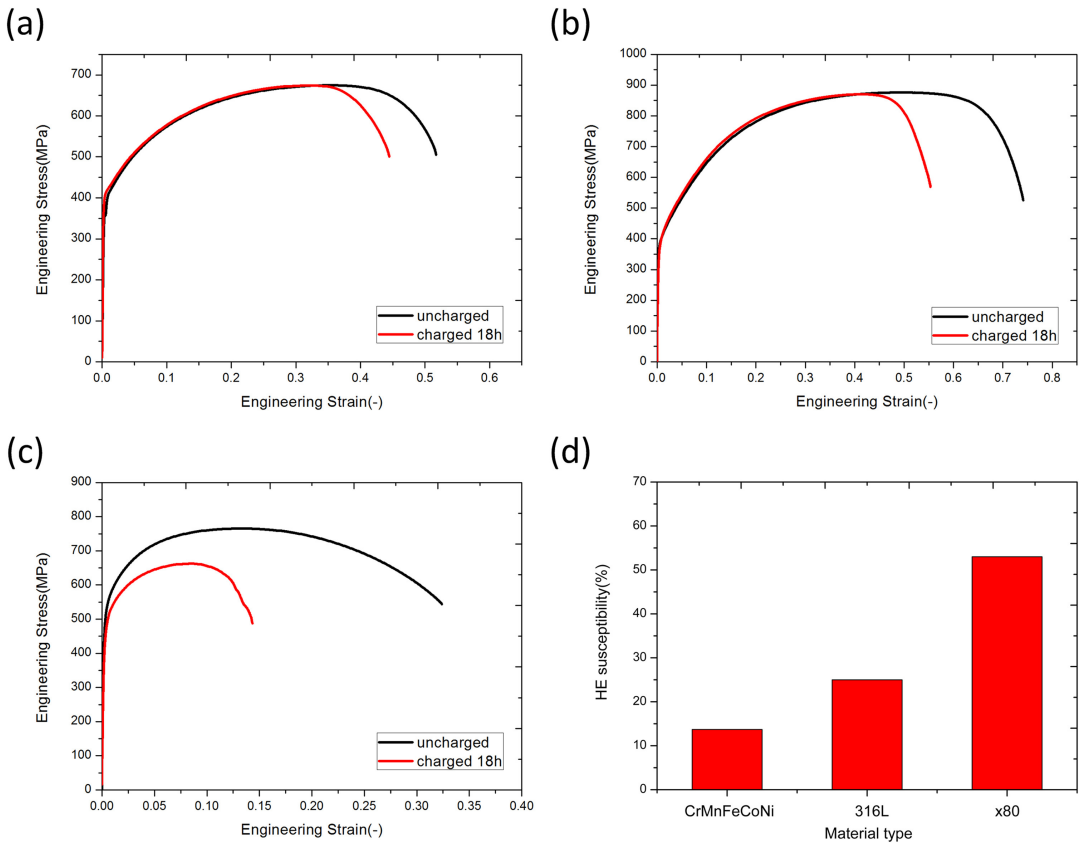


Figure 1.3: Stress-strain relation of (a) CrMnFeCoNi HEA, (b) 316L stainless steel, and (c) X80 PS with and without hydrogen charging at room temperature. (d) The susceptibility of the materials to hydrogen embrittlement [15].

Chapter 2 reviews current knowledge of hydrogen solubility and diffusion in chemically complex alloys, and includes selected studies on pure Fe to illustrate key diffusion modelling methodologies. Chapter 3 details the entire framework, including the KMC, MLSR, uMLIP, and MD methods used. Chapter 4 presents the results from the KMC simulations, the developed models, application to real alloy systems and validation with MD simulations. Chapter 5 discusses the implications and limitations of the approach. Chapter 6 summarises the key findings and Chapter 7 provides recommendations for future research.

2

State of the Art

The complex energy landscapes of MPEAs pose a significant challenge for understanding hydrogen behaviour through experimental methods alone. Therefore, computational studies have become essential in exploring hydrogen solubility and diffusion within these disordered systems.

Hydrogen solubility in MPEAs is a critical starting point of this chapter, as it determines whether hydrogen can be thermodynamically accommodated in the lattice and in which interstitial sites the hydrogen sits. These insights are essential to the modelling of hydrogen diffusion. Studies using density functional theory (DFT) and ML reveal that solubility strongly depends on local chemical environments and lattice distortion, which are very prevalent in MPEAs.

Hydrogen diffusion is inherently influenced by the energy landscape. Traditional assumptions of uniform diffusion barriers fail in disordered system. Computational techniques such as MD and KMC in combination with ML can be used to simulate hydrogen trajectories over extended time and length scales. ML models trained on DFT data have been shown to be successful in predicting energy barriers based on the local chemical environment. These ML models in combination with KMC enable accurate determination of diffusion coefficients and trapping sites while being relatively computational efficient.

2.1. Solubility of hydrogen in disordered systems

Hydrogen solubility in disordered alloys such as MPEAs is heavily influence by the local chemical environment, interstitial site geometry, lattice distortion and hydrogen concentration. Across various DFT studies it is consistently observed that H atoms preferentially sit in the tetrahedral (T) sites rather than octahedral (O) sites in BCC alloys. When initially placed in O-sites, H migrate to nearby T-sites upon relaxation, indicating that O-sites are energetically unstable at low hydrogen concentrations. This behaviour is seen in systems such as TiZrNbHfTa, MoNbTaW, WTaVCr, and TiNbZ, where a significant majority of hydrogen atoms introduced into O-sites spontaneously migrate to T-sites upon relaxation [28–31].

The solution energy (SE) of hydrogen depends not only on the site type but also on the local chemical environment. In TiZrNbHfTa, H SEs were negative for all tested interstitials, confirming that hydrogen accommodation is thermodynamically favourable. Average values were found to be slightly less negative for O-sites, though with greater variability [28]. In WTaVCr, SEs for hydrogen in T-sites ranged from –0.39 to 0.58 eV, depending on the surrounding

atoms. Hydrogen exhibited stronger affinity for sites near V and Cr atoms, while sites rich in W were associated with higher solution energies. This order of affinity (V > Cr > Ta > W) is opposite to what is typically observed in pure metals, illustrating the influence of local lattice strain and site volume in disordered alloys [30].

Geometric factors also play an important role. In MoNbTaW, the H SEs varies strongly with the volume of the interstitial polyhedron, with larger volumes generally corresponding to lower energies, and thus better H accommodation. This dependence on lattice distortion appears more pronounced in BCC alloys than in FCC alloys, where the denser packing leads to reduced variability. These findings emphasize that hydrogen site preference cannot be predicted solely based on chemical composition because local distortion and spatial availability are equally important [29].

The interstitial site preference of H changes with the concentration. At low concentrations, H occupies T-sites and causes little disruption. However, at high H concentrations, phase transformations can occur. In TiZrNbHfTa, increasing the hydrogen-to-metal (H/M) ratio from 0.6 to 2.0 leads to transitions from BCC to body-centred tetragonal (BCT) and eventually FCC at very high H concentrations. These phase transformations are temperature dependent, the TiZrNbHfTa FCC hydride is only stable below 550 K and only forms when H/M reaches 2.0. At lower H concentrations, the BCT phase becomes energetically favoured at lower temperatures [28]. Similar phase transitions are observed in TiNbZr, where occupancy of O-sites correlates with increasing c/a ratio due to anisotropic lattice expansion, indicating a shift from BCC to BCT [31].

The thermodynamics of hydrogen release is also influenced by structural disorder. In TiZrNbHfTa, the decomposition of the hydride occurs over a wide temperature range due to the wide range of local hydrogen environments. A second decomposition peak appears at higher temperatures when hydrogen atoms are bound to vacancies, which act as strong trapping sites. The presence of hydrogen was shown to lower vacancy formation energies, and vacancies, in turn, allowed hydrogen to stabilize in O-sites that would otherwise be energetically unfavourable [28]. However, in WTaVCr, despite the higher vacancy concentration compared to pure W, the trapping capacity is lower (only 3 to 5 hydrogen atoms per vacancy compared to 12 in pure W) suggesting reduced risk of hydrogen-induced damage [30].

To guide the discovery of promising hydrogen-soluble alloys, data-driven approaches have become increasingly important. Nefzi et al. [32] combined DFT and ML to explore over 8,000 equimolar quinary alloy compositions, identifying 568 with favourable thermodynamics for hydrogen storage. Alloys containing Ti, Zr, Hf, Nb, Mg, and V were most often identified as promising materials, with Zr, Ti, and Hf appearing in more than 60% of materials. These elements are known to enhance hydrogen affinity due to their relatively large atomic size and electronic properties [32].

The MLSR method further contributes to understanding hydrogen solubility in disordered systems. Korostelev et al. [33] used interpretable models based on structural, electronic, and vibrational descriptors to predict hydrogen absorption energy across HEAs and intermetallics. The most influential factors were d-band centre, interstitial pore size, and phonon frequency. These models confirmed that lower d-band centres and larger local volumes lead to stronger hydrogen absorption. Unlike conventional black-box machine learning, these models allow physical interpretation and are more transferable across alloy systems [33].

In summary, hydrogen solubility in disordered systems is governed by a complex interplay of interstitial site geometry, local chemical composition, lattice distortion, and concentration. In

BCC alloys, T-sites remain energetically preferred under most conditions, but at high hydrogen concentrations, O-sites begin to stabilise and drive phase transitions. Machine learning and MLSR methods offer new ways to predict favourable compositions and understand solubility mechanisms.

2.2. Diffusion in disordered systems

For the relatively simple BCC α-Fe system, Hasan et al. [34] used MD simulations to investigate the effect of H concentration (0.01-5 at.%) and temperature (350-900 K) on diffusivity. The interatomic interactions were described by an Fe-H Embedded Atom Method (EAM) potential developed by Ramasubramaniam et al. [35] and Mendelev et al. [36]. It was shown that at high concentrations, the diffusivity of H is significantly reduced, this effect is more pronounced at low temperatures. Consequently, it is proven that H diffusion can be influenced by other H interstitials. Moreover, it was found that at higher concentrations (1-5%) and below certain temperatures, H clusters form and the host Fe structure changes to either FCC or amorphous. This leads to non-Arrhenius behaviour at high concentrations.

Using the same BCC Fe system with the same EAM potential, Zhou et al. [37] employed MD simulations to study the effect of nanometre sized grains on the H diffusion coefficient. It was found that hydrogen diffusivity decreases significantly as grain size is reduced. This reduction appears primarily due to trapping effects at grain boundaries (GBs) and triple junctions. While GBs already serve as strong hydrogen traps, triple junctions become increasingly dominant at smaller grain sizes and lower temperatures. In contrast to FCC systems such as Ni, where GBs can act as fast diffusion paths, GBs in BCC Fe are strong barriers to hydrogen migrations.

To investigate the effect of tilt GBs and open surfaces on H diffusion in both BCC α-Fe and FCC γ-Fe of H diffusion, Smirnova and Starikov [38] used MD simulations. These simulations utilised a Fe-Cr-H potential developed by Starikov et al. [39] to model the interatomic interactions. For the BCC GBs there is a significant drop in diffusion coefficient compared to the bulk BCC, this drop is especially evident at low temperatures. Consequently, the GBs in BCC do not provide fast diffusion channels, while simultaneously acting as strong H trapping sites. Conversely, for the FCC GBs there is an increase in diffusion coefficient compared to the bulk FCC. Furthermore, H diffusion on surfaces was studied. For the BCC surfaces, the (100) surface showed a slight increase in diffusivity, while the (110) surface showed a slight decrease in diffusivity at high temperatures, but a greater decrease at room temperature. For the FCC surfaces, both the (100) and (110) showed diffusion coefficients order of magnitudes higher that the bulk FCC.

To capture the diversity of GB structures in polycrystalline materials, Sun et al. [40] developed a combined MD and ML approach. High-throughput MD simulations, using the EAM potential of Ramasubramaniam et al. [35], generated hydrogen diffusion data for a wide range of BCC α-Fe GB configurations, and a graph neural network was trained to predict diffusion coefficients from atomic configurations. The model showed that any structural disorder leads to a decrease in H diffusivity. In some configurations, diffusivity dropped to less than 0.1% of the bulk value. A multi-scale averaging scheme was developed, where local predictions across a material volume were used to estimate macroscopic H diffusivity. This approach proved capable of predicting effective diffusion in micrometre-scale volumes using only nanometre-scale structural data, and demonstrated how grain-level microstructure can dominate macroscopic H transport.

Time and length scale limitations of traditional MD have driven the development of more efficient long-timescale methods. Tang et al. [41] introduced a reinforcement learning (RL)

framework for simulating H diffusion in FCC Cu, Ni, and CrCoNi MEAs. Traditional KMC simulations require an event table with energy minima and transition pathways, correctly determining this event table is no trivial task, especially for complex alloys. To conduct KMC simulations without an events table off-lattice KMC was developed. In off-lattice KMC the atoms can move continuously in space, and instead uses an algorithm to find saddle points and diffusion pathways on-the-fly. Another method is temperature accelerated dynamics (TAD), where the transition pathways are explored using high temperature MD simulations. In both TAD and off-lattice KMC, the transition pathway is explored by random sampling. The high configurational space of complex alloys requires a large amount of random sampling, which limits the system size and timescale. Therefore, Tang et al. developed a RL method to guide the transition pathway sampling on chemically complex potential energy surface (PES). Not all nearby saddle points have to be sampled, instead a parametrized NN model is used to predict the direction of atomic motion with the high probability pathway, resulting in a reduced "transition energy landscape" (TEL). Their simulations reproduced experimental results in pure metals and extended the method to chemically complex systems, achieving several orders-of-magnitude efficiency improvement over off-lattice KMC.

In chemically complex alloys, such as MoNbTaW RMPEAs, Shuang et al. [42] developed a framework combining ML force fields (MLFF), neural network-driven KMC (NN-KMC), and symbolic regression. The MLFF, trained on DFT data, was used to compute hydrogen diffusion barriers across 287 alloy compositions. Diffusion coefficients were predicted using NN-KMC, and the dependence on composition, short-range order (SRO), and temperature was analysed. They found that elements with negative hydrogen solution energies, such as Nb and Ta, act as strong traps and dominate the diffusional properties of H. Alloys with large composition gradients between H-favouring (Nb, Ta) and H-repelling (W, Mo) elements exhibited lower diffusion coefficients due to deep trapping sites. This was especially apparent in non-equimolar alloys like $Ta_{20}W_{80}$, where small Ta clusters embedded in a W matrix created severe diffusion bottlenecks. Interestingly, the random mixing of elements in the equimolar MoNbTaW alloy leads to a relatively uniform (but roughened) energy landscape, each H repelling element is accompanied by an H favouring element, this prevents the formation of deep trapping sites such of which can be seen in the $Ta_{20}W_{80}$ alloy. Moreover, They found that the presence of H favouring elements significantly governs hydrogen diffusivity. For the MoNbTaW alloys these are Nb and Ta, in other systems BCC transition elements such as V, Ti, Zr and Hf have similar H favouring deep trapping effects due to their low solution energies [42].

Furthermore, the effect of chemical short-range order (SRO) on the diffusivity of 287 alloy compositions of MoNbTaW was investigated by Shuang et al. [42]. They found that in the $Ta_{20}W_{80}$ alloy, through SRO, the introduction of a repulsive action of the Ta-Ta atoms and an attraction between Ta-W atoms leads to an altered atomic arrangement, which increases the H SE near the Ta atoms. This reduces the effectiveness of the deep trapping sites, and in turn increases the H diffusivity. This effect is also noticed when considering low concentration of Nb, the other H-favouring element in the MoNbTaW alloys. In fact, the concentration of Nb and Ta is a measure of the effect that SRO would have on the diffusivity. The diffusivities of systems including SRO deviate from those of RSS when the composition of Nb and Ta is low. It was found that, for the equimolar MoNbTaW alloy, SRO has negligible effect on the diffusivity [42]. Furthermore, the effects of SRO diminish upon increasing temperature. This is due to the increased thermal energy that overcomes tendency to form an ordered system. More specifically, above 1000 K the MoNbTaW alloy tends to a RSS [43].

Zhou et al. [14] used ML and KMC to study non-equimolar FCC FeCoNiCrMn HEAs. They trained neural networks using DFT-derived solution energies and descriptors of local chemical

environments to predict hydrogen diffusivity. By systematically optimizing alloy compositions, they found that increasing Co and Mn content tends to reduce hydrogen diffusivity, whereas Fe and Ni promote faster diffusion. Their polynomial regression model captured these trends with high accuracy ($R^2 = 0.86$), offering a quantitative link between elemental ratios and diffusivity. This study demonstrated that even within a single compositional family, diffusion can be tuned predictably using data-driven models.

Although this thesis focuses on H diffusion, many of the computational techniques used, such as KMC and CI-NEB in combination with ML are used in the context of vacancy diffusion in MPEAs. In vacancy systems, the local atomic environment changes dynamically with each atomic swap, requiring real-time prediction of migration barriers using ML models trained on CI-NEB data [27, 44–46]. In contrast, hydrogen diffusion involves interstitial migration through a static lattice, allowing precomputed diffusion barriers to be reused during KMC simulations. Despite this difference, both systems face similar challenges in descriptor design, managing energy landscape roughness, and capturing the effects of SRO. Studies on vacancy diffusion in FCC MPEAs, such as CrCoNi and NiCoCr, have shown that SRO can strongly reduce diffusivity and introduce non-linear time dependencies in mean squared displacement [47, 48].

In summary, the studies reveal that hydrogen diffusion in disordered systems is sensitive to structural and chemical complexities. In BCC α-Fe, factors such as H concentration, temperature, grain size, and grain boundary character significantly influence diffusivity, with structural disorder and trapping effects at interfaces reducing mobility. While BCC grain boundaries and triple junctions act as strong traps, FCC systems exhibit enhanced surface and grain boundary diffusion. Recent advances, including high-throughput MD, ML, and RL-based KMC methods, have enabled efficient exploration of diffusion pathways in chemically complex alloys. In MPEAs, hydrogen diffusivity in bulk materials is strongly influenced by local chemical environments, particularly the presence of H-favouring elements like Nb and Ta. Moreover, structural features such as grain boundaries strongly influence H diffusion in BCC systems, but current studies are restricted to pure iron. These structural and composition-driven deep trapping effects further underline the impact of atomic-scale features on diffusion.

2.3. Research focus and motivation

Hydrogen diffusion in chemically complex alloys, such as MPEAs, is a relatively unexplored area. While numerous studies have addressed H diffusion in pure metals or simpler binary alloys, the interplay between chemical disorder and diffusion in MPEAs remains less understood. In particular, how the statistical properties of the H diffusion barrier distributions influence macroscopic diffusivity is not yet well established.

Existing state of the art methods, such as DFT, MD, and KMC, in combination with ML approaches, provide accurate insights but are computationally expensive, especially when dealing with the large configuration space of disordered systems. These methods typically require large datasets of expensive DFT calculations or long simulation times to converge meaningful diffusion coefficients, making them impractical for high-throughput alloy screening or rapid analysis.

This thesis addresses that limitation by developing a comprehensive computational framework for H diffusion in BCC MPEAs. The core idea is to link the diffusion coefficient to the statistical properties of the diffusion barrier distribution, specifically, the mean saddle point energy (μ), the standard deviation of the saddle point energies (σ_s), and the standard deviation of the well-energies (σ_w). Instead of determining the diffusivity directly through costly simulations,

this approach enables rapid estimation based on these barrier statistics.

To achieve this, a MLSR approach is applied to learn expressions of the form $D = f(\mu_s, \sigma_w, \sigma_s)$, trained on diffusion data generated from KMC simulations where the energy landscape is described through these three parameters. Both data-driven and physics-informed approaches will be explored. Moreover, by employing uMLIPs in combination with CI-NEB calculations, the statistical properties of diffusion barriers can be extracted efficiently for varying complex BCC systems and used as input for the model. This bypasses the need for full-scale KMC, MD, or DFT-based simulations, offering a fast method to estimate H diffusivity in disordered systems.

Ultimately, this research aims to provide a comprehensive framework that captures how atomic-scale disorder affects macroscopic H diffusion in BCC alloys, enabling both fundamental understanding and practical alloy evaluation. This framework will be evaluated using H diffusion coefficients obtained through MD simulations of H trajectories. In future work, this framework can be extended to FCC MPEAs.

3

Methodology

The computational framework of this thesis comprises three parts: (1) KMC simulations; (2) building analytical expressions through MLSR; (3) application of expressions through uMLIP and CI-NEB calculations for new materials and validation through MD simulations. An overview of the computational framework is presented in Figure 3.1.

First, KMC (Section 3.1) is employed to create a dataset of H diffusion coefficients based on a reference system that approximates the complex energy landscape of BCC MPEAs (Section 3.1.3). This approximation enables the creation of model materials, whose diffusion barrier distributions are characterised by three input parameters: μ, σ_s, and σ_w. Second, MLSR (Section 3.2) is employed to construct analytical expressions from the KMC dataset. Both purely data-driven and a physics-informed approaches are explored. Third, these analytical expressions are applied to several chemically complex BCC alloys. CI-NEB calculations (Section 3.4), performed using an uMLIP (Section 3.3), provide the input parameters required by the MLSR expressions. This enables prediction of H diffusion behaviour in these alloys without relying on DFT, KMC, MD, or additional ML model training. The predictions are validated against diffusivity data from MD simulations (Section 3.5), using the same uMLIP as in the CI-NEB calculations.

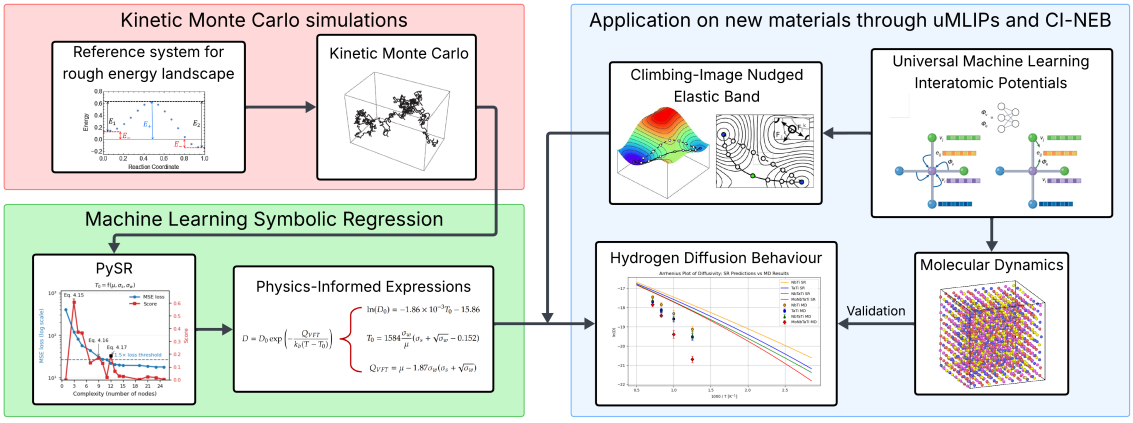


Figure 3.1: Overview of computational framework. Images from ref. [49–52].

3.1. Kinetic Monte Carlo method

KMC is a computational method in which the rates of certain events, such as diffusional jumps, can be modelled. These rates are related to the probability that an event will occur. A system in which diffusion occurs is stochastic in nature, with discrete steps in between events, i.e. the system randomly changes configurations, with some average time between the events. If the rates at which the configuration changes is known, the KMC method can be applied to model the system [50].

3.1.1. The KMC algorithm

Typically in a system more than one event can occur, the probability that a certain event takes place is proportional to its rate over all other rates of the other possible events. Consider a system in which N species can undergo M transitions with each transition at a rate r. The probability of a species undergoing a transition k is thus proportional to r_k. The total activity can thus be defined by the sum over all rates, this can be seen in Eq. 3.1 [50].

$$A(t) = \sum_{k=1}^{M} r_k(t) \tag{3.1}$$

The probability that a certain species l undergoes a transition is thus its specific rate over the activity, as can be seen in Eq. 3.2 [50].

$$P(l) = \frac{r_l}{A(t)} \tag{3.2}$$

In order to randomly choose an event to occur, all probabilities (r_l/A) can be put consecutively on a line with the length equal to its probability. The length of the line is taken as 1, which corresponds to the activity. A random number between 0 and 1 is chosen and depending on where the chosen number sits on the line, a certain event takes place. An example of this can be seen in Figure 3.2, in which four possible diffusional jumps can be made. In other words, a certain event m takes place if it satisfies Eq. 3.3, where R_1 is a random number between 0 and 1 [50].

$$\frac{\sum_{k=1}^{m-1} r_k}{A} < R_1 < \frac{\sum_{k=1}^{m} r_k}{A}, \tag{3.3}$$

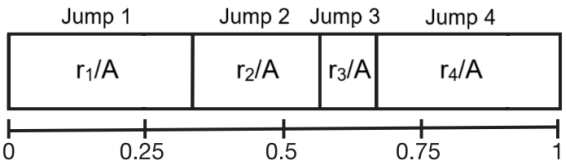


Figure 3.2: Picking an event in KMC using a random number between 0 and 1 [50].

The time step Δt of a transition in the KMC simulation is related to the activity. The activity, as defined earlier, depends on the transition rates of the possible events for a given configuration. When the transition rates are high, the activity is large, and the corresponding time step is small. Thus, the time step represents the elapsed simulation time until the next event occurs [50].

The time step is derived from the probability that no event occurs within a given time interval t, denoted as $g(t)$. Since $g(t)$ cannot be known beforehand, it is replaced by a stochastic variable with the same average properties. As this probability is restricted to values between 0 and 1, it can be represented by a random number R_2 drawn from a uniform distribution on [0,1]. Substituting R_2 for $g(t)$ yields Eq. 3.4, which links the time step directly to the activity and the stochastic nature of event selection [50].

$$\Delta t = -\frac{1}{A(t)} \ln R_2 \tag{3.4}$$

The KMC method consists of the following algorithm, called the Bortz-Kalos-Lebowitz algorithm or the N-fold way, as described by LeSar [50] and first proposed in 1975 by Bortz et al. [53]:

1. At time t, determine all possible events that can occur and add up all the rates of the events to determine the activity (Equation 3.1).
2. Calculate the probability of each event, using Equation 3.2.
3. Create a list of events, with the weight being their probability, such as in Figure 3.2.
4. Pick a random number $0 < R_1 < 1$ to determine which event takes place.
5. Enact the chosen event, which changes the configuration of the system.
6. Advance the time by Δt, which is calculated using Equation 3.4.
7. Repeat the steps.

The KMC method is not suitable for all systems because it relies on knowing all possible events of the system and the rates at which these occur. If one of the more critical events is missing, the outcome of the simulation is most likely not a good representation of the system [50].

3.1.2. KMC for modelling of atomic diffusion

The KMC method is very suitable to model atomic diffusion. The rates of atomic jumps can be calculated through Equation 3.5, in which ν_0 is the attempt frequency, ΔE_{ij} the diffusion barrier between site i and j, T the temperature, and k_b the Boltzmann constant. The attempt frequency is in the order of $10^{13} s^{-1}$. A previous study determined the attempt frequency for hydrogen in BCC metals to be $1.5 \times 10^{13}\ s^{-1}$ by minimizing the difference of diffusion coefficients between KMC and MD simulations for pure Mo, Nb, Ta, and W [42].

$$r_i = \nu_0 \exp(-\frac{\Delta E_{ij}}{k_b T}) \tag{3.5}$$

Pure geometrically, BCC lattices have 12 T-sites in their unit cell, thus 6 T-sites per atom. However, in a random HEA, not all sites are reachable due to unfavourable solution energy. Hydrogen atoms preferentially sit in interstitial T-sites in BCC lattices (Section 2.1). Consequently, H diffusion occurs through jumps between T-sites. Figure 3.3a shows the possible migration paths for an interstitial hydrogen atom in a BCC system. The hydrogen atom has two types of paths, a direct path ($T_1 - T_2$) and an indirect path ($T_1 - O - T_3$) through an O-site. The indirect path is observed to be straight, while the direct path curves towards the O-site. Figure 3.3b shows the diffusion barriers associated with each type of path, the $T_1 - T_2$ direct path has a lower diffusion barrier than the indirect path [54]. Hydrogen in a T site has four nearest neighbours, thus each

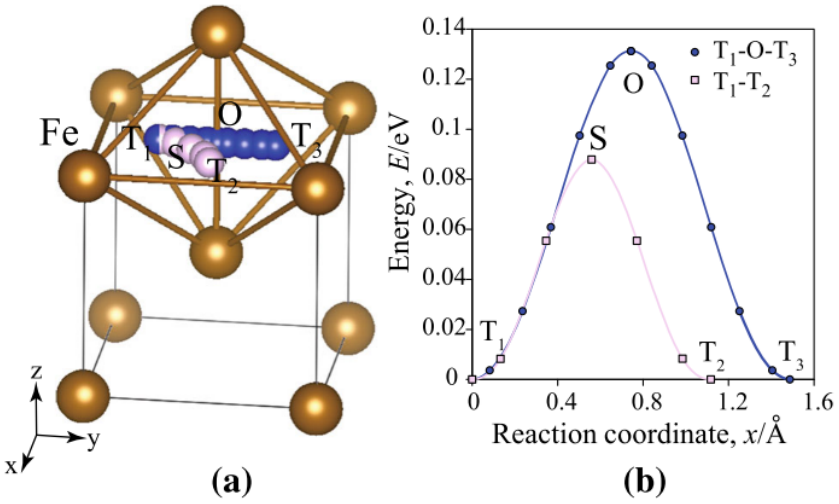


Figure 3.3: (a) Migration paths of hydrogen in a BCC lattice. O, T, and S represent the octrahedral site, tetrahedral site and the saddle point. (b) Diffusional barriers for $T_1 - T_2$ direct and $T_1 - O - T_3$ indirect paths [54].

hydrogen atom has four possible direct paths. The rates for each possible diffusion path have to be calculated, after which the KMC algorithm as described in Section 3.1.1 can be applied.

