Off-lattice kinetic Monte Carlo in the Fe-H System

Conor John Williams

Supervisor: Dr E I Galindo-Nava

PhD in Scientific Computing

University of Cambridge

A report submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Til gamans

Abstract

We develop an off-lattice kinetic Monte Carlo model, based on the dimer saddle-point finding method, suitable for simulating the interaction of hydrogen with crystal defects in iron (such as vacancies, grain-boundaries and dislocations), over timescales not achievable with classical methods including molecular dynamics. The model is used to interrogate the mechanisms of hydrogen embrittlement, causing severe loss of ductility in steels with complicated microstructures, focusing on hydrogen-vacancy systems.

Acknowledgements

Table of contents

Chapter 1

Introduction

In this project we develop enhancement to the off-lattice kinetic Monte Carlo (OLKMC) method. Our main contribution is an error-tolerant replacement for topological analysis of atomic local-environments. The motivation for our research is building a general simulation framework capable of modelling the complex interactions between crystal defects and hydrogen (H) in iron (Fe), into the timescales required to study the mechanisms of hydrogen embrittlement.

We apply our OLKMC implementation to study the diffusion of vacancy clusters in the presence of hydrogen. We demonstrate OLKMC is capable of reaching embrittlement timescales, of-the-order-of seconds, while simultaneously resolving the atomic motion of H-atoms. Through OLKMC we are able to study the atomic mechanisms through which H impedes the the diffusion of vacancy clusters. These are important first steps towards modelling more the complex H-defect interactions a required to gain a full understanding of hydrogen embrittlement.

The remainder of this report is laid out as follows. In Section 1.1 we introduce the α–Fe-H system and summarise the hydrogen embrittlement literature. In Chapter 2 we conduct a detailed review into the OLKMC literature. In Chapter 3 we recap the details of the empirical potentials underlying this work, propose an alternative confidence parameter for saddle-point search completeness and introduce our ITLEE classification method while discussing its efficient implementation. In Chapter 4 we detail our results simulating the diffusivity of various vacancy clusters and discuss the mechanisms of their diffusion. Finally, in Chapter 5 we make our conclusions, highlight our findings and propose avenues for future work.
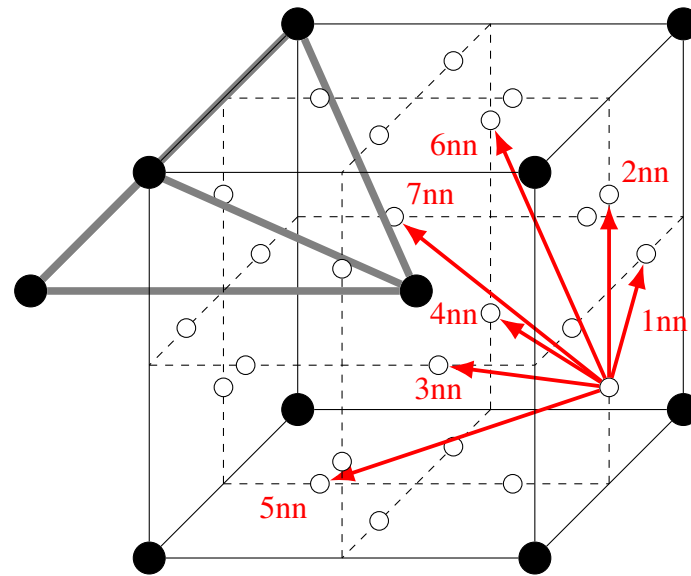
Fig. 1.1 Diagram of α–Fe body-centred cubic (BCC) unit cell (black dots), alongside the tetrahedral interstices (unfilled circles); thick grey lines highlight their tetrahedral shape. The lattice parameter is 2.8557Å. Hydrogen atoms preferentially occupy the tetrahedral interstices with a dissolution energy of approximately 0.3eV [73, 37]; marked with red arrows are the seven nearest-neighbour (nn) tetrahedral-tetrahedral adjacencies. The corresponding nn separations are: 1.01, 1.43, 1.75, 2.02, 2.26, 2.47 and 2.67Å, respectively.

1.1 The Fe-H system

It has been known for over 100 years [38, 75] that the presence of hydrogen (H) in metals – particularly steels/iron – can severely reduce ductility, leading to catastrophic failure below the yield-stress. The processes that cause these effects are collectively termed *hydrogen embrittlement* (HE).

Despite a century of research the core mechanisms of HE have yet to be fully understood and are still a topic of active research/debate [4]. The difficulty in understanding HE stems from its multi-scale nature; a full description of HE requires understanding of H-adsorption, H-diffusion/transport, and (most crucially) hydrogen interaction/influence with/on crystal defects. These processes span many orders of length/time scales, which presents challenges when isolating/connecting the impact of H at the atomistic scale to the macroscopic results.

Steel is a critical engineering material with almost 1.8 millions tonnes produced annually; as steels become stronger they become more susceptible to the effects of HE [49]. This, alongside the potential of the future hydrogen-economy [21], reinforces the need for a comprehensive understanding of HE in order to design better HE-resistant alloys.

In the primary phase of mild steels, BCC α–Fe (see Fig. 1.1), hydrogen has a low solubility but a high diffusivity [29]. Dissolved hydrogen can bind strongly to lattice defects [43] such as vacancies (V), dislocations, grain boundaries (GB), nano-voids, interstitials, etc. This binding reduces hydrogen mobility/diffusivity and has corresponding – but poorly understood – effects on the defect. It is common to distinguish between freely *diffusible hydrogen* in the lattice and *trapped hydrogen* bound to crystal defects. Theoretical frameworks, such as McNabb and Foster's kinetic model [55] and Oriani's kinetic-equilibrium trapping theory [64], attempt to abstract the atomic interactions. These go some way to understanding the impact of traps on H diffusivity (when appropriate physical parameters are available) however, fail to elucidate the causes of hydrogen embrittlement. Furthermore, quantum effects due to the low mass of the hydrogen nucleus may also play a significant role in H-transport/trapping [66], particularly for temperatures significantly below room temperature. This poses serious difficulties for purely classical theories.

1.1.1 Embrittlement mechanisms

A breadth of mechanisms for HE have been proposed, most of them revolve around the interactions between hydrogen and crystal defects. We shall briefly discuss a selection of the more prominent/successful ideas, for a more complete description see Dear and Skinner [18] and Barrera et al. [4]. Many of these mechanisms are supported by bodies of experimental work. As few are orthogonal to each other, it is likely that a full description of HE contains a combination of two/three of these mechanisms (alongside some yet undiscovered).

Hydrogen-induced decohesion (HID)   is one of the oldest explanations of HE; Pfeil [68] suggests that the presence of hydrogen in the lattice weakens the Fe-Fe bonding and reduces the strain at which decohesion occurs. Reports that H segregates to crack-tips [22] (due to the surrounding stress/strain fields) offers some support towards HID. The theory has been used to explain increases in crack-tip-opening angles during of hydrogen loaded samples [86, 30] however, direct experimental evidence for this decohesion has not been provided.

Adsorption-induced decohesion (AIDE)   likens HE to stress corrosion cracking and is motivated by experimental observations of crack growth in H-environments at crack-velocities faster than H-diffusion. Lynch [51] suggest the adsorption of H at the crack tip facilitates dislocation nucleation/emission. These dislocations assist the crack growth, which occurs through the linkage of microvoids ahead of the crack, this is sometimes referred to as *hydrogen assisted cracking* (HAC).

Hydrogen-enhanced localised plasticity (HELP) postulates that – possibly through the formation of Cottrell-like atmospheres [3] – hydrogen increases the mobility of dislocations causing increased plasticity [9], microvoid coalescence and ductile fracture. This plasticity is localised to regions containing high hydrogen concentrations. Unlike in AIDE, HELP predicts dissolved hydrogen ahead of the crack tip enhances dislocation mobility. Evidence for this increased mobility has been presented in FCC metals [24, 50, 76] but is limited in BCC metals. The H-enhanced mobility of dislocations could be due to core-effects or elastic screening.

Hydrogen-enhanced strain-induced vacancy (HESIV) formation originally proposed by Nagumo [59], explains HE as a consequence of hydrogen triggering a superabundance of vacancies [54, 61], which coalesce to form nanovoids/microvoids and offer a low-energy pathway for crack propagation. Vacancies clearly play a role in HE however, the scope of their effects and the interactions between V-H complexes and other defects is not completely clear.

Chapter 2

Literature review

In this chapter we establish the role of atomistic modelling in understanding hydrogen embrittlement before exploring in detail the background and cutting-edge of OLKMC modelling. In particular, we focus on the improvements to saddle-point searching that have been developed over the past two decades and break-down the methods used to recycle them between equivalent local-environments.

2.1 Atomistic modelling

The need for theoretical/computational modelling in the study of the Fe-H system stems from the inherent difficulty in experimental observations of atomic hydrogen [44]. The low solubility and high diffusivity [29] of H in BCC iron, combined with the small 'nucleus' and low electron density make direct experimental observations using electron-microscopy extremely challenging. Instead, techniques such as thermal desorption analysis (TDA) [71], electro-permeation (EP) experiments [70] and and atom probe tomography (APT) [82] are employed. Many of these methods (with the notable exception of APT) are unable to directly investigate HE on the atomic scale thus, we must fall back to computation/theory to unravel the atomic mechanisms that cause HE.

Many different modelling techniques have been used to investigate HE over varied assumptions and time/length scales. On the smallest length scales density functional theory (DFT) is used to study hydrogen preferred binding sites [17, 33, 62] and occasionally combined with molecular dynamic (MD) in *ab initio* MD to study H diffusion at the highest accuracy [78]. Additionally, work has been done using path-integral MD [45, 42] to explore H diffusion in iron while incorporating quantum effects.

While quantum effects are known to be important in the Fe-H system at lower temperatures [45], much success has been had modelling much larger systems using classical approaches. The most popular of these is MD and its accelerated-variants using semi-empirical potentials (reviewed in Section 3.1). This has enabled the study of H-defect kinetics, such as with grain-boundaries [88, 84] and dislocations [81, 83].

2.1.1   Monte-Carlo methods

Molecular dynamics (MD) simulations must resolve atomic vibrations in order to accurately track the dynamics of atom-scale systems. This imposes a significant computational effort as the integration time-step must be of-the-order of these vibrations. Hence, even using today's computers, MD simulation timescales rarely exceed $\mathcal{O}(100\mu s)$. For systems with kinetics dominated by rare events, much computational time is wasted (especially at low temperatures) while the system explores potential-energy (PE) basins. Monte-Carlo (MC) methods overcome this barrier by ignoring the explicit phase-space trajectory and instead focusing on the basin→basin transitions. This can significantly accelerate simulations.

The first MC method was the Metropolis–Hastings (MH) algorithm [32], employed in studying magnetic crystal domains. In the MC framework basins of the PES are represented as states in a Markov chain, linked together by the mechanisms between them and their corresponding (time-independent) transition probabilities:

$$P_{ij}: \text{probability of transition } i \rightarrow j \tag{2.1}$$

In MH MC the energy change between states is used to compute the transition-probability according to the Boltzmann distribution at equilibrium. This constrains MC to measuring only equilibrium properties of the system. To overcome this limitation the (traditional) Kinetic Monte Carlo (KMC) framework uses instead the rate constants:

$$\Gamma_{ij}: \text{probability per unit time of transition } i \rightarrow j \tag{2.2}$$

to describe the transition probabilities. This enables KMC to probe non-equilibrium phenomena, using the rejection-free n-fold way algorithm [12]; with the system in state i, the next state k is selected as the solution to:

$$\sum_{j=1}^{k-1} \Gamma_{ij} < \rho_1 \sum_{j=1}^{n} \Gamma_{ij} \leq \sum_{j=1}^{k} \Gamma_{ij} \tag{2.3}$$

where $\rho_1 \in (0, 1]$ is a uniform random number and $j, k \in \{1, 2, \ldots, n\}$. The rate constants model a Poisson process and therefore the time elapsed during a single MC step is [11]:

$$\Delta t = \frac{-\ln{(\rho_2)}}{\sum_{j=1}^{n} \Gamma_{ij}} \tag{2.4}$$

where $\rho_2 \in (0, 1]$ is a second uniform random number. Both equations require knowledge of the complete set of non-zero rate constants connected to a given state. In KMC these rate constants are typically pre-computed and the mechanisms known *a priori*. Furthermore, in order to make building a rate-catalogue (a list of all non-zero rate mechanisms exiting a given state) tractable, most KMC models require embedding the system onto a lattice. This further specialises the method and makes it particularly difficult to apply to partially-ordered/disordered systems. Nevertheless, KMC has been used successful in studies of HE [11, 23], notably the time-scales reached are often in the-order-of seconds.

2.1.2   Off-lattice kinetic Monte-Carlo

Kinetic MC methods, which build/adapt their rate catalogues on-the-fly, seek to overcome the limitations of traditional KMC and go by a variety of names including: *"adaptive"*, *"off-lattice"*, *"on-the-fly"*, *"self-learning"* and *"self-evolving"*. These methods are partitioned into two major categories: those that modify the rate-constants of the mechanisms based on the atomistic local environment (LE) [72] and those that dynamically discover mechanisms and calculate their rate constants by searching the local potential-energy surface (PES). The latter are a substantially more powerful set of methods, they can be used to investigate systems when the mechanisms are too complex to predict *a priori*. We shall use the term *off-lattice* kinetic Monte-Carlo (OLKMC) to refer exclusively to these mechanisms-constructing methods.

The OLKMC method was originally described by Henkelman and Jónsson [35]. A set of atoms is represented in continuous space and a potential-energy function is supplied that encodes the interaction between atoms:

$$U\colon \mathbb{R}^{3n} \to \mathbb{R} \tag{2.5}$$

the negative gradient of the PE is the corresponding force-field. Henkelman and Jónsson obtain the set of accessible mechanisms by exploring the potential energy surface (PES) using the *dimer* method [34]. The dimer method finds saddle-points connected to some initial basin without requiring a second target-basin (unlike the nudged elastic band method [39] which requires both) and does not require calculation or inversion of the complete Hessian matrix

(unlike classical minimum-mode following methods [13]). The dimer method is reviewed in detail in Section 2.1.3.

After finding all the saddle-points surrounding some initial basin, small displacements along the minimum-mode at each saddle-point and subsequent energy minimisations are made. Hence, mechanisms – pathways linking the initial-basin→SP→final-basin – are found on-the-fly. By applying the harmonic transition state theory (HTST) approximation, the rate constant connecting basins $i \rightarrow j$ via the single, first-order, saddle-point ‡ is described by the Arrhenius equation [87]:

$$\Gamma_{ij}^{\text{TST}} = \tilde{\nu}_{ij} e^{-\beta \left(E^{\ddagger} - E^{i} \right)} \tag{2.6}$$

where $\beta = \frac{1}{k_B T}$, $\tilde{\nu}_{ij}$ is the attempt frequency or Arrhenius prefactor and $E^{\ddagger}$, E^{i} are the energies (computed using the force-field) of the system at the saddle point and state i, respectively. It is common to set $\tilde{\nu}_{ij} = \nu$ a constant [56, 85]. The HTST approximation holds well for metals away from their melting point as the atoms are rigidly held in place and the basins well approximated by a quadratic expansion [87].

After the mechanisms have been discovered and their rate constants computed the simulation precedes as standard KMC; a mechanism is selected according to Eq. (2.3) and time advanced as prescribed by Eq. (2.4). At this stage no effort is made to deal with the flickering-problem (the simulation getting stuck inside a group of states linked by low-energy-barrier/fast-rates) and – although Henkelman and Jónsson [35] note that mechanisms are typically local – no attempt is made to recycle the work that is done by saddle-point searches done on similar environments.

Augmentations

At its core OLKMC is a combination of two simple processes: saddle-point searching and KMC. However, the design-space for optimisation is broad. One such optimisation is the introduction of active volumes (AVs) by Xu, Osetsky, and Stoller [93] in their self-evolving atomistic KMC (SEAKMC). These are defect-containing regions of the simulation, as defined by high per-atom potential-energy, in which the SP searches and minimisations are carried out. Active volume boundaries are automatically updated to account for long-range interactions. The method massively reduces the computational complexity for large, defect-sparse systems by reducing the number of atoms included in the force-field evaluation. However, this acceleration could be gained by using local force-fields and pruning high-barrier events from the event list. This would allow for finer control of which events are excluded and remove the difficulty of needing to define appropriate boundaries for the AVs (likely problem dependant).

Later work by Béland et al. [7] compare SEAKMC [93, 92] to OLKMC methods using both dimer and ART-*nouveau* [6] SPS methods. They establish the dimer method is marginally faster (requires fewer force calls) and importantly captures a broader distribution of saddle-points. All the methods rely on HTST but, make very different assumptions about the locality of events. Convergence of dynamic behaviour between the methods is reassuring and gives confidence that on-the-fly KMC methods produce physical results. The authors make an interesting suggestion to initialise SP searches with MD coordinates.

Rejection based MC sampling, as used in the original Metropolis–Hastings (MH) algorithm [32] is suggested by Ruzayqat and Schulze [77] as an alternative to N-fold way algorithm for OLKMC. They first partition the mechanisms into N sets, each set corresponding to an atom; then calculate a rate-sum upper-bound for each set using a nearest-neighbour approximation; select a set using the N-fold way algorithm; launch saddle-point searches to find the mechanisms in the chosen set; finally possibly reject or accept one of these mechanisms. This method has the advantage of constraining SP searches to a local region but, requires an estimate of the upper-bound rate that must be fitted (as function of number of 1st, 2nd... nearest neighbours) for each problem. This, alongside the additional error that comes with making an estimation, make rejection sampling a less promising avenue for optimisation than exploring the locality of mechanisms.

2.1.3   Saddle-point finding and the dimer method

One of the key aspects of OLKMC is the efficiency of the saddle-point search (SPS) procedure. Minimum-mode following methods, originally described by Cerjan and Miller [13], find saddle-points of the PES by climbing from a local minima to an adjacent SP. This is achieved by inverting the component of the force parallel to the minimum eigen-mode, $\hat{N}$, of the PES:

$$\boldsymbol{F}_{\text{eff}} = -\boldsymbol{\nabla}U_0 + 2\hat{N}\boldsymbol{\nabla}U_0^{\mathsf{T}}\hat{N} \tag{2.7}$$

Subsequently, translating along this force maximises the energy along the minimum-mode and minimises the energy along all other modes. Hence, converging to a local SP [41]. Unfortunately, explicit calculation of the minimum-mode from the full Hessian matrix requires compute-intensive matrix operation that scale as $\mathcal{O}(n^3)$. Additionally, just evaluating the Hessian requires second derivatives of the PE, which may not always be available/tractable.

