Elastodiffusion and cluster mobilities using kinetic Monte Carlo simulations: fast first-passage algorithms for reversible diffusion processes

Manuel Athènes,[1] Savneet Kaur,[1] Gilles Adjanor,[2] Thomas Vanacker †,[1] and Thomas Jourdan[1]

[1] *CEA, DEN, Service de Recherches de Métallurgie Physique, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France*

[2] *Groupe Métallurgie, MMC, EDF, Les Renardières, 77818 Moret-sur-Loing, France*

The microstructural evolution of metals and alloys is governed by the diffusion of defects over complex energy landscapes. Whenever metastability occurs in atomistic simulations, well-separated time scales emerge making it necessary to implement event-based kinetic models at larger scales. The crucial task then involves characterizing the important events contributing to mass transport. We herein describe fast first-passage algorithms based on the theory of absorbing Markov chains assuming that defects undergo reversible diffusion. We show that the absorbing transition rate matrix can be transformed into a symmetric definite-positive matrix enabling us to implement direct and iterative sparse solvers. The efficiency of the approach is demonstrated with direct computations of elastodiffusion properties around a cavity in Aluminum and Monte Carlo computations of cluster diffusivity in low alloyed Manganese steels.

I. INTRODUCTION

Kinetic Monte Carlo (KMC) simulations [1] are extensively used in materials science to predict the microstructural evolution of alloy systems driven out of equilibrium or to compute atomic transport properties, either at equilibrium or in a nonequilibrium steady state. A KMC method is traditionally implemented whenever the physical model system of interest is governed by a master equation which corresponds to a high-dimensional ordinary differential equation over the discrete or discretized (meshed) space. This equation can not be solved directly through time-stepping because of the combinatorial explosion in the number of variables to deal with. Any KMC method consists of simulating a single kinetic trajectory among the many possible ones. In materials science, KMC methods may treat events [2–4], objects [5, 6] or atoms [7].

The efficiency of the KMC method is drastically reduced whenever the transition matrix describing the evolution of the system exhibits a wide spectrum. In this situation, the system transitions a huge number of times between configurations separated by small energy barriers. These connected configurations form trapping basins from which the average escape time is much larger than the characteristic time for crossing the small barriers inside the basins. This issue is recurrent in KMC simulations. Cavities may form under irradiation and remain stable over a long period of time due to the low vacancy emission rate resulting from the strong attraction between cavities and neighboring vacancies. Kinetic trapping may also be caused by the formation of dynamically stable clusters of Manganese or Copper substitutional atoms and vacancies in α-iron. These solute clusters migrate slowly without dissociating owing to numerous atomic rearrangements.

Several ways of improving the KMC method are currently implemented in the literature. First, the encountered events may be tabulated for later reuse [8–16], which avoids repeatedly evaluating the same transition rates. This way of proceeding is particularly relevant whenever stable and saddle point energies are costly to evaluate as in off-lattice simulations [17]. Transition rates associated to tabulated events are then retrieved on the fly.

To reduce the dimensionality of the original master equation and further speedup the KMC simulations, nonlocal events involving mobile defects may be tabulated and randomly selected using the appropriate rule [18, 19]. Avoiding conflicts between defects evolving in parallel requires spatial protection of defects and exact time synchronization. Spatial protection serves to prevent distant walkers from colliding or conversely to enable two neighboring defects to recombine. To satisfy the time synchronization requirement, the theory of absorbing Markov chains [20] is used to draw first-passage times, and paths to distant states located on the periphery of the protection, which acts as an absorbing sink [18, 21, 22]. Drawing a first-passage time and escape from the exact probability distributions may be achieved through the direct factorization of the absorbing transition rate matrix [22] or through its eigenvalue decomposition [18, 21, 23–25].

The former randomization technique is based on the probabilistic interpretation of the factorization in term of paths [19, 26, 27]. This interpretation is implicitly invoked in the matrix method [28–30] to compute correlation factors associated with vacancy-solute exchanges in dilute alloy models for any crystallographic structure. These correlation factors are crucial quantities giving access to diffusion coefficients. They are obtained by inverting a matrix, which amounts to summing over all paths between two consecutive vacancy-solute exchanges. Green functions used in atomic transport theory [31, 32] also appear as pseudo-inverses of transition rate matrices

† Deceased 12 September 2018

and may also be interpreted as geometric sums of path probabilities.

The latter randomization technique consists of entirely computing the evolution operator for transient states, a matrix exponential. The approach was extended to Gaussian random walks in continuous three dimensional spaces using a Green function formalism [33–35]. In this framework, the probability that the walker is still in its protected volume appears as a series of decaying (real) exponential functions. At times large enough, the infinite sum can be safely truncated retaining only a limited number of the slowest eigen-modes because the contributions of the fastest eigen-modes rapidly decay with time. In these studies [33–35], the spectral decomposition is analytically tractable for the considered symmetric diffusion operators.

The symmetry property assumed in Ref. [33–35] entails that atomic transport is mediated by defects whose diffusion is reversible at equilibrium, i.e. the involved diffusion processes obey detailed balance even though the defect concentration may be out of equilibrium, as for instance after a quench or an irradiation cascade. This assumption is satisfied in many materials of practical interest. Even the state-to-state evolution of far from equilibrium glasses can be well approximated by a Markov chain that does obey detailed balance. A noticeable exception involves alloy systems subjected to steady irradiation, temperature gradients or chemical potential gradients. At the atomic scale, a consequence of reversibility is that the discrete transition rate matrix can be symmetrized by similarity transformation [31, 36]. For absorbing Markov processes, reversibility of diffusion is conditional upon the fact that the system is still located in a transient state. This guarantees that the transition rate matrix restricted to transient states is similar to a symmetric definite negative and that the transient evolution operator is the sum of decaying exponential functions. In this article, we investigate the computational implications of the conditional reversibility of the involved diffusion processes.

After introducing absorbing Markov chain theory for conditionally reversible processes in Section II, we illustrate several features of the approach on two realistic problems: the elasto-diffusion of vacancies in the neighborhood of cavities [37] in pure Aluminum and the diffusion of small vacancy-Mn clusters in α-iron. New features involve the ability to compute transition currents over large physical volumes and to approximate the evolution operator through projection on reduced subspaces, two aspects little discussed in the literature on lattice-based Markov processes, to our knowledge.

II. THEORY AND METHODS

A. Master equation and evolution operator

The phase space is considered to be discrete and is denoted by Ω. States describing the system correspond to the locations of atoms and defects (such as vacancies) on a crystalline lattice. Our knowledge about the current state of the system is materialized by a probability vector, i.e. a probability distribution over Ω. The time evolution of the probability vector $\mathbf{p}_t \equiv \mathbf{p}(t)$ is governed by the following master equation

$$\dot{\mathbf{p}}_t^T = \mathbf{p}_t^T \mathbf{K}, \tag{1}$$

where $\mathbf{K}$ stands for the Markov matrix of transition rates, assumed here to be time-independent: K_{ij} is the rate of transition from state i to state j (off-diagonal elements only). The standard convention is used to define the diagonal elements as $K_{ii} = -\sum_{i\neq\ell} K_{i\ell}$. Superscript T stands for transpose. The evolution operator, obtained formally from solutions of the master equation, can be expressed as an exponential of the Markov matrix of transition rates ($t_1 > t_0$)

$$\mathbf{P}(t_0, t_1) = \exp\left[\int_{t_0}^{t_1} \mathbf{K} dt\right] = \exp\left[(t_1 - t_0)\mathbf{K}\right]. \tag{2}$$

Matrix element $P_{ij}(t_0, t_1)$ is the probability to find the system in state j at t_1 given that it was in state i at time t_0. This operator fully characterizes the time evolution of the probability vector: $\mathbf{p}^T(t_1) = \mathbf{p}^T(t_0)\mathbf{P}(t_0, t_1)$. As defined, the evolution operator belongs to the class of stochastic matrices such that $\sum_\ell P_{i\ell} = 1$ and $P_{ij} \geq 0$ for any i, j, t and τ. This property entails conservation of the total probability. Besides, the stationary distribution satisfies $\boldsymbol{\rho}^T\mathbf{P}(t_1, t_0) = \boldsymbol{\rho}^T$, it is a left-eigenvector of the evolution operator associated with eigenvalue one.

If known, the evolution operator can be used to sample transitions between any two states and over arbitrary time intervals $\tau = t_1 - t_0$. In practice, the evolution operator can only be solved for small subspaces delimited by artificial absorbing boundaries. Substantial simulation speed-ups can be achieved by sampling transitions to distant states located on the absorbing perimeters of encountered trapping basins [18, 19, 21, 22, 25]. Prior to presenting the theory of absorbing Markov chains, which will be used to formulate such accelerated KMC algorithms, we recall the standard KMC method.

B. Standard Kinetic Monte Carlo

Standard KMC methods avoid exponentiating the transition rate matrix. The evolution operator is instead linearized to get a simple stochastic matrix and a randomization procedure is invoked to draw the time at which the event occurred. The simplest form of such

matrix is

$$\mathbf{P}^{\rm lin} = \boldsymbol{\mathcal{I}} + \tau \mathbf{K} \tag{3}$$

where $\boldsymbol{\mathcal{I}}$ is the identity operator and τ is a positive time step that must be lower or equal to $-1/K_{ii}$ for all states $S_i \in \Omega$. Time randomization then consists in drawing a time in the exponentially decaying distribution of rate τ^{-1}. Since $P^{\rm lin}_{ii} \geq 0$, it is possible that no transition has occurred after time incrementation. In practice, a different stochastic matrix is implemented, so as to guarantee a KMC transition at each step. Letting $\mathrm{Diag}(\mathbf{K})$ denote the diagonal matrix composed of the diagonal elements of $\mathbf{K}$, the following stochastic matrix is rather used

$$\mathbf{P}^{\rm std} = \boldsymbol{\mathcal{I}} - \mathrm{Diag}(\mathbf{K})^{-1}\mathbf{K}. \tag{4}$$

If system is currently located in state i, the exponentially decaying distribution of rate $-K_{ii}$ must instead be used to sample the residence time, i.e a stochastic variable yielding the elapsed time. The mean residence time on i is then $-1/K_{ii}$, compensating for the fact that the stationary distribution of $\mathbf{P}^{\rm std}$ is proportional to $-\mathrm{Diag}(\mathbf{K})\boldsymbol{\rho}$, a left eigenvector associated with the eigenvalue equal to 1. In the following we consider the generic stochastic matrix of the form

$$\mathbf{P}^{(0)} = \boldsymbol{\mathcal{I}} + \mathrm{diag}\left(\boldsymbol{\tau}^{(0)}\right)\mathbf{K} \tag{5}$$

where $\boldsymbol{\tau}^{(0)}$ stands for an effective residence time vector such that $\tau_i^{(0)} \leq -1/K_{ii}$ for all i and $\mathrm{diag}\left(\boldsymbol{\tau}^{(0)}\right)$ denotes the diagonal matrix composed of the elements of $\boldsymbol{\tau}^{(0)}$.

C. KMC implementations

The natural way of implementing KMC algorithm does not involve explicitly forming the transition matrices appearing in (5). At each cycle, the possible transition events are tabulated and two random numbers $\mathtt{r}_1$ and $\mathtt{r}_2$ are drawn uniformly in $)0,1]$ interval. The selected transition ℓ satisfies the following double inequality

$$\sum_{l=1}^{\ell-1} K_{i,j(l)} < -\mathtt{r}_1 K_{ii} \leq \sum_{l=1}^{\ell} K_{i,j(l)},$$

where i and $j(l)$ denote the current state and the ending state associated with the l-th listed transition, while the elapsed time is simulated from $K_{ii}\ln \mathtt{r}_2$.
An alternative algorithm consists in assigning an independent Markov process and time clock to each diffusing or reacting entity, while keeping a time ordered list of events up to date. Let κ_i^d stand for the total transition rate of the d-th diffusing entity from state i. At each KMC cycle, the time of the master clock is incremented to the time of the next event and the corresponding transition is performed. After an event occurred, a limited number of events needs being annihilated, created or re-sampled. This way of proceeding is statistically equivalent because the distribution of the minimum of exponential random variables is exponential with rate $\sum_d \kappa_i^d = -K_{ii}$. This alternative algorithm is easier to implement on a parallel computer architecture [38, 39]. It is currently implemented in EKMC methods [2, 3], wherein distant binary collisions can easily be simulated using simplified laws assumed to be mutually independent.

D. First-passage kinetic Monte Carlo

FPKMC is a statistically exact EKMC algorithm in which spatial protections is introduced to ensure that binary collisions are handled rigorously [21, 25, 33–35]. Figure 1 depicts the principle of the first-passage approach applied to the diffusion and collision of two vacancies in presence of trapping precipitates. FPKMC computations of the mobilities of vacancy-Manganese clusters in Iron are reported in Sec. IV. FPKMC technique requires forming the transition rate matrices appearing in (5).

E. Absorbing Markov processes

For the ease of exposition, trapping states are labeled from 1 to N and are called *transient*. Perimeter states connected to the transient states are pooled together into a single absorbing state labeled by index $N+1$. Transitions from a transient state to any perimeter states are thus replaced by a single transition to the absorbing macro-state with an overall transition rate cumulating the transition rates towards pooled peripheral states. Transitions from the macro-state to transient states or any other states are no more permitted. As defined, the absorbing Markov process coincides with the original Markov process as long as it remains located inside the trapping basin. As a result, the N trapping states become transient and the absorbing macro-state, acting like an artificial sink trapping the system infinitely, remains the only *recurrent* state of the system. The system being initially in a transient state, states beyond the perimeter states can not be reached and need not being considered. The Markov matrix for the absorbing process is thus defined as

$$\mathbf{K}^{\rm a} = \left(\begin{array}{c|c} -\mathbf{A} & \mathbf{A}\vec{\mathbf{1}} \\ \hline \vec{\mathbf{0}}^T & 0 \end{array}\right) \tag{6}$$

where $\mathbf{A}$ is a $N \times N$ matrix such that $A_{ij} = -K_{ij}$ and $\vec{\mathbf{1}} = (1, \cdots, 1)^T$ the N-dimensional column vector whose components all equal one. Vector $\vec{\mathbf{0}} = (0, \cdots, 0)^T$ is similarly defined and $\mathbf{I}$ will stand for the $N \times N$ identity matrix. The associated evolution operator reads

$$\mathbf{P}^{\rm a}(t) = \exp[\mathbf{K}^{\rm a} t]. \tag{7}$$

With these definitions, $\mathbf{K}^{\rm a}$ is a proper transition rate matrix and likewise $\mathbf{P}^{\rm a}$ is a proper stochastic matrix.