The number of needed steps can either be determined by setting the KMC simulation to stop after a certain time is reached or after a certain mean square displacement (MSD) is reached. In alloys with a large number of trapping sites, hydrogen diffusion is non-uniform and the residence time is undetermined. Thus, using a distance-based criterium is preferred to get reliable data, with an MSD criterion of 1×10^5 Å^2 having been used in a previous study [42]. Two common approaches to obtain diffusivity through the hydrogen trajectory is to either: (1) conduct multiple independent diffusion simulations and taking an averaged MSD through: MSD $= \left\langle \left| \vec{x}(t) - \vec{x_0} \right|^2 \right\rangle$, where $\vec{x}(t)$ is the position of the H atom at time t and $\vec{x_0}$ is its initial position [50], or (2) perform large-scale diffusion simulations and taking a converged MSD. In both cases, Eq. 3.6 is used to obtain diffusivity through MSD for n-dimension lattices.

For H diffusion in chemically complex alloys with strong trapping sites, where diffusion is highly non-uniform, the two commonly used approaches are not suitable. The MSD–time plot is often curved and noisy rather than cleanly linear. This limits the use of Eq. 3.6, which assumes a linear MSD–time relation. A previous study of H diffusion in pure aluminium using MD simulations used a fixed time interval (similar to the residence time), in which the distance over that time is recorded and cumulatively summed to obtain a linear MSD versus time curve [55]. For MPEAs, since the residence time is undetermined and repetitive jumping within a trapping site is not effective diffusion, instead it has been proposed to use an effective trajectory, which represents the jumps between (and thus the size of) effective trapping sites [42]. It was found that a small jump based-criterion (d_c) overestimates the diffusivity, while for $d_c \geq 20$ Å a plateau is reached in the D versus d_c plot [42]. Therefore, this work uses a jump based-criterion of 20 Å to determine the effective diffusion trajectory.

$$D = \frac{MSD}{2nt} \tag{3.6}$$

3.1.3. Reference system for rough energy landscapes

The rough energy landscape of diffusion processes in complex alloys can be decomposed into two simplified models that highlight different types of disorder. These two simplified models can be seen in Figure 3.4, where the the (c) general mixed model is divided into (a) random trap (RT), and (b) random barrier (RB) models [49]. Thomas and Patala used these models in the context of (substitutional) vacancy diffusion in MPEAs. For interstitial diffusion of hydrogen, the same reference system can be used.

In the RT model the saddle point (transition-state) energies (s) are of equal magnitude, while the site-energies (w) are considered to have a Gaussian distribution with a standard deviation of σ_w and a mean value of μ_w. An important aspect of the RT model is that every transition out of a certain state has the same diffusion barrier E. This model was compared to a situation with an uniform distribution of barriers with the same mean, this led to the analytical solution $D/D_u = \exp(-(\sigma_w/k_BT)^2)$ where D_u is the diffusion coefficient of a system with uniform barriers. Using this expression, the reduction of the diffusion coefficient can be explained by the trapping of the diffusing species by low energy sites due to the addition of the σ_w standard deviation term [49].

On the other hand, the RB model describes an energy landscape where the site-energies (w) have a fixed value and the transition-state energies (s) are randomly distributed with a standard deviation of σ_s and a mean value of μ_s. An important aspect of the RB model is the symmetry of diffusion barriers, i.e. if a transition from i to j is considered, then the associated diffusion barrier of i to j is equal to that of j to i, thus $E_{i\to j} = E_{j\to i}$ [49].

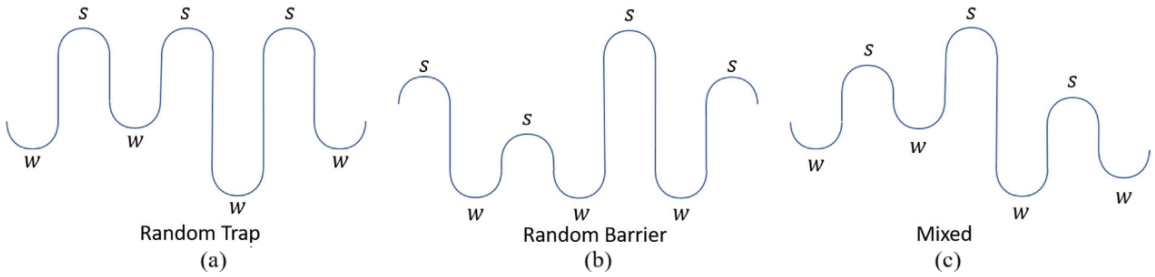


Figure 3.4: Energy landscapes with (a) random trap, (b) random barrier, and (c) mixed models. s and w show the transition-state (saddle point) energies and site-energies, respectively [49].

Figure 3.5a shows the energy associated to a jump from an initial state (left) to a final state (right), computed from one NEB calculation for vacancy migration in a CoNiFeCrMn alloy. In a rough energy landscape there is no well-defined reference energy, both w_i and w_f can be chosen as the reference. Thus, the symmetric and anti-symmetric components of a certain jump can be considered. As can be seen in Figure 3.5a, E_+ is the symmetric barrier component and E_- is the anti-symmetric barrier component. The energy reference is taken as the mean of the initial and final states ($E_+ = (E_{A\to B} + E_{B\to A})/2$ and $E_- = (E_{A\to B} - E_{B\to A})/2$). With these definitions, E_+ is the contribution of RB to the mixed model and E_- (the site-energy difference) is the contribution of RT to the mixed model. Figure 3.5b shows the distribution of E_+ and E_- of the CoNiFeCrMn alloy through 2971 NEB calculations [49].

Using the definition as shown in Figure 3.5a it can be said that $s = E_+$ and $w_i = -w_j = E_-$. Thus the diffusion barrier ΔE_{ij} can be rewritten as $s_{ij} - w_i$, where s_{ij} is the diffusion barrier between site i and j, and w_i is the site-energy of site i when considering a jump from i to j [49]. With the same reasoning, a backward jump from j to i can be written as $\Delta_{E_{ji}} = s_{ij} - (-w_j)$,

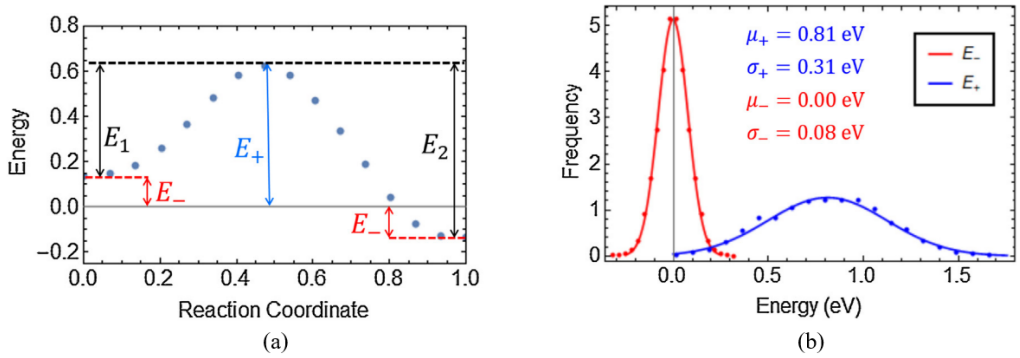


Figure 3.5: (a) Definition of symmetric (E_+) and anti-symmetric (E_-) barrier components. (b) Distribution of E_+ and E_- computed from 2971 NEB calculations of the CoNiFeCrMn MPEA [49].

considering that $s_{ij} = s_{ji}$ when s is defined by E_+, and $w_i = -w_j$ when w is defined by E_-. Equation 3.5 can thus be rewritten as:

$$r_i = \nu_0 \exp(-\frac{s_{ij} - w_i}{k_b T}) \tag{3.7}$$

Considering the definition of the symmetric and anti-symmetric components of the diffusion barrier, a division has to be made between forward and backward hops, as for a forward hop $\Delta E_{ij} = E_+ - E_- = s_{ij} - w_i$, while for the backward hop $\Delta E_{ji} = E_+ + E_- = s_{ij} + w_i$. The choice for the forward hop to have a negative sign in front of w might seem significant, but in reality it does not matter if w is subtracted or added for the forward or backward hop. Considering that the symmetric and anti-symmetric barriers are defined so that $\mu_w = 0$, meaning there is equal probability that w is positive or negative. For H diffusion in BCC alloys, each T-site has four possible diffusional paths: three forward jumps defined as $\Delta E = s - w$, and one backward jump to the previous site defined as $\Delta E = s + w$, where w has equal probability of being positive or negative.

Using this reference system, a model system can be created where all diffusion barriers of an BCC alloy can be assigned through random sampling from a Gaussian distribution of s and w using the parameters μ, σ_s, and σ_w.

3.1.4. Model systems of H in BCC RMPEA

Using the reference system as described in Section 3.1.3, model systems can be created that describe the different H diffusion barrier distributions based on the parameters μ, σ_s, and σ_w. In these model systems, the H diffusion barrier distribution is approximated to be a Gaussian distribution. This approximation for both s and w can be seen in Figure 3.6, which compares the true barriers of equimolar MoNbTaW as determined by ref. [42] with the Gaussian approximation, both the true and approximate distributions have the same μ, σ_s, and σ_w values. It can be seen that for the saddle point distribution, the Gaussian approximation overestimates the number of low energy diffusion barriers. Using the approximation, a large number of diffusion coefficients can be determined for varying BCC materials with Gaussian-like H diffusion barrier distributions. The H diffusion coefficient can be determined using KMC method as described in Section 3.1.1 and 3.1.2.

When the diffusion barrier distribution is approximated as Gaussian using the parameters μ, σ_s, and σ_w, a small number of negative diffusion barriers is inevitable. Although these barriers

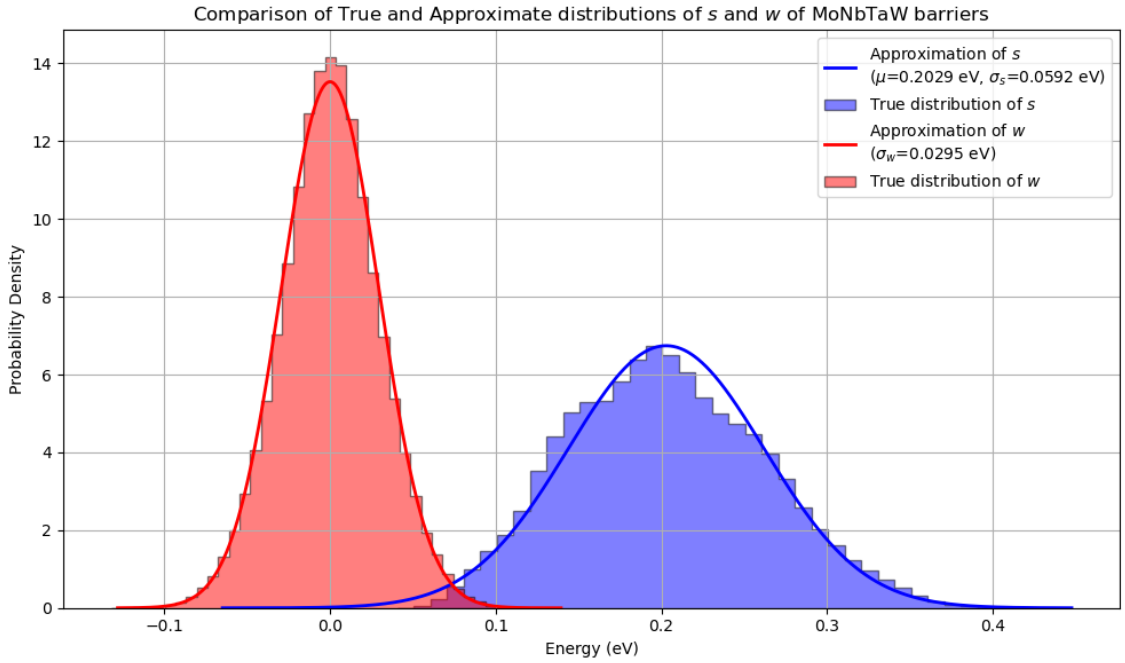


Figure 3.6: True barrier distribution of equimolar MoNbTaW from ref. [42] compared to the Gaussian approximation, both the true and approximate distributions have the same μ, σ_s, and σ_w values. For s (blue), the Gaussian approximation places larger probability in the low energy tail relative to the true distribution.

are non-physical, they do not affect the KMC algorithm, as negative values can be interpreted as zero barriers, representing unstable sites. To ensure physical relevance, the selected range of μ, σ_s, and σ_w combinations was constrained such that, even in the worst-case scenario, no more than 1.5% of the barriers are negative. These ranges are listed in Table 3.1. The steps of μ, σ_s, and σ_w are taken as 0.01 eV, leading to a total of 352 distinct model materials. This selection was informed by an analysis of the μ, σ_s, and σ_w values from 287 binary, ternary, and quaternary compositions of the MoNbTaW alloy reported in a previous study [42]. In contrast, another earlier study employed a looser criterion, allowing up to 5% negative barriers [49], which was found to be too broad for accurately modelling H diffusion in this work.

In total, 352 combinations of barrier distributions (Table 3.1, with steps of 0.01 eV for all parameters) and thus 352 different model materials were simulated across 11 different temperatures ranging from 350 to 2000 K. The minimum temperature of 350 K was chosen, as below this nuclear quantum effects play a role, which significantly decreases the height of the diffusion barrier and changes the diffusion pathway [56]. This effect cannot be captured by the KMC algorithm. Moreover, for combinations with large values of σ_s and σ_w, difficulties arise in obtaining a reliable H diffusion coefficient below 350 K due to extensive H trapping.

The model materials are created by considering two Gaussian distributions, of the saddle points and well-energies, based on the parameters μ, σ_s, and σ_w. Each of the 192,000 unique tetrahedral-tetrahedral diffusion paths in a $20 \times 20 \times 20$ BCC supercell are assigned a saddle point and a well-energy based on the random sampling of the two Gaussian distributions. The forward and backward diffusion barriers are defined as $\Delta E = s - w$ and $\Delta E = s + w$, respectively.

KMC is a stochastic process; therefore, each combination of model material and temperature was simulated five times to obtain a reliable average. In total, 19,360 KMC simulations were performed on the DelftBlue supercomputer provided by the Delft High Performance

Computing Centre [57]. A stopping criterion of MSD = 10^5 Å^2 is used to ensure that a large enough sample of jumps is considered.

Table 3.1: Range of σ_s and σ_w for different μ values to obtain distributions with less than 1.5% negative barriers. The steps of μ, σ_s, and σ_w are taken as 0.01 eV, leading to 352 distinct combinations of energy landscape distributions.

Mean barrier μ (eV)	**σ_s range (eV)**	**σ_w range (eV)**
0.16, 0.17	$0 \leq \sigma_s \leq 0.05$	$0 \leq \sigma_w \leq 0.05$
0.18, 0.19	$0 \leq \sigma_s \leq 0.05$	$0 \leq \sigma_w \leq 0.06$
0.20 – 0.23	$0 \leq \sigma_s \leq 0.06$	$0 \leq \sigma_w \leq 0.06$

3.2. Machine learning symbolic regression

MLSR is a powerful tool to create human-interpretable mathematical functions to describe large datasets [58]. In traditional regression methods, parameter optimisation begins with a predefined model as the algorithm's starting point. For example, the linear regression model is based on the assumption that the dependent variables and regression is linear. However, in MLSR, no prior assumption of the specific form of the function is needed. Instead, a mathematical expression space with candidate function building blocks are given. These building blocks include: mathematical operators, state variables, constants, and analytical functions. MLSR then searches through these building blocks to find the most optimal solution, thus both model structure and model parameters are optimized [59].

MLSR is a type of genetic programming, in which either a mutation, crossover, simplification, or optimization of constants of the parent function can occur to evolve the function [58]. The evolved expressions are then evaluated on fitness through either the sum of the squares (SSE) or the MSE [60]. The better the score, the larger probability that the function will be selected for reproduction. This selection rule agrees with the "survival-of-the-fittest" rule, as good features are more likely to be inherited into subsequent generations, which eventually converges to an optimal solution [59]. Figure 3.7a shows a mutation operation applied to an expression tree, this operator takes one parent structure and randomly substitutes a subtree with a randomly generated structure. Figure 3.7b shows the crossover operation, this operator takes two winners of the selection process as parents and creates a randomly generated offspring based on subtrees from both parent functions [58, 59]. Crossover is usually the dominant operation. Mutations are more aggressive than the crossover operation, since it adds a degree of randomness to the system. However, it is essential to have a finite chance of mutation, as it avoids being trapped in local minima [59].

PySR is a powerful open-source library for practical MLSR, developed by Cranmer [58]. Figure 3.8 shows the inner loop of PySR. A population of expressions is randomly subsampled. A tournament is performed among this subsample by determining the fitness of the functions, the winner of the tournament is selected through some probability, in which the fittest functions have the highest probability of being chosen for breeding. The breeding is done through either: mutation, crossover, simplification, or optimisation, the original parent function is copied and remains in the population. The new function then replaces the oldest of the population, PySR replaces the oldest instead of the weakest (as is often done in other genetic algorithms), as it is found to prevent early convergences, which would mean the population specialises too quickly and thus gets stuck in local minima. PySR is computationally efficient, as it allows for the parallelisation of the tournaments of the subsamples. Furthermore, PySR includes a simulated annealing step in the selection of the winner of the tournament. The probability for rejection is

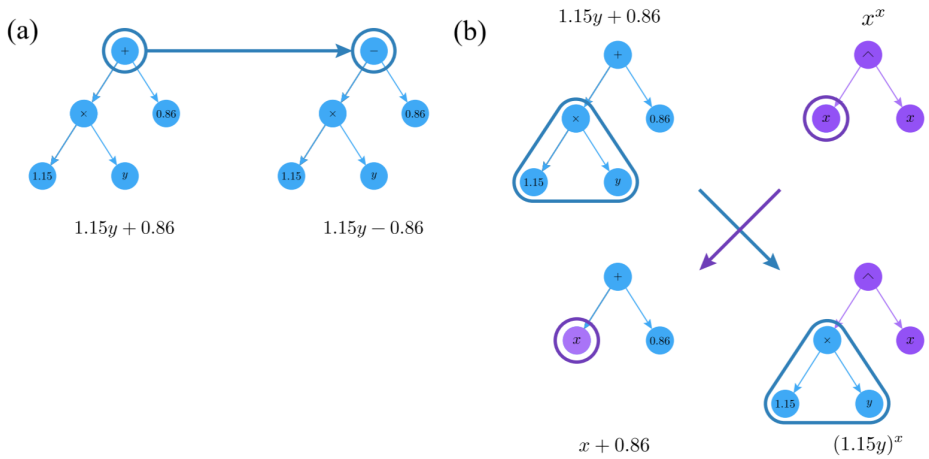


Figure 3.7: (a) Mutation operation applied to an expression, (b) Crossover operation between two expression trees [58].

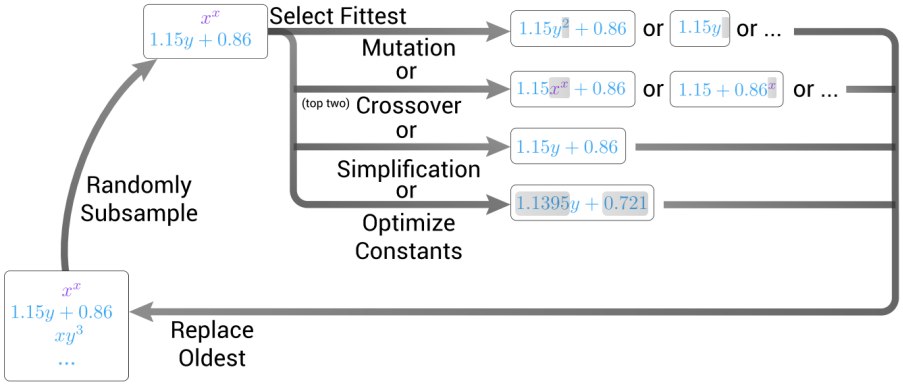


Figure 3.8: The inner loop of PySR [58].

calculated as $p = \exp(\frac{L_F - L_E}{\alpha T})$, in which L_F and L_E are the fitness of the mutated and original function, respectively, α a hyper-parameter, and the temperature $T \in [0, 1]$. This modification allows for the evolution to alternate between high temperature and low temperature phases, with high temperature increasing the diversity and low temperature narrowing in on the fittest individuals. The α-parameter controls the scale of the temperature, with $\alpha \to \infty$ being a regular tournament selection and $\alpha \to 0$ rejecting all mutations with lower fitness than the original function. The simulated annealing significantly speeds up the search process [58].

The MLSR algorithm using the PySR package is employed to create analytical expressions from the dataset obtained through KMC simulations (Section 3.1) in which the H diffusion coefficient is a function of the diffusion barrier distributions. Two approaches are taken: a data-driven approach and a physics-informed approach. For the data-driven approach, both direct and indirect models are made. The direct data-driven approach directly expresses D and as a function of the distribution of diffusion barriers and the temperature, $D = f(\mu, \sigma_s, \sigma_w, T)$, while the indirect approach expresses $\ln(D)$ as a function of the same parameters. The indirect model is explored to force the expressions to follow the $\exp(\cdot)$ form found in Arrhenius-type activation functions. Such formulations yield expressions that can closely resembles the dataset. However, the data-driven approach does not contain much information about the physics involved in varying the statistical barrier distribution and would thus be hard to interpret. Therefore, more

interpretable physics-informed expressions are made. The physics-informed approach uses the Vogel–Fulcher–Tammann (VFT) model [61], which can be seen in Equation 3.8. This model can be used to describe super-Arrhenius behaviour where $\ln(D)$ is not linear with respect to the temperature, especially at lower temperatures. This non-linearity is quantified by the Vogel temperature T_0, below this theoretical temperature diffusion ceases. The parameter Q_{VFT} represents the effective activation energy, and the pre-exponential factor D_0 is a measure of the diffusivity at infinite temperature [42]. The three parameters of the VFT model are learned through MLSR as analytical functions of the parameters μ, σ_s, and σ_w, which describe the energy barrier distribution. The physics-informed approach leads to a better understanding of the behaviour at varying distributions of diffusion barriers.

$$D = D_0 \exp\left(-\frac{Q_{VFT}}{k_b(T - T_0)}\right) \tag{3.8}$$

3.3. Universal machine learning interatomic potentials

Interatomic potentials express a system's potential energy as a function of the atomic positions. Traditional interatomic potentials are derived from physical insight into the nature of chemical bonds. Machine learning interatomic potentials (MLIPs) instead part with the physical insight and aim to predict the potential energy of the system by numerical interpolation generated by quantum mechanical calculations [62].

Electronic structure methods based on direct quantum mechanical treatment of the electrons, such as DFT, are the most accurate way to calculate energies and forces of both pure elements and multicomponent systems. However, the major limit of DFT calculations are that they are computationally expensive, as they scale with N^3, in which N is the number of atoms. Consequently, static DFT calculations are limited to a few hundred atoms and ab initio MD (AIMD) can only be run for a few hundred picoseconds. The length and time scales needed to model many materials processes, such as diffusion, require much larger scales due to the need to statistically average over many atoms and events, which is not achievable with DFT calculations. Traditional interatomic potentials provide a solution by greatly reducing computational cost by simplifying the interatomic potential to a fixed potential which is used in all atomic configurations. This simplification, however, comes at the price of reduced accuracy [62].

MLIPs can bridge the gap between accurate but slow DFT calculations and fast but inaccurate traditional interatomic potentials. State of the art MLIPs often achieve $<$ 5meV atom^{-}1 in energies and $<$ 0.1 eV Å^{-1} in forces, which is almost an order of magnitude better than traditional interatomic potentials [51].

MLIPs can be categorised into two types: specialised MLIPs (sMLIPs) and universal MLIPs (uMLIPs). sMLIPs are designed to specific material systems. They are trained on DFT data for a particular composition and structure. This narrow focus allows sMLIPs to achieve high accuracy and high computational efficiency within their applicability. The downside is that these potentials require large system specific datasets and are not able to extrapolate outside of the training regime [63]. uMLIPs are ready-to-use models that are widely applicable across the periodic table with DFT-level accuracy. uMLIPs offer a way to replace computationally expensive DFT calculations, which is needed for training of sMLIPs. In recent years, more accurate uMLIPs have been released. It has been found that certain uMLIPs models, trained on existing datasets, can replace costly DFT calculations for training of sMLIPs and can even outperform certain sMLIPs. Furthermore, it has been shown that uMLIPs offer an advantage

over sMLIPs where extrapolation in case of limited data is needed [64]. This extrapolative capability is particularly valuable in cases where data is limited, such as in sampling the large configuration space of MPEAs in this study.

To perform the CI-NEB calculations (Section 3.4) and MD simulations (Section 3.5), uMLIPs are employed that are developed using the Graph Atomic Cluster Expansion (GRACE) framework [65]. Specifically, the GRACE-FS-OMAT potential, trained on the OMat24 dataset, is used for all CI-NEB and MD simulations due to its favourable balance between accuracy and computational efficiency. For benchmarking, additional CI-NEB calculations are performed using two other uMLIPs: the GRACE-FS-OAM potential, also based on the OMat24 dataset but further fine-tuned on the sAlex and MPTrj datasets [66, 67], and the GRACE-2L-OMAT potential, a more complex two-layer version trained on the OMat24 dataset. The GRACE-2L-OMAT potential is only used to compute a subset of CI-NEB diffusion barriers due to its higher computational cost [64]. All simulations are conducted using the LAMMPS package [68].

3.4. Climbing-image nudged elastic band

The NEB method is a way to find the minimum energy path (MEP) of an energy landscape between two equilibrium configurations. The potential energy maximum along the MEP is the saddle point, which is used to estimate the activation energy barrier [69]. The MEP is found by constructing images of the system between the initial and final state, this can be seen in Figure 3.9. A spring interaction, imitating an elastic band, is added between images to ensure continuity of the path. The forces acting on the images are minimized, which optimizes the elastic band and brings the band to the MEP. A distinct feature of the NEB method is a force projection to ensure that the spring forces do not interfere with the convergence of the elastic band to the MEP. The tangent of the path needs to be estimated at each image and every iteration in order to decompose the true force and the spring force into parallel and perpendicular components. The perpendicular component of the true force and the parallel component of the spring force are included. The force projection is known as nudging. The spring forces control the spacing of the imaged along the band. It is essential to use this force projection, otherwise the spring forces tend to prevent the band from following a curved MEP due to corner-cutting. In the NEB method the strength of the spring forces can be varied by several orders of magnitude without affecting the equilibrium position of the band [70].

The climbing-image NEB method is a variant of the NEB method containing a small modification without significant computational effort. The information about the shape of the MEP is retained, but a convergence to a saddle point is also obtained. After sufficient iterations with regular NEB, the image with the highest energy is identified, this image is promoted to climb up to the transition state. The climbing-image moves up the energy surface along the elastic band and down the energy surface perpendicular to the band. This leads to the saddle point, as the saddle point is the maximum along the reaction coordinate and a minimum in all directions perpendicular to that [70].

The CI-NEB method is applied to equimolar $4 \times 4 \times 4$ BCC supercells created using Atomsk [72] with periodic boundary conditions containing 128 atoms with one hydrogen atom inserted into the structure at a tetrahedral site. The lattice constant is determined by relaxing a $12 \times 12 \times 12$ BCC supercell containing 3456 atoms with the same composition as the smaller supercell, with the larger supercell used to ensure that the calculated lattice constant is representative of the bulk crystal. The interatomic interactions are described using uMLIPs (Section 3.3), which are chosen due to their universal applicability across a broad range of elements and alloy compositions, making them suitable for studying the vast compositional space of BCC MPEAs without requiring additional training on DFT data. Two configurations are prepared,

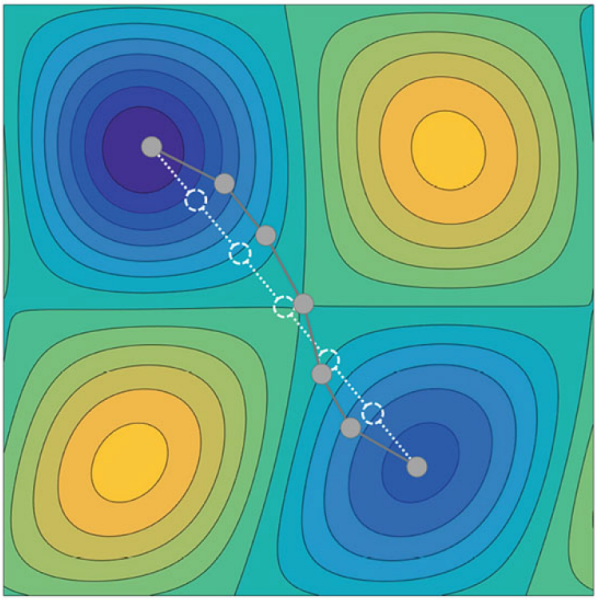


Figure 3.9: NEB method to relax a chain of connected images to determine the minimum energy path between two minima. The white path marks the initial set of images, while the grey path shows the minimized configuration. The spheres represent the images [71].

corresponding to the hydrogen atom at the initial and final positions of the specific diffusional jump. The $4 \times 4 \times 4$ supercell is initially unrelaxed, but the structures will be relaxed during the NEB process, thereby introducing lattice strain. The CI-NEB algorithm using LAMMPS [68] with 5 images, a spring constant of 0.1 eV/Å^2, and a force tolerance of 0.01 eV/Å is run on all 1536 unique T-T paths of the supercell. The initial, maximum, and final energy are recorded to compute the forward and backward barriers, leading to 3072 diffusion barriers. The results are shown in Section 4.4.

3.5. Molecular dynamics

MD is a method to study the dynamics of atoms and molecules over time in a force field. It is based on the principles of classical mechanics to model atomic systems, calculating the forces on atoms and solving Newton's equations to track their resulting motion [50]. The total force on each atom is computed using interatomic potentials. Integration algorithms, such as the Verlet algorithm, are used to evolve the atomic system over time. The Verlet algorithm updates atomic positions by using the current positions and accelerations at time t, along with the positions from the previous time step [73]. There is a fine balance between accuracy and computational cost in choosing the optimal time step. The force is assumed to be constant over the time interval, which leads to larger numerical errors and unstable particle motion with larger time step. However, a smaller time step leads to a greater computational cost, as each time step requires a force evaluation which consumes the largest amount of computational cost [50].

The major limitation to the MD method is the computational efficiency. The computational cost scales roughly with the number of atoms. Consequently, MD simulations are limited to the nanosecond range at best [50]. It often occurs that the phenomena of interest, such as long-term interstitial diffusion, takes place in time scales longer than the atomic vibrations that control the maximum time step that can be used [74].

