

## Following atomistic kinetics on experimental timescales with the kinetic Activation–Relaxation Technique



Normand Mousseau<sup>a,b,\*</sup>, Laurent Karim Béland<sup>e</sup>, Peter Brommer<sup>c</sup>, Fedwa El-Mellouhi<sup>f,g</sup>, Jean-François Joly<sup>d</sup>, Gawonou Kokou N'Tsouaglo<sup>a</sup>, Oscar Restrepo<sup>a,g</sup>, Mickaël Trochet<sup>a</sup>

<sup>a</sup> Département de physique, Université de Montréal, Case postale 6128, succursale centre-ville, Montréal, QC H3C 3J7, Canada

<sup>b</sup> Theoretical Physics, Utrecht University, Utrecht, The Netherlands

<sup>c</sup> Centre for Predictive Modelling, School of Engineering, University of Warwick, Library Road, Coventry CV4 7AL, UK

<sup>d</sup> Carleton University, Ottawa, ON, Canada

<sup>e</sup> Materials Science and Technology Division, Oak Ridge National Laboratory, TN, USA

<sup>f</sup> Qatar Environment and Energy Research Institute, Doha, Qatar

<sup>g</sup> Texas A&M University at Qatar, Doha, Qatar

### ARTICLE INFO

#### Article history:

Received 17 September 2014

Received in revised form 18 November 2014

Accepted 21 November 2014

Available online 24 December 2014

#### Keywords:

Defects

Diffusion

Activated dynamics

Kinetic Monte Carlo

Self assembly

### ABSTRACT

The properties of materials, even at the atomic level, evolve on macroscopic time scales. Following this evolution through simulation has been a challenge for many years. For lattice-based activated diffusion, kinetic Monte Carlo has turned out to be an almost perfect solution. Various accelerated molecular dynamical schemes, for their part, have allowed the study on long time scale of relatively simple systems. There is still a need, however, for methods able to handle complex materials such as alloys and disordered systems. Here, we review the kinetic Activation–Relaxation Technique (k-ART), one of a handful of off-lattice kinetic Monte Carlo methods, with on-the-fly cataloging, that have been proposed in the last few years.

© 2014 Elsevier B.V. All rights reserved.

## 1. Introduction

Computational materials science covers a wide range of time and length scales: from the electronic motion to crack propagation and aging. Among the challenges that this discipline faces is the need to cover multiple scales in a self-coherent fashion in order to relate with experiments and real-life phenomena. In spite of considerable efforts that have spanned decades, no complete answer to this challenge exists. Nevertheless, progress takes place and partial solutions are being proposed and applied, not necessarily to generic problems, but at least to specific classes of technologically-relevant materials. Irrespective of these unwelcome restrictions, these developments constantly open up new areas of study to the field.

Here, we are concerned with bridging the time gap between simulations and experiments while preserving a detailed description at the atomic level. This concern is not new. Over the past four decades, a number of groups have proposed methods to achieve this goal [1–7]. In this review, we focus on a recently proposed

algorithm, the kinetic Activation–Relaxation Technique (k-ART) [8,9], a method that bridges the time domain reaching, at experimentally-relevant temperatures, a second and event more. k-ART describes the atomic motions as particles diffuse and affect microscopic and macroscopic properties of complex materials dominated by activated processes. More precisely, k-ART is an off-lattice kinetic Monte Carlo method with on-the-fly cataloging capabilities. First proposed in Ref. [8], it has since been applied to a number of systems – including implanted crystalline silicon, amorphous silicon, crystalline iron and, more recently, alloys, showing a rich versatility and providing important insights by following the atomistic off-lattice motion of these complex systems over a time scale of up to one second and more [9–12]. In this review, we present an overview of k-ART (Section 2), that will allow the reader to gain a general understanding of the method, a review of recent applications (Section 3) and a more detailed description of our implementation (Section 4).

Of course, a number of other accelerated methods have been put forward over the years, many of which are described elsewhere in this special issue, and we briefly compare them to k-ART in Section 5. We will show, in particular, that k-ART is uniquely suited for the simulation of a large number of complex systems where

\* Corresponding author at: Département de physique, Université de Montréal, Case postale 6128, succursale centre-ville, Montréal, QC H3C 3J7, Canada.

E-mail address: [normand.mousseau@umontreal.ca](mailto:normand.mousseau@umontreal.ca) (N. Mousseau).

atoms are not confined to crystalline lattice position and where elastic deformations are important.

## 2. An overview of k-ART

Kinetic ART is a stochastic event-based kinetic algorithm. This means that its time scale is set by the nature of the local energy landscape and not by the integration of the equations of motion as in molecular dynamics. When the right conditions are met, the trajectory produced with this method is as valid and physical as with molecular dynamics. These conditions are well-defined by the transition state theory (TST), developed independently by Eyring [13] and Evans and Polanyi [14] (see Ref. [15] for a historical perspective). Fig. 1 presents the flowchart for this algorithm, which is detailed in the rest of this section.

### 2.1. The fundamentals of kinetic Monte Carlo

Here, we are concerned with the time evolution of a system characterized by relatively deep local metastable states, i.e. states isolated from others by energy barriers that are high with respect to the temperature. In this case, the time evolution is controlled by rare barrier crossings that bring the system from a well-defined local minimum into a new one. This rare-event process ensures that hops from local minimum to local minimum are completely uncorrelated so that the escape event selected at a time  $t_n$  is perfectly independent from that at time  $t_{n-1}$  (see Fig. 2). It also ensures that no two hops take place at exactly the same time, so that events can be uniquely ordered in time. In this case, the dynamical role of phonons – or thermal vibrations – becomes limited to decorrelating the hops and to provide a rare boost through thermal fluctuations. Thermal vibrations allow therefore the system to gain sufficient energy in a given direction and to reach a transition state leading to a new minimum. When these criteria are met, the transition state theory establishes how one can go to compute the escape rates from a local minimum (see, for example, Voter et al. [15]). With these rates at hand, as described below, it is then possible to generate physically-based kinetic trajectories that are fully described by minimum-energy configurations along the way and the escape rates towards neighboring states.

The basic idea behind kinetic Monte Carlo (KMC), introduced by Bortz et al. [16] and first applied to materials by Voter [17], is that when a system is trapped into an energy-minimum separated by

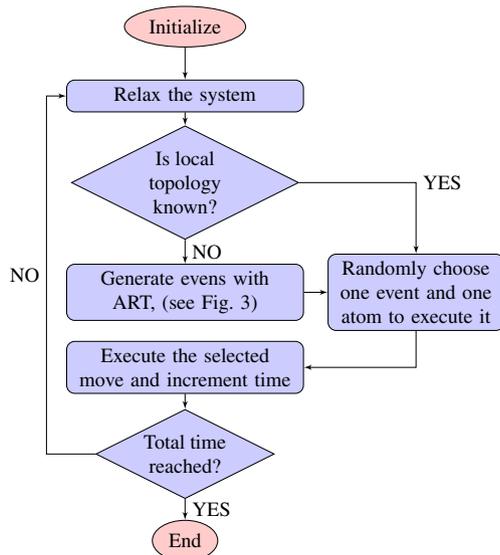


Fig. 1. Flowchart of the implementation of k-ART structure.

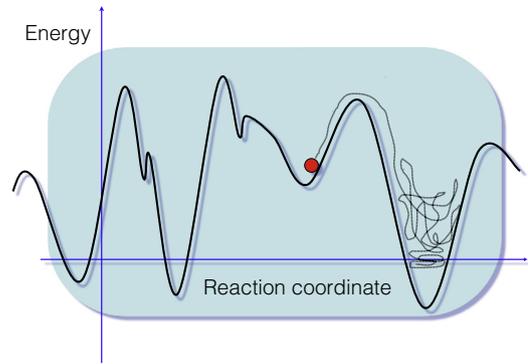


Fig. 2. Schematic representation of the energy landscape surrounding a deep local minimum. The thin black line represents a trajectory through the landscape. Since the barriers leading out of the local minimum in the center are high with respect to the temperature, the system spends a long time in this basin, and has lost memory of its past states by the time it manages to escape. It is therefore possible to fully characterize this state by the minimum-energy configuration, at the bottom of this well, and the curvature around this minimum.

large enough barriers with respect to the temperature, all memory from its previous states is forgotten before it leaves for a new states. This decorrelation is fully justified by the fundamentally chaotic nature of (deterministic) dynamical trajectories (see [18], for example). In this case, it is possible to simply consider that the pathway is constructed of uncorrelated steps each determined at random from a set of events  $\{j\}$  associated to a minimum  $i$ , each defined by its own rate,  $R_{i-j}$ .

The time evolution through those internal stochastic processes, is simply driven by the master equation

$$\frac{\partial P_i(t)}{\partial t} = \sum_j (P_j(t)R_{j-i} - P_i(t)R_{i-j}), \quad (1)$$

which gives us the probability  $P_i$  of finding the system in the minimum  $i$  at time  $t$ . Here,  $j$  runs over all states except  $i$ .

Since the probability that any event takes place is considered random, the escape rate  $r_e^i$  from the local minimum is simply given by the sum of the rates

$$r_e^i = \sum_j R_{i-j}, \quad (2)$$

with the first-passage escape time out of the minimum taken from a Poisson distribution characterized by this escape rate,

$$t_e = -\frac{\ln \mu}{r_e^i}, \quad (3)$$

where  $\mu$  is a random number taken from  $[0, 1]$  at which point an event  $j$  is produced with a probability given by its relative rate:

$$P_{i-j} = \frac{R_{i-j}}{r_e^i}. \quad (4)$$

Algorithmically, KMC consists therefore only of three steps:

1. From a given configuration, identify all escape events and their associated rate.
2. Compute the first-passage escape time (Eq. (3)).
3. Select an event at random with the proper probability (Eq. (4)) and move the system accordingly.

Those simple three steps allow us to solve equation (Eq. (1)) through a stochastic approach. The elegance of this method lies in its simplicity but also in the fact that, contrary to most Monte Carlo algorithms, selected events are always accepted.

For lattice-based problems with short-range interactions, the implementation of KMC is straightforward as rates are readily defined at the root of the model itself.

The application of KMC to materials is more difficult, however, mostly because of the difficulty associated with identifying events and computing the associated rates. k-ART addresses this issue at the expense of heavier computational load compared to standard KMC, still offering considerable speed-ups compared to standard molecular dynamics (MD) which remains the gold standard due to its versatility and simplicity.

