Off-Lattice Kinetic Monte Carlo Methods

Mickaël Trochet, Normand Mousseau, Laurent Karim Béland, and Graeme Henkelman

Contents

M. Trochet (✉) · N. Mousseau (✉)
Departement de physique and Regroupement québécois sur les matériaux de pointe, Université de Montréal, Montréal, QC, Canada
e-mail: mickael.trochet@umontreal.ca; normand.mousseau@umontreal.ca; nm@normandmousseau.com

L. K. Béland (✉)
Department of Mechanical and Materials Engineering, Queen's University, Kingston, ON, Canada
e-mail: laurent.beland@queensu.ca

G. Henkelman (✉)
Department of Chemistry and the Institute for Computational and Engineering Sciences, University of Texas at Austin, Austin, TX, USA
e-mail: henkelman@utexas.edu

W. Andreoni, S. Yip (eds.), *Handbook of Materials Modeling*,
https://doi.org/10.1007/978-3-319-42913-7_29-1

Abstract

Exact modeling of the dynamics of chemical and material systems over experimentally relevant time scales still eludes us even with modern computational resources. Fortunately, many systems can be described as rare event systems where atoms vibrate around equilibrium positions for a long time before a transition is made to a new atomic state. For those systems, the kinetic Monte Carlo (KMC) algorithm provides a powerful solution. In traditional KMC, mechanism and rates are computed beforehand, limiting moves to discretized positions and largely ignoring strain. Many systems of interest, however, are not well-represented by such lattice-based models. Moreover, materials often evolve with complex and concerted mechanisms that cannot be anticipated before the start of a simulation. In this chapter, we describe a class of algorithms, called off-lattice or adaptive KMC, which relaxes both limitations of traditional KMC, with atomic configurations represented in the full configuration space and reaction events are calculated on-the-fly, with the possible use of catalogs to speed up calculations. We discuss a number of implementations of off-lattice KMC developed by different research groups, emphasizing the similarities between the approaches that open modeling to new classes of problems.

1 Introduction

Modeling the dynamics of chemical and material systems is a fundamental challenge for computational scientists. While the equations of motion of atomic scale systems have been known since the days of Newton, their integration over experimentally relevant time scales still eludes us even given modern computational resources. With empirical potentials, we can achieve nanoseconds of simulation time per day, and with ab initio methods, we are limited to picoseconds. But in most chemical and material applications, we are interested in the human time scales of seconds to minutes, which are relevant for applications including catalysts and batteries. Bridging the time scale gap between what we can model with molecular dynamics (MD) and practical applications is key to making molecular simulations relevant.

Fortunately, many systems in chemistry and material science are what can be described as rare event systems. In this case, atoms in the material vibrate around equilibrium positions for a long time before a transition is made to a new atomic

configuration, or state. This might be the hopping of a Li atom between sites in a battery material or a bond-breaking event in a catalytic reaction. With microscopic kinetics taking place on time scale well separated from phonons, rare event systems open the door to the application of various numerical solutions based on the transition state theory that are not available generally: instead of modeling the vibrational dynamics, which occurs on a femtosecond time scale, transition state theory can be used to average over the thermal motion and calculate the rate of the slowest rare event of interest. If all such important rates can be found, the state-to-state evolution of the system can be calculated on the time scale of the rare events.

A powerful approach for modeling the evolution of a system when the rates are known is the kinetic Monte Carlo (KMC) method. In traditional KMC, the mechanism and the rate of every possible event are required before a calculation is started. This requirement essentially limits atomic KMC simulations to be defined on a lattice where interactions between atoms and atomic motion can be defined discretely. In KMC, a table of all possible events is made, and a single event is chosen with a probability proportional to its rate. The amount of time that evolves between KMC steps is, on average, given by the inverse of the sum of the rates. Thus, each KMC step can be accomplished with just a couple of random numbers and the bookkeeping required to keep track of the possible events; this computational efficiency allows for KMC simulations of millions or billions of events and time scale orders of magnitude longer than any single elementary event.

Obvious limitations of traditional KMC are (i) the need to know all possible events a priori and (ii) the representation of atomic configurations on a lattice. Many atomic systems of interest, however, especially in the presence of defects or disorder, are not well-represented by a lattice-based model. Additionally, materials often evolve with complex and concerted mechanisms that cannot be anticipated before the start of a simulation; it is the evolution of the system into unanticipated configuration via unexpected events that make simulations most interesting.

The purpose of this chapter is to describe a class of algorithms, called off-lattice or adaptive KMC, which relaxes both limitations of traditional KMC. Specifically, atomic configurations are represented in the full configuration space, and reaction events are calculated on-the-fly so that the KMC event table is not fixed but rather adapts as the simulation progresses. Thus, off-lattice KMC can be applied to a much wider range of interesting systems, and the calculated evolution of the system can reveal unexpected dynamics that were not anticipated by the modeler.

There are now a number of implementations of off-lattice KMC developed by different research groups. Later in the chapter, we will discuss some of the differences in philosophy and specific algorithms, but a primary objective here is to highlight what is common between off-lattice KMC methods and emphasize that what may appear to be different methods with different names are in fact often minor variations on a common theme.

2 Off-Lattice KMC

Off-lattice KMC was inspired by a number of closely related earlier methods. In an approach by Sørensen et al., high-temperature MD simulations were used to determine possible reaction mechanisms between a probe tip and a metal surface (Sørensen et al. 1996). These pathways were refined using the nudged elastic band (NEB) method (Jónsson et al. 1998; Henkelman and Jónsson 2000; Henkelman et al. 2000) to determine the energy barriers and rates at the lower temperature of interest. While the authors did not fully sample a rate table for KMC, their work anticipates the philosophy of off-lattice KMC and, in fact, was later developed into an accelerated molecular dynamics method called temperature-accelerated dynamics (Sørensen and Voter 2000). A second method with close connections to off-lattice KMC is the activation-relaxation technique (ART) (Mousseau and Barkema 1998b) in which local arrangements of atoms were driven away from their minimum positions on the potential energy landscape, over ridges, and into neighboring basins of attraction. Again, no KMC event tables were calculated, but the idea of sampling reaction pathways from local displacements of atoms became a key idea for off-lattice KMC methods. ART was later extended into an off-lattice KMC method called kinetic ART, the details of which will be discussed here (El-Mellouhi et al. 2008; Béland et al. 2011).

The earliest implementation of off-lattice KMC that we know of was in 2001 by Henkelman and Jónsson (2001). In this method, local atomic displacements were made to initiate minimum-mode following saddle point searches; these searches iteratively find the lowest mode of the Hessian and follow this mode to a saddle point.

Over the 15 years since the first introduction of off-lattice KMC, there have been a number of significant improvements to the algorithm. Much of the development has been devoted to identifying and implementing efficient methods for storing and reusing information about reactive events that have been calculated. Another key issue, which will be discussed later, is the so-called low-barrier or flickering problem in which fast events limit the overall time scale that can be achieved by off-lattice KMC. Today, there are a number of approaches associated with different aspects of off-lattice KMC methods. Yet, they can generally be described in the single framework outlined in Fig. 1. In brief, all off-lattice KMC methods start from an initial state and explore configuration space around the state to find reactive events which lead to adjacent states. While not absolutely essential, existing methods tend to focus on finding saddle points on the boundary of the initial state and use the harmonic approximation to TST to calculate the rate of each possible reaction mechanism. Current saddle points can be found using a number of ways: minimum mode following saddle point searches, high-temperature MD, the recycling of saddle point information from previous steps, or informed by a database of topologies or structures. As new reaction mechanisms are found, they are added to the rate table and passed to any database being used to store event information. Also, as the configuration space around the initial state is searched for possible reaction events, and the KMC event table is constructed, an estimation of

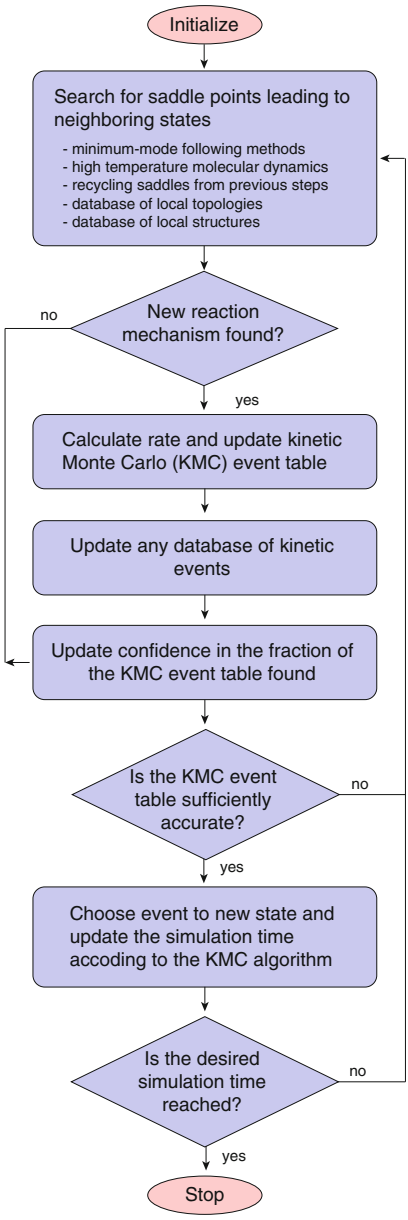

Fig. 1 Flow chart showing the general structure of off-lattice KMC algorithms

the completeness of the event table is updated. When there is sufficient confidence in the event table for escaping the current state, a KMC move is made to a new state, and the process is repeated.

The various off-lattice KMC implementations differ in the specifics in how saddle points are found and stored for reuse and how low barriers are managed. A number of the different philosophies and methods will be discussed next.