MD is implemented in LAMMPS [68] for H diffusion of chemically complex BCC structures

using the GRACE-FS-OMAT uMLIP [65] and run on the DelftBlue supercomputer provided by the Delft High Performance Computing Centre [57]. $10 \times 10 \times 10$ BCC supercells containing 2000 atoms of equimolar composition are created using Atomsk [72]. First, the atomic positions are relaxed with a force criteria of 0.01 eV/Å or stopped if 1000 steps is reached. After which, each structure without H is run for 10,000 steps at 1 fs per timestep at a specific temperature using the isothermic-isobaric (NPT) ensemble, this ensemble maintains constant temperature at an applied pressure of 0 bar. From these initial simulations the average lattice constant at that temperature is obtained in order to rescale the box size for the main simulation of H diffusion.

Second, a singular H atom is inserted into the rescaled box at a random T-site. The MD simulations are run for 5,000,000 steps using a timestep of 0.5 fs, leading to a runtime of 2.5 nanoseconds. A small timestep is chosen to capture the rapid interactions of the H atom. For this simulation the canonical (NVT) ensemble is used for improved stability, as box size fluctuations can interfere with the MSD calculation. In the NVT ensemble the volume is fixed, the equilibrium volume at that temperature is already known from the previous simulation with NPT ensemble.

Each MD simulation is run 5 times with different initial configurations with random elemental distribution in order to obtain statistical accuracy. Thus, in total 10 ns is simulated for each temperature for each considered alloy system. The temperatures range from 600 to 1400 K in steps of 200 K to capture both low and high temperature behaviour. From the MD simulations the MSD is extracted and the diffusion coefficient is calculated in the same way as described in Section 3.1.2. The diffusion coefficients obtained through MD are compared to the diffusion coefficients obtained through the MLSR and CI-NEB framework, the results are shown in Section 4.4.

4

Results

This chapter presents the main findings of this work, focusing on how the statistical characteristics of energy barriers influence hydrogen diffusivity in complex energy landscapes. Through KMC simulations and MLSR expressions, the effects of both symmetric (saddle point) and asymmetric (well-energy) disorder are investigated. The results are discussed in relation to diffusion coefficients, activation energies, and temperature-dependent behaviour, and are ultimately applied to real alloy systems using uMLIPs. The application to real alloy systems is carried out using CI-NEB calculations to extract diffusion barrier distributions, which are then used to predict diffusivity through MLSR expressions and validated by direct MD simulations.

4.1. Diffusivity of hydrogen

In this section, KMC simulations are used to study hydrogen diffusion in model systems characterised by statistical distributions of diffusion barriers. The aim is to understand how variability in diffusion barriers affects overall hydrogen mobility. The influence of each parameter is explored in isolation and in combination, across a range of temperatures.

Paths and trajectories of H atoms

The KMC simulations are performed in a $20 \times 20 \times 20$ BCC supercell containing 192,000 unique paths that are connected through T-sites, where each T-site is connected to four other T-sites. This connectivity can be seen in Figure 4.1 for a $4 \times 4 \times 4$ BCC supercell. Each unique path is assigned a saddle point s and well-energy w based on their Gaussian distributions, defined by the the mean barrier height μ, the standard deviation of the saddle points σ_s, and the standard deviation of the well-energies σ_w.

Figure 4.2 compares visitation-frequency maps for two trajectories that are run for 200,000 KMC steps at $T = 350\,\text{K}$ with $\mu = 0.20\,\text{eV}$: (a) a random walk ($\sigma_s = \sigma_w = 0\,\text{eV}$) and (b) a rough energy landscape ($\sigma_s = \sigma_w = 0.06\,\text{eV}$). In the random walk case, visits are broadly uniform, all sites can be equally preferred. In the rough energy landscape, the H atom repeatedly moves among a few sites (dark clusters), which act as effective trapping regions. It can be seen that the random walk has a much more evenly distributed trajectory. The H atom has no bias for specific sites and is equally likely to visit any site. In contrast, for the rough energy landscape the atom spends most of its time moving within a few regions that act as effective traps. It periodically escapes to another cluster, giving rise to hotspots with high visitattion frequency across the supercell. This trapping lowers the diffusivity because the atom experiences long residence times in low energy regions and only rarely makes long range hops, yielding a

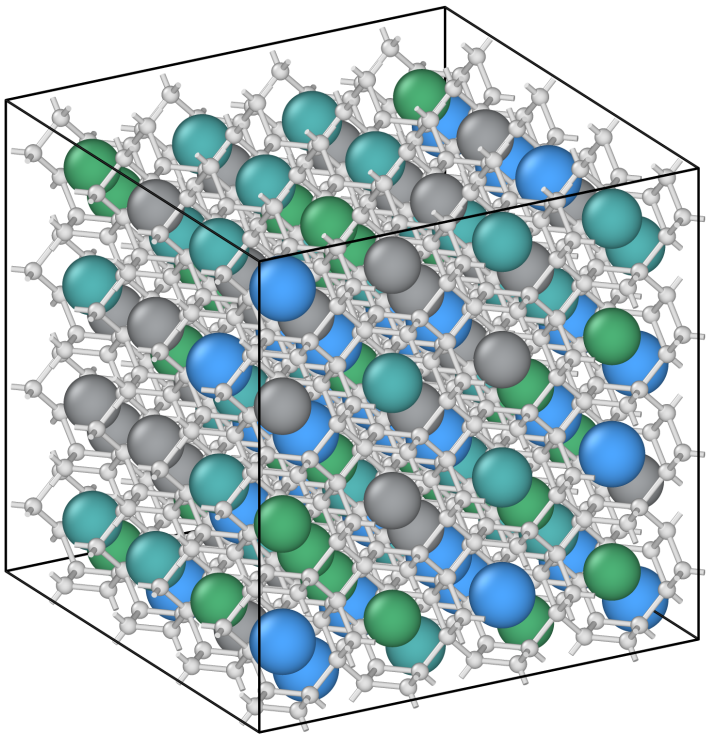


Figure 4.1: All possible T-sites and connected paths of H in a BCC lattice. A $4 \times 4 \times 4$ BCC supercell is shown, the KMC simulations utilise a $20 \times 20 \times 20$ BCC supercell.

shallower MSD–time slope and thus a smaller D, which can be seen from the MSD-time plot in Figure 4.3a. In this MSD-time plot, both simulations are run until an MSD of $\geq 10^5$ Å is reached.

For the rough energy landscape (blue line, Figure 4.3a), the MSD-time curve alternates between steep and shallow segments. The vertical spacing between points represents the distance that is travelled between effective trapping sites (with a jump-based criterion of 20 Å, as mentioned in Section 3.1.2), while the horizontal spacing reflects the residence time between each 20 Å jump. Longer residence times in traps correspond to shallower segments, while the travelling between trapping regions lead to steeper segments. The alternating nature suggests the behaviour that is seen in Figure 4.2b, where atoms hop between clusters of deep traping regions. By contrast, in the random walk case (red line, Figure 4.3a) the residence time between successive points is nearly constant, leading to an MSD–time slope of uniform magnitude between segments. The difference between the two cases is also reflected in the distribution of MSD gradients ($\Delta\text{MSD}/\Delta t$) in Figure 4.3b. For the rough energy landscape case (blue), there is a shift towards lower gradients compared to the random walk (red) case. In addition, for the rough energy landscape, still some gradients are high, which suggests the hopping between clusters of trapping regions.

Influence of energy landscape parameters on diffusivity

In the following part, the influence of the statistical energy landscape parameters μ, σ_s, and σ_w on diffusivity will be investigated. For this, diffusivity will be reported in a normalised form (D/D_u), this is the ratio of the diffusion coefficient D to that of a uniform barrier distribution D_u, where $\sigma_s = \sigma_w = 0$. This ratio reflects the effect of barrier fluctuations on diffusion and is a dimensionless quantity. Values of $D/D_u > 1$ indicate that diffusion is enhanced by the

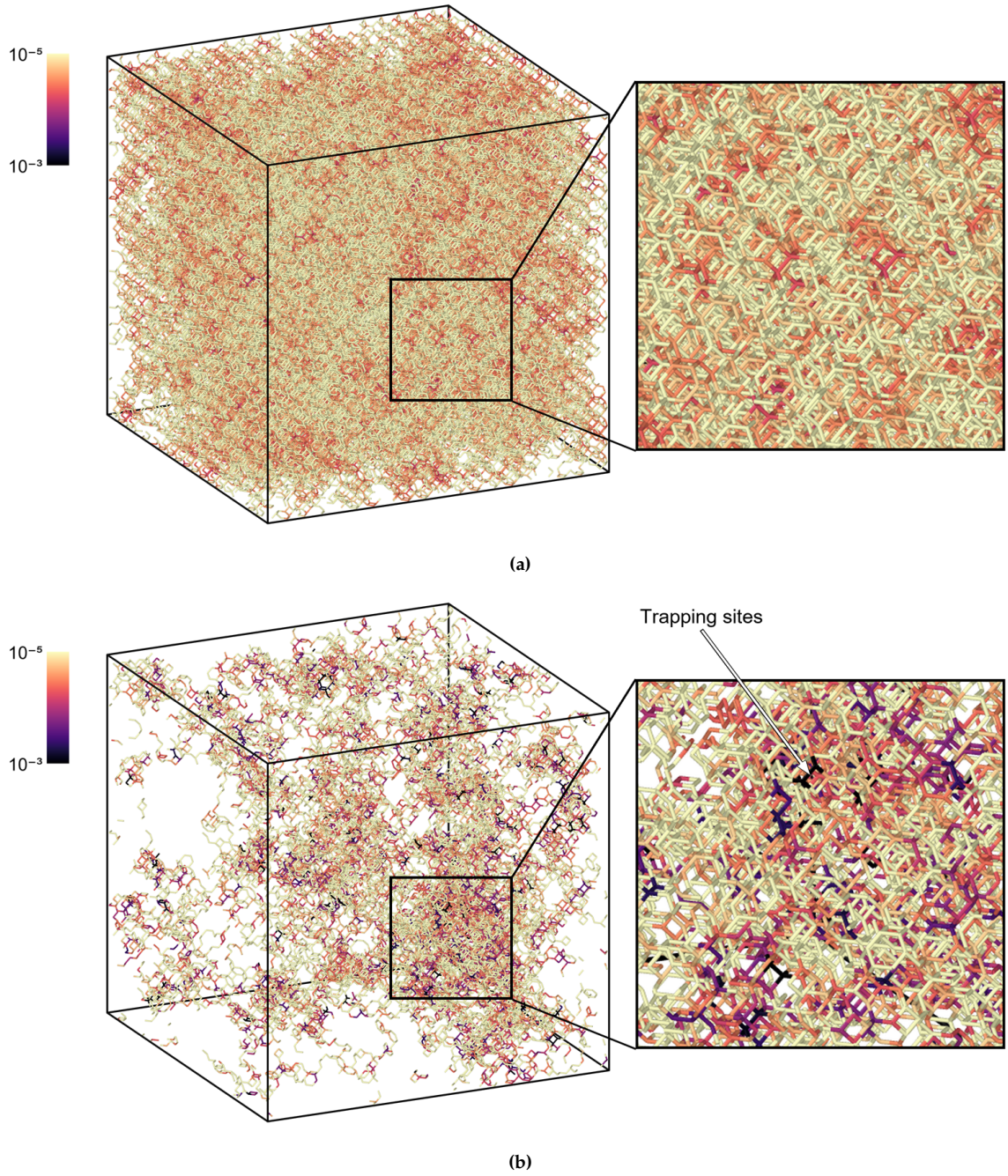


Figure 4.2: Visitation frequency (log scale) of sites for: (a) a random walk ($\sigma_s = 0$ eV, $\sigma_w = 0$ eV), and (b) a rough energy landscape ($\sigma_s = 0.06$ eV, $\sigma_w = 0.06$ eV). Both KMC simulations are run for 200,000 steps at 350 K and with $\mu = 0.20$ eV.

presence of fluctuations, while values of $D/D_u < 1$ indicate a reduction in diffusion.

The results of all KMC simulations are presented in Figure 4.4 and 4.5. Figure 4.4 shows the isolated effect of σ_s and σ_w on the normalised diffusion coefficient, while Figure 4.5 shows the coupled effect of σ_s and σ_w, where the x-axis shows the increase in σ_w and the colour legend shows the increase in σ_s. The lines represent the trend of diffusivity upon increasing σ_w, where the σ_s parameter is held constant. At each point on the x-axis, the variation in D/D_u corresponds to different values of μ. Thus, all KMC simulations at those temperatures are

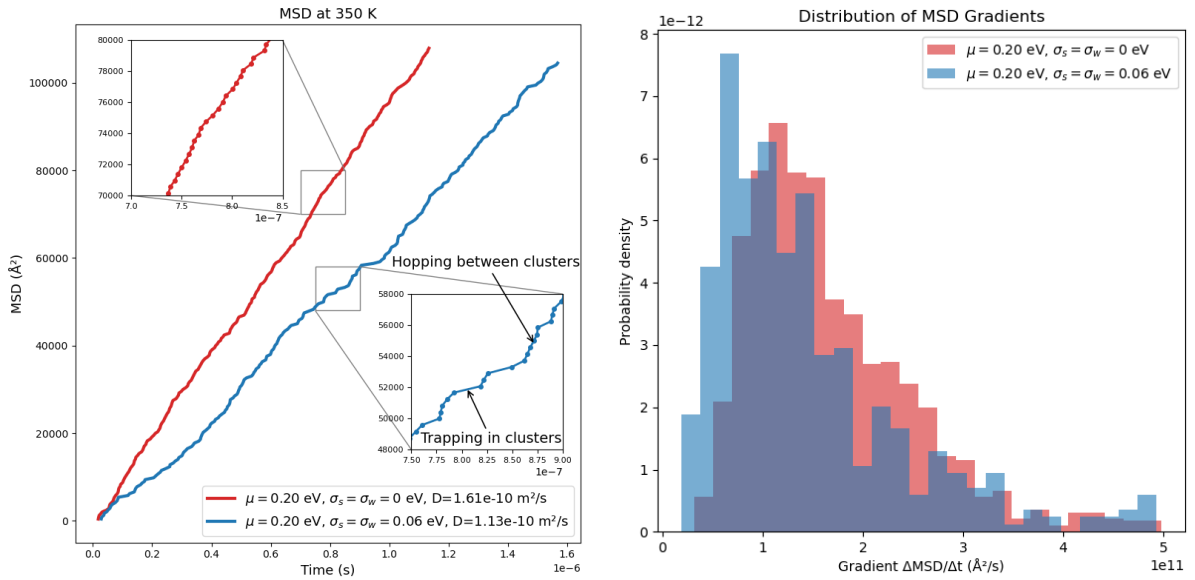


Figure 4.3: MSD at $T = 350$ K and $\mu = 0.20$ eV comparing a random walk (red: $\sigma_s = \sigma_w = 0$ eV) and a rough energy landscape (blue: $\sigma_s = \sigma_w = 0.06$ eV). (a) MSD-time curve, (b) distribution of ΔMSD/Δt gradients. High gradients represent movement between trapping clusters, while low gradients suggest trapping in these clusters.

shown in the figure.

The effect of σ_w on the diffusivity is dependent on the temperature. It can be seen that at low temperature (350 - 600 K), moderate σ_w enhances diffusivity, while larger values of σ_w decrease diffusivity (Figure 4.4). A possible explanation to this is that, at low to moderate σ_w, some barriers become lower than the mean diffusion barrier, this creates preferred low energy pathways. The probability that at least one barrier is unusually low increases with σ_w. However, upon increasing σ_w, excessive asymmetry also creates high barrier paths, increasing the fraction of sites where the backward direction and all three forward directions are blocked due to high energy. This traps the atom in certain low energy regions and reduces diffusivity at higher σ_w values. Thus, for enhanced diffusion, some energy barriers have to be low, causing fast diffusion pathways. But, if σ_w becomes too large, the probability of sites where all paths have high energy barriers increases and diffusivity is reduced. This specifically happens for the asymmetric component, as a low forward barrier becomes a high backward barrier after the jump. When all other paths outside the trapping region have high energy, the atom is blocked and moves within the low energy region. This can be further seen in the first row of Figure 4.6, where the accessible diffusion barrier spectra is plotted for two isolated σ_w cases. It can be seen that the spectra slightly shifts to the left and narrows. However, even at low temperatures, large barriers are accessed. Indicating that occasionally the atom must hop over these higher barriers after long residence times at trapping sites, leading to reduced diffusivity. At high temperature ($\geq$ 1200 K), the thermal activation energy is large enough to average out the forward and backward asymmetries of the paths, therefore, the effect of σ_w diminishes, leading to D/D_u approaching 1.

The isolated effect of σ_s from Figure 4.4 shows a slight increase in diffusivity ($D/D_u \approx 1.1$) upon increasing σ_s. An increase in σ_s broadens the barrier distribution, but at low temperature diffusion is dominated by the preferred low s pathways, increasing the diffusivity. The atom does not get trapped in places where all barriers are large, because the jump into that site

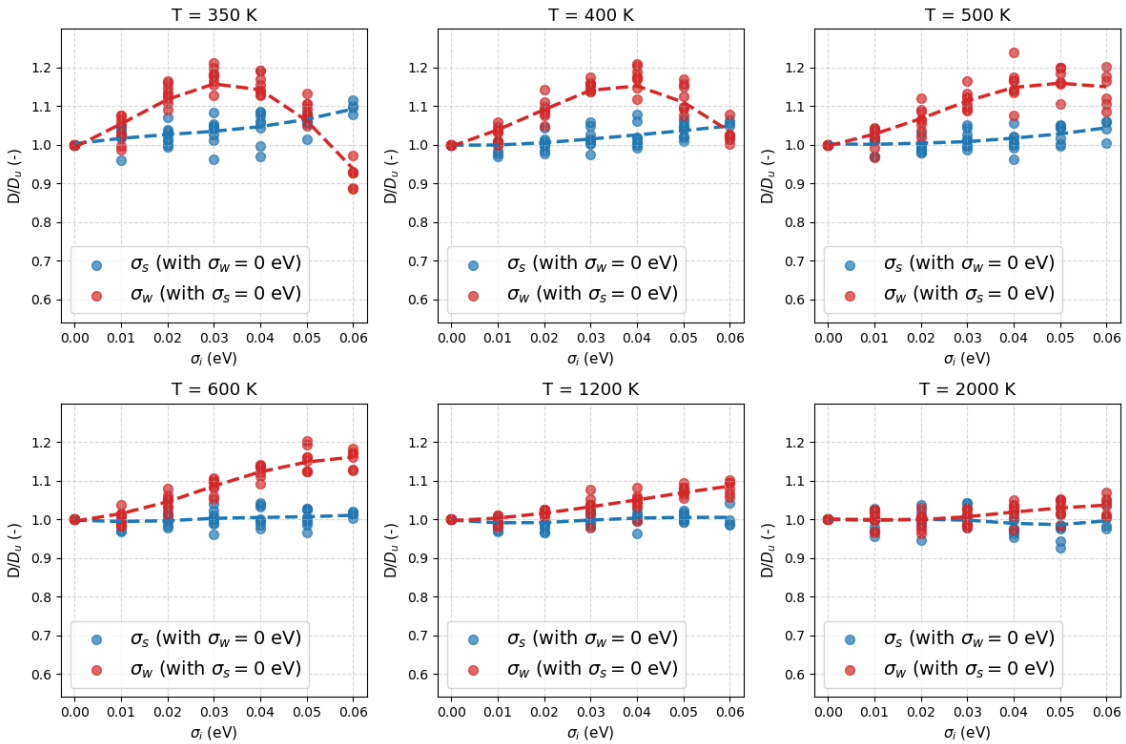


Figure 4.4: Normalised diffusion coefficient (D/D_u), a dimensionless quantity, plotted against σ_w (red symbols/line, with $\sigma_s = 0$) and σ_s (blue symbols/line, with $\sigma_w = 0$), for all μ values across different temperatures. Both parameters are shown within each subplot to enable direct comparison. D_u is the diffusion coefficient for a uniform barrier distribution ($\sigma_s = \sigma_w = 0$)

should also be a high barrier, and thus unlikely, as supposed to the asymmetric component. The analysis of the accessible diffusion barrier spectra (Figure 4.6, second row) supports this, showing a strong leftward shift towards lower barriers at low temperature due to the preference for low energy barriers, where low barrier pathways dominate because of their exponentially higher hopping rates. At high temperature this effect diminishes and D/D_u approaches 1.

However, when both σ_s and σ_w are present, as in Figure 4.5, an increasing σ_s instead decreases diffusivity when σ_w is moderate to large. While, an increase in σ_s at low σ_w increases diffusivity. With increasing σ_s, the probability of encountering sites where all available hopping paths have high barriers increases, thus significantly reducing diffusivity. This mechanism relies on the asymmetric component to allow the atom to enter these regions, thus this increase in trapping sites that are encountered are not seen in the isolated σ_s case. In the accessible diffusion barrier spectra of the third row of Figure 4.6 the combined effect of σ_s and σ_w can be seen. The spectra resemble the isolated σ_w case, suggesting a similar underlying diffusion mechanism dominated by asymmetric barrier fluctuations. However, the distribution shifts even more strongly toward lower barriers when both disorders are present, demonstrating intensified energetic selection. Simultaneously, the presence of high energy barriers increases due to symmetric disorder, enhancing the trapping effect and thus further suppressing diffusion.

The effect of μ on the normalised diffusion coefficient D/D_u is negligible, as D_u equally scales as D upon increasing μ. This suggests that the mean barrier height does not significantly increase the effects of σ_s and σ_w. The effect of μ on D, however, is significant. Figure 4.7 shows the effect of increasing μ on diffusivity at various temperatures and various values of σ_s and σ_w (the unplotted σ is held at zero, thus showing isolated effect of σ_s/σ_w). It can be seen that at

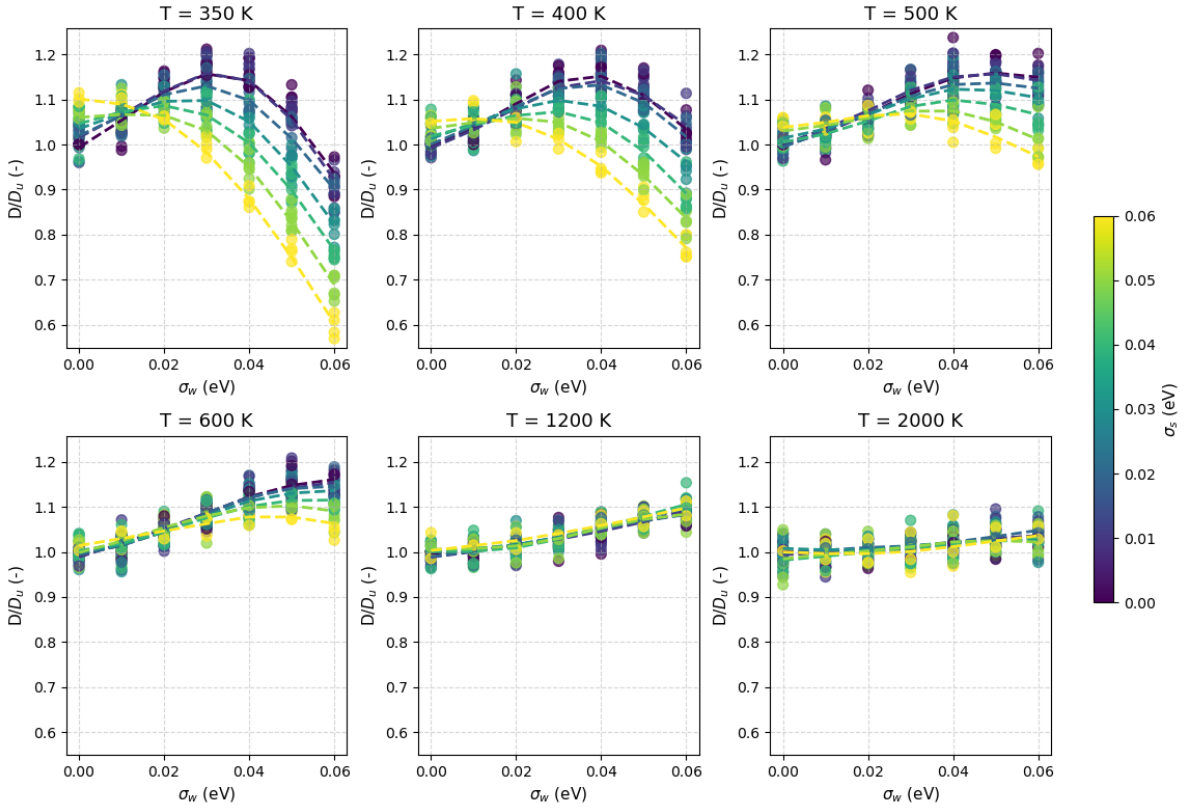


Figure 4.5: Normalised diffusion coefficient (D/D_u), a dimensionless quantity, plotted against σ_w with σ_s indicated in the colour legend, for all μ values across different temperatures. The lines indicate trends of increasing σ_w, where σ_s is held constant. D_u is the diffusion coefficient for a uniform barrier distribution ($\sigma_s = \sigma_w = 0$).

low temperatures, increasing μ from 0.16 to 0.23 eV lowers diffusivity by an order of magnitude and in a non-linear way. At higher temperatures, the decrease in diffusivity is much less strong and becomes linear with respect to μ. This is a clear consequence of the increase in thermal activation energy at higher temperatures.

In conclusion, at low temperature the asymmetric component σ_w is mainly responsible for trapping of the atom in regions that are surrounded by high energy barriers. Asymmetry allows the atom to reach these sites as the forward and backward barrier are low and high, respectively. These trapping regions reduce diffusivity. The symmetric component σ_s enhances the trapping of atoms when in combination with moderate to high σ_w, reducing diffusivity. However, σ_s alone increases diffusivity due to preferred low s pathways. Thus, it is crucial to both consider σ_s and σ_w. In all cases, at high temperature, the thermal activation energy is sufficiently large to reduce the effect of the trapping mechanism of σ_w, as the atom is able to more easily leave trapping sites, this causes D/D_u to approach 1.

4.2. Super-Arrhenius behaviour of hydrogen

Super-Arrhenius behaviour refers to the deviation from Arrhenius linearity in plots of $\ln(D)$ versus $1/T$. This curvature is well captured by the Vogel–Fulcher–Tammann (VFT) model (Eq. 3.8) [61]. In this model, the Vogel temperature T_0 sets the scale of the non-linear drop. As T reduces to T_0, the diffusivity tends to zero. The parameter Q_{VFT} plays the role of an effective activation energy that controls the steepness of the $1/(T - T_0)$ dependence, while the pre-exponential factor D_0 sets the high temperature diffusivity scale [42]. To further analyse

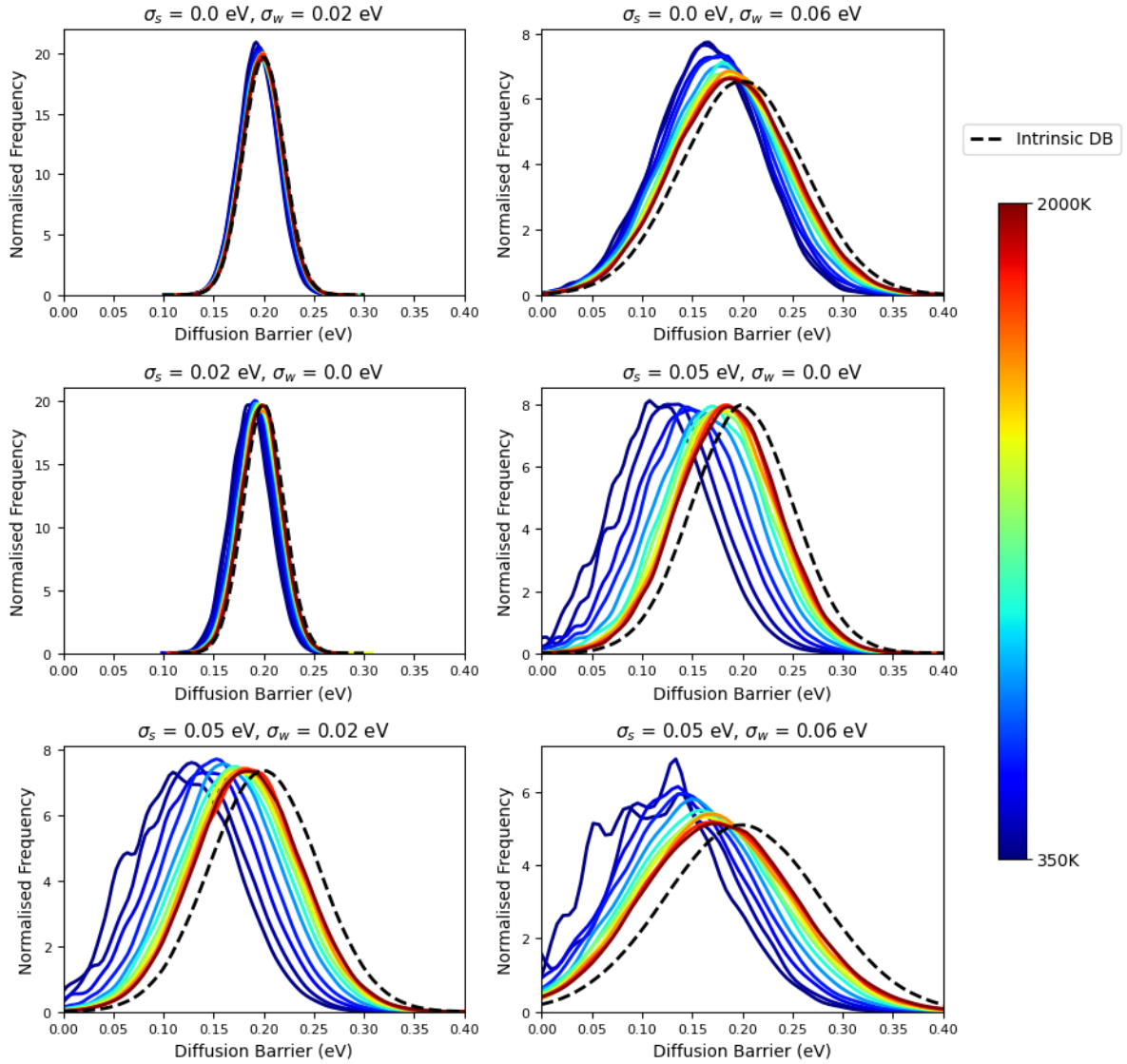


Figure 4.6: Accessible diffusion barrier spectra for various model systems with varying symmetric (σ_s) and asymmetric (σ_w) disorders at $\mu = 0.20$ eV. The spectra represent barriers actually encountered by the atom during diffusion at temperatures ranging from 350 K to 2000 K. First row: isolated σ_w effect; Second row: isolated σ_s effect; Third row: combined effects of σ_s and σ_w.