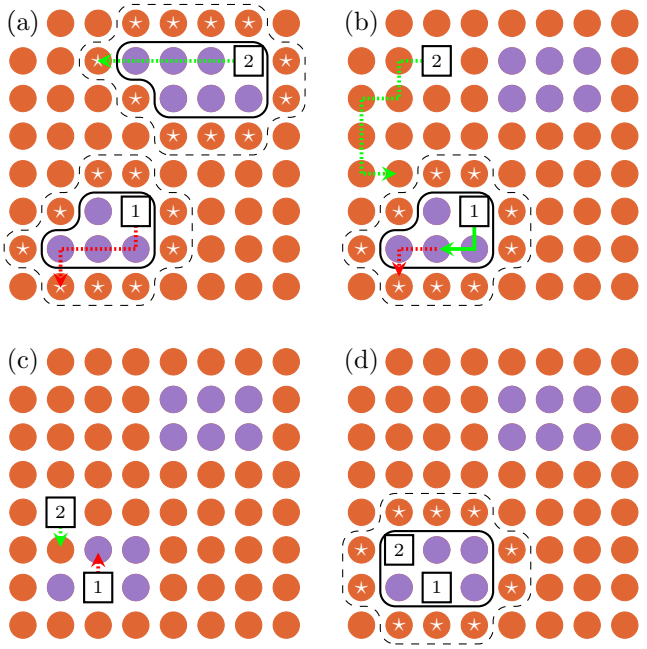


FIG. 1: Schematic of FPKMC method for two vacancies evolving on a square lattice symbolizing a FeMn alloy. Fe and Mn atoms are displayed in orange and violet. Vacancies V^1 and V^2, represented by the two labeled squares, are initially trapped inside Mn clusters. Exit sites for the trapped vacancies are indicated the stars. First-passage events are represented by dotted arrows. Events occurring first and second are colored in green and red, respectively. The sequence of events is as follows: (a) The two V-Mn clusters are spatially protected; First-passage times $t^1_{\rm fp}$ and $t^2_{\rm fp}$ are drawn (here $t^2_{\rm fp} < t^1_{\rm fp}$); (b) V^2 dissociates from Mn cluster, diffuses and collides with spatial protection of V^1 at time $t^2_{\rm col}$ (here $t^2_{\rm col} < t^1_{\rm fp}$); A no-passage event [21, 25, 33–35], materialized by the green solid arrow, is generated for synchronization; (c) V^2 attaches to the V^1-Mn cluster (nearest neighbor interactions are assumed); (d) a diffusing entity composed of two vacancies is created and spatially protected. Vacancy locations inside the thick solid line correspond to states that are referred to as *transient* in the theory of absorbing Markov processes. *Absorbing* states are those with the vacancy located on a starred site, beyond the solid line and before the dashed line.

Their components in each row summing to zero and one respectively:

$$\mathbf{K}^{\rm a}\begin{pmatrix}\vec{\mathbf{1}}\\ 1\end{pmatrix}=\begin{pmatrix}\vec{\mathbf{0}}\\ 0\end{pmatrix}\quad\text{and}\quad\mathbf{P}^{\rm a}\begin{pmatrix}\vec{\mathbf{1}}\\ 1\end{pmatrix}=\begin{pmatrix}\vec{\mathbf{1}}\\ 1\end{pmatrix}. \tag{8}$$

Consequently, for any probability vector $\boldsymbol{\pi}_t \equiv \boldsymbol{\pi}(t)$ evolving according to the master equation $\dot{\boldsymbol{\pi}}_t^T = \boldsymbol{\pi}_t^T \mathbf{K}^{\rm a}$ the probability to find the walker in one of the $N+1$ states is conserved over time and we have $\boldsymbol{\pi}_t^T = \boldsymbol{\pi}_0^T \exp[\mathbf{K}^{\rm a} t]$. We are now going to express the evolution operator (7) as a function of the exponential of $\mathbf{A}$. We first notice that the powers of minus the absorbing transition rate matrix writes for $h \geq 1$

$$(-\mathbf{K}^{\rm a})^h=\left(\begin{array}{c|c}\mathbf{A} & -\mathbf{A}\vec{\mathbf{1}}\\ \hline \mathbf{0}^T & 0\end{array}\right)^h=\left(\begin{array}{c|c}\mathbf{A}^h & -\mathbf{A}^h\vec{\mathbf{1}}\\ \hline \mathbf{0}^T & 0\end{array}\right). \tag{9}$$

This relation enables us to rearrange the matrix exponential as:

$$\begin{aligned}\exp[\mathbf{K}^{\rm a}t]&=\left(\begin{array}{c|c}\mathbf{I} & \mathbf{0}\\ \hline \mathbf{0}^T & 1\end{array}\right)+\sum_{h=1}^{+\infty}\left(\begin{array}{c|c}\mathbf{A}^h & -\mathbf{A}^h\vec{\mathbf{1}}\\ \hline \mathbf{0}^T & 0\end{array}\right)\frac{(-t)^h}{h!}\\ &=\left(\begin{array}{c|c}\exp[-\mathbf{A}t] & (\mathbf{I}-\exp[-\mathbf{A}t])\,\vec{\mathbf{1}}\\ \hline \mathbf{0}^T & 1\end{array}\right),\end{aligned} \tag{10}$$

where we substituted back the two matrix exponentials for the series in the two upper blocks of Eq. (10). The probability of being in state $j \leq N$ at time t starting from state $i \leq N$ is

$$P^{\rm a}_{ij}(t) = \mathbf{e}_i^T \exp(-\mathbf{A}t)\,\mathbf{e}_j \tag{11}$$

where $\mathbf{e}_i$ denotes the ith standard basis vector.

F. Conditional reversibility

We assume here that the original Markov process obeys detailed balance, which implies that the probability flux from state i to state j is equal to the reverse flux. The i-to-j probability flux is defined as the product of the stationary probability ρ_i to be in state i and the rate K_{ij} of transitioning to state j. The equation of detailed balance thus writes

$$\rho_i K_{ij} = \rho_j K_{ji} \tag{12}$$

When condition (12) is satisfied, the stationary probability vector $\boldsymbol{\rho}$ of the reference Markov process is usually associated with an equilibrium Gibbs-Boltzmann distribution. It corresponds to the left eigenvector of the Markov rate matrix for eigenvalue 0, since we have $\boldsymbol{\rho}^T\mathbf{K} = \mathbf{0}^T$.

The detailed balance condition (12) can be recast with respect to the absorbing Markov process considering the allowed transitions between the N transient states ($1 \leq i, j \leq N$):

$$\sqrt{\rho_i}A_{ij}/\sqrt{\rho_j} = \sqrt{\rho_j}A_{ji}/\sqrt{\rho_i}. \tag{13}$$

The following symmetric matrix is defined from (13),

$$A^B_{ij} = s_i s_j \sqrt{\rho_i}A_{ij}/\sqrt{\rho_j} = A^B_{ji}, \tag{14}$$

where the scaling factors s_i are strictly positive. They are numerical parameters reflecting the degree of freedom in the construction of stochastic matrices for KMC simulations based on (5). To specify this statement, we cast transformation (14) into matrix form resorting to the $N \times N$ diagonal matrices

$$\mathbf{S}=\sum_{i=1}^{N} s_i \mathbf{e}_i\mathbf{e}_i^T,\quad \mathbf{R}=\sum_{i=1}^{N}\frac{1}{\sqrt{\rho_i}}\mathbf{e}_i\mathbf{e}_i^T, \tag{15}$$

which are both diagonal definite positive. Hence matrices $\mathbf{S}$, $\mathbf{R}$ and $\mathbf{B} = \mathbf{S}^2$ commute and are invertible, enabling one to define

$$\mathbf{A}^B = \mathbf{S}\mathbf{R}^{-1}\mathbf{A}\mathbf{R}\mathbf{S} = (\mathbf{S}\mathbf{R})^{-1}\mathbf{B}\mathbf{A}(\mathbf{S}\mathbf{R}). \tag{16}$$

Scaling matrix $\mathbf{B}$ acts like a preconditioner. Its aim is to decrease the condition number of $\mathbf{A}^B$, which will be the main matrix in the first-passage problems investigated in the following. $\mathbf{B}$-scaling is introduced in the formalism for the sake of generality. Setting $\mathbf{B}$ to identity amounts to disabling the explicit preconditioning, as done in most literature studies and in Ref. [22] in particular. This setup also arises in the linearized KMC method based on (3), up to the τ limiting factor. In other works [19, 26, 27, 40], $\mathbf{B}$ is set to $\mathrm{Diag}(\mathbf{A})^{-1}$. This setup arises in the standard KMC method based on (4). Noticeably, it entails that $s_i = 1/\sqrt{A_{ii}}$ and $A^B_{ii} = 1$ for all i. We carry out a comparative study between the two mentioned setups in Sec. III B.

Matrix $\mathbf{SR}$ serves to make a diagonal similarity transformation and to formulate a generalized symmetric eigenvalue problem. Setting $\mathbf{B}$ to identity allows us to conclude that $\mathbf{A}$ is similar to a symmetric matrix $\mathbf{A}^I$. Similarity transformations preserving spectral properties and the spectrum of symmetric matrices being real, we conclude that the eigenvalues of $\mathbf{A}$ are real. By applying Gershgorin circle theorem to $\mathbf{A}$, we also conclude that they are positive. Eventually none of the eigenvalues are equal to zero, otherwise a stationary distribution over transient states would possibly be established, which is excluded.

G. Spectral decomposition of the evolution operator

Matrices $\mathbf{A}$ and $\mathbf{B}^{-1}\mathbf{A}^B$ being similar, we deduce that the spectrum of $\mathbf{A}$ can be obtained by solving the generalized symmetric eigenvalue problem (GSEP)

$$\mathbf{A}^B\boldsymbol{\varphi}_k = \mathbf{B}\boldsymbol{\varphi}_k\lambda_k. \tag{17}$$

Sorting the eigenvalues $(\lambda_1, \cdots, \lambda_N)$ of (17) in ascending order and letting $\boldsymbol{\Phi} = (\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \cdots, \boldsymbol{\varphi}_N)$ be a B-orthonormal basis of eigenvectors, the GSEP can be cast in the following matrix form

$$\mathbf{A}^B\boldsymbol{\Phi} = \mathbf{B}\boldsymbol{\Phi}\boldsymbol{\Lambda}. \tag{18}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix with diagonal elements Λ_{ii} equal to λ_i. Left multiplying both sides of (18) by $\boldsymbol{\Phi}^T$ and then right-multiplying (18) again by $\boldsymbol{\Phi}^T$ yields the two relations

$$\boldsymbol{\Phi}^T\mathbf{A}^B\boldsymbol{\Phi} = \boldsymbol{\Lambda}, \qquad \mathbf{A}^B = \mathbf{B}\boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^T\mathbf{B} \tag{19}$$

where we simplified resorting to B-orthogonality of $\boldsymbol{\Phi}$:

$$\boldsymbol{\Phi}^T\mathbf{B}\boldsymbol{\Phi} = \mathbf{I}, \qquad \boldsymbol{\Phi}\boldsymbol{\Phi}^T = \mathbf{B}^{-1}. \tag{20}$$

Using the generalized spectral theorem (19) and reverting relation (16), the absorbing matrix may therefore be decomposed as $\mathbf{A} = \mathbf{R}\mathbf{S}\boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^T\mathbf{S}\mathbf{R}^{-1}$. Expanding the matrix exponential of $\mathbf{A}$ in series, substituting $\mathbf{B}$ for $\mathbf{S}^2$ and invoking B-orthogonality (20) eventually yield

$$\exp\left[-\mathbf{A}t\right] = \mathbf{R}\mathbf{S}\boldsymbol{\Phi}\exp\left[-\boldsymbol{\Lambda}t\right]\boldsymbol{\Phi}^T\mathbf{S}\mathbf{R}^{-1}. \tag{21}$$

To express components of the matrix exponential, it is practically convenient to introduce a few additional notations. We denote the scaling and rescaling vectors composed of the diagonal elements of $\mathbf{S}$ and $\mathbf{R}$ by $\mathbf{s}$ and $\mathbf{r}$, respectively. We have $\mathbf{s} = \mathbf{S}\vec{\mathbf{1}}$ and $\mathbf{r} = \mathbf{R}\vec{\mathbf{1}}$. Letting $\odot$ and $\oslash$ symbols stand for element-wise multiplication and division, we also define two sets of rescaled basis vectors $\mathbf{g}_i = \mathbf{e}_i \odot \mathbf{s} \odot \mathbf{r}$ and $\mathbf{d}_j = \mathbf{e}_j \odot \mathbf{s} \oslash \mathbf{r}$. Then, resorting to the scalar products $g_i^h = \boldsymbol{\varphi}_h^T\mathbf{g}_i$ and $d_j^h = \boldsymbol{\varphi}_h^T\mathbf{d}_j$, the evolution operator (11) reads

$$P^{\mathrm{a}}_{ij}(t) = \sum_{h=1}^{N} g_i^h d_j^h \exp\left(-\lambda_h t\right),$$

where $(i,j) \leq N$. The survival probability after duration t given that the system was prepared in state $i \leq N$ at $t_0 = 0$ corresponds to the probability of not having been absorbed during the elapsed time, i.e. the probability of remaining in one of the N transient states,

$$\begin{aligned} p^{\mathrm{s}}_i(t) &\triangleq \sum_{j=1}^{N} P^{\mathrm{a}}_{i,j}(t) \\ &= (\mathbf{e}_i^T, 0)\exp\left[\mathbf{K}^{\mathrm{a}}t\right]\begin{pmatrix}\vec{\mathbf{1}}\\ 0\end{pmatrix} \\ &= \mathbf{e}_i^T\exp\left[-\mathbf{A}t\right]\vec{\mathbf{1}} \\ &= \sum_{h=1}^{N} a_i^h \exp\left(-\lambda_h t\right) \end{aligned} \tag{22}$$

where the weighing coefficients read

$$a_i^h = g_i^h \sum_{j=1}^{N} d_j^h. \tag{23}$$

Since matrix $\boldsymbol{\Phi}^T$ is B-orthogonal, the a_i^h coefficients sum to one:

$$\sum_{h=1}^{N} a_i^h = \sum_{j,h=1}^{N} \frac{r_i}{r_j}\Phi_{ih}s_h^2\Phi_{jh} = \sum_{j=1}^{N}\frac{r_i}{r_j}I_{ij} = 1.$$

This feature is consistent with the fact that initially the survival probability of the absorbing process is one : $p^{\mathrm{s}}_i(0) = 1$.