3 Search for Saddle Points Leading to Neighboring States

Even for relatively simple systems such as diffusing interstitials in metals, the complexity of the energy landscape is such that it is not possible to identify by hand all mechanisms (Henkelman and Jónsson 1999; Marinica et al. 2011). Open-ended event search methods are therefore necessary to identify the diffusion mechanisms and their associated barriers. There are two main classes of open-ended methods: (1) minimum-mode following methods and (2) molecular-dynamics based methods. Furthermore, local structures and topologies can be classified in order to identify where saddle points searches should be performed and also in order to classify and reuse these events in subsequent steps.

3.1 Minimum-Mode Following Methods

The minimum-mode following approach stabilizes first-order saddle points by finding the direction of lowest curvature on the potential energy surface, inverting the force in that direction, and relaxing the system, guided by these modified forces. A number of such methods have been proposed. The method used by Henkelman and Jónsson was named the dimer method because two images, separated by a finite displacement, were used to approximate the local curvature (Henkelman and Jónsson 1999). The dimer method was later shown to be equivalent to a method developed at the same time in the Wales' group, called the hybrid-eigenvector following method (Munro and Wales 1999). In fact, both methods use a Raleigh-Ritz quotient for iteratively determining the lowest curvature mode and a force inversion along this mode to stabilize first-order saddle points. Around the same time, ART was proposed, using a force projection approach to move the conformation toward saddle point (Barkema and Mousseau 1996). For better stability and convergence, the force projection was replaced by a Lanczos-based algorithm a few years later, forming ART nouveau (ARTn) (Malek and Mousseau 2000). It was subsequently further optimized, as described in details in Machado-Charry et al. (2011). Through benchmarking (Chill et al. 2014a) and a mathematical analysis (Zeng et al. 2014), it is now clear that these minimum-mode following methods are essentially equivalent in terms of computational efficiency for finding saddle points and that they differ mostly in their specific implementation.

A typical minimum-mode following search occurs in three steps:

1. From a local minimum, an atom and possibly its neighborhood are displaced in a direction that can be initiate randomly or systematically which aim to push the system out of the local harmonic state. To avoid collisions between

atoms, the system can be partially relaxed in the hyperplane perpendicular to the displacement. This procedure can also be repeated until the lowest eigenvalue becomes negative or falls below a given threshold (typically between -1 and $-10\,\text{eV/Å}^2$ for bulk semiconductors and metals).
2. The system is iteratively maximized along the negative (lowest) eigenvector and minimized in the hyperplane orthogonal to this direction. In the dimer and hybrid-eigenvector following methods, this is accomplished by following an effective force with the component along the negative mode inverted. Any optimizer, such as L-BFGS (Nocedal 1980), can then be used to converge to a first-order saddle point. In the ARTn method, the system is moved along the direction of the negative eigenvalue away from the initial minimum, and the energy is minimized in the orthogonal hyperplane at each iteration. In ARTn, if at any point the lowest eigenvalue becomes positive, iteration is stopped, and a new event search is launched, going back to (1). If the lowest eigenvalue becomes positive in the dimer method, the system is pushed up along the lowest positive mode until a negative mode is recovered. For either of these methods, a saddle point is considered located when the total force on the system falls below a given threshold (typically 0.01 eV/Å) with a negative lowest eigenvalue.
3. From the saddle point, the system is displaced along the negative mode and relaxed to find the connecting final minimum, completing the event.

Although the initial deformation is often limited to an atom and its surroundings, all the atoms in the box are allowed to respond to this change and to move, avoiding constraints on the nature and the size of the transition state. As the system leaves the harmonic basin and converges onto a first-order saddle point, however, many atoms initially displaced fall back close to their original position as events tend to be local in nature. In ARTn typically 50% of the time the negative eigenvalue is lost as the system is pushed along the eigenvector associated with the lowest eigenvalue, and it falls back into the initial minimum, as the structure of the energy landscape can include shoulders and bumps that do not lead to transition state. Any attempt to eliminate these lost events requires large initial deformations that typically bring the system to saddle points that are not directly connected to the initial minimum, breaking the continuity of the trajectory. With the dimer method, when a negative mode is lost, the system follows the lowest mode up the potential until a negative mode is recovered. While this approach allows minimum-mode following searches to push through positive curvatures, there is a greater chance of finding saddles that are disconnected from the minimum. The recently developed κ-dimer method, (Xiao et al. 2014) which uses the isocontour curvature to detect boundaries of the initial state, largely eliminates the problem of disconnected saddles.

The discovery of disconnected saddles is not always a problem. Wales and collaborators, for example, use large deformations to facilitate the construction of a transition matrix and can be then used to extract global kinetic information (Wales 2002). For other applications, however, it is essential to produce a continuous trajectory, and all transition states are tested to ensure that they are directly connected to the initial minimum.

As long as the lowest eigenvalue remains negative, moving to the saddle point is straightforward. While a convergence force of 0.01 eV/Å is generally chosen, ensuring convergence of the energy barrier to less than 0.01 eV within a few hundred force evaluations, the convergence criterion can be tightened.

3.2 High-Temperature Molecular Dynamics

An alternative to minimum-mode following methods for finding saddle points is to perform high-temperature MD initiated within the current state and use periodic quenching to see if a transition has been made. When a transition to a new state is detected, a double-ended saddle search method is used to find the minimum energy path between the two states. An efficient strategy for this is to use the climbing-image nudged elastic band method (Henkelman and Jónsson 2000; Henkelman et al. 2000) and then optionally, to save computational time, switch to a minimum-mode following search from within the neighborhood of the saddle.

The MD approach for finding saddle points from an initial state is very similar to the temperature-accelerated dynamics method (Sørensen and Voter 2000). The difference is that temperature-accelerated dynamics uses high-temperature MD to find for the first escape that would take place at the low temperature of interest, using harmonic TST to extrapolate to the low-temperature escape time. For off-lattice KMC, high-temperature MD is used to find all of the low-energy escape mechanisms and rates. What is common to both methods is that the use of MD to search for saddles can provide a rigorous confidence measure in the accuracy of the simulation. In the case of off-lattice KMC, this confidence measure is the fraction of the rate table which has been found using the MD saddle searches (Chill and Henkelman 2014).

Each MD saddle search typically takes more computational work to find a saddle than a minimum-mode following search. For off-lattice KMC, however, the cost of finding any one saddle is less important than the cost of evaluating the full rate table. For that, MD and minimum-mode following searches are competitive. If there is sufficient knowledge about the system to target minimum-mode following searches, they can find the rate table more efficiently. However, if there are many processes available to the system, MD searches can selectively find those which are most important because events with a higher rate are found more rapidly than slow events with MD searches (Chill and Henkelman 2014). Independent of efficiency, however, the main benefit of the MD searches is the confidence measure that it provides and the simplicity of having just one parameter to set, the high temperature, although anharmonicity may induce a false sense of completeness as discussed below.

4 Classifying Local Off-Lattice Environments

It is a good practice to classify local structures in order to identify those on which the saddle point searches should be centered, as well as to catalog events to be reused for future use. Several strategies for classifying structures and storing saddle point information are described here.

4.1 Lattice-Based Classification

To be useful, structural classification must be able to reduce a wide range of global conformations onto a finite set of locally defined variables. When atomic motion is limited to a discrete set of positions, there is no ambiguity when comparing two states: they're the same or they're not. For off-lattice calculations, atomic positions can occupy a continuous range of values. In effect, it is highly unlikely for two different global conformations to present identical local environments. To classify and compare structures in an off-lattice system, it is possible to define a discretization procedure that will map the continuous array of solutions into a discrete set of states while ensuring that these states share sufficient similarity when it comes, for example, to their list of diffusion mechanisms.

Pattern recognition approaches, such as the one adopted by Trushin et al. (2005) with the self-learning kinetic Monte Carlo method (SLKMC), are a step away from standard predetermined catalogs, allowing to treat a wider range of conformations and, therefore, better include local strain effects. Nevertheless, this algorithm still requires an underlying lattice to ensure the discretization of the configurations and events, and, while barriers are constructed on the fly, the event catalog consists only of single-atom nearest-neighbor hops characterized using the drag method.

With the self-evolving atomistic kinetic Monte Carlo (SEAK-MC), Xu, Osetsky, and Stoller Xu et al. (2011) propose an off-lattice approach limited to near crystalline configurations with an on-the-fly event searching step. A particular feature of this approach is the construction of *active volumes* associated with defects or noncrystalline environments. These defects are identified using a geometric criterion. Focusing on these defective regions reduces the computational effort for constructing an event catalog. Dimer (or ARTn Béland et al. 2015a) searches are used to search for mechanisms, starting from each defect in the active volume. After a KMC step, only the defects in the affected active volume are sampled for diffusion; events associated with other regions are simply carried over from the previous step. Although SEAK-MC can handle more complex environments than SLKMC, because of its reliance on defects to identify active volume, the method is not applicable to fully disordered configurations. Since it does not reconstruct all barriers after each step, moreover, SEAK-MC does not fully take into account long-range elastic effects.