H diffusion in chemically complex alloys, the KMC dataset was fitted with the VFT model. Figure 4.8 shows how the fitted parameters (D_0, Q_{VFT}, T_0) correlate with the energy landscape parameters.

The pre-exponential factor D_0, which ranges from 1.1×10^{-7} to 1.35×10^{-7} m^2/s in the dataset, influences the diffusion rate at infinite temperature. It is found that D_0 is not significantly affected by either μ or σ_s, with very weak correlations observed ($R^2 = 0.01$ and 0.02, respectively). However, a substantial negative correlation ($R^2 = 0.73$) is observed with σ_w.

For activation processes following Arrhenius type relations, the Meyer–Neldel rule states that the pre-exponential factor correlates with the activation energy E in a predictable way: $\ln(D_0) = a + bE$, in which a and b are constants [75]. A previous study on H diffusion in BCC MPEAs found a correlation ($R^2 = 0.48$) between $\ln(D_0)$ and the effective activation energy Q_{VFT} that follows this Meyer–Neldel rule [42]. In this work, the pre-exponential factor is more

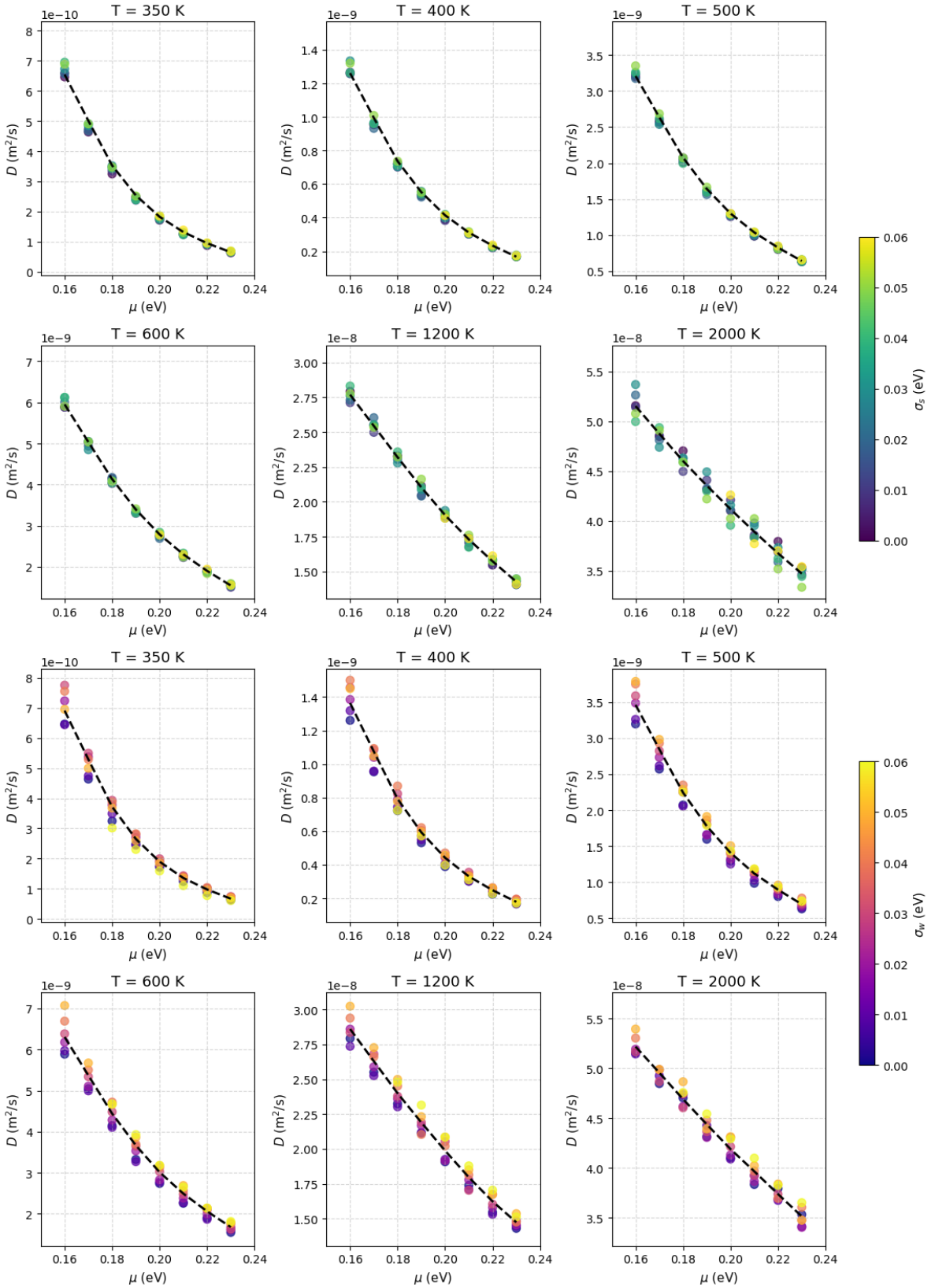


Figure 4.7: Effect of μ on the diffusion coefficient D at various temperatures. Results are shown for isolated cases of symmetric component ($\sigma_s > 0, \sigma_w = 0$) and asymmetric component ($\sigma_w > 0, \sigma_s = 0$).

strongly correlated with the Vogel temperature T_0 ($R^2 = 0.82$), as can be seen in Figure 4.9. However, a Meyer–Neldel–type positive relation between D_0 and Q_{VFT} still exists, but weaker ($R^2 = 0.10$). In the reference system used here, both D_0 and T_0 are strongly determined by σ_w, while Q_{VFT} is more strongly correlated with μ (Figure 4.8). Since D_0 and T_0 respond to the same parameter, a strong relation between them naturally appears. By contrast, σ_w only weakly reduces Q_{VFT}, so the D_0–Q_{VFT} trend is comparatively shallow.

With σ_w acting as the principal parameter that controls the non–linear, super–Arrhenius drop in diffusivity (through its strong correlation with T_0), the fitted D_0 co–varies with T_0 to accommodate the curvature, which produces the pronounced D_0–T_0 trend. Accordingly, it is expected that the D_0–T_0 relation observed here is primarily a fit-dependent effect, whereas the D_0–Q_{VFT} correlation reflects the more general, physics-based behaviour according to the Meyer-Neldel rule. Additionally, the spread of D_0 is relatively small in this dataset, and because the maximum temperature of the KMC simulations is 2000 K (below the high-temperature limit), D_0 is largely set by the fit (and its covariance with T_0), making the D_0-T_0 relation appear more prominently.

The effective activation energy parameter Q_{VFT} (ranging from 0.13 to 0.23 eV) demonstrates a strong, almost linear correlation ($R^2 = 0.81$) with μ, indicating that higher mean barriers significantly increase the energy required for diffusion. Interestingly, σ_s has minimal influence on the activation energy, as seen by a negligible correlation ($R^2 \approx 0$). In contrast, σ_w exhibit a small negative correlation ($R^2 = 0.11$) with Q_{VFT}, suggesting that asymmetry lowers the effective activation energy by creating preferential pathways with lower barriers.

Finally, the Vogel temperature T_0 (ranging from 0 to 76 K) indicates the temperature below which diffusion rapidly decreases, approaching a trapped state. This parameter does not correlate significantly with either μ or σ_s, but shows a strong positive correlation ($R^2 = 0.76$) with σ_w. Thus, increasing asymmetry significantly raises T_0, reinforcing earlier observations from Section 4.1 that σ_w strongly enhance trapping at low temperatures.

The diffusion barrier distribution formed by μ, σ_s, and σ_w is a perfect Gaussian distribution. As a consequence, it overestimates the number of low energy barriers (Figure 3.6). With such energy distributions, more alternative easy hops are available when the H gets trapped. Therefore, the super-Arrhenius behaviour is reduced and consequently the range of T_0 values compared to barriers with real (skewed) barrier distributions from ref. [42] is smaller. The maximum T_0 found in this work is 76 K, while that for real barriers is approximately 200 K [42].

4.3. Machine learning symbolic regression models

The dataset that is obtained through KMC simulations is used to create analytical expressions using MLSR that link the properties at the nanoscale (distribution of energy barriers) to the diffusion coefficient. The PySR framework [58] is used to learn the expressions. Two approaches are taken: (1) a data-driven approach, in which diffusivity is expressed as a function of the barrier distribution parameters and the temperature, and (2) a physics-informed approach, in which the KMC dataset is fitted to the VFT model that is able to describe super-Arrhenius behaviour, after which the parameters of the VFT model are expressed using the barrier distribution parameters.

4.3.1. Data-driven approach

For the data-driven approach, all 352 model systems with varying values of μ, σ_s, and σ_w are considered. The range of values taken can be seen in Table 3.1, with steps of 0.01 eV. Since the diffusion coefficient of each system is evaluated at 11 distinct temperatures, the dataset

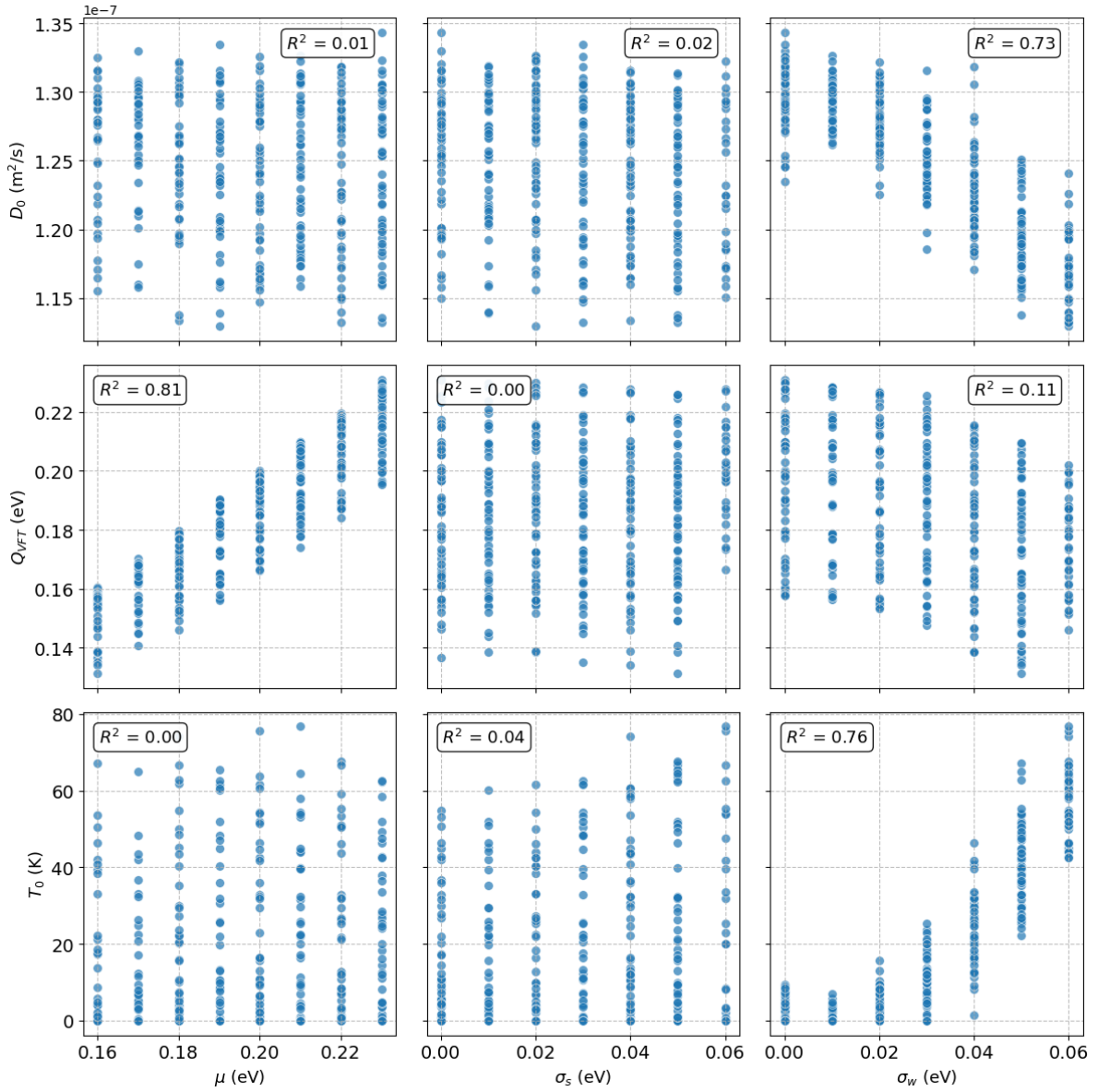


Figure 4.8: Correlations between statistical barrier properties of 352 systems and VFT model parameters.

comprises a total of 3872 data points. A training/test split of 80/20 is taken, meaning that the expressions are learned on 80% of the data and further evaluated on the 20% of unseen data to prevent overfitting.

For the data-driven approach, two MLSR models are developed. In the first, expressions are constructed where D is learned directly as a function of μ, σ_s, σ_w, and T. In the second, the natural logarithm of D is modelled instead, with $\ln(D)$ expressed as a function of the same variables. These models are referred to as the direct and the indirect data-driven approach, respectively. The logarithmic transformation rearranges the expressions so that D appears in an $\exp(\cdot)$ form, resembling the Arrhenius-type representation of activation energies. As will be shown in the following sections, this leads to significantly improved predictive accuracy.

The models are run for 1000 iterations with a population size of 50, meaning 50 candidate expressions are created and evaluated per iteration. The MSE loss function is used to evaluate

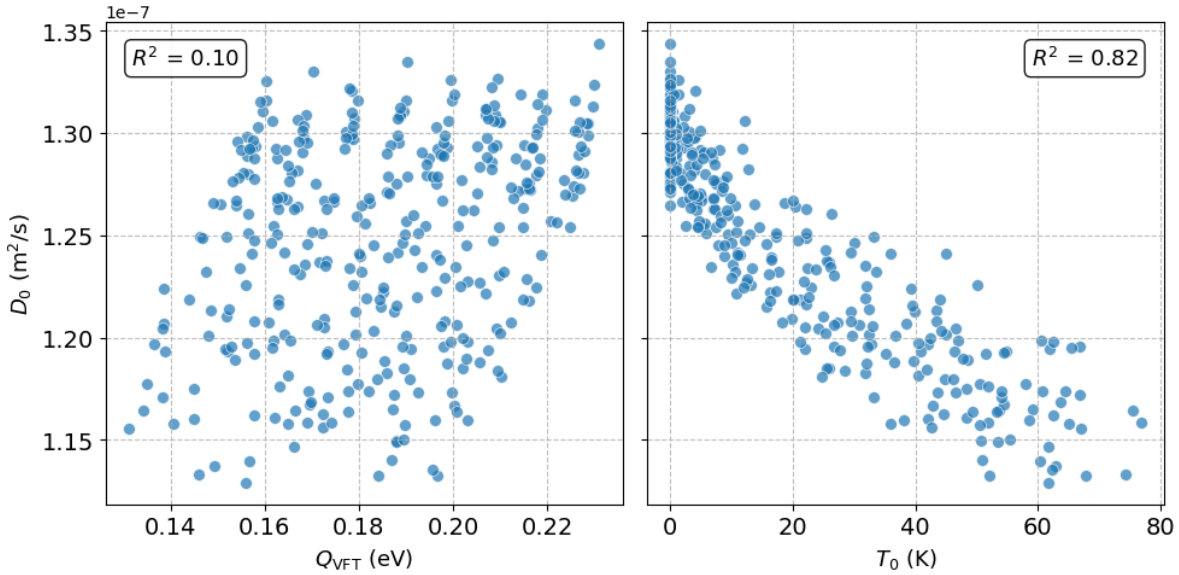


Figure 4.9: Correlations between the pre-exponential fatcor D_0 and the VFT model parameters Q_{VFT} and T_0.

the expressions. Expressions are constructed using binary operators (+, -, ×, ÷, ^) and unary functions (log, sqrt, exp) because these match the forms commonly used in diffusion models. In particular, exp(·) captures Arrhenius activation energy and temperature dependence, while log, exponents, and $\sqrt{\cdot}$ represent scaling and simple power laws. Trigonometric functions (sin, cos, tan) are omitted to preserve interpretability in the H diffusion context. In addition, units are given to the variables to ensure consistency in the expressions and the expressions were able to learn the Boltzmann constant (k_B = 8.6173e−5 eV/K) to improve unit consistency. Moreover, a penalty is given to expressions that do not adhere to correct units.

The MLSR algorithm is set to output multiple expressions of varying complexity up to a maximum of 35 operations (or nodes). No constraint is set to the maximum tree depth (nested operations). To compare the output expressions, PySR uses a performance metric called score, defined in Eq. 4.1, where the complexity corresponds to the number of nodes of the expression. The score represents drops in the loss-complexity curve, in other words, an expression with a high score minimises the loss and number of nodes, simultaneously [58, 76]. Model selection is performed using one of three options: "accuracy", which selects the expression with the lowest loss regardless of score; "score", which selects the expression with the highest score regardless of loss; and "best", which chooses the expression with the highest score among those whose loss is within a factor of 1.5 of the most accurate model [77]. In this work, the "best" option is used, since it balances predictive accuracy and interpretability by combining both loss and score in the selection process.

$$\text{score}_i = -\log(\text{loss}_i/\text{loss}_{i-1})/(\text{complexity}_i - \text{complexity}_{i-1}) \tag{4.1}$$

It should be noted that the expressions reported here have been manually algebraically simplified from the raw PySR output to enhance readability and interpretability, while preserving their functional form and predictive accuracy. As a result, the number of nodes of the expressions presented may not directly correspond to the complexity values assigned by PySR.

Diffusivity as a function of statistical barrier distribution parameters and temperature
For the direct data-driven approach, 21 expressions were generated, with complexities ranging from 1 to 34. Figure 4.10 shows how both the loss and the score (as defined in Eq. 4.1) vary with complexity. Although lower losses correspond to higher predictive accuracy, they are generally achieved by more complex expressions, which reduces interpretability. The score therefore provides a more meaningful basis for model selection, as it balances accuracy against simplicity. The expressions with complexity 3, 5, and 9 (Eq. 4.2) achieve a high score with a decent loss. However, relative to the most accurate expression, the losses of these expressions all lie above the 1.5× threshold and are therefore excluded by the "best" model-selection rule. Beyond these, Eq. 4.3 (complexity 14) is the first expressions that involves the parameter σ_w, however it plays a minor role and does not improve loss significantly compared to the less complex Eq. 4.3 expression.

The expression with complexity 17 has good score, but lies just above the threshold line. The expression with complexity 19 (Eq. 4.4), lies below the threshold line and has a similar form. Using the "best" selection rule, Eq. 4.4 is chosen, as it has the highest score among the expressions with loss $\leq$ 1.5× the lowest found loss. What is interesting is that this expression does not involve σ_w, but does improve loss compared to Eq. 4.3, which did involve σ_w. Moreover, an even complexer expression can be seen in Eq. 4.5 with complexity 29. This expression does involve σ_w, but only in a minor role compared to μ and T. Across, all 21 expressions generated, σ_s does not appear in any of the derived expressions, showing that its influence on diffusivity is not recognised within the data-driven approach. This suggests that while σ_s and σ_w may play roles in the underlying physics, their contributions are overshadowed by μ and T in the direct data-driven setting, as seen by the pronounced influence of μ relative to σ_s and σ_w in Figure 4.7. This highlights a key limitation of a direct data-driven approach: secondary parameters such as σ_s and σ_w are not adequately captured, even though they have physical relevance, as found in Section 4.2. This motivates the use of approaches where some constraint is given to the MLSR model to learn expressions of a specific form, such as in the following section for $\ln(D)$, where the form $\exp(\cdot)$ is enforced. And in the physics-informed approach where the VFT model is enforced, as will be shown in Section 4.3.2. This guides the model to include these secondary parameters.

All expressions were able to include the constant k_B in order to keep units consistent, particularly in the form $\frac{\mu}{k_B T}$, which is common in Arrhenius type activation functions. However, the $\exp(\cdot)$ function is only found in Eq. 4.3, and is not in the form commonly found in Arrhenius relations ($\exp\left(\frac{-\mu}{k_B T}\right)$). Furthermore, the units of all expressions are consistent, considering that PySR assigns units to constants. For example, in Eq. 4.2 the parameters $\frac{k_B T}{\mu}$ together have no unit, while the constants 6.106×10^{-8} and 1.139×10^{-8} have the unit m^2/s, which coincides with the unit of D.

The predictive performance of the expressions can be further evaluated using parity plots, shown in Fig. 4.11. These compare the MLSR predictions (D_{MLSR}) against the KMC reference data (D_{KMC}) across different expression complexities. The expression with complexity 19 (Eq. 4.4) provides the best overall balance of accuracy and interpretability, consistent with its highest score among the expressions with loss $\leq$ 1.5× the minimum found loss. However, there is a large deviation between D_{MLSR} and D_{KMC} at lower diffusivities (Fig.4.11c). The more complex expression (Fig.4.11d) shows some improvement in predictive accuracy, but this comes at the expense of increased complexity and reduced interpretability. For the least complex expressions considered (Eqs. 4.2 and 4.3), a quite significant deviation in the low D (low T) is found (Figures 4.11a-b).

$$D = \left(\frac{k_B T}{\mu}\right) 6.106 \times 10^{-8} - 1.139 \times 10^{-8} \tag{4.2}$$

$$D = -1.148 \times 10^{-8} + \left(\frac{k_B T \exp(0.634\sigma_w)}{\mu}\right) 6.018 \times 10^{-8} \tag{4.3}$$

$$D = \left(\frac{3.079\,\mu}{k_B\,T^2} + \frac{k_B\,T}{\mu}\right) 6.539 \times 10^{-8} - 1.4529 \times 10^{-8} \tag{4.4}$$

$$D = 6.5668 \times 10^{-8} \frac{k_B}{\mu} \left(T - \frac{1.1178}{\sigma_w + 0.03781} + \frac{\mu^2}{6.5668 \times 10^{-8} k_B T^3}\right) - 1.4079 \times 10^{-8} \tag{4.5}$$

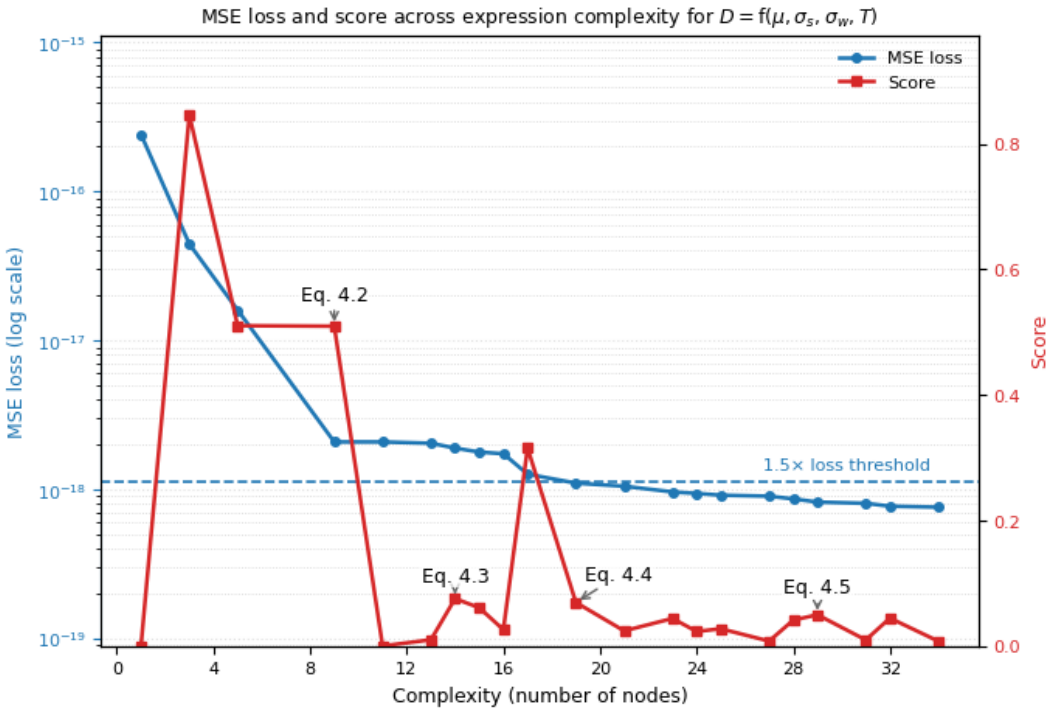


Figure 4.10: Loss (blue, log scale; left axis) and score (red; right axis) as functions of expression complexity. The blue dashed line marks the 1.5 × minimum-loss threshold used by the "best" model-selection rule. A lower loss indicates better performance, while a higher score (Eq. 4.1) is preferable to optimize accuracy vs. complexity.

Logarithm of diffusivity as a function of statistical barrier distribution parameters and temperature

For the indirect ($\ln(D)$) data-driven approach, 19 expressions in varying complexity up to 35 are learned. The loss-complexity and score-complexity curves can be seen in Figure 4.12. Four of these expressions can be seen in Eqs. 4.6, 4.7, 4.8, and 4.9, with complexities 7, 19, 24, and 29, respectively. The parity plots of these expressions can be seen in Figure 4.13. From this, the lowest complexity expression of $\ln(D)$ considered here Eq. 4.6, outperforms the highest complexity expression considered in the direct D data-driven approach (Eq. 4.5 with complexity 29). With even further improvements with higher complexity, ultimately achieving an expression that almost perfectly fits the data with test $R^2 = 0.999$ and MAE = 3.31e−10 m^2/s for the complexity 29 expression (Eq 4.9 and Figure 4.13d).

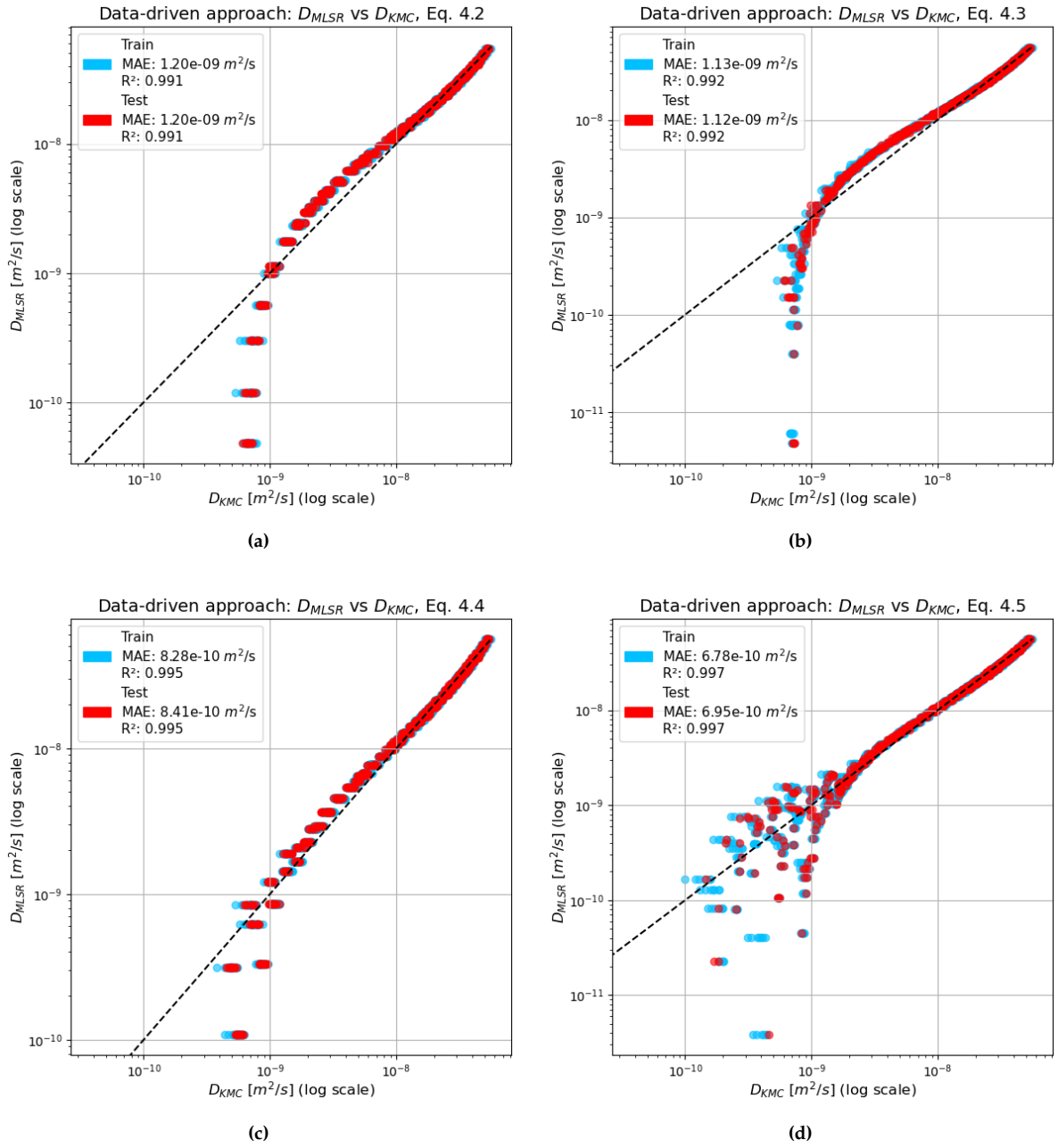


Figure 4.11: Parity plots comparing D_{MLSR} from data-driven MLSR expressions with D_{KMC} from KMC simulations at different expression complexities. (a) Eq. 4.2, complexity 9. (b) Eq. 4.3, complexity 14. (c) Eq. 4.4, complexity 19. (d) Eq. 4.5, complexity 29.

In all but the lowest complexity expression considered, the MLSR model successfully incorporated all statistical barrier distribution parameters. However, these expressions suffer from limited interpretability. Specifically, they do not clearly reveal how σ_s and σ_w influence diffusivity in any physically meaningful way. To address this limitation, the next section introduces the physics-informed approach, in which parameters from the VFT model are related to the statistical barrier distribution parameters, thereby providing more insight into the effect of σ_s and σ_w.