KMC simulations require drawing first-passage times with the appropriate statistics. This may be achieved by sampling a random number $\mathtt{r}_2 \propto \mathcal{U}_{]0,1]}$ (that is uniformly distributed $]0,1]$ and looking for t_{fp} satisfying $p^{\mathrm{s}}_i(t_{\mathrm{fp}}) = \mathtt{R}$. This way of proceeding requires evaluating the survival distribution though the spectral decomposition of $\mathbf{A}$. To

avoid collision with another absorbing process interacting with the transient states, it is sometimes necessary to stop the simulation at a given time $t_{\rm col}$ for synchronizing the Markov processes. This task, depicted in Fig. 1.b, involves the ability to sample the so-called *no-passage* distribution [21, 25, 33–35].

H. No-passage and quasi-stationary distributions

The no-passage distribution is the conditional probability to find the system in transient state j at time t given that it was initially in transient state i and that the process has not been absorbed yet. It is obtained by dividing the probability of being in j by the survival probability $p_i^{\rm s}(t)$:

$$P_{ij}^{\rm np}(t) = \frac{\mathbf{e}_i^T \exp\left[-\mathbf{A}t\right] \mathbf{e}_j}{\mathbf{e}_i^T \exp\left[-\mathbf{A}t\right] \vec{\mathbf{1}}}.$$

The quasi-stationary distribution over transient states corresponds to the probability vector $\mathbf{q}$ that is reached asymptotically in time by the no-passage distribution [41]. This asymptotic distribution is independent of the initial distribution, arbitrarily set equal to state i in the limit below

$$q_j = \lim_{t\to\infty} P_{ij}^{\rm np}(t) = \frac{d_j^1}{\sum_{\ell=1}^N d_\ell^1}$$

The quasi-stationary distribution being proportional to $\boldsymbol{\varphi}_1 \odot \mathbf{s} \oslash \mathbf{r}$, is also a left eigenvector of the lowest eigenvalue of the transition rate matrix $\mathbf{A}$.

I. Expected values of first-passage times, residence times and exit probabilities

The complementary of the survival probability, $p_i^{\rm a}(t) = 1 - p_i^{\rm s}(t)$, corresponds to the probability of having been absorbed. Its time derivative at t is positive and equal to the probability density of exiting at t. Since $p_i^{\rm a}(0) = 0$ and $p_i^{\rm a}(\infty) = 1$, the absorbing probability $p_i^{\rm a}(t)$ coincides with the cumulative distribution of the time probability of first passage. The mean first-passage time from i is the time expected with respect to the first-passage probability distribution. It can be formally obtained through integration by part:

$$\begin{aligned}\tau_i^{(N)} &= \int_0^\infty t \frac{d}{dt} p_i^{\rm a}(t) dt \\ &= \left[t p_i^{\rm a}(t) - t\right]_0^\infty - \int_0^\infty \left\{p_i^{\rm a}(t) - 1\right\} dt \\ &= \int_0^\infty \mathbf{e}_i^T \exp\left(-\mathbf{A}t\right) \vec{\mathbf{1}} dt = \mathbf{e}_i^T \mathbf{A}^{-1} \vec{\mathbf{1}}. \end{aligned} \tag{24}$$

The mean-first passage time satisfies $\mathbf{A}\boldsymbol{\tau}^{(N)} = \vec{\mathbf{1}}$ and is always more rapidly obtained by solving the linear system of equations involving the definite symmetric matrix $\mathbf{A}^B$:

$$\mathbf{A}^B \mathbf{x}^B = \mathbf{b}^B. \tag{25}$$

Symmetrizing the linear system entails scaling the right-hand side vector $\mathbf{b}^B$ and rescaling back the obtained solution. This is done by resorting to relation (16) between $\mathbf{A}$ and $\mathbf{A}^B$. Setting $\mathbf{b}^B$ to $\mathbf{s} \oslash \mathbf{r}$ yields the mean first-passage times as $\boldsymbol{\tau}^{(N)} = \mathbf{x}^B \odot \mathbf{s} \odot \mathbf{r}$.

The mean residence time in transient state j knowing that the system started from state i is given by the time integral

$$\begin{aligned}\int_0^\infty dt P_{ij}(t) &= \int_0^\infty \mathbf{e}_i^T \exp\left(-\mathbf{A}t\right) \mathbf{e}_j dt \\ &= \mathbf{e}_i^T \mathbf{A}^{-1} \mathbf{e}_j. \end{aligned} \tag{26}$$

The residence time vector associated with initial distribution $\boldsymbol{\pi}$ reads $\boldsymbol{\theta}^{(N)T} = \boldsymbol{\pi}^T \mathbf{A}^{-1}$ or, after taking the transpose, $\boldsymbol{\theta}^{(N)} = \mathbf{A}^{-T}\boldsymbol{\pi}$. As for the mean first-passage problem, the symmetric linear system (25) is invoked. However, it is solved using a different right-hand side vector and with transposed scalings. Setting $\mathbf{b}^B$ to $\boldsymbol{\pi} \odot \mathbf{s} \odot \mathbf{r}$ eventually provides us with mean residence times via $\boldsymbol{\theta}^{(N)} = \mathbf{x}^B \odot \mathbf{s} \oslash \mathbf{r}$.

The absorbing probability at state ℓ is the sum of the probability flux from neighboring connected states and of the initial source contribution, reflected by the identity matrix $I_{i\ell}$ below:

$$\begin{aligned}P_{i\ell}^{(N)} &= I_{i\ell} + \sum_{j=1}^N \int_0^\infty dt P_{ij}^{\rm a}(t) K_{j\ell} \\ &= I_{i\ell} + \sum_{j=1}^N \mathbf{e}_i^T \mathbf{A}^{-1} \mathbf{e}_j K_{j\ell} \end{aligned} \tag{27}$$

The probability $P_{i\ell}^{(N)}$ is equal to 0 if state ℓ is transient ($\ell < N$) or not connected to any transient state. It is non-zero for transitions to the peripheral states pooled in the absorbing macro-state.

The stochastic matrix $P_{i\ell}^{(N)}$ yields the transition probabilities used in first-passage or mean first-passage KMC methods. The latter variant method, implemented in Sec. IV and referred to as factorized KMC, increments the elapsed time by the mean first-passage time. The mean of the first-passage times can be used when a simulated walker (defect) never collides with another walker. In this case, the Markov process needs not being interrupted and synchronization is not required. Note that the stochastic matrix and residence time vector can be extended so as to encompass transitions from non-transient states $i > N$:

$$P_{ij}^{(N)} = P_{ij}^{(0)} + \sum_{\ell=1}^N P_{i\ell}^{(0)} P_{\ell j}^{(N)} \tag{28}$$

$$\tau_i^{(N)} = \tau_i^{(0)} + \sum_{\ell=1}^N P_{i\ell}^{(0)} \tau_\ell^{(N)} \tag{29}$$

where $\mathbf{P}^{(0)}$ and $\boldsymbol{\tau}^{(0)}$ are defined in (5).

J. Rank-one update

Let us assume that we have already identified $n-1$ trapping states, turned them transient by computing the transition probabilities $\mathbf{P}^{(n-1)}$ together with the mean times $\boldsymbol{\tau}^{(n-1)}$ and $\boldsymbol{\theta}^{(n-1)}$, and eventually performed a distant move. In practice, it may happen that the selected peripheral state is also a trapping state, in the sense that the system will later return to this peripheral state with a high probability. Fortunately, stochastic matrix $\mathbf{P}^{(n)}$ needs not being computed again from scratch. It is possible to perform a rank-one update of the stochastic matrix $\mathbf{P}^{(n-1)}$ by directly adding the selected peripheral state to the list of transient state. Based on $\mathbf{P}^{(n-1)}$, the probability of a transition from i to j via the new transient state labeled n is

$$P_{ij}^{(n)} = P_{ij}^{(n-1)} + P_{in}^{(n-1)} \sum_{f=0}^{+\infty} \left[P_{nn}^{(n-1)}\right]^f P_{nj}^{(n-1)}, \quad j > n,$$
$$P_{ij}^{(n)} = 0 \qquad\qquad\qquad\qquad\quad j \leq n,$$

where the sum accounts for the probabilities of all possible round-trips from n. Note that the updating rule involves both transient and peripheral states as starting states i. It is also possible to update the mean time vectors directly to get $\boldsymbol{\tau}^{(n)}$ and $\boldsymbol{\theta}^{(n)}$. The mean first-passage time to make a non-local transition starting from state n to any state $j > n$ is updated by accounting for the mean time spent performing flickers from n

$$\begin{aligned}\tau_n^{(n)} &= \sum_{f=0}^{+\infty} \left[P_{nn}^{(n-1)}\right]^f \left[1 - P_{nn}^{(n-1)}\right] (f+1)\tau_n^{(n-1)} \\ &= \tau_n^{(n-1)} \Big/ \left(1 - P_{nn}^{(n-1)}\right). \end{aligned} \tag{31}$$

Concerning the mean first-passage time to make the non-local transitions from any state i to any state $j > n$, one must account for a possible transition to state n, which eventually leads to

$$\tau_i^{(n)} = \tau_i^{(n-1)} + P_{in}^{(n-1)} \tau_n^{(n)}. \tag{32}$$

The updating rule obviously covers the case i equal to n.

K. Path factorization and space-time randomization

Path factorization [22] consists of directly constructing stochastic matrix $\mathbf{P}^{(N)}$ by repeatedly applying rank-one updates starting from stochastic matrix $\mathbf{P}^{(0)}$ defined in (5). The factorization may involve on-the-fly re-indexing. The mean first-passage time vector needs to be initialized. For all relevant states i, $\tau_i^{(0)}$ may be set to either $\tau = \min_{i\in\Omega}(-1/K_{ii})$ or $-1/K_{ii}$. The repeated updates can also be performed on the mean residence time vector $\boldsymbol{\theta}^{(n-1)}$. The starting vector $\boldsymbol{\theta}^{(0)}$ is initially set to $\boldsymbol{\tau}^{(0)} \odot \boldsymbol{\pi}$ and the updating rule (32) becomes:

$$\theta_j^{(n)} = \theta_j^{(n-1)} + \tilde{P}_{nj}^{(n-1)} \theta_n^{(n)}, \tag{33}$$

where the involved probability is defined from detailed balance and reads

$$\tilde{P}_{nj}^{(n-1)} = \frac{\rho_j \tau_n^{(0)}}{\rho_n \tau_j^{(0)}} P_{jn}^{(n-1)}. \tag{34}$$

This quantity corresponds to the canceled probability to eliminated states $j < n$, otherwise it is $P_{nj}^{(n-1)}$, the usual absorbing probability to the states $j \geq n$ that are not yet eliminated. Hence, setting j equal to n in (33) yields

$$\theta_n^{(n)} = \theta_n^{(n-1)} \Big/ \left(1 - P_{nn}^{(n-1)}\right). \tag{35}$$

Note that updating rules (31) and (35) exhibit a similar form.

The usefulness of path factorization is that it can be used not only to compute the expected values of the first-passage and residence times, but also to draw these two random variables from their exact distributions. Such randomization thus obviates the need to perform an eigenvalue decomposition. Algorithm 3 described in Appendix A and illustrated in Sec. III C implements time randomization based on the probabilistic interpretation of the factorization.

Path factorization amounts to constructing the triangular matrices involved in LU or LDL^T decomposition as well as their inverses, as shown in Appendix B. It is shown that the repeated application of updating rules (32) and (33) to obtain $\boldsymbol{\tau}^{(N)}$ and $\boldsymbol{\theta}^{(N)}$ from $\boldsymbol{\tau}^{(0)}$ and $\boldsymbol{\theta}^{(0)}$ amounts to directly applying the inverted triangular factors on $\boldsymbol{\tau}^{(0)}$ and $\boldsymbol{\theta}^{(0)}$. Noticeably, transitions to new transient states are removed within path factorization in the same way as matrix elements are canceled through Gaussian elimination. The latter technique is crucial ingredient for direct solvers. A direct sparse solver is implemented in Sec. III B to investigate the algorithmic complexity on the vacancy elastodiffusion problem. The solver computes the solutions $\boldsymbol{\tau}^{(N)}$ and $\boldsymbol{\theta}^{(N)}$ of (25) through forward and backward substitutions based on the symmetric triangular factor without its explicit inversion. The advantage of programming the rank-one updates is that the transition rate property of the transformed matrices can be preserved, as explained by Wales *et al.* in Ref. [27, 40]. Hence, the approach is more robust, albeit much slower, than available direct solvers. Preservation of probability fluxes is achieved by imposing that the flicker probability $P_{nn}^{(n-1)}$ from n and the corresponding escape probability, $D_{nn} = \sum_{j>n} P_{nj}^{(n-1)}$, exactly sum to one after each elimination. Path factorization is found more appropriate for studying trapping of vacancies in small Mn clusters in iron. It is used to accelerate KMC measurements of Mn-cluster diffusion coefficients in Sec. IV.

III. ELASTO-DIFFUSION OF VACANCIES IN ALUMINIUM

In the first application, we consider the diffusion of a single vacancy around a cavity in Aluminium. The crystalline structure is face-centered cubic. The coordination number is $Z = 12$. The vacancy formation energy is $E_V^{\rm f} = 0.67\,{\rm eV}$. It corresponds to the energy difference for displacing a vacancy from a free surface to the bulk. In our model, it determines the interaction energy of two neighboring vacancies. We consider nearest-neighbor pair interactions only and set $E_{VV}^{\rm nn} = -2E_V^{\rm f}/Z$. We also set $E_{V\,{\rm Fe}}^{\rm nn} = 0$ and $E_{\rm FeFe}^{\rm nn} = 0$. These interaction energies entail that nearest-neighbor vacancy pairs are energetically favored. The model describing thermally activated jumps of Aluminum atoms into next nearest-neighbor vacancy accounts for the elastic field created by the cavity [37]. The dipole tensor associated with the vacancy has been computed using electronic structure calculations. It is assumed to be independent of the induced elastic strain, i.e., higher order terms in the fast-multiple expansion of elastic interactions are neglected.

A. Vacancy emission flux from cavity

We focus on the emission of a single vacancy from a cavity. In this set-up, we neglect some dynamical effects and assume that only a single vacancy can migrate and be emitted from a static cavity. The vacancy emission rate could conceivably be impacted by a dynamically evolving cavity. The cavity is composed first of 2243 vacancies. Cavity sites are located inside a sphere of radius 20.7 Å. The mobile vacancy is considered to be initially equilibrated at temperature $T = 600\,{\rm K}$ on the sites of the first shell of the cavity. Trapping results from the fast intra-shell vacancy jumps and from the immediate reconnection of the vacancy after it just disconnected from it. A total of $N = 259320$ transient states are used to characterize the vacancy emission properties, which correspond to the vacancy sites located inside the protective sphere of radius 101 Å and centered on the cavity center.

