

Numerical analysis of first-passage processes in finite Markov chains exhibiting metastability

Daniel J. Sharpe  and David J. Wales **Department of Chemistry, University of Cambridge, Lensfield Road, and Cambridge CB2 1EW, United Kingdom*

(Received 9 February 2021; revised 28 May 2021; accepted 29 May 2021; published 6 July 2021)

We describe state-reduction algorithms for the analysis of first-passage processes in discrete- and continuous-time finite Markov chains. We present a formulation of the graph transformation algorithm that allows for the evaluation of exact mean first-passage times, stationary probabilities, and committor probabilities for all nonabsorbing nodes of a Markov chain in a single computation. Calculation of the committor probabilities within the state-reduction formalism is readily generalizable to the first hitting problem for any number of alternative target states. We then show that a state-reduction algorithm can be formulated to compute the expected number of times that each node is visited along a first-passage path. Hence, all properties required to analyze the first-passage path ensemble (FPPE) at both a microscopic and macroscopic level of detail, including the mean and variance of the first-passage time distribution, can be computed using state-reduction methods. In particular, we derive expressions for the probability that a node is visited along a direct transition path, which proceeds without returning to the initial state, considering both the nonequilibrium and equilibrium (steady-state) FPPEs. The reactive visitation probability provides a rigorous metric to quantify the dynamical importance of a node for the productive transition between two endpoint states and thus allows the local states that facilitate the dominant transition mechanisms to be readily identified. The state-reduction procedures remain numerically stable even for Markov chains exhibiting metastability, which can be severely ill-conditioned. The rare event regime is frequently encountered in realistic models of dynamical processes, and our methodology therefore provides valuable tools for the analysis of Markov chains in practical applications. We illustrate our approach with numerical results for a kinetic network representing a structural transition in an atomic cluster.

DOI: [10.1103/PhysRevE.104.015301](https://doi.org/10.1103/PhysRevE.104.015301)

I. INTRODUCTION

The analysis of first-passage processes [1–4], concerning the evolution of a system until a specified target state is hit, is of fundamental interest in the theory of stochastic dynamics. The usual dynamical observable is the mean first-passage time (MFPT), defined as the expected time for trajectories to hit the target state [5–8]. The set of possible paths and their associated probabilities, for a given initial occupation probability distribution, defines the (nonequilibrium) first-passage path ensemble (FPPE) [9–17]. Finite Markov chains [18–24], in which a dynamical process is modeled as a sequence of memoryless jumps between the nodes of a network [25], are a class of discrete-state stochastic model that have been widely adopted in diverse disciplines. First-passage processes in Markov chains can be used to model stochastic phenomena as varied as biomolecular conformational transitions [26–33], animal movement to a foraging site within an ecosystem [34], and the sequence of events leading to a stock market crash in economics [35].

The dynamical properties of interest characterizing the FPPE for a Markov chain can in principle be computed by solving a corresponding system of linear equations [36,37]. However, the required computations, including eigendecomposition or matrix inversion operations, are liable to encounter

numerical issues arising from finite precision when the Markov chain exhibits metastability [38]. For Markov chains featuring a separation of characteristic timescales, the subdominant eigenvalue of the underlying transition probability (rate) matrix approaches unity (zero) [39,40], respectively, and the system is therefore severely ill-conditioned [41–44]. Moreover, in general, it is nontrivial to apply preconditioning techniques to improve the numerical stability of sparse linear algebra methods [42,45]. In realistic applications, there is typically a rare event that is of particular interest, which is the first-passage process that we wish to analyze [38,46–59]. This situation motivates the development of alternative procedures that have inherent numerical stability, so that the fundamental dynamical properties of a Markov chain can be computed robustly.

In the present work we focus on state-reduction methods [60] to derive numerically stable algorithms for the analysis of Markov chain dynamics [61–64]. These methods proceed via elimination of the nodes in a Markov chain, while preserving averages for the dynamical properties of interest, and may also employ a back substitution phase to restore the eliminated nodes in turn [65]. State-reduction algorithms have been formulated to compute the stationary distribution [66,67], MFPTs [68], moments of the FPT distribution [69], and the group inverse [36,37] of an irreducible Markov chain [70–72].

We present a convenient formulation of the graph transformation (GT) algorithm [73–78] that allows for the

*dw34@cam.ac.uk

simultaneous determination of the MFPTs, stationary probabilities, and committor probabilities [79–82] for *all* nonabsorbing nodes of a Markov chain in a single computation (Sec. II), as well as the absorption probabilities [18]. We also show that a state-reduction algorithm can be designed to compute the expected number of times that nodes are visited along first-passage paths (Sec. III B). Other quantities that characterize the global and local properties of the FPPE, such as the variance of the FPT distribution and the variances in the number of times that nodes are visited, can be determined from this information (Sec. III). We derive expressions for the probabilities that nodes are visited along reactive transition paths, which proceed directly to the absorbing state without returning to the initial state, considering both the nonequilibrium and equilibrium (i.e., steady-state) FPPEs (Sec. III D). The reactive visitation probabilities quantify the dynamical relevance of individual nodes, and therefore allow us to identify the key mechanisms for productive transitions and the bottleneck nodes that mediate the dominant pathways. The theory of committor and reactive visitation probabilities, and of reactive fluxes along individual edges of the network, is generalizable to the case where there are multiple *taboo* states [83–85]. Separation of the dynamics into competing first-passage processes associated with alternative target states is frequently of interest in models featuring several attractors.

The use of the state-reduction algorithms is illustrated with numerical results for a kinetic network representing a solid-solid structural transition in a model atomic cluster, for which standard linear algebra methods are unable to compute any of the aforementioned dynamical quantities (Sec. IV). Hence, our methodology provides a viable means to analyze first-passage processes in Markov chains exhibiting rare event dynamics, at both a microscopic and macroscopic level of detail. The computations were performed using our DISCOTRESS software [86], further information on which is provided in the Appendix.

II. LU DECOMPOSITION FORMULATION OF GRAPH TRANSFORMATION

A. Markov chain dynamics

We consider discrete-time Markov chains (DTMCs) parameterized by a transition probability matrix $\mathbf{T}(\tau)$, where $i \leftarrow j$ transitions have probabilities $T_{ij}(\tau)$ and are associated with a fixed lag time τ [18], and continuous-time Markov chains (CTMCs) parameterized by a transition rate matrix \mathbf{K} [20]. The off-diagonal elements of \mathbf{K} are the $i \leftarrow j$ transition probabilities per unit time in the limit of an infinitesimally small time step, and the diagonal elements are $K_{jj} = -\sum_{\gamma \neq j} K_{\gamma j}$, so that the columns of the matrix sum to zero [87]. The $i \leftarrow j$ transition probabilities for a CTMC are the elements $P_{ij} = K_{ij} / \sum_{\gamma \neq j} K_{\gamma j}$ of the branching probability matrix \mathbf{P} , and the waiting time for the $i \leftarrow j$ transition is drawn from an exponential distribution with mean $\tau_j = 1 / \sum_{\gamma \neq j} K_{\gamma j}$ [88]. We shall denote the transition probability matrix, either $\mathbf{T}(\tau)$ or \mathbf{P} , by \mathbf{T} for generality. We consider a Markov chain with state space \mathcal{S} partitioned into the set of absorbing nodes \mathcal{A} and the set of *transient* (non-absorbing) nodes \mathcal{Q} , so that $\mathcal{S} \equiv \mathcal{Q} \cup \mathcal{A}$ [89]. That is, we

consider Markov chains where the set \mathcal{A} must eventually be reached when the process is initialized at any node of the set $\mathcal{Q} \equiv \mathcal{A}^c$ [18]. If the nodes of the set \mathcal{A} are not absorbing in the underlying model, so that it is possible to reach any node of the network from any other node, then the Markov chain is said to be *irreducible* [36,90–92]. We may then define the stationary probability distribution (column) vector $\boldsymbol{\pi}$, which satisfies the global balance equations $\mathbf{T}(\tau)\boldsymbol{\pi} = \boldsymbol{\pi}$ and $\mathbf{K}\boldsymbol{\pi} = \mathbf{0}$ [87].

A key dynamical quantity characterizing a first-passage process is the $\mathcal{A} \leftarrow j$ mean first-passage time (MFPT) \mathcal{T}_{Aj} , defined as the expected time at which a trajectory first hits the state \mathcal{A} , given that it was initialized at node j [20,93,94]. Let the time associated with a particular $\mathcal{A} \leftarrow j$ first-passage trajectory be denoted by t_{FPT} . Then the MFPT is defined as

$$\mathcal{T}_{Aj} = \langle t_{\text{FPT}} \rangle = \int_0^\infty t_{\text{FPT}} p(t_{\text{FPT}}) dt_{\text{FPT}}. \quad (1)$$

Here, $p(t_{\text{FPT}})$ is the first-passage time (FPT) distribution [15]

$$p(t_{\text{FPT}}) = \Pr\{\xi(t_{\text{FPT}}) \in \mathcal{A}, \xi(0 \leq t < t_{\text{FPT}}) \notin \mathcal{A} \mid \xi(t=0) = j\}, \quad (2)$$

where $\xi(t)$ denotes the node of the Markov chain that is occupied for the first-passage path $\xi = (\mathcal{A} \leftarrow i_n \leftarrow \dots \leftarrow i_1 \leftarrow j)$ at time t , where $j, i_1, \dots, i_n \notin \mathcal{A}$ [95]. The MFPTs satisfy a first-step relation [22],

$$\mathcal{T}_{Aj} = \tau_j + \sum_{\gamma \notin \mathcal{A}} T_{\gamma j} \mathcal{T}_{A\gamma}. \quad (3)$$

Therefore, in principle, the MFPTs $\mathcal{T}_{Aj} \forall j \notin \mathcal{A}$ can be determined by solving Eq. (3) using any appropriate linear algebra method, such as Gauss-Siedel iteration [96] or successive over-relaxation [97]. The MFPT for a transition from an initial set $\mathcal{B} \subseteq \mathcal{A}^c$, associated with a specified initial occupation probability distribution vector $\mathbf{p}(0)$, is then obtained simply as a weighted average,

$$\mathcal{T}_{AB} = \sum_{b \in \mathcal{B}} p_b(0) \mathcal{T}_{Ab}. \quad (4)$$

However, for Markov chains exhibiting metastability, the linear system of equations in Eq. (3) can be severely ill-conditioned, so that standard dense linear algebra methods may experience a severe propagation of numerical error arising from finite precision [76,77], and Krylov subspace methods [45,98] may fail to converge [42,99].