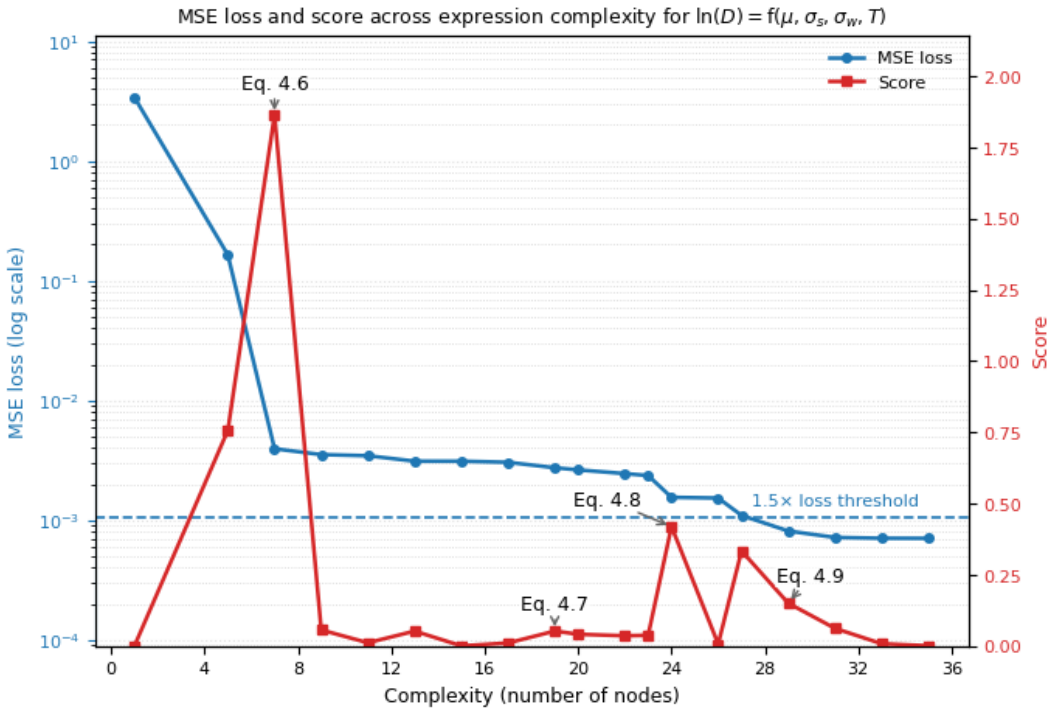


Figure 4.12: Loss (blue, log scale; left axis) and score (red; right axis) as functions of expression complexity. The blue dashed line marks the 1.5 × minimum-loss threshold used by the "best" model-selection rule. A lower loss indicates better performance, while a higher score (Eq. 4.1) is preferable to optimize accuracy vs. complexity.

$$\ln(D) = -11580\left(\frac{\mu}{T} + 0.001366\right) \tag{4.6}$$

$$\ln(D) = \left(-0.001369 + \frac{0.9426\mu}{14.26 + \dfrac{0.3238\,\sigma_s\sigma_w}{k_B} - T}\right)\frac{1}{k_B} \tag{4.7}$$

$$\ln(D) = \frac{-0.9576\mu}{k_B T}\exp\left(\frac{0.9608}{(0.1437 - \sigma_w - \sigma_s)T}\right) + \frac{\sigma_w}{0.5272} - 15.92 \tag{4.8}$$

$$\ln(D) = \frac{\mu}{T}\exp\left(-\frac{\dfrac{1.753}{T\exp(-21.09\sigma_w)\,(0.320 - \sigma_s) - 14.72} - 1.128\sigma_w}{1.029\,k_B}\right) - 15.90 \tag{4.9}$$

4.3.2. Physics-informed approach

In the physics-informed approach, the parameters $\ln(D_0)$, Q_{VFT}, and T_0 are learned as a function of the diffusion barrier distribution parameters μ, σ_s, and σ_w. For these expressions, 352 data points are used, as each model system consists of one set of VFT parameters. Just like in the data-driven approach, a 80/20 training/test split is used to evaluate if overfitting has occurred. Moreover, the same MSE loss function, binary and unary operators, number of iterations, population size, and maximum number of nested expressions are taken as the data-driven

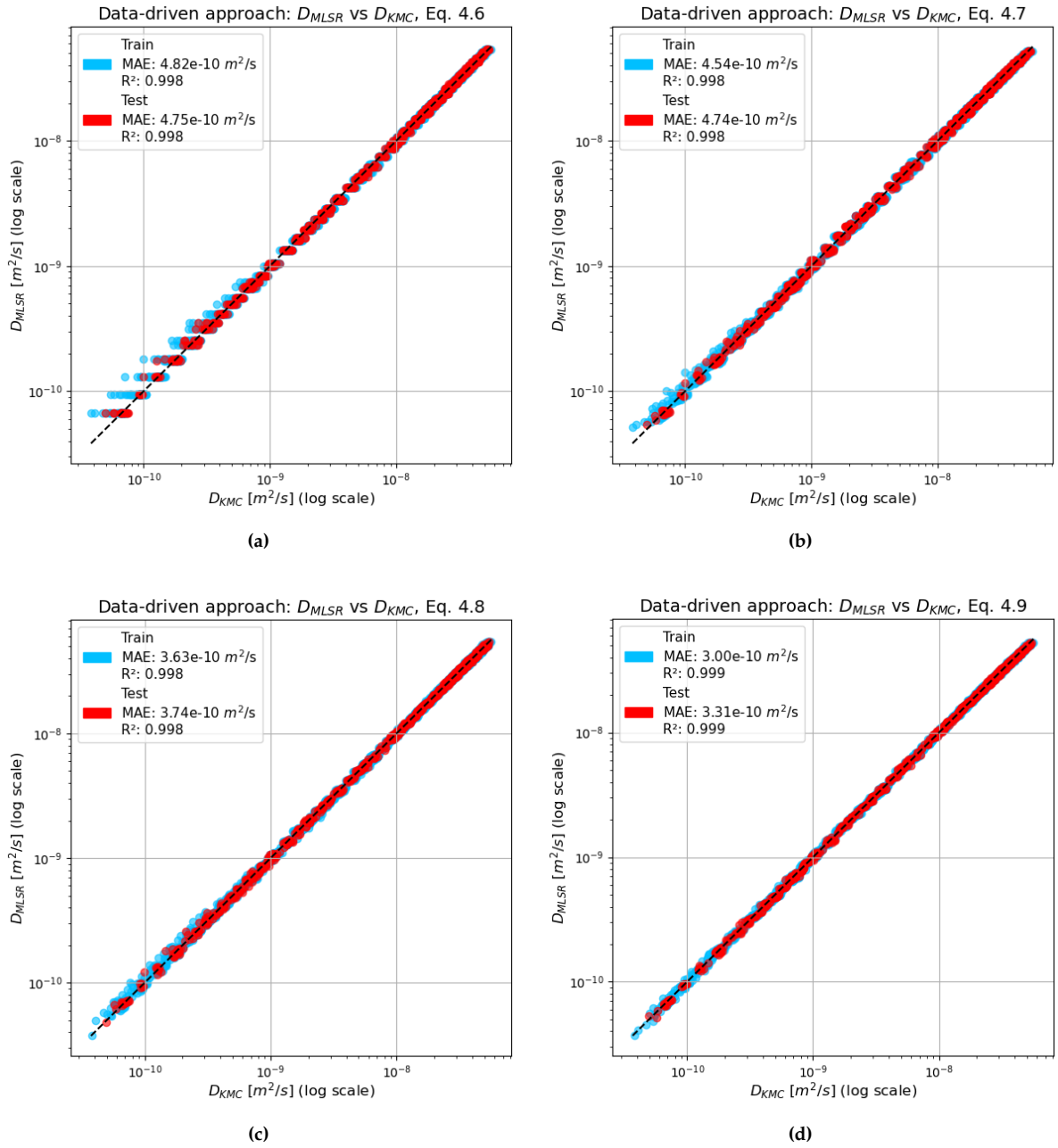


Figure 4.13: Parity plots comparing D_{MLSR} from data-driven MLSR expressions with D_{KMC} from KMC simulations at different expression complexities. (a) Eq. 4.6, complexity 7. (b) Eq. 4.7, complexity 19. (c) Eq. 4.8, complexity 24. (d) Eq. 4.9, complexity 29.

approach. However, for the physics-informed approach, a maximum amount of nodes (i.e. complexity) of 25 is taken instead of the 35 taken in the data-driven approach, this is done to ensure the expressions remain interpretable.

Expressions for the pre-exponential factor

For the pre-exponential factor D_0, two approaches are taken: (1) Learning $\ln(D_0)$ as a function of the statistical barrier distribution parameters μ, σ_s, and σ_w. (2) Learning $\ln(D_0)$ as a function of the the VFT parameters Q_{VFT} and T_0. The choice to learn $\ln(D_0)$ instead of D_0 is motivated by the Meyer-Neldel rule, in which the form $\ln(D_0) = a + bE$ appears in Arrhenius type relations, as explained in Section 4.2.

For the first approach, 16 expressions were generated by the MLSR model, the loss-complexity plot can be seen in Figure 4.14a. The equation with the highest score is Eq. 4.10, with a complexity of 5. In addition to the highest score, this expression also lies below the $\leq 1.5\times$ loss threshold and is thus selected by the "best" selection criteria. In this expression, $\ln(D_0)$ is a linear function of σ_w. Beyond this, Eq. 4.11 (complexity 7) improves the loss by including the square of σ_w. And Eq. 4.12 with complexity 10 further improves the loss by also considering σ_s. All expressions are of a similar form, in which a constant around -15.8 is further reduced by the parameters that describe the deviation in energy landscape distribution.

$$\ln(D_0) = -1.789\sigma_w - 15.847 \tag{4.10}$$

$$\ln(D_0) = -30.08\sigma_w^2 - 15.862 \tag{4.11}$$

$$\ln(D_0) = -7.366\,\sigma_w^{3/2} - 5.836\,\sigma_s^2 - 15.849 \tag{4.12}$$

For the second approach, where $\ln(D_0)$ is learned as a function of Q_{VFT} and T_0, 15 expressions were generated. Figure 4.14b shows the loss-complexity plot of this MLSR model. Eq. 4.13 (complexity 5) has the highest score and lies below the loss threshold, thus this expression is determined to be the best by the PySR model selection. In this expression $\ln(D_0)$ is a linear fit of T_0. Two more complex expression can be seen in Eq. 4.14 (complexity 8), and 4.15 (complexity 18). Eq 4.14 uses a similar form as Eq. 4.13, but the square root of T_0 is considered instead. In Eq. 4.15, $\ln(D_0)$ is instead a compound function of both T_0 and Q_{VFT}.

The second approach, with $\ln(D_0) = \mathrm{f}(Q_{VFT}, T_0)$ achieves an overall lower loss (Figure 4.14b) then the first approach with $\ln(D_0) = \mathrm{f}(\mu, \sigma_s, \sigma_w)$ (Figure 4.14a). Indicating that the pre-exponential factor is best expressed by the two other VFT parameters. Since Q_{VFT} and T_0 can both be expressed as a function of the statistical barrier distribution parameters, $\ln(D_0) = \mathrm{f}(Q_{VFT}, T_0)$ is in fact an indirect function of μ, σ_s, and σ_w, that allows for more complexity without increasing the length of the expressions. This explains the better performance of the $\ln(D_0) = \mathrm{f}(Q_{VFT}, T_0)$ MLSR expressions.

The link between D_0 and T_0 reported in Section 4.2 is reflected directly in the physics-informed results. In the first approach, the highest scoring expressions for $\ln(D_0)$ depend almost exclusively on σ_w (Eqs. 4.10–4.12). Independently, T_0 also varies mainly with σ_w (Figure 4.8). Eliminating σ_w between these relations yields a near-linear link between $\ln(D_0)$ and T_0, which matches the second approach (Eq. 4.13), with only small improvements from adding Q_{VFT} (Eqs. 4.15).

As mentioned in Section 4.2, the super-Arrhenius curvature is set mainly by σ_w through its strong link to T_0, the VFT fit then rescales D_0 to accommodate that curvature within the measured range, making D_0 and T_0 partially covariant. Together with the narrow spread of D_0 (and the lack of data at the high-temperature limit), this produces a tighter D_0–T_0 relation than D_0–Q_{VFT} and explains why the Q_{VFT} relation is less strong, despite the general Meyer–Neldel expectation that $\ln(D_0)$ scales linearly with Q_{VFT} (Section 4.2). Despite this, still the expressions that scale $\ln(D_0)$ linearly are selected as the "best", however instead of the activation function the parameters more linked to the roughness of the energy landscape (σ_w and T_0) are seen.

$$\ln(D_0) = -0.001861\,T_0 - 15.865 \tag{4.13}$$

$$\ln(D_0) = -0.0170\sqrt{T_0 + 1.913} - 15.832 \tag{4.14}$$

$$\ln(D_0) = \frac{k_B}{0.501 - 0.125\,T_0} - 0.0442\,\sqrt{Q_{VFT}\,T_0 + 0.6535} - 15.820 \tag{4.15}$$

Expressions for the effective activation energy

The MLSR model of the effective activation energy Q_{VFT} resulted in 15 expressions. The loss-complexity plot can be seen in Figure 4.14c. The expression with the highest score by far is shown in Eq. 4.16 (complexity 5), in this expression Q_{VFT} is simply μ reduced by a linear function of σ_w. However, the loss of this function is not below the 1.5 × minimum loss threshold of the model selection. Instead, the model selection picks Eq. 4.17 (complexity 8) as the best expression, in this expression μ is reduced by a more complex function of σ_s and σ_w. Beyond this, a more complex expression is shown in Eq. 4.18 (complexity 19), this expression is more complex but does not improve loss significantly.

All Q_{VFT} expression have a similar form, the effective activation energy is expressed as a reduction of the mean barrier parameter μ by the statistical parameters that determine the width (i.e. roughness) of the energy landscape. This can be understood from the fact that a Gaussian distribution of energy barriers allows the system, especially at low temperatures, to follow diffusional paths with lower than mean energy barriers. This can visually be seen by the accessible diffusion barrier spectra of Figure 4.6, in which the introduction of a wider distribution of energy barriers leads to a decrease of the mean value of the accessed barriers, which corresponds to Q_{VFT} in the VFT model. Moreover, since σ_w has a greater effect on diffusivity than σ_s, σ_w also shows a greater influence on Q_{VFT}. In fact, according to Figure 4.8, isolated σ_s should have negligible effect on Q_{VFT}. However, the expressions with more complexity (Eqs. 4.17- 4.18) show a minor influence of σ_s. More specifically in Eq. 4.18, the parameter σ_s is coupled with σ_w. In other words, the effect of σ_s is amplified by σ_w, further reducing diffusivity. This was also seen in Figure 4.5 in Section 4.1.

$$Q_{VFT} = \mu - 0.4489\sigma_w \tag{4.16}$$

$$Q_{VFT} = \mu - 0.0790\,\sigma_s - 8.445\,\sigma_w^2 \tag{4.17}$$

$$Q_{VFT} = \mu - 0.8839\,\sigma_s\,(\sigma_s + \sigma_w) - 7.3246\,\sigma_w\,(\sigma_w + 0.005729) - 0.0007064 \tag{4.18}$$

Expressions for the Vogel temperature

The MLSR model of the Vogel temperature T_0 generated 16 expressions. The loss-complexity curve can be seen in Figure 4.14d. The expression with the highest score can be seen in Eq. 4.19 (complexity 3). In this expression, T_0 is a simple linear function of σ_w. Eq. 4.20 (complexity 9) is an expression with the introduction of the σ_s parameter, this expression achieves a loss of just above the 1.5 × loss threshold. Based on the "best" model selection rule, Eq. 4.21 (complexity 12) is chosen, as it has the highest score among the expressions below the loss threshold. In this expression, T_0 is a function of all three statistical barrier distribution parameters.

T_0 reflects the non-linear slowdown of diffusion at low temperature, this slowdown occurs due to increased trapping at low temperature. Therefore, T_0 reflects the amount of H atoms trapped at low temperature. It can be seen that an increase in σ_s and σ_w increases T_0 and thus

increases the amount of trapping H atoms experience. Furthermore, an increase in μ leads to a reduction in T_0 and thus a reduction in trapping, this can be explained by the fact that the relative influence of the rough energy landscape increases upon decreasing μ. In other words, the lower the mean value, the lower the energies of low energy outliers that introduce the trapping of H atoms, which increases the trapping behaviour.

$$T_0 = 713\sigma_w \tag{4.19}$$

$$T_0 = \frac{1525\,\sigma_w^2}{\left(0.1295 - \sigma_s\right)} \tag{4.20}$$

$$T_0 = \frac{\sigma_w}{\mu}\left(33529\,\sigma_w\,(\sigma_s + 0.09207) - 44.96\right) \tag{4.21}$$

Performance of selected expressions

The physics-informed expressions further evaluated are those that achieve the highest score (Eq. 4.1) within the 1.5× threshold relative to the lowest test loss (the "best" model selection rule). These expressions (Eqs. 4.13, 4.17, and 4.21) are marked by black symbols in Fig. 4.14 and are assessed with parity plots in Fig. 4.16. Performance metrics (R^2 and MAE for both train and test) are shown in Table 4.1.

For the two approaches of $\ln(D_0)$, the model $\ln(D_0) = f(Q_{VFT}, T_0)$ provides a better match to the data than $\ln(D_0) = f(\mu, \sigma_s, \sigma_w)$ (Figs. 4.16a–4.16b and Table 4.1). When $\ln(D_0)$ is learned directly on σ_w, multiple energy landscapes that share the same σ_w return identical $\ln(D_0)$ values, producing horizontal bands in the parity plot. In contrast, the T_0-based model captures a better dependence (except near the boundary case $T_0 = 0$ K), leading to a more diverse scatter about the parity line.

For $Q_{VFT} = f(\mu, \sigma_s, \sigma_w)$, Eq. 4.17 achieves the highest score among models within the loss threshold. Its parity plot (Fig. 4.16c) shows excellent agreement, with test $R^2 = 0.986$ and MAE = 2.10 meV (Table 4.1). For $T_0 = f(\mu, \sigma_s, \sigma_w)$, Eq. 4.21 is selected by the same criterion, it has the highest score among those within the loss threshold. Furthermore, the corresponding parity plot (Figure 4.16d) shows strong agreement.

Finally, combining the VFT model (Eq. 3.8) with the MLSR–derived expressions from Eqs. 4.13, 4.17, and 4.21 yields a single predictive physics-informed model for diffusivity. Its performance, benchmarked against the KMC diffusivities across all temperatures, is shown in Fig. 4.15 and summarised in the last row of Table 4.1 (test $R^2 = 0.999$, MAE = 3.33e−10 m^2/s). The direct data-driven approach achieves test $R^2 = 0.995$ and MAE = 8.41e−10 m^2/s for the expression best scoring among those with a loss threshold $\leq 1.5\times$ the minimum loss (complexity 19) (Eq. 4.4 and Figure 4.11c) and test $R^2 = 0.997$ with MAE = 6.95e−10 m^2/s for the larger complexity 29 expression (Eq. 4.5 and Figure 4.11d). The indirect ($\ln(D)$) data-driven approach significantly improves the predictive accuracy by enforcing the form $\exp(\cdot)$ onto the MLSR model. The lowest complexity model considered (Eq. 4.6, Figure 4.13a, complexity 7) achieves test $R^2 = 0.998$ and MAE = 4.75e−10 m^2/s, outperforming the most complex direct data-driven expression. Increasing complexity (Eq. 4.9, Figure 4.13d, complexity 29) further improves the predictive accuracy to test $R^2 = 0.999$ and MAE = 3.31e−10 m^2/s, which is comparable to the physics-informed approach.

The main error of the direct data-driven approach lies in the low diffusivity range, where the diffusivity prediction deviates strongly from the dataset, this is apparent from the parity plots

of Figure 4.11. The direct data-driven approach is only able to significantly include σ_w when very large expressions (Eq. 4.5) are considered, and even then the influence of σ_w is minimal and σ_s is not included all together. In contrast, the indirect $\ln(D)$ model was able to more significantly include the σ_s and σ_w parameters and accurately describe the dataset by forcing the expressions to have the form $\exp(\cdot)$, which is found in Arrhenius-type relations. However, the drawback of this method is that the input parameters give no insight into the way σ_s and σ_w influence diffusivity in a physically meaningful way. Whereas the physics-informed approach is able to reproduce the full diffusivity range while also providing meaningful context about the diffusion behaviour though D_0, Q_{VFT}, and T_0.

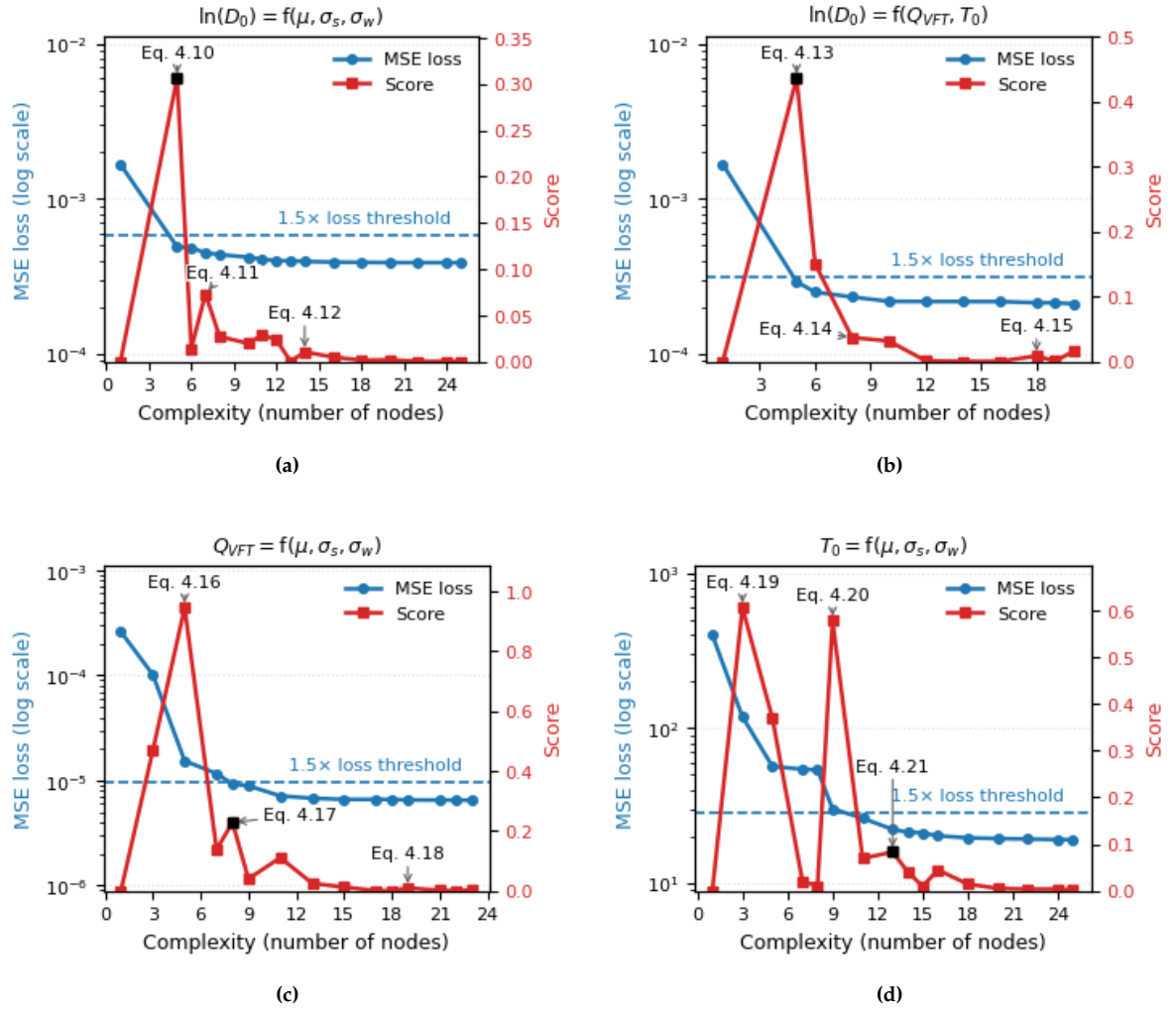


Figure 4.14: Loss–complexity curves for the physics-informed approach. Plots (a)–(d) show the MSE loss (blue, log scale; left axis) and the PySR score (red; right axis) as functions of expression complexity. The dashed blue line indicates the 1.5× loss threshold relative to the lowest-loss model used by the "best" selection rule. The black marker denotes the expression selected by "best" (highest score among models with loss ≤ 1.5× the minimum). Labelled points (Eq. 4.6–4.17) correspond to the expressions discussed in the text. The score is defined in Eq. 4.1.

4.4. Application on complex BCC alloys using uMLIPs

After establishing the relationship between statistical barrier distribution parameters and H diffusivity in model systems, the next step is to evaluate whether this framework can be meaningfully applied to real materials. In this section, the MLSR expressions developed from KMC simulations are applied to a set of chemically complex BCC alloys. To do so, diffusion

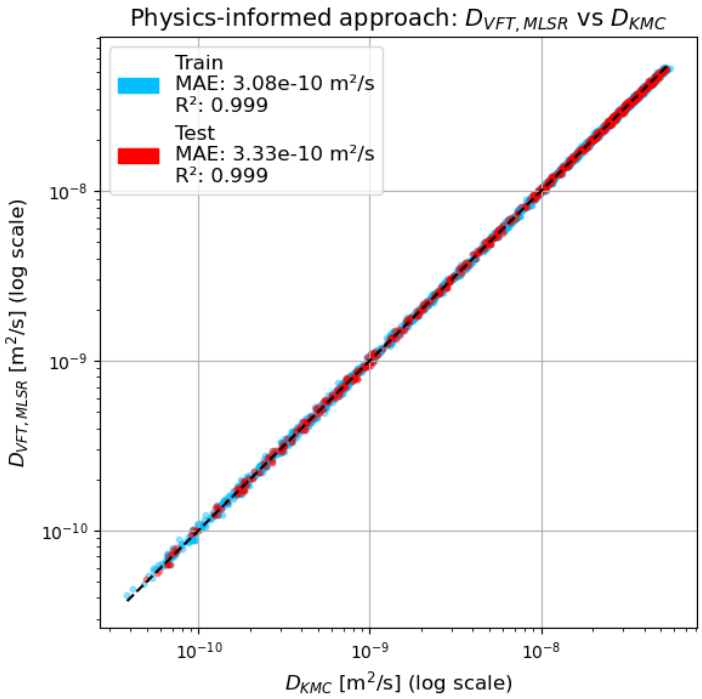


Figure 4.15: Parity plot of the physics-informed $D_{VFT,MLSR}$, where $D_{VFT,MLSR}$ is calculated (Eq 3.8) from VFT parameters $D_{0,MLSR}$, $Q_{VFT,MLSR}$, and $T_{0,MLSR}$ using Eq. 4.13, 4.17, and 4.21

Table 4.1: Performance metrics of MLSR expressions for the data-driven and physics-informed approaches. While certain MLSR expressions are written in terms of $\ln(D)$ or $\ln(D_0)$, all reported performance metrics are evaluated and presented in terms of D and D_0, in order to maintain consistency.