We first resorted to algorithm 1 to construct matrix $\mathbf{A}^B$. We next solved equation (25) by implementing a direct solver (see Sec. III B) to obtain the mean residence time vector $\boldsymbol{\theta}^{(N)}$ associated with the imposed initial distribution $\boldsymbol{\pi}$ and through appropriate rescaling (see Sec. II I). We next define the mean probability currents $\theta_j K_{j\ell} - \theta_\ell K_{\ell j}$ between both transient and absorbing states, where $\boldsymbol{\theta}$ stands for an extended mean residence time vector coinciding with $\boldsymbol{\theta}^{(N)}$ on transient states and canceling elsewhere. Probability currents are a practical tool serving to characterize not only nonequilibrium steady states [36, 42], but also transient nonequilibrium regimes as presently. Letting $\hat{\mathbf{r}}_i$ denote the three-dimensional lattice coordinates of the vacancy associated with state i, the vacancy flux at $\hat{\mathbf{r}}_i$ is defined as the three-dimensional vacancy current density:

$$\hat{\boldsymbol{\phi}}_j = \frac{1}{2v} \sum_\ell \left(\theta_j K_{j\ell} - \theta_\ell K_{\ell j}\right) \left(\hat{\mathbf{r}}_\ell - \hat{\mathbf{r}}_j\right).$$

The site volume v is uniform and the half factor stems from the fact that adjacent transitions are counted twice and entails an average of the fluxes entering and leaving lattice site $\hat{\mathbf{r}}_j$. Note that reversible dynamics obeying detailed balance, probability currents and vacancy fluxes vanish at equilibrium. The absorbing probability to state $\ell > N$ given initial distribution $\boldsymbol{\pi}$ is eventually computed by plugging the residence time vector into relation (27), which yields

$$\pi_\ell^a = \sum_{i \le N} \pi_i P_{i\ell}^{(N)} = \sum_{i \le N} \theta_i^{(N)} K_{i\ell}. \tag{36}$$

Residence times and vacancy fluxes are displayed in Fig. 2 for sites j located in the (001) plane intercepting the cavity center. Vacancy fluxes along [001] direction inside this particular (001) plane cancel due to the reflective system symmetry. The absorbing probabilities to the peripheral states are displayed in Fig. 3 for the emission of a single vacancy from the (hidden) cavity located at the center.

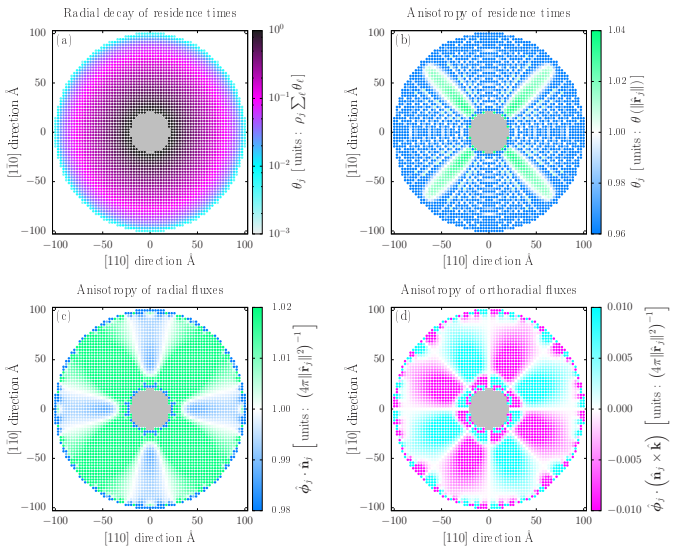


FIG. 2: Emission of a single vacancy from the centered gray cavity of radius 20.7 Å to the absorbing sphere of radius 101 Å: Panels (a) and (b) display the residence times using two distinct color scales; Panels (c) and (d) display the radial and orthoradial vacancy fluxes, respectively. Coordinates of displayed sites satisfy $\hat{\mathbf{r}}_j \cdot \hat{\mathbf{k}} = 0$ where $\hat{\mathbf{k}}$ is the normalized basis vector orthogonal to (001). The Euclidean norm $\|\hat{\mathbf{r}}_j\|$ corresponds to the distance to the cavity center and $\theta\left(\|\hat{\mathbf{r}}_j\|\right)$ is the average residence time on the centered sphere of radius $\|\hat{\mathbf{r}}_j\|$ and surface area $4\pi\|\hat{\mathbf{r}}_j\|^2$. Vector $\hat{\mathbf{n}}_j = \hat{\mathbf{r}}_j/\|\hat{\mathbf{r}}_j\|$ is the normalized radial vector. The cross product $\hat{\mathbf{n}}_j \times \hat{\mathbf{k}}$ corresponds to the orthoradial direction of the flux at $\hat{\mathbf{r}}_j$.

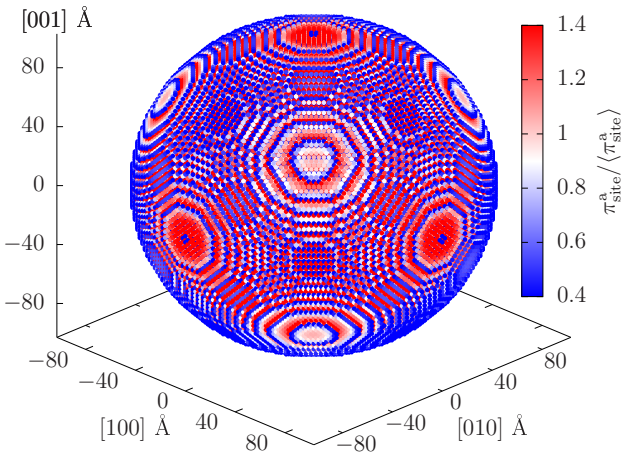


FIG. 3: Anisotropy of peripheral site absorption for vacancies emitted from a small Aluminum cavity. Absorbing probability vector is $\mathbf{p}^{aT} = \boldsymbol{\pi}^T\mathbf{P}^{(N)}$. Red and blue coloration respectively indicates larger and lower than average for the displayed site absorbing probabilities. The inverse mean probability corresponds to the number of peripheral sites. Cavity and protection radii are 20.7 Å and 101 Å, respectively.

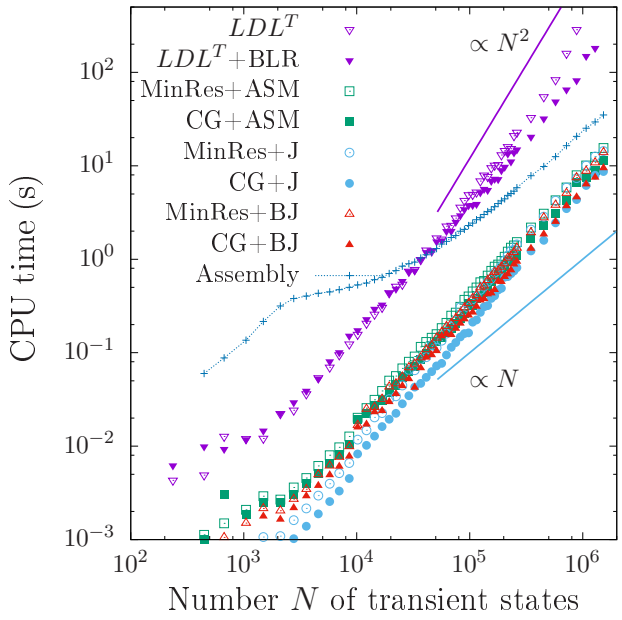


FIG. 4: Comparative cost of direct and iterative linear solvers. See text for details on the solvers. Simulations are performed on an Intel 4-Cores i5-4310U processor running at 2.00GHz with 8GB memory. OpenMP shared-memory parallelism is enabled for LDL^T factorization.

In Fig 2.a, we have scaled the residence times with respect to the corresponding equilibrium distribution for comparison. We observe that residence time distribution coincides with the equilibrium distribution on trapping sites located on the first two shells of the cavity. However, the former distribution become considerably smaller than the latter one as the vacancy moves away from the cavity. The emission anisotropy is clearly evidenced in Fig 2.b wherein the residence times have been rescaled with respect to their spherical averages. The vacancy preferentially resides along ⟨100⟩ crystalline directions. This trend induces an identical anisotropy of the radial flux observed in Fig. 2.c. Preferential emission paths along ⟨100⟩ crystalline directions may result from the orthoradial components of the flux that move the vacancy away from ⟨100⟩ directions beyond the first outer shells as evidenced in Fig. 2.d. The emission anisotropy is not due to the nearest-neighbor chemical interactions between the vacancy and the cavity but to the elastic interactions. This property is confirmed by the fact that isotropic fluxes are obtained when elastic interactions are switched off.

The strong local heterogeneity observed in Fig. 3 for the absorbing probabilities is attributed to the varying numbers of interconnections between transient and absorbing states. However, smaller modulations are clearly visible at larger scale indicating that absorption is more important along ⟨100⟩ directions and smaller along ⟨110⟩ directions, in agreement with the measured vacancy fluxes.

Further, we discuss about the computational aspects of the sparse linear solvers tested in the vacancy emission problem.

B. Computational complexity and stability

For problems involving the hopping of a defect on a lattice, the absorbing transition rate matrix is sparse and contains a maximum of $Z+1$ nonzero elements per rows ($Z = 12$), while its size may exceed several millions in practice. Such linear systems are efficiently solved using either a multi-frontal sparse direct solver based on LDL^T decomposition [43, 44] or sparse iterative solvers based on Krylov subspace projection (KSP) methods [45] based on PETSc software [46]. We first compare the costs of solving the linear system using iterative solvers to those of using direct solvers.

As for sparse iterative solvers, conjugate gradient (CG) is the appropriate method when the matrix is symmetric definite positive, however we also tested the minimum residual (MinRes) method that is adapted to general symmetric matrices. Three preconditioners were tested: the additive Schwarz method (ASM) and the Jacobi and Block-Jacobi methods (B and BJ, respectively).

As for the multi-frontal sparse LDL^T factorization, we installed Version 5.2 of MUMPS [43, 44], which enables OpenMP threading and implements several ordering packages to construct the elimination tree. Among them, we selected METIS for its efficiency. The block-low rank (BLR) compression of the factors was also tested. Enabling this option reduces both storage and number of operations by an amount inversely proportional to the tolerance on the solution. The tolerance control parameter was set to 10^{-5} which yielded a good trade-off between performance and accuracy. Simulations are per-

formed for the emission problem in which the emitted vacancy reaches a protective sphere of increasing radii. The largest matrix size is nearly 2.10^6 (number of transient states). Results are displayed in Fig. 4. Scaling matrix $\mathbf{B}$ is first set to identity.

We observe that the iterative solver performs better than the direct solver, by a factor of 10-20. We obtain a square complexity for the standard direct solver, as expected for a sparse matrix describing transitions or connections within a 3-dimensional space. BLR becomes more beneficial the larger the matrices. For 10^6 transient states, BLR is 2.5 times faster and requires 15 times less operations than the standard factorization. The observed complexity of the iterative solvers is between square and linear with the combination of conjugate gradient and Jacobi preconditioning being the most efficient.

CPU costs for computing the transition rates and assembling $\mathbf{A}^B$ matrix are also reported in Fig. 4. They are represented by the dotted line referred to as "Assembly". Asymptotically, the overhead cost grows linearly with system size. It however remains larger than the one taken by any iterative solver for all simulated sizes. The preliminary calculation of transition rates is in fact substantial and should certainly be optimized in future KMC applications.

KSP methods for sparse symmetric linear system [45] allow to solve first-passage problems over large volume, with matrix sizes that could not be investigated before. Note that implementing iterative KSP methods with the original non symmetric matrix, for instance resorting to generalized minimal residual method, increases the computational cost by one-to-two orders of magnitude and requires more memory by the same amount. The deterioration of efficiency results from the orthogonalization procedure that must be performed with respect to all Krylov basis vectors. At variance, with symmetric matrices orthogonalization is performed with respect to the two last vectors, omitting occasional re-orthogonalizations aiming at preserving accuracy.

Further, we investigate the effect of scaling matrix $\mathbf{B}$ on the condition number of $\mathbf{A}^B$, denoted by $\kappa\left(\mathbf{A}^B\right)$, and on the accuracy of the direct solver (LDL^T factorization with MUMPS). The 1-norm of the obtained residual vectors with respect to the absorbing rate matrix serves as a measure of accuracy. The 1-norms and condition numbers are evaluated for a series of temperatures ranging from 160 K to 600 K and are displayed in Fig. 5 wherein $\mathbf{\Delta}$ stands for $\mathrm{Diag}(\mathbf{A})^{-1}$ and $\mathbf{B}$ is set to $\mathbf{I}$ or $\mathbf{\Delta}$. We notice that the latter scaling variant improves mainly the numerical accuracy for mean first-passage time calculations. Besides, it systematically yields the lowest condition numbers $\kappa\left(\mathbf{A}^B\right)$. Note that the direct solver fails below 180 K, compared to 250 K for the best iterative solver (conjugate gradient). When the solvers fail to converge, the probability flux is not preserved and negative times may even be returned. As Wales *et al.* [27, 40], we believe that this issue is due to round-off errors and too large differences between the diagonal and non-diagonal

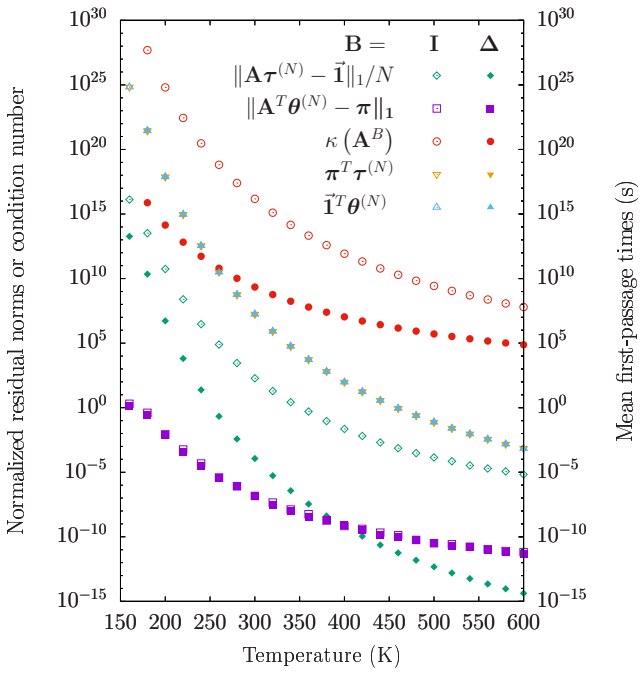


FIG. 5: Effect of scaling matrix $\mathbf{B}$ on accuracy of residence and first-passage time calculations. The residual norms and condition numbers are displayed for two setups: $\mathbf{B}$ equal to $\mathbf{I}$ and $\mathbf{\Delta} = \mathrm{Diag}(\mathbf{A})^{-1}$. The corresponding mean-first passage times, evaluated from $\boldsymbol{\pi}^T\boldsymbol{\tau}^{(N)}$ and $\boldsymbol{\theta}^T\vec{\mathbf{1}}$, are displayed for comparison. Note that $\|\vec{\mathbf{1}}\|_1 = N$ and $\|\boldsymbol{\pi}\|_1 = 1$.

elements. For stiff problems, like the one involving the diffusivity of Mn-clusters in IV, path factorization is to be implemented to guarantee that the special structure of the transition matrix is preserved during eliminations.