In the next two subsections, we describe the k-ART approach to generating and classifying events around each minimum, providing the basic information needed for KMC. In the last paragraph, we give a short overview of a few other aspects of the algorithm that are necessary to make the code efficient.

2.2. An off-lattice KMC approach with on-the-fly catalog building capabilities

Kinetic ART lifts three restrictions generally associated with standard KMC implementation to materials. First, it introduces an efficient approach for on-the-fly updating of the event catalog. Second, it provides a route to regrouping and classifying similar off-lattice events without limiting the set of possible activated mechanisms. Third, it ensures that short- and long-range elastic deformations are fully taken into account in the barrier height.

The first restriction is lifted with the use of an open-ended algorithm for finding transition with knowledge of only one end-point (local minimum). A number of algorithms have been proposed in the past to accomplish this task, many of them being very similar in their spirit and their efficiency [19–21]. Here, we use ART *nouveau* [22–24], an efficient Lanczos-based Activation–Relaxation Technique that was developed and improved over the past 20 years in our group (see, for a comparison with other methods, Refs. [24,28] as well as a few papers in this issue).

The cataloging of off-lattice events is a crucial step in the development of k-ART. Achieving an optimal balance between a sufficient catalog for off-lattice events and sufficiently long timescales was the essence of the algorithmic developments in k-ART. For this, it is necessary to develop a simplified scheme that reduces the complexity of off-lattice geometries: in k-ART, the classification provides a discrete list of approximate positions for transition states available from a given local minimum. The exact rate must be obtained after full reconstruction of the transition state and precise convergence to the specific saddle point. This last step ensures, in addition, that short- and long-range elastic deformations are fully accounted for.

More precisely, k-ART poses that similar geometries will have the same set of events, which will differ in barrier height in local rearrangements. It is therefore possible to assign a set of generic events to each family of geometries and, for each configurations, reconstruct from those the specific events.

Geometries, especially in complex materials with off-lattice positions, are difficult to discretize. k-ART uses the fact that, for most systems dominated by activated dynamics, there is a one-to-one correspondence between local topology and local geometry, within the limits discussed below. Topology being in essence discrete, it is straightforward – once this transformation is adopted – to move off-lattice with the same ease as standard KMC algorithms.

Details matter, of course. In the rest of this section, we review the general ideas behind k-ART. Specific implementation details and options can be found in Section 4.

2.3. Getting the events: the Activation–Relaxation Technique (ART *nouveau*)

Except in very simple systems, it is difficult to identify by inspection all possible pathways connecting a local minimum with its neighbors. Even for the diffusion of a metallic atom on a (100)

surface, where hops between neighboring sites appear as the most efficient diffusion pathways, Kellogg and Feibelman showed, 25 years ago, that an exchange-mechanism where the surface atom moves into the bulk, pushing the neighboring atom out on the surface could also be energetically favorable, multiplying considerably the number of pathways that must be taken into accounts for the kinetics (see Fig. 4) [25]. A decade later, Henkelman and Jónsson showed that the exchange moves could lead to much further effective surface displacements with a similar energy barrier [1].

In general, therefore, it is preferable to use open-ended search methods for generating the list of possible diffusion events surrounding a local minimum on a high-dimensional energy landscape. Over the years, a number of such methods have been proposed [19,22,26,27], most very similar in spirit but with differing implementation details [28].

The open-ended method implemented here, ART *nouveau* [23,24,29], generates events through three fundamental steps: starting from a local minimum, first leave the harmonic basin, second converge to saddle point, and third relax to a new local minimum stage. In more detail, the ART algorithm consists of the following steps (see also Fig. 3):

1. Consider  $V(q)$  to be the potential energy of a system of  $N$  particles (with  $q = \{\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N\}$  being a 3N-dimensional vector) and its 3N-dimensional force  $F(q) = -dV(q)/dq$ . The system is first brought into a local-energy minimum configuration,  $q^n$ , using a standard energy minimization approach (FIRE [30] or conjugate gradient minimization). At the start of ART *nouveau*, therefore,  $F(q^n) = 0$ .
2. We produce an initial random unitary displacement  $u$  and update positions to  $q^n \rightarrow q' = q^n + \delta x_A u$ . This is the initial deformation. A typical value of  $\delta x_A$  is 0.1–0.2 Å.

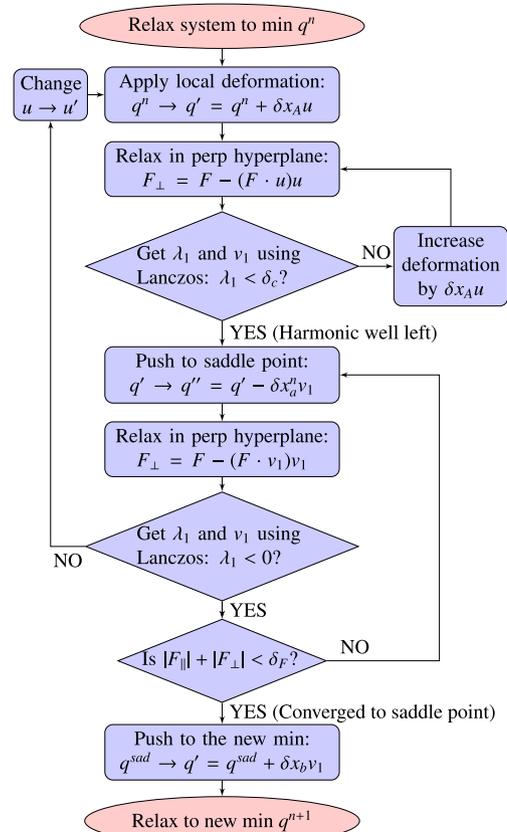
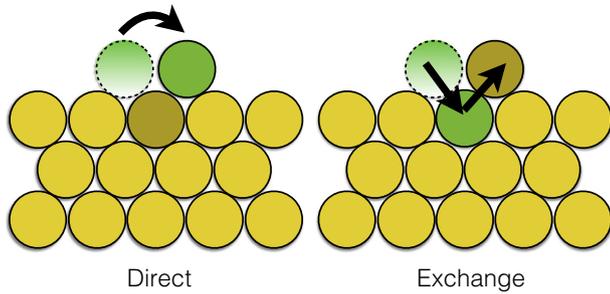


Fig. 3. Flowchart of the implementation of ART *nouveau* at each  $n$  KMC-step to generate events for k-ART.



**Fig. 4.** Diffusion pathways for a Al atom on a (100) surface. (a) Direct mechanism: hop over the surface atoms with an activation barrier of 0.37 eV; (b) Exchange mechanisms: Feibelman's exchange pathway with an activation barrier of 0.23 eV (energies from Ref. [1]).

3. We apply a few minimization steps (typically one to three) in the hyperplane perpendicular to the initial deformation in order to avoid collisions, but not enough to fully relax the system. For this, we calculate forces and separate them into perpendicular and parallel components  $F = F_{\perp} + F_{\parallel}$ . Then, we relax  $q'$  using only the  $F_{\perp}$  component that can be obtained by using the scalar product as  $F_{\perp} = F - (F \cdot u)u$ .
4. At the end of each iteration, we compute the lowest eigenvalue  $\lambda_1$  of the Hessian matrix using the Lanczos algorithm [31]. If  $\lambda_1 > \delta_c$ , a new deformation iteration is done, i.e., the cycle with the previous steps is repeated until  $\lambda_1 < \delta_c$  (where  $\delta_c$  is a negative certain threshold value set to for example  $\delta_c = -1$  eV/Å), to ensure that the system is really outside the harmonic well, or until the maximum number of iterations is reached, in which case we start a new search.
5. The configuration is now out of the harmonic well and convergence to a nearby saddle point is possible. For that, we move the system along the direction of the lowest normalized eigenvector,  $v_1$ , i.e.,  $q' \rightarrow q'' = q' - \delta x_a^n v_1$ , away from the initial minimum. According to the method proposed by Cancès et al. [32] the quantity,  $\delta x_a^n$ , needs to become smaller as the saddle point is approached, therefore, it is set to  $\delta x_a^n = \min\left(2\delta x_a, \frac{|F_{\parallel}|}{\max(|\lambda_1|, 0.5)}\right)$ .
6. As in the previous step (3), after each push along the direction of negative curvature  $q''$  must be relaxed, but now in the hyperplane orthogonal to  $v_1$ . Therefore, we minimize the energy of the system using  $F_{\perp} = F - (F \cdot v_1)v_1$ . This step is repeated until the conditions  $|F_{\perp}| + |F_{\parallel}| < \delta_F$  and  $\lambda_1 < 0$  be met. If  $\lambda_1 > 0$ , indicating that there is no longer a negative eigenvalue, the activation is stopped and a new event search is attempted.
7. When the convergence condition is met, we consider that the system has reached a first-order saddle point,  $q^{sad}$ . As a final step the system must be relaxed to the new local minimum  $q^{n+1}$ . To do that, we push the configuration slightly over the saddle point,  $q^{sad} \rightarrow q' = q^{sad} + \delta x_b v_1$ , and minimize the total energy. Here,  $\delta x_b = \pm 0.1$ , is the fraction of distance (initial-saddle) that the system is pushed over the saddle point (its sign must be chosen according to the dot product,  $(q^n - q^{sad}) \cdot v_1$ , to avoid going back to the previous minimum). The relaxation is finished when the magnitude of the  $3N$ -dimensional force be  $|F| < 10^{-4}$  eV/Å.

A few general points should be noted:

1. Since events taking place in materials at low temperature are local in nature, it is more efficient to apply an initial local random deformation on an atom and its surroundings.

This allows to diversify more rapidly the set of generated events without affecting the nature of those that are found.

2. Since we stop the activation when the lowest eigenvalue becomes positive, we require that this eigenvalue falls below a negative threshold before considering that the system has left its original harmonic basin. This limits the probabilities that the negative eigenvalue vanish in the first iterations of the activation phase.
3. The Lanczos algorithm is a stable algorithm that provides a reliable lowest eigenvalue and corresponding eigenvector with only 12 to 15 force evaluations irrespective of the dimensional of the system studied. As such, Lanczos is particularly efficient for large systems.

Because ART nouveau typically converges to a saddle point within a few hundred force evaluations, it can be used with *ab initio* codes [24,33,34] as well as with empirical potentials [35–38].