B. Stochastic complements and the graph transformation algorithm

The $\mathcal{A} \leftarrow j$ MFPT for a transition from a particular node $j \notin \mathcal{A}$ can be computed robustly using stochastic complementation [100–102]. Let us partition the transition probability matrix as

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{\mathcal{Z}\mathcal{Z}} & \mathbf{T}_{\mathcal{Z}\mathcal{N}} \\ \mathbf{T}_{\mathcal{N}\mathcal{Z}} & \mathbf{T}_{\mathcal{N}\mathcal{N}} \end{bmatrix}, \quad (5)$$

where $\mathcal{Z} \equiv \mathcal{A} \cup \{j\}$ and $\mathcal{N} \equiv \mathcal{Z}^c \equiv \mathcal{Q} \setminus \{j\}$. In Eq. (5), $\mathbf{T}_{\mathcal{Z}\mathcal{N}}$ contains the probabilities for transitions from nodes of the set \mathcal{N} to the set \mathcal{Z} , and the other blocks are defined similarly.

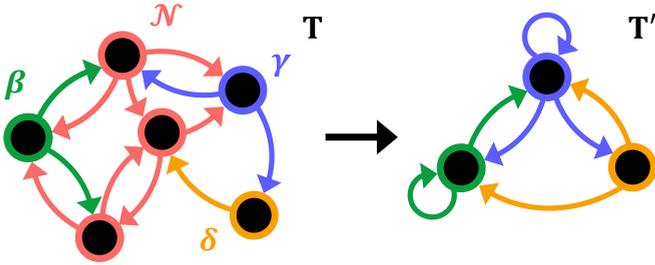


FIG. 1. Illustration of a renormalization operation performed on the stochastic matrix \mathbf{T} to yield the censored Markov chain \mathbf{T}' . To eliminate the block of nodes \mathcal{N} , the renormalized transition probabilities [Eq. (6)] in the resulting censored Markov chain (or *stochastic complement* [100]) must account for transitions that proceed via \mathcal{N} in the original network. Hence, the renormalized stochastic matrix includes a direct $\gamma \leftarrow \beta$ transition that is not present in the original network, which corresponds to the collective probabilities of all possible $\gamma \leftarrow \mathcal{N} \leftarrow \beta$ paths. By the same reasoning, the stochastic complement contains nonzero probabilities for $\beta \leftarrow \beta$, $\beta \leftarrow \gamma$, $\gamma \leftarrow \gamma$, and $\gamma \leftarrow \delta$ transitions. The probability for the $\delta \leftarrow \gamma$ transition does not increase with renormalization, and likewise there is no $\delta \leftarrow \beta$ transition in the derived stochastic complement, since there are no such indirect transitions proceeding via \mathcal{N} in the original Markov chain. The waiting times associated with all three of the retained nodes are increased in the censored Markov chain [Eq. (7)], since the transition probabilities to \mathcal{N} are nonzero for each of these nodes.

The diagonal blocks $\mathbf{T}_{\mathcal{Z}\mathcal{Z}}$ and $\mathbf{T}_{\mathcal{N}\mathcal{N}}$ are therefore square substochastic matrices [91,103]. The stochastic complement for the nodes in \mathcal{Z} is defined as [100–102]

$$\mathbf{T}'_{\mathcal{Z}\mathcal{Z}} = \mathbf{T}_{\mathcal{Z}\mathcal{Z}} + \mathbf{T}_{\mathcal{Z}\mathcal{N}}(\mathbb{I}_{\mathcal{N}\mathcal{N}} - \mathbf{T}_{\mathcal{N}\mathcal{N}})^{-1}\mathbf{T}_{\mathcal{N}\mathcal{Z}}, \quad (6)$$

where $\mathbb{I}_{\mathcal{N}\mathcal{N}}$ is the $|\mathcal{N}| \times |\mathcal{N}|$ -dimensional identity matrix. Equation (6) defines renormalized transition probabilities for the nodes in \mathcal{Z} . This transformed system is sometimes referred to as a *censored* Markov chain, because the renormalized probabilities correspond to the values that would be observed if the transitions within \mathcal{N} were obscured (see Fig. 1) [104–106]. With $\mathcal{Z} \equiv \mathcal{A} \cup \{j\}$, the only remaining transitions associated with node j in the renormalized network are to nodes of the set \mathcal{A} , and the $j \leftarrow j$ self loop. If the waiting times of the nodes in the set \mathcal{Z} , contained in the $|\mathcal{Z}|$ -dimensional vector $\boldsymbol{\tau}_{\mathcal{Z}}$, are renormalized according to [78,107],

$$\boldsymbol{\tau}'_{\mathcal{Z}} = \boldsymbol{\tau}_{\mathcal{Z}} + \boldsymbol{\tau}_{\mathcal{N}}(\mathbb{I}_{\mathcal{N}\mathcal{N}} - \mathbf{T}_{\mathcal{N}\mathcal{N}})^{-1}\mathbf{T}_{\mathcal{N}\mathcal{Z}}, \quad (7)$$

then the MFPT for the $\mathcal{A} \leftarrow j$ transition after eliminating the $|\mathcal{Q}| - 1$ nodes of set \mathcal{N} is given by [76]

$$\begin{aligned} \mathcal{T}_{\mathcal{A}j} &= [\boldsymbol{\tau}'_{\mathcal{Z}}]_j (1 - [\mathbf{T}'_{\mathcal{Z}\mathcal{Z}}]_{jj}) \sum_{n=1}^{\infty} n [\mathbf{T}'_{\mathcal{Z}\mathcal{Z}}]_{jj}^{n-1} \\ &= \frac{[\boldsymbol{\tau}'_{\mathcal{Z}}]_j}{1 - [\mathbf{T}'_{\mathcal{Z}\mathcal{Z}}]_{jj}}. \end{aligned} \quad (8)$$

Here, we have used the fact that $(1 - [\mathbf{T}'_{\mathcal{Z}\mathcal{Z}}]_{jj})^{-1}$ is the expected number of $j \leftarrow j$ transitions before node j is exited plus one for the final escape step [18].

The renormalization of the mean waiting (or lag) times [Eq. (7)] accounts for the expected number of transitions within the set of nodes to be eliminated, \mathcal{N} . That is, the updated waiting time for the j th node, $j \notin \mathcal{N}$, includes a contribution corresponding to the average of all $\mathcal{N}^c \leftarrow \mathcal{N} \leftarrow j$ paths that leave j , enter \mathcal{N} , and exit to \mathcal{N}^c . Because this contribution is not specific to which node \mathcal{N}^c is hit upon leaving \mathcal{N} , the GT algorithm [Eqs. (6) and (7)] preserves the *average* MFPT to the set of absorbing nodes \mathcal{A} , and not the individual MFPTs to particular absorbing nodes $a \in \mathcal{A}$ [76,107]. The renormalized $i \leftarrow j$ transition probabilities do preserve the probabilities of the (censored) paths to individual absorbing nodes, because the updated transition probabilities exactly account for the probability to transition from j to i via \mathcal{N} [73–75,100].

The mean first-passage path length for the $\mathcal{A} \leftarrow j \notin \mathcal{A}$ transition can be derived by direct analogy to the MFPT. In general, the two are not related for a CTMC parameterized by the branching probability matrix \mathbf{P} , because the mean waiting times for nodes are nonuniform [88]. For a path on the original Markov chain, an $i \leftarrow j$ transition increments the path length by one, and so the initial mean number of steps to exit node j , ℓ_j , are all unity. When nodes of the set \mathcal{N} are eliminated, an $i \leftarrow j$ transition on the renormalized network also includes steps taken within the censored region \mathcal{N} . Hence, the $\{\ell_j\}$ can be renormalized by analogy with Eq. (7), and the mean $\mathcal{A} \leftarrow j$ path length is then given by a relation analogous to Eq. (8), again with ℓ_j replacing τ_j .

We can also apply GT renormalization analogous to Eqs. (6) and (7) to eliminate nodes $n = 1, 2, \dots, |\mathcal{Q}| - 1$ one at a time, where $n \notin \mathcal{A}$ [73–78]. In a variation of this iterative procedure, where all $|\mathcal{Q}|$ transient nodes are now to be eliminated, the network of the $(n - 1)$ th iteration can be related to the network of the n th iteration via

$$T_{ij}^{(n-1)} = T_{ij}^{(n)} - U_{in}L_{nj}, \quad (9)$$

where the matrix \mathbf{U} has elements

$$U_{in} = T_{in}^{(n-1)} - \delta_{in}, \quad (10)$$

with δ_{in} the Kronecker delta, and the matrix \mathbf{L} has elements

$$L_{nj} = T_{nj}^{(n-1)} / (1 - T_{nn}^{(n-1)}). \quad (11)$$

In practice, the equivalence $1 - T_{nn} \equiv \sum_{\gamma \neq n} T_{\gamma n}$ is exploited to avoid subtraction operations and thus maintain numerical stability [61–64]. This iterative version of the GT algorithm can be thought of as a LU decomposition of a stochastic matrix [90,108]. The LU decomposition formulation of the GT algorithm [Eqs. (9) and (10)] gives $T_{nj}^{(n)} = 0$ and $T_{in}^{(n)} = T_{in}^{(n-1)} / (1 - T_{nn}^{(n-1)})$, thereby removing transitions to the eliminated node, n , and renormalizing the $i \leftarrow n$ transition probability to account for self-transitions. Hence, renormalization using Eqs. (9) and (10) preserves transitions from eliminated to noneliminated nodes, but not *vice versa*. Comparing Eqs. (7) and (11), the renormalized waiting time for the j th node in the censored Markov chain at the n th iteration can be written as

$$\tau_j^{(n)} = \tau_j^{(n-1)} + \tau_n^{(n-1)}L_{nj}. \quad (12)$$

The $\mathcal{A} \leftarrow j$ MFPT for the $j \equiv |\mathcal{Q}|$ th (transient) node, i.e., the last node to be eliminated, is therefore obtained straightforwardly as the associated renormalized waiting time in the censored Markov chain for which only the $|\mathcal{A}|$ absorbing nodes remain noneliminated. Although the preservation of transitions from eliminated nodes is not necessary to compute the $\mathcal{A} \leftarrow j$ MFPT, the formulation of the GT algorithm in Eqs. (9) and (10) can be used to compute other dynamical quantities such as committor probabilities, as we show in the following section.