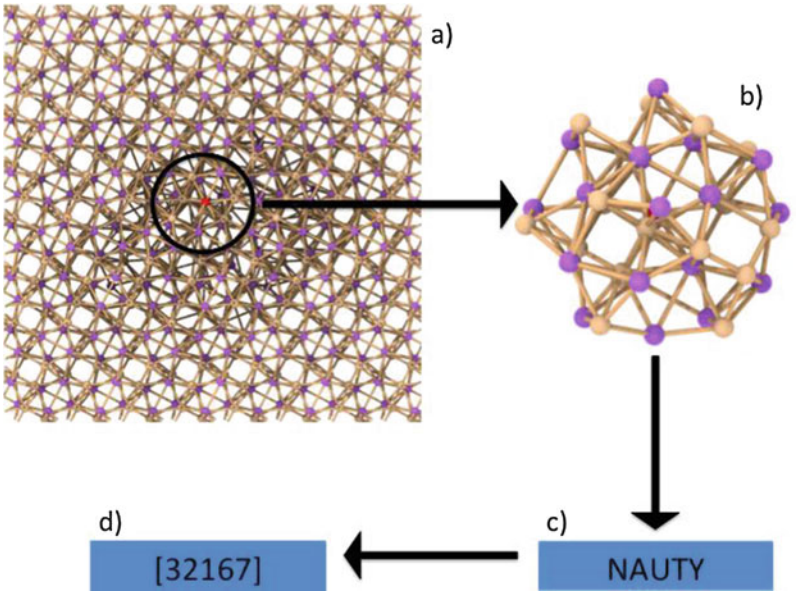

Fig. 2 The topological classification of each atom, in k-ART, is determined by its local environment. (**a**) An atom and its neighborhood are first identified; (**b**) vertices are drawn between atoms within a cutoff distance generally selected between first and second neighbor; (**c**) the resulting graph is then sent to NAUTY, which returns an identifier characterizing the automorphic class its belongs too and a set of transformation onto a reference graph

4.2 Topological Classification

Introducing a topological classification, the kinetic activation-relaxation technique (k-ART), first published in 2008 (El-Mellouhi et al. 2008), lifts these limitations and is the first fully off-lattice KMC approach with on-the-fly cataloging.

K-ART attributes a topological key to each atom in a cell based on its local environment (see Fig. 2): all neighboring atoms within a sphere centered on the reference atom are considered as vertices of a graph, with edges drawn between atoms within a specific cutoff, typically, but not always, set between first and second nearest distances. As a function of the system's complexity, the sphere radius is generally set between 6 and 8 Å, including around 50 to 80 atoms. The graph, as generated, is then sent to NAUTY, a rich topological analysis code developed over many decades by McKay (McKay et al. 1981; McKay and Piperno 2014) used as a library by k-ART. NAUTY returns a key characteristic of the graph's automorphic class as well as a mapping list into a reference graph. This list is used by k-ART to map specific local environments onto reference configurations found in the event catalog.

For each key, a series of ART nouveau event searches is launched centered on the atom associated with this key. Saddle points with different topological keys are considered unique and stored. The topological key associated with a specific atom defines the list of possible events it can undergo. Using the mapping list, it is

possible to assign a correspondence between the specific environment and the atoms in the catalog associated with the generic event. Using this correspondence and the atomic positions of the generic configuration at the saddle point, one can reconstruct an approximate saddle point for the specific configuration. In the presence of strain, this saddle point will be close to the reference event, but not exactly the same, and a few steps of ART nouveau relaxation are needed to converge onto the specific saddle point, providing a precise information on the geometry and energetics of the transition state.

K-ART's basic assumption poses that local minima in the same automorphic class share a unique set of events, characterized by the topological classifications, with the geometrical details at the saddle point depending on the elastic deformation. This allows k-ART to store events based on the local topology and reconstruct their specific geometry at these specific points on the energy landscape. This assumption is valid most of the time for three reasons: (i) the correspondence is set to work only at specific points on the energy landscape – minima and first-order saddle points; (ii) the correspondence holds therefore only for a given forcefield or, equivalently, a unique energy landscape; and (iii) the local graph is reconstructed after it is embedded within the larger two- or three-dimensional space and attached to the rest of the network. This one-to-one correspondence between topology and geometry can fail. However, this failure can be identified readily as activated configurations reconstructed from a reference geometry will not show a first-order saddle point in their vicinity. In general this error indicates that more than one geometries is associated with the topology. To lift this degeneracy, k-ART reduced the edge cutoff criterion until these geometries are assigned to different classes.

PESTO, by Louis Vernon (Vernon et al. 2011; Vernon 2010, 2012), is a variation on adaptive KMC, which has integrated a number of k-ART's features. While topological classification in k-ART is based on the local environment of *atoms*, PESTO is based on the local environment of *defects*. This is a three-step process. First, defects are identified using a variety of schemes including comparison to a reference lattice. These defects can be point or extended defects. Second, NAUTY is used to identify the defect's topology. While k-ART looks at cluster of atoms centered on individual atoms, PESTO looks at cluster of atoms centered on defects. Third, the positions of the atoms in this cluster are compared to those of previously stored clusters for this topology. If they match those of the stored cluster, the configuration is considered to be known. If they do not, the configuration is considered to be new and it is stored.

In 2016, Alexander and Schuh developed a version of k-ART that used a systematic and orderly search and classified transition based on atomic motion vectors, rather than a topological classification. Displacement vectors of each atom between the saddle point and the initial state and between the final and the initial states are stored for future comparisons. Comparison of two transitions is performed by comparing the x, y, and z displacement of each atoms of both transitions. The goal pursued by Alexander and Schuh is to evaluate the convergence of the residence time by assessing the completeness of search for any configuration in the potential energy landscape.

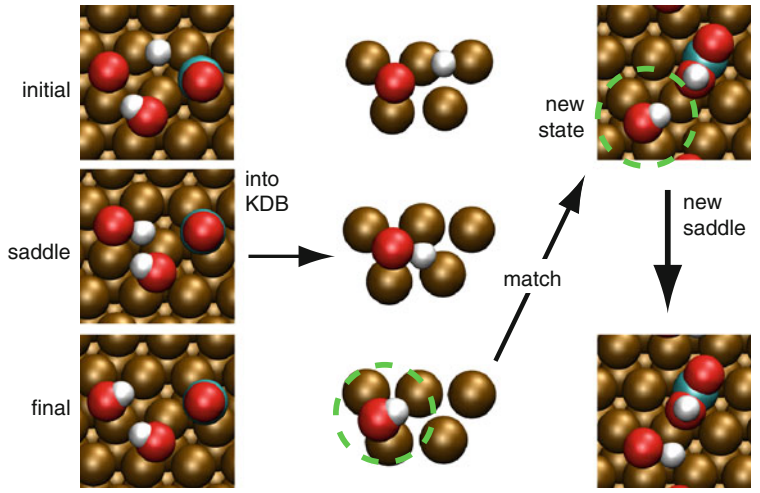

Fig. 3 Kinetic database example of OH and CO reacting on Au(111). When a new reaction mechanism is found, the configurations of the moving atoms and their first neighbors are stored in the database. In any subsequent state, the database is queried to see if any known initial or final states match

4.3 Geometric Classification

An alternative to characterizing local environments by the bonding topology is to directly use the atomic geometry of the atoms which participate in a reactive event. In the kinetic database (KDB) approach (Terrell et al. 2012), all atoms that move by more than a specified distance are considered part of the reaction mechanisms. The position of those atoms in the initial, saddle, and final states, as well as their direct neighbors, is stored in the KDB. Figure 3 shows a calculation of CO oxidation on Au(111) in the presence of hydroxyl, using forces from density functional theory (Ojifinni et al. 2008). When a hydroxyl formation mechanism is discovered, the local geometry of the reacting atoms are stored in the KDB. Later in the simulation, or in an entirely different simulation, the KDB tries to match the reaction mechanisms stored to the current geometry. Specifically, matching is done using a geometric fingerprint of each atom in the database structure (initial or final state geometries). Using a depth-first tree search, neighboring atoms in both the database and current configuration are matched. If a complete match is made, the optimal alignment is calculated, and a score is assigned to each candidate, based upon the difference in atomic positions between those in the database and those in the current configuration. In Fig. 3, the final hydroxyl structure in the database was matched to a hydroxyl on the surface in a subsequent step in the off-lattice KMC simulation. The KDB is then used to predict possible saddle points by moving the local atoms to their saddle point configuration from the database. These suggested saddle geometries are used as initial configurations for minimum-mode following

searches so that the true saddle point and rate of the event in the new environment are calculated. Importantly, the KDB is used to provide approximate saddle point configurations for the current geometry based upon what has been seen before. Only if an exact saddle can be found is the mechanism entered into the rate table. In this way, there is no additional approximation for off-lattice simulations using the KDB. As the database stores a greater number of kinetic events, saddle suggestions become more accurate and converge more rapidly to true saddles, when they exist, so that the cost of the off-lattice KMC simulation approaches that of KMC.

There are some key differences between the topological classification of reactive events used in k-ART and the geometric classification in the KDB. K-ART has the advantage that matching structures from the database to the current simulation is extremely efficient. Also, if there is sufficient trust in the catalog of events in the database, then new saddle searches need only to be done when an unknown topology is reached. In the KDB, on the other hand, the structure matching is not discretized; geometries which are close, but not exact, are used to suggest saddle point structures. When these structures are converged, the exact geometry of the saddle and the activation energy are calculated. If no saddle point is found, the loss is only the computational time required for the saddle search. Both methods can be used to store and recover information that has been learned to accelerate off-lattice simulations.

4.4 Challenges in the Event Reconstruction

Reconstruction of saddle points is crucial in any scheme that creates a generic catalog built either on geometrical or topological keys as symmetry operations and elastic deformations in an off-lattice system must be taken account.

To describe this challenge, we focus here on the topological approach used in k-ART, which creates a catalog based on reference geometries at the initial minimum and saddle point and uses the knowledge of current geometry at the minimum and the correspondence between the reference and the current topologies at the minimum to reconstruct a new saddle point.