ART nouveau has been applied as a standalone technique for sampling energy landscapes as well as for finding minimum-energy pathways of disordered systems, materials and biological systems. For sampling energy landscapes, a large number of searches are launched from each local energy minimum of interest in order to reconstruct the energy landscape structure. Such an approach was used, for example, to study the energy landscape of crystalline materials with defects [31,33,34] as well as that of amorphous silicon and glasses [36,38–42]. Coupled with a Metropolis accept/reject criterion based on the energy-difference between the final and the initial minimum in an event, ART nouveau was also used extensively to generate protein folding and aggregation trajectories [43–47] and identify relaxation pathways in complex materials [35,37,48,49].

#### 2.4. Getting the rates: transition state theory

Now that we know how to generate minimum-energy pathways connecting a local minimum to nearby ones, KMC requires that we evaluate the transition rate  $R_{i \rightarrow j}$  associated with each of these events.

In transition state theory (TST), the escape rate from a basin is defined as the equilibrium flux through a dividing surface between two basins. Basins can be single minima, such as in the systems we study here, or much more complex structures, such as those discussed in Ref. [50], for example.

Defining  $\Delta W$  as the free energy difference between the bottom of the basin and the dividing surface, the TST rate is given by

$$R_{i \rightarrow j} = \kappa \nu e^{-\frac{\Delta W}{k_B T}}, \quad (5)$$

where  $T$  is temperature,  $k_B$  the Boltzmann constant,  $\nu$  the attempt frequency, i.e., the frequency at which the system moves in the direction of the dividing surface and  $\kappa$  is the transmission coefficient, the probability for a system to cross into a new basin once it has reached the dividing surface.

More formally, the attempt frequency is defined as

$$\nu = \sqrt{\frac{k_B T}{2\pi m}} \left[ \int_{\text{well}} e^{-(W(x) - W(x_m))/k_B T} dx \right]^{-1}, \quad (6)$$

with  $W(x)$ , the potential of mean-force,

$$W(x) = \int_{x_m}^x \langle f(\lambda) \rangle_{\lambda=y} dy, \quad (7)$$

representing the averaged force on the reaction coordinate  $\lambda$  measured along the reaction pathway, while the transmission coefficient,

$$\kappa = \langle \Theta[x(t) - x_b] - \Theta[x(-t) - x_b] \rangle, \quad (8)$$

is the thermodynamical average over pathways for times  $t$  much larger than vibrational times ( $\Theta$ , here, is the Heaviside function,

$x_b$  defines the boundary of the local minimum and  $x(t)$ , the position along trajectory).

For simple energy basins, with a single energy minimum, it is possible to simplify the equation. First, we can separate the free energy difference in terms of entropy and configurational energy:

$$R_{i-j} = \Gamma_0 e^{-\Delta E/k_B T}, \quad (9)$$

where  $\Delta E = E_{\text{saddle}} - E_{\text{min}}$ , the energy difference between the saddle point and local minimum and

$$\Gamma_0 = \kappa \nu e^{\Delta S/k_B}, \quad (10)$$

with the entropy difference  $\Delta S$  also computed between the local minimum and the dividing surface. The entropy difference can be computed either through thermodynamic integration or the quasi-harmonic approximation [51]. In bulk materials, however, it is generally considered that the entropy varies relatively little between events [39,52]. Similarly, since saddle points tend to be relatively simple, the transmission coefficient  $\kappa$  is close to 0.5 and the attempt frequency is defined by typical phonon frequencies.

Most KMC simulations, therefore, use a fixed prefactor  $\Gamma_0$  that is set to  $10^{12}$  to  $10^{13}$  Hz [6,8,53]. The rate used here and in most KMC applications is therefore simply:

$$R_{i-j} = \Gamma_0 e^{-\Delta E/k_B T}, \quad (11)$$

with  $\Gamma_0 = 10^{13}$  Hz.

While this approximation, which we follow, is generally reasonable, recent calculations suggest that this might not always be the case, especially in glassy systems [41]. More work remains to be done to establish the generality of this finding in view of other results supporting a fairly constant and harmonic prefactor far away from melting.

### 2.5. Classifying the events: a topological approach

The two previous steps – generating events and computing associated rates – are sufficient for implementing an off-lattice kinetic Monte Carlo approach. This is the algorithm already proposed more than 10 years ago by Henkelman and Jónsson [1], and used for a number of problems since then [6,54].

While the method does work, there is something apparently wasteful by not attempting to reuse the events found earlier for further simulations. When there are only a few relevant barriers, this aspect is not fundamentally an issue. However, for large and complex systems, the burden to start from scratch at every step might become heavy. Especially, since it remains necessary to develop criteria to ensure that degenerate events are not included in the event list. A part of the required classification work is therefore already accomplished, but simply thrown away. Starting from this algorithm, the net computational overhead for cataloging becomes almost negligible once an efficient method is identified.

Distinguishing events is not quite enough, however, to generate a catalog, especially when there are no restrictions on atomic positions. In order to be useful, a catalog must propose reusable events, i.e. events that can be assigned to a discretized space. The discretization can take many forms. Various KMC algorithms have proposed different ways to partially handle off-lattice positions through geometrical analysis typically done with respect to a crystalline order [5,6].

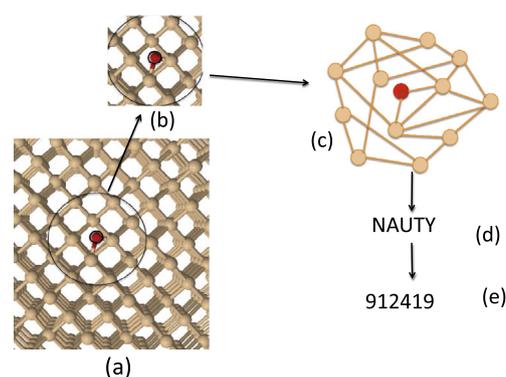
k-ART proposes a unique topological approach to event classification that completely does away with crystalline positions, allowing the algorithm to work as easily with ordered and disordered configurations. This lifts the limitations found in other approaches, which are associated with the use of a crystalline reference state for classifying all events.

The local topological environment surrounding each atom is characterized independently. For each atom, a surrounding region is selected, that typically involves between 40 and 60 atoms, and a graph is constructed by linking nearby atoms, generally using a cut-off set between the first and second-neighbor distance. This classification is used to identify both minima and transition states. It is based on the hypothesis that structures belonging to the same automorphic group share the same list of events, topologically speaking (see Fig. 5).

This hypothesis is equivalent to requesting that (1) the embedding of the local graph in a spatial environment is unique and (2) the set of local minimum geometries mapping to the same automorphic group is directly related to the sets of transition states surrounding this minimum. The first requirement is much less exacting that it appears at first: k-ART is applicable in a regime dominated by activated events and slow dynamics, which implies a temperature well below melting, and defined by specific force-fields. This regime imposes that states are characterized by well-defined metastable local minima. It also means that it is possible to move from one local minimum geometry to any other associated with the same topology by simply deforming the boundary conditions without ever passing by a state that does not belong to this group.

If this is the case, one can expect that the same will hold for the environment surrounding the local minimum including the transition state. With this condition, we obtain a unique correspondence between the topological classification and the geometry, allowing us to construct a valid and efficient catalog.

As it turns out, testing this hypothesis is straightforward: it suffices to ascertain, for each event, whether or not the saddle points predicted by the catalog for a given local minimum can be reconstructed. Since saddle points are, by definition, metastable states, their existence is sufficient to demonstrate the catalog's validity. As discussed in more details below, as a general rule we have found that this correspondence applies to a wide range of systems, both covalent and metallic in bulk, defective, and surface environments. Of course, it would still be possible for new saddle points to arise in a given topology without affecting those already in the catalog. We sometimes observe this in disordered systems with the appearance of very low saddle points (0.01 eV or so), as the geometry is changed without affecting the topology. However, the appearance of a significant saddle point will affect the surrounding energy landscape and move already existing saddlepoints around.



**Fig. 5.** To discretize the configuration k-ART uses a topological description. (a) We select an atom of the configuration and its neighborhood. (b) Bonds are drawn between these atoms resulting in (c) a connectivity graph which is sent (d) through NAUTY, which returns (e) the automorphic class id of the graph as well as the correspondence between its vertices and those of a reference graph. Here, the central atom is a two-coordinated oxygen in a c-Si matrix.

### 3. Selected applications

The full k-ART was applied to a number of systems ranging from crystalline metals to amorphous semiconductors. A simplified version of the algorithm, implemented by Jiang et al. was also recently used to study the stability and mobility of carbon tri-interstitial clusters in cubic SiC [55]. To help better understand the method, we review here a few selected recent applications.

#### 3.1. Di-vacancy diffusion: the importance of elastic effects

Let us first look at an apparently simple problem: the diffusion of a di-vacancy in Stillinger–Weber *c*-Si [56]. We consider here a 510-atom box with periodic boundary conditions. We start from two vacancies positioned at random in the simulation box and relaxed at  $T = 0$  and launched a 5000-step k-ART simulation at 500 K with a prefactor of  $\Gamma_0 = 10^{13}$  Hz.

Since the elastic interaction is important, the aggregation into the divacancy ground state occurs in less than 30 ns, leading to a structure with a formation of  $E_{F2Va} = 4.59$  eV, 0.71 eV below the two isolated vacancies and in agreement with literature [57].

If we consider only the topologies sampled from the ground state (2Va), we identify five characteristic topologies and 19 characteristic events: 2Va to 2Ve, corresponding to vacancy position at the first to fifth nearest-neighbor positions. The top panel of Fig. 6 shows the relative energy, as measured from the ground state, and the occurrence probability. We note, in particular, that 2Va, 2Vb and 2Vc are the most probable states.

The bottom panel shows the transition energies separating the most common of these states. We note, in particular, that the transition state connect 2Va to 2Vb (vacancies in second neighbor position) and 2Vc (in third neighbor position) is degenerate. A closer inspection reveals that the degeneracy is complete: the saddle connecting 2Va to 2Vb is the same, within 0.01 Å and 0.01 eV, to that connecting 2Va to 2Vc. This degenerate state dominates completely the divacancy kinetics, leading to two-rate diffusion: as the divacancies hop between 2Vb and 2Vc, with a barrier of 0.24 eV (0.44 eV from 2Vc to 2Vb), these defects move rapidly across the box to be trapped only when the two vacancies move into the nearest-neighbor position, which is isolated by a 1 eV barrier.

