A Path Factorization Approach to Stochastic Simulations

Manuel Athènes\(^1\) and Vasily V. Bulatov\(^2\)

\(^{1}\)CEA, DEN, Service de Recherches de Métallurgie Physique, F-91191 Gif-sur-Yvette, France\quad \(^{2}\)Lawrence Livermore National Laboratory, Livermore, California 94551, USA

A common issue limiting the efficiency of stochastic simulation algorithms is the numerical slowing-down that results from the occurrence of metastable states. Here, an exact algorithm based on a path factorization of the linearized propagator is proposed and applied to anomalous diffusion and to the precipitation of impurities via vacancy diffusion, two model systems in which metastability is important. Paths escaping from the metastable basins can be directly generated owing to the algorithm, which results in a dramatic acceleration of the simulations.

Kinetic Monte Carlo (kMC) is a stochastic algorithm used extensively in physics, chemistry and biology to simulate the time evolution of complex systems described by a discrete master equation. Regrettably, simulations based on this algorithm become prohibitively expensive whenever the spectrum of the transition rate matrix, \(M\), is broad. The simulated dynamical system encounters sets of states from which it escapes rarely, i.e. with an overall rate that is much lower than the fast transition rates. When \(M\) is independent of time, the metastability issue can be partially alleviated by performing non-local transitions to absorbing boundaries surrounding the trapping basins responsible for the numerical slowing down. In practice, the propagator of system must be solved in each trapping basin delimited by an absorbing boundary through a partial eigenvalue decomposition of \(M\). The most serious deficiency of the approach is that the overhead associated with the submatrix diagonalization is high and that the size and shape of the trapping basin must be known \emph{a priori}. A radically different way to analyse a dynamical system within a trapping basin is to perform a Metropolis Monte Carlo sampling in path space. The path-sampling approach has been applied to discrete master equations and enables one to generate the distribution of first-passage times for escaping a basin. However, the method is limited to transition paths of short duration, requires a time-consuming preliminary equilibration in path-space and is thus inappropriate for simulating multi-scale kinetics. In this letter, we propose an algorithm for directly sampling paths escaping from a trapping basin. This will result in an enhanced flexibility and a dramatic improvement in the computational efficiency because neither matrix diagonalization nor equilibration in path space are required. Our approach consists first in factoring the escaping paths generated by an approximate propagator and second in drawing a first-passage time through an exact randomization procedure.

The method is based on the propagator matrix, obtained by integrating the master equation from \(t\) to \(t + \tau\):

\[
P(t, t + \tau) = \exp \left( \int_t^{t+\tau} M ds \right) = \exp (\tau M). \tag{1}
\]

Entry \(P_{\beta\gamma}(t, t + \tau)\) yields the probability that the system is in state \(\gamma\) at \(t + \tau\) given that it was in state \(\beta\) at time \(t\) earlier. Accordingly, off-diagonal entries \(M_{\beta\gamma}\) correspond to the rate for transitioning from state \(\beta\) to state \(\gamma\). Diagonal entries satisfy \(M_{\beta\beta} = -\sum_{\gamma \neq \beta} M_{\beta\gamma}\). This condition ensures that \(P\) is a stochastic matrix: whatever \(\beta, \gamma, t\) and \(\tau\), we have \(\sum_{\nu} P_{\beta\nu} = 1\) and \(P_{\beta\gamma} \geq 0\). Besides, \(P(t, t)\) coincides with the identity matrix, denoted by \(I\).

To avoid computing the matrix exponential, we approximate the propagator by linearizing \(P\) as follows

\[
P^{(0)}(t, t + \tau) = I + \tau M. \tag{2}
\]