Equation(s)	Output	Input	Training		Test	
			R^2	MAE	R^2	MAE
Data-driven approach						
4.4	D [m^2/s]	μ [eV], T [K]	0.995	8.28e−10 m^2/s	0.995	8.41e−10 m^2/s
4.5	D [m^2/s]	μ [eV], T [K]	0.997	6.78e−10 m^2/s	0.997	6.95e−10 m^2/s
4.6	$\ln(D)$ [-]	μ [eV], T [K]	0.998	4.82e−10 m^2/s	0.998	4.75e−10 m^2/s
4.9	$\ln(D)$ [-]	μ [eV], σ_s [eV], σ_w [eV], T [K]	0.999	3.00e−10 m^2/s	0.999	3.31e−10 m^2/s
Physics-informed approach						
4.10	$\ln(D_0)$ [–]	σ_w [eV]	0.731	2.07e−09 m^2/s	0.797	1.95e−09 m^2/s
4.13	$\ln(D_0)$ [–]	T_0 [K]	0.825	1.72e−09 m^2/s	0.827	1.89e−09 m^2/s
4.17	Q_{VFT} [eV]	μ, σ_s, σ_w [eV]	0.989	2.02e−03 eV	0.986	2.10e−03 eV
4.21	T_0 [K]	μ, σ_s, σ_w [eV]	0.945	3.62 K	0.964	3.13 K
3.8, 4.13, 4.17, 4.21	D [m^2/s]	D_0, Q_{VFT}, T_0	0.999	3.08e−10 m^2/s	0.999	3.33e−10 m^2/s

barrier distributions are obtained through CI-NEB calculations using uMLIPs. From these distributions, the statistical model parameters (μ, σ_s, and σ_w) are obtained, allowing the MLSR expressions to predict temperature-dependent diffusivities. The accuracy of these predictions

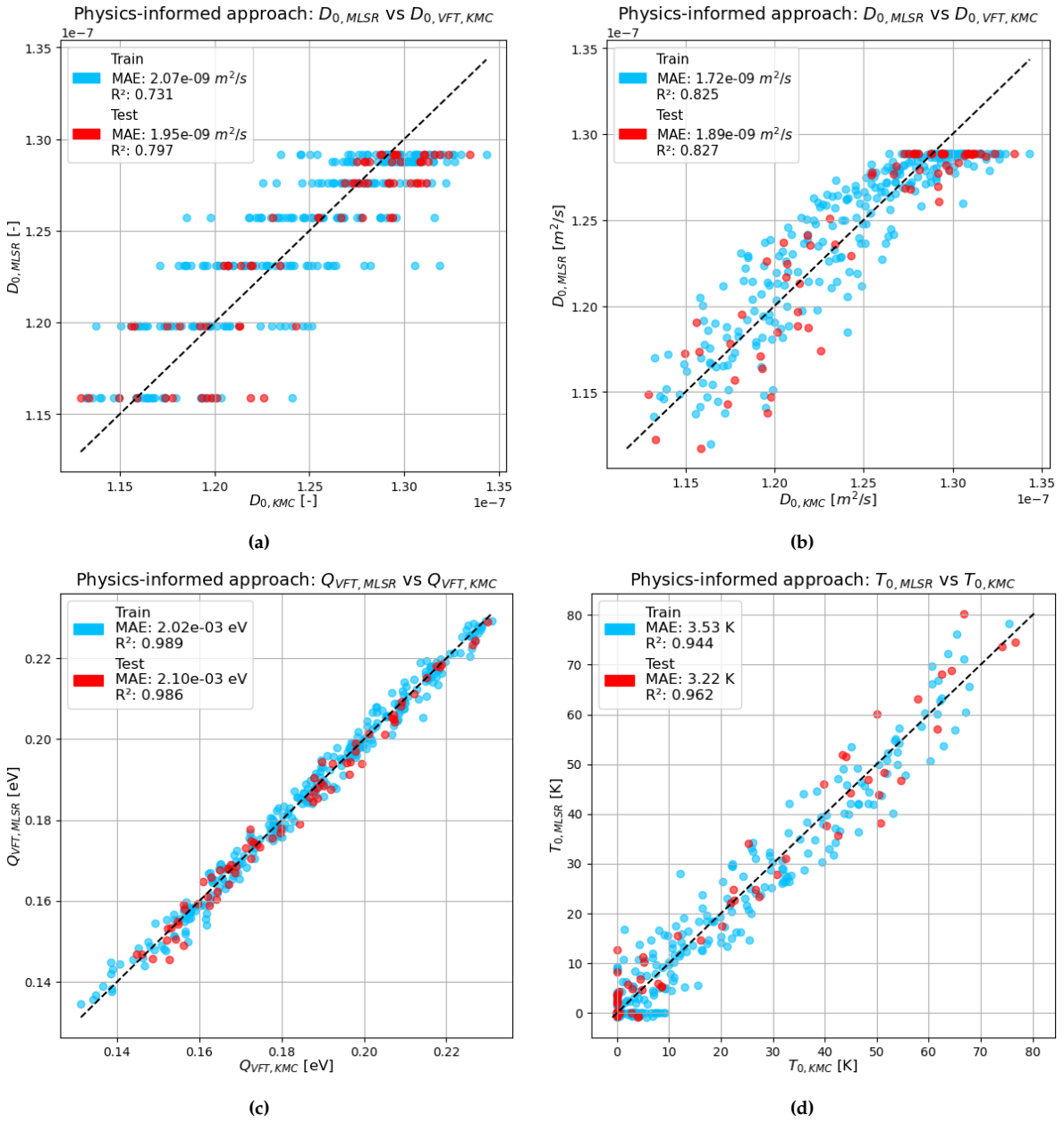


Figure 4.16: Parity plots of physics-informed MLSR expressions for VFT model parameters. (a) D_0 as a function of σ_w (Eq. 4.10). (b) D_0 as a function of T_0 (Eq. 4.13). (c) Q_{VFT} as a function of μ, σ_s, and σ_w (Equation 4.17). (d) T_0 as a function of μ, σ_s, and σ_w (Equation 4.21). Although Eqs. 4.10 (a) and 4.13 (b) are expressed in terms of $\ln D_0$, the corresponding parity plots are presented in terms of D_0 for clarity.

is then assessed through direct MD simulations. This evaluation, applying CI-NEB data to the MLSR expressions and validating against MD, serves to assess the limitations and practical applicability of the developed methodology for complex alloy design.

All CI-NEB calculations and MD simulations are performed using the GRACE-FS-OMAT uMLIP implemented in LAMMPS [68]. For comparison, two other models are benchmarked against CI-NEB calculations: GRACE-FS-OAM, which shares the same architecture and OMat24 pretraining as GRACE-FS-OMAT but is further fine-tuned on the sAlex and MPTrj datasets [66, 67]; and GRACE-2L-OMAT, which uses a different (two-layer) architecture while being trained on the same OMat24 dataset as GRACE-FS-OMAT without further fine-tuning. Among

these, FS-OMAT and FS-OAM have comparable computational cost, but FS-OMAT delivers higher accuracy (as will be discussed in Section 5.4). Moreover, the 2L-OMAT potential is too computationally large for routine CI-NEB (and MD) use and was therefore applied only to a small subset of CI-NEB barriers for evaluation.

4.4.1. CI-NEB calculations and application of MLSR expressions

The diffusion barrier distributions of 12 equimolar BCC alloys were calculated (Figure 4.17). The equimolar NbTi, TaTi, NbTaTi, and MoNbTaTi alloys (highlighted in red in Figure 4.17) are selected to be evaluated through MD simulations, as they have the most appropriate Gaussian-like barrier distributions. The results of the CI-NEB calculations of the four selected complex BCC alloys can be seen in Figure 4.18, in which the coloured barrier distributions are the CI-NEB calculations. In the MoNbTaTi-H system of Figure 4.18d some barriers with zero energy can be seen, these represent H diffusion paths where the initial and/or final T-site is unstable, causing the H atom to relax to the same position and therefore a zero-energy barrier is obtained. This is a limitation of the uMLIP and will be further discussed in Section 5.4.

From the obtained barrier distributions using CI-NEB, the three model system parameters μ, σ_s, and σ_w are extracted using the reference system as described in Section 3.1.3, the grey overlay of barriers in Figure 4.18 represents the barrier distributions using the reference system. The values obtained using the reference system and the corresponding parameters of the VFT model using the physics-informed MLSR expressions (Eqs. 4.13, 4.17, 4.21) can be seen in Table 4.2. The Arrhenius plot of the four alloys using the VFT parameters can be seen in Figure 4.19a.

Table 4.2: Model system parameters obtained from energy barrier distribution through CI-NEB calculations and the associated VFT model parameters ($\ln(D_0)$ is dimensionless).

Alloy system	μ [eV]	σ_s [eV]	σ_w [eV]	$\ln(D_0)$	Q_{VFT} [eV]	T_0 [K]
NbTi-H†	0.145	0.024	0.042	-15.93	0.128	34.3
TaTi-H†	0.156	0.035	0.053	-15.98	0.130	61.4
NbTaTi-H	0.166	0.033	0.045	-15.94	0.146	39.0
MoNbTaTi-H‡	0.169	0.036	0.063	-16.02	0.133	84.1

† μ outside the dataset range ([0.16, 0.23] eV) .

‡ σ_w outside the dataset range ([0, 0.06] eV).

4.4.2. Evaluation with MD simulations

The results of the MD simulations for validation of the four complex BCC alloys across 800, 1000, 1200, and 1400 K can be seen in Figure 4.19, in which (a) the Arrhenius plot and (b) the parity plot of D_{MLSR} and D_{MD} are shown. Additionally, MD simulations at 600 K were done, however, within the simulation time of 2.5 ns, the MSD did not grow sufficiently large in order to obtain a reliable diffusion coefficient, and are therefore omitted from the results. The MSD is determined through the method described in Section 3.1.2, which is the same method used for the KMC simulations. The only difference is that a time-based stopping criteria (2.5 ns) is used instead of a MSD-based stopping criteria. Each simulation was done 5 times using different initial configurations with random element distribution, the error bars in Figure 4.19 represent the standard deviation between simulation results. The range of temperatures of the KMC database on which the MLSR expressions are learned is 350 to 2000 K. However, for MD simulations, obtaining reliable diffusivity data below 600 K is computationally expensive due to the extensive simulation time required to obtain sufficiently large MSD values, particularly in systems with rough energy landscapes due to frequent trapping events. On the other end, temperatures above 1400 K are expected to reliably follow Arrhenius behaviour and thus offer

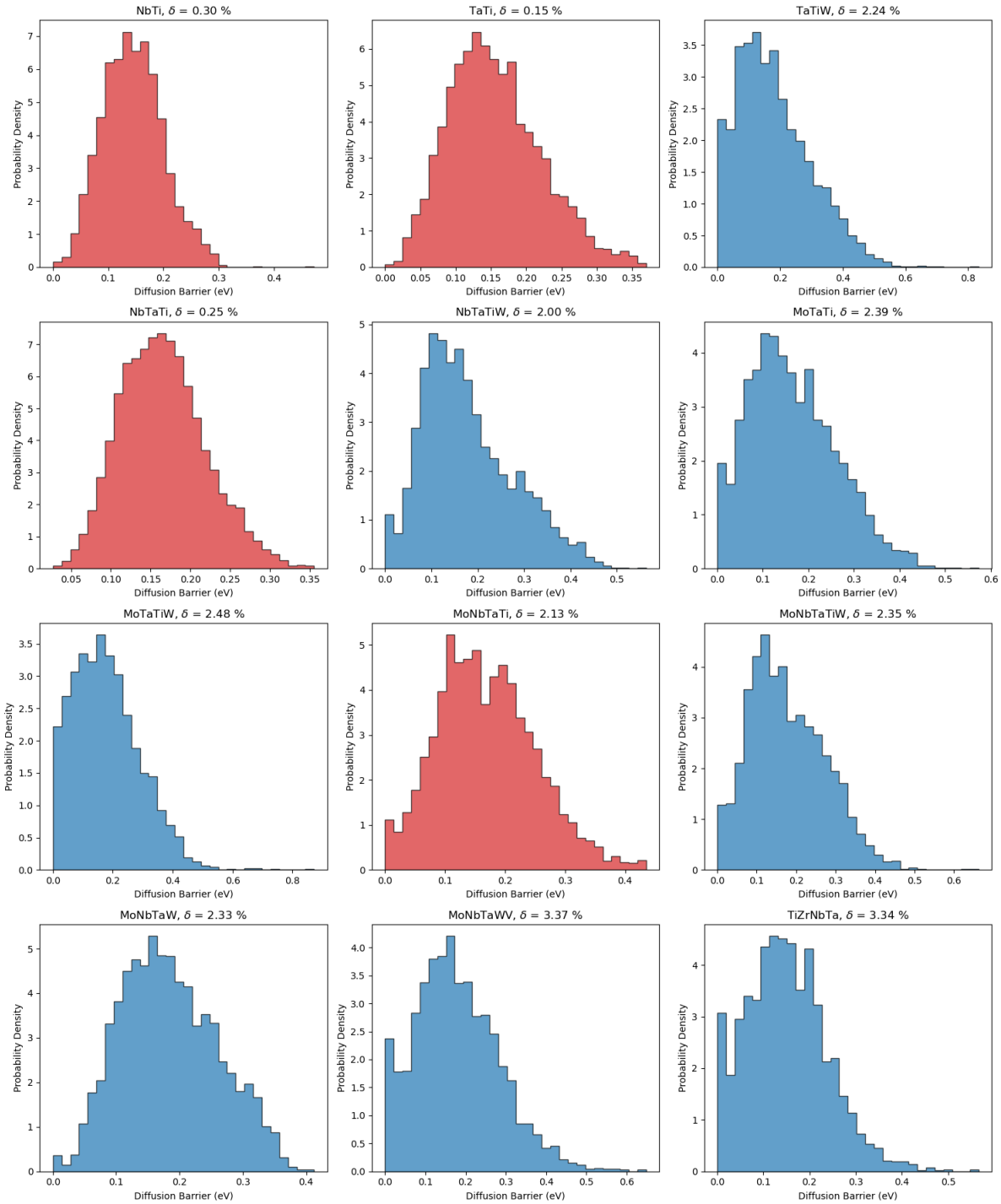
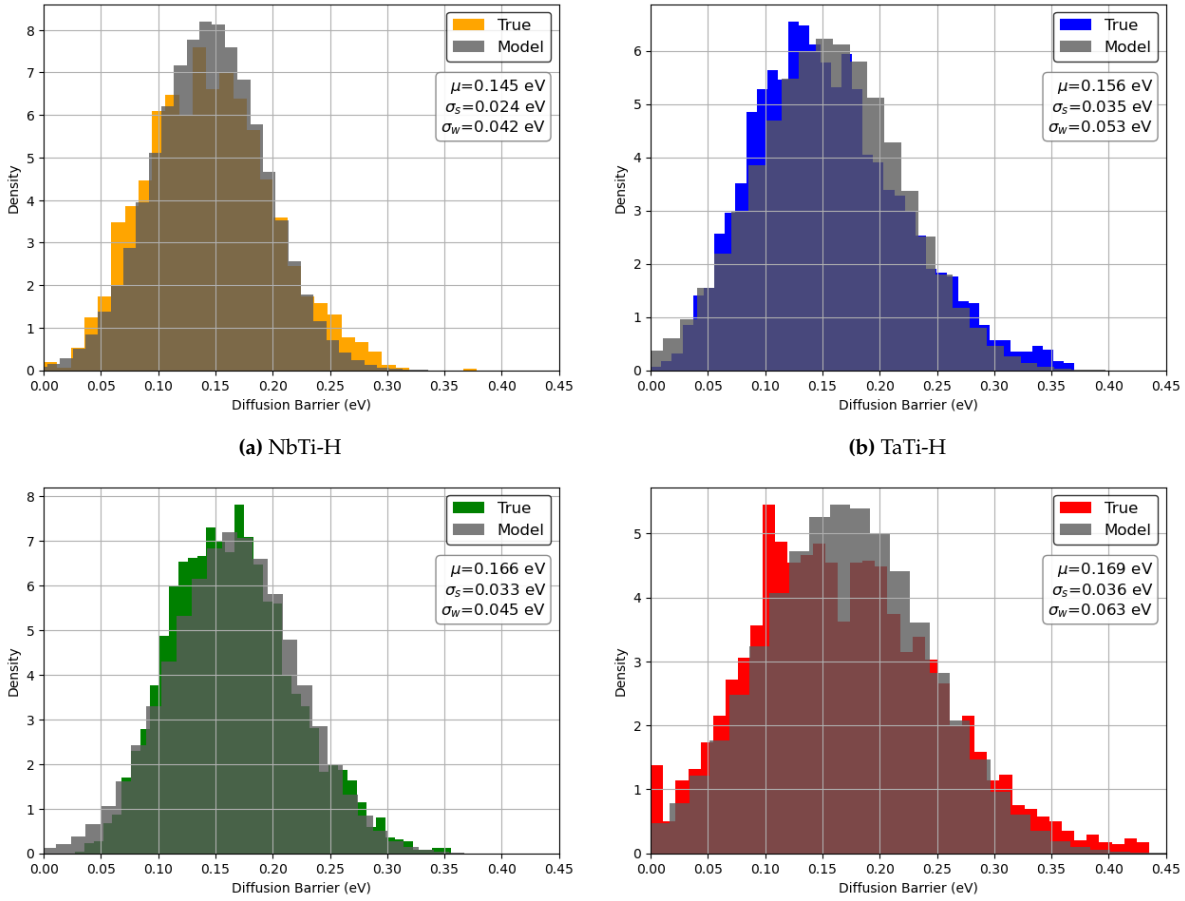


Figure 4.17: Diffusion barrier distributions of 12 BCC MPEAs, obtained through 1536 CI-NEB calculations per alloy using the GRACE-FS-OMAT uMLIP. The red highlighted barriers have the best Gaussian-like distribution and are used for further evaluation. The parameter δ is the lattice strain.

limited additional insight, making them less critical for evaluation.

From the comparison between the predictions using MLSR and the MD simulations, it can be seen that the MLSR expressions overestimate the diffusion coefficient. The MAE between MLSR and MD is on the order of 10^{-8} m^2/s for all temperatures (Table 4.3), but this MAE is relatively large at low temperatures, where the diffusivity decreases due to reduced thermal



(a) NbTi-H **(b)** TaTi-H

(c) NbTaTi-H **(d)** MoNbTaTi-H

Figure 4.18: Diffusion barrier distributions of (a) NbTi-H, (b) TaTi-H, (c) NbTaTi-H, and (d) MoNbTaTi-H. True barriers (coloured) are obtained through 1536 NEB calculations; the model system (gray) is obtained by fitting μ, σ_s, and σ_w to the respective true barriers.

energy. From Figure 4.19 it can be seen that the MLSR expressions consistently rank the alloys from highest to lowest diffusivity as: NbTi - TaTi - NbTaTi - MoNbTaTi. A similar trend is found in the MD simulations, however, for the MD simulations the TaTi and NbTaTi alloys lie very close together, within the margin of error given by the error bars. For the MLSR expressions, the MoNbTaTi and NbTaTi alloys lie closer together.

The Gaussian approximation of the distribution of energy barriers leads to more low energy barriers then the barriers directly calculated through CI-NEB, as can be seen in Figure 4.18. Low energy barriers are needed for trapping, but these should be relatively rare and surrounded by high energy barriers to reduce diffusivity. If there are too many low energy barriers, diffusivity is enhanced because the trapping effect is reduced as there are more pathways for the H atom to escape the trapping region. Consequently, diffusivity obtained through the MLSR expressions is greater then those obtained through MD simulations, especially at lower temperatures where the trapping effect dominates diffusivity, as also seen in Figure 4.5.

As a consequence of the reduced trapping effect of barrier distributions that are described using a Gaussian distribution, the super-Arrhenius behaviour, which is represented by the parameter T_0, is reduced. There is still a non-linearity between $\ln(D)$ and T^{-1}, but since the diffusivity is especially overestimated al low temperature compared to high temperature, the relative decrease of the diffusivity over temperature is reduced. A previous study determined that, for

the equimolar MoNbTaW alloy, T_0 = 163.5 K, based on a MLIP trained via DFT, in combination with CI-NEB calculations and KMC simulations [42]. Using the reference system as described in Section 3.1.3 on the barriers of MoNbTaW, μ = 0.203 eV, σ_s = 0.059 eV, and σ_w = 0.030 eV are found. Using Equation 4.21 to calculate the T_0 parameter, a value of T_0 = 18.2 K is obtained.

This discrepancy in the T_0 value highlights the limitations of the Gaussian approximation at low temperatures. Although the MLSR method yields a low global MAE of 1.44e−08 m^2/s (Table 4.3) compared to MD, which is sufficient to relatively quickly determine if an alloy may exhibit low H diffusivity at high temperature, the deviation at low temperature is too significant to effectively capture super-Arrhenius behaviour. Therefore, this model does not provide sufficient accuracy for reliable determination of diffusivity at low temperatures.

The mean diffusion barrier values for NbTi and TaTi (Table 4.2) are lower than the range used in the KMC simulations (Table 3.1), which provided the database from which the MLSR expressions were learned. Moreover, the MoNbTaTi system exhibits a σ_w value beyond that dataset range (Table 3.1). Nevertheless, the MAEs between the MLSR expressions and the MD simulations are comparable across all four alloy systems (Table 4.3), indicating that the MLSR expressions maintain similar predictive accuracy even outside the training range of the KMC-derived database.

Table 4.3: Mean absolute errors (MAE) of D in m^2/s between MLSR predictions and MD simulations for each alloy system at different temperatures.

Alloy system	800 K	1000 K	1200 K	1400 K	All temperatures
NbTi-H†	1.23e−8	1.47e−8	1.59e−8	1.42e−8	1.43e−8
TaTi-H†	1.17e−8	1.46e−8	1.73e−8	1.64e−8	1.50e−8
NbTaTi-H	0.93e−8	1.20e−8	1.49e−8	1.26e−8	1.22e−8
MoNbTaTi-H‡	1.16e−8	1.65e−8	1.74e−8	1.60e−8	1.54e−8
All alloys	1.12e−8	1.45e−8	1.64e−8	1.48e−8	1.42e−8

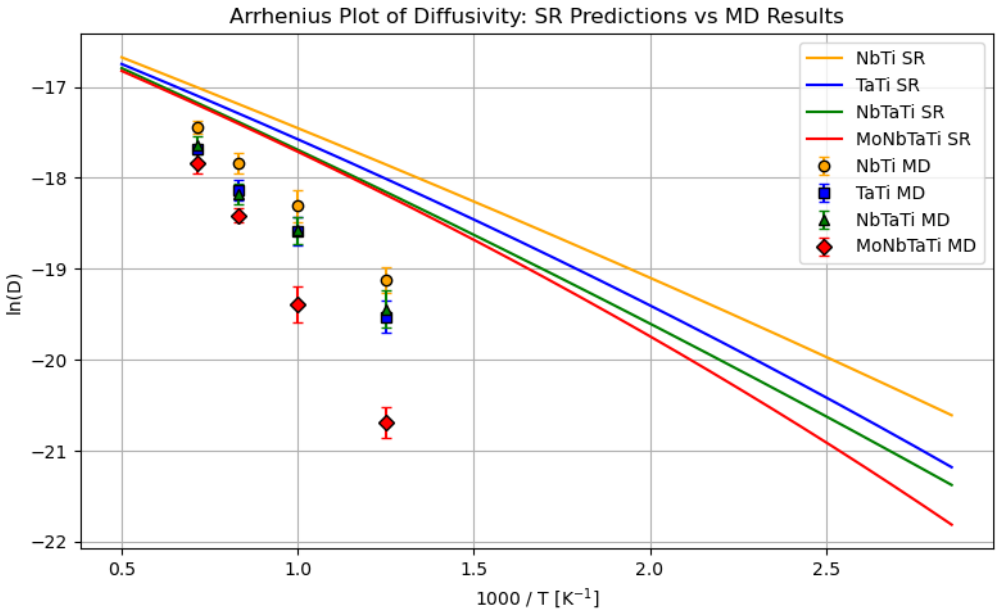
† μ outside the dataset range ([0.16, 0.23] eV).

‡ σ_w outside the dataset range ([0, 0.06] eV).

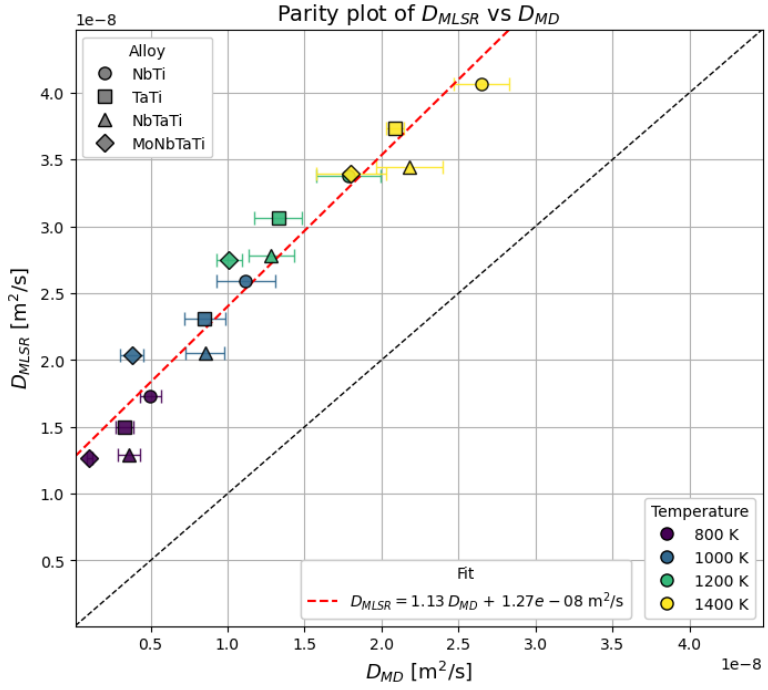
The H trajectories from the MD simulations are further analysed to correlate the local atomic environment with diffusion behaviour. Trapping events are identified by applying a displacement threshold of 1 Å over 0.5 ps time intervals: if the H atom does not move at least 1 Å within that period, it is considered trapped. This threshold was selected based on an analysis of the number of trapped frames as a function of the displacement criterion. A value of 1 Å was found to be sufficiently large to disregard atomic vibrations as meaningful motion, yet small enough to exclude transitions between distinct interstitial sites.

To determine the local atomic environment during trapping, specifically the four metallic neighbours forming the T-site, a cut-off radius of ∼ 2.4 Å was used. This value was chosen by plotting the frequency of nearest-neighbour counts as a function of cut-off radius and selecting the radius at which the number of frames with four nearest neighbours was maximised (≈ 70% of trapped frames having four nearest neighbours).

The elemental composition of the local atomic environment surrounding the trapped H atom shows consistent trends across all alloys and temperatures as can be seen in Figure 4.20. For the quaternary MoNbTaTi alloy, Ti dominates the trapping environment, accounting for over 60% of nearest neighbours at all temperatures. The influence of Ti slightly decreases at higher temperatures, indicating that the local environment of trapping sites becomes more varied at higher temperatures. Moreover, Nb and Ta account for approximately 15% of nearest



(a)



(b)

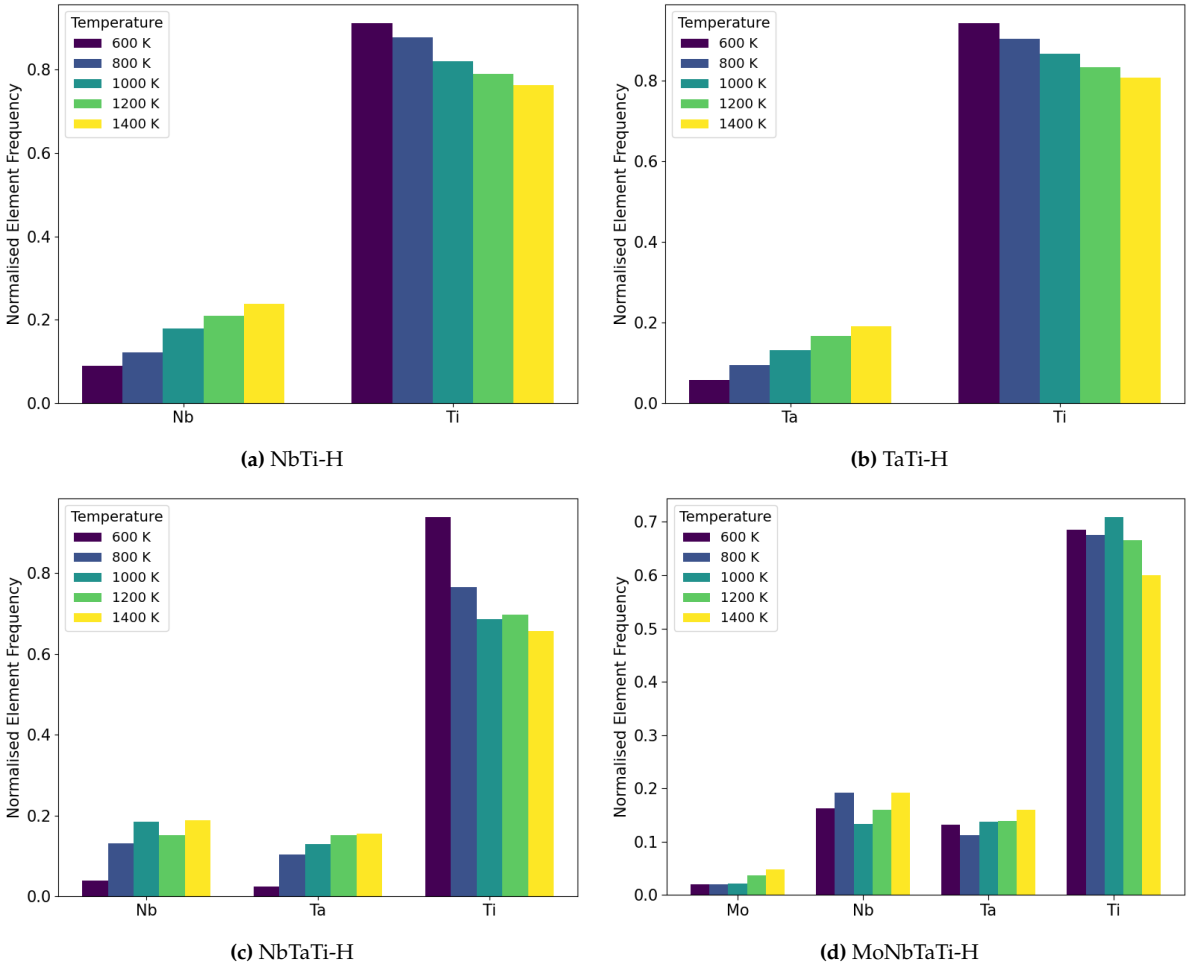
Figure 4.19: Evaluation of results obtained from MLSR with CI-NEB and MD. (a) ln(D) vs 1000/T from MLSR and MD. (b) Parity plot of D_{MLSR} and D_{MD} with a linear fit $D_{MLSR} = 1.13 D_{MD} + 1.27\text{e}{-}08$ m^2/s.

neighbours each. In contrast, Mo appears a negligible amount.

For the ternary NbTaTi alloy, Ti again plays a dominant role, although slightly more evenly distributed at elevated temperatures. For the binary alloys, NbTi and TaTi, the strong trapping effect of Ti is reinforced, with a slightly more favourable trapping effect of Nb over Ta. The strong trapping tendency is in line with prior literature, which shows that Ti has strong affinity towards H [32]. Moreover, it is known that both Nb and Ta are H favouring elements, while

Mo is an H repelling element [42]. Overall, the MD results suggest that the relative influence on trapping follows the trend: Ti ≫ Nb > Ta ≫ Mo, which is in good agreement with prior research [32, 42].

Figure 4.20 plots the normalised element frequency, thus the occurrence of the elements divided by the total number of trapped frames. Figure 4.21 instead plots the total number of trapped frames for each alloy over temperature. The decrease in trapped frames further emphasizes that trapping reduces at high temperature due to increased thermal activation energy.



(a) NbTi-H

(b) TaTi-H

(c) NbTaTi-H

(d) MoNbTaTi-H

Figure 4.20: Normalised element frequency of the local atomic environment of trapping sites of (a) NbTi-H, (b) TaTi-H, (c) NbTaTi-H, and (d) MoNbTaTi-H.

4.4.3. Computational cost

Comparing the computational cost of the MLSR method with MD simulations, each MD simulation required ~ 1024 core-hours (64 h on 16 cores). Because MD must be run at multiple temperatures to obtain diffusivity as a function of temperature, the four temperatures used here (800, 1000, 1200, 1400 K) exceed 4000 core-hours per alloy, not counting additional runs for statistical averaging. In contrast, the CI-NEB calculations used to input the MLSR expressions are performed once per alloy, the barriers are temperature independent, and required only ~ 12.5 core-hours (2.5 h on 5 cores). After this one-time cost, the MLSR model can evaluate diffusivity between 350 and 2000 K (within the dataset) without any additional cost. This represents a >99.5% reduction in computational cost for the four temperature case, and even more when considering more temperatures.