Detailed analysis of vacancy and interstitial Si systems is underway, which supports the conclusion of a recent characterization of interstitial diffusion in Fe that even for self-defects, the richness of the energy landscape is much larger than was expected only a few years ago [31].

Already, the fact that k-ART can automatically identify this complex saddle, which had not been observed previously, demonstrates both the need for automated sampling methods that can go beyond the researcher's intuition into the kinetic pathways as, even for this simple system, unexpected moves take place, and the necessity of including correctly elastic effects, essential for trapping the vacancies.

#### 3.2. Relaxation of ion-bombarded Si

k-ART is not limited to simple materials or small cells. Using local force calculations coupled with global energy relaxation and the topological classification, it is possible to simulate the relaxation of ion-implanted crystal over experimentally time scales.

Ion-implantation of *c*-Si plays an important role in micro-electronic device manufacturing and constitutes a model problem to understand the order–disorder transition. In the low-keV range, initial damage accumulation in such a system is well described by MD, as well as by the binary collision approximation. The initial stages of damage recovery during annealing are well characterized by MD for timescales reaching a few ns. On-lattice simulations of the long-time evolution of the defects produced by ion bombardment, based on the local density of bond-defects (i.e. the IV-pair model [58,59]), accurately predict the stability of amorphous regions in the material. However, there are significant discrepancies between the kinetics predicted by the model and the results of nanocalorimetry experiments. Thus, fully atomistic long-time simulations capable of handling off-lattice positions and elastic effects were necessary to capture the mechanisms leading to damage recovery, which motivated the use of k-ART [11].

Simulation boxes were prepared by running MD calculations of 3 keV self-implantations in *c*-Si at 300 K, using the Stillinger–Weber [60] potential as described in [61]. Three such MDs were performed for 1–10 ns. These MD-generated configurations were used as starting points for k-ART. Three k-ART runs were performed for each MD-generated initial configuration, for a total of nine simulations.

The results of these runs is shown in Fig. 7. The potential energy of the system decreases approximately proportionally to the logarithm of time. The interplay of the point defects and small defect complexes involved in this problem thus led to an ubiquitous relaxation behavior in disordered materials: logarithmic relaxation. These isothermal annealing simulations were shown to be in very good agreement with nanocalorimetric experiments, which suggests the validity of the mechanisms involved in the k-ART simulations.

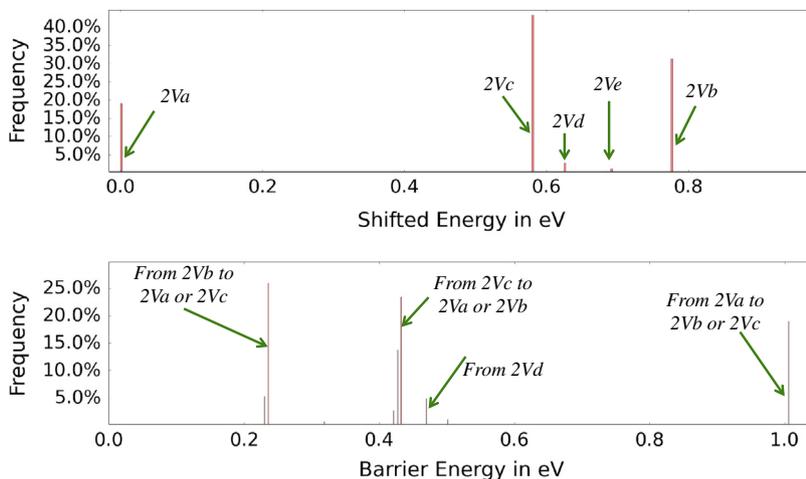
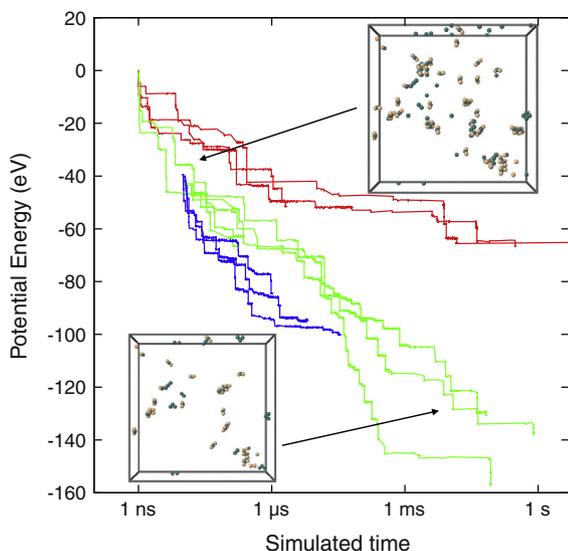


Fig. 6. Histogram of shifted energy and activation barrier found in a 5000-step run at 500 K starting from the divacancy ground state.



**Fig. 7.** Evolution of a 27 000-atom Si box after the implantation of a 3 keV Si atom. After a first 1–10 ns MD simulation, we plot the energy evolution simulated with k-ART for a number of independent runs on three different samples (blue, red and green). Also shown: two snapshots representing the defective atoms at the onset of the simulation and after about 1 s of atomistic simulation. The data was first report in Ref. [11]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Damage recovery was associated to the aggregation, reconfiguration and annihilation of self-interstitial atoms and vacancies [11]. A very broad distribution of activation barriers were executed during each time-frame, with no correlation between forward and backward barriers. Relaxation was rate-limited by the presence of high-energy barriers in all time-frames, which usually did not directly lead to damage recovery, but rather unlocked the system, which would otherwise have been trapped in a region of the potential energy landscape with no accessible relaxation events. This novel description of relaxation, made possible only with the help of second-long simulations highly defective systems, dubbed *replenish and relax*, was found to apply also, for example, to fully disordered amorphous materials.

### 3.3. Iron: application to metallic environment

k-ART was also applied to study vacancy aggregation in Fe [10]. To avoid duplication, we refer the reader to another paper published in this special issue of Computational Materials Science by Bèland et al. who compare a number of accelerated methods, including k-ART using this system as a model reference.

### 3.4. Amorphous silicon: a long-time simulation of a fully disordered system

Simulating the structural evolution of disordered solids represents a unique challenge for off-lattice KMC methods. Unlike crystalline systems where there is a reference crystalline geometry that could be used to simplify cataloging, disordered systems are entirely composed of “defective configurations” so that each atom’s environment is unique.

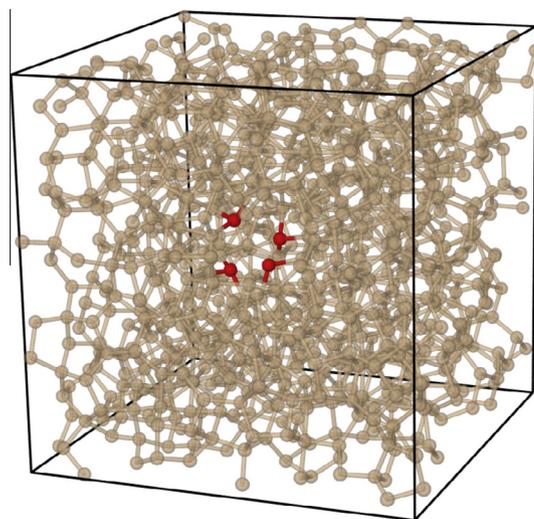
Because of this complexity, k-ART is currently the only KMC-based method that can be applied to amorphous and disordered systems. Even though, because of structural complexity, constructing an event catalog for a disordered solid is much more costly in terms of computational resources than for a defective crystal, k-ART is perfectly capable of handling these systems correctly and can even reach simulation timescales on the order of a second at room temperature.

In a recent study [12], we used k-ART to study the stability of vacancy-like defects in amorphous silicon (*a*-Si) at 300 K. For the last 30 years, a number of experimental results have suggested that vacancy-like defects exist and play an important role in the structural relaxation of *a*-Si [62–68]. Unlike in *c*-Si, however, the disordered nature of *a*-Si makes it possible for single isolated coordination defect (dangling or floating bond) to exist.

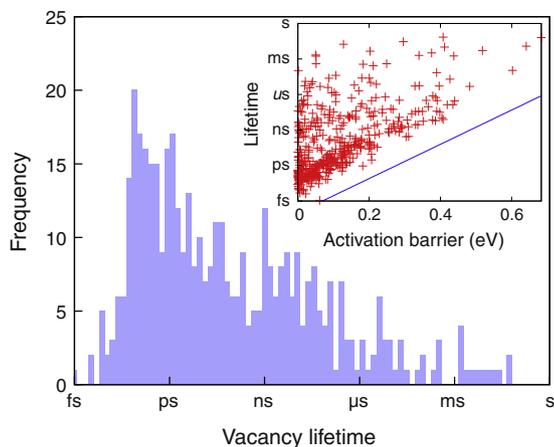
In view of this additional richness in local defect, the debate has long ported on the presence and stability of bond versus point defects: do vacancy-like defects (four correlated dangling bonds) dominate in *a*-Si, for energetic reasons, or does the larger entropic contribution of isolated bond defects favor the latter? Unfortunately, most molecular simulation studies done on the subject were limited to timescales of less than one ns [69–73], much too short to provide a complete picture of the phenomenon at low temperatures.

Here we used the unique capabilities of k-ART to generate atomistic trajectories over long time scales. Models containing a single vacancy were created by removing an atom from an initial well-relaxed *a*-Si model. One thousand models were made by removing sequentially every atom of the initial 1000 atom model. Vacancy-like defects are defined as a void surrounded by a cluster of four dangling bonds (missing bond of an under-coordinated atom) as shown in Fig. 8. These simulations use a modified version of the Stillinger–Weber potential optimized to correctly describe *a*-Si [74]. Of the initial vacancies, 547 are unstable and disappear during the energy minimization. k-ART simulations were launched on the remaining 453 models and the diffusion and lifetime of these vacancy-like defects were studied. Since the initial configurations are close to one another (only 40–50 atoms are displaced from the initial unperturbed model), the event catalog generated from one of the configuration was reused to speed up the simulations the remaining models.