We choose \(\tau\) to be positive and lower than \(k_{\beta} = -M_{\beta\beta}\) whatever \(\beta\) so that all diagonal entries of \(P^{(0)}\) remain positive. Hence, the linearized propagator is a stochastic matrix that can be used to generate paths, those to be factored. The path factorization is done iteratively by combining hop sequences and summing over all possible numbers of hops. It amounts to eliminating connections or states from the connectivity graph associated with the propagator, and also serves to compute mean-first passage times. Here, Wales’s formulation is adopted because of its simplicity. We refer to the \(N\) states \((a_n)^N\) that are iteratively eliminated from the connectivity graph by their labels sorted in ascending order in \(E = \{ 1, 2, \ldots, N \}\). We next define \(P^{(n)}\) as a set of transformed stochastic matrices such that the transition probability \(P^{(n)}_{\beta\nu}\) is zero for \(\nu \leq n\) whatever \(\beta\). Let also define two matrices \(U\) and \(L\) whose columns are initially set to zero. The iterative procedure aims at constructing \(P^{(N)}\) from \(P^{(0)}\). At the \(n\)-th iteration, one first fills the \(n\)-th row of \(U\) and the \(n\)-th column of \(L\) by setting \(U_{n\gamma} = P^{(n-1)}_{n\gamma} - I_{n\gamma}\) and \(L_{\beta n} = -P^{(n-1)}_{\beta n}/U_{nn}\) and next compute \(P^{(n)}\) as follows

\[
P^{(n)}_{\beta\gamma} = P^{(n-1)}_{\beta\gamma} + L_{\beta n} U_{n\gamma}. \tag{3}
\]

The specification that \(P^{(n-1)}_{\beta\nu} = 0\) for \(\nu \leq n - 1\) implies that \(U\) remains upper triangular after the \(n\)-th row addition and that the equality \(L_{\beta n} U_{n\nu} = -P^{(n-1)}_{\beta\nu}\) is satisfied for \(\nu \leq n\). It results that \(P^{(n)}_{\beta\nu} = 0\) for \(\nu \leq n\), as required. This property holds for \(0 \leq n \leq N\), by induction on \(n\).
The factorization \(3\) corresponds to Gauss-Jordan elimination method, as shown in Supplemental Material [19]. Besides, the canceled probability of transitioning from \(\beta\) to \(n\) is taken into account in the transition probabilities \(P_{\beta\gamma}^{(n)}\) owing to the quantity \((\gamma > n)\)

\[
\mathcal{L}_{\beta n}U_{n\gamma} = P_{\beta\gamma}^{(n-1)} \sum_{f=0}^{+\infty} \left[ P_{n\gamma}^{(n-1)} \right] f P_{n\gamma}^{(n-1)}. \tag{4}
\]

The sum of the probability products represents the probability of transitioning from \(\beta\) to \(\gamma > n\) via \(n\) with an arbitrary number of consecutive flickers (round-trips) from \(n\) and using \(P^{(n)}\). The additional contributions ensure that the transformed matrices \(P^{(n)}\) remain stochastic. One can then perform a stochastic simulation with \(P^{(n)}\) and monitor the transition numbers in an associated matrix \(H^{(n)}\); entry \(H_{\beta\gamma}^{(n)}\) is used to count the number of transitions from \(\beta\) to \(\gamma\). Let now explain how to obtain \(H^{(n-1)}\), the matrix similarly counting the transition numbers in a stochastic process based on \(P^{(n-1)}\) without actually performing the simulation using \(P^{(n-1)}\). The ratio of transition probabilities defined by \((\gamma > n)\)

\[
R_{\beta\gamma}^{(n)} = \frac{P_{\beta\gamma}^{(n-1)}}{P_{\beta\gamma}^{(n)}} = 1 - \frac{\mathcal{L}_{\beta n}U_{n\gamma}}{P_{\beta\gamma}^{(n-1)}} \tag{5}
\]

is the conditional probability that the system performs a direct transition from \(\beta\) to \(\gamma\) using \(P^{(n-1)}\) given that it has transitioned from \(\beta\) to \(\gamma\) in a stochastic simulation performed with \(P^{(n)}\). For \(\beta = n\), \(R_{\gamma n}^{(n)}\) is independent of \(\gamma\) and is equal to \(1 - P_{n\gamma}^{(n-1)}\), the escape probability from \(n\). It is thus possible to generate \(H^{(n-1)}\) by performing a stochastic simulation with \(P^{(n)}\), harvesting \(H^{(n)}\) and drawing random variates in binomial distributions whose exponents and coefficients are respectively given by the entries of \(H^{(n)}\) and \(R^{(n)}\). This also suggests the possibility of directly drawing \(H^{(0)}\) from \(H^{(N)}\) generated using a stochastic simulation with \(P^{(N)}\) by iterating the randomizing procedure on \(n\) in reverse order from \(N\) to 1.