Another advantage of performing the factorization is that additional solutions can be obtained at a much smaller cost using forward/backward substitution. The factorization can be reused to compute mean residence times over each site given any new initial conditions or to generate first-passage times directly from the exact distribution through randomization. To validate the latter time randomization procedure, we show that it is possible to reconstruct the survival probability distribution from a sample of first-passage times.

C. Survival probabilities and spectral truncation

Equation (22) yields the survival probabilities at time t with respect to the set of initial distributions $\{\mathbf{e}_i\}_{i\leq N}$. For the particular initial distribution $\boldsymbol{\pi}$, the probability becomes $S(t) = \boldsymbol{\pi}^T\mathbf{p}^{\mathrm{s}}(t)$. Defining and plugging the scalar product $\alpha_h = \sum_{i\leq N}\pi_i a_i^h$ into the survival probability yields

$$S(t) = \sum_{h=1}^{N} \alpha_h \exp\left[-\lambda_h t\right]. \tag{37}$$

Considering the vacancy emission problem again, we entirely solved the eigenvalue problem for a small system containing 236 transient states and 13 immobile vacancies in the central cavity of radius 4.04 Å. The protective sphere radius is 10.1 Å. The default dense solver from Lapack library was used. The survival probability and the distribution of first-passage log-times are reported in Fig. 6 for reference. We next run Algorithm 1 and 2 of Appendix A to make the factorization of **BA** and Algorithm 3 to generate a sample of 10^5 first-passage time to the protective sphere. The survival probability distribution reconstructed from the generated sample of first-passage times is reported in Fig. 6. It perfectly matches with the reference distribution obtained from (37), which validates the time randomization approach.

Interestingly, a perfect agreement is also observed when the survival probability is evaluated retaining only the lowest eigenvalue associated with the quasi-stationary distribution. This suggests that it is possible to truncate the spectral decomposition above a certain threshold and approximate the survival probability retaining the k first terms:

$$S_k(t) = \sum_{h=1}^{k} \alpha_h \exp\left[-\lambda_h t\right]. \tag{38}$$

The truncation error can be directly quantified at time $t = 0$ since we know that $S(0) = 1$. The time-integral of the error can also be quantified from the ratio

$$\mathcal{T}_k = \frac{\sum_{h=1}^{k} \alpha_h/\lambda_h}{\sum_{h=1}^{N} \alpha_h/\lambda_h}, \tag{39}$$

where the denominator formally corresponds to the mean first-passage time $\boldsymbol{\pi}^T \boldsymbol{\tau}^{(N)}$ and is thus rather computed from a linear solve.

For the large considered systems, the k lowest eigenvalues and their associated eigenvectors are efficiently extracted by performing reverse iterations using the KrylovSchur method [47, 48] and the factored matrix. This amounts to extracting the largest eigenvalues of the inverse matrix. Calculations are performed using SLEPc software [49, 50]. We investigate the effect of truncating the spectral decomposition on two computational setups: (i) the emission of a single vacancy from a cavity of radius 20.7 Å to a protective sphere of radius 101 Å (same conditions as in Fig. 2 and 3) and (ii) the absorption by the cavity of a single vacancy initially located at a distance of 57.13 Å from cavity center in [110] direction. In setup (ii), there is no protective sphere and the cell is periodically replicated along ⟨100⟩ directions with periodicity length 80.8 Å. The absorbing macro-state is reached whenever the hopping vacancy becomes connected to the cavity. Setup (ii) entails 34801 transient states compared to 259320 for setup (i).

The survival probabilities and first-passage distributions are reported for various truncation threshold in Figs. 7 and 8. We observe that the trapping kinetics is

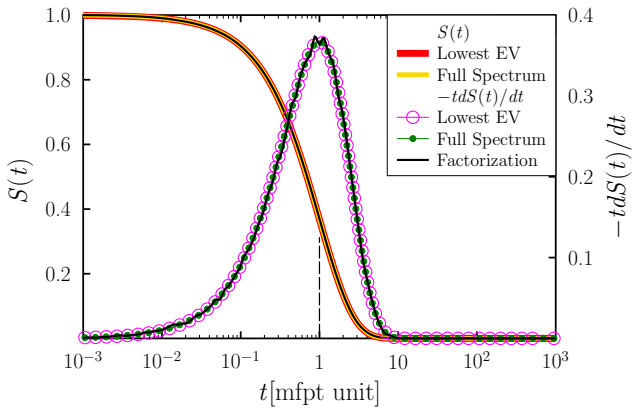


FIG. 6: Emission of a single vacancy from cavity (4.04 Å radius) to protective sphere (10.1 Å radius). The survival probabilities (left axis) and first-passage distributions (right axis) are evaluated from the quasi-stationary distribution (Lowest Eigenvalue), the exact distribution (Full Spectrum) and randomization (Factorization). Time is given in unit of mean first-passage time (mfpt).

governed by the quasi-stationary distribution [41] for the vacancy emission process, and for the vacancy absorption at times larger than the mean first-passage time. At short times, a substantial portion of 10% of the decaying exponentials needs to be included to faithfully reproduce the early stages of the absorption kinetics.

To rationalize this trend, the convergence of the truncated and reduced quantities $\mathcal{T}_k$ and $S_k(0)$ are displayed in panel (a) and (b) of Fig. 9, respectively. We observe that truncation errors are lower in the estimation of the mean first-passage time than of the initial survival probability. Convergence is non monotonous and proceeds by plateaus, suggesting the predominance of specific modes. To evidence them, scatter plots of computed eigenvalues and associated α_k factors are shown in Fig. 10. We indeed observe that many more modes with large eigenvalues contribute in the absorption problem compared to the emission one. Furthermore, the large spectral gap between λ_1 and λ_2 explains the fast time-decay of the truncation error on the survival probability. At times larger than the mean-first passage time, the quasi-stationary distribution is reached. The early stage absorption kinetics is the most problematic to compute from spectral decomposition because many modes contribute. In this situation, we observe that it is far more efficient to compute the probability vector $\pi_t^T = \pi_0^T \exp[-\mathbf{A}t]$ at a given time t using a krylov subspace method [51] for evaluating the application of a vector on a matrix function. Here, we applied the scaled initial probability to the exponential of $-t\mathbf{A}^I$ and then reverted the scaling as follows:

$$\pi_t = \left\{\exp\left[-t\mathbf{A}^I\right](\pi_0 \odot \mathbf{r})\right\} \oslash \mathbf{r} \tag{40}$$

This method however becomes less efficient than the truncated eigenvalue decomposition method as time increases. At half the mean first-passage time of the ab-

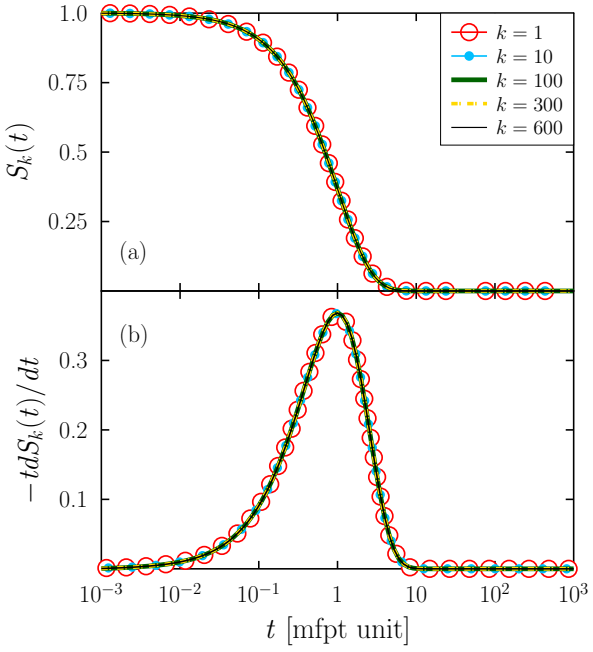


FIG. 7: Emission of a single vacancy from the cavity (20.5 Å radius) to the protective sphere (101 Å radius). Survival probabilities and first-passage distributions are respectively plotted in panels (a) and (b) using various truncation thresholds k.

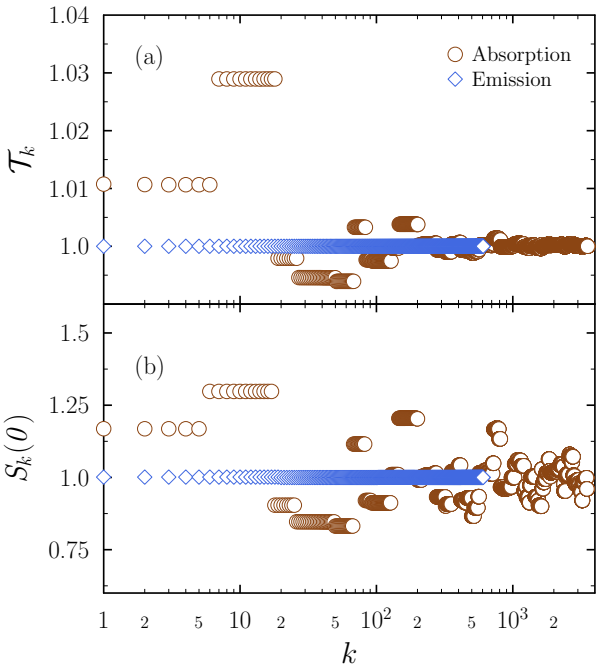


FIG. 9: Effect of retaining the k lowest modes for approximating the reduced mean first-passage time $\mathcal{T}_k$ [panel (a)] and the initial survival probability [panel (b)].

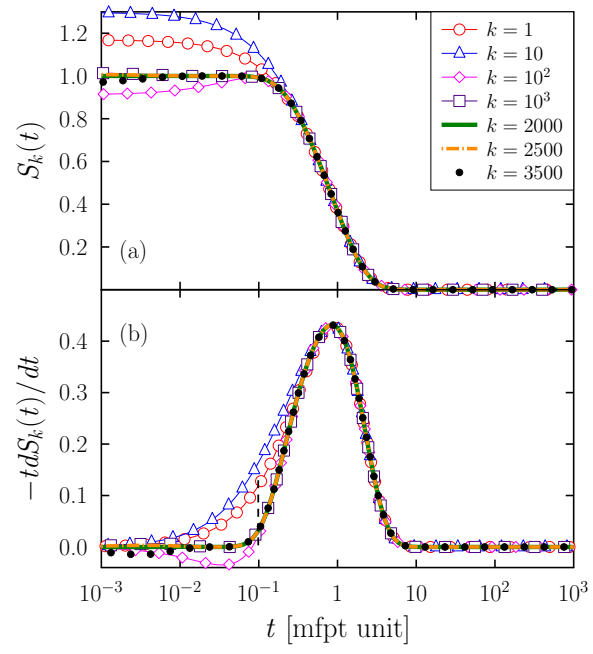


FIG. 8: Absorption of a single vacancy by the cavity (20.5 Å radius). Survival probability distribution [panel (a)] and first-passage distribution [panel (b)] as a function of time.

sorption kinetics, the QSD already yields an excellent approximation. This one is less efficiently extracted using the forward iterations of Ref. [51] than reverse iterations within Krylov-Schur method. The open question to address is how to combine both approaches optimally.

Note that the absorption kinetics is paradoxically easier to simulate using KMC simulations because the energetic basin of attraction is precisely the absorbing sink. Extensive KMC simulations have been performed for the present absorption problem in Ref [37] for calculating sink strengths of various cavities and dislocations.

IV. DIFFUSION OF MN-V CLUSTERS IN α-IRON

In this second application, we illustrate how path factorization can be implemented in kinetic Monte Carlo simulations to compute diffusion coefficients in FeMn system and how additional simulation speedups can be obtained by storing and efficiently retrieving the factorizations in hash tables. Simulation aims at computing the diffusivity of small Mn clusters. The enhanced mobility of solute clusters impacts the early stage of phase separation kinetics in quenched alloys [52], and is also suspected to be responsible for the anomalous incubation times observed in some Aluminium commercial alloys [53].

The simulation box contains 10^3 unit cells with two

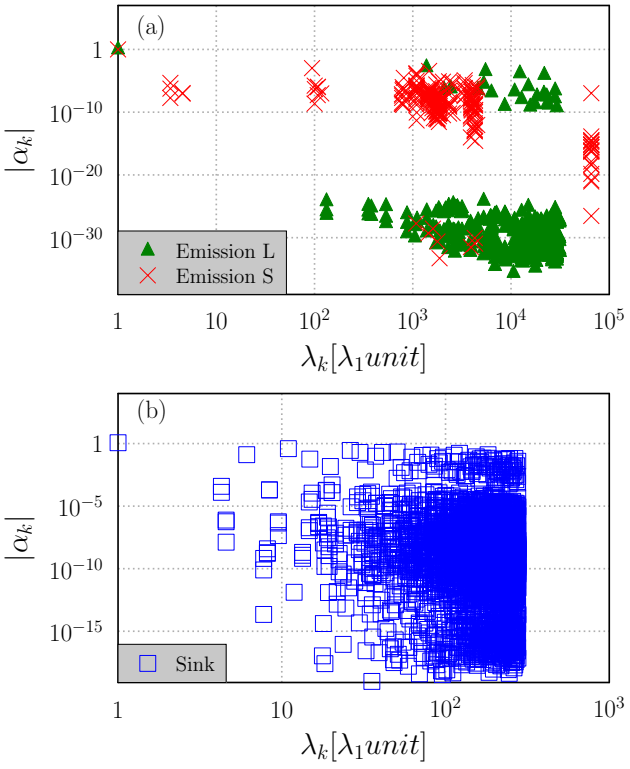


FIG. 10: Absolute values of α_k contributing coefficients displayed as a function of the eigenvalues scaled relatively to the lowest eigenvalue λ_1.

nodes per cell. The crystalline structure is body centered cubic and periodic boundary conditions are used. Interaction energies of Fe and Mn atoms and vacancies have been deduced from electronic structure calculations and are given in Ref. [54, 55]. Below 700 K, Mn atoms tend to form a single cluster that rarely dissociates during the simulations. This is due to their thermodynamic stability and to the high emission barriers. The two following algorithms are implemented and tested:

- The standard kinetic Monte Carlo algorithm denoted by KMC: At each cycle, a single vacancy transitions to one of its nearest neighbour sites, i.e. exchanges with a nearest neighbour atom. Time is incremented by the mean residence time on the previously occupied site;

- The factorized KMC algorithm denoted by F-KMC: The vacancy makes a non-local transition and escapes the trapping basin based on the path factorization algorithm. The set of transient states (the trap) encompasses the initial vacancy state and all states that can be reached via vacancy-Mn exchanges exclusively. The physical time is increased by the mean first passage time associated with the non-local escaping transition. It corresponds to the kinetic path sampling algorithm of Ref. [22].