C. Committor and absorption probabilities from graph transformation

If we define an initial macrostate $\mathcal{B} \subset \mathcal{A}^c$, then the transition probabilities for the renormalized Markov chain where all nodes of the intervening set $\mathcal{I} \equiv (\mathcal{A} \cup \mathcal{B})^c$ have been eliminated by the LU decomposition formulation of the GT algorithm relate straightforwardly to the committor probabilities. The $\mathcal{A} \leftarrow \mathcal{B}$ committor probability for the j th node, q_j^+ , is defined as the probability that a trajectory at node j first hits the target macrostate \mathcal{A} before returning to the initial macrostate \mathcal{B} [79–82]. By definition, $q_{a \in \mathcal{A}}^+ = 1$ and $q_{b \in \mathcal{B}}^+ = 0$ [93]. The committor probabilities of all other nodes satisfy a first-step relation [77,96],

$$q_j^+ = \sum_{\gamma \notin \mathcal{B}} T_{\gamma j} q_\gamma^+. \quad (13)$$

For the renormalized Markov chain where all nodes of the set \mathcal{I} have been eliminated according to Eqs. (9) and (10), with transition probabilities T'_{ij} , the only transitions from nodes of the set \mathcal{I} are directly to either of the endpoint macrostates \mathcal{A} or \mathcal{B} . The $\mathcal{A} \leftarrow \mathcal{B}$ committor probability for the j th node is therefore given straightforwardly by

$$q_j^+ = \sum_{a \in \mathcal{A}} T'_{aj} \equiv T'_{aj} = 1 - T'_{Bj}. \quad (14)$$

An analogous expression yields the committor probabilities for the reverse ($\mathcal{B} \leftarrow \mathcal{A}$) direction.

The definition of the committor probability in Eq. (14) is readily extended to the case where there are multiple *taboo* macrostates where visits are forbidden [83,84]. Let us define the set of macrostates $\mathcal{H} \equiv \{\mathcal{H}_1 \cup \dots \cup \mathcal{H}_N\} \subset \mathcal{S}$, which forms a subset of the complete state space. We wish to determine the probability of hitting a particular target macrostate \mathcal{H}_k before hitting any of the taboo nodes of the set $\mathcal{H} \setminus \mathcal{H}_k$, when the process is initialized at node $j \in \mathcal{H}^c$. The j th node is associated with separate committor probabilities corresponding to each of the first-passage processes defined by permuting the state \mathcal{H}_k that is considered to be the target. We denote by $q_j^{\mathcal{H}_k}$ the committor probability for the $\mathcal{H}_k \leftarrow \mathcal{H}^c$ transition, with all nodes of the set $\mathcal{H} \setminus \mathcal{H}_k$ considered taboo. These committor probabilities satisfy $q_j^{\mathcal{H}_1} + \dots + q_j^{\mathcal{H}_N} = 1 \forall j$. The committor probabilities of all nodes with respect to all first-passage processes can be determined efficiently by solving a single system of linear equations using the GT approach. That is, the transition probabilities of the renormalized network for which all nodes of the set \mathcal{H}^c have been eliminated using the LU decomposition formulation of the GT algorithm [Eqs. (9) and (10)] yield the various committor probabili-

ties via Eq. (14). We can then define a net reactive flux to the target state along an $i \leftarrow j$ edge for each first-passage process [109–113]. Specifically, for an irreducible Markov chain at equilibrium, the $i \leftarrow j$ net flux to the target macrostate \mathcal{H}_k when all nodes of the set $\mathcal{H} \setminus \mathcal{H}_k$ are taboo is

$$f_{ij}^{\mathcal{H}_k} = \begin{cases} \pi_j T_{ij} (q_i^{\mathcal{H}_k} - q_j^{\mathcal{H}_k}), & \text{if } q_i^{\mathcal{H}_k} > q_j^{\mathcal{H}_k}, \\ 0, & \text{otherwise,} \end{cases} \quad (15)$$

where π_j is the stationary (equilibrium occupation) probability of node j .

For an $\mathcal{A} \leftarrow \mathcal{B}$ transition, if we also eliminate the nodes $b \in \mathcal{B}$ of the initial macrostate according to Eqs. (9) and (10), thus leaving only the nodes of the absorbing macrostate \mathcal{A} , then the final renormalized transition probabilities are

$$T''_{aj} = T'_{aj} + T'_{aj} T'_{jj} / (1 - T'_{jj}) = T'_{aj} / T'_{Aj} \equiv B_{aj}. \quad (16)$$

Here, we have denoted the absorption (hitting) probability, i.e., the probability that trajectories initialized at the j th transient node, $j \in \mathcal{Q}$, will be absorbed at the a th absorbing node, $a \in \mathcal{A}$, by B_{aj} . The sum over absorbing nodes $a \in \mathcal{A}$ for B_{aj} is unity for all nodes j .

D. Extension of graph transformation with a backward pass phase

Following $|\mathcal{Q}| - 1$ renormalization steps of the standard formulation of the GT algorithm [Eqs. (6) and (7)] to eliminate a single node at each iteration, the network only has a single noneliminated transient node $j \equiv |\mathcal{Q}| \notin \mathcal{A}$, and the $\mathcal{A} \leftarrow j$ MFPT is given by

$$\mathcal{T}_{Aj} = \frac{\tau_j^{(j-1)}}{1 - T_{jj}^{(j-1)}}. \quad (17)$$

Working backwards to undo the GT procedure, in the previous iteration the first-step relation [Eq. (3)] gives

$$\mathcal{T}_{A,j-1} = \tau_{j-1}^{(j-2)} + \mathcal{T}_{A,j-1} T_{j-1,j-1}^{(j-2)} + \mathcal{T}_{Aj} T_{j,j-1}^{(j-2)},$$

$$\text{so } \mathcal{T}_{A,j-1} = \frac{\tau_{j-1}^{(j-2)} + \mathcal{T}_{Aj} T_{j,j-1}^{(j-2)}}{1 - T_{j-1,j-1}^{(j-2)}}, \quad (18)$$

and we can therefore determine $\mathcal{T}_{A,j-1}$ from \mathcal{T}_{Aj} if we save the necessary quantities from iteration $j - 2$. In general, we have the following expression to compute the MFPT for the $\mathcal{A} \leftarrow n$ transition, where n is the node that was eliminated at the n th iteration of the forward pass phase:

$$\begin{aligned} \mathcal{T}_{An} &= \tau_n^{(n-1)} + \sum_{n \leq \gamma \leq |\mathcal{Q}|} \mathcal{T}_{A\gamma} T_{\gamma n}^{(n-1)}, \\ &\tau_n^{(n-1)} + \sum_{n < \gamma \leq |\mathcal{Q}|} \mathcal{T}_{A\gamma} T_{\gamma n}^{(n-1)} \\ \text{or } \mathcal{T}_{An} &= \frac{\tau_n^{(n-1)} + \sum_{n < \gamma \leq |\mathcal{Q}|} \mathcal{T}_{A\gamma} T_{\gamma n}^{(n-1)}}{1 - T_{nn}^{(n-1)}}. \end{aligned} \quad (19)$$

Note the change in the summation range between the two lines of Eq. (19). Hence, we can work backwards and compute \mathcal{T}_{An} from $\mathcal{T}_{A\gamma}$ with $\gamma = n + 1, n + 2, \dots, |\mathcal{Q}|$ if we save $\tau_n^{(n-1)}$ and $T_{\gamma n}^{(n-1)}$ for $n \leq \gamma \leq |\mathcal{Q}|$ during the forward pass phase. An analogous scheme can be written for the committor probabilities, which obey a first-step relation [Eq. (13)] of the same form as that for MFPTs [Eq. (3)] [93].

Using the LU decomposition formulation of the GT algorithm [Eqs. (9) and (10)], we can write a more concise expression for the $\mathcal{A} \leftarrow n$ MFPT, where node n was eliminated at the n th iteration of the forward pass phase. This expression requires the renormalized probabilities for transitions from, and waiting time for, the n th node at the iteration where this node was eliminated, as well as the MFPTs for transitions from nodes that were eliminated after node n in the forward pass phase:

$$\mathcal{T}_{\mathcal{A}n} = \tau_n^{(n)} + \sum_{\gamma \notin \mathcal{A}} \mathcal{T}_{\mathcal{A}\gamma} T_{\gamma n}^{(n)}. \quad (20)$$

This equation follows from the fact that the $\gamma \leftarrow n$ transition probabilities for eliminated nodes $\gamma \leq n$ vanish in $\mathbf{T}^{(n)}$ by construction from Eq. (9).