1. Atoms of the current minimum configuration are matched with the reference minimum using the correspondence between the two graphs drawn by NAUTY.
2. A reference frame is then constructed, by comparing the relative positions of the various atoms in the two configurations.
3. Atoms in the current minimum are then displaced using this reference frame, according to the difference between the reference saddle point and energy minimum.
4. The energy and forces are evaluated on the reconstructed activated state. If the activation energy difference between the reconstructed and the reference state and the absolute forces in the reconstructed state are below given thresholds, the reconstruction is considered successful ARTn is applied to converge to the nearby saddle. If not, then the possibility that many correspondences between the

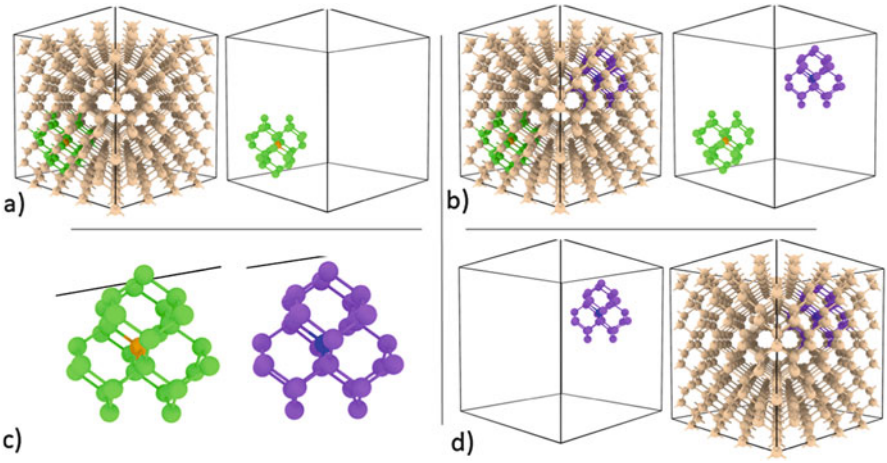

Fig. 4 Reconstruction of a saddle point from the topological classification. (**a**) ART nouveau generates an event from the initial topology of a central atom (red sphere). (**b**) Once ART nouveau search is done, k-ART reconstructs each event associated with the same initial topology: topological correspondence of two central atoms (blue and red spheres). (**c**) Applying geometrical symmetry is needed to the initial structure of the second central atom (blue sphere) as well as the saddle point and final structures. (**d**) Finally, guess structures (i.e., saddle point and final applied of the second central atom) are then refined using ART nouveau to include elastic deformations

reference and current graphs exist is considered and a number of permutations in the correspondence labeling are attempted, based on various reconstructed reference frames.

5. When no attempted reconstruction is accepted, k-ART concludes that the correspondence between topology and geometry for this environment is not unique. The cutoff for defining edges between atoms is then modified, and the topology is split. If a new topology is found in the way, it is then populated by events, if a known topology is generated, the algorithm starts over for generating a specific event (Fig. 4).

The success in reconstruction depends closely on the graph. For compact systems, such as bulk metals or semiconductors, a graph containing between 50 and 70 atoms, with edges drawn between nearest neighbor atoms, is sufficient. For an anisotropic system such as graphene, however, it is necessary to define long-enough cutoff to ensure that planes are linked. Reconstruction also requires that a specific saddle point related to the generic one exist. For high-enough barriers, typically above 0.1 eV, a deformation strong enough to make such a barrier disappear would lead to a change in topology. For low barriers, particularly surrounding unstable points, this is not always the case. To avoid this problem, one can systematically recreate the event catalog associated with a given topology when such a low-energy barrier disappears, either automatically cross or simply ignore them.

5 Confidence in the Completeness of the Saddle Point Catalog

An important parameter for off-lattice KMC is determining how many saddle point searches must be done. Too few searches will lead to incorrect kinetics, but searches can be computationally wasteful if all important transitions are already known. Some methods – such as k-ART and SEAK-MC – perform a finite number of searches for each defect that is encountered. In the case of k-ART, the event catalog is reused for all know topologies; new searches are performed to complete the catalog or to complete it, if a new topology is added. New searches are also performed to improve the event catalog, for very common topologies (searchers are added every log-10 times a topology is found). In SEAK-MC, the event catalog for a given defect is destroyed after the execution of one of those events and is rebuilt from scratch every time a defect is encountered.

When using minimum-mode following searches to find saddles, the completeness of the KMC catalog can be estimated from the statistics of how often new saddles are found. For example, if an increasing number of searches are done without finding any new events, there should be a growing confidence that the rate table is complete. Importantly, however, there can be a wide range in the frequency at which different saddles are found, and this bias can lead to unquantifiable uncertainties in the completeness of the catalog (Henkelman and Jónsson 1999).

Another way to estimate the completeness of the event table is to use MD saddle searches, as discussed previously. The key difference between MD searches and minimum-mode following searches is that the probability of finding any event using MD is proportional to its rate. Since we are interested in rare events, basin-constrained MD at an artificially high temperature is used to determine possible reaction mechanisms, and harmonic TST is used to calculate the rate of the events at the low temperature of interest. To the extent that harmonic TST holds, one can calculate a well-defined estimator for the fraction of the rate table that is missing, based upon the amount of MD time that is used to determine the rate table and the events that have been found (Chill and Henkelman 2014).

It should be noted that while high-temperature MD searches can provide an estimate of the completeness of the catalog, the high-temperature event catalog may miss events for which the rate decreases as temperature increases. Such non-Arrhenius behavior would be due to strong entropic effects. Examples of such behavior include thermal stabilization of nano-voids (Perez et al. 2013) and dislocations (Kim and Tadmor 2014).

6 The Low-Barrier Problem and Coarse Graining

While off-lattice KMC is able to coarse grain over the fast atomic vibration time scale and model the slower time scale of the state-to-state dynamics, it is ubiquitous to have another separation of time scales between the fastest state-to-state events and

slower time scales of interest. This so-called low-barrier problem describes how off-lattice KMC simulations can spend all of the computational time on fast events so that they are not able to reach long time scales.

A key idea that has been used to bridge the gap between fast and slow events is the framework of the Monte Carlo with adsorbing Markov chains (MCAMC), which was described by Novotny in 1995 (Novotny 1995). In this pioneering work, Novotny showed how the kinetics of a system that was partitioned into a set of transient states and adsorbing states could be solved to give a specific time for the transition from a transient state to a specific adsorbing state. Remarkably, this was shown to be possible without adding any approximations beyond the rates that are used in KMC regardless of the partitioning between transient and adsorbing states. This means that in an off-lattice KMC simulation, any set of states can be identified as transient, and the time to transition to a neighboring adsorbing state can be calculated using the MCAMC approach without additional approximations.

The exact calculation of transition times from transient states to adsorbing states requires an iterative set of matrix calculations each with dimension of the size of the total number of states involved. Novotny also showed that it was possible to simplify this calculation and instead calculate the average escape time from the transient states via the first moment of the escape time distribution (Novotny 2001). This approximation was later referred to as the mean-rate method (MRM) by Puchala et al. (2010) since the escape time from the set of transient states is characterized by a single rate. The mean-rate approximation works well when there is a separation of time scales between the transient and the adsorbing states. In this regime, the set of transient states has been referred to as a superbasin. The mean-rate method reduces the computational cost of escaping a superbasin to a single matrix inversion.

The MCAMC approach allows off-lattice KMC simulations to switch, at any point, between a KMC description of the state-to-state dynamics to a transition between any specified state of transient states to the neighboring adsorbing states. When MCAMC is done exactly, any set of transient states can be chosen, and so different definitions of the set of transient states have been proposed. On one extreme, all visited states can be considered transient so that every transition is made to a new state. This strategy was described in the Markov web, proposed by Boulougouris and co-workers (Boulougouris and Frenkel 2005; Boulougouris and Theodorou 2007). While advantageous in terms of being able to visit new states as rapidly as possible, this comes at the cost of losing the state-to-state detail of the trajectory to states that have been visited. Additionally, for simulations which are out of equilibrium and explore new parts of configuration space, including all states in the transient space add unnecessary cost to the matrix operations. Another possibility is to count the number of times that each state has been visited and combine states into a superbasin when the visit-count exceeds a specified threshold (Chill et al. 2014b). It is also possible to consider the energy of the states and the saddles between them (Pedersen et al. 2012) or combine states into a superbasin when they are connected by rates above a specified value, as in the basin-autoconstructing MRM (bac-MRM) (Brommer et al. 2014). However the superbasins are defined, they can grow as more transient states are explored.

7 How to Reduce Computational Costs

The main computational costs of a typical on-the-fly KMC run are associated with the evaluation of forces and energies and the construction of neighbor lists. Neighbor-list costs can be decreased using standard cell and Verlet neighbor list methods. Computational cost reduction, for force evaluations, requires taking advantage of the fact that the reference states, in KMC, are effectively at zero temperature, allowing algorithms to exploit the local nature of activated events. Given the large literature on the neighbor-list cost-reduction techniques, this section will focus on the forces.

A number of studies (Mousseau and Barkema 1998b; Pedersen and Luiser 2014; Xu et al. 2015; Guteŕrez et al. 2016) found that initial deformations applied to search for saddle points should be local in order to successfully lead faster to a diverse set of activated events. No more than a few hundred atoms should initially be involved in the event search, and usually less. As the positions of atoms are optimized while reaching the saddle point, more atoms will be deterministically displaced (typically, a few thousand); an even larger number will be displaced as the system is relaxed into a new minimum. Notably, extended cascade-type concerted events can take place during this phase (Béland and Mousseau 2013). In other words, while relaxation should be performed globally, activation can generally be performed locally to maximize the probability of successfully finding saddle points.