A common approach for finding saddle-points linking adjacent basins is the nudged elastic band (NEB) [39] method. Unfortunately, this requires prior-knowledge of the final-state which is not available in OLKMC.
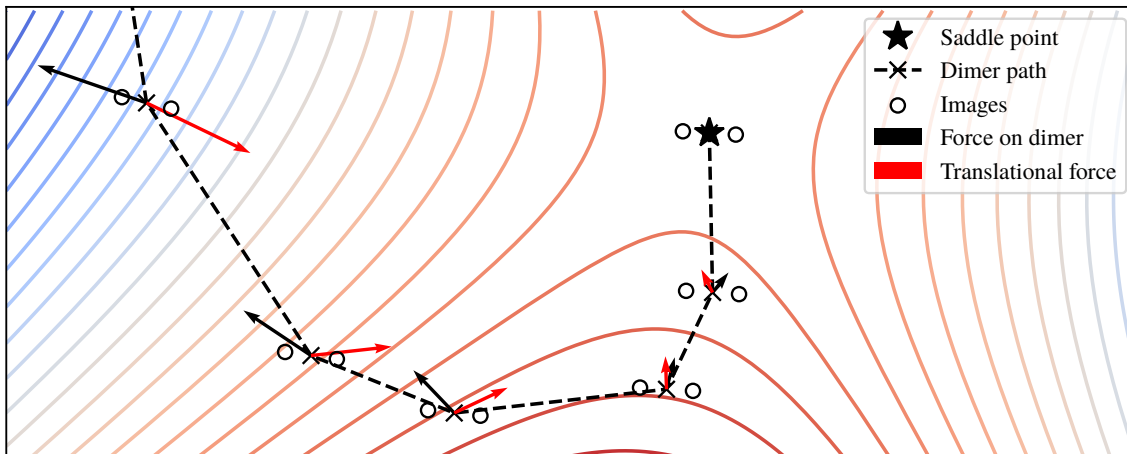
Fig. 2.1 Diagram of a dimer converging to a SP in a model two-dimensional LEPS potential coupled to an harmonic oscillator [39]. The dimer is rotated between each step to minimise its rotational PE, aligning its axis (line between images) with the minimum-mode. The displacement between image-pairs for each dimer has been exaggerated for clarity. The force and the effective translational force (Eq. (2.7)) on the dimer are marked for each step.

Alternatively, Henkelman and Jónsson's [34] dimer method, can efficiently compute the minimum-mode of the Hessian by avoiding simultaneous computation of other eigen-modes. This is achieved through a construction called the *dimer*, formed from a pair of images of the system displaced by a small amount. Successive rotations, minimising the total PE, orient the dimer with lowest curvature mode. During rotation, they approximate the curvature along the dimer-axis, this only requires first-derivatives of the PE and avoids matrix operations. Figure 2.1 summarises how the dimer converges to a SP. By varying the initial orientation of the dimer and making small random perturbations to the initial state, different saddle points can be found.

In our implementation, developed in previous work [90], we predominantly follow the formulation of the dimer method presented by Kästner and Sherwood [41], which contains several optimisations over the original method – notably this includes Heyden, Bell, and Keil's [36] improvements:

- Working with the potential gradients at the centre and one end of the dimer (as opposed to each end), such that the central gradient can be reused during rotations. This halves the number of force calculations during dimer translations and reduces the number of force calculations during rotation from six to four.

- A Fourier series expansion of the curvature during rotation and approximation of its second derivative, enabling an improved guess for the dimer rotation angle. This accelerates convergence to the minimum-mode.

- Exploiting the planarity of the rotated dimer to extrapolate one of the gradients at the dimer end, thus saving a force call per rotation.

- Introduction of the limited-memory formulation of the Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) optimiser [48, 63] to determine the plane of rotation. This is demonstrated to converge faster than original conjugate gradient (CG) method.

- Introduction of L-BFGS optimiser during the dimer translation step to predict both the step-size and direction. This elides the force calculations during the line search and, being a quasi-Newton method, can lead to superlinear convergence once sufficiency close to the SP [63, 41].

Kästner and Sherwood [41] do not fully describe the coupling of their dimer method to the L-BFGS optimiser however, their results clearly demonstrate improvement. This is confirmed by Zeng, Xiao, and Henkelman [95] who unified several minimum-mode following algorithms under one mathematical framework. They prove all investigated methods are bounded in efficiency by the Lanczos method [47]. In numerical tests the L-BFGS dimer method was shown to converge almost as fast as the Lanczos method but require fewer force evaluations. This was the basis for our choice to use the dimer method.

In our implementation developed in previous work [90], we diverge slightly from the formulation of Kästner and Sherwood [41] during the dimer translation step. We still use the L-BFGS algorithm for determining the translation direction and step size but, introduce a trust-radius based approach to limit the step-size. The maximum step size, s_{trust}, is scaled according to the success of the previous steps; the projection of the effective gradient on the search direction is calculated after a step:

$$P = -\boldsymbol{F}_{\text{eff}}^{\mathsf{T}}\boldsymbol{p} \tag{2.8}$$

where $\boldsymbol{p}$ is the approximate Newton step, computed using the L-BFGS method. An ideal step length would have $P = 0$. Hence, we increase s_{trust} when $P < -P_{\text{tol}}$ and decrease s_{trust} when $P > P_{\text{tol}}$. Additionally, we bound s_{trust} such that $s_{\text{min}} < s_{\text{trust}} < s_{\text{max}}$.

2.1.4   Completeness of the rate catalogue

A key assumption (and point of criticism) in OLKMC is the confidence in the completeness of the rate catalogue. This is a fundamentally-hard problem as we are attempting to estimate the support size of an arbitrary and unknown distribution from a finite and probably *small* sample. It is clearly impossible to estimate the support size without some underlying assumptions about the shape of the distribution. Very general assumptions have been explored from a mathematical perspective [65, 91] however, we shall focus on the literature relevant to OLKMC where the assumptions can be stronger.

Most implementations of OLKMC follow Henkelman and Jónsson [34] and launch a fixed number of SPS per basin – usually with some initial validation runs to establish what a large enough number of searches is. This has two clear flaws: some states encountered later in the simulation may have many mechanisms and thus, require many more SPS; it provides no quantification of the certainty with which we have found all relevant saddle-points. Correspondingly, as the sum of the rate-constants is used in Eq. (2.4), there should be an error on the elapsed time (this idea is explored fully by Wang [89] in their reliable KMC (R-KMC) method). Additionally, if some fast mechanisms are missed, the simulated kinetics could become unphysical.

The first steps toward quantifying the error in the rate catalogue was Xu and Henkelman's [94] introduction of a confidence parameter to asses the probability a certain fraction of the saddle-points surrounding a basin have been found:

$$C = \frac{1}{\alpha N_r} \tag{2.9}$$

with N_r the number of consecutive repeated saddle-points discovered and the minimum probability of finding a SP: $P_{\min} = \frac{\alpha}{N_p}$. They derive there confidence parameter through analysis of the uniform-probability case (which has long been known to be false [35, 2]) but, incorporate non-uniform probabilities through the α parameter. This only shifts the burden to estimating α. Moreover, their confidence parameter makes no use of most of the information extractable from the sample (e.g. estimation of the relative probabilities of the sample saddle-points). Finally, their confidence-parameter only provides a confidence in having found a fraction of the saddle-points; as previously noted, missing any could be crucial. It would be interesting to apply their ideas but, use the fraction of total basin-escape rate instead of the number of saddle-points found, this could gives a direct error in the rate-sum.

Chill and Henkelman [15] explore using high temperature MD (with periodic minimisations) as an alternative to SP searches. They develop an error estimator for the completeness of

rate catalogue using the discovery-time – the total MD simulated time. This gives the most quantitative estimates thus far as, the probability of MD finding a transition is proportional to its transition rate. Unfortunately 4-6 times more force-field evaluations per SP are required compared to the dimer method. This is still a very promising avenue as, the potential for biasing the MD simulation away from previously discovered saddle-points could reduce the number of repeated SP discoveries.

Chatterjee [14] conducts a more rigorous analysis, building KMC networks from MD data. They focus on formulating a validity time for the KMC network (how long a KMC simulation running in derived network could be simulated before significant leakage to unknown and unexplored states). They are able to estimate the total rate to unknown states, see Bhute and Chatterjee [8] for more detail. Their framework is a two stage model, MD followed by KMC. In OLKMC+MD-SPS would iterate between the two hence, their model would need to be adapted but, could provide an improved error-analysis over Chill and Henkelman's [15]

An alternative approach is taken by Alexander and Schuh [2] to quantify the completeness of their OLKMC rate catalogue. They use a deterministic search-method choosing starting-points uniformly spaced on concentric shells/3-spheres (perturbing a single atom) instead of randomly perturbing the initial basin. It is hoped that deterministic-searches result in a better sampling of the adjacent saddle-points. Unfortunately, they demonstrate only a weak relationship between perturbation direction and resulting SP. Furthermore, perturbing a single atom could result in biased searches. Finally, they advocate quantifying the error in the basin residence-time as a system dependant parameter, this is not easily generalisable.

To overcome some of the difficulties we have discussed; we develop a confidence parameter inspired by Xu and Henkelman [94] in Section 3.3.

2.1.5   Saddle-point recycling

The most computationally expensive element of OLKMC is the SPS procedure; each SPS requires many hundreds of calls to the force-field and many SPS must be carried out to ensure the completeness of the KMC catalogue. Due to the local nature of mechanisms, it was soon realised that most of these SPS are unnecessary. For example, in a section of perfect lattice the local environment (LE) around each atom is identical hence, the mechanisms that can occur at each atom are identical. Secondly, consider two atoms sufficiently far apart; a local mechanism centred on one will likely not change the LE around the second hence, its accessible mechanisms remain the same. Finally, many atoms are in LE's differing only by an Euclidean transformation (of the form $x \mapsto Rx + c$, with R an orthogonal matrix) hence, their mechanisms are related by

the same transformation. Multiple methods have been developed to reduce the cost of building the KMC catalogue by exploiting this locality.

System-wide methods

The simplest form of SP recycling, described by Xu and Henkelman [94], accelerates SPS in a new state by attempting to combine the saddle-points from the previous state with the current state. Letting the system of n atoms be represented by the vector $\boldsymbol{R} \in \mathbb{R}^{3n}$, then the vectors $\boldsymbol{R}^{\text{init}}$, $\boldsymbol{R}_i^{\text{sp}}$ and $\boldsymbol{R}_i^{\text{final}}$ are the initial, saddle and final states of the system after undergoing the ith mechanism. An initial guess for position of the kth atom, in the ith saddle-point, after undergoing the jth mechanism is:

$$\boldsymbol{R}_{i,k}^{\text{sp, recycle}} = \begin{cases} \boldsymbol{R}_{i,k}^{\text{sp}} & \left\| \boldsymbol{R}_{j,k}^{\text{final}} - \boldsymbol{R}_{j,k}^{\text{init}} \right\| < dr \\ \boldsymbol{R}_{i,k}^{\text{final}} & \text{otherwise} \end{cases} \tag{2.10}$$

where dr is a small tolerance controlling how far an atom has to move during a mechanism for it to be considered significant. These guesses are then (quickly) refined using a few iterations of a saddle-point search procedure to converge to the true saddle-points of the new state.

   This accelerates the building of the rate catalogue for a given MC step by extrapolating the rate catalogue of the previous MC step. However, it does not allow for the propagation of mechanisms from many MC steps previous and – as it acts on the entire system, $\boldsymbol{R}$ – does not exploit any translational or rotational symmetries of the internal LE's. Furthermore, it still requires calling the SPS procedure for every possible mechanisms in the catalogue at each step.

Local methods

More advanced methods seek to classify the LE around each atom in the system. It is then possible to associate mechanisms entirely within a LE. They can then be cached and if an equivalent LE is discovered, instead of launching new SPS, the mechanisms can be reconstructed from the cached information. Two key challenges arise with these methods, classifying the LE's into discrete categories and transforming mechanisms between equivalent LE's.

   Classifying LE's is a form of pattern recognition; Shah et al. [80] develop a 2D scheme for adatom diffusion on the FCC (111) surface for use alongside an adaptive KMC algorithm. Unfortunately, due to the presence of the (111) surface being used as a reference frame, the resulting LE categorisation method is not suitable for general OLKMC. They use symmetry transformations to reduce the size of the catalogue by searching for a finite number of

symmetrically transformed LE's. This highlights the need for a classification scheme that is invariant under Euclidean transformations. This is because in OLKMC there are an infinite number of possible rotations/reflections; they could not all be systematically stored in the catalogue.

Nandipati et al. [60] extend the 2D scheme [80] to 3D and generalised it to function without a reference frame. They explicitly wished to avoid the errors associated with the precision of the atomic positions in off-lattice methods. They associated mechanisms with the LE of the most displaced atom. Electing to use a local Cartesian grid, fine enough such that each cell has binary occupation. Unwrapping this grid, a single binary number can then encode the complete LE. This method does not exploit the symmetry of the LE, instead the authors store in the catalogue several additional symmetry-transformed versions of the LE. This requires more memory overhead and overlays a lattice frame of reference. Finally, they do nothing to account for small errors in the atoms positions near cell boundaries, this means some local environments may have multiple keys associated with them, a source of inefficiency.

Moving toward an Euclidean transformation invariant classification, Konwar, Bhute, and Chatterjee [46] present a system that stores the LE of an atom at r_i as $\left\{ r_{ij} \mid r_{ij} < r_{\text{env}} \right\}$. LE's centred on atoms i and i' are then considered equivalent when:

$$\forall j \; \exists j' \text{ such that } \left\| r_{ij} - r_{i'j'} \right\| < \delta r \tag{2.11}$$

This method gracefully allows for some error on the positions of atoms in a LE. Equivalence can be tested for after using their *atom-by-atom matching* procedure; this is computationally expensive hence, they first prune the search-space by using some discrete heuristics (species of central atom, number of atoms, discretised RDF). Although their definition of LE equivalence is invariant under Euclidean transformations their *atom-by-atom matching* does not take into account rotations/reflections. Therefore, many symmetry-related LE's are still classified as unique/inequivalent. Additionally, (transformation invariant) information is available in the full set of atomic separations inside each LE that is currently unexploited.

Topological methods

Topological methods are a form of local SP recycling introduced by El-Mellouhi, Mousseau, and Lewis [56] that fully exploit the symmetry of LE's, enabling mechanisms to be reconstructed from LE's related by arbitrary euclidean transformations. The generalised topological classification process is summarised graphically in Fig. 2.2. In brief, atoms in the LE are used to draw a graph; atoms become nodes and atoms considered bonded ($r_{ij} < r_{\text{edge}}$) are connected with an
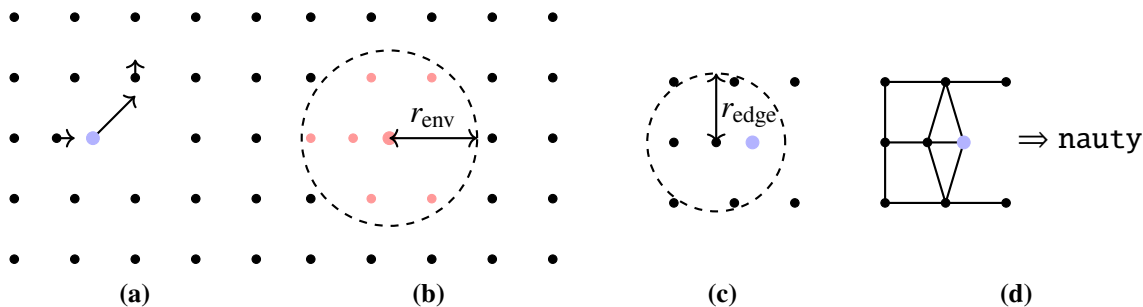
Fig. 2.2 Summary of the (simplified) topological mechanism caching process. Fig. 2.2a the minima→minima displacement vectors of a (hypothetical) mechanism are marked; an H atom (blue dots) bound to a vacancy in the α-Fe lattice (black dots) migrates to an adjacent site. Fig. 2.2b highlights (in red) the atoms within r_{env} of the maximally-displaced/central atom. Fig. 2.2c shows the set of atoms which constitute the LE in isolation alongside r_{edge} the maximum edge length. Finally, Fig. 2.2d is the (coloured) graph, built from the LE, that is passed to `nauty` [53] which produces the topological key, canonical form and reconstruction transformations then stored in the catalogue.

edge. The question of whether LE's are equivalent is then a question of whether their graph representations are isomorphic. This is, in general, a problem in its own complexity class GI ∈ NP which is not known to be in either P or NP-complete [27]. Fortunately, there exists implementations such as McKay and Piperno's `nauty`[1] software which can solve this problem in polynomial time for many graphs.

Using `nauty` a LE can be transformed into a canonical ordering and a discrete topological-key generated, from the canonical adjacency matrix, G, of corresponding graph. This key can then be used in an associative key-value container/catalogue to store mechanisms in a canonical form that can be reconstructed onto equivalent LE's.

We adopted this methodology in previous work [90] however, topological classification relies on a one-to-one correspondence between topology and geometry. Initially we found this correspondence to break-down in the Fe-H system due to the small size of the H-atom and small displacements during mechanisms. We overcame these problems: firstly extending the definition of the adjacency matrix:

$$G_{ij} = \begin{cases} 1 & r_{ij} \leq r_i^{edge} \\ 0 & \text{otherwise} \end{cases} \tag{2.12}$$

i.e allowing r_{edge} to vary for each atom, to ensure atoms near the surface of the LE are fully specified. Secondly, colouring each atom as the pair formed from the atoms' atomic number

---

[1]https://pallini.di.uniroma1.it/

and the local sum:

$$\left\lfloor c \sum_j G_{ji} r_{ij} \right\rfloor \tag{2.13}$$

with c a problem-dependant scaling constant. This encodes much more of the information contained within $\{ r_{ij} \}$ into the (di)graph.

These extensions allowed us to restore the one-to-one correspondence between topology and geometry. Unfortunately, before (by choosing r_{edge} between the first and second nearest neighbour distances in α-Fe) the classification was less sensitive (but not completely tolerant) of small errors in the atomic positions. With the above modifications however, infinitesimal perturbations in position can result in many keys being associated with the same geometry. This, alongside the more fundamental problem of having no explicit/quantitative link between topological keys and the similarity of LE's, pushed us to move away from a topological classification scheme. Section 3.2 details our alternative solution to the problem.

2.1.6 Superbasins and the low-barrier problem

As previously eluded to in Section 2.1.2 a common issue encountered during OLKMC simulations is the *flickering/low-barrier problem* [56, 85]. This occurs when a collection of basins – often called a *superbasin* – are connected by a series of fast mechanisms. It requires many MC steps to escape from a superbasin. As the rate-sum, $\sum_{j=1}^{n} \Gamma_{ij}$ in Eq. (2.4), is very large during this period, the simulated time advances very slowly.

Two exact solutions to this problem are presented by Fichthorn and Lin [25]; the key insight is the partitioning of states into *transient* and *absorbing* sets, followed by analytically solving the motion inside the transient states. They use the average simulation time-step as a criterion for detecting when a superbasin is encountered: $\langle \Delta t \rangle_n < \Delta t_{min}$. Other analytical solutions have been proposed [69, 6] that compute the mean superbasin-escape time.

Contrastingly, simpler TABU-like [31] methods that ban recent-transitions have been employed [56, 16]. These have been shown to be thermodynamically sound providing the total number of KMC steps is much greater than the oldest banned transition.

Ramasubramaniam et al. [72] take a further simplified approach to superbasins; combining states connected by *fast* mechanisms into a single state and ignoring all internal superbasin kinetics. Although this is clearly not exact they have some success measuring bulk diffusion coefficients.