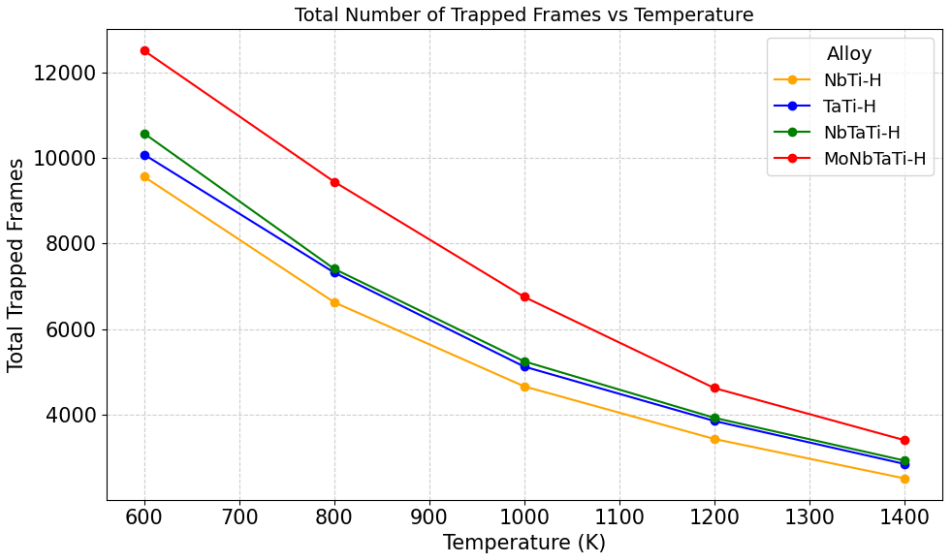


Figure 4.21: Total number of trapped frames vs temperature for the four alloys.

5

Discussion

5.1. Hydrogen diffusion behaviour from KMC

The effect of σ_s and σ_w on diffusivity is complex. No studies of interstitial H diffusion have been performed using this reference system. The explanations to the phenomena presented in Section 4.1 are plausible and motivated by the obtained KMC data, but they cannot be confirmed as the definitive mechanisms.

Although the reference system is a practical way to describe the rough energy landscape, it oversimplifies the true complex interactions of H with the rough energy landscape. More specifically, with this reference system, only the local character of individual jumps is considered. Thus, the symmetric and antisymmetric components of the diffusion barrier might be a better description than σ_w and σ_s, as variability in s and w implies a global character. To include the global character of the rough energy landscape, a more accurate description of the distribution of s and w could be made by including μ_w. In that case, a set reference system that is globally tied to the system has to be made, instead of defining the zero energy reference (μ_w) for each jump. Moreover, the Gaussian approximation, which leads to an overestimation of the low energy barriers, underestimates the amount of trapping at low temperature. The extra low energy barriers allow the H atom to prematurely leave certain trapping regions. This leads to weakened super-Arrhenius behaviour when compared to the true energy barrier (Section 4.2). Extra parameters that describe the rough energy landscape, specifically the low energy tail, are expected to increase the amount of trapping, thus increasing super-Arrhenius behaviour and improving accuracy compared to MD simulations (Section 4.4.2).

The Meyer-Neldel rule (Section 4.2), which states that $\ln(D_0)$ and Q_{VFT} are related in a predictable way ($\ln(D_0) = a + bQ_{VFT}$) [75], is only weakly present in this dataset (Figure 4.9). Instead, such a relation is found with $-T_0$ (Eqs. 4.13 and 4.20). This is attributed to D_0 acting more as a fitting parameter to accommodate the curvature introduced by T_0 instead of a true physics determined relation through the Meyer-Neldel rule. An improvement to this could possibly be done by including higher temperature KMC simulations (up to 5000 K), even exceeding the melting temperature of the alloys, in order to better represent the true D_0 at the high temperature limit.

5.2. Machine learning symbolic regression models

MLSR is a powerful tool to create expressions that follow a dataset. PySR outputs multiple expressions with varying complexity. With this, a trade-off can be made between accuracy

and complexity. Based on the score evaluation metric (Eq. 4.1), both the loss and complexity can be optimised simultaneously, leading to an expression with minimal complexity but high accuracy. Relying on just the score metric is not sufficient, expressions should be manually inspected to ensure an expression with a sufficiently low loss for the application is chosen. The 1.5× minimum loss threshold used in this work leads to expressions that remain interpretable while adequately describing the dataset.

The direct data-driven approach achieves high global accuracy, but systematically misses the low D (low T) regime. More specifically for the lowest complexity data-driven expression (Eq. 4.2), with test $R^2 = 0.991$ and MAE = 1.20×10^{-9} m^2/s, the diffusivity is underestimated by an order of magnitude (Figure 4.11a). Increasing complexity (Eq. 4.4), with test $R^2 = 0.995$ and MAE = 8.41×10^{-10} m^2/s, improves the fit at low D, but comes at the cost of increased expression complexity, and still does not fully capture the extreme low D data. The deviation in the low diffusivity range of the data-driven approach can be attributed to the difficulty it had finding the $\exp(\cdot)$ form. In addition, it had difficulty incorporating σ_s and σ_w into the expressions, σ_s was never found in the expressions, while σ_w only showed up in meaningful capacity at high complexities (Eq. 4.5). Across the dataset, D is largely set by T and μ, consequently, the direct data-driven approach reaches a local minimum.

A solution to this local minimum problem is to constrain the model to follow some form. For the indirect data-driven approach, this was done by learning $\ln(D)$ instead of D, thereby effectively forcing the expressions to follow $\exp(\cdot)$. This significantly improves predicative accuracy and is able to include the secondary parameters σ_s and σ_w. However, this indirect data-driven approach leads to expressions that are hard to interpret, and do not provide additional insights into H diffusion. In contrast, the physics-informed approach, where the VFT model (Eq. 3.8) is learned as a function of the statistical parameters, does give meaningful insight through the VFT parameters D_0, Q_{VFT}, and T_0. This approach was able to achieve test $R^2 = 0.999$ and MAE = 3.33×10^{-10} m^2/s, almost perfectly fitting the dataset, just as the direct data-driven approach.

A comparison of the direct data-driven approach (Eq. 4.5, complexity 29), the indirect data-driven approach (Eq. 4.9, complexity 29) and the physics-informed approach (Eqs. 4.13, 4.17, 4.21) with the KMC data can be seen in Figure 5.1, showing that both the indirect data-driven approach and the physics-informed approach almost perfectly match the KMC data across all temperatures of the dataset (350-2000 K).

When validated against MD on real alloys, the physics-informed model maintains similar MAEs even when μ or σ_w fall outside the KMC training ranges (Table 4.3, μ: NbTi, TaTi, and σ_w: MoNbTaTi). This suggests that extrapolation in μ (and to a smaller extend σ_w) is possible. Nevertheless, a broader evaluation across more systems is needed to verify the extrapolative capability.

5.3. Extending the framework to FCC systems

The framework of this thesis is applied to BCC systems. For extending this framework to FCC systems, a similar approach can be taken. The reference system used in this work would need to be adjusted to reflect the diffusional paths of FCC systems. The diffusional paths of pure FCC Fe can be seen in Figure 5.2. In contrast to the T-T path of BCC lattices, in FCC lattices the most energetically favourable is the O-T-O path [54]. Thus, to apply the framework to FCC systems, a new reference system has to be created. This reference system should incorporate the mean values and deviations of the saddle points and the energies of the metastable T-sites, as well as the well-energies of the O-sites.

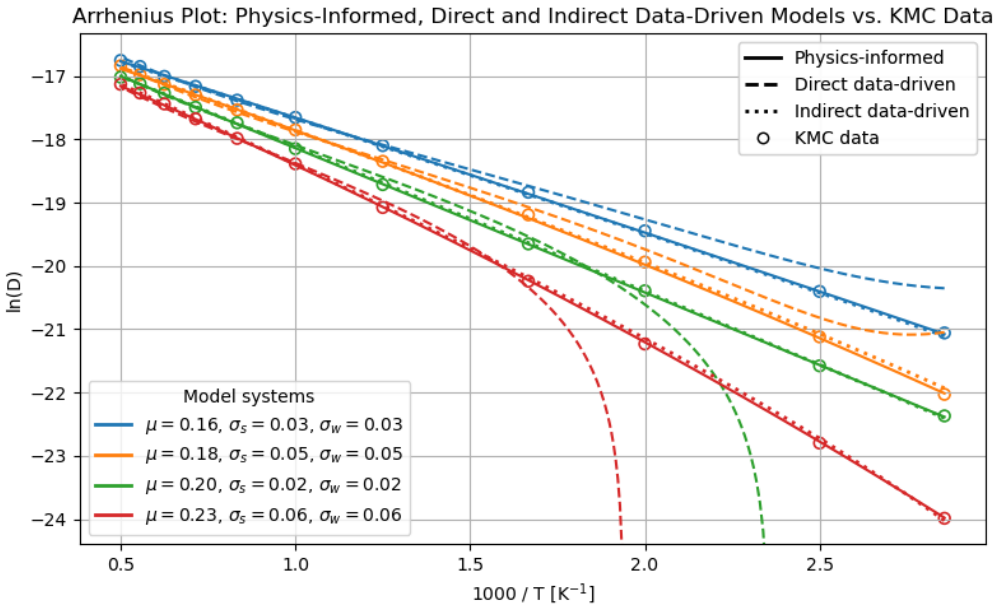


Figure 5.1: Comparison of the data-driven approach (Eq. 4.5) and the physics-informed approach (Eqs. 4.13, 4.17, and 4.21) with KMC data. The physics-informed approach captures the low T curvature and reduces the systematic underestimation of D in the low diffusivity regime.

The rest of the framework stays conceptually the same. KMC simulations can be performed on synthetic energy landscape, after which physics-informed MLSR expressions using the VFT form can be learned on the KMC dataset. Those expressions can then be applied by extracting the FCC barrier statistics from CI-NEB calculations. The learned coefficients will be FCC-specific (reflecting the different connectivity), but the methodology and its interpretability remain the the same. A challenge with FCC systems is that the magnetic state influences the diffusion barriers. In practice, using an antiferromagnetic description lowers the O-T-O diffusion barrier [54]. For the framework this implies that the FCC reference dataset should be generated for the state of interest.

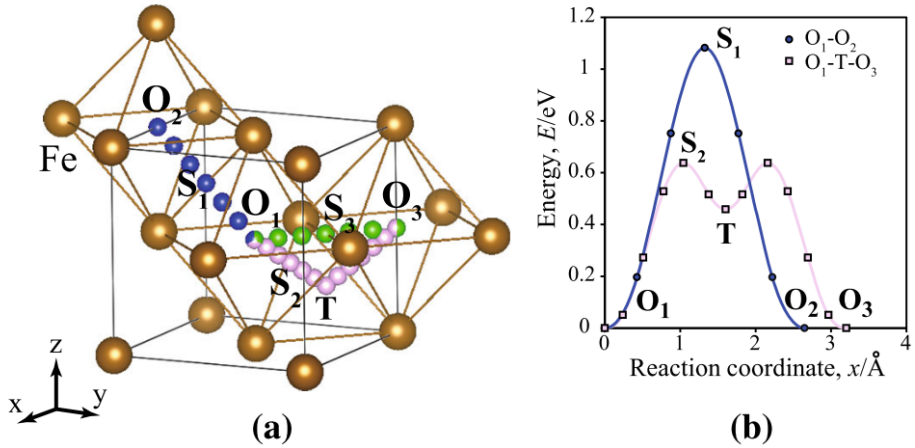
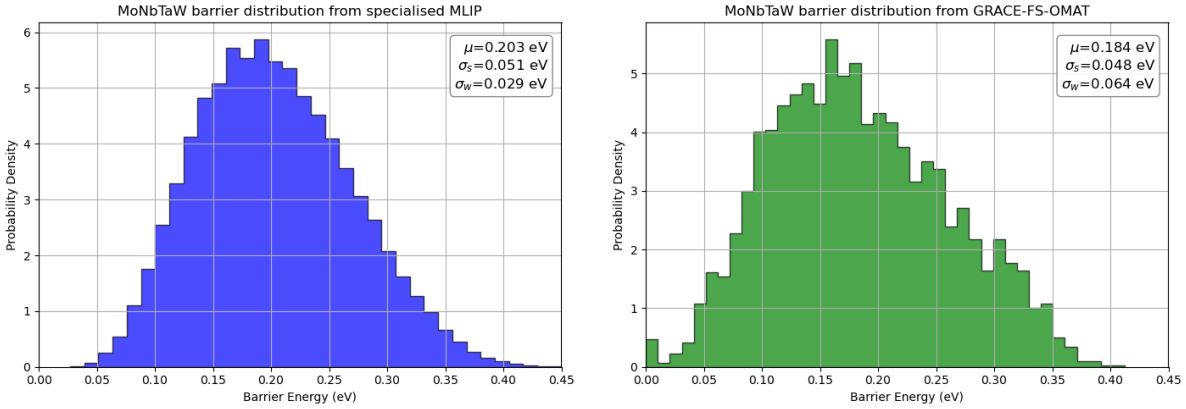


Figure 5.2: (a) Diffusional paths of H in pure FCC Fe, (b) the energy profiles of the O-O direct and O-T-O indirect paths, the indirect path is more favourable due to its lower energy [54].

5.4. Accuracy of uMLIPs

As mentioned in Section 4.4 and shown in Figure 4.18, certain diffusional pathways calculated using the CI-NEB method result in either the forward and/or backward barrier being zero. Further investigation reveals that this occurs because, during the relaxation of the structure during the CI-NEB calculation, the H atom in either the initial or the final structure moves away from the geometrically determined T-site. Consequently, that T-site is determined to be unstable for H to reside. This could either be a real unstable site due to the rough energy landscape of the alloy, or this instability can be a consequence of the uMLIP not being able to accurately describe the energies at that T-site. Figure 5.3 compares the diffusion barrier distributions of the equimolar MoNbTaW system. The results from the CI-NEB calculations of ref. [42], obtained using a MLIP trained on DFT data, are compared with those obtained using the GRACE-FS-OMAT uMLIP. The uMLIP predicts several zero-energy barriers, whereas the specialised MLIP trained directly on DFT data of MoNbTaW-H systems does not. This difference suggests that the zero-energy barriers are not intrinsic to the material and are therefore, by extension, unlikely to be present in the similar systems studied in this work. Instead, these unstable H sites likely arise from the limited accuracy of the uMLIP. Moreover, from the barrier distribution using the uMLIP, $\sigma_w = 0.064$ eV is found, while the specialised MLIP from ref. [42], $\sigma_w = 0.029$ eV is found (Figure 5.3). The larger value of σ_w predicted by the uMLIP indicates a systematic overestimation of site-to-site energy differences compared to specialised MLIP. Physically, this corresponds to a rougher well-energy landscape, implying greater asymmetry than actually present in the material. Such an overestimation is either linked to specific local atomic environments being underrepresented (or absent) in the uMLIP's training dataset or insufficient accuracy of the GRACE-FS model itself.



(a) **(b)**

Figure 5.3: Barrier distributions of MoNbTaW from (a) ref. [42], obtained using a specialised MLIP trained on DFT data, and (b) the universal MLIP GRACE-FS-OMAT employed in this work.

It is found that increasing chemical complexity and lattice strain lead to an increase in zero-energy barriers (Figure 4.17). Consequently, the use of these uMLIPs for calculating energy barrier distributions via CI-NEB should be restricted to systems containing, at most, four elements with similar lattice constant. However, these uMLIPs are continuously being refined, and future improvements may reduce the number of zero-energy barriers, extending their applicability to more complex systems.

As mentioned earlier, three different uMLIPs are tested using CI-NEB calculations on chemically complex BCC alloys. Two GRACE-FS potentials are compared: (1) the OMAT potential, trained on the OMat24 database, and (2) the OAM potential, also trained on the OMat24 database but

further fine-tuned on the sAlex and MPTrj datasets [66, 67]. Additionally, a more complex two-layer potential, GRACE-2L-OMAT, is compared with the GRACE-FS-OMAT potential.

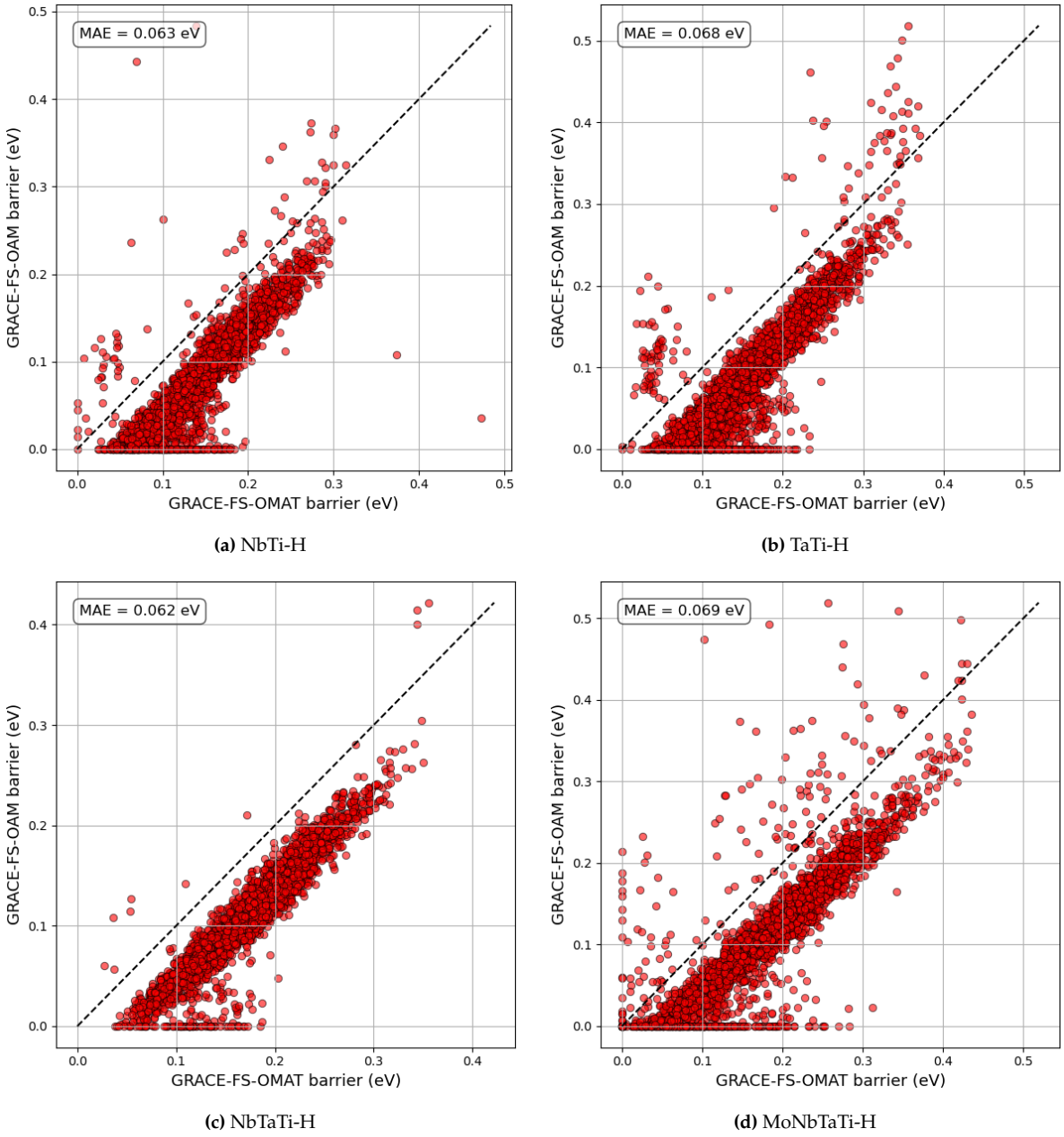
The results of the comparison of the GRACE-FS-OMAT and OAM potentials can be found in Figure 5.4. For each alloy and potential combination, 1536 CI-NEB calculations outputting 3072 diffusion barriers were performed. It can be seen that the OAM potential has much more zero-energy barriers, meaning that the potential was unable to correctly determine the minimum energy path and thus the associated saddle point. In addition, the OAM potential has more outliers with very high energy barrier, although some high energy barriers can be expected, values greater than 0.4 eV for these systems seem unphysical, especially if there is large discrepancy between the OMAT and OAM potential. Moreover, it can be seen that the OAM potential has an overall lower energy for the diffusion barriers, indicating that the OAM potential is less suitable to capture the saddle points using CI-NEB calculations. The OMat24 dataset contains non-equilibrium configurations, while sAlex and MPTrj contain (near-) equilibrium configurations [66]. The fine-tuning on (near-) equilibrium configurations leads to a decrease in the ability to correctly capture the saddle point.

A recent study by Deng et al. [78] offers a possible explanation for this behaviour. They identified a systematic softening of the potential energy surface (PES) in uMLIPs, which leads to underestimation of energies and forces. This softening stems from training primarily on (near-) equilibrium configurations, limiting the ability of the model to extrapolate accurately to high-energy states such as saddle points of diffusion barriers in complex alloys. As a result, configurations that are physically stable may be incorrectly predicted as unstable, producing artificial zero-barriers. Since the softening error is systematic, Deng et al. also found that it can be mitigated by fine-tuning the models with a small number of out-of-distribution, high energy configurations.

Therefore, it is deemed that potentials trained on the OMat24 dataset without the fine-tuning on sAlex and MPTrj datasets are more suitable for diffusion barrier calculations of the chemically complex BCC alloys investigated in this work.

Two GRACE potentials with different architectures, GRACE-FS and GRACE-2L, both trained on the OMat24 dataset, are compared in Figure 5.5 for the MoNbTaTi-H system, which is the most chemically complex system considered in this work. A random subset of 45 CI-NEB calculations was performed out of the 1536 possible paths in the supercell, yielding 90 diffusion barriers. Due to the greater computational demand of the 2L potential [64], only these 45 CI-NEB calculations were carried out. The GRACE-2L potential is chosen as a high-accuracy baseline, as recent studies show that such uMLIPs are able to attain near-DFT level accuracy [64, 79]. It can be seen the values are scattered across the parity line with a slight tendency for the 2L potential to obtain a lower diffusion barrier. From the results of these 90 diffusion barriers it is deemed that, for these systems and calculations, the FS potential is comparable to the 2L potential. Moreover, due to the computational demand of the 2L potential, MD simulations of complex systems with $10 \times 10 \times 10$ BCC supercells (as performed in Section 4.4) are not possible. Therefore, the FS model is used, however, the scatter between the two potentials indicates that the calculated diffusion barriers are very dependent on the potential used and that a large enough sample size of calculations is needed to obtain reliable μ, σ_s, σ_w parameters.

It is expected that the mean of the diffusion barriers in real materials is higher than that of the results of the CI-NEB calculations using the GRACE uMLIPs. This is motivated by the presence of the systematic softening, as explained by Deng et al. [78]. This can to some extend also be seen from the difference in mean value between the equimolar MoNbTaW barriers determined by the specialised MLIP trained on DFT data of ref. [42] and the GRACE-FS-OMAT



(a) NbTi-H **(b)** TaTi-H

(c) NbTaTi-H **(d)** MoNbTaTi-H

Figure 5.4: Parity plots of (a) NbTi-H, (b) TaTi-H, (c) NbTaTi-H, and (d) MoNbTaTi-H, 3072 diffusion barriers obtained through 1536 NEB calculations using GRACE-FS-OAM and GRACE-FS-OMAT uMLIPs.

potential in Figure 5.3. Therefore, it is expected that H diffusion in these MPEAs using the GRACE-FS potential is overestimated compared to the real material. Moreover, the larger value of the σ_w parameter suggests that the GRACE uMLIPs are not able to describe the depth of the well-energy in some sites. If either the depth of the initial or the final site is not as deep, the value of s is overestimated, leading to a larger value of σ_w.

Despite the limited accuracy of the GRACE-FS uMLIPs, the comparison between the MLSR expressions through CI-NEB calculations and the MD simulations gives insight into the relative accuracy of the MLSR expressions. The barrier statistic parameters μ, σ_s, σ_w are calculated using the same potential with the same error as the MD simulations Therefore, agreement between MLSR and MD demonstrates internal consistency under a common potential, but it

does not establish absolute accuracy with respect to DFT or experimental data. An external check would require benchmarking with higher accuracy MLIPs. For example, the GRACE-2L uMLIP or DFT-trained sMLIPs. However, the GRACE-2L and similar larger uMLIPs are too computationally expensive to use on these large-scale diffusion simulations through MD, and DFT-trained models require computationally expensive DFT training data and limit the ability to sample the compositional space of BCC MPEAs due to the system specific nature of these specialised MLIPs. In addition, the results could be compared to experimental data, however, no experimental H diffusion data on BCC MPEAs is available at this time.

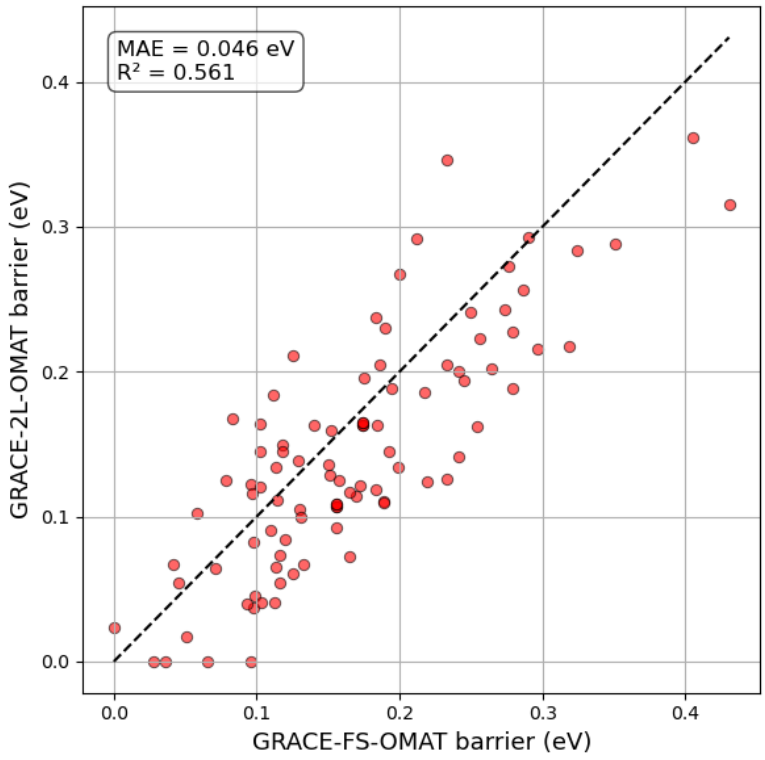


Figure 5.5: Parity plot of 90 diffusion barriers of MoNbTaTi-H system using GRACE-2L and GRACE-FS, both trained on OMat24.

5.5. Limitations of the model

Although the developed framework effectively captures the relationship between statistical features of energy landscapes and hydrogen diffusivity through MLSR expressions, and demonstrates good agreement with MD simulations at elevated temperatures, it has several limitations that constrain its accuracy and general applicability across broader material systems.

1. The assumption that the distribution of energy barriers follows a Gaussian profile simplifies the description of complex diffusion environments. However, this approximation tends to overestimate the number of low energy barriers compared to those derived from CI-NEB calculations. As a result, the model underestimates the extent of trapping at low temperatures, leading to an overestimation of diffusivity in the super-Arrhenius regime. This discrepancy is evident in the significantly underestimated Vogel temperature T_0 when compared to literature results for systems like MoNbTaW-H. The inability to capture the correct tail behaviour of the barrier distribution limits the accuracy of the framework for predicting low temperature diffusion behaviour.

2. The framework assumes a fully random atomic distribution and does not account for SRO. SRO can alter local environments and associated diffusion pathways, leading to deviations from the predicted statistical behaviour. Since diffusion barriers are sampled randomly based on statistical parameters and are therefore treated as globally independent variables, the framework cannot describe ordered or partially ordered systems. For the alloys studied in Section 4.4, it was found that equimolar NbTaTi has close to no SRO at high and low temperature [80]. However, for equimolar MoNbTi, which is similar to the MoNbTaTi alloy studied in this work, some degree of SRO (particularly influenced by Mo) is found at low temperature (300 K), while the degree of SRO is close to zero at high temperature (1673 K) [80].
3. All simulations and predictions are based on bulk configurations, thus excluding the influence of grain boundaries, surfaces, dislocations, or other defects. In BCC α-Fe, it was found that grain boundaries and surfaces can act as strong trapping sites, leading to altered diffusion behaviour in polycrystalline materials compared to single crystals [37, 38, 40]. The reduction of diffusivity due to defect is not captured by the framework.
4. The framework uses just one H atom per supercell structure, thus the applicability is limited to low H concentrations. At high H concentrations in MPEAs, O-sites begin to be occupied and could drive phase transitions from BCC to BCT to FCC [28, 31]. For H diffusion in BCC α-Fe, non-Arrhenius behaviour has been observed at elevated H concentrations (1–5%). This deviation arises from the formation of H clusters below certain temperatures, which cause a structural transformation to either FCC or amorphous phases, thereby reducing diffusivity in a non-Arrhenius manner [34].
5. KMC simulations used to generate the dataset are based on classical hopping mechanisms, which become less accurate below approximately 350 K due to the onset of nuclear quantum effects that modify diffusion pathways and lower the diffusion barrier [56]. This limitation restricts the temperature range used to learn the MLSR expressions to above 350 K, and therefore limit the framework applicability to above this temperature.
6. Although fast and computationally efficient, the GRACE-FS-OMAT uMLIP has limitations in describing highly strained or chemically complex systems. Unstable hydrogen sites (i.e. zero-energy barriers) increasingly appear in CI-NEB calculations as the number of elements or local lattice strain increases. This results from inaccuracies in the uMLIP, possibly due to systematic softening, as shown in comparisons with a more specialised MLIP trained directly on DFT data. Consequently, the model is currently limited to systems with moderate complexity (up to four elements with similar atomic sizes and lattice constants). The ongoing development of more accurate uMLIPs may increase the applicability of the framework in the future.
7. The MLSR expressions are derived from model systems without elemental identities. Therefore, a fixed lattice constant of 3.244 Å, corresponding to MoNbTaW, is used to convert H trajectories to MSD. While this value represents a reasonable approximation for refractory BCC alloys, variations in real materials may influence the absolute diffusivity. However, the relaxed lattice constants of the studied alloys show only minor deviations: 3.279 Å for NbTi, 3.267 Å for TaTi, 3.290 Å for NbTaTi, and 3.237 Å for MoNbTaTi. These small differences suggest that the fixed lattice constant is a valid and practical approximation for the systems considered in this work.