Search algorithms can exploit the locality of activated events (Mousseau and Barkema 1998b). The simplest way is to impose activation volumes. For instance, Xu et al. (2011, 2015) showed that a spherical active volume of 4 lattice parameters is sufficient to capture the activation barrier for vacancy diffusion in BCC iron within 0.001 eV, while an active volume of 6 lattice parameters was sufficient for the case of the dumbbell interstitial, and an active volume of 7.5 Å could capture the kinetics of 37-interstitial clusters (Xu et al. 2013). In other words, saddle searches can yield accurate results by calculating forces and energies over a subset of a few hundred atoms. It is also possible to find saddle points using small active volumes and then refine them using increasingly larger ones.

Imposing strict active volumes does have drawbacks. For instance, such an approach imposes static long-range elastic interactions, preventing a dynamical reaction of the whole system. To avoid this bias, it is possible to dynamically update a list of active atoms during the event search (Béland et al. 2011; Joly et al. 2012) defined by those atoms with a minimum force acting upon them and their neighbors. As the search progresses, an increasing number of atoms will have non-negligible forces. This procedure can effectively make computational time scale sublinearly with system size, with no strong assumptions to be made about the size of the active volume involved during event searches (Joly et al. 2012).

Such an approach, while elegant, is difficult to implement with standard force field libraries such as LAMMPS (Plimpton 1995). In this case, one turns to a partial use of active volume. After a global energy minimization of the whole structure, to ensure that all elastic effects are fully incorporated in the local minimum, a sphere

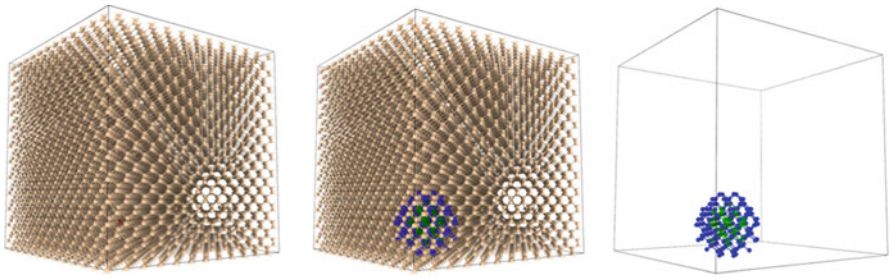

Fig. 5 Graphical description of the local force calculations. Two spheres are centered on the same atom associated with a given topology. In the inner sphere (green atoms), all atoms can move following the activation. Atoms in the outer sphere (blue atoms) are fixed and contribute to ensuring that the correct forces apply to the inner atoms

is drawn around the central atom in the topology. An inner radius defines the set of atoms that are allowed to move during activation; a second, outer radius includes fixed atoms needed to ensure that the forces are accurate in the inner part. As a general rule, the inner sphere contains between 500 and 2000 atoms, while the outer shell is determined by the force field cutoff radius (Fig. 5) (Trochet et al. 2017). As with large enough active volumes, the error on the energy is small, and generic events are generated without further global relaxation at the saddle point nor the final minimum. Specific events, for their part, see by default the barrier fully relaxed at the saddle point after a first local relaxation, with the final configuration also relaxed globally. For systems with simple localized defects, it is possible to avoid a global relaxation at the saddle point for specific events, provided that the active volume is large enough, reducing the global relaxation only to accepted minimum states. The parameters to ensure a given error threshold must then be assessed specifically for each system. As a general rule, however, the use of local forces on systems containing 10,000 to 1 million atoms makes the event search almost order 1, considerably accelerating the algorithm without any significant loss in information or precision (Raine et al. 2017).

For the simulation of large systems, parallelization is crucial. An advantage of MC is that every event search can be performed independently, which means that many computer cores can simultaneously generate events to be added to the catalog. As mentioned in Sect. 7, one can accelerate the saddle searches by exploiting the locality of events. By increasing the number of "worker" computer cores proportionally to the number of saddle points to be generated, and by considering a local subset of atoms during saddle searches, the runtime between KMC steps will be approximatively independent of system size, if event management and KMC overhead costs are negligible.

8 Advantages of MC: Tricks and Shortcuts Available

Monte Carlo methods present a major advantage over molecular dynamics: it is much easier to use tricks to focus on the problems of interest, since the system's evolution is event-based, instead of being continuous, offering a much better control on the rules that are used. In this section, we review a number of approximations, solutions, and tricks that make these off-lattice approaches more competitive that could be expected through a simplistic evaluation of their computational costs.

The Low-Barrier Problem We have already presented the approach for handling processes with low barriers. Such methods are crucial for long-time simulations of complex materials, with many low-barrier events. By solving analytically the in-basin kinetics, it is possible to effectively perform billions of jumps without having to compute them directly, allowing the simulation to reach the relevant time scale instead of being slowed down to a halt by irrelevant mechanisms.

Constant prefactor The use of constant prefactor is not essential, and some of us compute the prefactor for every event using the harmonic approximation. Nevertheless, for most compact system, the prefactor varies relatively little compared with the barrier, so it can be given a constant value, leaving the cost of evaluating a barrier to converging to a saddle point (Valiquette and Mousseau 2003). Clearly, however, as shown by Koziatek et al. (2013), the harmonic approximation fails in systems with significant density fluctuations, and it is generally necessary to, at least, demonstrate the validity of this approximation before using it. Prefactors can be affected by temperature, when barriers are low with respect to $k_B T$, as discussed previously. Overcoming this limitation requires likely to turn to thermodynamical integration, which is computationally heavy and has not yet been automated.

Restricting event searches to specific environments To decrease computational costs, it is possible to prevent off-lattice KMC calculations from searching for events in very stable environments. For example, when we are interested by a phenomena occurring on time scales of vacancy or interstitial diffusion, identifying events from perfectly crystalline environment, with barriers that are many eVs high, does not contribute to the dynamics. In some case, also, we can focus on specific regions, near a defect, for example, so that we may want to ignore the rest of the box, knowing that it will not contribute to the kinetics of interest. Off-lattice KMC calculations can therefore be instructed to ignore all crystalline topologies, specific atomic species, or regions when constructing its catalog.

Biasing pathways While the previous shortcuts do not significantly affect the kinetics of the system on the time scale selected, it is also possible to give up the correct kinetics in exchange for exploring specific pathways. This can be done in many ways. First, a general bias in a given direction can be imposed, either through a selection bias from a complete event catalog or through the construction of an

event catalog that only allows moves along a given pathway, specific mechanisms or a general direction. It is also possible to hand select, at every step, the event of interest among the list and evolve the system along a biased path.

The list of tricks with Monte Carlo approaches is largely limited only by the researcher's imagination. In many ways, it is through these shortcuts that allow a better focus on the important physics that KMC methods are most useful that, in addition to accelerating the simulation, they provide a much clearer picture of the fundamental mechanisms dominating specific processes.

9 Applications

Off-lattice KMC methods have been used to study a wide range of system, including metal on metal diffusion (Henkelman and Jónsson 2001), interstitials and vacancy clusters in c-Si (El-Mellouhi et al. 2008; Trochet et al. 2015), methanol decomposition on Cu, (Xu et al. 2009) Fe (Brommer et al. 2014; Restrepo et al. 2016) and Ni (Mahmoud et al. 2018), ion-implanted relaxation in Si (Béland et al. 2013; Béland and Mousseau 2013; Jay et al. 2017), ion- and neutron-irradiated metals (Béland et al. 2015b; Lu et al. 2016) and alloys (Béland et al. 2016; Lu et al. 2016; Osetsky et al. 2016), grain boundary diffusion in Cu (Pedersen et al. 2009), hydrogen diffusion in Al grain boundaries, (Pedersen and Jónsson 2009) a solid-solid phase transformation in Mo (Duncan et al. 2016) defects in amorphous Si (Joly et al. 2013), Li impurities in Si (Trochet and Mousseau 2017) and C impurities in Fe (Restrepo et al. 2016, 2017), and many more.

We review here a few applications that represent some of the strengths and limitations of these methods.

9.1 Loop Transformation in FeCr

Off-lattice KMC is a powerful tool to simulate the long-time kinetics of point defects and small defect clusters. However, it is not limited to these relatively simple problems and can be used to capture slow kinetics involving extended defects. A good example of such a problem is the transformation of $1/2\langle 111\rangle$ prismatic dislocation loops in Fe into $\langle 100\rangle$ prismatic dislocation loops. $1/2\langle 111\rangle$ and $\langle 100\rangle$ loops in bcc Fe and FeCr are common interstitial-type defects observed in neutron- and ion-irradiated samples. The $1/2\langle 111\rangle$ loops are known to be fast one-dimensional diffusers, which are closely related to void swelling. $\langle 100\rangle$ loops are largely immobile. The ratio of formation of $1/2\langle 111\rangle$ to $\langle 100\rangle$ loops is directly linked to void-swelling rates. However, collision cascades simulations in bulk Fe predict the formation of $1/2\langle 111\rangle$ loops, but not of $\langle 100\rangle$ loops. There must be a post-cascade mechanism that permits the transformation of the post-cascade defects into $\langle 100\rangle$ loops. High-temperature MD simulations – more than 1000 K – of two $1/2\langle 111\rangle$ loops intersecting and transforming into a single loop suggested that this reaction solely leads to one large $1/2\langle 111\rangle$ loop (Terentyev et al. 2008). Using

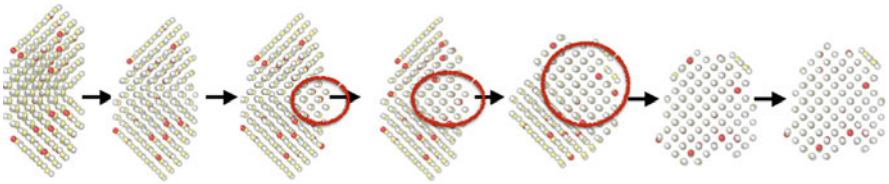

Fig. 6 An illustration of the transformation of two 1/2⟨111⟩ interstitial clusters into a ⟨100⟩ loop in FeCr (10 at. % Cr). The red circles indicate the subregion that is changing orientation within the interstitial cluster. Gray spheres are Fe atoms not sitting on bcc lattice sites, orange spheres are Cr atoms not sitting on bcc lattice sites, and yellow atoms are empty bcc lattice sites. The simulation cell contains 16074 atoms. (The figure is adapted from Béland et al. 2015a)

off-lattice KMC in Fe (Xu et al. 2013) and FeCr (Béland et al. 2015a), a novel mechanism for the transformation of two 1/2⟨111⟩ into a ⟨100⟩ was discovered, which is illustrated in Fig. 6. At 600 K, the waiting time was 200 ns in off-lattice KMC – i.e., a 0.73 eV overall activation barrier. This mechanism was later confirmed by molecular dynamics, as reported in Xu et al. (2013). The simulations in FeCr (10 at. % Cr) also indicated that Cr decoration of the interstitial clusters favor the transformation to a large 1/2⟨111⟩, relative to the situation in pure Fe. This was a good example of a situation where off-lattice KMC was able to predict a mechanism that helped guide further MD simulations.