The above derivation shows that we can calculate the MFPTs for *all* nonabsorbing nodes in a backward pass phase of the GT algorithm, by iteratively undoing the steps of the renormalization procedure and computing the MFPT for the newly restored node. This is the idea behind the extended GTH (EGTH) algorithm of Hunter [94,114]. As we have shown, it is not necessary to store the complete transition matrices at each iteration in the course of the forward (elimination) phase of the algorithm, and instead only a subset of waiting times and transition probabilities are required. Another convenient way to implement the backward pass phase of the algorithm that avoids excessive memory usage is to exploit the analogy between the GT algorithm and LU decomposition [Eq. (9)], in which case the MFPTs for restored nodes are computed via Eq. (20). This procedure has the advantage of simultaneously yielding the committor and absorption probabilities via Eqs. (14) and (16), respectively. The overall procedure is given as pseudocode in Algorithm 1 and illustrated in Fig. 2. The steps of the GTH algorithm [66,67] to compute the stationary distribution of an irreducible Markov chain can also be readily incorporated into this procedure.

There are numerous practical factors to consider in optimizing the efficiency and memory usage of state-reduction algorithms. Prioritizing renormalization of nodes with the lowest number of connections can speed up the calculation significantly [76]. The GT procedure in our PATHSAMPLE program uses a compressed row storage scheme at the start of a calculation for the MFPT, when the transition matrix is sparse. The GT renormalization adds nonzero probabilities as nodes are eliminated, and the program switches to dense storage when more than 2% of the elements became nonzero, if there are fewer than 11,000 remaining nodes. Our DISCOTRESS software [86] is designed similarly, employing a sparse data structure to keep memory requirements manageable and avoid unnecessary floating point operations when the network is large.

III. EXPECTED NUMBER OF NODE VISITS AND NODE VISITATION PROBABILITIES FOR FIRST-PASSAGE AND TRANSITION PATHS

A. Fundamental matrix of an absorbing Markov chain

Consider the substochastic $|\mathcal{Q}| \times |\mathcal{Q}|$ -dimensional matrix $\mathbf{T}_{\mathcal{Q}\mathcal{Q}}$ whose elements are the probabilities for transitions

within the set $\mathcal{Q} \equiv \mathcal{A}^c$. All nodes represented in $\mathbf{T}_{\mathcal{Q}\mathcal{Q}}$ must be transient. That is, it must be possible to reach the absorbing set of nodes \mathcal{A} from any node in \mathcal{Q} . Then the inverse $\mathbf{N}_{\mathcal{Q}\mathcal{Q}} = \mathbb{I}_{\mathcal{Q}\mathcal{Q}} + \mathbf{T}_{\mathcal{Q}\mathcal{Q}} + \mathbf{T}_{\mathcal{Q}\mathcal{Q}}^2 + \dots = (\mathbb{I}_{\mathcal{Q}\mathcal{Q}} - \mathbf{T}_{\mathcal{Q}\mathcal{Q}})^{-1}$ exists, and is called the *fundamental matrix* of the absorbing Markov chain [21]. Since the fundamental matrix of a reducible Markov chain with $|\mathcal{Q}|$ transient nodes is always a $|\mathcal{Q}| \times |\mathcal{Q}|$ -dimensional square matrix, in the following we will use the notation \mathbf{N} for brevity.

The element N_{ij} of the fundamental matrix is the expected number of times that the i th node is visited along a first-passage path initialized from node j [115]. Many more dynamical properties of interest can be written straightforwardly in terms of \mathbf{N} [18]. For example, the variance in the number of times that node i is visited prior to absorption when trajectories are initialized from node j is given by the relevant element of the matrix [18]

$$\mathbf{N}^{(2)} = \mathbf{N}(2\mathbf{N}_d - \mathbb{I}) - (\mathbf{N} \circ \mathbf{N}), \quad (21)$$

where \circ denotes the Hadamard (i.e., element-wise) product, and \mathbf{N}_d is the matrix whose only nonzero elements are the diagonal elements of \mathbf{N} . A general expression for the n th moment of this distribution, $\mathbf{N}^{(n)}$, is derived in Ref. [18]. The probabilities H_{ij} that the i th node is visited along first-passage paths initialized from the j th node, excluding the initial occupation of j , also follow directly from the N_{ij} elements. The mean number of visits to node i for such first-passage paths, N_{ij} , must be equal to the probability of hitting node i , multiplied by the mean number of visits to i prior to absorption for paths starting from i , plus one if i is the initial node:

$$N_{ij} = \delta_{ij} + H_{ij}N_{ii} \Rightarrow H_{ij} = (N_{ij} - \delta_{ij})/N_{ii}. \quad (22)$$

The above condition in matrix form is

$$\mathbf{H} = \mathbf{N}_d^{-1}(\mathbf{N} - \mathbb{I}). \quad (23)$$

The absorption probabilities B_{ij} are the elements of the matrix $\mathbf{B} = \mathbf{T}_{\mathcal{A}\mathcal{Q}}\mathbf{N}$, where $\mathbf{T}_{\mathcal{A}\mathcal{Q}}$ is the matrix of probabilities for transitions from \mathcal{Q} to \mathcal{A} . The absorption probabilities are nonzero only for $i \in \mathcal{A}$, $j \notin \mathcal{A}$.

The $\mathcal{A} \leftarrow \mathcal{B}$ MFPT can be obtained from \mathbf{N} via

$$\mathcal{T}_{\mathcal{A}\mathcal{B}} = \sum_{j \in \mathcal{Q}} \sum_{b \in \mathcal{B}} p_b(0) N_{jb} \tau_j, \quad (24)$$

for an initial probability distribution $\mathbf{p}(0)$ localized in \mathcal{B} , $\sum_{b \in \mathcal{B}} p_b(0) = 1$. Higher moments of the FPT distribution can also be determined given the elements of the fundamental matrix. For a DTMC, the waiting times for nodes are fixed and equal to the lag time τ . It can be shown that the vector with elements $\mathcal{V}_{\mathcal{A}j}$, i.e., the variance of the FPT distribution for transitions from the j th (transient) node, is [69]

$$\mathcal{V}_{\mathcal{A}} = ((2\mathbf{N}^T - \mathbb{I})\boldsymbol{\ell} - (\boldsymbol{\ell} \circ \boldsymbol{\ell}))\tau^2, \quad (25)$$

where $\boldsymbol{\ell} = \mathbf{N}^T \mathbf{1}_{\mathcal{Q}}$ is the vector of mean first-passage path lengths, and $\mathbf{1}_{\mathcal{Q}}$ is the $|\mathcal{Q}|$ -dimensional column vector with all elements equal to unity. See Ref. 18 for a derivation. In the continuous-time case, Eq. (25) gives the variances of the FPT distributions for transitions from nodes of the Markov chain parameterized by the linearized transition probability matrix [108] $\mathbf{T}_{\text{lin}}(\tau) = \mathbb{I} + \tau\mathbf{K}$, where $\tau \leq \min\{-K_{jj}^{-1} : \forall j\}$, for which the mean waiting times are uniform, $\tau_j \equiv \tau \forall j$ [115].