In Fig. 9, we see the distribution vacancy lifetimes for the 447 models that lost their vacancy through an activated event while the inset shows the activated barrier of the annihilation event as a function of vacancy lifetime. While most of the vacancies are annihilated rapidly (only 43 remain after 1  $\mu$ s), it is important to note that 6 vacancies are still present after 0.1 s of simulated time (0.6% of the original 1000 created) when the simulations are stopped. The wide distribution of activation barriers for the annihilation events seen in the inset shows that there is no strong



**Fig. 8.** A single vacancy created by removing an atom from a perfectly coordinated amorphous silicon model. In red is shown the atoms that are undercoordinated. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 9.** Vacancy lifetime distribution of the vacancies which were annihilated by an activated event (not including those lost during the initial minimization). The inset shows the distribution of the vacancies' lifetimes as a function of the annihilation event's activation barrier. This data was first reported in Ref. [12].

correlation between the activation barrier of the annihilation event and the lifetime of the vacancy. Rather, the longer lifetimes are often caused by the need of the system to proceed with a energetically costly local rearrangement around the vacancy site before the final annihilation step.

In conclusion, a dangling or floating bond (over-coordinated atom) is still present around the vacancy site. Either because the annihilation step did not remove all of the dangling bonds or because new coordination defects were created during the annihilation. Finally, full diffusion of the vacancies as a cluster of four dangling bond was not seen for more than one step. In all cases, the cluster was partially annihilated soon after it moved.

These results demonstrates the strength of k-ART in treating complex materials. The limiting factor, here, is the number of low-energy barriers. This is why the selected initial state was well-relaxed, allowing the simulation to reach long time scales. The reuse of the event catalog was also crucial to be able to launch 1000 simulations with time scales of up to 1 s. The catalog is also useful during the simulation, as the system explores basins and flickers of different energy levels. It is also useful to identify which states should be recomputed and which are still valid, saving also considerable time, although less, clearly, than in a simpler environment. More work remains to be done to fully characterize the help provided by the catalog for such complex systems.

#### 4. Details of implementation

Now that we have shown what k-ART can do, it is useful to go back to the algorithm and look at its implementation in more detail (see Fig. 1).

##### 4.1. Generating events

For cataloging purposes, each generated event is assigned to the topology of the atom that moves most from the initial configuration to the saddle point. While this is a somewhat arbitrary assignment, it is uniquely defined, limiting the number of degenerate events stored in the catalog.

##### 4.2. Constructing the catalog

The current implementation of the topological classification is straightforward. First, the discretization of the events and the

generation of the key that will be used in the catalog follows three steps:

1. Select all atoms within a sphere of radius  $R_c$  around the central atom; we typically use a 6–7 Å cut-off.
2. Generate a connectivity list using a cut-off distance,  $r_c$ , that is generally between first and second neighbor distance.
3. Send the graph to NAUTY [75] to obtain a unique key associated with the automorphic group it belongs to.

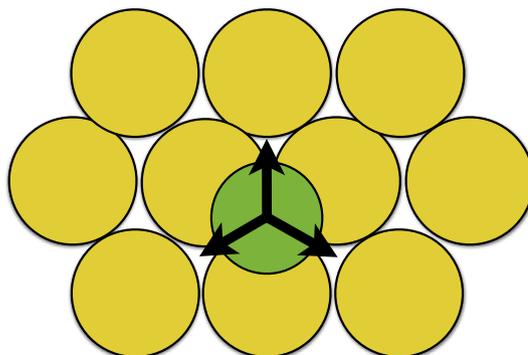
These steps are applied anytime a topology label is sought. The topology of each atom in the initial configuration is first determined. A preset number of ART nouveau event searches, typically 50, is then launched on each topology, selected at random between the atoms belonging to the same class.

For alloys involving atoms of different sizes, it is preferable to define the size of the graph by the number of vertices, i.e. the number of atoms included in the local neighborhood. This ensures that the amount of information used for the topological classification does not depend on the local chemistry. Since such an approach requires computing and ordering distances between neighboring atoms, it is best to limit its application to cases where it is essential.

Once an event is generated, a topology label is obtained for the saddle point and the final energy minimum, and the event is compared with the current catalog to establish whether this is a new event or not. If not, the event is discarded and a new search is launched. If so, a unique event key is generated by hashing the topology index of the initial, saddle and final states and it is added to the generic event catalog. This phase is repeated until a minimum number of event searches has been launched on every identified topology.

During the catalog building, it is necessary to handle correctly the symmetries of the system. In crystalline environment, many directions are equivalent. The topological analysis does not distinguish, for example, between an atom jumping to the right, left, top or bottom on a face-centered cubic (FCC) lattice (see Fig. 10). These three directions, however, must be included in the catalog to ensure the right diffusion mechanisms. To do so, we keep track of the real-space configuration-specific atomic displacement from the initial state to the saddle point.

If a second event on the same atom is found with the same topological index and displacement but with a different direction of motion, we identify the move as belonging to this class of events and we generate two events, with the same topology but a different motion. To ensure that sampling catches all the symmetries, we then add a fixed number of searches to the standard value and continue the standard event searches. This automated



**Fig. 10.** Cartoon representation of an atom diffusion on a surface of a FCC lattice. Jumps indicated by the arrow share the same topological index. To distinguish them, it is therefore necessary to add a displacement vector component.

approach has the benefit of not forcing the preliminary identification of local symmetries. It is sturdy enough, however, to have provided the right answers both with ion-implanted Si and vacancy and carbon diffusion in Fe.

This approach respects the overall philosophy of k-ART, which attempts to be as automated as possible in order to minimize the information that must be provided at the onset of the simulation and, therefore, biases that could be introduced.

#### 4.2.1. Storing events

As discussed above, each event is given a unique label obtained by hashing the topological ID of the initial, saddle and final states. To avoid clashes, labels are stored in a large table with a predefined size. When hash collisions occurs, the label number is incremented by one until an empty spot is found.

Avoiding hash tag clashes does not affect the fundamentals of the algorithm since events are linked directly to the initial topology. Nevertheless, it is preferable to limit clashes as they require a number of steps that must be repeated every time we need to check whether a newly generated event is already in the catalog. Since the number of elements can grow significantly, particularly for alloys and disordered systems – for the simulation on  $\alpha$ -Si, for example, this table include between 300.000 and 350.000 elements – it is useful to set this table to as large a size as possible.

#### 4.3. Reconstructing events: the specific events

Once event searches are finished both for new initial and for recurring topologies, we can construct the event table for the current configuration. First, all generic events associated with the active topologies are copied. Then a first rough rate is computed. This rate serves to determine which events should be specifically reconstructed and fully relaxed for current configurations.

Constructing a specific event from a generic one is a fundamental part of k-ART algorithm.

1. A correspondence table is generated using the topological transformations provided by NAUTY, that links all topological graphs to a unique reference structure for each automorphic group. This information identifies the relation between the atom list associated with the generic event and the specific representation for a local environment in its current conformation.
2. The saddle-point geometry is reconstructed for the local environment based on the knowledge of the associated initial minimum as well as the reference (generic) state initial minimum and saddle point geometries.
3. ART nouveau activation steps are then applied to converge the reconstructed geometry to the exact saddle point, allowing all the atoms in the box to relax to their optimal position.

#### 4.4. Interfacing with LAMMPS

The k-ART package is forcefield-independent. It can be applied to any system as long as the force used are continuous and derivable, so that the Hessian is uniquely defined. To simplify its use, k-ART is interfaced to the LAMMPS library [76,77], giving access to a large array of empirical forcefields and parameters. With this interface, it is relatively straightforward to port any additional forcefield to k-ART.

Beyond the systems discussed here, we have tested k-ART with LAMMPS using Stillinger–Weber and Tersoff potentials for Si, EAM and MEAM potentials for pure and alloyed metals and ReaxFF for silica mixtures. In each case, we could reproduce MD results in the short-time regime.

While formally, k-ART is not limited to empirical descriptions and can also be used with more accurate force calculations, the computation cost of the method is still prohibitive for most applications and the current package has not been tested with this approximation level for interactions.

#### 4.5. Accelerating the simulation

As with any KMC approach, k-ART involves considerable data management and calculations. If, for simple systems, the algorithm as described above offers a clear acceleration with respect to MD for a simple systems, any computational gain is welcome in order to study larger and more complex systems, such as those presented in the previous section.

##### 4.5.1. Handling flickers: the bac-MRM method

Flickers can cause significant slowing down in the time scale sampled using KMC approaches: since the computational costs associated with generating a step is independent of the barrier height, low-energy barriers can easily critically slow down any KMC method. To avoid this well-known limitation, it is necessary to handle these low-barrier events with a different algorithm. For k-ART, we have developed the basin-autoconstructing Mean Rate Method (bac-MRM) [9] which is a basin-based acceleration scheme build upon the Mean-Rate Method introduced by Puchala et al. [78].

A flickering state is defined as a pair of local minima with similar energy and connected by a barrier below a given energy threshold. A basin represents a set of connected flickering states.

The bac-MRM computes the average escape time from this basin to nearby states, taking fully into account the energy of each flickering state and the exact barriers for leaving the basin. To compute the escape probability to a state  $x$ , we must first determine the transition probability matrix  $\underline{T}$  given by:

$$T_{ji} = \frac{R_{i-j}}{\sum_k R_{i-k}} = \tau_i^1 R_{i-j}. \quad (12)$$

Here,  $R_{i-j}$  is the rate going from basin state  $i$  to basin state  $j$  and is normalized by the sum over all escape rates which corresponds to  $\tau_i^1$ , the mean residence time in state  $i$  each time it is visited.

Defining  $\Theta_i(t)$  as the occupation probability of a state  $i$  at time  $t$ , we can compute the in-basin occupation probability vector as

$$\Theta^{\text{sum}} = \sum_{m=0}^{\infty} \underline{T}^m \Theta(0) = (\mathbf{1} - \underline{T})^{-1} \Theta(0), \quad (13)$$

where the components of the initial vector,  $\Theta(0)$ , are given by  $\Theta_i(0) = \delta_{is}$ , with  $s$ , the starting state.

The mean residence time in basin state  $i$  before leaving the basin is therefore simply:

$$\tau_i = \tau_i^1 \Theta_i^{\text{sum}}. \quad (14)$$

That is all we need to obtain the escape rates from the basin  $i$ ,

$$\langle R_{i-j} \rangle = \frac{\tau_i}{\sum_k \tau_k} R_{i-j}, \quad (15)$$

with  $k$  summing over all basin states. These rates are then inserted into the full rate matrix on a par with all other activated events to determine the next time step, which allows the KMC algorithm to take these basins directly into account, ensuring that the correct kinetics is followed.