9.2 Phase Transformation in Mo

Figure 7 illustrates both the success and limitations of the off-lattice KMC method for the simulation of a transition between a complex A15 phase to the lower-energy BCC state in bulk Mo (Duncan et al. 2016). In this calculation, the interatomic interactions are described by an embedded atom method potential (Zhou et al. 2001). The formation and dissolution of complex (topologically close-packed) phases are important both for understanding the hardness and fracture of Ni-based superalloys (Sinha 1972). In Fig. 7a the phase transition from A15 to BCC is observed with AKMC to occur at 300 K over time scales of microseconds. The atoms are colored according to a common-neighbor analysis (Faken and Jónsson 1994) to better visualize the interface (gray) between the A15 (red) and BCC (blue) phases.

The simulation in Fig. 7a involves only a few hundreds of atoms, and yet we can already see a system-size-related problem caused by the disorder at the phase boundary. Specifically, as shown by the disconnectivity graph in Fig. 7b, there are many states connected to the initial state by low barriers which form a superbasin. All of these states must be enumerated in the AKMC simulation, as well as the rates between them. Only then a MCAMC move can be used to find a higher barrier process leading to interface motion, to the final state indicated. An analysis of the superbasin states shows that most transitions within the superbasin involve groups of atoms switching from A15- to BCC-coordinated. The fundamental problem is that

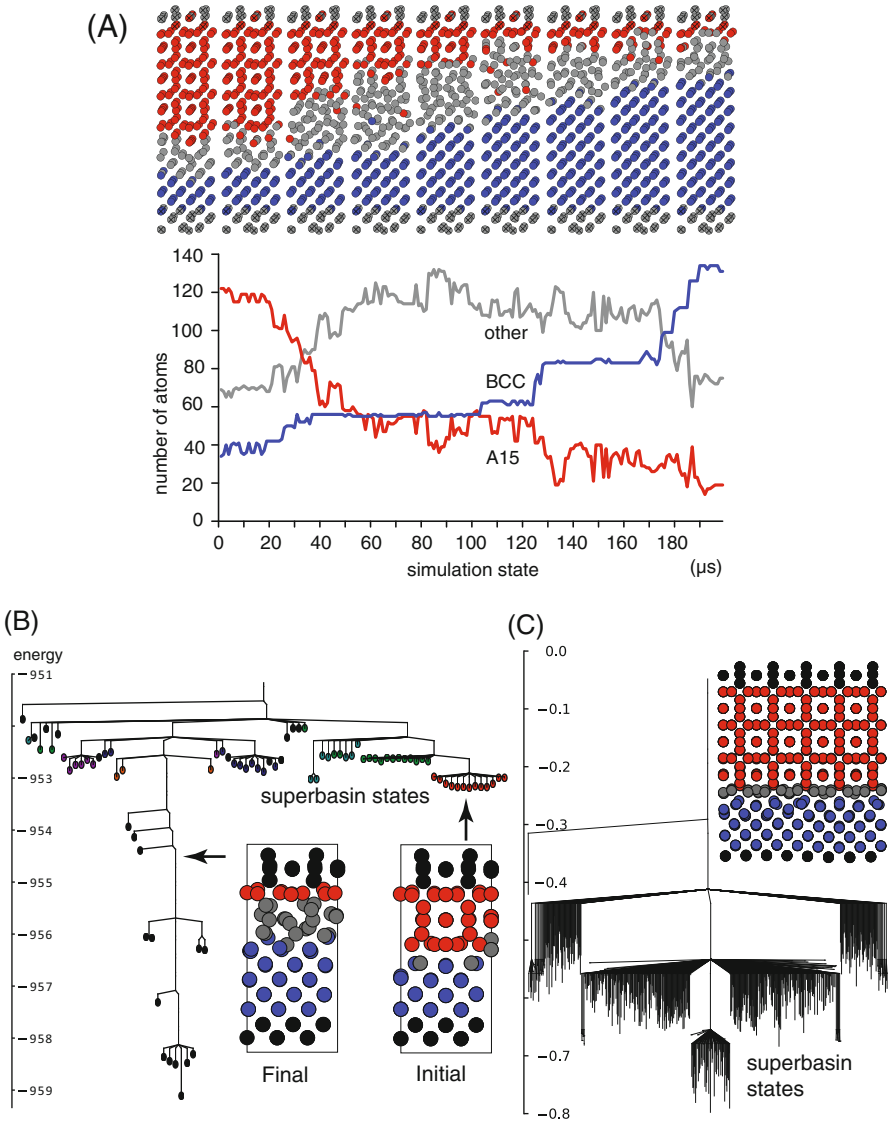

Fig. 7 (**a**) AKMC simulation of a transition from the A15 complex phase in Mo to the lower-energy BCC phase, which occurs on the timescale of microseconds at room temperature. (**b**) A disconnectivity graph showing that disorder at the interface between the phases gives rise to many states connected by low barriers. (**c**) As the system is made larger, the number of states in the superbasins grows combinatorially, making the simulation intractable for simulating even a single-layer transition

when there are several such independent groups of flickering atoms, the total number of states in the superbasin grows combinatorially with the number of groups. This

catastrophic situation is shown clearly in Fig. 7c where the supercell is increased and the total number of states in the superbasin is so large that even a single-layer transition between A15 and BCC cannot be observed with our off-lattice KMC machinery.

10 Discussion

Off-lattice KMC approaches have allowed to study the long-time dynamics of systems that were long off-limits, including systems with many defects, alloys, interfaces, grain boundaries, and even fully disordered materials. These methods have also demonstrated that, even for systems that appeared simple, unexpected mechanisms could play a significant role and that even lattice-based methods would benefit from constructing their event catalog using unbiased search methods.

10.1 Limits

While open-end saddle point search methods such as ART, ART nouveau, and the dimer methods are powerful, we still lack a fundamental theoretical support for establishing the completeness of the generated event catalog. How many event searches should one launch by atom? When can one be certain that all relevant events are found? From experience, we know that, with ART nouveau, lower-energy barrier are found much more often than very high barriers. For a method such as k-ART, this is certainly positive. Moreover, comparison with MD and other methods, when possible, suggest that it recovers all previously identified mechanisms. Yet, it is not possible to assess, even statistically, what the error is with these methods contrary to what can often be obtained by the MD-based accelerated approaches of Voter and colleagues (see discussion above). Theoretical bounds or limits would certainly help greatly to ensure that these simulations do capture the essential physics.

While off-lattice with or without on-the-fly catalog building KMC approaches can reach time scales inaccessible to MD, these methods remain much heavier than lattice-based atomistic KMC and are typically limited to a tens of thousands of KMC steps, not counting, for course, analytically handled flickers. These methods can therefore only be applied to systems where physically relevant mechanisms are largely dominated by local mechanisms.

For example, vacancy-induced solute diffusion in metal, where one must wait for a vacancy to diffuse near a solute to see the atom jump, is much too expensive for these types of simulations. To overcome these limitations, one could solve directly in the equilibrium distribution of vacancies of the system for the time step between configurations with vacancies near a solute or couple off-lattice KMC with lattice-based KMC so that the vacancy diffusion away from solutes can be solved efficiently while the elastic effects and interactions between solutes and points defects are addressed exactly with off-lattice description. Clearly, other possibilities exist, and

work will have to be performed in this field to assess the best way of coupling these scales.

Also, the total number of defects that can be effectively handled by off-lattice KMC is typically limited to a few hundred defects. Increasing the number of defects increases the computational cost in two ways. (1) Events must be assigned to each additional defect. While efficient recycling of events can minimize the associated computational cost, such recycling is not always an option. For example, in concentrated alloys, the number of local configurations grows combinatorially with the number of elements, which limits the usefulness of recycling events. (2) As the number of defects increases, the total rates increases as well, which reduces the waiting time of each KMC step. More KMC steps – i.e., more computational resources – are necessary to reach the same time scales. While synchronous (Martínez et al. 2008) and asynchronous (Shim and Amar 2006) KMC parallelization techniques offer a possible solution to this problem, this is still an unresolved challenge for off-lattice KMC.