To show how full randomization is achieved in practice, the following preliminary definitions are required. The binomial law of trial number \(h \in \mathbb{N}\) and success probability \(r\) is denoted by \(B(h, r)\). The probability mass 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 \(NB(h, 1 - p)\). The probability mass 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. The gamma law of shape parameter \(h\) and time-scale \(\tau\) is denoted by \(\Gamma(h, \tau)\). \(C_{\alpha}\) denotes the categorical laws whose probability vector is the \(\alpha\)-th line of \(P^{(N)}\) if \(\alpha < N\) or of the stochastic matrix obtained from \(P^{(0)}\) by eliminating the single state \(\alpha > N\). The symbol \(\sim\) means “is a random variate distributed according to the law that follows”.

Let \(\alpha\) denote the current state of the system. The cyclic structure of the algorithm, referred to as kinetic path sampling (kPS), is

a. compute \(P^{(N)}\) by iterating \(3\) on \(n\) from 1 to \(N\), and label the states connected to \(E\) in ascending order from \(N + 1\) to \(N_c\) through appropriate permutations;

b. define \(A\), the set of absorbing states and \(T = A \cup E\), the set of noneliminated transient states; set \(H^{(N)}\) to the null matrix and \(h\) to the null vector;

c. draw \(\gamma \sim C_{\alpha}\); increment \(H_{\beta\gamma}^{(n)}\) by one if \(\alpha \in E\) or increment the hop number \(h_\alpha\) by one if \(\alpha \in T\); \(\alpha\) becomes \(\gamma\); if \(\alpha \in A\) go to (d) otherwise repeat (c);

d. iterate in reverse order from \(n = N\) to 1:

i. for \(\beta \in E \setminus \{n\}\) and \(\gamma \in \{n+1, ..., N_c\}\) draw

\[
H_{\beta\gamma}^{(n-1)} \sim B\left( H_{n\gamma}^{(n-1)}, P_{n\gamma}^{(n-1)} \right);
\]

ii. for \(\beta \in E \setminus \{n\}\) count the new hops from \(n\)

\[
H_{\beta\gamma}^{(n-1)} = \sum_{\gamma \in \{n+1, ..., N_c\}} H_{\beta\gamma}^{(n)} - H_{\beta\gamma}^{(n-1)};
\]

iii. for \(\gamma \in \{n+1, ..., N_c\}\) count the hops from \(n\) to \(\gamma\)

\[
H_{\gamma n}^{(n-1)} = H_{\gamma n}^{(n)} + \sum_{\beta \in E \setminus \{n\}} H_{\beta\gamma}^{(n)} - H_{\beta\gamma}^{(n-1)};
\]

iv. compute \(h_n = \sum_{\gamma \in \{n+1, ..., N_c\}} H_{\gamma n}^{(n-1)}\), the number of hops from \(n\), and draw the flicker number

\[
H_{\gamma n}^{(n-1)} \sim NB\left( h_n, 1 - P_{n\gamma}^{(n-1)} \right);
\]

v. store \(T_{n} = h_n + H_{\gamma n}^{(n-1)}\), the number of transitions from \(n\), and deallocate \(H^{(n)}\), \(P^{(n)}\) and \(R^{(n)}\);

e. for \(\ell \in T\) store \(T_{\ell} = h_{\ell} + f_{\ell}\) with \(f_{\ell} \sim NB(h_{\ell}, 1 - P_{\ell\ell}^{(0)})\);

f. evaluate \(T^A = \sum_{\ell \in E \cup T} T_{\ell}\), the total number of flickers and hops associated with the path generated in (3);

g. increment the physical time by \(t^h \sim \Gamma(T^A, \tau)\).