In previous work [90] we attempted to use a TABU method however, during testing we discovered this method produced large biases in the results. Therefore, we extended Puchala, Falk, and Garikipati's [69] *mean rate method* (MRM) by further partition the mechanisms into transient and absorbing categories. This allowed us to apply the MRM cleanly to OLKMC without having to fully explore all the states in a superbasin.

2.2   Summary

In summary, OLKMC is an unbiased tool being successfully applied to study the kinetics of various metal-systems. A multitude of avenues for optimisations (massively-parallel implementations, improved SP recycling, swarm based SPS methods, etc) are still unexplored, many of which could allow for simulations that further our current understanding of HE. Although, the accuracy of the underlying potentials will always be a limiting factor, they are an ever improving field. Off-lattice KMC allows for the exploration of previously inaccessible time-scales at atomic fidelity. As such, it is the perfect tool to explore the uncertain and complex mechanisms controlling HE.

Chapter 3

Theory

This chapter is devoted to the theory supporting our OLKMC implementation. In Section 3.1, we summarise the semi-empirical potentials employed in this work and present analytical expressions for their first and second derivatives. Section 3.2 details the invariant and tolerant local-environment classification scheme we have developed to replace topological methods during saddle-point recycling. Finally, in Section 3.3 we detail an alternative confidence parameter used as a stopping criterion for the SPS procedure.

3.1 Embedded atom method potentials

The limiting level of theory in OLKMC is the underlying potential. Unfortunately, during SPS many calls to the force-field are required thus, the potential also becomes the performance-limiting factor. Hence, *ab initio* approaches are out of reach if we want to reach HE timescales – of-the-order-of seconds.

The embedded atom method (EAM) – originally developed by Finnis and Sinclair [26] – is a fast, well tested, semi-empirical model of the potential energy of a collection of atoms. EAM potentials are sufficiently short range to enable fast MD simulations. Although they are not without criticism [58], EAM potentials have become well established in the literature.

We use the slight variation presented and fitted by Ramasubramaniam, Itakura, and Carter [73]. They generalise the EAM embedding function and fit to first-principles (DFT) measurements and experimental data. They fit to wide variety of targets (perfect crystal, point defect, surface measurements, etc) and their potential provides good reproduction of several crystal defect structures, including the sixfold symmetry of screw dislocation cores. We also include the modifications by Song and Curtin [81] who, introduce additional H-H repulsion to reduce the hydrogen clustering observed in the original potentials.

Adopting the notation of Greek superscripts for atom indexes (and later Roman subscripts for vector components) the EAM potential form is:

$$U = \frac{1}{2} \sum_{\substack{\alpha,\beta \\ \alpha \neq \beta}} V^{\alpha\beta} + \sum_{\beta} F^{\beta} \tag{3.1}$$

with $V^{\alpha\beta}$ a symmetric pair-potential function (such that $V^{\alpha\beta} = V^{\beta\alpha}$) acting on $r^{\alpha\beta} = \left\| \boldsymbol{r}^{\alpha\beta} \right\| = \left\| \boldsymbol{r}^{\beta} - \boldsymbol{r}^{\alpha} \right\|$ the atomic separations. Additionally, F^{β} is the embedding function acting on ρ^{β}, the electron density of the βth atom:

$$\rho^{\beta} = \sum_{\alpha \neq \beta} \phi^{\alpha\beta} \tag{3.2}$$

where $\phi^{\alpha\beta}$ describes the electron density of the αth atom acting at the βth atom. Strictly speaking, the embedding and potential functions are empirically chosen and do not represent physical electron densities [26]. Finally, the functional forms of V and ϕ are chosen such that:

$$V\left(r^{\alpha\beta}\right) = \phi\left(r^{\alpha\beta}\right) = 0 \quad \text{when} \quad r^{\alpha\beta} > r_{\text{cut}} \tag{3.3}$$

Hence, the potentials are described as *local* and the atoms within a particular cut-off radius are described as the central atoms' *neighbours*.

Due to the complexity of the analytic forms of the fitted functions and the high cost of evaluating them, we pre-compute and tabulate the 9 EAM functions. During the simulation we pre-compute and store the natural cubic spline coefficients [1] from the tabulated values, then reconstruct the smooth functions and their differentials appropriately.

3.1.1   First derivatives

The gradient of the potential is required for saddle-point finding methods; noting (see Appendix A.1 for a derivation):

$$\frac{\partial r^{\alpha\beta}}{\partial r_j^{\gamma}} = \left(\delta^{\gamma\beta} - \delta^{\gamma\alpha} \right) \hat{r}_j^{\alpha\beta} \tag{3.4}$$

the first derivative of the potential takes the form:

$$\frac{\partial U}{\partial r_j^\gamma} = \frac{1}{2} \sum_{\substack{\alpha,\beta \\ \alpha \neq \beta}} \dot{V}^{\alpha\beta} \left(\delta^{\gamma\beta} - \delta^{\gamma\alpha} \right) \hat{r}_j^{\alpha\beta} + \sum_\beta \dot{F}^\beta \sum_{\alpha \neq \beta} \dot{\phi}^{\alpha\beta} \left(\delta^{\gamma\beta} - \delta^{\gamma\alpha} \right) \hat{r}_j^{\alpha\beta}$$

$$= \sum_{\alpha \neq \gamma} \dot{V}^{\alpha\gamma} \hat{r}_j^{\alpha\gamma} + \sum_{\substack{\alpha,\beta \\ \alpha \neq \beta}} \dot{F}^\beta \dot{\phi}^{\alpha\beta} \left(\delta^{\gamma\beta} - \delta^{\gamma\alpha} \right) \hat{r}_j^{\alpha\beta}$$

$$= \sum_{\alpha \neq \gamma} \left(\dot{V}^{\alpha\gamma} + \dot{F}^\gamma \dot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \dot{\phi}^{\gamma\alpha} \right) \hat{r}_j^{\alpha\gamma} \qquad (3.5)$$

where we use an over-dot to denote differentiation and have applied the antisymmetry of $\boldsymbol{r}^{\alpha\beta}$.

The sum runs-over the neighbours of gamma. The gradient can be efficiently computed using two loops over all atoms, see Algorithm 3.1 for pseudocode.

Algorithm 3.1 Two-loop EAM gradient.

> **procedure** GRAD
>> **for** each atom, γ **do**   ▷ Loop 1.
>>> Compute and store ρ^γ by summing over neighbours of γ ▷ Equation (3.2).
>> **for** each atom, γ **do**   ▷ Loop 2.
>>> Compute $\frac{\partial U}{\partial r_j^\gamma}$ by summing over neighbours of γ ▷ Equation (3.5).

3.1.2   Second derivatives and Vineyard theory

Although most OLKMC implementations make the constant pre-factor approximation in Eq. (2.6) it is possible to explicitly calculate $\tilde{v}_{ij}$ using HTST by assuming the PES is quadratic near the saddle point:

$$\tilde{v}_{ij} = \frac{\prod_{k=1}^N v_k^i}{\prod_{k=1}^{N-1} v_k^\ddagger} \qquad (3.6)$$

where $v_k^\ddagger$, v_k^i are the real normal-mode frequencies at the saddle point and state i respectively. However, Chill and Henkelman [15] identify large variations in the harmonic pre-factor during Al adatom diffusion events (5.8×10^{12} to $2.0 \times 10^{14} \text{s}^{-1}$ for top three events). This is evidence to start calculating $\tilde{v}_{ij}$, rather than relying on constant approximation.

In order to do this we shall require the explicit Hessian of the potential energy; differentiating Eq. (3.5) we obtain:

$$\frac{\partial^2 U}{\partial r_i^\eta \partial r_j^\gamma} = \sum_{\alpha \neq \gamma} \hat{r}_j^{\alpha\gamma} \frac{\partial}{\partial r_i^\eta} \left(\dot{V}^{\alpha\gamma} + \dot{F}^\gamma \dot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \dot{\phi}^{\gamma\alpha} \right)$$

$$+ \sum_{\alpha \neq \gamma} \left(\dot{V}^{\alpha\gamma} + \dot{F}^\gamma \dot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \dot{\phi}^{\gamma\alpha} \right) \frac{\partial \hat{r}_j^{\alpha\gamma}}{\partial r_i^\eta} \tag{3.7}$$

which after careful manipulation (see Appendix A.2) resolves to:

$$\frac{\partial^2 U}{\partial r_i^\gamma \partial r_j^\gamma} = \ddot{F}^\gamma \mu_i^\gamma \mu_j^\gamma + \sum_{\alpha \neq \gamma} \left[A^{\alpha\gamma} \delta_{ij} + \left(A^{\alpha\gamma} - B^{\alpha\gamma} - \ddot{F}^\alpha \dot{\phi}^{\gamma\alpha} \dot{\phi}^{\gamma\alpha} \right) \hat{r}_i^{\alpha\gamma} \hat{r}_j^{\gamma\alpha} \right] \tag{3.8}$$

$$\left. \frac{\partial^2 U}{\partial r_i^\eta \partial r_j^\gamma} \right|_{\eta \neq \gamma} = (B^{\eta\gamma} - A^{\eta\gamma}) \hat{r}_i^{\eta\gamma} \hat{r}_j^{\gamma\eta} - A^{\eta\gamma} \delta_{ij} + \ddot{F}^\eta \dot{\phi}^{\gamma\eta} \mu_i^\eta \hat{r}_j^{\eta\gamma} - \ddot{F}^\gamma \dot{\phi}^{\eta\gamma} \hat{r}_i^{\eta\gamma} \mu_j^\gamma - O_{ij}^{\eta\gamma} \tag{3.9}$$

where we have introduced the electron-density dipole:

$$\mu_i^\beta = \sum_{\alpha \neq \beta} \dot{\phi}^{\alpha\beta} \hat{r}_i^{\alpha\beta} \tag{3.10}$$

symmetric tensors A, B such that:

$$A^{\alpha\beta} = \frac{\dot{V}^{\alpha\beta} + \dot{F}^\beta \dot{\phi}^{\alpha\beta} + \dot{F}^\alpha \dot{\phi}^{\beta\alpha}}{r^{\alpha\beta}} \quad \text{and} \quad B^{\alpha\beta} = \ddot{V}^{\alpha\gamma} + \dot{F}^\gamma \ddot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \ddot{\phi}^{\gamma\alpha} \tag{3.11}$$

and overlap term:

$$O_{ij}^{\eta\gamma} = \sum_{\alpha \neq \gamma, \eta} \ddot{F}^\alpha \dot{\phi}^{\gamma\alpha} \dot{\phi}^{\eta\alpha} \hat{r}_i^{\eta\alpha} \hat{r}_j^{\alpha\gamma} \tag{3.12}$$

Similarly to the gradient, this can be efficiently computed by looping twice over all the neighbours of each atom, see Algorithm 3.2 for pseudocode.

3.2 Invariant and tolerant local-environment equivalence

The objective of our classification scheme is to design a method to efficiently store a catalogue of LE's, each with an associated set of mechanisms (represented by a displacement vector,

Algorithm 3.2 Two-loop EAM Hessian.

procedure HESS
    for each atom, γ **do**   ▹ Loop 1.
        Compute and store ρ^γ and $\boldsymbol{\mu}^\gamma$ ▹ Equation (3.2), Equation (3.10).
    for each atom, γ **do**   ▹ Loop 2.
        Build a list, n, of the neighbours of γ.
        Iterating n: compute $\frac{\partial^2 U}{\partial r_i^\gamma \partial r_j^\gamma}$ ▹ Equation (3.8).
        Iterating n: compute contributions to $\left.\frac{\partial^2 U}{\partial r_i^\eta \partial r_j^\gamma}\right|_{\eta\neq\gamma}$ ▹ Equation (3.9).
        Iterating all pairs, (α,β), in n: compute contributions to $O_{ij}^{\alpha\beta}$ ▹ Equation (3.12).

minim→sp→minima, for each atom in the LE). It must be possible to efficiently match a new LE to an equivalent LE in the catalogue and compute the required set of transformations to map the stored mechanisms onto the new LE. Ideally the equivalence should be invariant under infinitesimal perturbations of atomic positions and must be invariant under Euclidean transformation and permutation of identical atoms of/in the LE.

In order to overcome the aforementioned difficulties when dealing with the small errors in the atomic positions we move away from the graph-based representation of the atoms. Instead we represent a LE centred on the (coloured) atom/point $p_0 = \left\{ C_0^P \in \mathbb{Z}, \boldsymbol{p}_0 \in \mathbb{R}^3 \right\}$ as the (coloured) point-set:

$$P = \{ p_0, p_1, \ldots, p_n \mid \|\boldsymbol{p}_0 - \boldsymbol{p}_i\| < r_{\text{env}} \} \tag{3.13}$$

where, without loss of generality, we set the centroid of P to the origin.
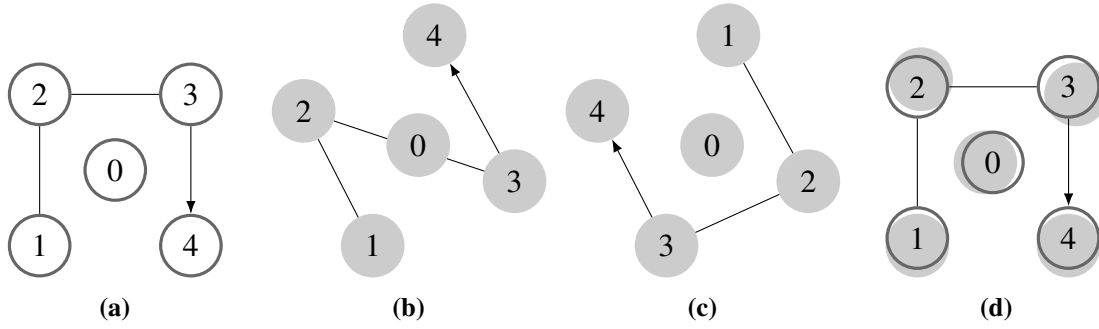
3.2.1 Tolerant equivalence

Algorithm 3.3 Checks to see if Eq. (3.14) holds for point sets P and Q.

Require: P and Q contain the same number, $N = 1 + n$, of points.
    function ARE_EQUIV(P, Q)
        $\boldsymbol{O} \leftarrow$ ROTOR_ONTO(P, Q)   ▹ Algorithm 3.5.
        $\Delta^2 \leftarrow \sum_{k=0}^{n} \|\boldsymbol{p}_k - \boldsymbol{O}\boldsymbol{q}_k\|^2$
        return $\Delta^2 \leq \delta^2$

Fig. 3.1 Figure 3.1a, a stored (reference) LE centred on atom labelled 0. Figure 3.1b, an unclassified LE, equivalent to the reference LE but with: small perturbations to all the atomic positions, a 30° rigid-body rotation and a permutation of the atomic labels. Figure 3.1c, the result of applying Algorithm 3.6 to the unclassified LE. Figure 3.1d, the result of applying Algorithm 3.5 to find the optimal rotation/reflection to map Fig. 3.1c onto the (overlaid) reference LE. We see all the atoms are close enough that the LE's can be considered equivalent.

Now the question of determining if two LE's, P and Q, (of the same size) are *equivalent* is the same as asking if there exists a matrix O and permutation π such that:

$$\sum_{i=0}^{n} \left\| p_i - O q_{\pi(i)} \right\|^2 \leq \delta^2 \tag{3.14}$$

and:

$$C_i^P = C_{\pi(i)}^Q \tag{3.15}$$

subject to the constraints:

$$O O^\mathsf{T} = O^\mathsf{T} O = I \quad \text{and} \quad \pi(0) = 0 \tag{3.16}$$

where δ is the maximum point-error/inter-point separation, as well as the maximum ℓ_2 norm or "distance" between the point-sets. Additionally, we define $\Delta_i = \left\| p_i - O q_{\pi(i)} \right\|$. The equivalence described is represented graphically in Fig. 3.1 and can be tested using Algorithm 3.3. Allowing O to be a general orthogonal (rather than pure rotation) matrix accounts for reflections. The permutation is required as there is no guarantee the neighbours of a point will be discovered in any particular order. Finding the optimal π, R is a strongly constrained variation of the well-studied point-set registration problem [97, 52].

The choice of δ controls how similar two LE's must be before they are considered equivalent; δ must be small enough to distinguish geometries that are local-minima of the atomic potential.

This requires at minimum $\delta < r_{min}$, with $r_{\min}$ the minimum interatomic separation. In the Fe-H system this implies we require $\delta \leq 1\text{Å}$ – corresponding to the smallest H-H separation.

An even stronger bound on δ can be established through the connection to U, the potential energy; Taylor-expanding about a converged extrema:

$$\Delta U \approx \Delta x^{\mathsf{T}} \cancelto{0}{\nabla U} + \frac{1}{2} \Delta x^{\mathsf{T}} H \Delta x \tag{3.17}$$

$$\approx \frac{1}{2} \Delta x^{\mathsf{T}} Q \Lambda Q^{\mathsf{T}} \Delta x \tag{3.18}$$

where in the second line, we have applied the eigendecomposition [10, p. 80] to the real symmetric Hessian, forming Λ the diagonal matrix of eigenvalues and Q the orthogonal matrix of eigenvectors. Noting an orthogonal transformation does not change the magnitude of a vector; A (weak) upper-bound on ΔU near a basin can be constructed from Eq. (3.18):

$$\Delta U \leq \frac{\lambda_{\max}}{2} \left\| Q^{\mathsf{T}} \Delta x \right\|^2$$

$$\leq \frac{\lambda_{\max}}{2} \left\| \Delta x \right\|^2$$

$$\leq \frac{\lambda_{\max}}{2} \delta^2 \tag{3.19}$$

where $\lambda_{\max}$ is the maximum eigenvalue of H and the third line follows from Eq. (3.14). If two LE's are equivalent (satisfying Eq. (3.14) and Eq. (3.15) under the constraints of Eq. (3.16)) and δ is small enough such that the mechanisms are transferable, then the energy barriers of the reconstructed mechanisms should be of-the-order-of ΔU off the true energy barriers. Ultimately, choosing a smaller value of δ increases the accuracy of the simulation, at the expense of increasing the number of SPS required. Hence, we should choose the largest δ ensuring ΔU is much less than the minimum relevant energy barrier.

Typically in the Fe-H system, with $r_{\mathrm{env}} = r_{\mathrm{cut}} = 6\text{Å}$, $N \approx 65$ and (using the perfect lattice as an order of magnitude estimation) $\lambda_{\max} \approx 10\mathrm{eV}\,\text{Å}^{-2}$. Therefore, we choose $\delta = 0.01\text{Å}$ resulting in an energy tolerance of approximately $\Delta U \leq 5 \times 10^{-4}\mathrm{eV}$. In practice we expect $\Delta U \ll 5 \times 10^{-4}\mathrm{eV}$ as Eq. (3.19) assumes Δx is parallel to the largest eigenvector of H which is unlikely. We see ΔU is much less than the energy barrier for H diffusion (0.0376eV), typically the fastest mechanism in the Fe-H system, justifying our choice of δ.