Another fundamental limit of on-the-fly KMC is that it probes the potential energy surface. If the problem of interest involves significant entropic contributions to the free energy, exploring the potential energy surface might lead to incorrect predictions. On the other hand, many free energy landscape-based methods suffer from the "curse of dimensionality" (Althorpe et al. 2016); to be effective, these accelerated methods necessitate that a proper – and relatively small – set of reaction coordinates be inputed. In the future, we can imagine on-the-fly KMC being used to find promising reaction coordinates to be used as input for accelerated free energy methods.

10.2 Future Developments

Beyond these questions, a number of other developments should be undertaken to improve the efficiency of off-lattice KMC methods.

Moving beyond master-worker parallelization As larger systems with more defects are simulated, the computational overhead of the master-worker parallelization scheme of current off-lattice KMC codes will become a bottleneck. The way forward may be to switch to a decentralized parallelization scheme. While it entails challenges in regard to cataloging, recycling, and load-balancing, such an approach could significantly increase the scalability of off-lattice KMC.

Recycling basins to handle flickers The absorbing Markov chain algorithm described above – i.e. the superbasins – can be define states either as global configurations – cf. Sect. 6 – or local configurations – (Fichthorn and Lin 2013). The latter can provide a significant acceleration over the former as it can decouple the kinetics of non-interacting defects. This local treatment of superbasins is implemented in a number of the current off-lattice codes. A further acceleration would be to reuse superbasins built previously if the configurations are encountered

again. This would increase the bookkeeping costs but can largely use the same geometrical and topological classification tools that are currently used to recycle events.

Extended defects Off-lattice KMC can handle certain extended defects. For example, it was used to simulate the propagation of a grain boundary in Mo (Duncan et al. 2016) and to simulate the transformation of prismatic dislocation loops in bcc Fe and FeCr (Xu et al. 2013; Béland et al. 2015a). However, the kinetics of dislocation lines and their interactions with other extended defects has not been captured yet by off-lattice KMC. This is an important challenge, since MD cannot simulate such kinetics at strain rates consistent with experiments. However, building an event catalog for a dislocation line – which contains thousands of interconnected possible sites where events may be take place – is a daunting task.

11 Conclusion

The last years have seen considerable developments with respect to accelerated atomistic methods, with access to ever more powerful computers and the introduction of new algorithms. These tools are giving modelers access to questions that could not even be asked in a recent past, increasing their interest for communities that are more turned to applications rather than methodological developments.

This additional interest will be beneficial to the field as it attracts new researchers with original ideas, knowledge, and skills, which will result in accelerated developments. Yet, as we have shown here, in spite of some limitations and questions, even in the current implementations, off-lattice kinetic Monte Carlo methods can deliver new insights for a wide range of problems dominated by activated diffusion.

The codes are available, and there is no reason today not to try them!

12 Code Availability

Various version of ART nouveau are available at http://normandmousseau.com. The kinetic ART package can be obtained freely by writing to MT or NM. The EON code is available at http://henkelmanlab.org/eon/.

Acknowledgements This work was supported in part by a grant from the Natural Science and Engineering Research Council of Canada. MT and NM are grateful to Calcul Québec and Compute Canada for providing extensive computer time and computer access. The work in Austin was supported by the National Science Foundation (CHE-0645497, CHE-1152342, and CHE-1534177) and the Welch Foundation (F-1841). Sustained computational resources have been provided by the Texas Advanced Computing Center.

References

Alexander KC, Schuh CA (2016) Towards the reliable calculation of residence time for off-lattice kinetic Monte Carlo simulations. Model Simul Mater Sci Eng 24(6):65014. http://stacks.iop.org/0965-0393/24/i=6/a=065014?key=crossref.38e788234d74209ed2f8ad8b6b21fa51, https://doi.org/10.1088/0965-0393/24/6/065014

Althorpe S, Angulo G, Astumian RD, Beniwal V, Bolhuis PG, Brandão J, Ellis J, Fang W, Glowacki DR, Hammes-Schiffer S et al (2016) Application to large systems: general discussion. Faraday Discuss 195:671–698

Barkema GT, Mousseau N (1996) Event-based relaxation of continuous disordered systems. Phys Rev Lett 77(21):4358–4361

Béland LK, Mousseau N (2013) Long-time relaxation of ion-bombarded silicon studied with the kinetic activation-relaxation technique: microscopic description of slow aging in a disordered system. Phys Rev B 88(21):214201

Béland LK, Brommer P, El-Mellouhi F, Joly JF, Mousseau N (2011) Kinetic activation-relaxation technique. Phys Rev E 84(4):046704. https://doi.org/10.1103/PhysRevE.84.046704

Béland LK, Anahory Y, Smeets D, Guihard M, Brommer P, Joly JFF, Pothier JcC, Lewis LJ, Mousseau N, Schiettekatte F, Postale C, Centre-ville S (2013) Replenish and relax: explaining logarithmic annealing in ion-implanted c-Si. Phys Rev Lett 111(10):1–5. https://doi.org/10.1103/PhysRevLett.111.105502, http://arxiv.org/abs/1304.2991

Béland LK, Osetsky YN, Stoller RE, Xu H (2015a) Interstitial loop transformations in FeCr. J Alloys Compd 640:219–225

Béland LK, Osetsky YN, Stoller RE, Xu H (2015b) Slow relaxation of cascade-induced defects in Fe. Phys Rev B 91(5):054108

Béland LK, Samolyuk GD, Stoller RE (2016) Differences in the accumulation of ion-beam damage in Ni and NiFe explained by atomistic simulations. J Alloys Compd 662:415–420

Boulougouris GC, Frenkel D (2005) Monte Carlo sampling of a Markov web. J Chem Theory Comput 1:389–393

Boulougouris GC, Theodorou DN (2007) Dynamical integration of a Markovian web: a first passage time approach. J Chem Phys 127:084903

Brommer P, Béland LK, Joly JF, Mousseau N (2014) Understanding long-time vacancy aggregation in iron: a kinetic activation-relaxation technique study. Phys Rev B 90(13):1–9 https://doi.org/10.1103/PhysRevB.90.134109

Chill ST, Henkelman G (2014) Molecular dynamics saddle search adaptive kinetic Monte Carlo. J Chem Phys 140:214110

Chill ST, Stevenson J, Ruhle V, Shang C, Xiao P, Farrell J, Wales D, Henkelman G (2014a) Benchmarks for characterization of minima, transition states and pathways in atomic systems. J Chem Theory Comput 10:5476–5482

Chill ST, Welborn M, Terrell R, Zhang L, Berthet JC, Pedersen A, Jónsson H, Henkelman G (2014b) Eon: software for long time scale simulations of atomic scale systems. Model Simul Mater Sci Eng 22:055002

Duncan J, Harjunmaa A, Terrell R, Drautz R, Henkelman G, Rogal J (2016) Collective atomic displacements during complex phase boundary migration in solid-solid phase transformations. Phys Rev Lett 116(3):035701

El-Mellouhi F, Mousseau N, Lewis L (2008) Kinetic activation-relaxation technique: an off-lattice self-learning kinetic Monte Carlo algorithm. Phys Rev B 78(15):153202. https://doi.org/10.1103/PhysRevB.78.153202

Faken D, Jónsson H (1994) Systematic analysis of local atomic structure combined with 3D computer graphics. Comput Mater Sci 2:279–286

Fichthorn KA, Lin Y (2013) A local superbasin kinetic Monte Carlo method. J Chem Phys 138(16):164104

Gutefrez M, Argaéz C, Jónsson H (2016) Improved minimum mode following method for finding first order saddle points. J Chem Theory Comput 13:125–134

Henkelman G, Jónsson H (1999) A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. J Chem Phys 111:7010–7022

Henkelman G, Jónsson H (2000) Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. J Chem Phys 113:9978–9985

Henkelman G, Jónsson H (2001) Long time scale kinetic Monte Carlo simulations without lattice approximation and predefined event table. J Chem Phys 115(21):9657–9666. https://doi.org/10.1063/1.1415500, http://link.aip.org/link/JCPSA6/v115/i21/p9657/s1&Agg=doi

Henkelman G, Uberuaga BP, Jónsson H (2000) A climbing image nudged elastic band method for finding saddle points and minimum energy paths. J Chem Phys 113:9901–9904

Jay A, Raine M, Richard N, Mousseau N, Goiffon V, Hemeryck A, Magnan P (2017) Simulation of single particle displacement damage in silicon part II: generation and long time relaxation of damage structure. IEEE Trans Nucl Sci 64(1):141–148. https://doi.org/10.1109/TNS.2016.2628089, http://ieeexplore.ieee.org/document/7742370/

Joly JF, Béland LK, Brommer P, El-Mellouhi F, Mousseau N (2012) Optimization of the kinetic activation-relaxation technique, an off-lattice and self-learning kinetic Monte-Carlo method. J Phys Conf Ser 341:012007. https://doi.org/10.1088/1742-6596/341/1/012007, http://stacks.iop.org/1742-6596/341/i=1/a=012007?key=crossref.dfb01ebf3ff94111aa93a5794b3384f8

Joly JF, Béland LK, Brommer P, Mousseau N (2013) Contribution of vacancies to relaxation in amorphous materials: a kinetic activation-relaxation technique study. Phys Rev B 87(14):144204. http://link.aps.org/doi/10.1103/PhysRevB.87.144204

Jónsson H, Mills G, Jacobsen KW (1998) Nudged elastic band method for finding minimum energy paths of transitions. In: Berne BJ, Ciccotti G, Coker DF (eds) Classical and quantum dynamics in condensed phase simulations. World Scientific, Singapore, pp 385–404

Kim WK, Tadmor EB (2014) Entropically stabilized dislocations. Phys Rev Lett 112(10):105501

Koziatek P, Barrat JL, Derlet P, Rodney D (2013) Inverse Meyer-Neldel behavior for activated processes in model glasses. Phys Rev B 87:224105. https://doi.org/10.1103/PhysRevB.87.224105