In summary, MRM separates the trajectory into basin states and states outside the basing to which system can access, and accelerates the simulation by averaging over all possible jumps between basin states, yielding the correct probability to exit a basin to a certain absorbing state.

#### 4.5.2. Event reusing

Event reusing in k-ART takes place at two levels with the goal of minimizing the cost of each KMC step: First, the cataloging of generic events. Second, the on-the-fly short-term storing of specific events.

Generic events are associated with the computed specific topology centered on the atom that move the most between the initial and the saddle states. For each local-minimum topology, a list of events is stored along with an example of geometry at the initial minimum, saddle state and final minimum, accompanied by the associated topologies, as well as the forward and backward activation energy and the cut-off used. This last number is generally constant except when it is necessary to lift a degeneracy in the correspondence between geometry and topology.

Once generated, the catalog can be reused as a starting point for other simulations on the same or related system, accelerating the overall efficiency of the method. For example, the catalog generated for defect diffusion in *c*-Fe can be used as a starting point for studying dislocation diffusion or C diffusion in Fe. This time saving feature was essential for the study of vacancy diffusion in amorphous silicon (*a*-Si) [12]: generating a event catalog for a *a*-Si 1000-atom cell required about 2000 h. Reusing the same initial catalog for the 1000 different configurations studies saved about 2 000 000 h of calculations and made the project manageable.

During the simulation, the lowest-energy events are fully reconstructed and converged at the saddle point to correctly include all elastic effects in the kinetics. While the number of events depends on the system studied, this step generally involves refining between ten and one hundred events to their exact transition state using ART nouveau. It is therefore useful to store the reconstructed specific events when moving to the next step and to reuse the transition states as initial configurations for a new convergence, when appropriate.

Such event reusing can cut the computational time by 20–80% depending on the system studied and the distribution of low-energy barrier events with very low overhead.

#### 4.5.3. Event and force parallelization

In order to take advantage of modern computer architecture, most actions in k-ART are regrouped and managed through a list of tasks to accomplish. At the beginning of each step, a list of event searches is first established on the master node. These are then dispatched to the worker nodes, that return with either a successful event or a failure; the master node analyses these returns independently of each other and can therefore send a new task without waiting for all nodes to be finished. A similar organization is applied when constructing refining specific events. In this case, again, a list is first constructed and attributed in sequence to available nodes.

Such an approach allows the use of a flexible number of processors, simplifying submissions and offering a significant gain in human time that depends on the complexity of the system studied as well as on the simulation stage. For example, the initial event catalog used for following energy relaxation in a 1000-atom *a*-Si box is generated in about 10 h using 200 cores. Since, during the k-ART simulation, only a small number of atoms change their environment, 10–50 cores, depending on the system, is generally optimal to ensure close to 100% usage. For example, the optimal number of cores for a few defects in *c*-Si is about 10 while it is between 10 and 20 for the 50-vacancy Fe system.

It is also possible to increase efficiency by also parallelizing the forces, taking advantage of LAMMPS's parallelization, for example [76]. We are currently testing our implementation of this multiple level parallelization which should be available in the distributed package.

#### 4.5.4. Local forces

Contrary to MD, where events can take place in parallel, and the time increases linearly with the number of steps, irrespective of system size, the sequential nature of KMC means that the average time step decreases as  $1/N$ , where  $N$  is the number of atoms, for a homogeneous system and, therefore, the number of steps needed to achieve a constant time scale increases like  $N$ .

With, MD, however, the force must be computed on all atoms at every step, an  $\mathcal{O}(N)$  operation. This is not the case for k-ART. Since activated events are local in nature, it is possible, when using empirical potentials, to restrict the force calculations to the atoms involved in the event and their affected neighborhood, for an  $\mathcal{O}(1)$  computational cost.

Therefore, when using a local force calculations, the overall computational costs for reaching a given timescale goes like  $\mathcal{O}(N)$  for k-ART, the same as for MD.

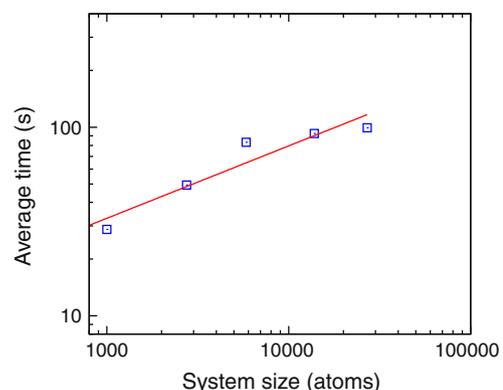
We have used this approach for simulations of a 27 000-atom of ion-implanted Si adding a full  $\mathcal{O}(N)$  force calculation for the final convergence of specific events and accepted energy minima, to ensure that long-range elastic effects are fully included [11]. As shown in Ref. [79], for the system studied, this raised the computation cost to  $\mathcal{O}(N^{0.4})$ , slightly more than MD but still well below 1 (see Fig. 11).

#### 4.6. Rigid lattice

To isolate effects stemming from long-range interactions from defects, we modified k-ART to intentionally suppress any influence from the atomic structure outside the local environment used for topological classification. This is achieved by embedding the atoms inside the topological classification sphere with radius  $R_c$  in a defect-free “generic” bulk crystal. This so-called *generic embedding* (GE) [80] is obviously not equivalent to skipping the specific event refinement described in Section 4.3: In standard k-ART, generic events are searched and found in a certain arrangement of all atoms and may thus include a bias introduced by long-range interactions.

The GE procedure corresponds to simulating defects in a rigid lattice, as no elastic deformation can propagate outside the topology cutoff. Inside the cutoff sphere, off-lattice positions are still permitted. This makes generic embedding simulations much more sophisticated than standard KMC codes which usually require on-lattice positions throughout the system.

The events obtained in this fashion are truly “generic”, as the same barrier energy can be used for all representations of a



**Fig. 11.** Scaling with system size of the average time necessary for generating a generic event search. Symbols represent the simulation data and the red line is a least-square fit of  $f(N) = 2.3N^{0.4}$  (reproduced from Ref. [79]). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

topology environment. Consequently, these events are not refined to specific events, which significantly reduces computational effort. The overhead for creating the embedding environment is negligible compared to the cost of searching for events. This speeds up GE simulations compared to their full counterparts, particularly once a system has explored most of the topologies, and a significant fraction of computational effort is devoted to refining the already known events.

It should be noted that GE requires the system to show some resemblance to the bulk crystal, for example by having a significant fraction of atoms identifying as bulk atoms through their topology key. It is thus most useful to study elastic effects on the evolution of defects in a crystalline system.

Generic embedding has been successfully applied to vacancy agglomeration in bcc iron [80], where it was shown that long-range elastic interactions affect the kinetics in this system primarily on intermediate time scales (milliseconds). There, these interactions seem to trigger quick cascades of energy release, which happens significantly slower if long-range effects are suppressed. On longer time scales however, both kinds of simulations lead to similar final states in this comparatively compact system. In more open (e.g. covalently bound) materials, elastic interactions decay less rapidly with distance and larger effects are to be expected.

## 5. Discussion: strengths and limitations of k-ART

k-ART is not the only algorithm attempting to lift these restrictions. Henkelman and Jónsson introduced an off-lattice KMC approach with on-the-fly identification of events in 2001 [1]. Without catalog, this method was applied only to relatively simple systems but already demonstrated the need to handle correctly elastic effects. More recently, a number of groups have proposed approaches to lift some limitations of KMC [6,81].

Given the number of accelerate methods that have been proposed recently, many of which are discussed in this special issue, it is useful to establish the strengths and limitations of k-ART.

### 5.1. Local versus global moves

As a general rule, all events generated with ART nouveau, including the activation towards a new saddle point, are defined in the full  $3N$ -dimension space of the configuration box. Events can therefore involve as few or as many atoms as necessary and, for energy minima and saddle points, the elastic effects are exactly taken into account (within, of course, the limits of the potential used). While the initial random deformation is generally applied to a few atoms, in order to facilitate the sampling of a large number of events, the whole system can respond and the final event can involve any atom in the system, preventing biases that could limit the generation of specific event classes.

In practice, activated events tend to see only a few atoms change their bonding environment, as should be expected in the temperature regime where KMC approximations are valid: well below melting and in systems for which barriers are high with respect to  $k_B T$ . However, long-range elastic deformation are observed and their effects are directly measurable in the observed barrier heights for same events in varying mid- and long-range environments. However, these mid- and long-range effects are particularly important at the transition state, since, by construction, this point implies a less favorable state, with larger elastic deformations than at the minimum.

Even when using local forces during the activation, in order to accelerate the algorithm, we make sure that the final relaxation step, once converged to the saddle point or to a new minimum, is performed on the full  $3N$ -coordinate energy landscape.

While the kinetic ART does not define or restrict event size, it is important to ensure that the size of the region defining the topological label is sufficiently big so that it contains all atoms changing their neighboring environment in the last events otherwise, the event reconstruction will not work properly. Since, in general, less than 10 atoms change their coordination during events for the systems presented above, a sphere containing 50–60 atoms is completely sufficient to ensure that there is no even indirect limitation to event sizes within the algorithm.

This also means that the usefulness of the topological classification is limited to cases when events are relatively local in nature. In highly strained systems, with avalanche-like motion, involving possible hundreds of atoms or more, the approach we use here is not applicable. In that case, however, it is still possible to perform the simulation, as long as a new catalog is constructed after such large scale event.

### 5.2. Ensuring detailed balance

One of the advantages of using a catalog is that it is possible to enforce strict detailed balance by adding all events to the catalog both in forward and reverse direction. While it is not possible to demonstrate that the event list is complete, this simple step, coupled with full relaxation of relevant events, is sufficient to make sure that (i) every trajectory is fully reversible and (ii) detailed balance is respected locally for each event even when also using the bac-MRM flicker handling algorithm.

We emphasize that local detailed balance cannot be assured in the absence of a catalog as only a complete sampling of events would guarantee that all reversible pathways are found.