The choice of δ is continually validated during a simulation. If δ is too large then, following a mechanism reconstruction, a relaxation of the lattice will result in a large energy change.
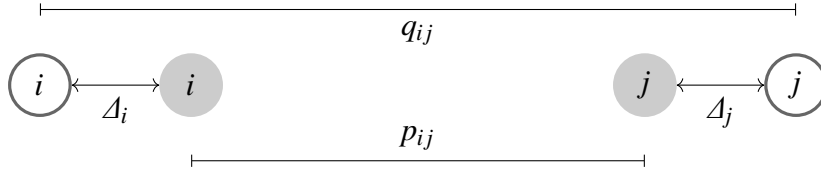
Fig. 3.2 Diagram showing the orientation of two pairs of points i and j in point-sets P (dark grey) and Q (white) that maximises $\Delta_i^2 + \Delta_j^2$ the sum of the square inter-point separations.

If/when this is detected δ can be adjusted. Conversely, if no such energy changes are detected δ can be increased to try and increase the performance of the simulation.

3.2.2 Point-set registration

Algorithm

To find the required permutation to determine equivalence, we recursively order Q; at each recursion searching for a point that ensures the intra-point separations to all previously ordered points in Q, are the same as (within tolerance of) those in P. Only requiring the intra-point separations, $r_{ij} = \lVert r_i - r_j \rVert$, in P and Q (which are invariant under Euclidean transformations of P and Q), this method can match the order of the points in LE's that are related by arbitrary rotations/reflections *before* solving for the rotation/reflection.

A constraint between the intra-point separations, p_{ij} and q_{ij}, can be obtained by studying Fig. 3.2:

$$\left| p_{ij} - q_{ij} \right| \leq \Delta_i + \Delta_j \tag{3.20}$$

maximising $\Delta_i + \Delta_j$ subject to the constraint from Eq. (3.14): $\Delta_i^2 + \Delta_j^2 \leq \delta^2$, we find:

$$\left| p_{ij} - q_{ij} \right| \leq \sqrt{2}\delta \tag{3.21}$$

which can be used to match pairs of points in Q to P and *vice-versa*. Each new point must satisfy Eq. (3.21) for all previously matched point, this can be tested for using Algorithm 3.4.

Once the order of the points in P and Q match we must solve:

$$\min_{O \in M_{3,3}(\mathbb{R})} \sum_{i=0}^{n} \lVert p_i - O q_i \rVert^2 \quad \text{s.t.} \quad OO^\mathsf{T} = O^\mathsf{T}O = I \tag{3.22}$$

Algorithm 3.4 Checks that Eq. (3.21) holds $\forall j < i$.

Require: `P[]` and `Q[]` are of length i or greater.
    function MATCH_UP_TO(`P[]`, `Q[]`, i)
        for $j \leftarrow 0, \ldots, (i-1)$ **do**
            if $|\text{NORM}(\text{P}[i], \text{P}[j]) - \text{NORM}(\text{Q}[i], \text{Q}[j])| > \sqrt{2}\delta$ **return** `false`
        return `true`

this is equivalent to the *orthogonal Procrustes problem* [79] which can be efficiently solved using the singular value decomposition (SVD). For completeness, the full method is included in Algorithm 3.5.

Algorithm 3.5 Solves the orthogonal Procrustes problem [79, 96]. Returns the orthogonal matrix which rotates/reflects Q onto P.

Require: P and Q contain the same number, $N = 1 + n$, of points.
    function ROTOR_ONTO(P, Q)
        $\boldsymbol{H} \leftarrow \sum_{k=0}^{n} \boldsymbol{q}_k \boldsymbol{p}_k^{\mathsf{T}}$
        Compute $\boldsymbol{U}, \boldsymbol{V}$ from the SVD of $\boldsymbol{H}$ such that $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathsf{T}}$
        return $\boldsymbol{V}\boldsymbol{U}^{\mathsf{T}}$

The full order-matching and equivalence testing method is detailed in Algorithm 3.6; once the algorithm has matched the orders and colours it checks if the permutation satisfies Eq. (3.14). This is required as there may be degenerate permutations satisfying Eq. (3.21) $\forall i, j$ but not satisfying Eq. (3.14). Additionally, we pre-order points, sorting them by their distance from the origin. This reduces the number of points that must be searched through to find the next matching point.

Complexity analysis

A key consideration for the usefulness of Algorithm 3.6 is its time complexity; for two randomly permuted point-sets the expected time complexity is:

$$\mathcal{O}(P_i) = \begin{cases} N + d_i \left(\mathcal{O}\left(M_i^{\mathsf{t}}\right) + \mathcal{O}(P_{i+1}) \right) + (N - d_i)\left(\mathcal{O}\left(M_i^{\mathsf{f}}\right) \right) & 1 \leq i < N \\ N & \text{otherwise} \end{cases} \tag{3.23}$$

where d_i (possibly a function of N) is the number of (degenerate) points for which MATCH_UP_TO returns `true`, each triggering d_i recursions; $\mathcal{O}\left(M_i^{\mathsf{t}}\right)$ is the time complexity of MATCH_UP_TO when it returns `true` and $\mathcal{O}\left(M_i^{\mathsf{f}}\right)$ the corresponding time complexity when it returns `false`.

Algorithm 3.6 Attempts to permute elements of Q such that Eq. (3.14) and Eq. (3.15) hold. Returns `true` if P and Q are equivalent.

Require: `P[]` and `Q[]` are of length $N = 1 + n$.
  function PERMUTE_ONTO($P[], Q[], i$)
    if $i \geq N$ **return** ARE_EQUIV(P, Q)   ▷ Algorithm 3.3, recursion terminates.
    for $j \leftarrow i, \ldots, n$ **do**   ▷ Try to find the ith matching point.
      if COL(P[i]) = COL(Q[j])
        SWAP(Q[i], Q[j])
        if MATCH_UP_TO(P, Q, i) **and** PERMUTE_ONTO(P, Q, $i + 1$) **return** `true`
        SWAP(Q[i], Q[j])
    return `false`

It is important to note in 3D the distance from four non-coplanar points fully specifies a unique point, hence:

$$\mathcal{O}\left(M_i^{\text{t}}\right) = N \quad \text{and} \quad \mathcal{O}\left(M_i^{\text{f}}\right) = 1 \tag{3.24}$$

By the same logic we see that d_i is independent of N and when $i \geq 4$ a constant, d; substituting into Eq. (3.23) and simplifying, the time complexity of Algorithm 3.6 is:

$$\mathcal{O}(P_i) = \begin{cases} N + d\mathcal{O}(P_{i+1}) & 1 \leq i < N \\ N & \text{otherwise} \end{cases} \tag{3.25}$$

a geometric progression summing to:

$$\mathcal{O}(P) = \begin{cases} \frac{N}{1-d} & d < 1 \\ N^2 & d = 1 \\ Nd^N & d > 1 \end{cases} \tag{3.26}$$

We must highlight the linear complexity regime, when $d < 1$, would imply that PERMUTE_ONTO must return `false` thus, in general with $d \leq 1$ we conclude $\mathcal{O}(P) = N^2$. Values of d greater than 1 rapidly trigger exponential complexity which makes the run-time of Algorithm 3.6 unfeasible.

Due to the tolerance in intra-point distances and assuming an approximately uniform point density ρ, we expect:

$$d = \frac{4}{3}\pi \left(\sqrt{2}\delta\right)^3 \rho \tag{3.27}$$

Therefore, in order to avoid exponential complexity, we require $d \leq 1$. For the case of α-Fe with lattice constant a and two atoms per unit cell, this requires:

$$\delta \leq a \left(\frac{3}{16\sqrt{2}\pi}\right)^{\frac{1}{3}} < 1\text{Å} \tag{3.28}$$

which is easily satisfied by our choice of δ. Smaller values of δ still reduce the constant and linear terms hidden in Eq. (3.26) which can be dominant for small N.

Heuristics

With the algorithms detailed thus far, a catalogue could be built that satisfied our requirements for LE classification however, although we ensure the condition of Eq. (3.28), a call to Algorithm 3.6 still takes $\approx 10\mu s$ with a LE containing 65 atoms. Hence, as we may call Algorithm 3.6 for every LE in the catalogue when encountering a new LE, this becomes prohibitively expensive. To reduce the search-space we partition the catalogue into sub-catalogues each indexed by a *discrete* key, k_d:

$$k_d \colon P \to \left\{ C_0^P, \{ m^\alpha \} \right\} \tag{3.29}$$

where m^α is the number of points in P of the αth colour. k_d can be used as the key to a hash-table (or other suitable key–value store) enabling $\mathcal{O}(1)$ look-up of the sub-catalogues.

The sub-catalogues may still become very large, especially in systems where all points are the same colour. As a simulation progresses, we can sort the order of the LE's in each sub-catalogue by its occurrence count. This significantly decreases the look-up time for a typical LE as many systems have most points in the same LE and only a small number of "active" points (e.g near defects) in rare LE's.

To further accelerate searches of the sub-catalogues we introduces a second *fuzzy* key, k_f, the collection of ordered-sets/lists:

$$k_f \colon P \to \left\{ \{ p_{0i} \mid i > 0 \}_{\leq}^{\alpha}, \{ p_{ij} \mid i > 0, j > i \}_{\leq}^{\alpha, \beta \leq \alpha} \right\} \tag{3.30}$$

where p_{ij} denotes the intra-point distances between points i and j and the superscripts α, β indicate the colour of the points in the point pair. For example, in the α-Fe system there are two possible point colours hence, k_f contains five ordered lists. Two each containing the intra-point distances between points of a particular colour and the central point; a further three lists containing the intra-point distances between pairs of atoms coloured H-H, Fe-H/H-Fe, and Fe-Fe.

By construction, k_f is invariant under Euclidean transformations and permutations of the points in P. Two fuzzy keys can be compared for equivalence using Algorithm 3.7. It is important to note Algorithm 3.7 is a non-transitive binary relation.

Algorithm 3.7 Compare two fuzzy keys for equivalence.

Require: P and Q have matching discrete keys.
> **function** FUZZY_EQUIV(k_f^P, k_f^Q)
>> **for** each pair of ordered lists `p[]`, `q[]` in k_f^P, k_f^Q **do**
>>> **for** each pair of elements p, q in `p[]`, `q[]` **do**
>>>> **if** $|p - q| > \sqrt{2}\delta$ **return** `false`
>>
>> **return** `true`

If we store alongside each LE in the catalogue its fuzzy key, before comparing LE's with Algorithm 3.6 we require their fuzzy keys to be equivalent. Specifically, it is necessary but not sufficient for their fuzzy keys to be equivalent for Algorithm 3.6 to return `true`. In practice, equivalence of fuzzy keys is a very strong pre-conditioner for Algorithm 3.6. As comparison of fuzzy keys is orders of magnitude faster (typically taking $\approx$ 100ns) this substantially accelerates searching the sub-catalogues.

Summary

The full method for classifying a LE, represented by the point-set Q, and reconstructing the mechanisms discovered by previous SPS at an equivalent LE proceeds as follows:

1. Compute the discrete, k_d, and fuzzy, k_f, keys of Q.

2. Find the sub-catalogue corresponding to k_d.

3. Search the sub-catalogue; for each LE denoted P:

   (a) If the fuzzy keys of P and Q are equivalent.

   (b) If (using Algorithm 3.6) Q can be ordered and rotated/reflected to match P.

(c) End the search and return P, the equivalent LE.

4. If a match, P, was found, the mechanisms associated with P, can be reconstructed onto Q by multiplying their atomic displacement-vectors by $\boldsymbol{O}^{\mathsf{T}}$.

5. Otherwise, Q represents a new LE; append Q to the sub-catalogue and launch SPS centred on the LE in order to discover any mechanisms associated with it.

Our invariant and tolerant LE equivalence (ITLEE) classification scheme satisfies the requirements laid out in Section 3.2. Equivalence is fully invariant under: Euclidean transformations of the LE; permutations of identical atoms and most crucially small perturbations of the atomic positions. Our classification scheme uses all the information encoded within the set of $N(N-1)$ inter-atomic separations in a LE. We have provided heuristics to enable efficient implementations of our scheme. Furthermore, we have justified our choice of $\delta = 0.01\text{Å}$, connecting δ to the error in the reconstructed energy barriers.

3.3 Confidence parameter for catalogue completeness

We wish to develop a confidence parameter that gives us a measure of the certainty that we have found all relevant saddle-points surrounding some basin. The confidence parameter should give us some idea of when we should stop searching for new saddle-points. Importantly, we want to allow for the possibility that the probability of discovering a SP is non-uniform (it has been demonstrated that high-energy saddle-points are less likely to be discovered).

Suppose we have been searching for saddle-points for some time and have discovered a collection of *known* saddles. We choose to investigate:

$$P\,(\text{found all saddles} \mid m \text{ consequtive failed searchs (CFS) })\qquad(3.31)$$

where, in this context, a 'failed' search is one that fails to discover an *unknown* SP (e.g. finds a known SP). We conceptualise this like a ball-and-urn problem. There exists two possibilities either: we have been found all the saddles or there exists some number of unknown saddles that occupy a fraction, ϵ, of the remaining probability space. In the former case the probability of m CFS is trivially unity, while in the latter:

$$P\,(m \text{ CFS } \mid \epsilon) = (1 - \epsilon)^m\qquad(3.32)$$

Applying the Bayes–Price [5] theorem to Eq. (3.31) and substituting in the above:

$$P(\text{found all} \mid m \text{ CFS}) = \frac{P(\text{found all})}{P(\text{found all}) + P(\text{some remain}) \cdot (1 - \epsilon)^m} \quad (3.33)$$

Hence, accepting the zero-information prior $P(\text{found all}) = P(\text{some remain})$, becomes:

$$P(\text{found all} \mid m \text{ CFS}) = \frac{1}{1 + (1 - \epsilon)^m} \quad (3.34)$$

We note, it is possible to estimate the relative probability of finding a known SP by keeping a tally, f_i, of the number of times SP i has been discovered by a SP search. We make the assumption that the probability distribution of finding a SP is sufficiently-smooth such that the maximum probability of finding an unknown SP is close to the probability of the rarest known SP. Hence, a lower-bound on Eq. (3.34) can be constructed assuming there is only a single unknown saddle remaining:

$$P(\text{found all saddles} \mid m \text{ CFS}) \gtrless \frac{1}{1 + \left(\dfrac{1}{1 + \frac{f_{\min}}{\sum f_i}} \right)^m} \quad (3.35)$$

Once the above surpasses a confidence limit, e.g. 95%, we end the search for new saddle-points.

Chapter 4

Results and discussion

4.1 Vacancy cluster diffusion

In previous work [90] we used OLKMC with topological LE identification to simulate vacancy cluster diffusivities. Hence, to verify our ITLEE classification we initially repeat then extend these measurements to larger and more complex clusters. Previously we used a K-Means clustering algorithm [67] to explicitly track the position of vacancies during the simulation. This is difficult to extend to larger clusters hence, we choose to use the mean-squared displacement (MSD):

$$\text{MSD} = \frac{1}{N} \sum_{\alpha}^{N} \left\| \boldsymbol{r}_{t=0}^{\alpha} - \boldsymbol{r}_{t}^{\alpha} \right\|^2 \tag{4.1}$$

which is directly related to the simulated and effective diffusivities, D_{sim} and D_{eff} respectively [74]:

$$D_{\text{sim}} = \frac{\text{MSD}}{6t} \quad \text{and} \quad D_{\text{eff}} = \frac{D_{\text{sim}}}{x_d} \tag{4.2}$$

with x_d the defect concentration. This has the additional benefit of averaging over many more atoms.

For our first investigation we insert vacancy-clusters of various sizes into a 6^2 unit-cell supercell and track the MSD of the Fe atoms over time. The results are presented in Fig. 4.1 and summarised in Table 4.1. Most immediately in Fig. 4.1 we see a clear convergence to Eq. (4.2) for all clusters, once long-enough timescales have been reached. This timescale is important as it is a property of the system (not simulation method). Off-lattice KMC is capable
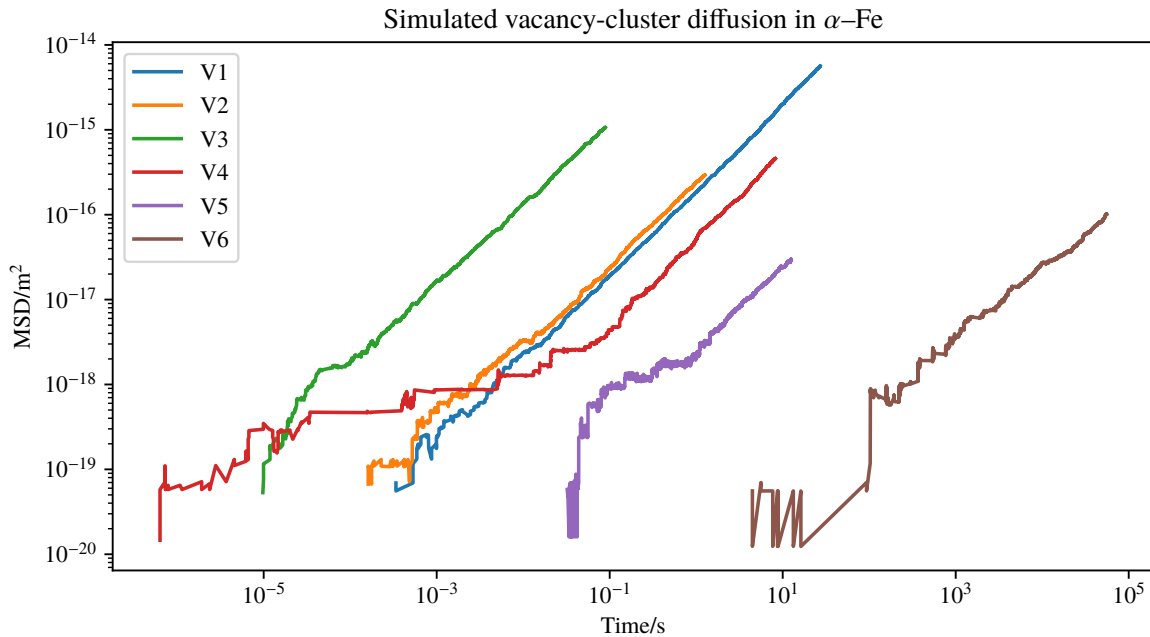
Fig. 4.1 Vacancy clusters diffusing in a perfect 6^3 unit-cell supercell at 300K. This required approximately 100 hours of CPU-time on a quad-core Intel® Xeon® at 3.10GHz.

of reaching diffusive timescales for all the cluster investigated, with convergence requiring between 5×10^{-4}s and 5×10^5s. This is made possible by a combination of the ITLEE caching and the periodic boundary conditions; after a series of mechanisms the clusters are moved to a topologically equivalent state just rotated/translate. Therefore, SP searches were only required during the initial *learning* phase of the simulation. This highlights the need for saddle-point recycling in any OLKMC simulation.

The expected behaviour for cluster diffusivity is larger clusters becoming less mobile. This trend is visible in Table 4.1 but, the the diffusivities for the V2 and V3 clusters seem to reverse the trend. We shall now discuss each cluster's diffusion mechanisms in order to explain the results.