Lu C, Jin K, Béland LK, Zhang F, Yang T, Qiao L, Zhang Y, Bei H, Christen HM, Stoller RE et al (2016) Direct observation of defect range and evolution in ion-irradiated single crystalline Ni and Ni binary alloys. Sci Rep 6:19994

Plimpton S (1995) Fast Parallel Algorithms for Short-Range Molecular Dynamics, J Comp Phys, 117, 1–19

Machado-Charry E, Béland LK, Caliste D, Genovese L, Deutsch T, Mousseau N, Pochet P (2011) Optimized energy landscape exploration using the ab initio based activation-relaxation technique. J Chem Phys 135(3):034102. https://doi.org/10.1063/1.3609924, http://www.ncbi.nlm.nih.gov/pubmed/21786982

Mahmoud S, Trochet M, Restrepo OA, Mousseau N (2018) Study of point defects diffusion in nickel using kinetic activation-relaxation technique. Acta Mater 144:679–690. https://doi.org/10.1016/j.actamat.2017.11.021, http://www.sciencedirect.com/science/article/pii/S1359645417309643

Malek R, Mousseau N (2000) Dynamics of Lennard-Jones clusters: a characterization of the activation-relaxation technique. Phys Rev E 62(6):7723–7728. https://doi.org/10.1103/PhysRevE.62.7723

Marinica MC, Willaime F, Mousseau N (2011) Energy landscape of small clusters of self-interstitial dumbbells in iron. Phys Rev B 83(9):094119. https://doi.org/10.1103/PhysRevB.83.094119

Martínez E, Marian J, Kalos MH, Perlado JM (2008) Synchronous parallel kinetic Monte Carlo for continuum diffusion-reaction systems. J Comput Phys 227(8):3804–3823

McKay BD, Piperno A (2014) Practical graph isomorphism, II. J Symb Comput 60:94–112. https://doi.org/10.1016/j.jsc.2013.09.003, http://www.sciencedirect.com/science/article/pii/S0747717113001193

McKay BD et al (1981) Practical graph isomorphism. Congr Numer 30:45–87

Mousseau N, Barkema GT (1998b) Traveling through potential energy landscapes of disordered materials: the activation-relaxation technique. Phys Rev E 57:2419–2424. https://doi.org/10.1103/PhysRevE.57.2419

Munro LJ, Wales DJ (1999) Defect migration in crystalline silicon. Phys Rev B 59:3969
Nocedal J (1980) Updating quasi-Newton matrices with limited storage. Math Comput 35:773–782
Novotny MA (1995) Monte Carlo algorithms with absorbing Markov chains: fast local algorithms for slow dynamics. Phys Rev Lett 74:1–5
Novotny MA (2001) A tutorial on advanced dynamic monte carlo methods for systems with discrete state spaces. In: Stauffer D (ed) Annual reviews of computational physics IX. Wold Scientific, Singapore, pp 153–210
Ojifinni RA, Froemming NS, Gong J, Pan M, Kim TS, White J, Henkelman G, Mullins CB (2008) Water-enhanced low-temperature CO oxidation and isotope effects on atomic oxygen-covered Au (111). J Am Chem Soc 130(21):6801–6812
Osetsky YN, Béland LK, Stoller RE (2016) Specific features of defect and mass transport in concentrated FCC alloys. Acta Mater 115:364–371
Pedersen A, Jónsson H (2009) Simulations of hydrogen diffusion at grain boundaries in aluminum. Acta Mater 57:4036–4045
Pedersen A, Luiser M (2014) Bowl breakout: escaping the positive region when searching for saddle points. J Chem Phys 141(2):024109
Pedersen A, Henkelman G, Schiøtz J, Jónsson H (2009) Long time scale simulation of a grain boundary in copper. New J Phys 11:073034
Pedersen A, Berthet JC, Jónsson H (2012) Simulated annealing with coarse graining and distributed computing. Lect Notes Comput Sci 7134:34–44
Perez D, Luo SN, Voter AF, Germann TC (2013) Entropic stabilization of nanoscale voids in materials under tension. Phys Rev Lett 110(20):206001
Puchala B, Falk ML, Garikipati K (2010) An energy basin finding algorithm for kinetic Monte Carlo acceleration. J Chem Phys 132(13):134104. https://doi.org/10.1063/1.3369627, http://www.ncbi.nlm.nih.gov/pubmed/20387918
Raine M, Jay A, Richard N, Goiffon V, Girard S, Member S, Gaillardin M, Paillet P, Member S (2017) Simulation of single particle displacement damage in silicon part I: global approach and primary interaction simulation. IEEE Trans Nucl Sci 64(1):133–140. https://doi.org/10.1109/TNS.2016.2615133, http://ieeexplore.ieee.org/document/7582531/
Restrepo OA, Mousseau N, El-Mellouhi F, Bouhali O, Trochet M, Becquart CS (2016) Diffusion properties of Fe-C systems studied by using kinetic activation-relaxation technique. Comput Mater Sci 112:96–106. https://doi.org/10.1016/j.commatsci.2015.10.017, http://www.sciencedirect.com/science/article/pii/S0927025615006643, http://linkinghub.elsevier.com/retrieve/pii/S0927025615006643
Restrepo OA, Becquart CS, El-Mellouhi F, Bouhali O, Mousseau N (2017) Diffusion mechanisms of C in 100, 110 and 111 Fe surfaces studied using kinetic activation-relaxation technique. Acta Mater 136:303–314. https://doi.org/10.1016/j.actamat.2017.07.009, http://www.sciencedirect.com/science/article/pii/S135964541730558X
Shim Y, Amar JG (2006) Hybrid asynchronous algorithm for parallel kinetic Monte Carlo simulations of thin film growth. J Comput Phys 212(1):305–317
Sinha AK (1972) Topologically close-packed structures of transition metal alloys. Prog Mat Sci 15:81
Sørensen MR, Voter AF (2000) Temperature-accelerated dynamics for simulation of infrequent events. J Chem Phys 112:9599–9606
Sørensen MR, Jacobsen KW, Jónsson H (1996) Thermal diffusion processes in metal-tip-surface interactions: contact formation and adatom mobility. Phys Rev Lett 77:5067–5070
Terentyev D, Malerba L, Klaver P, Olsson P (2008) Formation of stable sessile interstitial complexes in reactions between glissile dislocation loops in BCC Fe. J Nucl Mater 382(2):126–133
Terrell R, Welborn M, Chill ST, Henkelman G (2012) Database of atomistic reaction mechanisms with application to kinetic Monte Carlo. J Chem Phys 137:014105
Trochet M, Mousseau N (2017) Energy landscape and diffusion kinetics of lithiated silicon: a kinetic activation-relaxation technique study. Phys Rev B 96(13):134118. https://doi.org/10.1103/PhysRevB.96.134118

Trochet M, Béland LK, Joly JF, Brommer P, Mousseau N (2015) Diffusion of point defects in crystalline silicon using the kinetic activation-relaxation technique method. Phys Rev B 91(22):224106. https://doi.org/10.1103/PhysRevB.91.224106

Trochet M, Sauvé-Lacoursière A, Mousseau N (2017) Algorithmic developments of the kinetic activation-relaxation technique: accessing long-time kinetics of larger and more complex systems. J Chem Phys 147(15):152712. https://doi.org/10.1063/1.4995426

Trushin O, Karim A, Kara A, Rahman TS (2005) Self-learning kinetic Monte Carlo method: application to Cu(111). Phys Rev B 72(11):115401. https://doi.org/10.1103/PhysRevB.72.115401

Valiquette F, Mousseau N (2003) Energy landscape of relaxed amorphous silicon. Phys Rev B 68:125209. https://doi.org/10.1103/PhysRevB.68.125209

Vernon LJ (2010) Modelling the growth of TiO2. Ph.D. thesis, Loughborough University

Vernon LJ (2012) PESTO: potential energy surface tools. https://github.com/louisvernon/pesto

Vernon L, Kenny SD, Smith R, Sanville E (2011) Growth mechanisms for ${\mathbf{TiO}}_{2}$ at its rutile (110) surface. Phys Rev B 83(7):75412. https://doi.org/10.1103/PhysRevB.83.075412

Wales DJ (2002) Discrete path sampling. Mol Phys 100:3285–3305

Xiao P, Wu Q, Henkelman G (2014) Basin constrained κ-dimer method for saddle point finding. J Chem Phys 141:164111

Xu L, Mei DH, Henkelman G (2009) Adaptive kinetic Monte Carlo simulation of methanol decomposition on Cu(100). J Chem Phys 131:244520

Xu H, Osetsky YN, Stoller RE (2011) Simulating complex atomistic processes: on-the-fly kinetic Monte Carlo scheme with selective active volumes. Phys Rev B 84(13):132103. https://doi.org/10.1103/PhysRevB.84.132103

Xu H, Stoller RE, Osetsky YN, Terentyev D et al (2013) Solving the puzzle of <100> interstitial loop formation in BCC iron. Phys Rev Lett 110(26):265503

Xu H, Stoller RE, Béland LK, Osetsky YN (2015) Self-evolving atomistic kinetic Monte Carlo simulations of defects in materials. Comput Mater Sci 100:135–143

Zeng Y, Xiao P, Henkelman G (2014) Unification of algorithms for minimum mode optimization. J Chem Phys 140:044115

Zhou XW, Wadley HNG, Johnson RA, Larson DJ, Tabat N, Cerezo A, Petford-Long AK, Smith GDW, Clifton PH, Martens RL, Kelly TF (2001) Atomic scale structure of sputtered metal multilayers. Acta Mater 49:4005–4015