Because a single vacancy is used in simulations, the time for performing a transition does not need to be drawn in its first-passage distribution and its expected value is used. This amounts to performing conditioning over time [56] and aims at reducing the statistical variance of the estimated diffusion coefficients. We consider here the diffusion coefficient of solute Mn atoms, defined as the three-dimensional average of half the time derivative of the mean square displacement (MSD)

$$\mathrm{D}(X,T) = \frac{1}{6}\lim_{t\to\infty}\frac{\mathrm{d}}{\mathrm{d}t}\,\langle\|\hat{\mathbf{x}}(t)-\hat{\mathbf{x}}(0)\|^2\rangle \qquad (41)$$

where X is the number of Mn atoms, T is temperature and $\hat{\mathbf{x}}(t)$ is the solute displacement vector at time t. With non-local events and conditioning performed over time, the time variable is replaced by the product of ℓ, the number of involved jumps, and $\widehat{\tau}^L = \frac{1}{L}\sum_{h=1}^{L}\tau_h$, the mean first-passage time averaged over a sample of size L generated using KMC or F-KMC. The solute diffusion coefficient is then estimated resorting to the following estimator

$$\widehat{D}_\ell^L = \frac{\frac{1}{L-\ell}\sum_{h=1}^{L-\ell}\|\hat{\mathbf{x}}_{h+\ell}-\hat{\mathbf{x}}_h\|^2}{6\ell\widehat{\tau}^L} \qquad (42)$$

where $\hat{\mathbf{x}}_{\ell+h}$ is the solute displacement vector after $\ell+h$ jumps.

Simulations are carried out for temperatures T ranging from 300 K to 1200 K and numbers X of Mn atoms increasing from 1 to 60. For each (X,T) pair, a series of ten runs of eight hours are performed using a Gold-6140 Intel Xeon processor running at 2.30 GHz. The computed diffusion coefficients and their average over the 10 runs are displayed in Fig. 11. For better visualization, a rescaling has been done using the high temperature activation energy for Mn monomer diffusion ($X = 1$) at 600 K. The diffusivity of V-Mn_X clusters increases with increasing X before tapering off for all temperatures lower than 800 K. Furthermore, the diffusivity maximum increases with temperature, suggesting the presence of a maximum at Mn content that could not be simulated. A similar increase trend has been reported in FeCu system using standard KMC simulations (see Fig.9 in Ref. [57]), although the temperature dependence of the maximum could not be investigated due to a severe vacancy trapping in Cu clusters.

With increasing Mn content, F-KMC simulations failed to converge. This is because path factorization becomes too costly. Besides, at the lower temperatures, the system gets trapped in super-basins containing several vacancy-cluster shapes. To understand the limitations of the current F-KMC algorithm and quantify the potential speedups of future developments, we tested two additional features in F-KMC algorithm, namely

- A dictionary-enhanced version denoted by FD-KMC: Hash tables are used to store computed data

about cluster shapes and factorizations. The goal is to retrieve the stored information when needed to avoid performing the same factorizations many times. This algorithm is described in chapters 5 and 6 of Ref. [58].

- A graph-enhanced version denoted by FDG-KMC: The nodes of the graph correspond to the cluster keys that have been added to the dictionary. The edges of the graph correspond to the previously encountered non-local transitions. The goal is to save computational resources by making transitions from one cluster shape to another one in the graph without recalculating the cluster key. This algorithm is described in chapter 7 of Ref. [58].

Simulations at 600 K with increasing Mn cluster sizes have been performed using the four algorithms and their relative efficiencies are displayed in Fig. 12. The efficiency of KMC algorithm relative to F-KMC is observed to decrease with increasing Mn content. This trend already reported in [22] for FeCu is attributed to the increase of vacancy trapping with cluster size. The observed increase of FD-KMC efficiency with increasing Mn content is explained by the concomitant increase in factorization costs: it is more and more advantageous to store and retrieve the factored matrices, as their sizes and computational costs increase. When cluster sizes exceed 40, it is also beneficial to connect the various clusters resorting to a graph using FGD-KMC algorithm. This trend results from the fact that the kinetics repeatedly visit a few cluster shapes, as observed in FeCu system [22]. These simulations show that the use of hash tables and graphs are also beneficial for KMC simulations on a rigid lattices. A point left for future developments involves the optimal deallocation of entries rarely looked up whenever the dictionary memory reaches a given threshold.

V. DISCUSSION AND CONCLUSION

To summarize, the theory of absorbing Markov chains is applied to characterize rare events occurring when the diffusion process is trapped within a finite set of states. The initial probability distribution corresponding to the trapped system acts as an emitting source while the peripheral states of the trap become an artificial absorbing sink. In this framework, the theory yields formal expressions for the transient evolution operator, the source-to-sink probability fluxes and the mean residence times on transient states. Whenever the original non-absorbing diffusion process is reversible, we show that the associated first-passage problem can be greatly simplified. The absorbing process then inherits a reversibility property that is conditional on that the dynamics has not reached the absorbing state. It transiently satisfies Kolmogorov's criterion: the probability of any circular sequence of transient states is equal to that of the time-reversed sequence,

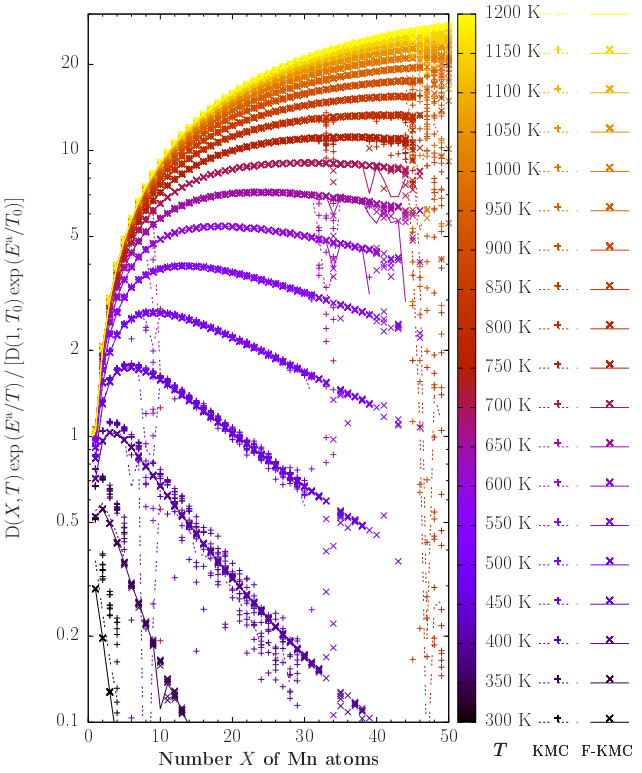


FIG. 11: Diffusion coefficients of Mn solute atoms as a function of cluster size for various temperatures. For clarity, scaling is done relative to monomer diffusivity at T_0 and its high temperature activation energy $E^a = 0.646$ eV. Reference temperature is $T_0 = 600$ K and reference diffusivity is $D(1, T_0) = 6.716 \cdot 10^{-15} m^2/s$ at vacancy concentration of 5‰.

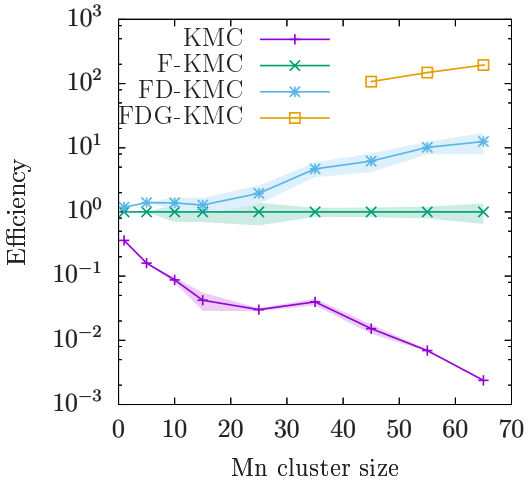


FIG. 12: Relative efficiencies of the algorithms as a function of cluster size. Efficiency is evaluated as the ratio of the simulated physical time to that of F-KMC for a given wall-clock time. Filling around curves corresponds to the 68% confidence interval.

even though probability currents are nonzero. This conditional reversibility entails in particular that the absorbing transition rate matrix is similar to a symmetric definite negative matrix and that the transformation matrix exhibits a simple diagonal form.

To compute mean first-passage times, exit probabilities and source-to-sink probability fluxes in relatively small problems, with less than 10^3 transient states, we implemented the path factorization technique. This technique corresponds to a direct and robust method for solving linear problems based on Gaussian elimination. Its robustness stems from the preservation of the transition rate property during the graph transformation [27, 40], i.e. the rank-one updates. Besides, by interpreting the factorization in terms of paths, a randomization procedure is formulated enabling first-passage times and exits to be drawn directly from the exact distributions. The acceleration in KMC simulations employing path factorization is substantial and makes it possible to compute the mobility of kinetically stable Mn clusters in iron down to the operating temperatures of pressurized reactors/vessels.

Characterizing the transient distribution associated with any absorbing processes is also an important problem to solve, as no-passage distributions serve to synchronize defects evolving in parallel within first-passage KMC simulations [33–35]. Evaluating the distribution requires the knowledge of both eigenvalues and eigenvectors of a symmetric positive definite matrix. In practical applications, we observe that the evolution on transient state is governed only by a fraction of the eigenspectrum. The most contributing mode is the one possessing the lowest eigenvalue, and its eigvenvector corresponds to the quasi-stationary distribution. For the studied emission problem, which exhibits a strong energetic trapping, the quasi-stationary distribution overwhelmingly predominates and correctly describes the transient evolution. For the absorption problem where trapping is essentially entropic, a small but substantial fraction of the slow modes are observed to contribute to the no-passage distribution and to govern the slow decay of the survival probability. For problems with more than 10^4 transient states, the transition rate matrix becomes sparse and iterative solvers are to be used to evaluate the eigenvalues and their contributions to no-passage distributions on the fly. We advocate to perform reverse iterations for extracting eigenvalues in ascending order. This can be achieved by iteratively applying the inverted matrix resorting to the LDL^T factorization, which amounts to extracting the largest eigenvalues of the inverted matrix. For very large trap sizes, we show that resorting to a direct multi-frontal LDL^T solver (possibly combined with block low-rank compression) makes it possible to perform sink strength calculations for the absorption of a vacancy from a cavity and also to compute vacancy emission rates from the cavity. Computations can be done using millions of transient states per processor, allowing us to investigate realistic cavity concentrations in irradiated or quenched Aluminum.

To conclude, symmetrizing the transition rate matrix associated with reversible diffusion processes enables one to apply efficient linear and eigenvalue solvers and to characterize the important rare events governing the long-term microstructural evolution of alloys, such as cluster mobilities, sink strengths and associated first-passage distributions. All these physical quantities are crucial input parameters for larger-scale simulations employing object/event KMC methods or rate equation cluster dynamics. The dependence of cluster mobilities on their size and temperature can possibly be included in larger scale models.

Acknowledgments

Fruitful discussions with Patrick Amestoy, Pascal Bellon, Vasily Bulatov, Olivier Boiteau, Jérôme Creuze, Christophe Domain, Luca Messina, Baptiste Pannier, Frédéric Soisson and Dallas Trinkle are gratefully acknowledged.

[1] A. F. Voter, in *Radiation Effects in Solids*, edited by K. E. Sickafus, E. A. Kotomin, and B. P. Uberuaga (Dordrecht: Springer Netherlands, 2007), vol. 235 of *NATO Science Series II: Mathematics, Physics and Chemistry*, chap. 1, pp. 1–23.

[2] J.-M. Lanore, *Simulation de l'évolution des défauts dans un réseau par le méthode de monte-carlo*, Radiation Effects **22**, 153 (1974).

[3] J. Dalla Torre, J.-L. Bocquet, N. V. Doan, E. Adam, and A. Barbu, *Jerk, an event-based kinetic monte carlo model to predict microstructure evolution of materials under irradiation*, Philosophical Magazine **85**, 549 (2005).

[4] P. Terrier, M. Athènes, T. Jourdan, G. Adjanor, and G. Stoltz, *Cluster dynamics modelling of materials: A new hybrid deterministic/stochastic coupling approach*, Journal of Computational Physics **350**, 280 (2017).

[5] N. Soneda, S. Ishino, A. Takahashi, and K. Dohi, *Modeling the microstructural evolution in bcc-fe during irradiation using kinetic monte carlo computer simulation*, Journal of Nuclear Materials **323**, 169 (2003), proceedings of the Second IEA Fusion Materials Agreement Workshop on Modeling and Experimental Validation.

[6] L. Malerba, C. S. Becquart, and C. Domain, *Object kinetic monte carlo study of sink strengths*, Journal of Nuclear Materials **360**, 159 (2007).

[7] M. Athènes, P. Bellon, G. Martin, and F. Haider, *A monte-carlo study of {B2} ordering and precipitation via vacancy mechanism in b.c.c. lattices*, Acta Materialia **44**, 4739 (1996).