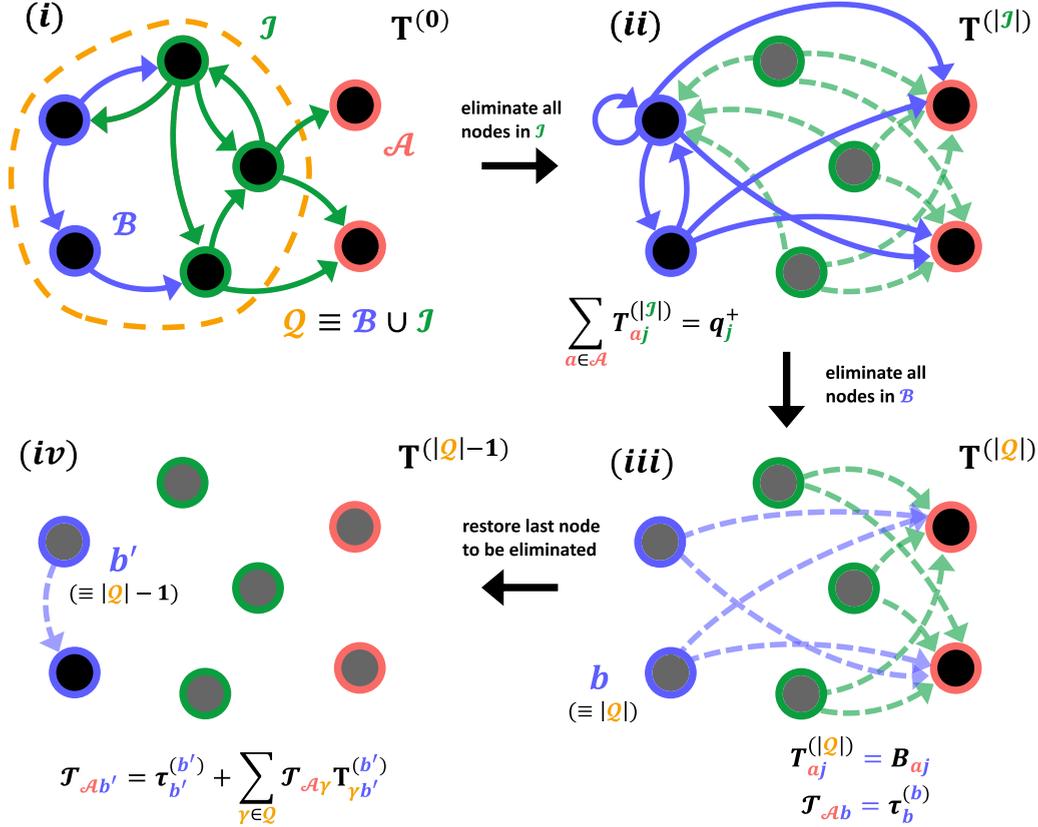


FIG. 2. Illustration of the LU decomposition formulation of the graph transformation (GT) algorithm with a backwards pass phase (Algorithm 1), which computes the committor probabilities and MFPTs for all transient nodes, as well as the absorption probabilities. The steps of the GTH algorithm [66,67] to compute the stationary distribution can also be incorporated into this procedure. Nodes that have been eliminated by renormalization (see Fig. 1) are shown as transparent. LU decomposition [Eqs. (9) and (10)] is used to renormalize transition probabilities, so that transitions from eliminated to noneliminated nodes are preserved (such connections are indicated by a transparent, dashed line). (i) A Markov chain for which the state space \mathcal{S} is divided into the set of absorbing nodes \mathcal{A} and the set of transient nodes \mathcal{Q} . The set of transient nodes is further divided into an initial macrostate, \mathcal{B} , and the set of intervening nodes, \mathcal{I} . (ii) In the first stage of the forward pass phase of the algorithm, all nodes of the state \mathcal{I} are iteratively eliminated by renormalization [Eqs. (9) and (10)], with the mean waiting (or lag) times for nodes renormalized according to Eq. (12). In the censored Markov chain where only nodes of the set $\mathcal{A} \cup \mathcal{B}$ remain noneliminated, the sum of transition probabilities from the j th transient node to absorbing nodes is the $\mathcal{A} \leftarrow \mathcal{B}$ committor probability for node j , q_j^+ [Eq. (13)]. (iii) In the remainder of the forward pass phase, the nodes of the initial state \mathcal{B} are iteratively eliminated. In the censored network where only absorbing nodes remain noneliminated, the renormalized $i \leftarrow j$ transition probabilities from transient to absorbing nodes are the absorption probabilities B_{ij} [Eq. (16)]. The MFPT for the $\mathcal{A} \leftarrow b$ transition, where $b \in \mathcal{B}$ was the last node to be eliminated, is equal to the renormalized waiting time for the b th node in the final censored Markov chain. (iv) In the backwards pass phase, eliminated nodes are iteratively restored using the \mathbf{L} and \mathbf{U} matrices that were constructed during the forward pass phase, and the MFPTs for transitions from transient nodes are computed by a recursive formula [Eq. (20)]. The figure shows the first step of this phase, in which the final node to be eliminated, b , for which the $\mathcal{A} \leftarrow b$ MFPT has previously been calculated, is restored. The $\mathcal{A} \leftarrow b'$ MFPT, where b' is the node that was eliminated before b , has two contributions. The first term is the renormalized mean waiting (or lag) time for the b' th node in the censored Markov chain of the b' th iteration. The second contribution corresponds to transitions to noneliminated transient nodes (here, $b \leftarrow b'$) of the relevant renormalized network.

B. Fundamental matrix of an absorbing Markov chain computed using state reduction

Inversion of the Markovian kernel $\mathbb{I}_{\mathcal{Q}\mathcal{Q}} - \mathbf{T}_{\mathcal{Q}\mathcal{Q}}$, which is required to compute the fundamental matrix \mathbf{N} of an absorbing Markov chain, is numerically unstable when the transition matrix features metastable macrostates. Therefore, as for the computation of MFPTs and committor probabilities in Markov chains exhibiting rare event dynamics (Sec. II), we wish to devise an inherently stable algorithm to robustly compute \mathbf{N} , and hence many additional dynamical properties

of interest. To this end, we define the augmented matrix

$$\mathbf{N}^* = \begin{pmatrix} \mathbf{T}_{\mathcal{Q}\mathcal{Q}} & \mathbb{I}_{\mathcal{Q}\mathcal{Q}} \\ \mathbb{I}_{\mathcal{Q}\mathcal{A}} & \mathbf{0}_{\mathcal{Q}\mathcal{A}} \\ \mathbf{T}_{\mathcal{A}\mathcal{Q}} & \mathbf{0}_{\mathcal{A}\mathcal{Q}} \end{pmatrix}, \quad (26)$$

where $\mathbf{0}_{\mathcal{Q}\mathcal{A}}$ is the $|\mathcal{Q}| \times |\mathcal{A}|$ -dimensional null matrix. Evidently, \mathbf{N}^* does not relate to a stochastic matrix, since the column sums corresponding to transition probabilities from transient nodes necessarily exceed unity. Nonetheless, if we proceed to compute the analog of the stochastic complement

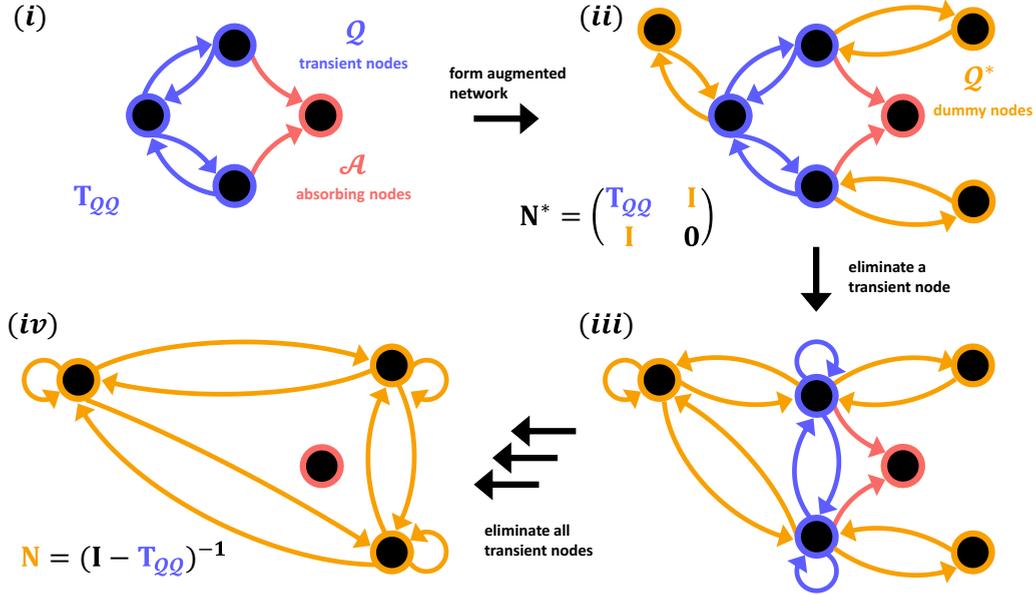


FIG. 3. Illustration of the numerically stable state-reduction procedure to compute the fundamental matrix \mathbf{N} for an absorbing Markov chain, effectively performing a matrix inversion operation on the Markovian kernel $\mathbb{I}_{\mathcal{Q}\mathcal{Q}} - \mathbf{T}_{\mathcal{Q}\mathcal{Q}}$. (i) The state space $\mathcal{S} \equiv \mathcal{Q} \cup \mathcal{A}$ of the Markov chain is divided into sets of transient and absorbing nodes, denoted \mathcal{Q} and \mathcal{A} , respectively. The substochastic matrix $\mathbf{T}_{\mathcal{Q}\mathcal{Q}}$ only includes transition probabilities between transient nodes (blue), and does not include absorbing nodes (red). (ii) Dummy nodes (yellow), of the set \mathcal{Q}^* , are partnered with transient nodes via forward and reverse edges with weights equal to unity. (iii) State reduction [cf. Eq. (6)] is used to eliminate transient nodes either iteratively [Eq. (27)] or in blocks [Eq. (28)]. The updated $i \leftarrow j$ edge weights account for paths that proceed via the eliminated nodes. Transitions from dummy to absorbing nodes do not have a meaningful interpretation and are not required in the algorithm, so can be ignored. (iv) The $i \leftarrow j$ edge weights in the network where only the dummy nodes remain are the elements N_{ij} of the fundamental matrix \mathbf{N} .

[Eq. (6)] corresponding to the remaining network when all nodes represented in $\mathbf{T}_{\mathcal{Q}\mathcal{Q}}$ are eliminated from \mathbf{N}^* , then we obtain the fundamental matrix \mathbf{N} .

Specifically, the proposed state-reduction algorithm to compute the fundamental matrix \mathbf{N} of an absorbing Markov chain is as follows. For each of the transient nodes in the network, of the set \mathcal{Q} , we introduce a dummy partner node. Thus, we have the augmented state space $\mathcal{S}^* \equiv \mathcal{S} \cup \mathcal{Q}^*$, where \mathcal{Q}^* denotes the set of dummy nodes, with $|\mathcal{Q}^*| \equiv |\mathcal{Q}|$. Each dummy node is connected to its transient partner by forward and reverse edges, both with weights equal to unity [cf. Eq. (26)]. When all of the transient nodes have been eliminated, via the analog of a stochastic complement [Eq. (6)], the weights of the $i \leftarrow j$ edges in the remaining network that correspond to transitions between dummy nodes are the elements N_{ij} of the fundamental matrix \mathbf{N} . In the nodewise iterative formulation of this procedure, upon eliminating a single transient node n , the edge weights for transitions between all remaining nodes in the augmented network are updated according to

$$N_{ij}^* \leftarrow N_{ij}^* + \frac{N_{in}^* N_{nj}^*}{\sum_{\gamma \in \mathcal{S} \setminus \{n\}} N_{\gamma n}^*} \quad \forall i, j \in \mathcal{S}^* \setminus \{n\}. \quad (27)$$

This state-reduction procedure is illustrated in Fig. 3 and given as pseudocode in Algorithm 2. Equation (27) shows the advantage of including the absorbing nodes in the augmented state space. In particular, if the probabilities for transitions

from transient to absorbing nodes are renormalized in the course of the algorithm, then the total probabilities for transitions from transient to nondummy nodes remain conserved, and equal to unity. Hence, since only the transient nodes are eliminated, all subtraction operations can be avoided by exploiting the relation $1 - T_{nn} = \sum_{\gamma \neq n} T_{\gamma n}$, where \mathbf{T} denotes the stochastic matrix of the censored Markov chain comprising the noneliminated nodes of the state space \mathcal{S} . Using this trick, the state-reduction algorithm is numerically stable [61–64].