V1 cluster

The single vacancy diffuses by $\frac{1}{2}\langle 111 \rangle$ vacancy hops with an activation energy of 0.65eV and kinetic pre-factor 8.38×10^{13}Hz. This mechanism is sketched in Fig. 4.2a and Fig. 4.2b. The energy barrier and mechanism are in good agreement with the literature [74, 57].
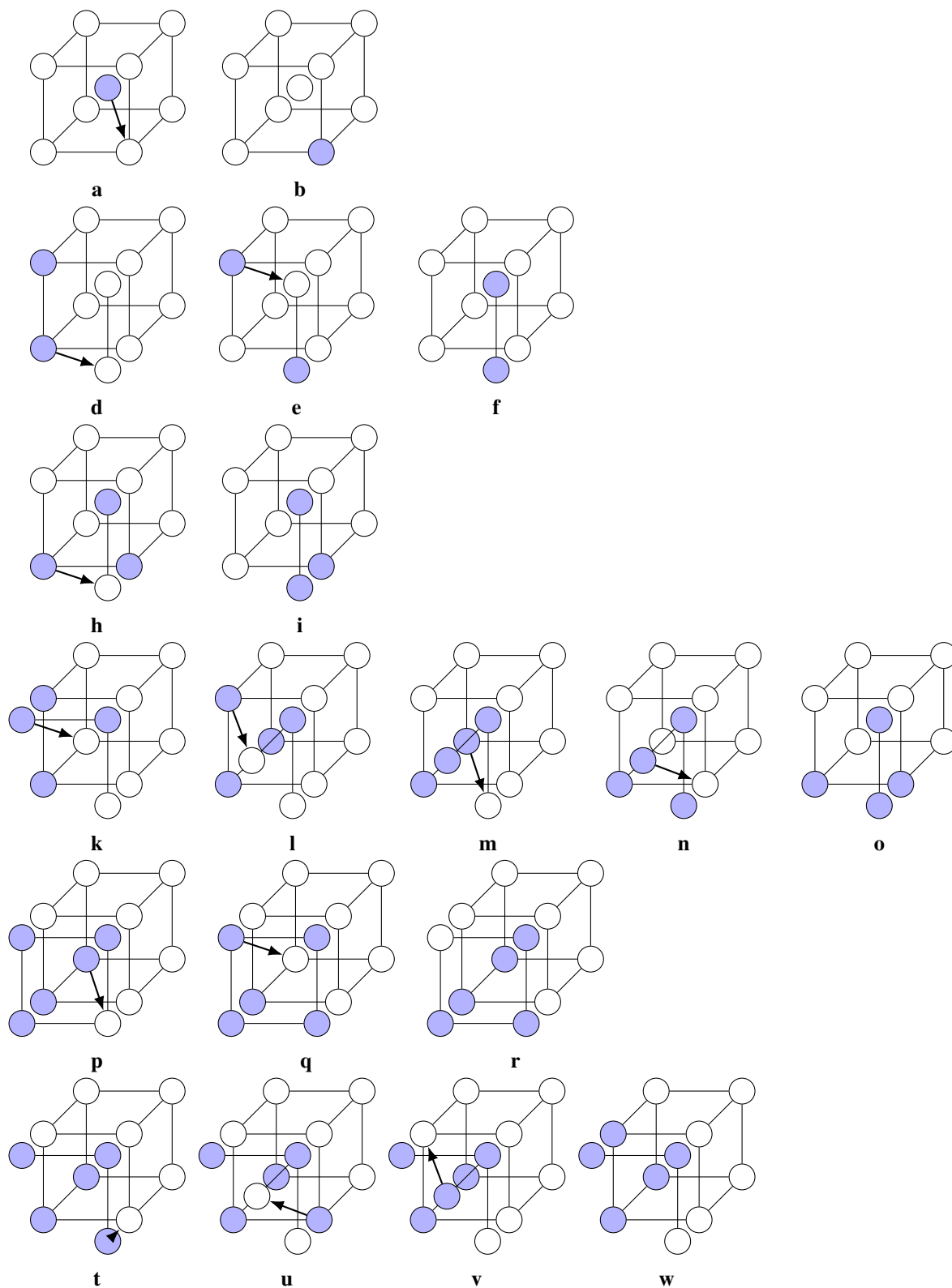
Fig. 4.2 Mechanisms identified for vacancy-cluster diffusion in the α–Fe lattice. Blue circles mark a vacancy; white circles represent an occupied lattice site and arrows mark the path of the vacancy during a mechanism. Perturbations to the lattice have been omitted for clarity. See the text for a full description.

Table 4.1 Summary of results for various vacancy clusters diffusing in the α–Fe lattice. All diffusivities have a fractional error less than 1 in 100. The quoted energy-barriers and corresponding harmonic pre-factor are for the largest barrier during the sequence of mechanisms.

Cluster	ΔE^{max}/eV	$\nu/10^{13}$Hz	Simulated time/s	D_{eff}/m^2 s^{-1}
V1	0.65	8.38	27.2	3.42×10^{-17}
V2	0.65	10.4	1.26	3.99×10^{-17}
V3	0.48	5.22	0.0891	2.04×10^{-15}
V4	0.68	3.41	8.23	9.04×10^{-18}
V5	0.70	6.82	12.5	4.09×10^{-19}
V6	0.73	6.87	55600	2.98×10^{-22}

V2 cluster

The lowest-energy configuration for V2 is the second nearest neighbour (2nn) pair followed by the 1nn then 4nn orientations [20]. Two mechanisms are identified for V2 diffusion. The predominant mechanism observed was oscillations between the 2nn and 4nn states. This is sketched in Fig. 4.2d to Fig. 4.2f, the corresponding energy barriers and kinetic pre-factors are 0.65eV, 8.38×10^{13}Hz and 0.44eV, 7.52×10^{13}Hz. The 2nn pathway may be expected to be the dominant mechanisms, as one may predict the transition to the lower energy 1nn state to have a lower energy-barrier. However, the 2nn to 1nn transition has an energy barrier of 0.72eV and kinetic pre-factor 2.97×10^{13}Hz. This increased energy-barrier makes it kinetically less-favourable.

The V2 diffusion barrier is very close to the diffusion barrier for the single vacancy, this may be an artefact of the EAM potential as *ab initio* studies typically predict an energy barrier 0.05–0.11eV [19, 28] lower. Nevertheless, this is in agreement with other works that use semi-empirical potentials [57].

The combination of a near-identical energy barrier, slightly higher kinetic pre-factor and additional (albeit slower) 2nn pathway result in the V2 cluster having a marginally higher diffusivity compared to the single vacancy.

Off-lattice KMC has identified the counter-intuitive predominant diffusion mechanisms of the V2 cluster. Although, this could have been captured with traditional KMC and careful DFT analysis, here it arises naturally without any special modifications.

V3 cluster

As previously hinted, the V3 cluster defies the expectation and is the most mobile cluster with a diffusivity almost a thousand times higher than a single vacancy at 300K. This is due to the minimum energy configuration (MEC) – Fig. 4.2h – permitting a mechanism/vacancy-hop with an energy barrier of 0.48eV and kinetic pre-factor 5.22×10^{13}Hz that immediately reforms the MEC but displaced/rotated. Similarly to the single vacancy, V3 can diffuse without changing its shape. The mechanism is detailed in Fig. 4.2h and Fig. 4.2i. The simulated MEC matches theoretical predictions [28], as does the mechanism [28].

V4 cluster

At the V4 cluster the mobility of the clusters begins to decrease. This is due to the high energy barrier, 0.68eV, required to break apart the MEC, sketched in Fig. 4.2k. Once the MEC has disassociated a series of diffusion mechanisms are possible. One such mechanism is detailed in Fig. 4.2k to Fig. 4.2o. The energy barriers during these mechanisms are: 0.68eV, 0.51eV, 0.51eV, and 0.47eV; while the corresponding kinetic pre-factors are: 3.41×10^{13}Hz, 5.38×10^{13}Hz, 5.81×10^{13}Hz and 3.46×10^{13}Hz; respectively. Hence, the decreased diffusivity, compared to the single vacancy, is predominantly due to the increased energy-barriers and additional steps required to diffuse.

V5 and V6 clusters

Similar to the V4 clusters the V5 and V6 clusters continue to become less mobile as their size increases. For the V5 cluster, two mechanisms are detailed in Fig. 4.2p to Fig. 4.2k and Fig. 4.2t to Fig. 4.2w. For both mechanisms the limiting barrier to diffusion is the dissociation of the MEC with a barrier of 0.70eV and a harmonic prefactor of 9.86×10^{13}Hz.

The V6 cluster was particularly difficult to analyse due to the dynamic superbasin acceleration that was used. During the simulation the tolerance for a mechanism to be considered transient rose to approximately 0.6eV. This meant that many vacancy hops could occur between the time-steps outputted during the simulation. Although the kinetics remain accurate, the state-to-state dynamics inside the superbasin(s) are lost. This enabled the simulation to reach the very long timescales required to observe the diffusive behaviour of the V6 cluster which would not have otherwise been achievable.

Table 4.2 Summary of cluster diffusivities with the addition of a single hydrogen atom during the simulations in Fig. 4.3, Fig. 4.4, Fig. 4.5 and Fig. 4.6.

Cluster	Diffusivity, $D/\mathrm{m^2\,s^{-1}}$		
	Without H	With H	D_{eff} of H
V1	3.42×10^{-17}	5.16×10^{-19}	1.11×10^{-16}
V2	3.99×10^{-17}	3.27×10^{-18}	2.04×10^{-16}
V3	2.04×10^{-15}	1.54×10^{-15}	-
V4	9.04×10^{-18}	7.03×10^{-18}	3.94×10^{-16}

4.1.1   Summary

Off-lattice KMC has successfully been applied to study the diffusion of vacancy clusters. The mechanisms predicted and diffusivity-trends match those seen in the literature [20, 28, 57] however, they have all been predicted *a priori* by the highly general OLKMC framework. The requirement of solving the flickering problem has been emphasized by our study of the larger cluster – particularly V5 and V6. The long timescales reached are achieved through our SP recycling method based on our ITLEE classification, this can typically completely elide SPS during the later stages of a cluster-diffusion simulations.

4.2   Cluster-hydrogen complex diffusion

To build upon the cluster diffusivity results and highlight OLKMC's capability of resolving events occurring at vastly different timescales we add a single hydrogen atom into several of the clusters forming Vn-H complexes. The resulting diffusivities are summarised in Table 4.2. We shall now discuss each complex in detail.

V1-H complex

Figure 4.3 details the MSD of the Fe and H during the complex diffusion alongside the H-free data from Fig. 4.1 for reference. Three regimes are visible in Fig. 4.3. Initially, below 10^{-9}s, the H atom explores the 6 bound states inside the vacancy, the energy-barrier for these transitions is 0.061eV with a kinetic pre-factor of 3.12×10^{12}Hz. Once the 24 internal mechanisms have been explored the superbasin analytical acceleration kicks in and mechanisms that disassociate the complex begin to occur. This is visible as a large time-discontinuity in Fig. 4.3.
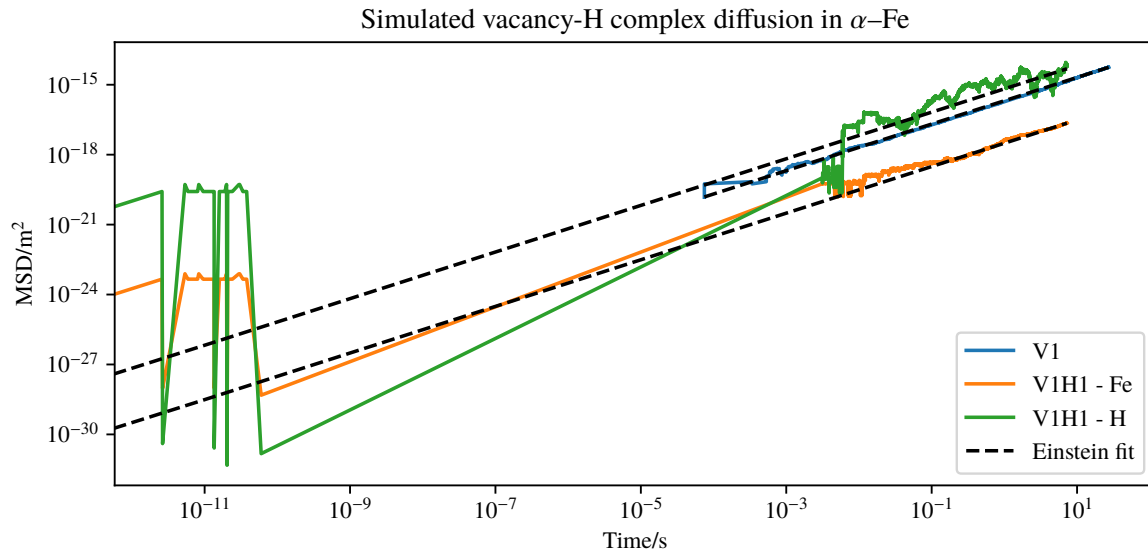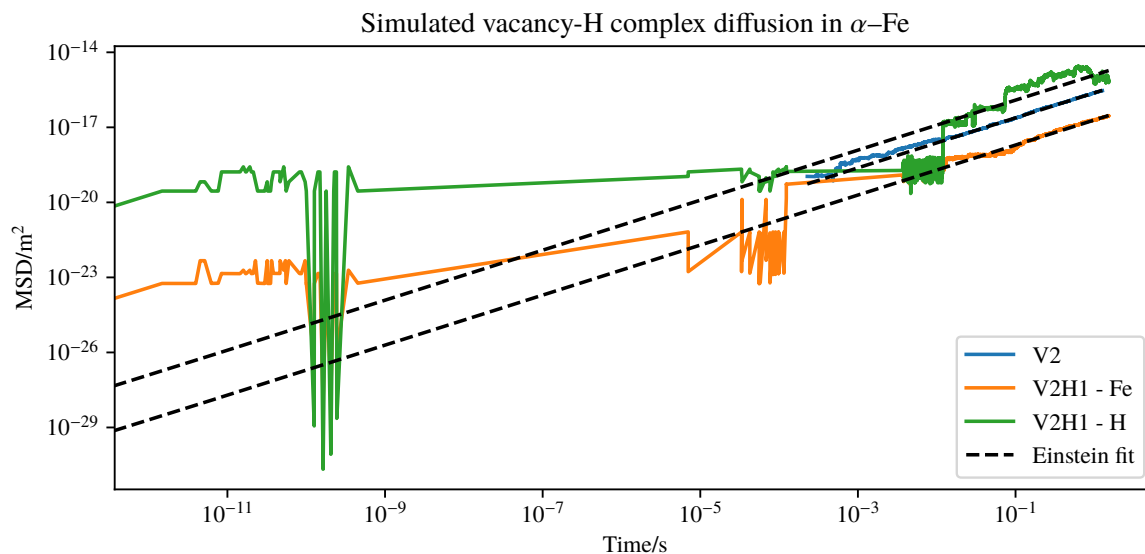
Fig. 4.3 V1-H complex diffusing in a perfect 6^3 unit-cell supercell at 300K. Dashed lines are fits to Eq. (4.2)

The second regime occurs around 5×10^{-3} s, here the MSD of the H and Fe are approximately the same. This occurs as the cluster+H co-diffuse via two mechanisms. The first is the $\frac{1}{2}\langle 111 \rangle$ vacancy hop into the complex with an activation energy of 0.84eV and kinetic pre-factor 2.33×10^{14} Hz, followed by recapture of the H atom. The barrier is elevated compared to the single vacancy due to the steric hindrance provided by the H atom but, this is partially compensated for by the increase in the kinetic prefactor. The second mechanisms is limited by the 0.52eV barrier for H de-trapping with a kinetic pre-factor of 2.38×10^{12} Hz. The escaped H-atom then "pushes" a Fe atom into the vacancy and becomes trapped in the newly formed vacancy.

The third regime, above 6×10^{-3} s, occurs once the H-atom begins to completely de-trap from the vacancy, diffuse through the lattice and re-trapping at a vacancy then repeating the cycle. This is visible as a large jump in the MSD of the H-atom in Fig. 4.3. This occurs through tetrahedral-tetrahedral interstitial hops with an energy barrier of 0.048eV and kinetic pre-factor 5.52×10^{12} Hz.

This culminates in an overall reduction in diffusivity as the co-diffusion mechanism competes with the complete de-trapping and the $\frac{1}{2}\langle 111 \rangle$ hop's barrier is elevated. The two order-of-magnitude difference in the harmonic pre-factors is evidence that, the constant pre-factor approximation should not be made for multi-element systems.
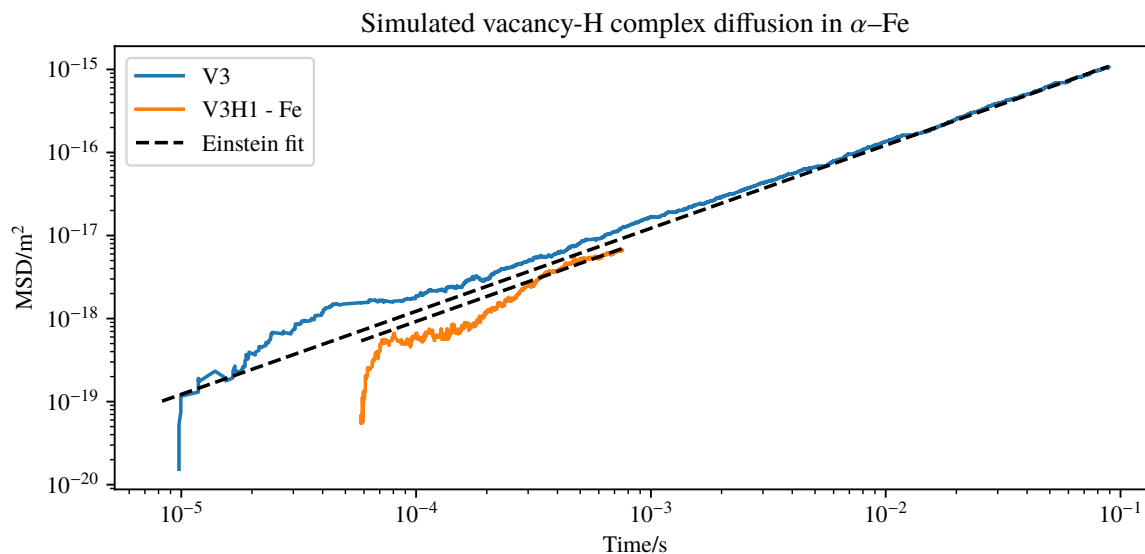
Fig. 4.4 V2-H complex diffusing in a perfect 6^3 unit-cell supercell at 300K. Dashed lines are fits to Eq. (4.2)

Unlike the MSD of the Fe atoms the MSD of the H atoms does not converge to a smooth line in Fig. 4.3. This is because there is only a single H-atom so the fractional error in the MSD remains constant.

V2-H complex

The V2-H complex follows a similar three regime behaviour as the V1-H complex however, the V2 cluster contains 14 trap sites connected by 56 internal mechanisms. The energy barrier for direct formation of the 4nn state remains close to the H free case however, this mechanism is severely suppressed as the H atom remains in the unmoved vacancy. This results in the 4nn configuration being overwhelmingly likely to collapse into the original 2nn configuration due to the steric hindrance from the H atom, hence no diffusion occurs. Additionally, the H-atom increase the barrier of directly forming the 1nn configuration, suppressing diffusion via 1nn–2nn oscillations.