### 5.3. Ensuring completeness

While detailed balance is strictly enforced in kinetic ART, it is not possible to guarantee the completeness of the catalog. Indeed, saddlepoint search method available today are not systematic so that completeness is never assured. Completeness is also missing from most lattice-based KMC simulations. This means that important pathways can be missed and it is impossible to account for these possible systematic errors when evaluating kinetic pathways and timescales.

In spite of this formal limitation, extensive searches on well-known (and relatively simple) systems have shown that with sufficient sampling, it is possible to recover all important and kinetically relevant mechanisms [23,31]. Moreover, these methods also provide a much richer event catalog than lattice-based simulation, increasing the accuracy of the results. k-ART and similar methods represent therefore a significant step forward in addition to accessing a new range of physical phenomena.

Yet, as long as it is not possible to evaluate and quantify the catalog's completeness, a correct error measure remains impossible and efforts must be continued to address this fundamental issue.

### 5.4. Topological classification

An important element of the applicability of kinetic ART to complex materials is the use of a topological classification for cataloging. In order to be effective, there must be a close relationship between local geometry and local topology. In simple terms, this means that the topology defines a unique geometry for a local region when the corresponding connectivity graph is embedded with geometrical constraints defined by (i) the atomic positions of the network outside of the local region, (ii) a given forcefield and (iii) a connectivity list.

In practice, this correspondence requirement is less stringent than it appears as the graph is first reconstructed using a reference

geometry belonging to the same automorphic group and the nearby initial minimum state. More precisely, the saddle point and the final energy minimum are reconstructed using both the initial energy minimum and a reference saddle and final energy minimum states, adding a strong “proximity” constraints that limits clashes and high-energy configurations.

The reconstruction of saddle point configurations imposes a strict test on the correspondence validity as it is an unstable state. It is therefore easy to establish when the reconstruction has failed. If that happens, it means that more than one geometry correspond to the same topology. To resolve the problem, it is possible to change the cut-off used to define the topology until we recover a unique correspondence.

It is not necessary to ensure that the correspondence is bijective. More than one topology can therefore correspond to the same geometry, within the limits of elastic deformations. While such degeneracy adds unnecessary complexity to the event catalog, it does not affect the kinetics nor the list of events associated with a local minimum as each atom is defined by a single topology.

The generic topologically-based cataloging approach should be applicable to all activated-based kinetics defined by discrete sets of minima provided that the connectivity graph is sufficiently rich, with cut-offs that could vary, for example, with the local environment. In practice, this means that systems with a very rough energy landscape might be more difficult to handle with this approach and would require to adapt cut-offs more specifically.

This roughness, however, does not always reflect a complex diffusion landscape. This is the case, for example, of silica,  $\text{SiO}_2$ , a network glass where, at  $T = 0$ , the oxygen atoms can occupy many states of similar energy. Formally, therefore, it is very difficult to reconstruct uniquely the oxygen position using a topological classification. However, since the exact position of the oxygen atoms does not impact the diffusion barriers, it is possible to neglect the degenerate oxygen positions without affecting the physics.

### 5.5. Handling flickers

As described above, the bac-MRM computes the average escape time out of a basin of flickering states. While this solution represents a huge efficiency gain with respect to straightforward KMC, building the basin can be time consuming. However, as soon as the system goes out of the basin and evolves into a new state, one has to rebuild the complete basin, which can slow considerably the simulation.

To limit this cost, we can first store the whole basin in memory until a new one is found. This way, if the system falls back into it, after a few steps, we can at once find a new escape, with minimal computational cost.

It is also possible, when the studied system is large and flickers are far apart geometrically, to introduce *local basins*, independent basins that can be treated as individual events and limit the information loss between events. This way, when an accepted event takes place in one part of the system, only basins physically near the affected region are destroyed, the others remaining in the catalog. Because relevant barriers are reconstructed at each step, the introduction of local basins will not change the overall simulation kinetics as long as the interaction between these basins is small enough not to change the basin states and their event list. In practice, this means that we have applied this approach only for very large systems, counting at least many thousands of atoms. In this case, however, the efficiency gain is significant.

As a final comment, we note that the bac-MRM algorithm is much easier to implement than more traditional first-escape time approaches [82,78]. Recently, Athènes et al. proposed an efficient first-escape time flicker-handling algorithm that seems to be of similar computational cost as bac-MRM [83]. While our first

approach is statistically correct, the latter is exact on an event-per-event basis. Whether there are actual non-statistical differences between trajectories generated with these two approaches for real problems remains to be seen and we are currently testing this question.

## 6. Conclusion

The simulation of complex materials on experimentally relevant time scale remains a challenge. Recent developments, however, have made it possible to access this time regime in system dominated by activated mechanisms, i.e., mechanisms controlled by energy barriers high with respect to the temperature, by extending the application of kinetic Monte Carlo (KMC) methods to non-crystalline environments.

The kinetic Activation–Relaxation Technique (k-ART), an off-lattice KMC approach with topologically-based on-the-fly cataloging described in this review article, represents, at the moment, the most flexible algorithm for handling point defects, alloys and even fully disordered materials. We have shown here that it manages to reach time scales a million to a billion times longer than MD for the same systems, opening up the study of long-ignored phenomena.

As the more technical discussion shows, algorithms such as k-ART are complex and cannot be run as easily as molecular dynamics. However, as we learn from applying them, we can simplify their use and optimize much more readily their parameters. Clearly, we can expect to see a handful of other algorithms to be proposed in the coming years. In addition to providing new approaches for solving the problem, these will be helpful ensure the validity of the various methods.

In the meanwhile, however, k-ART is there and available, opening up a new era in computational materials sciences.

## 7. Code availability

The ART nouveau package is freely distributed here.<sup>1</sup> It can be linked to *first principle* codes and an example with SIESTA is provided. ART nouveau can also be used with empirical potentials either ported to the code or, when the ART nouveau package is linked against LAMMPS library, any forcefield available to LAMMPS [76,77].

The kinetic ART package is still in development, but is distributed on request. It can be used with the SW potential integrated in the code as well as linked to the LAMMPS library for a wider range of interactions. At this point, the k-ART package can be obtained by contacting the authors.

## Acknowledgments

This work has been supported by the Canada Research Chairs program and by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC). The work of OR and FE is supported by the Qatar National Research Fund (QNRF) through the National Priorities Research Program (NPRP 6-863-2-355). We are grateful to *Calcul Québec/Compute Canada (CQ/CC)* for generous allocations of computer resources. NM would also like to thank Gerard Barkema and the Institute for Theoretical Physics of Utrecht University, where a good part of the article was written, for their hospitality. Some of LKB's work was supported by the Center for Defect Physics, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences.

<sup>1</sup> See <http://normandmousseau.com>.