The theory of stochastic complements presented in Sec. II B can be leveraged to design a block formulation of this state-reduction algorithm, wherein multiple nodes are eliminated simultaneously [100]. Let us consider the elimination of a set of transient nodes $\mathcal{N} \subseteq \mathcal{Q}$, and denote the set of all remaining nonabsorbing nodes in the augmented state space as $\mathcal{Z} \equiv \mathcal{S}^* \setminus (\mathcal{A} \cup \mathcal{N})$. The augmented matrix [Eq. (26)] is then updated according to [cf. Eq. (6)]

$$\mathbf{N}_{\mathcal{Z}\mathcal{Z}}^* \leftarrow \mathbf{N}_{\mathcal{Z}\mathcal{Z}}^* + \mathbf{N}_{\mathcal{Z}\mathcal{N}}^* (\mathbb{I}_{\mathcal{N}\mathcal{N}} - \mathbf{T}_{\mathcal{N}\mathcal{N}})^{-1} \mathbf{N}_{\mathcal{N}\mathcal{Z}}^*, \quad (28)$$

where we have used the notation $\mathbf{N}_{\mathcal{Z}\mathcal{Z}}^*$ to explicitly indicate the dimensionality of the augmented network. For $\mathcal{N} \equiv \{n\}$, Eq. (28) reduces to Eq. (27). After eliminating all transient nodes, the resulting matrix $\mathbf{N}_{\mathcal{Z}\mathcal{Z}}^*$ is the fundamental matrix \mathbf{N} . This procedure is numerically stable if the blocks of nodes to be eliminated, \mathcal{N} , correspond to metastable macrostates, so that the matrix inversion operations for the Markovian kernels $\mathbb{I}_{\mathcal{N}\mathcal{N}} - \mathbf{T}_{\mathcal{N}\mathcal{N}}$ are not associated with significant numerical

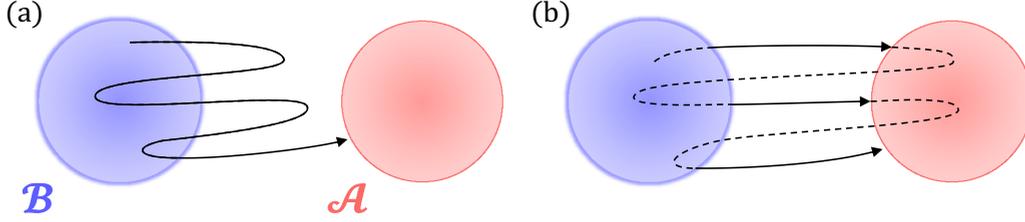


FIG. 4. Schematic depiction of the nonequilibrium and equilibrium (i.e., steady state) $\mathcal{A} \leftarrow \mathcal{B}$ first-passage and transition path ensembles (FPPE and TPE, respectively). (a) Trajectories of the nonequilibrium FPPE start within the initial state \mathcal{B} and are absorbed upon hitting the target state \mathcal{A} . The TPE (solid line) is the portion of the FPPE that transitions directly to \mathcal{A} from \mathcal{B} without revisiting \mathcal{B} . (b) For an irreducible Markov chain, we can also consider the equilibrium FPPE and TPE, which result from considering an infinitely long trajectory that continually transitions between the \mathcal{B} and \mathcal{A} states. The steady-state $\mathcal{A} \leftarrow \mathcal{B}$ TPE is the set of path segments that transition directly from \mathcal{B} to \mathcal{A} at equilibrium. The steady-state MFPT is the inverse of the rate at which trajectories that last visited the initial state hit the target state.

error. This formulation of the algorithm therefore requires a careful partitioning of the nodes into appropriate communities, but leads to improved time complexity [107,116].

C. Reactive and nonreactive segments of the first-passage path ensemble

Knowledge of the committor probabilities, $\{q_j^+\}$ [Eq. (14)], and the fundamental matrix of the absorbing Markov chain, \mathbf{N} (Sec. III A), can be exploited to divide the $\mathcal{A} \leftarrow \mathcal{B}$ first-passage path ensemble (FPPE) into nonreactive ($\mathcal{B} \leftarrow \mathcal{B}$) and reactive (direct $\mathcal{A} \leftarrow \mathcal{B}$) segments [Fig. 4(a)] [17]. This division allows for a more detailed analysis of the FPPE at a nodewise level of detail, beyond the results outlined in Sec. III A. The reactive segments of the FPPE, which correspond to the transition path ensemble (TPE) [109–113], are particularly insightful to understand the characteristics of the productive $\mathcal{A} \leftarrow \mathcal{B}$ process. In Sec. III D, we will derive novel analytical results for key dynamical properties characterizing the influence of individual nodes on direct $\mathcal{A} \leftarrow \mathcal{B}$ transitions, such as the probability that a particular node is visited along a reactive (transition) path [Eq. (44)]. Since we have shown that both the $\{q_j^+\}$ and \mathbf{N} can be computed robustly by state-reduction methods (Secs. II C and III B, respectively), we can likewise compute the derived properties by a numerically stable route. In the remainder of the current section, we formally introduce the factorization of first-passage paths into reactive and nonreactive trajectory segments.

The expected number of times that a nonabsorbing node is visited along a first-passage path is simply an average of the mean number of visits when starting from the initial state $\mathcal{B} \subseteq \mathcal{Q}$, taken over the initial occupation probability distribution localized within this set:

$$\theta_j = \sum_{b \in \mathcal{B}} p_b(0) N_{jb} \quad \forall j \in \mathcal{Q}, \quad (29)$$

with $\sum_{b \in \mathcal{B}} p_b(0) = 1$. Absorbing nodes can only be visited once along a particular first-passage path, with probability B_{aj} , and the average over the initial distribution is

$$\theta_a = \sum_{j \notin \mathcal{A}} p_j(0) B_{aj} \quad \forall a \in \mathcal{A}. \quad (30)$$

The first-step relation for the elements of the fundamental matrix of the absorbing Markov chain, which includes only

transient nodes of the set \mathcal{Q} , is [18]

$$N_{ij} = \delta_{ij} + \sum_{\gamma \in \mathcal{Q}} [\mathbf{T}_{\mathcal{Q}\mathcal{Q}}]_{j\gamma} N_{i\gamma}. \quad (31)$$

Note that this expression does not have the same form as the first-step relations for the MFPTs, committor probabilities, or absorption probabilities [cf. Eq. (3)], and therefore the state-reduction algorithms presented in Secs. II B–II D cannot be used to compute the $\{\theta_j\}$. Since $\mathbf{T}_{\mathcal{Q}\mathcal{Q}}\mathbf{N} = \mathbf{N}\mathbf{T}_{\mathcal{Q}\mathcal{Q}}$, which follows from writing \mathbf{N} as a geometric progression in $\mathbf{T}_{\mathcal{Q}\mathcal{Q}}$, we can rewrite the first-step relation [Eq. (31)] and sum over the initial distribution within \mathcal{B} to obtain

$$\begin{aligned} \sum_{b \in \mathcal{B}} p_b(0) N_{jb} &= \sum_{b \notin \mathcal{B}} p_b(0) \delta_{jb} + \sum_{\gamma \in \mathcal{Q}} [\mathbf{T}_{\mathcal{Q}\mathcal{Q}}]_{j\gamma} \sum_{b \in \mathcal{B}} p_b(0) N_{\gamma b} \\ \text{so } \theta_j &= p_j(0) + \sum_{\gamma \in \mathcal{Q}} [\mathbf{T}_{\mathcal{Q}\mathcal{Q}}]_{j\gamma} \theta_\gamma \quad \forall j \in \mathcal{Q}. \end{aligned} \quad (32)$$

The absorption probability matrix is $\mathbf{B} = \mathbf{T}_{\mathcal{A}\mathcal{Q}}\mathbf{N}$, so for absorbing nodes we have

$$\begin{aligned} \theta_a &= \sum_{b \in \mathcal{B}} p_b(0) B_{ab} = \sum_{\gamma \in \mathcal{Q}} [\mathbf{T}_{\mathcal{A}\mathcal{Q}}]_{a\gamma} \sum_{b \in \mathcal{B}} p_b(0) N_{\gamma b} \\ &= \sum_{\gamma \in \mathcal{Q}} [\mathbf{T}_{\mathcal{A}\mathcal{Q}}]_{a\gamma} \theta_\gamma \quad \forall a \in \mathcal{A}. \end{aligned} \quad (33)$$

Hence, the $\{\theta_j\}$ satisfy the following system of linear equations:

$$\theta_j = p_j(0) + \sum_{\gamma \in \mathcal{Q}} T_{j\gamma} \theta_\gamma \quad \forall j \in \mathcal{S}, \quad (34)$$

where $p_j(0) = 0$ for $j \notin \mathcal{B}$. Equation (34) can be solved directly by standard linear algebra methods, but the $\{\theta_j\}$ are most robustly determined via Eqs. (29) and (30) when \mathbf{N} is computed using the state-reduction algorithm described in Sec. III B.

We can now break down properties of the FPPE into contributions from reactive and nonreactive path segments [17,112,117]. Recall that nodes not belonging to endpoint states are members of the intervening set $\mathcal{I} \equiv (\mathcal{A} \cup \mathcal{B})^c$. A reactive path from \mathcal{B} is one that leaves \mathcal{B} and reaches \mathcal{A} without returning to \mathcal{B} [110]. Nonreactive paths contain nodes from $\mathcal{B} \cup \mathcal{I}$ in first-passage path segments starting in \mathcal{B} up to the final escape from \mathcal{B} before reaching \mathcal{A} . The

average numbers of visits to the j th node along nonreactive and reactive paths, $\bar{\theta}_j$ and $\tilde{\theta}_j$, respectively, are given by [17]

$$\bar{\theta}_j = \theta_j(1 - q_j^+), \quad (35a)$$

$$\tilde{\theta}_j = \mu_j + \theta_j q_j^+, \quad (35b)$$

where

$$\mu_j = 1_{\mathcal{B}}(j) \bar{\theta}_j \sum_{\gamma} T_{\gamma j} q_{\gamma}^+, \quad (36)$$

is the probability that a reactive path left the initial state \mathcal{B} from node $j \in \mathcal{B}$. Here, $1_{\mathcal{B}}(j)$ is the indicator function for the initial region, equal to unity for $j \in \mathcal{B}$ and zero otherwise, which ensures that the initial probability distribution in Eq. (36) is contained within \mathcal{B} . Let $\partial\mathcal{B} \subseteq \mathcal{B}$ denote the boundary nodes of the initial set, i.e., nodes of the initial set for which a direct transition to a node of the set \mathcal{B}^c exists. Then $\sum_{b \in \partial\mathcal{B}} \mu_b = 1$ and $\mu_j = 0 \forall j \notin \partial\mathcal{B}$. Equation (35b) simply states that the expected number of times that a *reactive* trajectory, beginning at the boundary $\partial\mathcal{B}$ of the initial state \mathcal{B} , visits a node $j \notin \mathcal{B}$ is the product of the expected number of times that any first-passage trajectory visits the j th node and the probability that a trajectory initialized from node j is reactive.