After this cycle, the system \(\alpha\) is in the boundary state reached in (3). The gamma law \(\Gamma(T^A, \tau)\) in (5) simulates the time elapsed after performing \(T^A\) consecutive Poisson processes of rate \(\tau^{-1}\). Indeed, after any hop or flicker performed with \(P^{(0)}\), the physical time must be incremented by a residence time drawn in the exponential distribution of decaying rate \(\tau^{-1}\) [20]. The Supplemental Material proves that the algorithm generates the correct distribution of first-passage times to \(A\) for \(N = 0\). The proof that it is correct for \(N \geq 1\) follows by induction on \(N\) resorting to the iterative structure of the algorithm. The way \(E\) and \(A\) are constructed at each cycle is specific to the application.
We first apply the algorithm to the anomalous diffusion of a defect on a disordered substrate [21] consisting of a $256 \times 256$ square lattice with periodic conditions and lattice parameter $a$. Each position $\alpha$ of the defect on the lattice defines a state of energy $E_\alpha$. The defect hops to nearest-neighbour sites exclusively with transition rates $M_{\alpha \beta} = \omega \exp\left[\frac{(E_\alpha - E_\beta)}{\kappa T}\right]$ where $\omega$ is an attempt frequency, $T$ the temperature and $E_\alpha$ the saddle energy between $\alpha$ and $\beta$. The saddles $\sigma$ associated with nearest-neighbour pairs $(\alpha, \beta)$ are located on an additional lattice. Let $\Omega(\alpha)$ denote the subset of states $\nu$ for which the Chebyshev distance $D_\infty(\alpha, \nu)$ is lower than $48a$. Subset $\Omega^s(\sigma)$ similarly contains states $\nu$ satisfying $D_\infty(\sigma, \nu) \leq 48a$. We set

$$E_\alpha = \epsilon \sum_{\nu \in \Omega(\alpha)} (a_\nu + b_\nu) \quad E_\nu^s = \epsilon \sum_{\nu \in \Omega^s(\sigma)} (a_\nu - b_\nu)$$

where $a_\nu$ and $b_\nu$ are independent and identically distributed random variables in $\{-1, 1\}$. The saddle energy surface is displayed in Fig. 1. Temperatures are given in $\epsilon/\kappa$ units. Local hops performed using kMC enables the defect to migrate on long distances at temperature $T = 2.5$ but fails to activate migration over the large saddle barriers of Fig. 1 at $T = 1$ in a reasonable computer timescale: the defect remains trapped in its basin around the site of coordinates $(127, 127)$. In kPS, the current state is eliminated first. The $n$-th state to be eliminated is the state $\gamma$ for which $P_t^{(n)}$ is largest. Here, $T = 0$ implying $A = \infty$. The magnifying contours displayed in Fig. 1 represent the absorbing boundary $\partial A$ obtained after the eliminations of $2^7$, $2^9$, $2^{11}$, $2^{12}$, $2^{13}$ and $2^{14}$ states starting from $(127, 127)$. To demonstrate the correctness of kPS, we eliminate $N = 2^{13}$ states and generate $10^4$ paths starting from $(127, 127)$ state and ending in $A$, using both kPS and kMC at $T = 2.5$. The perfect match between the two estimated distributions of first-passage times is shown in Fig. 2(a). The mean first-passage times to $A$ are plotted in Fig. 2(b) as a function of the number of eliminated states at $T = 2.5$ and $T = 1$, while the cost of both methods are displayed in Fig. 2(c) for comparison. We observe that the factorization cost scales as $N^3$ and exceeds that of kMC for $N > 2^{12}$ at $T = 2.5$. However, at $T = 1$, the speedup achieved by kPS, measured as the ratio of the computational cost of kPS to that of kMC, is four orders of magnitude for $N = 2^{15}$. At this low temperature, trapping becomes severe and standard kMC is inefficient. Note that a first-passage time to $A$ can still be generated from the diagonalization of the transition rate matrix $[3, 8]$. However, this computational approach is less simple than the randomization procedure (d) of the kPS algorithm.

![Energy surface of saddle points. The colorscale spans 36 $\epsilon$ units and artificial smoothing is used for better visualization. The defect is initially located at (127, 127).](image)

**FIG. 1:** Energy surface of saddle points. The colorscale spans 36 $\epsilon$ units and artificial smoothing is used for better visualization. The defect is initially located at (127, 127).