## References

- [1] G. Henkelman, H. Jónsson, *J. Chem. Phys.* **115** (2001) 9657.
- [2] T.F. Middleton, D.J. Wales, *J. Chem. Phys.* **120** (2004) 8134.
- [3] F. Hontinfinde, A. Rapallo, R. Ferrando, *Surf. Sci.* **600** (2006) 995–1003, <http://dx.doi.org/10.1016/j.susc.2005.12.025>.
- [4] L. Xu, D. Mei, G. Henkelman, *J. Chem. Phys.* **131** (2009) 244520, <http://dx.doi.org/10.1063/1.3281688>.
- [5] A. Kara, O. Trushin, H. Yildirim, T.S. Rahman, *J. Phys.-Condens. Mater.* **21** (2009) 084213, <http://dx.doi.org/10.1088/0953-8984/21/8/084213>.
- [6] H. Xu, Y. Osetsky, R. Stoller, *Phys. Rev. B* **84** (2011) 1.
- [7] D. Konwar, V.J. Bhute, A. Chatterjee, *J. Chem. Phys.* **135** (2011) 174103.
- [8] F. El-Mellouhi, N. Mousseau, L.J. Lewis, *Phys. Rev. B* **78** (2008) 153202.
- [9] L.K. Béland, P. Brommer, F. El-Mellouhi, J.-F. Joly, N. Mousseau, *Phys. Rev. E* **84** (2011) 046704.
- [10] P. Brommer, N. Mousseau, *Phys. Rev. Lett.* **108** (2012) 219601, <http://dx.doi.org/10.1103/PhysRevLett.108.219601>.
- [11] L.K. Béland, Y. Anahory, D. Smeets, M. Guihard, P. Brommer, J.-F. Joly, J.-C. Pothier, L.J. Lewis, N. Mousseau, F. Schiettekatte, *Phys. Rev. Lett.* **111** (2013) 105502, <http://dx.doi.org/10.1103/PhysRevLett.111.105502>.
- [12] J.-F. Joly, L.K. Béland, P. Brommer, N. Mousseau, *Phys. Rev. B* **87** (2013) 144204, <http://dx.doi.org/10.1103/PhysRevB.87.144204>.
- [13] H. Eyring, *Chem. Rev.* **17** (1935) 65, <http://dx.doi.org/10.1021/cr60056a006>.
- [14] M.G. Evans, M. Polanyi, *Trans. Faraday Soc.* **31** (1935) 875–894, <http://dx.doi.org/10.1039/TF9353100875>.
- [15] A.F. Voter, F. Montalenti, T.C. Germann, *Ann. Rev. Mater. Res.* **32** (1) (2002) 321.
- [16] A. Bortz, M. Kalos, J. Lebowitz, *J. Comput. Phys.* **17** (1) (1975) 10–18, [http://dx.doi.org/10.1016/0021-9991\(75\)90060-1](http://dx.doi.org/10.1016/0021-9991(75)90060-1).
- [17] A.F. Voter, *Phys. Rev. B* **34** (1986) 6819, <http://dx.doi.org/10.1103/PhysRevB.34.6819>.
- [18] L.E. Reichl, *A Modern Course in Statistical Physics*, Wiley-VCH, 2009.
- [19] J.P.K. Doye, D.J. Wales, *Z. Phys. D Atom. Mol. Cl.* **40** (1997) 194, <http://dx.doi.org/10.1007/s004600050192>.
- [20] G. Henkelman, B.P. Uberuaga, H. Jónsson, *J. Chem. Phys.* **113** (2000) 9901.
- [21] A. Goodrow, A.T. Bell, M. Head-Gordon, *J. Chem. Phys.* **130** (2009) 244108, <http://dx.doi.org/10.1063/1.3156312>.
- [22] G. Barkema, N. Mousseau, *Phys. Rev. Lett.* **77** (1996) 4358.
- [23] R. Malek, N. Mousseau, *Phys. Rev. E* **62** (2000) 7723.
- [24] E. Machado-Charry, L.K. Béland, D. Caliste, L. Genovese, T. Deutsch, N. Mousseau, P. Pochet, *J. Chem. Phys.* **135** (2011) 034102.
- [25] G.L. Kellogg, P.J. Feibelman, *Phys. Rev. Lett.* **64** (1990) 3143, <http://dx.doi.org/10.1103/PhysRevLett.64.3143>.
- [26] G. Henkelman, H. Jónsson, *J. Chem. Phys.* **111** (1999) 7010, <http://dx.doi.org/10.1063/1.480097>.
- [27] A. Samanta, E. Weinan, *J. Chem. Phys.* **136**, <http://dx.doi.org/10.1063/1.3692803>.
- [28] R.A. Olsen, G.J. Kroes, G. Henkelman, A. Arnaldsson, H. Jónsson, *J. Chem. Phys.* **121** (2004) 9776, <http://dx.doi.org/10.1063/1.1809574>.
- [29] E. Cancès, F. Legoll, M.C. Marinica, K. Minoukadeh, F. Willaime, *J. Chem. Phys.* **130**, <http://dx.doi.org/10.1063/1.3088532>.
- [30] E. Bitzek, P. Koskinen, F. Gaehler, M. Moseler, P. Gumbsch, *Phys. Rev. Lett.* **97** (2006) 170201, <http://dx.doi.org/10.1103/PhysRevLett.97.170201>.
- [31] M.C. Marinica, F. Willaime, N. Mousseau, *Phys. Rev. B* **83**, <http://dx.doi.org/10.1103/PhysRevB.83.094119>.
- [32] E. Cancès, F. Legoll, M.-C. Marinica, K. Minoukadeh, F. Willaime, *J. Chem. Phys.* **130** (11) (2009), <http://dx.doi.org/10.1063/1.3088532>.
- [33] F. El-Mellouhi, N. Mousseau, P. Ordejon, *Phys. Rev. B* **70** (2004) 1098, <http://dx.doi.org/10.1103/PhysRevB.70.205202>.
- [34] K. Levasseur-Smith, N. Mousseau, *Eur. Phys. J. B* **64** (2008) 165–172, <http://dx.doi.org/10.1140/epjb/e2008-00296-4>.
- [35] G.T. Barkema, N. Mousseau, *Phys. Rev. Lett.* **81** (1998) 1865.
- [36] F. Valiquette, N. Mousseau, *Phys. Rev. B* **68** (2003) 125209.
- [37] D. Rodney, C. Schuh, *Phys. Rev. Lett.* **102** (23) (2009) 1–4.
- [38] H. Kallel, N. Mousseau, F. Schiettekatte, *Phys. Rev. Lett.* **105** (2010) 045503.
- [39] Y. Song, R. Malek, N. Mousseau, *Phys. Rev. B* **62** (2000) 15680, <http://dx.doi.org/10.1103/PhysRevB.62.15680>.
- [40] D. Vanderbilt, X. Zhao, D. Ceresoli, *Thin Sol. Films* **486** (2005) 125, <http://dx.doi.org/10.1016/j.tsf.2004.11.232>.
- [41] P. Koziatek, J.-L. Barrat, P. Derlet, D. Rodney, *Phys. Rev. B* **87**, <http://dx.doi.org/10.1103/PhysRevB.87.224105>.
- [42] S. Swayamjyoti, J.F. Loeffler, P.M. Derlet, *Phys. Rev. B* **89**, <http://dx.doi.org/10.1103/PhysRevB.89.224201>.
- [43] G. Wei, N. Mousseau, P. Derreumaux, *J. Chem. Phys.* **117** (2002) 11379, <http://dx.doi.org/10.1063/1.1522373>.
- [44] G. Wei, N. Mousseau, P. Derreumaux, *Proteins* **56** (2004) 464, <http://dx.doi.org/10.1002/prot.20127>.
- [45] M. Yun, R. Lavery, N. Mousseau, K. Zakrzewska, P. Derreumaux, *Proteins* **63** (2006) 967, <http://dx.doi.org/10.1002/prot.20938>.
- [46] J.-F. St-Pierre, N. Mousseau, *Proteins* **80** (2012) 1883, <http://dx.doi.org/10.1002/prot.24085>.
- [47] L. Dupuis, N. Mousseau, *J. Chem. Phys.* **136** (2012) 035101, <http://dx.doi.org/10.1063/1.3671986>.
- [48] P. Ganster, L.K. Béland, N. Mousseau, *Phys. Rev. B* **86** (2012) 075408, <http://dx.doi.org/10.1103/PhysRevB.86.075408>.
- [49] Y. Fan, T. Iwashita, T. Egami, *Phys. Rev. E* **89** (2014) 062313, <http://dx.doi.org/10.1103/PhysRevE.89.062313>.
- [50] C. Dellago, P.G. Bolhuis, F.S. Csajka, D. Chandler, *J. Chem. Phys.* **108** (1998) 1964.
- [51] G. Boisvert, N. Mousseau, L. Lewis, *Phys. Rev. B* **58** (1998) 12667.
- [52] G. Boisvert, N. Mousseau, L.J. Lewis, *Phys. Rev. Lett.* **80** (1998) 203, <http://dx.doi.org/10.1103/PhysRevLett.80.203>.
- [53] A. Voter, *Rad. Eff. Sol.* (2007) 1.
- [54] J. Hernandez-Rojas, D. Wales, *J. Non-Cryst. Solids* **336** (2004) 218, <http://dx.doi.org/10.1016/j.jnoncrysol.2004.01.012>.
- [55] H. Jiang, C. Jiang, D. Morgan, I. Szlufarska, *Comput. Mater. Sci.* **89** (2014) 182, <http://dx.doi.org/10.1016/j.commatsci.2014.03.051>.
- [56] F. Stillinger, T. Weber, *Phys. Rev. B* **31** (1985) 5262, <http://dx.doi.org/10.1103/PhysRevB.31.5262>.
- [57] D. Caliste, P. Pochet, *Phys. Rev. Lett.* **97** (13) (2006) 135901, <http://dx.doi.org/10.1103/PhysRevLett.97.135901>.
- [58] L.A. Marqués, L. Pelaz, J. Hernández, J. Barbolla, G.H. Gilmer, *Phys. Rev. B* **64** (2001) 045214, <http://dx.doi.org/10.1103/PhysRevB.64.045214>.
- [59] L. Marqués, L. Pelaz, M. Aboy, L. Enriquez, J. Barbolla, *Phys. Rev. Lett.* **91** (2003) 135504.
- [60] F.H. Stillinger, T.A. Weber, *Phys. Rev. B* **31** (8) (1985) 5262–5271, <http://dx.doi.org/10.1103/PhysRevB.31.5262>.
- [61] J.-C. Pothier, F. Schiettekatte, L.J. Lewis, *Phys. Rev. B* **83** (2011) 235206, <http://dx.doi.org/10.1103/PhysRevB.83.235206>.
- [62] S. Roorda, W. Sinke, J. Poate, D. Jacobson, S. Dierker, B. Dennis, D. Eaglesham, F. Spaepen, P. Fuoss, *Phys. Rev. B* **44** (1991) 3702.
- [63] K. Laaziri, S. Kycia, S. Roorda, M. Chicoine, J.L. Robertson, J. Wang, S.C. Moss, *Phys. Rev. Lett.* **82** (17) (1999) 3460–3463.
- [64] S. Coffa, J.M. Poate, D.C. Jacobson, W. Frank, W. Gustin, *Phys. Rev. B* **45** (15) (1992) 8355–8358.
- [65] S. Coffa, J.M. Poate, D.C. Jacobson, A. Polman, *Appl. Phys. Lett.* **58** (1991) 2916.
- [66] S. Coffa, J.M. Poate, *Appl. Phys. Lett.* **63** (1993) 1080.
- [67] S. Coffa, F. Priolo, A. Battaglia, *Phys. Rev. Lett.* **70** (1993) 3756.
- [68] S. Coffa, D.C. Jacobson, J.M. Poate, F. Priolo, *Appl. Phys. A* **54** (1992) 481.
- [69] R. Lutz, L. Lewis, *Phys. Rev. B* **47** (1993) 9896.
- [70] E. Kim, Y. Lee, *Phys. Rev. B* **51** (1995) 5429.
- [71] E. Kim, Y. Lee, C. Chen, T. Pang, *Phys. Rev. B* **59** (1999) 2713.
- [72] C.R. Miranda, A. Antonelli, A.J.R. da Silva, A. Fazzio, *J. Non-Cryst. Solids* **338–340** (2004) 400.
- [73] X. Urli, C. Dias, L. Lewis, S. Roorda, *Phys. Rev. B* **77** (2008) 155204.
- [74] R. Vink, G.T. Barkema, W. van der Weg, N. Mousseau, *J. Non-Cryst. Solids* **282** (2001) 248.
- [75] B.D. McKay, *Congr. Numer.* **30** (1981) 45.
- [76] S. Plimpton, *J. Comput. Phys.* **117** (1995) 1.
- [77] Lammps web site. <<http://lammps.sandia.gov/index.html>>.
- [78] B. Puchala, M.L. Falk, K. Garikipati, *J. Chem. Phys.* **132** (2010) 134104.
- [79] J.-F. Joly, L.K. Béland, P. Brommer, F. El-Mellouhi, N. Mousseau, *J. Phys.: Conf. Ser.* **341** (2012) 012007.
- [80] P. Brommer, L.K. Béland, J.-F. Joly, N. Mousseau, *Understanding Long-Time Vacancy Aggregation in Iron: A Kinetic Activation–Relaxation Technique Study*, 2014. arXiv:1407.7694.
- [81] C. Scott, R. Smith, *Proc. Roy. Soc. A: Math. Phys. Eng. Sci.* **469** (2013) 20120480, <http://dx.doi.org/10.1098/rspa.2012.0480>.
- [82] M.A. Novotny, *Phys. Rev. Lett.* **74** (1995) 1424, <http://dx.doi.org/10.1103/PhysRevLett.74.1424>.
- [83] M. Athènes, V.V. Bulatov, *A Path Factorization Approach to Stochastic Simulations*, 2014. arXiv:1403.228.