The decomposition of the FPPE into reactive and nonreactive segments allows for analysis of the individual nodes and edges of the Markovian network that make significant and productive contributions to the $\mathcal{A} \leftarrow \mathcal{B}$ process. The flux along the $i \leftarrow j$ edge of the network is defined as $J_{ij} = \theta_j T_{ij}$ [17]. This flux can also be split into nonreactive and reactive contributions \bar{J}_{ij} and \tilde{J}_{ij} , respectively,

$$J_{ij} \equiv \theta_j T_{ij} = \bar{J}_{ij} + \tilde{J}_{ij}, \quad (37)$$

where the reactive flux along the $i \leftarrow j$ edge is given by [17]

$$\tilde{J}_{ij} = \frac{\tilde{\theta}_j T_{ij} q_i^+}{\sum_{\gamma} T_{\gamma j} q_{\gamma}^+}, \quad (38)$$

for $i \in \mathcal{I} \cup \mathcal{A}$, $j \in \mathcal{I}$, and when the set \mathcal{A} is reachable from both nodes i and j . Here, “reachable” means that a path to \mathcal{A} exists that passes through the \mathcal{I} set without hitting \mathcal{B} . Other than this condition, the derivation of Eqs. (34)–(38) does not assume that the Markov chain is ergodic [17]. \tilde{J}_{ij} is essentially the nonequilibrium analog of the stationary (i.e., equilibrium) reactive flux f_{ij}^+ [Eq. (15)].

D. Analysis of reactive paths

A further key dynamical property characterizing the ensemble of $\mathcal{A} \leftarrow \mathcal{B}$ transition (i.e., reactive) paths [112] is the conditional probability that the i th node is visited along a trajectory initialized from node j when the trajectory is reactive. We shall denote this quantity by \tilde{H}_{ij} . To simplify the notation in deriving this probability, we assume that there are no nodes in the set \mathcal{I} from which the absorbing macrostate \mathcal{A} is not reachable, since such nodes do not contribute to the reactive segment of the FPPE. Similarly, it is not necessary to consider nodes of the set $\mathcal{B} \setminus \partial\mathcal{B}$. If there are no such internal initial nodes, so that $\partial\mathcal{B} \equiv \mathcal{B}$, then we have $|\mathcal{Q}|$ nodes in the relevant set of transient nodes $\partial\mathcal{B} \cup \mathcal{I} \subseteq \mathcal{Q}$. For brevity, we

shall assume this to be the case, and we therefore consider the $|\mathcal{Q}| \times |\mathcal{Q}|$ -dimensional substochastic matrix $\mathbf{T}_{\mathcal{Q}\mathcal{Q}}$.

To derive the \tilde{H}_{ij} probabilities, we introduce the substochastic transition probability matrix for the reactive process on the set of (relevant) transient nodes, $\tilde{\mathbf{T}}_{\mathcal{Q}\mathcal{Q}}$ [117]. We also define the $|\mathcal{Q}|$ -dimensional vector of committor probabilities for the relevant transient nodes, $\mathbf{q}_{\mathcal{Q}}^+$, and the modified committor probability vector $\mathbf{q}_{\mathcal{Q}}^{+'}$, for which the elements corresponding to initial boundary nodes are nonzero, equal to $q_{b \in \partial\mathcal{B}}^{+'} = \sum_{\gamma \notin \mathcal{B}} T_{\gamma b} q_{\gamma}^+$. This probability is the probability that a trajectory is absorbed before *hitting* any node of the set \mathcal{B} [cf. Eq. (13)] [76]. Then the reactive transition probability matrix for transient nodes is [17]

$$\tilde{\mathbf{T}}_{\mathcal{Q}\mathcal{Q}} = \text{diag}(\mathbf{q}_{\mathcal{Q}}^+) \mathbf{T}_{\mathcal{Q}\mathcal{Q}} \text{diag}(\mathbf{q}_{\mathcal{Q}}^{+'})^{-1}, \quad (39)$$

and can be evaluated robustly by using a state-reduction algorithm to compute the committor probabilities (Sec. II C). We note again that internal initial nodes are discarded in this representation, since such nodes do not contribute to the reactive segment of the FPPE, i.e., $q_{b \in \mathcal{B} \setminus \partial\mathcal{B}}^{+'} = 0$. The corresponding fundamental matrix for the reactive process is

$$\tilde{\mathbf{N}}_{\mathcal{Q}\mathcal{Q}} = (\mathbb{I}_{\mathcal{Q}\mathcal{Q}} - \tilde{\mathbf{T}}_{\mathcal{Q}\mathcal{Q}})^{-1}, \quad (40)$$

and can be computed using the numerically stable state-reduction algorithm derived in Sec. III B. We shall henceforth omit the dimensionality subscripts from $\tilde{\mathbf{N}}$ for notational simplicity. The fundamental matrix for the reactive process provides a natural means to express the expected number of visits to a transient node along a reactive path:

$$\tilde{\theta}_j = \sum_{b \in \partial\mathcal{B}} \mu_b \tilde{N}_{jb} \quad \forall j \in \mathcal{Q}. \quad (41)$$

Recall that for absorbing nodes we simply have $\tilde{\theta}_a = \theta_a \forall a \in \mathcal{A}$. By analogy with the visitation probability matrix \mathbf{H} associated with the FPPE [Eq. (23)], we can calculate \tilde{H}_{ij} , the probability that a *reactive* trajectory will ever visit node i if it starts at node j for $i, j \in \mathcal{Q}$, not counting the occupancy of the initial node:

$$\tilde{N}_{ij} = \delta_{ij} + \tilde{H}_{ij} \tilde{N}_{ii}. \quad (42)$$

Hence, we obtain the matrix $\tilde{\mathbf{H}}$:

$$\tilde{\mathbf{H}} = \tilde{\mathbf{N}}_d^{-1} (\tilde{\mathbf{N}} - \mathbb{I}). \quad (43)$$

Thus, the reactive visitation probabilities \tilde{H}_{ij} can be determined robustly by using state-reduction algorithms to compute the committor probability vector, $\mathbf{q}_{\mathcal{Q}}^+$, and the fundamental matrix for the reactive process, $\tilde{\mathbf{N}}$.

We can similarly define a substochastic transition matrix corresponding to the nonreactive process on the set of transient nodes. This Markov chain is constructed so that “absorption” corresponds to the first-passage trajectory hitting a node at the boundary of the initial state for the final time, after which point the trajectory proceeds to be reactive [17]. Fundamental and visitation probability matrices for the nonreactive segment of the FPPE then follow by an analogous argument to the reactive case.

The probability that the j th noninitial transient node is visited along a reactive $\mathcal{A} \leftarrow \mathcal{B}$ transition path [118], r_j^+ , is an average of the \tilde{H}_{jb} elements with respect to the initial

occupation probability distribution for reactive trajectories μ [Eq. (36)],

$$r_j^+ = \sum_{b \in \partial \mathcal{B}} \mu_b \tilde{H}_{jb} \quad \forall j \in \mathcal{Q} \setminus \partial \mathcal{B}. \quad (44)$$

For initial boundary nodes $b \in \partial \mathcal{B}$, $\tilde{H}_{bj} = 0 \forall j$, since initial nodes cannot be revisited along reactive paths by definition. The probability that a node at the boundary of the initial state appears along a reactive path is therefore simply $r_b^+ = \mu_b$. The probability that an absorbing node $a \in \mathcal{A}$ appears along a reactive path is an average of the elements of the absorption probability matrix [Eq. (16)] weighted by the μ distribution; $r_a^+ = \sum_{b \in \partial \mathcal{B}} \mu_b B_{ab}$.

The reactive visitation probability r_j^+ provides detailed characterization of the FPPE at a microscopic level of detail. Nodes that have a high probability of being visited along reactive trajectories are those that mediate the dominant pathways for the overall productive transition. For an effective two-state system, nodes that are associated with a high r_j^+ probability, and which also have values for the committor probability q_j^+ close to 0.5, represent the dynamical bottleneck region of the network [119–121]. That is, these nodes constitute the transition state ensemble (TSE) [122]. Global dynamical quantities, including the $\mathcal{A} \leftarrow \mathcal{B}$ MFPT [123], are most sensitive to perturbations in the transition probabilities associated with these bottleneck nodes [124,125]. Therefore, the reactive visitation and committor probabilities are the central objects in understanding how the local dynamics at a small subset of nodes, namely, the TSE, modulate the slow, macroscopic dynamics. In general, for systems exhibiting multiple metastable macrostates, there are multiple TSEs that are the boundary regions between the metastable states, across which the committor probability changes sharply.