![Survival probability of the defect in the trap (solid lines), obtained from $-\frac{ds}{dt}$, the distribution of the logarithms of the simulated first-passage times (dashed lines); (b) evolution of the mean first-passage time as a function of $N$; (c) computational cost of kMC and kPS as a function of $N$ for the two temperatures. The cost unit is the cost of a kMC hop. The cost of a kMC path to $A$ is the mean number of kMC hops along the path.](image)

**FIG. 2:** (a) $S(t)$ is the survival probability of the defect in the trap (solid lines), obtained from $-\frac{ds}{dt}$, the distribution of the logarithms of the simulated first-passage times (dashed lines); (b) evolution of the mean first-passage time as a function of $N$; (c) computational cost of kMC and kPS as a function of $N$ for the two temperatures. The cost unit is the cost of a kMC hop. The cost of a kMC path to $A$ is the mean number of kMC hops along the path.

We now apply kPS to the modeling of Cu substitutional impurities in $\alpha$-Fe system, which precipitate via a vacancy diffusion mechanism at low enough temperatures. Interaction energies of the FeCu model system
have been obtained from electronic structure calculations [22]. The simulated lattice has periodic boundary conditions and contains $2^{21}$ sites. Its crystallographic structure is body centered. Vacancies having a substantially lower formation energy in Cu get trapped in Cu precipitates, making kMC inefficient below 500 K [22]. In each kPS cycle, set $E$ contains the possible positions of the vacancy in the cluster formed by the vacancy itself and the $N-1$ neighboring Cu atoms with which the vacancy can exchange without moving any Fe atom. Cu and Fe atoms are unlabeled. The $n$-th eliminated pivot in $E$ corresponds to the least connected entry of $P^{(n)}$. Set $T$ consists of states corresponding to the positions of the vacancy in the iron bulk. The vacancy reaches $A$ when it first enters another Cu cluster or the old cluster in a new conformation. A single vacancy is introduced and Cu concentration is set to 1.34 at.%. Three temperatures $T_0 = 273$ K, $T_1 = 373$ K and $T_2 = 473$ K are used. We define the integrated speed-up as the ratio of the mean physical time simulated by kPS to that obtained using kMC simulations and given an identical computer wall time. The integrated speed-up is plotted as a function of the averaged physical times simulated by kPS in Fig. 3.a. Averages are performed over 41 kinetics for damping the fluctuations of the physical time. We observe that the integrated speed-up increases with decreasing temperatures. This is expected as vacancy trapping is more severe at low temperatures. The two oblique segments in Fig. 3.a have a slope of 1 and enable one to report the integrated speed-ups (achieved by kPS) onto the averaged physical times simulated by kMC. The precipitation kinetics, monitored via the averaged Warren-Cowley short-range order (SRO) parameter [23], are shown in Fig. 3.b both for kPS and kMC simulations. Two ageing regimes are observed on the simulated timescale and temperature range: the first one corresponds to the incubation of critical copper clusters and the second one to the agglomeration of solute Cu atoms by Cu clusters migrating with the vacancy [22, 25]. With the aid of the vertical arrows pointing downward, one can compare the degree of phase separation obtained with kPS to that of kMC from the two SRO values corresponding to an identical computer wall time. An integrated speed-up of more than seven orders of magnitude is measured at $T_0$ from the pair of arrows displayed in solid lines. This shows the inability of kMC to simulate the agglomeration regime at this temperature. Note that the speed-up is optimal in the incubation regime and decreases as soon as kPS starts simulating the agglomeration regime. This trend results from the overhead associated with the matrix factorization which increases asymptotically as $N^3$. The cost of factorization will ultimately exceed that of kMC which scales linearly with $N$, the cluster size.

In conclusion, we proposed a kinetic path sampling algorithm based on a path factorization of the linearized propagator. The strength of the approach lies in its flexibility. The path factorization can be adapted to the trapping basin on-the-fly and used to estimate the cost of standard kMC from the mean first-passage number of kMC jumps. As a result, the path factorization of kPS can be stopped whenever its cost exceeds that of kMC. Finally, we note that transformed stochastic matrices have been used to generate non-local transitions in stochastics simulations [10, 23, 24]. In these studies, the physical time was updated using mean first-passage times instead of first-passage times, an approximation that introduces artificial correlations. The demonstrated ability to generate first passage times from the path factorization makes it possible to synchronize several subsystems evolving independently in protective cells [24]. Hence, the proposed approach should be particularly useful in phase-transformation kinetics in which few atoms, particles or defects undergo rapid fluctuations in confined portions of the space while the microstructure of the entire system evolves extremely slowly.