Alternatively, the V2-H complex diffuses similarly to the V1-H complex; the H atom first escapes the cluster then "pushes" an atom into the cluster. This can form either the 4nn or 1nn configuration with the H atom in the new vacancy, which then favours collapse into the translated 2nn state. The total energy barrier for the 4nn process is 0.66eV with a kinetic prefactor of 1.43×10^{12}Hz. Again, we see the fall in kinetic pre-factor makes a major contribution to the

Fig. 4.5 V3-H complex diffusing in a perfect 6^3 unit-cell supercell at 300K. Dashed lines are fits to Eq. (4.2). The simulation was cut short due to an implementation error.

fall in diffusivity, compared to the 2V cluster. The 1nn pathway has a total energy barrier of 0.75eV meaning it contributes very little to the complex's diffusion.

V3-H complex

The results gathered for the V3-H complex diffusion are presented in Fig. 4.5, unfortunately, due to an implementation error, the simulation was cut-short before full convergence was achieved. Additionally, as the de-trapping barrier remains high, de-trapping occurs over timescales around 10^{-2}s hence, the H-diffusion is omitted from Fig. 4.5. Due to the increased size of the V3 cluster, the addition of an H-atom into one of its 18 trapping sites does not significantly alter the energy barrier of the primary mechanism – sketched in Fig. 4.2h and Fig. 4.2i. The barrier of the conservative mechanism becomes 0.49eV and its kinetic pre-factor is 2.72×10^{14}Hz, However these are only available in when the H-atom at an extreme end of the MEC. This culminates in the diffusivity of the V3-H complex being fractionally lower than the V3 cluster.

V4-H complex

In Fig. 4.6, which shows the results for the V4-H complex diffusion, we see the same three regimes that were visible during the during the V1/V2 complexes. Similar to the V3-H complex, the increased size of the V4 cluster allows the same diffusion mechanisms to occur as for the
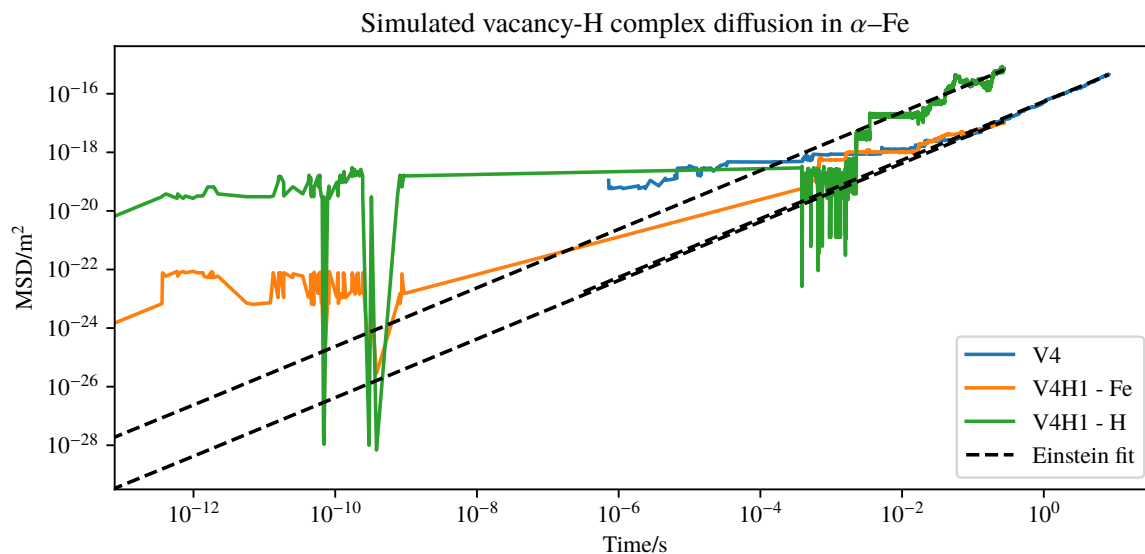
Fig. 4.6 V4-H complex diffusing in a perfect 6^3 unit-cell supercell at 300K. Dashed lines are fits to Eq. (4.2)

V4 cluster with very similar energy-barriers/pre-factors. This again results in a marginally decreased diffusivity when compared to the hydrogen free case.

4.2.1 Summary

The diffusivity of the hydrogen in the presence of vacancy clusters is presented in Table 4.2. The effective diffusivity of hydrogen increase for larger clusters. This could be due to falling de-trapping barriers, rising trapping barriers, increased cluster surface-area or a combination of these and possibly unknown effects.

The presence of hydrogen on vacancy clusters in α–Fe has been shown to reduce their diffusivity. This is in agreement with the literature however, some studies suggest adding more hydrogen atoms can reverse this trend [74]. The magnitude of the reduction in diffusivity reduces as clusters become larger. This could be because only a single H-atom was used hence, for larger clusters, mechanisms could occur without the H-atom directly interacting with them. Future work should investigate the effect of increasing the number of H-atoms in the supercell.

The multi-scale nature of vacancy-hydrogen complex diffusion is prominent in many of the figures in this section; OLKMC's ability to span many time-scales has again proven crucial in understanding these phenomenon. As more complex structures were simulated the diversity of mechanisms increased rapidly, modelling this with traditional KMC would quickly become intractable.

Chapter 5

Conclusion and future work

5.1 Conclusions

We have implemented a bleeding-edge off-lattice kinetic Monte Carlo simulator, incorporating optimisations and enhancements from across the literature as well as our invariant and tolerant local-environment equivalence classification scheme. We have investigated the diffusion of vacancy clusters, with and without the addition of hydrogen, in the α-Fe lattice. Our key contributions and results include:

- Development of the ITLEE method for saddle-point recycling, providing a rigorous connection between equivalence and energy-differences between environments. Furthermore, ITLEE is tolerant of small perturbations minimising the number of SP searches required.

- Applying OLKMC to autonomously identify the primary diffusion mechanisms of vacancy clusters up to the V5 cluster.

- Simulating the diffusivities of the first six vacancy clusters and accurately reproducing the mobility increase at V3 before returning to the expectation (reduced mobility for larger clusters).

- Simulating the effect of a single H atom on the diffusivity of vacancy clusters and probing the H-induced mechanisms causing this reduction – different for each complex. In all investigated cases H reduced the cluster diffusivity.

- During our OLKMC simulations we compute the harmonic pre-factor as-well-as energy barriers. We demonstrate pre-factors can vary by two orders-of-magnitude. Hence, we suggest all OLKMC simulations (particularly in multi-element systems) verify the constant pre-factor approximation carefully.

The Fe-H system, with its small interstitials, multi-stage mechanisms, frequent flickering-problems, varied harmonic pre-factors and sensitive energy-barriers present many challenges to model with OLKMC. Nevertheless, many of these have been overcome and OLKMC has proved an invaluable and capable tool for the study of these systems. This is extremely promising for the future of modelling defect-H interactions and improving our understanding of HE at the atomic scale.

5.2 Future work

Several optimisations to the OLKMC method/implementation are yet to be explored:

- We already create a catalogue of local environments hence, it would be possible to refine the energy-barriers of the most frequent/critical mechanisms using *ab initio* force-fields.

- Migrating to a distributed-memory programming model would allow us to take advantage of more parallelism during SP searches.

- Constructing a lattice approximation on-the-fly to elide global minimisations. This could allow OLKMC to become competitive with hand-tuned KMC when it is possible to learn all mechanisms quickly. (This is currently under active development with a collaborator)

- Augmenting the SP search procedure by taking into account knowledge of previously discovered saddle-points in each LE would reduce/eliminate redundant SP searches. This could be achieved through repulsive additions to the PES at known saddle-points or moving toward a particle swarm optimiser [40] with a repulsive component between dimers.

The last of these is a particularly-promising avenue for the acceleration of low symmetry supercell which are SPS intensive. Therefore, this will be considered a prime focus for future work.

The limiting factor underlying all OLKMC/MD simulations is the semi-empirical potentials. Development of new potentials, focusing on H-defect interactions, using machine-learning methods could also be done. This could improve both the speed and accuracy of OLKMC simulations.

Off-lattice KMC is just as capable of modelling any system for which suitable potentials exist, of particular interest are the Fe-C and Fe-H-C systems. The former could offer easier experimental verification while the latter may be equally important in-order to understand HE.

Finally, understanding HE will require modelling larger, more complex defects such as grain-boundaries and dislocations. Initially, studying their interactions with single H-atoms is the simplest path to build up a full understanding and should be on the horizon of future work. For dislocations, this will first require validating the mechanisms of H diffusion around an immobile dislocation, to later include strain-induced glide to determine how H affects dislocation mobility and *vice-versa*.

References

[1]   J. H. Ahlberg. *The Theory of Splines and Their Applications : Mathematics in Science and Engineering: A Series of Monographs and Textbooks, Vol. 38*. Saint Louis: Elsevier Science, 2016. ISBN: 9781483222950.

[2]   Kathleen C Alexander and Christopher A Schuh. "Towards the reliable calculation of residence time for off-lattice kinetic Monte Carlo simulations". In: *Modelling and Simulation in Materials Science and Engineering* 24.6 (Aug. 2016), p. 065014. DOI: 10.1088/0965-0393/24/6/065014.

[3]   Sang Bak, Muhammad Abro, and Dong Lee. "Effect of Hydrogen and Strain-Induced Martensite on Mechanical Properties of AISI 304 Stainless Steel". In: *Metals* 6.7 (July 2016), p. 169. DOI: 10.3390/met6070169.

[4]   O. Barrera et al. "Understanding and mitigating hydrogen embrittlement of steels: a review of experimental, modelling and design progress from atomistic to continuum". In: *Journal of Materials Science* 53.9 (Feb. 2018), pp. 6251–6290. DOI: 10.1007/s10853-017-1978-5.

[5]   Bayes and Price. "LII. An essay towards solving a problem in the doctrine of chances." In: *Philosophical Transactions of the Royal Society of London* 53 (Dec. 1763), pp. 370–418. DOI: 10.1098/rstl.1763.0053.

[6]   Laurent Karim Béland et al. "Kinetic activation-relaxation technique". In: *Physical Review E* 84.4 (Oct. 2011). DOI: 10.1103/physreve.84.046704.

[7]   Laurent Karim Béland et al. "Kinetic Activation–Relaxation Technique and Self-Evolving Atomistic Kinetic Monte Carlo: Comparison of on-the-fly Kinetic Monte Carlo algorithms". In: *Computational Materials Science* 100 (Apr. 2015), pp. 124–134. DOI: 10.1016/j.commatsci.2014.12.001.

[8]   Vijesh J. Bhute and Abhijit Chatterjee. "Building a kinetic Monte Carlo model with a chosen accuracy". In: *The Journal of Chemical Physics* 138.24 (June 2013), p. 244112. DOI: 10.1063/1.4812319.

[9]   H.K. Birnbaum and P. Sofronis. "Hydrogen-enhanced localized plasticity—a mechanism for hydrogen-related fracture". In: *Materials Science and Engineering: A* 176.1-2 (Mar. 1994), pp. 191–202. DOI: 10.1016/0921-5093(94)90975-x.

[10]   Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN: 0387310738.

[11]   D. Bombač et al. "Theoretical evaluation of the role of crystal defects on local equilibrium and effective diffusivity of hydrogen in iron". In: *Materials Science and Technology* 33.13 (Apr. 2017), pp. 1505–1514. DOI: 10.1080/02670836.2017.1310417.

[12] A.B. Bortz, M.H. Kalos, and J.L. Lebowitz. "A new algorithm for Monte Carlo simulation of Ising spin systems". In: *Journal of Computational Physics* 17.1 (Jan. 1975), pp. 10–18. DOI: 10.1016/0021-9991(75)90060-1.

[13] Charles J. Cerjan and William H. Miller. "On finding transition states". In: *The Journal of Chemical Physics* 75.6 (Sept. 1981), pp. 2800–2806. DOI: 10.1063/1.442352.

[14] Abhijit Chatterjee. "Uncertainty quantification of Kinetic Monte Carlo models constructed on-the-fly using molecular dynamics". In: *MRS Communications* 8.03 (May 2018), pp. 850–857. DOI: 10.1557/mrc.2018.93.

[15] Samuel T. Chill and Graeme Henkelman. "Molecular dynamics saddle search adaptive kinetic Monte Carlo". In: *The Journal of Chemical Physics* 140.21 (June 2014), p. 214110. DOI: 10.1063/1.4880721.

[16] Mykyta V. Chubynsky et al. "Exploiting memory in event-based simulations". In: *Journal of Non-Crystalline Solids* 352.42-49 (Nov. 2006), pp. 4424–4429. DOI: 10.1016/j.jnoncrysol.2006.01.117.

[17] W.A. Counts, C. Wolverton, and R. Gibala. "First-principles energetics of hydrogen traps in α-Fe: Point defects". In: *Acta Materialia* 58.14 (Aug. 2010), pp. 4730–4741. DOI: 10.1016/j.actamat.2010.05.010.

[18] Felicity F. Dear and Guy C. G. Skinner. "Mechanisms of hydrogen embrittlement in steels: discussion". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 375.2098 (June 2017), p. 20170032. DOI: 10.1098/rsta.2017.0032.

[19] F.G. Djurabekova et al. "Stability and mobility of small vacancy and copper-vacancy clusters in bcc-Fe: An atomistic kinetic Monte Carlo study". In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 255.1 (Feb. 2007), pp. 47–51. DOI: 10.1016/j.nimb.2006.11.009.

[20] F. Djurabekova et al. "Kinetics versus thermodynamics in materials modeling: The case of the di-vacancy in iron". In: *Philosophical Magazine* 90.19 (July 2010), pp. 2585–2595. DOI: 10.1080/14786431003662515.

[21] Yi Dou et al. "Opportunities and Future Challenges in Hydrogen Economy for Sustainable Development". In: *Hydrogen Economy*. Elsevier, 2017, pp. 277–305. DOI: 10.1016/b978-0-12-811132-1.00010-9.

[22] A. Drexler et al. "Hydrogen segregation near a crack tip in nickel". In: *Scripta Materialia* 194 (Mar. 2021), p. 113697. DOI: 10.1016/j.scriptamat.2020.113697.

[23] Yaojun A. Du, Jutta Rogal, and Ralf Drautz. "Diffusion of hydrogen within idealized grains of bcc Fe: A kinetic Monte Carlo study". In: *Physical Review B* 86.17 (Nov. 2012). DOI: 10.1103/physrevb.86.174110.

[24] P.J. Ferreira, I.M. Robertson, and H.K. Birnbaum. "Hydrogen effects on the interaction between dislocations". In: *Acta Materialia* 46.5 (Mar. 1998), pp. 1749–1757. DOI: 10.1016/s1359-6454(97)00349-2.

[25] Kristen A. Fichthorn and Yangzheng Lin. "A local superbasin kinetic Monte Carlo method". In: *The Journal of Chemical Physics* 138.16 (Apr. 2013), p. 164104. DOI: 10.1063/1.4801869.

[26] M. W. Finnis and J. E. Sinclair. "A simple empirical N-body potential for transition metals". In: *Philosophical Magazine A* 50.1 (July 1984), pp. 45–55. DOI: 10.1080/01418618408244210.

[27] Scott Fortin. "The Graph Isomorphism Problem". In: (1996). DOI: 10.7939/R3SX64C5K.

[28] Chu-Chun Fu et al. "Multiscale modelling of defect kinetics in irradiated iron". In: *Nature Materials* 4.1 (Dec. 2004), pp. 68–74. DOI: 10.1038/nmat1286.

[29] Yuh Fukai. *The metal-hydrogen system: basic bulk properties*. Berlin New York: Springer, 2005. ISBN: 978-3-540-28883-1.

[30] W. W. Gerberich et al. "The necessity of both plasticity and brittleness in the fracture thresholds of iron". In: *Philosophical Magazine A* 63.2 (Feb. 1991), pp. 363–376. DOI: 10.1080/01418619108204854.

[31] Fred Glover and Manuel Laguna. "Tabu Search". In: *Handbook of Combinatorial Optimization*. Springer US, 1998, pp. 2093–2229. DOI: 10.1007/978-1-4613-0303-9_33.

[32] W. K. Hastings. "Monte Carlo sampling methods using Markov chains and their applications". In: *Biometrika* 57.1 (Apr. 1970), pp. 97–109. DOI: 10.1093/biomet/57.1.97.

[33] Erin Hayward and Chu-Chun Fu. "Interplay between hydrogen and vacancies in α-Fe". In: *Physical Review B* 87.17 (May 2013). DOI: 10.1103/physrevb.87.174103.

[34] Graeme Henkelman and Hannes Jónsson. "A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives". In: *The Journal of Chemical Physics* 111.15 (Oct. 1999), pp. 7010–7022. DOI: 10.1063/1.480097.

[35] Graeme Henkelman and Hannes Jónsson. "Long time scale kinetic Monte Carlo simulations without lattice approximation and predefined event table". In: *The Journal of Chemical Physics* 115.21 (Dec. 2001), pp. 9657–9666. DOI: 10.1063/1.1415500.

[36] Andreas Heyden, Alexis T. Bell, and Frerich J. Keil. "Efficient methods for finding transition states in chemical reactions: Comparison of improved dimer method and partitioned rational function optimization method". In: *The Journal of Chemical Physics* 123.22 (Dec. 2005), p. 224101. DOI: 10.1063/1.2104507.

[37] D. E. Jiang and Emily A. Carter. "Diffusion of interstitial hydrogen into and through bcc Fe from first principles". In: *Physical Review B* 70.6 (Aug. 2004). DOI: 10.1103/physrevb.70.064102.

[38] WILLIAM H. JOHNSON. "On Some Remarkable Changes Produced in Iron and Steel by the Action of Hydrogen and Acids". In: *Nature* 11.281 (Mar. 1875), pp. 393–393. DOI: 10.1038/011393a0.

[39] Hannes Jónsson, Greg Mills, and Karsten W. Jacobsen. "Nudged elastic band method for finding minimum energy paths of transitions". In: *Classical and Quantum Dynamics in Condensed Phase Simulations*. WORLD SCIENTIFIC, June 1998. DOI: 10.1142/9789812839664_0016.

[40] Mudita Juneja and S. K. Nagar. "Particle swarm optimization algorithm and its parameters: A review". In: *2016 International Conference on Control, Computing, Communication and Materials (ICCCCM)*. IEEE, Oct. 2016. DOI: 10.1109/iccccm.2016.7918233.

[41] Johannes Kästner and Paul Sherwood. "Superlinearly converging dimer method for transition state search". In: *The Journal of Chemical Physics* 128.1 (Jan. 2008), p. 014106. DOI: 10.1063/1.2815812.

[42] Ivaylo H. Katzarov, Dimitar L. Pashov, and Anthony T. Paxton. "Fully quantum mechanical calculation of the diffusivity of hydrogen in iron using the tight-binding approximation and path integral theory". In: *Physical Review B* 88.5 (Aug. 2013). DOI: 10.1103/physrevb.88.054107.

[43] Anastasiia S. Kholtobina et al. "Hydrogen Trapping in bcc Iron". In: *Materials* 13.10 (May 2020), p. 2288. DOI: [10.3390/ma13102288](10.3390/ma13102288).