6

Conclusion

In this thesis, a computational framework was built to efficiently predict hydrogen diffusion in multi-principal element alloys (MPEAs). These materials, though promising for hydrogen embrittlement resistance and hydrogen storage, present difficulties in experimentally navigating their vast compositional space, therefore computational approaches are needed. The current state of the art in computational approaches for hydrogen diffusion in MPEAs uses machine learning interatomic potentials (MLIPs) trained on computationally expensive density functional theory (DFT) data to determine diffusion barriers through climbing-image nudged elastic band (CI-NEB) calculations. Then, the local atomic environment of the MPEAs are correlated to the diffusion barriers through machine learning (ML) models, such as neural networks. After which, kinetic Monte Carlo (KMC) simulations are performed using barriers obtained from these ML models to obtain hydrogen trajectories, of which the mean square displacement (MSD) and diffusion coefficient D is extracted.

While accurate, this approach is computationally intensive and difficult to scale across the vast compositional space of MPEAs. This thesis presents an alternative, more efficient framework by reframing the problem through the statistical properties of the energy landscape. By introducing a reference system that characterises the hydrogen diffusion energy landscape using three parameters, namely the mean diffusion barrier (μ), the standard deviation of the saddle point energies (σ_s), and the standard deviation of the well energies (σ_w), it was possible to construct a large dataset of KMC simulations. These simulations sampled a wide range of energy landscapes and enabled a direct connection between diffusivity and the underlying energy landscape parameters.

Machine learning symbolic regression (MLSR) was employed to derive analytical expressions. Two approaches were explored: a purely data-driven approach with two models of the form $D = \mathrm{f}(\mu, \sigma_s, \sigma_w, T)$ and $\ln(D) = \mathrm{f}(\mu, \sigma_s, \sigma_w, T)$, and a physics-informed approach based on the Vogel–Fulcher–Tammann (VFT) model, where the VFT parameters themselves were expressed as functions of the barrier statistics. The latter approach offered more interpretable expressions between local energy landscape roughness and macroscopic transport behaviour, especially under conditions of super-Arrhenius diffusion at low temperatures.

The framework was applied to real complex BCC alloy systems by extracting barrier statistics from CI-NEB calculations using universal MLIPs (uMLIPs), avoiding the need for training system-specific potentials. The predicted diffusivities were validated against molecular dynamics (MD) simulations, demonstrating reasonable agreement with the model. This

efficient and generalisable method enables rapid screening of MPEA compositions for hydrogen transport properties, previously limited by the high cost of simulations or experiments.

However, the framework contains several limitations. The assumption of Gaussian-distributed energy barriers simplifies the true complexity of MPEA energy landscapes. Specifically, the Gaussian approximation overestimates the number of low energy barriers, and therefore overestimates the diffusivity, especially at low temperature where super-Arrhenius behaviour occurs. Additionally, while uMLIPs enable rapid and broad applicability, their accuracy in capturing local chemical effects across all configurations remains uncertain and may introduce bias into the statistics due to systematic softening. Moreover, currently, uMLIPs are not able to be applied to systems with large lattice strain and very chemical complex environments.

Despite these limitations, the thesis provides a step toward interpretable and scalable models for hydrogen diffusion in disordered systems. By decoupling the diffusion process from the full atomic resolution of local environments and instead focusing on statistical descriptors, this work opens a pathway for fast and physically meaningful prediction of hydrogen diffusivity in complex alloys. Moreover, a significant reduction in computational cost is found compared to MD simulations.

In conclusion, this framework complements existing high-fidelity simulation methods by offering a lightweight, interpretable tool for rapid assessment of hydrogen diffusivity in MPEAs. It enables both fundamental insight into the role of energy landscape disorder and practical utility in alloy design for hydrogen-related applications. Future work should explore extending the framework to handle non-Gaussian distributions through more input parameters, therefore more accurately capturing low temperature super-Arrhenius behaviour.

7

Recommendations

1. The framework of this thesis show promise for application at higher temperatures. However, the assumption of a Gaussian energy landscape overestimates the diffusivity at low temperature, thus underestimating super-Arrhenius behaviour. Future studies could explore MLSR expressions derived from KMC datasets where the energy barrier distribution is parametrised with additional variables. In particular, accurately capturing low-energy barriers is essential for describing super-Arrhenius effects. While more parameters require larger datasets, MLSR methods are well-suited for learning these higher-dimensional expressions.
2. Potentials trained exclusively on the OMat24 dataset appear more suitable than OAM potentials for saddle-point calculations using the CI-NEB method, as the OMAT potential shows fewer unstable H sites and yields mean energy barriers that align more closely with expected values. In contrast, the additional fine-tuning of OAM potentials on (near-) equilibrium structures tends to amplify the softening effect [78]. Enhancing the ability of a uMLIP to describe non-equilibrium structures reduces the number of unstable H sites and improves accuracy. Consequently, fine-tuning on a small number of DFT calculations could further increase the reliability of H site and saddle-point descriptions.
3. Hasan et al. [34] demonstrated that high H concentrations in BCC Fe lead to non-Arrhenius diffusion behaviour driven by H clustering which causes phase transformations at lower temperatures. It is known that increasing H concentration in BCC MPEAs causes the occupation of O-sites, which then cause phase transformations [28, 31]. However, it is not known whether increasing H concentration similarly enhances super-Arrhenius behaviour in BCC MPEAs, and how this effect varies with alloy composition, lattice distortion, and the presence of strong trapping elements. Such studies should also account for the role of microstructural features, particularly grain boundaries, which act as strong traps in BCC systems and can significantly reduce effective diffusivity [37, 38, 40].
4. An interesting direction is the integration of bulk and grain boundary diffusivity data, as demonstrated by Sun et al. [40], to obtain a complete description of H diffusion in polycrystalline MPEAs. This approach would enable assessment of the relative contributions of bulk and interfacial diffusion under varying H concentrations and temperatures, offering a more complete picture of macroscopic diffusivity. Such work could also reveal how H–microstructure interactions influence diffusion.

References

[1] D. J. Soeder. "Fossil Fuels and Climate Change". In: *Fracking and the Environment: A Scientific Assessment of the Environmental Risks from Hydraulic Fracturing and Fossil Fuels*. Springer International Publishing, Nov. 2021. Chap. 9, pp. 155–185. ISBN: 9783030591212. DOI: 10.1007/978-3-030-59121-2.

[2] K. T. Møller, T. R. Jensen, E. Akiba, and H. w. Li. "Hydrogen - A sustainable energy carrier". In: *Progress in Natural Science: Materials International* 27.1 (Feb. 2017), pp. 34–40. ISSN: 17455391. DOI: 10.1016/j.pnsc.2016.12.014.

[3] I. Dincer and M. I. Aydin. "New paradigms in sustainable energy systems with hydrogen". In: *Energy Conversion and Management* 283 (May 2023). ISSN: 01968904. DOI: 10.1016/j.enconman.2023.116950.

[4] R. R. Wang, Y. Q. Zhao, A. Babich, D. Senk, and X. Y. Fan. *Hydrogen direct reduction (H-DR) in steel industry—An overview of challenges and opportunities*. Dec. 2021. DOI: 10.1016/j.jclepro.2021.129797.

[5] Our World in Data. *CO2 emissions by fuel or industry type*. 2024. URL: https://ourworldindata.org/grapher/co2-by-source.

[6] X. Li, J. Zhang, Y. Cui, M. B. Djukic, H. Feng, and Y. Wang. *Review of the hydrogen embrittlement and interactions between hydrogen and microstructural interfaces in metallic alloys: Grain boundary, twin boundary, and nano-precipitate*. June 2024. DOI: 10.1016/j.ijhydene.2024.05.257.

[7] J. Hoschke, M. F. W. Chowdhury, J. Venezuela, and A. Atrens. *A review of hydrogen embrittlement in gas transmission pipeline steels*. June 2023. DOI: 10.1515/corrrev-2022-0052.

[8] Y. S. Chen, C. Huang, P. Y. Liu, H. W. Yen, R. Niu, P. Burr, K. L. Moore, E. Martínez-Pañeda, A. Atrens, and J. M. Cairney. "Hydrogen trapping and embrittlement in metals – A review". In: *International Journal of Hydrogen Energy* (2024). ISSN: 03603199. DOI: 10.1016/j.ijhydene.2024.04.076.

[9] J. Andersson and S. Grönkvist. *Large-scale storage of hydrogen*. May 2019. DOI: 10.1016/j.ijhydene.2019.03.063.

[10] F. Marques, M. Balcerzak, F. Winkelmann, G. Zepon, and M. Felderhoff. *Review and outlook on high-entropy alloys for hydrogen storage*. Oct. 2021. DOI: 10.1039/d1ee01543e.

[11] E. P. George, D. Raabe, and R. O. Ritchie. *High-entropy alloys*. Aug. 2019. DOI: 10.1038/s41578-019-0121-4.

[12] J. W. Yeh, S. K. Chen, S. J. Lin, J. Y. Gan, T. S. Chin, T. T. Shun, C. H. Tsau, and S. Y. Chang. "Nanostructured high-entropy alloys with multiple principal elements: Novel alloy design concepts and outcomes". In: *Advanced Engineering Materials* 6.5 (2004), pp. 299–303. ISSN: 14381656. DOI: 10.1002/adem.200300567.

[13] W. Li, D. Xie, D. Li, Y. Zhang, Y. Gao, and P. K. Liaw. *Mechanical behavior of high-entropy alloys*. May 2021. DOI: 10.1016/j.pmatsci.2021.100777.

[14] X. Y. Zhou, J. H. Zhu, Y. Wu, X. S. Yang, T. Lookman, and H. H. Wu. "Machine learning assisted design of FeCoNiCrMn high-entropy alloys with ultra-low hydrogen diffusion coefficients". In: *Acta Materialia* 224 (Feb. 2022). ISSN: 13596454. DOI: 10.1016/j.actamat.2021.117535.

[15] Z. Pu, Y. Chen, and L. H. Dai. "Strong resistance to hydrogen embrittlement of high-entropy alloy". In: *Materials Science and Engineering: A* 736.July (2018), pp. 156–166. ISSN: 09215093. DOI: 10.1016/j.msea.2018.08.101.

[16] A. Liski, T. Vuoriheimo, J. Byggmästar, K. Mizohata, K. Heinola, T. Ahlgren, K. K. Tseng, T. E. Shen, C. W. Tsai, J. W. Yeh, K. Nordlund, F. Djurabekova, and F. Tuomisto. "Solubility of Hydrogen in a WMoTaNbV High-Entropy Alloy". In: *Materials* 17.11 (June 2024). ISSN: 19961944. DOI: 10.3390/ma17112574.

[17] M. Sahlberg, D. Karlsson, C. Zlotea, and U. Jansson. "Superior hydrogen storage in high entropy alloys". In: *Scientific Reports* 6 (Nov. 2016). ISSN: 20452322. DOI: 10.1038/srep36770.

[18] Z. Ding, Y. Li, H. Jiang, Y. Zhou, H. Wan, J. Qiu, F. Jiang, J. Tan, W. Du, Y. Chen, L. L. Shaw, and F. Pan. "The integral role of high-entropy alloys in advancing solid-state hydrogen storage". In: *Interdisciplinary Materials* 4.1 (Jan. 2025), pp. 75–108. ISSN: 2767-4401. DOI: 10.1002/idm2.12216. URL: https://onlinelibrary.wiley.com/doi/10.1002/idm2.12216.

[19] D. B. Miracle and O. N. Senkov. *A critical review of high entropy alloys and related concepts*. Jan. 2017. DOI: 10.1016/j.actamat.2016.08.081.

[20] O. N. Senkov, G. B. Wilks, J. M. Scott, and D. B. Miracle. "Mechanical properties of Nb25Mo25Ta 25W25 and V20Nb20Mo 20Ta20W20 refractory high entropy alloys". In: *Intermetallics* 19.5 (May 2011), pp. 698–706. ISSN: 09669795. DOI: 10.1016/j.intermet.2011.01.004.

[21] K. Rose, R. Devanathan, J. Perepezko, and G. Balasubramanian. *Design of High-Temperature Multi-Principal Element Alloys: accurate theory-guided design validated by high-throughput experiments*. Tech. rep. Dec. 2021. DOI: https://doi.org/10.2172/1839082. URL: http://www.osti.gov/bridge.

[22] S. Dixit, S. Rodriguez, M. R. Jones, P. Buzby, R. Dixit, N. Argibay, F. W. DelRio, H. H. Lim, and D. Fleming. *Refractory High-Entropy Alloy Coatings for High-Temperature Aerospace and Energy Applications*. Apr. 2022. DOI: 10.1007/s11666-022-01324-0.

[23] L. Zhuo, Y. Xie, and B. Chen. "A review on recent progress of refractory high entropy alloys: From fundamental research to engineering applications". In: *Journal of Materials Research and Technology* 33 (Nov. 2024), pp. 1097–1129. ISSN: 22140697. DOI: 10.1016/j.jmrt.2024.09.131.

[24] X. Li, X. Ma, J. Zhang, E. Akiyama, Y. Wang, and X. Song. *Review of Hydrogen Embrittlement in Metals: Hydrogen Diffusion, Hydrogen Characterization, Hydrogen Embrittlement Mechanism and Prevention*. June 2020. DOI: 10.1007/s40195-020-01039-7.

[25] Y. S. Chen, P. Y. Liu, R. Niu, A. Devaraj, H. W. Yen, R. K. Marceau, and J. M. Cairney. *Atom Probe Tomography for the Observation of Hydrogen in Materials: A Review*. Feb. 2023. DOI: 10.1093/micmic/ozac005.

[26] A. Raina, V. S. Deshpande, and N. A. Fleck. "Analysis of thermal desorption of hydrogen in metallic alloys". In: *Acta Materialia* 144 (Feb. 2018), pp. 777–785. ISSN: 13596454. DOI: 10.1016/j.actamat.2017.11.011.

[27] B. Xing, T. J. Rupert, X. Pan, and P. Cao. "Neural network kinetics for exploring diffusion multiplicity and chemical ordering in compositionally complex materials". In: *Nature Communications* 15.1 (Dec. 2024). ISSN: 20411723. DOI: 10.1038/s41467-024-47927-9.

[28] C. M. Moore, J. A. Wilson, M. J. Rushton, W. E. Lee, J. O. Astbury, and S. C. Middleburgh. "Hydrogen accommodation in the TiZrNbHfTa high entropy alloy". In: *Acta Materialia* 229 (May 2022). ISSN: 13596454. DOI: 10.1016/j.actamat.2022.117832.

[29] X. L. Ren, P. H. Shi, B. D. Yao, L. Wu, X. Y. Wu, and Y. X. Wang. "Hydrogen solution in high-entropy alloys". In: *Physical Chemistry Chemical Physics* 23.48 (Dec. 2021), pp. 27185–27194. ISSN: 14639076. DOI: 10.1039/d1cp04151g.

[30] T. R. Yang, Y. X. Wang, Y. H. Li, S. Zhao, H. Bo Zhou, F. Tian, and G. H. Lu. "Exploring the inhibitory effect of WTaVCr high-entropy alloys on hydrogen retention: From dissolution, diffusion to desorption". In: *Journal of Nuclear Materials* 601 (Dec. 2024). ISSN: 00223115. DOI: 10.1016/j.jnucmat.2024.155346.

[31] C. Wu, Y. Gong, C. Liu, X. Li, G. Gizer, C. Pistidda, F. Körmann, Y. Ma, J. Neugebauer, and D. Raabe. "Hydrogen accommodation and its role in lattice symmetry in a TiNbZr medium-entropy alloy". In: *Acta Materialia* 288 (Apr. 2025). ISSN: 13596454. DOI: 10.1016/j.actamat.2025.120852.

[32] Z. Nefzi, C. Barreteau, L. Levrel, and J.-C. Crivello. "Promising Multicomponent Cubic Alloys for Hydrogen Storage: Insights from First-Principles Calculations and Machine Learning". In: *ACS Applied Energy Materials* (Mar. 2025). ISSN: 2574-0962. DOI: 10.1021/acsaem.4c02728. URL: https://pubs.acs.org/doi/10.1021/acsaem.4c02728.

[33] V. Korostelev, J. Wagner, and K. Klyukin. "Simple local environment descriptors for accurate prediction of hydrogen absorption and migration in metal alloys". In: *Journal of Materials Chemistry A* 11.43 (Oct. 2023), pp. 23576–23588. ISSN: 20507496. DOI: 10.1039/d3ta04554d.

[34] M. A. A. Hasan, J. Wang, Y. C. Lim, A. Hu, and S. Shin. "Concentration dependence of hydrogen diffusion in α-iron from atomistic perspectives". In: *International Journal of Hydrogen Energy* 44.51 (2019), pp. 27876–27884. ISSN: 03603199. DOI: 10.1016/j.ijhydene.2019.09.007. URL: https://doi.org/10.1016/j.ijhydene.2019.09.007.

[35] A. Ramasubramaniam, M. Itakura, and E. A. Carter. "Interatomic potentials for hydrogen in α -iron based on density functional theory". In: *Physical Review B - Condensed Matter and Materials Physics* 79.17 (May 2009). ISSN: 10980121. DOI: 10.1103/PhysRevB.79.174101.

[36] M. I. Mendelev, S. Han, D. J. Srolovitz, G. J. Ackland, D. Y. Sun, and M. Asta. "Development of new interatomic potentials appropriate for crystalline and liquid iron". In: *Philosophical Magazine* 83.35 (Dec. 2003), pp. 3977–3994. ISSN: 14786435. DOI: 10.1080/14786430310001613264.

[37] X. Y. Zhou, J. H. Zhu, and H. H. Wu. "Molecular dynamics studies of the grain-size dependent hydrogen diffusion coefficient of nanograined Fe". In: *International Journal of Hydrogen Energy* 46.7 (Jan. 2021), pp. 5842–5851. ISSN: 03603199. DOI: 10.1016/j.ijhydene.2020.11.131.

[38] D. Smirnova and S. Starikov. "Atomistic study of hydrogen diffusion in presence of defects in bcc and fcc iron". In: *Computational Materials Science* 230.March (2023), p. 112433. ISSN: 09270256. DOI: 10.1016/j.commatsci.2023.112433. URL: https://doi.org/10.1016/j.commatsci.2023.112433.

[39] S. Starikov, D. Smirnova, T. Pradhan, I. Gordeev, R. Drautz, and M. Mrovec. "Angular-dependent interatomic potential for large-scale atomistic simulation of the Fe-Cr-H ternary system". In: *Physical Review Materials* 6.4 (Apr. 2022). ISSN: 24759953. DOI: 10.1103/PhysRevMaterials.6.043604.

[40] C. Sun, Y. Yuan, M. Xu, X. Li, M. Li, and C. Zhou. "A high-throughput molecular dynamics study with a machine-learning method on predicting diffusion coefficient of hydrogen in α-iron grain boundaries". In: *Materials Today Communications* 44 (Mar. 2025). ISSN: 23524928. DOI: 10.1016/j.mtcomm.2025.111997.

[41] H. Tang, B. Li, Y. Song, M. Liu, H. Xu, G. Wang, H. Chung, and J. Li. "Reinforcement Learning-Guided Long-Timescale Simulation of Hydrogen Transport in Metals". In: *Advanced Science* 11.5 (Feb. 2024). ISSN: 21983844. DOI: 10.1002/advs.202304122.

[42] F. Shuang, Y. Ji, Z. Wei, C. Dong, W. Gao, L. Laurenti, and P. Dey. "Decoding the hidden dynamics of super-Arrhenius hydrogen diffusion in multi-principal element alloys via machine learning". In: *Acta Materialia* 289 (2025). ISSN: 1359-6454. DOI: https://doi.org/10.1016/j.actamat.2025.120924.

[43] P. A. Santos-Florez, S. C. Dai, Y. Yao, H. Yanxon, L. Li, Y. J. Wang, Q. Zhu, and X. X. Yu. "Short-range order and its impacts on the BCC MoNbTaW multi-principal element alloy by the machine-learning potential". In: *Acta Materialia* 255 (Aug. 2023). ISSN: 13596454. DOI: 10.1016/j.actamat.2023.119041.

[44] B. Xu, J. Zhang, Y. Xiong, S. Ma, Y. Osetsky, and S. Zhao. "Mechanism of sluggish diffusion under rough energy landscape". In: *Cell Reports Physical Science* 4.4 (Apr. 2023). ISSN: 26663864. DOI: 10.1016/j.xcrp.2023.101337.

[45] B. Xing, W. Zou, T. J. Rupert, and P. Cao. "Vacancy diffusion barrier spectrum and diffusion correlation in multicomponent alloys". In: *Acta Materialia* 266 (Mar. 2024). ISSN: 13596454. DOI: 10.1016/j.actamat.2024.119653.

[46] B. Xu, J. Zhang, S. Ma, Y. Xiong, S. Huang, J. J. Kai, and S. Zhao. "Revealing the crucial role of rough energy landscape on self-diffusion in high-entropy alloys based on machine learning and kinetic Monte Carlo". In: *Acta Materialia* 234 (Aug. 2022). ISSN: 13596454. DOI: 10.1016/j.actamat.2022.118051.

[47] Y. Li, J. P. Du, S. Shinzato, and S. Ogata. "Tunable interstitial and vacancy diffusivity by chemical ordering control in CrCoNi medium-entropy alloy". In: *npj Computational Materials* 10.1 (Dec. 2024). ISSN: 20573960. DOI: 10.1038/s41524-024-01322-6.

[48] B. Xu, S. Ma, S. Huang, J. Zhang, Y. Xiong, H. Fu, X. Xiang, and S. Zhao. "Influence of short-range order on diffusion in multiprincipal element alloys from long-time atomistic simulations". In: *Physical Review Materials* 7.3 (Mar. 2023). ISSN: 24759953. DOI: 10.1103/PhysRevMaterials.7.033605.

[49] S. L. Thomas and S. Patala. "Vacancy diffusion in multi-principal element alloys: The role of chemical disorder in the ordered lattice". In: *Acta Materialia* 196 (Sept. 2020), pp. 144–153. ISSN: 13596454. DOI: 10.1016/j.actamat.2020.06.022.

[50] R. LeSar. *Introduction to computational materials science : fundamentals to applications*. Cambridge University Press, 2013, p. 414. ISBN: 9780521845878.

[51] T. W. Ko and S. P. Ong. *Recent advances and outstanding challenges for machine learning interatomic potentials*. Dec. 2023. DOI: 10.1038/s43588-023-00561-9.

[52] P. Cordier. *RheoMan - Nudged Elastic Band*. [Online; accessed 31. Jul. 2025]. Dec. 2021. URL: https://umet.univ-lille.fr/Projets/RheoMan/en/to-learn-more-about/nudged-elastic-band.php.html.

[53] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz. *A New Algorithm for Monte Carlo Simulation of Ising Spin Systems**. Tech. rep. 1975, p. 18. DOI: https://doi.org/10.1016/0021-9991(75)90060-1.

[54] K. Hirata, S. Iikubo, M. Koyama, K. Tsuzaki, and H. Ohtani. "First-Principles Study on Hydrogen Diffusivity in BCC, FCC, and HCP Iron". In: *Metallurgical and Materials Transactions A: Physical Metallurgy and Materials Science* 49.10 (Oct. 2018), pp. 5015–5022. ISSN: 10735623. DOI: 10.1007/s11661-018-4815-9.

[55] X. W. Zhou, F. El Gabaly, V. Stavila, and M. D. Allendorf. "Molecular Dynamics Simulations of Hydrogen Diffusion in Aluminum". In: *Journal of Physical Chemistry C* 120.14 (Apr. 2016), pp. 7500–7509. ISSN: 19327455. DOI: 10.1021/acs.jpcc.6b01802.

[56] H. Kimizuka, H. Mori, and S. Ogata. "Effect of temperature on fast hydrogen diffusion in iron: A path-integral quantum dynamics approach". In: *Physical Review B - Condensed Matter and Materials Physics* 83.9 (Mar. 2011). ISSN: 10980121. DOI: 10.1103/PhysRevB.83.094110.

[57] Delft High Performance Computing Centre (DHPC). *DelftBlue Supercomputer (Phase 2)*. https://www.tudelft.nl/dhpc/ark:/44463/DelftBluePhase2. 2024.

[58] M. Cranmer. "Interpretable Machine Learning for Science with PySR and SymbolicRegression.jl". In: (May 2023). URL: http://arxiv.org/abs/2305.01582.

[59] Y. Wang, N. Wagner, and J. M. Rondinelli. *Symbolic regression in materials science*. Sept. 2019. DOI: 10.1557/mrc.2019.85.

[60] D. Angelis, F. Sofos, and T. E. Karakasidis. *Artificial Intelligence in Physical Sciences: Symbolic Regression Trends and Perspectives*. July 2023. DOI: 10.1007/s11831-023-09922-z.

[61] L. S. Garca-Coln, L. F. del Castillo, and P. Goldstein. "Theoretical basis for the Vogel-Fulcher-Tammann equation". In: *Phys. Rev. B* 40 (10 Oct. 1989), pp. 7040–7044. DOI: 10.1103/PhysRevB.40.7040.

[62] Y. Mishin. "Machine-learning interatomic potentials for materials science". In: *Acta Materialia* 214 (Aug. 2021). ISSN: 13596454. DOI: 10.1016/j.actamat.2021.116980.

[63] R. Jacobs et al. *A practical guide to machine learning interatomic potentials – Status and future*. Mar. 2025. DOI: 10.1016/j.cossms.2025.101214.

[64] F. Shuang, Z. Wei, K. Liu, W. Gao, and P. Dey. "Universal machine learning interatomic potentials poised to supplant DFT in modeling general defects in metals and random alloys". In: *Machine Learning: Science and Technology* 6.3 (Sept. 2025). ISSN: 26322153. DOI: 10.1088/2632-2153/adea2d.

[65] A. Bochkarev, Y. Lysogorskiy, and R. Drautz. "Graph Atomic Cluster Expansion for Semilocal Interactions beyond Equivariant Message Passing". In: *Phys. Rev. X* 14.2 (June 2024), p. 021036. DOI: 10.1103/PhysRevX.14.021036.

[66] L. Barroso-Luque, M. Shuaibi, X. Fu, B. M. Wood, M. Dzamba, M. Gao, A. Rizvi, C. L. Zitnick, and Z. W. Ulissi. "Open Materials 2024 (OMat24) Inorganic Materials Dataset and Models". In: (Oct. 2024). URL: http://arxiv.org/abs/2410.12771.

[67] B. Deng. "Materials Project Trajectory (MPtrj) Dataset". In: (July 2023). DOI: 10.6084/m9.figshare.23713842.v2. URL: https://figshare.com/articles/dataset/Materials_Project_Trjectory_MPtrj_Dataset/23713842.

[68] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton. "LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales". In: *Comp. Phys. Comm.* 271 (2022), p. 108171. DOI: 10.1016/j.cpc.2021.108171.

[69] H. Jónsson, G. Mills, and K. W. Jacobsen. "Nudged elastic band method for finding minimum energy paths of transitions". In: World Scientific Pub Co Pte Lt, June 1998, pp. 385–404. DOI: 10.1142/9789812839664_0016.

[70] G. Henkelman, B. P. Uberuaga, and H. Jónsson. "Climbing image nudged elastic band method for finding saddle points and minimum energy paths". In: *Journal of Chemical Physics* 113.22 (Dec. 2000), pp. 9901–9904. ISSN: 00219606. DOI: 10.1063/1.1329672.

[71] R. E. Jones, C. R. Weinberger, S. P. Coleman, and G. J. Tucker. "Introduction to Atomistic Simulation Methods". In: *Multiscale Materials Modeling for Nanomechanics*. Ed. by C. R. Weinberger and G. J. Tucker. Cham: Springer International Publishing, 2016, pp. 1–52. ISBN: 978-3-319-33480-6. DOI: 10.1007/978-3-319-33480-6_1. URL: https://doi.org/10.1007/978-3-319-33480-6_1.

[72] P. Hirel. "Atomsk: A tool for manipulating and converting atomic data files". In: *Comput. Phys. Commun.* 197 (Dec. 2015), pp. 212–219. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2015.07.012.

[73] S. T. Murphy, K. Nordlund, Djurabekova F, and D. M. Duffy. "Brief review of basics of molecular dynamics". In: *Tools for investigating electronic excitation : experiment and multi-scale modelling*. Madrid, 2021. Chap. 8, pp. 199–225. ISBN: 978-84-09-36032-1. URL: http://hdl.handle.net/10138/337067.

[74] E. Martinez, D. Perez, V. Gavani, and S. Kenny. *Advanced atomistic algorithms in materials science*. Apr. 2018. DOI: 10.1557/jmr.2018.69.

[75] A. Yelon, B. Movaghar, and H. M. Branz. "Origin and consequences of the compensation (Meyer-Neldel) law". In: *Phys. Rev. B* 46 (19 Nov. 1992), pp. 12244–12250. DOI: 10.1103/PhysRevB.46.12244.

[76] M. Cranmer. *SymbolicRegression.jl*. [Online; accessed 31. Jul. 2025]. July 2025. URL: https://github.com/MilesCranmer/SymbolicRegression.jl.

[77] M. Cranmer. *API Reference - PySR*. [Online; accessed 18. Aug. 2025]. Aug. 2025. URL: https://astroautomata.com/PySR/api.

[78] B. Deng, Y. Choi, P. Zhong, J. Riebesell, S. Anand, Z. Li, K. J. Jun, K. A. Persson, and G. Ceder. "Systematic softening in universal machine learning interatomic potentials". In: *npj Computational Materials* 11.1 (Dec. 2025). ISSN: 20573960. DOI: 10.1038/s41524-024-01500-6.

[79] K. Liu, Z. Wei, W. Gao, P. Dey, M. H. F. Sluiter, and F. Shuang. "Heterogeneous Ensemble Enables a Universal Uncertainty Metric for Atomistic Foundation Models". In: (July 2025). URL: http://arxiv.org/abs/2507.21297.

[80] H. Zheng, L. T. Fey, X. G. Li, Y. J. Hu, L. Qi, C. Chen, S. Xu, I. J. Beyerlein, and S. P. Ong. "Multi-scale investigation of short-range order and dislocation glide in MoNbTi and TaNbTi multi-principal element alloys". In: *npj Computational Materials* 9.1 (Dec. 2023). ISSN: 20573960. DOI: 10.1038/s41524-023-01046-z.