[8] D. Mason, T. Hudson, and A. Sutton, *Fast recall of state-history in kinetic monte carlo simulations utilizing the zobrist key*, Computer Physics Communications **165**, 37

(2005).
[9] O. Trushin, A. Karim, A. Kara, and T. S. Rahman, *Self-learning kinetic monte carlo method: Application to cu(111)*, Phys. Rev. B **72**, 115401 (2005).
[10] F. El-Mellouhi, N. Mousseau, and L. J. Lewis, *Kinetic activation-relaxation technique: An off-lattice self-learning kinetic monte carlo algorithm*, Phys. Rev. B **78**, 153202 (2008).
[11] A. Kara, O. Trushin, H. Yildirim, and T. S. Rahman, *Off-lattice self-learning kinetic monte carlo: application to 2d cluster diffusion on the fcc(111) surface*, Journal of Physics: Condensed Matter **21**, 084213 (2009).
[12] L. K. Béland, P. Brommer, F. El-Mellouhi, J.-F. Joly, and N. Mousseau, *Kinetic activation-relaxation technique*, Phys. Rev. E **84**, 046704 (2011).
[13] G. Nandipati, A. Kara, S. I. Shah, and T. S. Rahman, *Off-lattice pattern recognition scheme for kinetic monte carlo simulations*, Journal of Computational Physics **231**, 3548 (2012).
[14] S. I. Shah, G. Nandipati, A. Kara, and T. S. Rahman, *Extended pattern recognition scheme for self-learning kinetic monte carlo simulations*, Journal of Physics: Condensed Matter **24**, 354004 (2012).
[15] G. Nandipati, N. Govind, A. Andersen, and A. Rohatgi, *Self-learning kinetic monte carlo simulations of al diffusion in mg*, Journal of Physics: Condensed Matter **28**, 155001 (2016).
[16] M. Trochet, A. Sauvé-Lacoursière, and N. Mousseau, *Algorithmic developments of the kinetic activation-relaxation technique: Accessing long-time kinetics of larger and more complex systems*, The Journal of Chemical Physics **147**, 152712 (2017).
[17] M. Trochet, N. Mousseau, L. K. Béland, and G. Henkelman, *Off-Lattice Kinetic Monte Carlo Methods* (Springer International Publishing, Cham, 2019), pp. 1–29, URL `https://doi.org/10.1007/978-3-319-42913-7_29-1`.
[18] M. A. Novotny, *Monte carlo algorithms with absorbing markov chains: Fast local algorithms for slow dynamics*, Phys. Rev. Lett. **74**, 1 (1995).
[19] M. Athènes, P. Bellon, and G. Martin, *Identification of novel diffusion cycles in b2 ordered phases by monte carlo simulation*, Philosophical Magazine A **76**, 565 (1997).
[20] S. Redner, *A guide to first-passage processes* (Cambridge University Press, Cambridge, 2001).
[21] G. Nandipati, Y. Shim, and J. G. Amar, *First-passage time approach to kinetic monte carlo simulations of metal (100) growth*, Phys. Rev. B **81**, 235415 (2010).
[22] M. Athènes and V. V. Bulatov, *Path factorization approach to stochastic simulations*, Physical Review Letters **113**, 230601 (2014).
[23] D. Gillespie, *Markov Processes: An Introduction for Physical Scientists* (Elsevier Science, 1991).
[24] M. A. Novotny, *A tutorial on advanced dynamic Monte Carlo methods for systems with discrete state spaces* (WORLD SCIENTIFIC, 2001), pp. 153–210, URL `http://www.worldscientific.com/doi/abs/10.1142/9789812811578_0003`.
[25] B. Puchala, M. L. Falk, and K. Garikipati, *An energy basin finding algorithm for kinetic monte carlo acceleration*, The Journal of Chemical Physics **132**, 134104 (2010).
[26] S. A. Trygubenko and D. J. Wales, *Graph transformation method for calculating waiting times in markov chains*, The Journal of Chemical Physics **124**, 234110 (2006).
[27] D. J. Wales, *Calculating rate constants and committor probabilities for transition networks by graph transformation*, The Journal of Chemical Physics **130**, 204111 (2009).
[28] J. G. Mullen, *Effect of bardeen-herring correlation on vacancy diffusion in anisotropic crystals*, Phys. Rev. **124**, 1723 (1961).
[29] J. R. Manning, *Correlation factors for impurity diffusion. bcc, diamond, and fcc structures*, Phys. Rev. **136**, A1758 (1964).
[30] J. Bocquet, *Correlation factor for diffusion in cubic crystals with solute–vacancy interactions of arbitrary range*, Philosophical Magazine **94**, 3603 (2014).
[31] D. R. Trinkle, *Automatic numerical evaluation of vacancy-mediated transport for arbitrary crystals: Onsager coefficients in the dilute limit using a green function approach*, Philosophical Magazine **97**, 2514 (2017).
[32] D. R. Trinkle, *Variational principle for mass transport*, Phys. Rev. Lett. **121**, 235901 (2018).
[33] T. Opplestrup, V. V. Bulatov, G. H. Gilmer, M. H. Kalos, and B. Sadigh, *First-passage monte carlo algorithm: Diffusion without all the hops*, Phys. Rev. Lett. **97**, 230602 (2006).
[34] T. Oppelstrup, V. V. Bulatov, A. Donev, M. H. Kalos, G. H. Gilmer, and B. Sadigh, *First-passage kinetic monte carlo method*, Phys. Rev. E **80**, 066701 (2009).
[35] A. Donev, V. V. Bulatov, T. Oppelstrup, G. H. Gilmer, B. Sadigh, and M. H. Kalos, *A first-passage kinetic monte carlo algorithm for complex diffusion–reaction systems*, Journal of Computational Physics **229**, 3214 (2010).
[36] R. K. P. Zia and B. Schmittmann, *Probability currents as principal characteristics in the statistical mechanics of non-equilibrium steady states*, Journal of Statistical Mechanics: Theory and Experiment **2007**, P07012 (2007).
[37] D. Carpentier, T. Jourdan, Y. L. Bouar, and M.-C. Marinica, *Effect of saddle point anisotropy of point defects on their absorption by dislocations and cavities*, Acta Materialia **136**, 323 (2017).
[38] Y. Shim and J. G. Amar, *Rigorous synchronous relaxation algorithm for parallel kinetic monte carlo simulations of thin film growth*, Phys. Rev. B **71**, 115436 (2005).
[39] E. Martínez, P. Monasterio, and J. Marian, *Billion-atom synchronous parallel kinetic monte carlo simulations of critical 3d ising systems*, Journal of Computational Physics **230**, 1359 (2011).
[40] J. D. Stevenson and D. J. Wales, *Communication: Analysing kinetic transition networks for rare events*, The Journal of Chemical Physics **141**, 041104 (2014).
[41] G. Di Gesu, T. Lelievre, D. Le Peutrec, and B. Nectoux, *Jump markov models and transition state theory: the quasi-stationary distribution approach*, Faraday Discuss. **195**, 469 (2016).
[42] M. Picciani, M. Athènes, J. Kurchan, and J. Tailleur, *Simulating structural transitions by direct transition current sampling: The example of LJ(38)*, Journal of Chemical Physics **135** (2011).
[43] P. R. Amestoy, I. S. Duff, J. Koster, and J.-Y. L'Excellent, *A fully asynchronous multifrontal solver using distributed dynamic scheduling*, SIAM Journal on Matrix Analysis and Applications **23**, 15 (2001).
[44] P. R. Amestoy, A. Guermouche, J.-Y. L'Excellent, and S. Pralet, *Hybrid scheduling for the parallel solution of linear systems*, Parallel Computing **32**, 136 (2006).
[45] Y. Saad, *Iterative Methods for Sparse Linear Systems*

(2003), chap. 6. and 7. Krylov Subspace Methods, Part I and II, pp. 151–244, URL https://epubs.siam.org/doi/abs/10.1137/1.9780898718003.ch6.

[46] S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, V. Eijkhout, W. D. Gropp, D. Kaushik, et al., *PETSc users manual*, Tech. Rep. ANL-95/11 - Revision 3.9, Argonne National Laboratory (2018).

[47] Y. Saad, *Numerical Methods for Large Eigenvalue Problems* (2011), chap. 6. Krylov Subspace Methods, pp. 125–162, URL https://epubs.siam.org/doi/abs/10.1137/1.9781611970739.ch6.

[48] P. Arbenz, *Lecture notes on solving large scale eigenvalue problems*, Tech. Rep., Computer Science Department, ETH Zürich (2016).

[49] V. Hernandez, J. E. Roman, and V. Vidal, *SLEPc: A scalable and flexible toolkit for the solution of eigenvalue problems*, ACM Trans. Math. Software **31**, 351 (2005).

[50] J. E. Roman, C. Campos, E. Romero, and A. Tomas, *SLEPc users manual*, Tech. Rep. DSIC-II/24/02 - Revision 3.11, D. Sistemes Informàtics i Computació, Universitat Politècnica de València (2019).

[51] M. Eiermann and O. Ernst, *A restarted krylov subspace method for the evaluation of matrix functions*, SIAM Journal on Numerical Analysis **44**, 2481 (2006).

[52] M. Athènes, P. Bellon, and G. Martin, *Effects of atomic mobilities on phase separation kinetics: a monte-carlo study*, Acta Materialia **48**, 2675 (2000).

[53] S. Pogatscher, H. Antrekowitsch, M. Werinos, F. Moszner, S. S. A. Gerstl, M. F. Francis, W. A. Curtin, J. F. Löffler, and P. J. Uggowitzer, *Diffusion on demand to control precipitation aging: Application to al-mg-si alloys*, Phys. Rev. Lett. **112**, 225701 (2014).

[54] L. Messina, M. Nastar, T. Garnier, C. Domain, and P. Olsson, *Exact ab initio transport coefficients in bcc* Fe $-$ *x (x =* Cr*,* Cu*,* Mn*,* Ni*, p,* Si*) dilute alloys*, Phys. Rev. B **90**, 104203 (2014).

[55] L. Messina, L. Malerba, and P. Olsson, *Stability and mobility of small vacancy–solute complexes in fe–mnni and dilute fe–x alloys: A kinetic monte carlo study*, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms **352**, 61 (2015), proceedings of the 12th International Conference on Computer Simulation of Radiation Effects in Solids, Alacant, Spain, 8-13 June, 2014.

[56] M. Athènes, *Conditioning and enhanced sampling schemes for simulating thermodynamic and kinetic properties of condensed matter*, Habilitation à diriger des recherches, Université Paris Saclay ; Université Paris Sud (2018), URL https://hal-cea.archives-ouvertes.fr/tel-01851686.

[57] F. Soisson and C.-C. Fu, *Cu-precipitation kinetics in* $\alpha-$Fe *from atomistic simulations: Vacancy-trapping effects and cu-cluster mobility*, Phys. Rev. B **76**, 214102 (2007).

[58] T. Vanacker, *Improvement of an advanced kinetic monte carlo algorithm through storing and recycling factorized transition matrices*, Master's thesis, Ecole Centrale Paris and KTH (2017), URL oai:DiVA.org:kth-239920.

[59] D. Mason, R. Rudd, and A. Sutton, *Stochastic kinetic monte carlo algorithms for long-range hamiltonians*, Computer Physics Communications **160**, 140 (2004).

Appendix A: Algorithmic implementation

1. Path factorization

Algorithm 1 is used to compute transition rates, to discriminate transient and absorbing states and to assemble the associated transition sub-matrices. Then, algorithm 2 is used to make the path factorization. Note that the characterization of transition rates and transient states can be done on the fly in algorithm 2. This requires a selection rule for next transient state based on the transformed transition probabilities, as done in Ref. [22] for simulating the anomalous diffusion of a defect on a disordered substrate.

Algorithm 1 Assembly of transition sub-matrices.

$N \leftarrow 1$; $N^{\rm tot} \leftarrow 1$;
initial transient state is indexed 1; $i \leftarrow N$;
while $i \leq N$ **do**
  list the Z_i possible transitions of S_i;
  for $\ell = 1, \cdots, Z_i$ **do**
    evaluate key for final state associated with ℓ-th transition;
    if new key **then**
      $N^{\rm tot} \leftarrow N^{\rm tot} + 1$; $j \leftarrow N^{\rm tot}$;
      add key and its state index j to dictionary;
      if state transient **then**
        $N \leftarrow N + 1$;
      end if
    else ▷ key exists
      retrieve state index j of existing key;
    end if
    evaluate K_{ij}, transition rate from S_i to S_j; ▷ for ℓ-th listed transition
  end for
  $i \leftarrow i + 1$
end while
re-order transient states from 1 to N and absorbing states from $N + 1$ to $N^{\rm tot}$;
construct $\boldsymbol{\tau}^{(0)}$, $\mathbf{P}^{(0)}$ and absorbing transient rate matrix $\mathbf{A}^B$ from $\mathbf{K}$;

2. Spacetime randomization

At the Nth rank-one update, stochastic probability matrix $\mathbf{P}^{(N)}$ subsumes all possible transitions involving the deleted states in the trapping basin $\mathbb{E} = \{1, \cdots, N\}$. Although any intermediate matrix $\mathbf{P}^{(n)}$ with $n \in \mathbb{E}$ can be used to randomly generate escapes from any state $i \in \mathbb{E}$, a trajectory generated using $\mathbf{P}^{(N)}$ is the simplest containing a single transition. On the other end, reverting back to a standard KMC simulation based on $\mathbf{P}^{(0)}$, a detailed escape trajectory that accounts for all transitions within $\mathbb{E}$ can be generated. Remarkably, it

Algorithm 2 path factorization [22] adapted from graph transformation [27] and an early version [19].

construct $\mathbf{P}^{(0)}$ and $\boldsymbol{\tau}^{(0)}$;
if flux enabled **then**
  $\boldsymbol{\theta}^{(0)} = \boldsymbol{\tau}^{(0)} \odot \boldsymbol{\pi}$;
end if
for n from 1 to N **do**
  if *adaptation enabled* **then**
    select new transient state and label it n
    re-order $\mathbf{P}^{(n)}$, $\boldsymbol{\tau}^{(n)}$ and possibly $\boldsymbol{\theta}^{(n)}$;
  end if
  compute $\mathbf{P}^{(n)}$ by performing rank-one update of $\mathbf{P}^{(n-1)}$;
  compute $\boldsymbol{\tau}^{(n)}$ by performing rank-one update of $\boldsymbol{\tau}^{(n-1)}$;
  if flux enable **then**
    compute $\boldsymbol{\theta}^{(n)}$ by performing rank-one update of $\boldsymbol{\theta}^{(n-1)}$;
  end if
end for
if *flux calculation enabled* **then**
  compute fluxes from $\boldsymbol{\theta}^{(n)}$
end if

is possible to efficiently construct statistically correct escape trajectories without ever performing any detailed (and inefficient) KMC simulation. Space-time randomization is based from the set of conditional probabilities defined for all i and for $j > n$

$$R_{ij}^{(n)} = P_{ij}^{(n-1)} / P_{ij}^{(n)}. \tag{A1}$$

A particular $R_{ij}^{(n)}$ yields the probability that an nth order transition from i to $j > n$ avoids site n when decomposed in term of $(n-1)$th order transitions. Prior to describing space-time randomization, the following preliminary definitions are required. The binomial law of trial number $h \in \mathbb{N}$ and success probability r is denoted by $\mathrm{B}(h, r)$. The probability of s successes is $\binom{h}{s} r^s (1-r)^{h-s}$. The negative binomial law of success number h and success (escape) probability $1-p$ is denoted by $\mathrm{NB}(h, 1-p)$. The probability of f failures before the h-th success is $\binom{f+h-1}{f} p^f (1-p)^h$ where p is the failure or flicker probability (flickers will correspond to round-trips from a given state). The gamma law of shape parameter h and time-scale τ is denoted by $j(h, \tau)$. C_α denotes the categorical laws whose probability vector is the α-th row of $\mathbf{P}^{(N)}$ if $\alpha \leq N$ or of the stochastic matrix obtained from $\mathbf{P}^{(0)}$. The symbol $\sim$ means "is a random variate distributed according to the law that follows". Let $\mathbb{A}$ denote the set of absorbing peripheral states. The set of states beyond the peripheral states (that are non transient and non absorbing) is $\overline{\mathbb{A} \cup \mathbb{E}}$, the complementary of the union of $\mathbb{A}$ and $\mathbb{E}$. State α denote the current state of the system.