This work was performed using HPC resources from GENCI-[CCRT/CINES] (Grant x2013096973). The authors wish to express their gratitude to J.-L. Bocquet, A. Donev, F. Soisson and T. Opplestrup for fruitful discussions.

[1] J.-M. Lanore, Rad. Effects 22, 153 (1974).
[2] A. Bortz, M. Kalos, J. Lebowitz, J. Comp. Phys. 17, 10 (1975).
[3] D. Gillespie, J. Chem. Phys. 81 2340 (1977).
[4] N. G. van Kampen, Stochastic Processes in Physics and Chemistry (Elsevier Science, 2007).
Here, we show that the path factorization of Eq. 3 corresponds to the Gauss-Jordan elimination method on matrix $\tau M$. Summing $P_{\beta\gamma}^{(v)} - P_{\beta\gamma}^{(v-1)} = L_{\nu\beta} U_{\nu\gamma}$ from $\nu = \beta$ to $\nu = n$ when $\beta \leq n$ or from $\nu = n+1$ to $\nu = \beta - 1$ when $n < \beta - 1$ yields the relation

$$P_{\beta\gamma}^{(n)} = P_{\beta\gamma}^{(\beta-1)} + \sum_{\nu \geq 0, \nu \leq n} L_{\nu\beta} U_{\nu\gamma} - \sum_{\nu > n, \nu < \beta} L_{\nu\beta} U_{\nu\gamma}. \quad (6)$$

Choosing $n > \max(\beta, \gamma)$ in Eq. (6) entails $P_{\beta\gamma}^{(n)} = 0$ and $(U^\# U)_{\beta\gamma} = I_{\beta\gamma}$, while setting $n = 0$ yields $LU = P^{(0)} - I = \tau M$. Factorizations of $I$ and $\tau M = P^{(0)} - I$ result from transformations below and above the eliminated pivot, respectively. This is the Gauss-Jordan pivot elimination method. Note that the information necessary for evaluating $R_{\beta\gamma}^{(n)}$ in Eq. 4 is easily and cheaply retrieved from the entries of $U$ and $L$.

Next, we prove that the algorithm sample the correct first-passage distribution when $\tau = 0$, i.e. $E = 0$. In (g), we take advantage of the distributivity of the Gamma law with respect to its shape parameter and decompose the first-passage time as follows:

$$t^h \sim \sum_{\ell \in \mathbb{T}} \Gamma(h \ell + f_\ell, \tau), \quad (7)$$

where $h_\ell$ and $f_\ell$ are the generated numbers of hops and flickers from $n$. For any state $\ell \in \mathbb{T}$ that is visited $h$ times, the probability of having $f$ flickers before the $h$-th escape from $\ell$ is $\tau^{-1}(n-1)^{f}(1-P_\ell)^h$, which corresponds to the probability mass of the negative binomial law of success number $n$ and success probability $1 - P_\ell$. The residence time associated with $h + f$ hops and flickers from state $\ell$ is distributed according to $\Gamma(h + f, \tau)$, the Gamma law of shape parameter $1 + f$ and time-scale $\tau$. The probability mass of $\Gamma(h + f, \tau)$ law at $t$ being $\tau^{-1} \frac{1}{(n-1)^{f}(n-1)^{f}} \left(t/\tau\right)^{n-1+f} \exp[-t/\tau]$, the overall probability to draw residence time $t$ for $h$ visits state $\ell$ is obtained by summing over the compound occurrences as follows

$$\frac{1}{(h-1)!} \left(\frac{\tau k_\ell}{\tau}\right)^h t^{h-1} \sum_{f=0}^{+\infty} \frac{1}{f!} \left[\frac{t}{\tau}\right]^f \exp[-t/\tau] = \frac{(k_\ell t)^{h-1}}{(h-1)!} k_\ell \exp[-k_\ell t],$$

thereby removing the dependence on $\tau$. We obtain the distribution of the Gamma law of shape parameter $h$ and time-scale $1/k_\ell$, which corresponds to the convolution of $h$ decaying exponential of rate $k_\ell$. This is the expected distribution for the times elapsed after $h$ consecutive Poisson processes of rate $k_\ell$. Note that the standard kMC algorithm simply draws an escape time according to $\Gamma(1, 1/k_\ell)$, the exponential law of rate $k_\ell$ for each visit of $\ell$, which is statistically equivalent. This amounts to prescribing a success probability of 1, which results in $\tau = 1/k_\ell$ and no failures.