[44] Sung Jin Kim and Jae-Won Lee. "Metallurgical evidence of stress induced hydrogen diffusion and corresponding crack nucleation behaviors of high-strength ferritic steel used in sour environment". In: *Journal of Materials Research and Technology* 8.5 (Sept. 2019), pp. 5036–5040. DOI: 10.1016/j.jmrt.2019.07.030.

[45] Hajime Kimizuka, Hideki Mori, and Shigenobu Ogata. "Effect of temperature on fast hydrogen diffusion in iron: A path-integral quantum dynamics approach". In: *Physical Review B* 83.9 (Mar. 2011). DOI: 10.1103/physrevb.83.094110.

[46] Dhrubajit Konwar, Vijesh J. Bhute, and Abhijit Chatterjee. "An off-lattice, self-learning kinetic Monte Carlo method using local environments". In: *The Journal of Chemical Physics* 135.17 (Nov. 2011), p. 174103. DOI: [10.1063/1.3657834](10.1063/1.3657834).

[47] C. Lanczos. "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators". In: *Journal of Research of the National Bureau of Standards* 45.4 (Oct. 1950), p. 255. DOI: 10.6028/jres.045.026.

[48] Dong C. Liu and Jorge Nocedal. "On the limited memory BFGS method for large scale optimization". In: *Mathematical Programming* 45.1-3 (Aug. 1989), pp. 503–528. DOI: [10.1007/bf01589116](10.1007/bf01589116).

[49] Gianfranco Lovicu et al. "Hydrogen Embrittlement of Automotive Advanced High-Strength Steels". In: *Metallurgical and Materials Transactions A* 43.11 (June 2012), pp. 4075–4087. DOI: [10.1007/s11661-012-1280-8](10.1007/s11661-012-1280-8).

[50] Gang Lu et al. "Hydrogen-Enhanced Local Plasticity in Aluminum: AnAb InitioStudy". In: *Physical Review Letters* 87.9 (Aug. 2001). DOI: 10.1103/physrevlett.87.095501.

[51] S.P. Lynch. "Hydrogen embrittlement and liquid-metal embrittlement in nickel single crystals". In: *Scripta Metallurgica* 13.11 (Nov. 1979), pp. 1051–1056. DOI: [10.1016/0036-9748(79)90202-3](10.1016/0036-9748(79)90202-3).

[52] Baraka Maiseli, Yanfeng Gu, and Huijun Gao. "Recent developments and trends in point set registration methods". In: *Journal of Visual Communication and Image Representation* 46 (July 2017), pp. 95–106. DOI: 10.1016/j.jvcir.2017.03.012.

[53] Brendan D. McKay and Adolfo Piperno. "Practical graph isomorphism, II". In: *Journal of Symbolic Computation* 60 (Jan. 2014), pp. 94–112. DOI: [10.1016/j.jsc.2013.09.003](10.1016/j.jsc.2013.09.003).

[54] R.B. McLellan and Z.R. Xu. "Hydrogen-induced vacancies in the iron lattice". In: *Scripta Materialia* 36.10 (May 1997), pp. 1201–1205. DOI: [10.1016/s1359-6462(97)00015-8](10.1016/s1359-6462(97)00015-8).

[55] A. McNabb and P. K. Foster. "Institute of Metals Division - A New Analysis of the Diffusion of Hydrogen in Iron and Ferritic Steels". In: *Materials* (Jan. 1963), p. 10.

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

[57] Luca Messina, Lorenzo Malerba, and Pär Olsson. "Stability and mobility of small vacancy–solute complexes in Fe–MnNi and dilute Fe–X alloys: A kinetic Monte Carlo study". In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 352 (June 2015), pp. 61–66. DOI: 10.1016/j.nimb.2014.12.032.

[58] Johannes J Möller and Erik Bitzek. "Comparative study of embedded atom potentials for atomistic simulations of fracture in α-iron". In: *Modelling and Simulation in Materials Science and Engineering* 22.4 (Apr. 2014), p. 045002. DOI: 10.1088/0965-0393/22/4/045002.

[59] M. Nagumo. "Hydrogen related failure of steels – a new aspect". In: *Materials Science and Technology* 20.8 (Aug. 2004), pp. 940–950. DOI: 10.1179/026708304225019687.

[60] Giridhar Nandipati et al. "Off-lattice pattern recognition scheme for kinetic Monte Carlo simulations". In: *Journal of Computational Physics* 231.9 (May 2012), pp. 3548–3560. DOI: 10.1016/j.jcp.2011.12.029.

[61] R. Nazarov, T. Hickel, and J. Neugebauer. "Ab initiostudy of H-vacancy interactions in fcc metals: Implications for the formation of superabundant vacancies". In: *Physical Review B* 89.14 (Apr. 2014). DOI: 10.1103/physrevb.89.144108.

[62] R. Nazarov, T. Hickel, and J. Neugebauer. "First-principles study of the thermodynamics of hydrogen-vacancy interaction in fcc iron". In: *Physical Review B* 82.22 (Dec. 2010). DOI: 10.1103/physrevb.82.224104.

[63] Jorge Nocedal. "Updating quasi-Newton matrices with limited storage". In: *Mathematics of Computation* 35.151 (Sept. 1980), pp. 773–773. DOI: 10.1090/s0025-5718-1980-0572855-7.

[64] R.A Oriani. "The diffusion and trapping of hydrogen in steel". In: *Acta Metallurgica* 18.1 (Jan. 1970), pp. 147–157. DOI: 10.1016/0001-6160(70)90078-7.

[65] Alon Orlitsky, Ananda Theertha Suresh, and Yihong Wu. "Optimal prediction of the number of unseen species". In: *Proceedings of the National Academy of Sciences* 113.47 (Nov. 2016), pp. 13283–13288. DOI: 10.1073/pnas.1607774113.

[66] Anthony T. Paxton and Ivaylo H. Katzarov. "Quantum and isotope effects on hydrogen diffusion, trapping and escape in iron". In: *Acta Materialia* 103 (Jan. 2016), pp. 71–76. DOI: 10.1016/j.actamat.2015.09.054.

[67] J.M Peña, J.A Lozano, and P Larrañaga. "An empirical comparison of four initialization methods for the K-Means algorithm". In: *Pattern Recognition Letters* 20.10 (Oct. 1999), pp. 1027–1040. DOI: 10.1016/s0167-8655(99)00069-0.

[68] Leonard Bessemer Pfeil. "The effect of occluded hydrogen on the tensile strength of iron". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 112.760 (Aug. 1926), pp. 182–195. DOI: 10.1098/rspa.1926.0103.

[69] Brian Puchala, Michael L. Falk, and Krishna Garikipati. "An energy basin finding algorithm for kinetic Monte Carlo acceleration". In: *The Journal of Chemical Physics* 132.13 (Apr. 2010), p. 134104. DOI: 10.1063/1.3369627.

[70] A. Raina, V. S. Deshpande, and N. A. Fleck. "Analysis of electro-permeation of hydrogen in metallic alloys". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 375.2098 (June 2017), p. 20160409. DOI: 10.1098/rsta.2016.0409.

[71] 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. DOI: 10.1016/j.actamat.2017.11.011.

[72] A. Ramasubramaniam et al. "Effect of atomic scale plasticity on hydrogen diffusion in iron: Quantum mechanically informed and on-the-fly kinetic Monte Carlo simulations". In: *Journal of Materials Research* 23.10 (Oct. 2008), pp. 2757–2773. DOI: 10.1557/jmr.2008.0340.

[73] Ashwin Ramasubramaniam, Mitsuhiro Itakura, and Emily A. Carter. "Interatomic potentials for hydrogen in α–iron based on density functional theory". In: *Physical Review B* 79.17 (May 2009). DOI: 10.1103/physrevb.79.174101.

[74] Sebastián Echeverri Restrepo, Henry Lambert, and Anthony T. Paxton. "Effect of hydrogen on vacancy diffusion". In: *Physical Review Materials* 4.11 (Nov. 2020). DOI: 10.1103/physrevmaterials.4.113601.

[75] Osborne Reynolds. "On the effect of acid on the interior of iron wire". In: *Journal of the Franklin Institute* 99.1 (Jan. 1875), pp. 70–72. DOI: 10.1016/0016-0032(75)90215-x.

[76] I.M. Robertson. "The effect of hydrogen on dislocation dynamics". In: *Engineering Fracture Mechanics* 68.6 (Apr. 2001), pp. 671–692. DOI: 10.1016/s0013-7944(01)00011-x.

[77] Hamza M. Ruzayqat and Tim P. Schulze. "A Rejection Scheme for Off-Lattice Kinetic Monte Carlo Simulation". In: *Journal of Chemical Theory and Computation* 14.1 (Dec. 2017), pp. 48–54. DOI: 10.1021/acs.jctc.7b00974.

[78] J. Sanchez et al. "Ab initio molecular dynamics simulation of hydrogen diffusion in α-iron". In: *Physical Review B* 81.13 (Apr. 2010). DOI: 10.1103/physrevb.81.132102.

[79] Peter H. Schönemann. "A generalized solution of the orthogonal procrustes problem". In: *Psychometrika* 31.1 (Mar. 1966), pp. 1–10. DOI: 10.1007/bf02289451.

[80] Syed Islamuddin Shah et al. "Extended pattern recognition scheme for self-learning kinetic Monte Carlo simulations". In: *Journal of Physics: Condensed Matter* 24.35 (Aug. 2012), p. 354004. DOI: 10.1088/0953-8984/24/35/354004.

[81] Jun Song and W. A. Curtin. "Atomic mechanism and prediction of hydrogen embrittlement in iron". In: *Nature Materials* 12.2 (Nov. 2012), pp. 145–151. DOI: 10.1038/nmat3479.

[82] Jun Takahashi, Kazuto Kawakami, and Toshimi Tarui. "Direct observation of hydrogen-trapping sites in vanadium carbide precipitation steel by atom probe tomography". In: *Scripta Materialia* 67.2 (July 2012), pp. 213–216. DOI: 10.1016/j.scriptamat.2012.04.022.

[83] A Tehranchi et al. "Hydrogen–vacancy–dislocation interactions in α-Fe". In: *Modelling and Simulation in Materials Science and Engineering* 25.2 (Dec. 2016), p. 025001. DOI: 10.1088/1361-651x/aa52cb.

[84] S. M. Teus and V. G. Gavriljuk. "Grain-Boundary Diffusion of Hydrogen Atoms in the α-Iron". In: *METALLOFIZIKA I NOVEISHIE TEKHNOLOGII* 36.10 (Sept. 2016), pp. 1399–1410. DOI: 10.15407/mfint.36.10.1399.

[85] Mickaël Trochet et al. "Off-Lattice Kinetic Monte Carlo Methods". In: *Handbook of Materials Modeling*. Springer International Publishing, 2018, pp. 1–29. DOI: 10.1007/978-3-319-42913-7_29-1.

[86] H. Vehoff and P. Neumann. "Crack propagation and cleavage initiation in Fe-2.6%-Si single crystals under controlled plastic crack tip opening rate in various gaseous environments". In: *Acta Metallurgica* 28.3 (Mar. 1980), pp. 265–272. DOI: 10.1016/0001-6160(80)90161-3.

[87] George H. Vineyard. "Frequency factors and isotope effects in solid state rate processes". In: *Journal of Physics and Chemistry of Solids* 3.1-2 (Jan. 1957), pp. 121–127. DOI: 10.1016/0022-3697(57)90059-8.

[88] Shuai Wang et al. "Effect of hydrogen environment on the separation of Fe grain boundaries". In: *Acta Materialia* 107 (Apr. 2016), pp. 279–288. DOI: 10.1016/j.actamat.2016.01.067.

[89] Yan Wang. "Reliable kinetic Monte Carlo simulation based on random set sampling". In: *Soft Computing* 17.8 (Feb. 2013), pp. 1439–1451. DOI: 10.1007/s00500-013-1013-y.

[90] Conor John Williams. "Developing an off-lattice kinetic Monte Carlo model to simulate hydrogen-crystal defect interactions in α–iron". Masters thesis [[unpublished]]. University of Cambridge, 2020.

[91] Yihong Wu and Pengkun Yang. "Chebyshev polynomials, moment matching, and optimal estimation of the unseen". In: *The Annals of Statistics* 47.2 (Apr. 2019). DOI: 10.1214/17-aos1665.

[92] Haixuan Xu, Yuri N Osetsky, and Roger E Stoller. "Self-evolving atomistic kinetic Monte Carlo: fundamentals and applications". In: *Journal of Physics: Condensed Matter* 24.37 (Aug. 2012), p. 375402. DOI: 10.1088/0953-8984/24/37/375402.

[93] Haixuan Xu, Yury N. Osetsky, and Roger E. Stoller. "Simulating complex atomistic processes: On-the-fly kinetic Monte Carlo scheme with selective active volumes". In: *Physical Review B* 84.13 (Oct. 2011). DOI: 10.1103/physrevb.84.132103.

[94] Lijun Xu and Graeme Henkelman. "Adaptive kinetic Monte Carlo for first-principles accelerated dynamics". In: *The Journal of Chemical Physics* 129.11 (Sept. 2008), p. 114104. DOI: 10.1063/1.2976010.

[95] Yi Zeng, Penghao Xiao, and Graeme Henkelman. "Unification of algorithms for minimum mode optimization". In: *The Journal of Chemical Physics* 140.4 (Jan. 2014), p. 044115. DOI: 10.1063/1.4862410.

[96] Z. Zhang. "A flexible new technique for camera calibration". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22.11 (2000), pp. 1330–1334. DOI: 10.1109/34.888718.

[97] Hao Zhu et al. "A Review of Point Set Registration: From Pairwise Registration to Groupwise Registration". In: *Sensors* 19.5 (Mar. 2019), p. 1191. DOI: 10.3390/s19051191.

Appendices

A.1   Partial vector-derivative of the distance between atoms

Here we derive Eq. (3.4), repeated below for convenience:

$$\frac{\partial r^{\alpha\beta}}{\partial r_j^{\gamma}} = \left(\delta^{\gamma\beta} - \delta^{\gamma\alpha}\right)\hat{r}_j^{\alpha\beta} \tag{A.1}$$

where $\boldsymbol{r}^{\alpha}$ is the position vector of the αth atom, $r^{\alpha\beta} = \left\|\boldsymbol{r}^{\alpha\beta}\right\| = \left\|\boldsymbol{r}^{\beta} - \boldsymbol{r}^{\alpha}\right\|$ and δ is the Kronecker delta. Starting from the definition of $r^{\alpha\beta}$:

$$\left[r^{\alpha\beta}\right]^2 = r_i^{\alpha\beta} r_i^{\alpha\beta}$$

$$\frac{\partial}{\partial r_j^{\gamma}}\left[r^{\alpha\beta}\right]^2 = \frac{\partial}{\partial r_j^{\gamma}} r_i^{\alpha\beta} r_i^{\alpha\beta}$$

$$r^{\alpha\beta}\frac{\partial r^{\alpha\beta}}{\partial r_j^{\gamma}} = r_i^{\alpha\beta}\frac{\partial r_i^{\alpha\beta}}{\partial r_j^{\gamma}}$$

$$\frac{\partial r^{\alpha\beta}}{\partial r_j^{\gamma}} = \hat{r}_i^{\alpha\beta}\frac{\partial r_i^{\alpha\beta}}{\partial r_j^{\gamma}} \tag{A.2}$$

then expanding $r_i^{\alpha\beta} = r_i^{\beta} - r_i^{\alpha}$:

$$\frac{\partial r^{\alpha\beta}}{\partial r_j^{\gamma}} = \hat{r}_i^{\alpha\beta}\left(\frac{\partial r_i^{\beta}}{\partial r_j^{\gamma}} - \frac{\partial r_i^{\alpha}}{\partial r_j^{\gamma}}\right) = \hat{r}_i^{\alpha\beta}\left(\delta^{\gamma\beta}\delta_{ij} - \delta^{\gamma\alpha}\delta_{ij}\right) = \left(\delta^{\gamma\beta} - \delta^{\gamma\alpha}\right)\hat{r}_j^{\alpha\beta} \tag{A.3}$$

we arrive at Eq. (3.4).

A.2 Analytical form of the EAM Hessian

Here we simplify Eq. (3.7). We begin by separating the first term:

$$
\frac{\partial^2 U}{\partial r_i^\eta \partial r_j^\gamma} = \sum_{\alpha \neq \gamma} \hat{r}_j^{\alpha\gamma} \frac{\partial \dot{V}^{\alpha\gamma}}{\partial r_i^\eta}
$$

$$
+ \sum_{\alpha \neq \gamma} \hat{r}_j^{\alpha\gamma} \frac{\partial}{\partial r_i^\eta} \left(\dot{F}^\gamma \dot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \dot{\phi}^{\gamma\alpha} \right)
$$

$$
+ \sum_{\alpha \neq \gamma} \left(\dot{V}^{\alpha\gamma} + \dot{F}^\gamma \dot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \dot{\phi}^{\gamma\alpha} \right) \frac{\partial \hat{r}_j^{\alpha\gamma}}{\partial r_i^\eta} \tag{A.4}
$$

then make progress on the partial-derivative in the final line:

$$
\frac{\partial \hat{r}_j^{\alpha\gamma}}{\partial r_i^\eta} = \frac{\partial}{\partial r_i^\eta} \frac{r_j^{\alpha\gamma}}{r^{\alpha\gamma}} = \frac{1}{r^{\alpha\gamma}} \frac{\partial r_j^{\alpha\gamma}}{\partial r_i^\eta} - r_j^{\alpha\gamma} \left(\frac{1}{r^{\alpha\gamma}} \right)^2 \frac{\partial r^{\alpha\gamma}}{\partial r_i^\eta} \tag{A.5}
$$

by applying Eq. (A.3) and $r_j^{\alpha\gamma} = r_j^\gamma - r_j^\alpha$ we find:

$$
\frac{\partial \hat{r}_j^{\alpha\gamma}}{\partial r_i^\eta} = \frac{1}{r^{\alpha\gamma}} \frac{\partial}{\partial r_i^\eta} \left(r_j^\gamma - r_j^\alpha \right) - \left(\frac{1}{r^{\alpha\gamma}} \right)^3 r_j^{\alpha\gamma} r_i^{\alpha\gamma} \left(\delta^{\eta\gamma} - \delta^{\eta\alpha} \right)
$$

$$
= \frac{1}{r^{\alpha\gamma}} \left(\delta^{\eta\gamma} - \delta^{\eta\alpha} \right) \delta_{ij} - \frac{1}{r^{\alpha\gamma}} \hat{r}_j^{\alpha\gamma} \hat{r}_i^{\alpha\gamma} \left(\delta^{\eta\gamma} - \delta^{\eta\alpha} \right)
$$

$$
= \frac{1}{r^{\alpha\gamma}} \left(\delta^{\eta\gamma} - \delta^{\eta\alpha} \right) \left(\delta_{ij} - \hat{r}_i^{\alpha\gamma} \hat{r}_j^{\alpha\gamma} \right) \tag{A.6}
$$