After implementation of Algorithm 3, the system has moved beyond the peripheral set and is disconnected to the trapping basin reached: the current state α belongs to $\overline{\mathbb{E} \cup \mathbb{A}}$ in item (6). The gamma law $j\left(T_n, \tau_n^{(0)}\right)$ in (23)

Algorithm 3 Kinetic path sampling [22]: subset $\mathbb{E} = \{1, \cdots, N\}$ encompasses the transient states and subset $\mathbb{A} = \{N+1, \cdots, N_{\mathrm{tot}}\}$ includes absorbing peripheral states; system is initially in state $\alpha \in \mathbb{E} \cup \mathbb{A}$

define $N_c \times N_c$ hoping matrix $\mathbf{H}^{(N)}$ and set its entries to zero;
while $\alpha \in \mathbb{A} \cup \mathbb{E}$ **do**
  draw $j \propto \mathrm{C}_\alpha$;
  increment $H_{\alpha,j}^{(N)}$ by one;
  $\alpha \leftarrow j$; ▷ move current state α to j
end while ▷ $\alpha \in \overline{\mathbb{A} \cup \mathbb{E}}$
for $n = N$ to 1: **do**
  evaluate $\mathbf{P}^{(n-1)}$;
  deallocate $\mathbf{P}^{(n)}$;
  for $i \in \{\mathbb{E} \cup \mathbb{A}\} \setminus \{n\}$ and $j \in \{n+1, ..., N_c\}$ **do**
    evaluate $R_{ij}^{(n)} = P_{ij}^{(n-1)} / P_{ij}^{(n)}$
    draw $H_{ij}^{(n-1)} \sim \mathrm{B}\left(H_{ij}^{(n)}, R_{ij}^{(n)}\right)$
  end for
  for $i \in \{\mathbb{E} \cup \mathbb{A}\} \setminus \{n\}$ **do** ▷ count new hops from i to n:
    $H_{in}^{(n-1)} = \sum_{j \in \{n+1, \ldots, N_c\}} H_{ij}^{(n)} - H_{ij}^{(n-1)}$
  end for
  for $j \in \{n+1, ..., N_c\}$ **do** ▷ count hops from n to j
    $H_{nj}^{(n-1)} = H_{nj}^{(n)} + \sum_{i \in \mathbb{E} \setminus \{n\}} H_{ij}^{(n)} - H_{ij}^{(n-1)}$
  end for
  $h_n = \sum_{j \in \{n+1, \ldots, N_c\}} H_{nj}^{(n-1)}$ ▷ count hop number from n,
  $H_{nn}^{(n-1)} \sim \mathrm{NB}\left(h_n, 1 - P_{nn}^{(n-1)}\right)$ ▷ draw flicker number
  $T_n = h_n + H_{nn}^{(n-1)}$ ▷ evaluate hop number from n,
  $\tilde{\theta}_n \sim j\left(T_n, \tau_n^{(0)}\right)$ ▷ convert to residence time in n,
  deallocate $\mathbf{H}^{(n)}$ and $\mathbf{R}^{(n)}$;
end for
evaluate first-passage time $\tilde{\tau}^{(N)} = \sum_{\ell \in \mathbb{E} \cup \mathbb{A}} \tilde{\theta}_\ell$ associated with the path generated in (2);
increment the physical time t by $\tilde{\tau}^{(N)}$.

simulates the time elapsed after performing T_n consecutive Poisson processes of rate $1/\tau_n^{(0)}$. Indeed, after any hop or flicker performed with $\mathbf{P}^{(0)}$, the physical time must be incremented by a residence time drawn in the exponential distribution of time-scale $\tau_n^{(0)}$. Note that algorithm (3) generalizes of the time randomization procedure proposed by Mason and coworkers [59] for the second-order residence time algorithm [19].

In practice, several transitions exiting $\mathbb{E}$ are typically recorded in the hopping matrix. As a result, the elapsed physical time is generated for several escaping trajectories simultaneously. The generated path may also return to the same trap several times prior to reaching another trap. In practice, the current path factorization is re-used as many times as necessary.

Appendix B: Reformulation of path factorization

1. LU decomposition

We herein establish the connection between rank-one updates of Sec. II J and the Gauss-Jordan elimination method on the scaled transition rate matrix $\mathbf{BA}$. Scaling matrix $\mathbf{B}$ is set to $\tau\mathbf{I}$ or $\mathrm{Diag}(\mathbf{A})^{-1}$ where $\tau = \min(1/A_{ii} : 1 \leq i \leq N)$.

More precisely, we show that path factorization entails decomposition $\mathbf{BA} = \mathbf{L}^+\mathbf{D}\mathbf{U}^+$ when the initial stochastic matrix $\mathbf{P}^{(0)}$ is set to $\mathbf{I}-\mathbf{BA}$. Matrix $\mathbf{D}$ is diagonal and its diagonal elements D_{nn} are equal to $1 - P_{nn}^{(n-1)}$, the escape probability from state n. The quantity $P_{nn}^{(n-1)}$ is the probability of a round-trip from the same state. We next define three matrices L_{ij}^+, U_{ij}^- and U_{ij}^+ whose entries are initially set to zero. At the nth update, the nth columns L_{in}^+ and U_{in}^- are filled by setting

$$L_{in}^+ = \begin{cases} \dfrac{P_{in}^{(n-1)} - I_{in}}{P_{nn}^{(n-1)} - 1} & \text{if } i \geq n \\ 0, & \text{if } i < n \end{cases}, \tag{B1}$$

$$U_{in}^- = \begin{cases} I_{in} & \text{if } i \geq n \\ P_{in}^{(n-1)} & \text{if } i < n \end{cases} \tag{B2}$$

The nth row U_{nj}^+ is filled as

$$U_{nj}^+ = \frac{P_{nj}^{(n-1)} - I_{nj}}{P_{nn}^{(n-1)} - 1}. \tag{B3}$$

We have

$$L_{in}^+ - U_{in}^- D_{nn}^{-1} = -\frac{P_{in}^{(n-1)}}{1 - P_{nn}^{(n-1)}}, \tag{B4}$$

$$D_{nn}U_{nj}^+ = P_{nj}^{(n-1)} - I_{nj}. \tag{B5}$$

Since $P_{ij}^{(n-1)} = 0$ for $j \leq n-1$, matrix $\mathbf{U}^+$ remains upper triangular after the nth row addition. The nth rank-one update of Sec. II J amounts to constructing $\mathbf{P}^{(n)}$ as follows:

$$P_{ij}^{(n)} = P_{ij}^{(n-1)} + \left(L_{in}^+ - U_{in}^- D_{nn}^{-1}\right) D_{nn} U_{nj}^+, \tag{B6}$$

where $P_{ij}^{(n)} = 0$ for $j \leq n$, as required. This property holds by induction up to $n = N$. Recall that the probabilities of transitions from i to j ($j > n$) subsume the canceled probability of all possible transitions from i to n. This ensures that the transformed matrices $\mathbf{P}^{(n)}$ remain stochastic.

Summing relation (B6) from $n = i$ to $n = N$ when $i \leq N$ yields the relation

$$P_{ij}^{(N)} = P_{ij}^{(i-1)} + L_{ii}^+ D_{ii} U_{ij}^+ - \sum_{n=i}^{N} U_{in}^- U_{nj}^+, \tag{B7}$$

$$= I_{ij} - \sum_{n=i}^{N} U_{in}^- U_{nj}^+, \tag{B8}$$

where we substituted I_{ij} for $P_{ij}^{(i-1)} + L_{ii}^+ D_{ii} U_{ij}^+$. Let $\mathbf{U}^+$ and $\mathbf{U}^-$ denote the $N \times N$ upper triangular submatrices obtained by restricting U_{ij}^+ and U_{ij}^- to the trapping states $i, j \leq N$. Since $P_{ij}^{(N)} = 0$ for $i, j \leq N$, we obtain

$$\mathbf{I} = \mathbf{U}^-\mathbf{U}^+, \tag{B9}$$

recalling that $\mathbf{I}$ is the $N \times N$ identity matrix. Relation (B9) entails that $\mathbf{U}^-$ is the inverse of $\mathbf{U}^+$. Summing relation (B6) from $\ell = 1$ to $\ell = i - 1$ yields the relation

$$P_{ij}^{(i)} = P_{ij}^{(0)} + \sum_{n=1}^{i} L_{in}^+ D_{nn} U_{nj}^+ - U_{ii}^- U_{ij}^+.$$

Since $P_{ij}^{(i)} = I_{ij} - U_{ii}^- U_{ij}^+$, we obtain the following factorization

$$\sum_{n=1}^{i} L_{in}^+ D_{nn} U_{nj}^+ = I_{ij} - P_{ij}^{(0)} = B_{ii} A_{ij}.$$

We eventually obtain the decomposition of the $N \times N$ scaled rate matrix $\mathbf{BA}$ into the product of a lower triangular matrix, a diagonal matrix $\mathbf{D} = \mathrm{diag}(D_{11}, D_{22}, \cdots, D_{NN})$ and an upper triangular matrix:

$$\mathbf{BA} = \mathbf{L}^+\mathbf{D}\mathbf{U}^+. \tag{B10}$$

Matrix $\mathbf{U}^-$ being the inverse of $\mathbf{U}^+$, inverting $\mathbf{A}$ from the factorization still requires inverting the lower triangular matrix $\mathbf{L}^+$. Let $\mathbf{L}^-$ denote the inverse of $\mathbf{L}^+$. $\mathbf{D}$ being diagonal and positive definite, its inverse, denoted below by $\mathbf{D}^{-1}$, exists. Matrices $\mathbf{L}^\pm$ can be written as products involving the following elementary matrices

$$\mathbf{L}_n^\pm = \mathbf{I} \pm \sum_{i \neq n} \mathbf{e}_i L_{in} \mathbf{e}_n. \tag{B11}$$

We have in particular $\mathbf{L}^+ = \mathbf{L}_1^+\mathbf{L}_2^+\cdots\mathbf{L}_N^+$ and $\mathbf{L}^- = \mathbf{L}_N^-\mathbf{L}_{N-1}^-\cdots\mathbf{L}_1^-$. From the matrix products above and property $\mathbf{L}_n^+\mathbf{L}_n^- = \mathbf{I}$, we deduce that $\mathbf{L}^+\mathbf{L}^- = \mathbf{I}$, hence $\mathbf{L}^-$ corresponds to the inverse of $\mathbf{L}^+$. The decomposition of $\mathbf{L}^-$ and $\mathbf{U}^-$ into product of triangular elemental matrices are used in the updating rule (32) to compute $\vec{\boldsymbol{\tau}}^{(N)}$, the vector of mean first-passage times. The sequential procedure (32) amounts to applying vector $\mathbf{b} = \mathbf{B}\vec{\mathbf{1}}$ on matrices $\mathbf{L}^-$, $\mathbf{D}^{-1}$ and $\mathbf{U}^-$, successively:

$$\mathbf{b} = \left(\tau_1^{(0)}, \tau_2^{(0)} \cdots, \tau_N^{(0)}\right)^T$$

$$\mathbf{L}^-\mathbf{b} = \left(\tau_1^{(0)}, \tau_2^{(1)} \cdots, \tau_N^{(N-1)}\right)^T$$

$$\mathbf{D}^{-1}\mathbf{L}^-\mathbf{b} = \left(\tau_1^{(1)}, \tau_2^{(2)} \cdots, \tau_N^{(N)}\right)^T$$

$$\mathbf{U}^-\mathbf{D}^{-1}\mathbf{L}^-\mathbf{b} = \left(\tau_1^{(N)}, \tau_2^{(N)} \cdots, \tau_N^{(N)}\right)^T.$$

Replacing the product $\mathbf{U}^{-}\mathbf{D}^{-1}\mathbf{L}^{-}$ by $\mathbf{A}^{-1}\mathbf{B}^{-1}$ in the last equation yields the expected expression for the mean first-passage time:

$$\mathbf{A}^{-1}\vec{\mathbf{1}} = \boldsymbol{\tau}^{(N)}.$$

As for time randomization (algorithm 3 of Appendix A), the information processed to evaluate the conditional probabilities $R_{ij}^{(n)}$ defined in (A1) can easily be retrieved by resorting to (B6) and the stored entries of $\mathbf{L}^{+}$, $\mathbf{D}$, $\mathbf{U}^{+}$ and $\mathbf{U}^{-}$.

2. Cholesky decomposition

Whenever the underlying Markov process is reversible, the symmetric positive definite matrix $\mathbf{A}^{B} = \mathbf{S}\mathbf{R}^{-1}\mathbf{A}\mathbf{R}\mathbf{S}$ can be defined, where diagonal matrix $\mathbf{R}$ is defined in (15) from the equilibrium distribution ρ. Cholesky decomposition can then be applied, yielding

$$\mathbf{A}^{B} = \mathbf{L}^{B}\mathbf{D}\mathbf{L}^{BT} \tag{B12}$$

where $\mathbf{L}^{B}$ is a $N \times N$ lower triangular matrix with ones on the diagonal. Since $\mathbf{D}$, $\mathbf{R}$ and $\mathbf{S}$ commute, the absorbing transition rate matrix writes

$$\mathbf{B}\mathbf{A} = (\mathbf{S}\mathbf{R})\mathbf{L}^{B}(\mathbf{S}\mathbf{R})^{-1}\mathbf{D}(\mathbf{S}\mathbf{R})\mathbf{L}^{BT}(\mathbf{S}\mathbf{R})^{-1} \tag{B13}$$

Comparing to $\mathbf{L}^{+}\mathbf{D}\mathbf{U}^{+}$ decomposition enables one to identify the following relations

$$\mathbf{L}^{+} = (\mathbf{S}\mathbf{R})\mathbf{L}^{B}(\mathbf{S}\mathbf{R})^{-1}$$
$$\mathbf{U}^{+} = (\mathbf{S}\mathbf{R})\mathbf{L}^{BT}(\mathbf{S}\mathbf{R})^{-1}$$

We deduce that $\mathbf{U}^{+} = (\mathbf{S}\mathbf{R})^{2}\mathbf{L}^{+T}(\mathbf{S}\mathbf{R})^{-2}$ and that the inverse of $\mathbf{L}^{+}$ is $\mathbf{L}^{-} = (\mathbf{S}\mathbf{R})^{2}\mathbf{U}^{-T}(\mathbf{S}\mathbf{R})^{-2}$.