**Supplemental Material**

Here, we show that the path factorization of Eq. 3 corresponds to the Gauss-Jordan elimination method on matrix $\tau M$. Summing $P_{\beta\gamma}^{(v)} - P_{\beta\gamma}^{(v-1)} = L_{\nu\beta} U_{\nu\gamma}$ from $\nu = \beta$ to $\nu = n$ when $\beta \leq n$ or from $\nu = n+1$ to $\nu = \beta - 1$ when $n < \beta - 1$ yields the relation

$$P_{\beta\gamma}^{(n)} = P_{\beta\gamma}^{(\beta-1)} + \sum_{\nu \geq 0, \nu \leq n} L_{\nu\beta} U_{\nu\gamma} - \sum_{\nu > n, \nu < \beta} L_{\nu\beta} U_{\nu\gamma}. \quad (6)$$

Choosing $n > \max(\beta, \gamma)$ in Eq. (6) entails $P_{\beta\gamma}^{(n)} = 0$ and $(U^\# U)_{\beta\gamma} = I_{\beta\gamma}$, while setting $n = 0$ yields $LU = P^{(0)} - I = \tau M$. Factorizations of $I$ and $\tau M = P^{(0)} - I$ result from transformations below and above the eliminated pivot, respectively. This is the Gauss-Jordan pivot elimination method. Note that the information necessary for evaluating $R_{\beta\gamma}^{(n)}$ in Eq. 4 is easily and cheaply retrieved from the entries of $U$ and $L$.

Next, we prove that the algorithm sample the correct first-passage distribution when $\tau = 0$, i.e. $E = 0$. In (g), we take advantage of the distributivity of the Gamma law with respect to its shape parameter and decompose the first-passage time as follows:

$$t^h \sim \sum_{\ell \in \mathbb{T}} \Gamma(h \ell + f_\ell, \tau), \quad (7)$$

where $h_\ell$ and $f_\ell$ are the generated numbers of hops and flickers from $n$. For any state $\ell \in \mathbb{T}$ that is visited $h$ times, the probability of having $f$ flickers before the $h$-th escape from $\ell$ is $\tau^{-1}(n-1)^{f}(1-P_\ell)^h$, which corresponds to the probability mass of the negative binomial law of success number $n$ and success probability $1 - P_\ell$. The residence time associated with $h + f$ hops and flickers from state $\ell$ is distributed according to $\Gamma(h + f, \tau)$, the Gamma law of shape parameter $1 + f$ and time-scale $\tau$. The probability mass of $\Gamma(h + f, \tau)$ law at $t$ being $\tau^{-1} \frac{1}{(n-1)^{f}(n-1)^{f}} \left(t/\tau\right)^{n-1+f} \exp[-t/\tau]$, the overall probability to draw residence time $t$ for $h$ visits state $\ell$ is obtained by summing over the compound occurrences as follows

$$\frac{1}{(h-1)!} \left(\frac{\tau k_\ell}{\tau}\right)^h t^{h-1} \sum_{f=0}^{+\infty} \frac{1}{f!} \left[\frac{t}{\tau}\right]^f \exp[-t/\tau] = \frac{(k_\ell t)^{h-1}}{(h-1)!} k_\ell \exp[-k_\ell t],$$

thereby removing the dependence on $\tau$. We obtain the distribution of the Gamma law of shape parameter $h$ and time-scale $1/k_\ell$, which corresponds to the convolution of $h$ decaying exponential of rate $k_\ell$. This is the expected distribution for the times elapsed after $h$ consecutive Poisson processes of rate $k_\ell$. Note that the standard kMC algorithm simply draws an escape time according to $\Gamma(1, 1/k_\ell)$, the exponential law of rate $k_\ell$ for each visit of $\ell$, which is statistically equivalent. This amounts to prescribing a success probability of 1, which results in $\tau = 1/k_\ell$ and no failures.