Substituting this back into Eq. (A.4) and splitting up the second line, we obtain:

$$
\frac{\partial^2 U}{\partial r_i^\eta \partial r_j^\gamma} = \sum_{\alpha \neq \gamma} \hat{r}_j^{\alpha\gamma} \ddot{V}^{\alpha\gamma} \frac{\partial r^{\alpha\gamma}}{\partial r_i^\eta}
$$

$$
+ \sum_{\alpha \neq \gamma} \hat{r}_j^{\alpha\gamma} \left(\dot{F}^\gamma \frac{\partial \dot{\phi}^{\alpha\gamma}}{\partial r_i^\eta} + \dot{\phi}^{\alpha\gamma} \frac{\partial \dot{F}^\gamma}{\partial r_i^\eta} \right)
$$

$$
+ \sum_{\alpha \neq \gamma} \hat{r}_j^{\alpha\gamma} \left(\dot{F}^\alpha \frac{\partial \dot{\phi}^{\gamma\alpha}}{\partial r_i^\eta} + \dot{\phi}^{\gamma\alpha} \frac{\partial \dot{F}^\alpha}{\partial r_i^\eta} \right)
$$

$$
+ \sum_{\alpha \neq \gamma} \left(\dot{V}^{\alpha\gamma} + \dot{F}^\gamma \dot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \dot{\phi}^{\gamma\alpha} \right) \frac{1}{r^{\alpha\gamma}} \left(\delta^{\eta\gamma} - \delta^{\eta\alpha} \right) \left(\delta_{ij} - \hat{r}_i^{\alpha\gamma} \hat{r}_j^{\alpha\gamma} \right) \tag{A.7}
$$

Taking Eq. (A.7), we expand the derivatives and contracting the final sum:

$$
\frac{\partial^2 U}{\partial r_i^\eta \partial r_j^\gamma} = \sum_{\alpha \neq \gamma} \hat{r}_j^{\alpha\gamma} \ddot{V}^{\alpha\gamma} \left(\delta^{\eta\gamma} - \delta^{\eta\alpha} \right) \hat{r}_i^{\alpha\gamma}
$$

$$
+ \sum_{\alpha \neq \gamma} \hat{r}_j^{\alpha\gamma} \left(\dot{F}^\gamma \ddot{\phi}^{\alpha\gamma} \frac{\partial r^{\alpha\gamma}}{\partial r_i^\eta} + \dot{\phi}^{\alpha\gamma} \ddot{F}^\gamma \frac{\partial \rho^\gamma}{\partial r_i^\eta} \right)
$$

$$
+ \sum_{\alpha \neq \gamma} \hat{r}_j^{\alpha\gamma} \left(\dot{F}^\alpha \ddot{\phi}^{\gamma\alpha} \frac{\partial r^{\gamma\alpha}}{\partial r_i^\eta} + \dot{\phi}^{\gamma\alpha} \ddot{F}^\alpha \frac{\partial \rho^\alpha}{\partial r_i^\eta} \right)
$$

$$
+ \delta^{\eta\gamma} \sum_{\alpha \neq \gamma} \left(\dot{V}^{\alpha\gamma} + \dot{F}^\gamma \dot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \dot{\phi}^{\gamma\alpha} \right) \frac{1}{r^{\alpha\gamma}} \left(\delta_{ij} - \hat{r}_i^{\alpha\gamma} \hat{r}_j^{\alpha\gamma} \right)
$$

$$
- \left(1 - \delta^{\eta\gamma} \right) \left(\dot{V}^{\eta\gamma} + \dot{F}^\gamma \dot{\phi}^{\eta\gamma} + \dot{F}^\eta \dot{\phi}^{\gamma\eta} \right) \frac{1}{r^{\eta\gamma}} \left(\delta_{ij} - \hat{r}_i^{\eta\gamma} \hat{r}_j^{\eta\gamma} \right) \tag{A.8}
$$

then contracting the first sum and expanding the derivatives a second time to arrive at:

$$
\frac{\partial^2 U}{\partial r_i^\eta \partial r_j^\gamma} = \delta^{\eta\gamma} \sum_{\alpha \neq \gamma} \ddot{V}^{\alpha\gamma} \hat{r}_i^{\alpha\gamma} \hat{r}_j^{\alpha\gamma} - \left(1 - \delta^{\eta\gamma} \right) \ddot{V}^{\eta\gamma} \hat{r}_i^{\eta\gamma} \hat{r}_j^{\eta\gamma}
$$

$$
+ \sum_{\alpha \neq \gamma} \hat{r}_j^{\alpha\gamma} \left(\dot{F}^\gamma \ddot{\phi}^{\alpha\gamma} \left(\delta^{\eta\gamma} - \delta^{\eta\alpha} \right) \hat{r}_i^{\alpha\gamma} + \dot{\phi}^{\alpha\gamma} \ddot{F}^\gamma \sum_{\beta \neq \gamma} \dot{\phi}^{\beta\gamma} \left(\delta^{\eta\gamma} - \delta^{\eta\beta} \right) \hat{r}_i^{\beta\gamma} \right)
$$

$$
+ \sum_{\alpha \neq \gamma} \hat{r}_j^{\alpha\gamma} \left(\dot{F}^\alpha \ddot{\phi}^{\gamma\alpha} \left(\delta^{\eta\alpha} - \delta^{\eta\gamma} \right) \hat{r}_i^{\gamma\alpha} + \dot{\phi}^{\gamma\alpha} \ddot{F}^\alpha \sum_{\beta \neq \alpha} \dot{\phi}^{\beta\alpha} \left(\delta^{\eta\alpha} - \delta^{\eta\beta} \right) \hat{r}_i^{\beta\alpha} \right)
$$

$$
+ \delta^{\eta\gamma} \sum_{\alpha \neq \gamma} \left(\dot{V}^{\alpha\gamma} + \dot{F}^\gamma \dot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \dot{\phi}^{\gamma\alpha} \right) \frac{1}{r^{\alpha\gamma}} \left(\delta_{ij} - \hat{r}_i^{\alpha\gamma} \hat{r}_j^{\alpha\gamma} \right)
$$

$$
- \left(1 - \delta^{\eta\gamma} \right) \left(\dot{V}^{\eta\gamma} + \dot{F}^\gamma \dot{\phi}^{\eta\gamma} + \dot{F}^\eta \dot{\phi}^{\gamma\eta} \right) \frac{1}{r^{\eta\gamma}} \left(\delta_{ij} - \hat{r}_i^{\eta\gamma} \hat{r}_j^{\eta\gamma} \right) \tag{A.9}
$$

wherein we split up the second and third lines before recombining them to yield:

$$
\begin{aligned}
\frac{\partial^2 U}{\partial r_i^\eta \partial r_j^\gamma} &= \delta^{\eta\gamma} \sum_{\alpha \neq \gamma} \ddot{V}^{\alpha\gamma} \hat{r}_i^{\alpha\gamma} \hat{r}_j^{\alpha\gamma} - (1 - \delta^{\eta\gamma}) \ddot{V}^{\eta\gamma} \hat{r}_i^{\eta\gamma} \hat{r}_j^{\eta\gamma} \\
&+ \sum_{\alpha \neq \gamma} \hat{r}_j^{\alpha\gamma} \dot{F}^\gamma \ddot{\phi}^{\alpha\gamma} (\delta^{\eta\gamma} - \delta^{\eta\alpha}) \hat{r}_i^{\alpha\gamma} + \hat{r}_j^{\alpha\gamma} \dot{F}^\alpha \ddot{\phi}^{\gamma\alpha} (\delta^{\eta\alpha} - \delta^{\eta\gamma}) \hat{r}_i^{\gamma\alpha} \\
&+ \sum_{\alpha \neq \gamma} \sum_{\beta \neq \gamma} \hat{r}_j^{\alpha\gamma} \dot{\phi}^{\alpha\gamma} \ddot{F}^\gamma \dot{\phi}^{\beta\gamma} (\delta^{\eta\gamma}) \hat{r}_i^{\beta\gamma} \\
&+ \sum_{\alpha \neq \gamma} \sum_{\beta \neq \alpha} \hat{r}_j^{\alpha\gamma} \dot{\phi}^{\gamma\alpha} \ddot{F}^\alpha \dot{\phi}^{\beta\alpha} (\delta^{\eta\alpha}) \hat{r}_i^{\beta\alpha} - \sum_{\alpha \neq \gamma} \sum_{\beta \neq \gamma} \hat{r}_j^{\alpha\gamma} \dot{\phi}^{\alpha\gamma} \ddot{F}^\gamma \dot{\phi}^{\beta\gamma} \left(\delta^{\eta\beta} \right) \hat{r}_i^{\beta\gamma} \\
&- \sum_{\alpha \neq \gamma} \sum_{\beta \neq \alpha} \hat{r}_j^{\alpha\gamma} \dot{\phi}^{\gamma\alpha} \ddot{F}^\alpha \dot{\phi}^{\beta\alpha} \left(\delta^{\eta\beta} \right) \hat{r}_i^{\beta\alpha} \\
&+ \delta^{\eta\gamma} \sum_{\alpha \neq \gamma} \left(\dot{V}^{\alpha\gamma} + \dot{F}^\gamma \dot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \dot{\phi}^{\gamma\alpha} \right) \frac{1}{r^{\alpha\gamma}} \left(\delta_{ij} - \hat{r}_i^{\alpha\gamma} \hat{r}_j^{\alpha\gamma} \right) \\
&- (1 - \delta^{\eta\gamma}) \left(\dot{V}^{\eta\gamma} + \dot{F}^\gamma \dot{\phi}^{\eta\gamma} + \dot{F}^\eta \dot{\phi}^{\gamma\eta} \right) \frac{1}{r^{\eta\gamma}} \left(\delta_{ij} - \hat{r}_i^{\eta\gamma} \hat{r}_j^{\eta\gamma} \right) \quad (A.10)
\end{aligned}
$$

finally, contracting the remaining Kronecker deltas, we arrive at an expression for the Hessian:

$$
\begin{aligned}
\frac{\partial^2 U}{\partial r_i^\eta \partial r_j^\gamma} &= \delta^{\eta\gamma} \sum_{\alpha \neq \gamma} \ddot{V}^{\alpha\gamma} \hat{r}_i^{\alpha\gamma} \hat{r}_j^{\alpha\gamma} - (1 - \delta^{\eta\gamma}) \ddot{V}^{\eta\gamma} \hat{r}_i^{\eta\gamma} \hat{r}_j^{\eta\gamma} \\
&+ \delta^{\eta\gamma} \sum_{\alpha \neq \gamma} \left(\dot{F}^\gamma \ddot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \ddot{\phi}^{\gamma\alpha} \right) \hat{r}_i^{\alpha\gamma} \hat{r}_j^{\alpha\gamma} - (1 - \delta^{\eta\gamma}) \left(\dot{F}^\gamma \ddot{\phi}^{\eta\gamma} + \dot{F}^\eta \ddot{\phi}^{\gamma\eta} \right) \hat{r}_i^{\eta\gamma} \hat{r}_j^{\eta\gamma} \\
&+ \delta^{\eta\gamma} \sum_{\alpha \neq \gamma} \sum_{\beta \neq \gamma} \hat{r}_j^{\alpha\gamma} \dot{\phi}^{\alpha\gamma} \ddot{F}^\gamma \dot{\phi}^{\beta\gamma} \hat{r}_i^{\beta\gamma} \\
&+ (1 - \delta^{\eta\gamma}) \left[\sum_{\alpha \neq \eta} \hat{r}_j^{\eta\gamma} \dot{\phi}^{\gamma\eta} \ddot{F}^\eta \dot{\phi}^{\alpha\eta} \hat{r}_i^{\alpha\eta} - \sum_{\alpha \neq \gamma} \hat{r}_j^{\alpha\gamma} \dot{\phi}^{\alpha\gamma} \ddot{F}^\gamma \dot{\phi}^{\eta\gamma} \hat{r}_i^{\eta\gamma} \right] \\
&- \sum_{\alpha \neq \gamma, \eta} \hat{r}_j^{\alpha\gamma} \dot{\phi}^{\gamma\alpha} \ddot{F}^\alpha \dot{\phi}^{\eta\alpha} \hat{r}_i^{\eta\alpha} \\
&+ \delta^{\eta\gamma} \sum_{\alpha \neq \gamma} \left(\dot{V}^{\alpha\gamma} + \dot{F}^\gamma \dot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \dot{\phi}^{\gamma\alpha} \right) \frac{1}{r^{\alpha\gamma}} \left(\delta_{ij} - \hat{r}_i^{\alpha\gamma} \hat{r}_j^{\alpha\gamma} \right) \\
&- (1 - \delta^{\eta\gamma}) \left(\dot{V}^{\eta\gamma} + \dot{F}^\gamma \dot{\phi}^{\eta\gamma} + \dot{F}^\eta \dot{\phi}^{\gamma\eta} \right) \frac{1}{r^{\eta\gamma}} \left(\delta_{ij} - \hat{r}_i^{\eta\gamma} \hat{r}_j^{\eta\gamma} \right) \quad (A.11)
\end{aligned}
$$

Equation (A.11) has two very-separate regimes for the block-diagonal and off-diagonal elements of the Hessian. Exploring the block-diagonal elements, for which $\eta = \gamma$:

$$\frac{\partial^2 U}{\partial r_i^\gamma \partial r_j^\gamma} + \sum_{\alpha \neq \gamma} \left(\ddot{V}^{\alpha\gamma} + \dot{F}^\gamma \ddot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \ddot{\phi}^{\gamma\alpha} + \ddot{F}^\alpha \dot{\phi}^{\gamma\alpha} \dot{\phi}^{\gamma\alpha} \right) \hat{r}_i^{\alpha\gamma} \hat{r}_j^{\alpha\gamma}$$

$$+ \ddot{F}^\gamma \left[\sum_{\beta \neq \gamma} \dot{\phi}^{\beta\gamma} \hat{r}_i^{\beta\gamma} \right] \left[\sum_{\alpha \neq \gamma} \dot{\phi}^{\alpha\gamma} \hat{r}_j^{\alpha\gamma} \right]$$

$$+ \sum_{\alpha \neq \gamma} \left(\dot{V}^{\alpha\gamma} + \dot{F}^\gamma \dot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \dot{\phi}^{\gamma\alpha} \right) \frac{1}{r^{\alpha\gamma}} \left(\delta_{ij} - \hat{r}_i^{\alpha\gamma} \hat{r}_j^{\alpha\gamma} \right) \tag{A.12}$$

which, if we introduce the electron-density dipole:

$$\mu_i^\beta = \sum_{\alpha \neq \beta} \dot{\phi}^{\alpha\beta} \hat{r}_i^{\alpha\beta} \tag{A.13}$$

and symmetric tensors A, B such that:

$$A^{\alpha\beta} = \frac{\dot{V}^{\alpha\beta} + \dot{F}^\beta \dot{\phi}^{\alpha\beta} + \dot{F}^\alpha \dot{\phi}^{\beta\alpha}}{r^{\alpha\beta}} \quad \text{and} \quad B^{\alpha\beta} = \ddot{V}^{\alpha\gamma} + \dot{F}^\gamma \ddot{\phi}^{\alpha\gamma} + \dot{F}^\alpha \ddot{\phi}^{\gamma\alpha} \tag{A.14}$$

becomes:

$$\frac{\partial^2 U}{\partial r_i^\gamma \partial r_j^\gamma} = \ddot{F}^\gamma \mu_i^\gamma \mu_j^\gamma + \sum_{\alpha \neq \gamma} A^{\alpha\gamma} \delta_{ij} + \left(A^{\alpha\gamma} - B^{\alpha\gamma} - \ddot{F}^\alpha \dot{\phi}^{\gamma\alpha} \dot{\phi}^{\gamma\alpha} \right) \hat{r}_i^{\alpha\gamma} \hat{r}_j^{\gamma\alpha} \tag{A.15}$$

Here the sum over α runs over all neighbours of γ and we require the density ρ^α of all of these neighbours. However, if we loop twice over all atoms, in the first pass we can compute $\boldsymbol{\mu}$, ρ for each atom and in the second pass evaluate the block-diagonal Hessian elements.

For the off-diagonal elements of the Hessian, for which $\eta \neq \gamma$, we can simplify Eq. (A.11) to:

$$
\begin{aligned}
\left. \frac{\partial^2 U}{\partial r_i^\eta \partial r_j^\gamma} \right|_{\eta \neq \gamma} = & - \left(\ddot{V}^{\eta\gamma} + \dot{F}^\gamma \ddot{\phi}^{\eta\gamma} + \dot{F}^\eta \ddot{\phi}^{\gamma\eta} \right) \hat{r}_i^{\eta\gamma} \hat{r}_j^{\eta\gamma} \\
& + \sum_{\alpha \neq \eta} \hat{r}_j^{\eta\gamma} \dot{\phi}^{\gamma\eta} \ddot{F}^\eta \dot{\phi}^{\alpha\eta} \hat{r}_i^{\alpha\eta} - \sum_{\alpha \neq \gamma} \hat{r}_j^{\alpha\gamma} \dot{\phi}^{\alpha\gamma} \ddot{F}^\gamma \dot{\phi}^{\eta\gamma} \hat{r}_i^{\eta\gamma} \\
& - \sum_{\alpha \neq \gamma, \eta} \dot{\phi}^{\gamma\alpha} \ddot{F}^\alpha \dot{\phi}^{\eta\alpha} \hat{r}_i^{\eta\alpha} \hat{r}_j^{\alpha\gamma} \\
& - \left(\dot{V}^{\eta\gamma} + \dot{F}^\gamma \dot{\phi}^{\eta\gamma} + \dot{F}^\eta \dot{\phi}^{\gamma\eta} \right) \frac{1}{r^{\eta\gamma}} \left(\delta_{ij} - \hat{r}_i^{\eta\gamma} \hat{r}_j^{\eta\gamma} \right)
\end{aligned}
\tag{A.16}
$$

which – after substituting in the gradient-component function and electron-density dipole – reduces to:

$$
\left. \frac{\partial^2 U}{\partial r_i^\eta \partial r_j^\gamma} \right|_{\eta \neq \gamma} = (B^{\eta\gamma} - A^{\eta\gamma}) \hat{r}_i^{\eta\gamma} \hat{r}_j^{\gamma\eta} - A^{\eta\gamma} \delta_{ij} + \ddot{F}^\eta \dot{\phi}^{\gamma\eta} \mu_i^\eta \hat{r}_j^{\eta\gamma} - \ddot{F}^\gamma \dot{\phi}^{\eta\gamma} \hat{r}_i^{\eta\gamma} \mu_j^\gamma - O_{ij}^{\eta\gamma} \tag{A.17}
$$

Here the first four terms are non-zero when $r^{\eta\gamma} < r_{\text{cut}}$ and depend on $\boldsymbol{\mu}^\eta$, $\boldsymbol{\mu}^\beta$, ρ^η and ρ^β – which can be computed on the first pass of a two-loop evaluation. However, the overlap term:

$$
O_{ij}^{\eta\gamma} = \sum_{\alpha \neq \gamma, \eta} \ddot{F}^\alpha \dot{\phi}^{\gamma\alpha} \dot{\phi}^{\eta\alpha} \hat{r}_i^{\eta\alpha} \hat{r}_j^{\alpha\gamma} \tag{A.18}
$$

is more complex as the sum runs-over the intersection of the neighbours of atoms η and γ thus, can be non-zero when $r^{\eta\gamma} < 2r_{\text{cut}}$.