The visitation probability of the j th node along reactive trajectories in Eq. (44) corresponds to the *nonequilibrium* TPE [17]. If the set of initial boundary nodes $\partial \mathcal{B}$ contains more than one node, then Eq. (44) differs from the result when the system is at a steady state, i.e., corresponding to the *equilibrium* TPE [112]. The two path ensembles are illustrated schematically in Fig. 4. In the steady-state regime, which exists if the Markov chain is irreducible [100], the $\mathcal{A} \leftarrow \mathcal{B}$ path ensemble has relaxed to equilibrium. The probability μ_j^{SS} that the reactive portion of trajectories began after the nonreactive trajectory segment hit node j is therefore dependent on the stationary distribution [17],

$$\mu_j^{\text{SS}} \propto 1_{\mathcal{B}}(j) \pi_j \sum_{\gamma} T_{\gamma j} q_{\gamma}^+. \quad (45)$$

Similar to the nonequilibrium case [Eq. (36)], this initial distribution for reactive trajectories at steady state satisfies $\sum_{b \in \partial \mathcal{B}} \mu_b^{\text{SS}} = 1$ and $\mu_j^{\text{SS}} = 0 \forall j \notin \partial \mathcal{B}$. That is, this distribution is localized at the boundary of the initial state. The visitation probability of the j th noninitial transient node along reactive trajectories for the equilibrium TPE is a weighted average of the elements of the $\tilde{\mathbf{H}}$ matrix [Eq. (43)] with respect to this initial occupation probability distribution,

$$r_j^{+, \text{SS}} = \sum_{b \in \partial \mathcal{B}} \mu_b^{\text{SS}} \tilde{H}_{jb} \quad \forall j \in \mathcal{Q} \setminus \partial \mathcal{B}. \quad (46)$$

In addition, $r_b^{+, \text{SS}} = \mu_b^{\text{SS}} \forall b \in \partial \mathcal{B}$ and $r_a^{+, \text{SS}} = \sum_{b \in \partial \mathcal{B}} \mu_b^{\text{SS}} B_{ab} \forall a \in \mathcal{A}$. Similarly, the average number of times that the transient node j is visited along reactive trajectories at steady state is a weighted average of the elements of the $\tilde{\mathbf{N}}$ matrix [Eq. (40)] with respect to the μ^{SS} distribution,

$$\tilde{\theta}_j^{\text{SS}} = \sum_{b \in \partial \mathcal{B}} \mu_b^{\text{SS}} \tilde{N}_{jb} \quad \forall j \in \mathcal{Q}, \quad (47)$$

and for absorbing nodes we have $\tilde{\theta}_a^{\text{SS}} = \sum_{b \in \partial \mathcal{B}} \mu_b^{\text{SS}} B_{ab} \forall a \in \mathcal{A}$.

Recall that within the state-reduction formalism, a single linear system of equations can be solved to robustly obtain the set of committor probabilities $\{\mathbf{q}^{\mathcal{H}_1}, \dots, \mathbf{q}^{\mathcal{H}_N}\}$, where each committor probability vector $\mathbf{q}^{\mathcal{H}_k}$ is associated with a different target macrostate $\mathcal{H}_k \in \mathcal{H}$, conditioned on all nodes of the set $\mathcal{H} \setminus \mathcal{H}_k$ being taboo (Sec. II C). This formulation allows us to compute all the dynamical properties that we have derived relating to the reactive segments of the FPPEs for alternative $\mathcal{H}_k \leftarrow \mathcal{H}^c$ transitions. This result is useful if, for example, we want to analyze transition paths associated with a particular sequence of events, which form a subset of the ensemble of all paths transitioning to the target state. This analysis can be achieved by setting states that are not involved in the paths of interest to be taboo. For instance, a Markov chain representing the folding transition of a protein may feature several competing mechanisms that can be distinguished on the basis of the intermediate metastable states that are visited [27,126–128]. By designating a particular intermediate state to be taboo, we can investigate the TPE specifically for transitions that proceed via alternative intermediate states.

IV. NUMERICAL RESULTS

We demonstrate the methodology outlined in Secs. II and III with numerical results for a kinetic network representing a structural transition for a cluster of 38 atoms bound by the Lennard-Jones potential (LJ₃₈) [129,130]. Specifically, we consider the transition from a structure based on an incomplete Mackay icosahedron (I_h) to a face-centered cubic (F) geometry, which was also analyzed in Ref. [38]. The network model was constructed by mapping the local minima and transition states of the underlying potential energy landscape to the nodes and edges of a CTMC, which consists of 885 nodes and 1126 bidirectional edges. The face-centered cubic (F) state is represented by the single node of the Markov chain with the largest stationary probability (lowest free energy), and the icosahedral (I_h) state is represented by the single node with lowest free energy belonging to a separate funnel on the landscape. Because these two competing low-energy nodes are separated by a large energy barrier, the $F \leftarrow I_h$ solid-solid transition becomes an increasingly rare event [59] with decreasing temperature. We employ standard reduced units for the LJ potential in the following analysis [129,130].

We analyze the Markov chain parameterized at a temperature of $T = 0.12$, which approximately coincides with the start of the regime where the kinetic network exhibits significant metastability. At this temperature, the number of internode transitions in $F \leftarrow I_h$ first-passage paths is

typically 10^8 or 10^9 , precluding the use of the standard kinetic Monte Carlo [131] (kMC) algorithm to sample the FPPE. Moreover, dense linear algebra methods to perform an eigen-decomposition of the Markov chain, to invert the Markovian kernel $\mathbb{I}_{\mathcal{Q}\mathcal{Q}} - \mathbf{T}_{\mathcal{Q}\mathcal{Q}}$ (required to compute the fundamental matrix \mathbf{N} and the stochastic complement [Eq. (6)]), or to solve relevant linear systems of equations [Eqs. (3), (13), and (34)] suffer from a severe propagation of numerical error arising from finite precision [116]. Similarly, iterative sparse linear algebra methods [42] fail to converge. Computational analysis of this Markov chain is therefore intractable without employing the state-reduction algorithms described in Secs. II and III.

Figure 5 shows the results of various state-reduction calculations to robustly compute the salient dynamical properties associated with individual nodes of the kinetic network, for the $F \leftarrow I_h$ transition. The Markov chain is visualized as a disconnectivity graph [132], where the interconvertibility of sets of nodes in the network is considered at decreasing threshold increments representing the available energy. A fork in the graph indicates that a transition between the sets of nodes requires energy exceeding the threshold, and the branches terminate at the energies of the corresponding nodes. MFPTs to the F state and committor probabilities were computed by the iterative LU decomposition formulation of the GT algorithm with a backward pass phase (Algorithm 1). The expected number of node visits along reactive paths, θ_j [Eq. (41)], were obtained from the fundamental matrix for the reactive process [Eq. (40)], computed using the state-reduction procedure given in Algorithm 2. The reactive visitation probabilities for nodes were determined from the elements of the reactive fundamental matrix and from the committor probabilities via Eqs. (39)–(44). The exact results from the state-reduction calculations were verified numerically by comparison with kinetic path sampling simulations [108,133], an advanced method to sample the numbers of internode transitions along trajectories that is unaffected by metastability [118].

Inspection of the committor probabilities [Fig. 5(a)] reveals that the network is effectively a two-state system, with the I_h and F nodes representing strong attractors that characterize the respective regions of the state space. That is, there are relatively few nodes with intermediate values for the committor probability ($q_j^+ \approx 0.5$), and instead the vast majority of nodes are strongly associated with relaxation to either the I_h or the F state (indicated by committor probabilities $q_j^+ \approx 0$ and $q_j^+ \approx 1$, respectively). The MFPTs to the F state are $\mathcal{T}_{F \leftarrow j} \approx 10^9$ for most nodes j , although there are a small number of nodes associated with extreme values for $\mathcal{T}_{F \leftarrow j}$. In particular, nodes separated from the F state by small energy barriers relax to the F state comparatively rapidly ($\mathcal{T}_{F \leftarrow j} \approx 10^4$), but there are nodes that constitute kinetic traps, for which transitions to F correspond to very long timescales ($\mathcal{T}_{F \leftarrow j} \approx 10^{18}$).

A striking feature of the network is the localization of the reactive dynamics to a small subset of nodes, demonstrating that there are strongly preferred pathways for the $F \leftarrow I_h$ transition at this temperature. On average, only around 10 % of nodes are visited more than once along a reactive $F \leftarrow I_h$ path, and the average number of visits is $\theta_j < 10^{-4}$

for around half of the nodes j [Fig. 5(c)]. The localization of the transition path ensemble is also evident from the reactive visitation probabilities r_j^+ for nodes [Fig. 5(d)]: less than 10 % of nodes are associated with values $r_j^+ > 0.1$, and only around half of the nodes have $r_j^+ > 10^{-5}$. The reactive visitation probabilities for the 10 % of nodes with the highest stationary probabilities are essentially negligible ($r_j^+ < 10^{-10}$). With decreasing temperature, the number of nodes associated with nonnegligible values for the reactive visitation probability becomes even smaller [118]. Moreover, the expected number of times that nodes are visited along $F \leftarrow I_h$ transition paths represents only a small fraction of the expected number of times that nodes are visited along first-passage paths. This result confirms that the majority of the MFPT is accounted for by unproductive “flickering” [108] within nodes that have a strong tendency to relax back to the I_h attractor node.

It is often insightful to closely examine the properties of specific nodes that play a critical role in the reactive dynamics. In Fig. 5, we highlight two metastable intermediate structures, M_1 and M_2 , that are particularly relevant to the $F \leftarrow I_h$ transition. The M_1 state is a somewhat disordered structure that is highly likely to be visited along a reactive $F \leftarrow I_h$ transition path ($r_{M_1}^+ \approx 0.9$), although trajectories at this node have a high probability of returning to the initial I_h state ($q_{M_1}^+ \approx 10^{-3}$). The M_1 state therefore represents a structure that (usually) must be located to successfully transition to the F from the I_h state, but this is an early step that does not modulate the slow dynamics. Large perturbations to the transition probabilities associated with the M_1 node would significantly affect the MFPT, but the global dynamics are not overly sensitive to small perturbations of this node, since the state does not constitute a limiting step in the rare event [124,125]. The M_2 state, a configuration that retains some of the symmetry of the incomplete icosahedral I_h state, is a true dynamical bottleneck node in the network. Around half of the reactive trajectories proceed to F via M_2 ($r_{M_2}^+ \approx 0.47$). Furthermore, the M_2 node is a member of the TSE [122] of nodes dividing the effective regions of attraction characterized by the I_h and F states. That is, trajectories reaching the M_2 state then have an approximately equal probability of first hitting either I_h or F , with the latter state slightly favoured ($q_{M_2}^+ \approx 0.63$). Since the M_2 node is likely to be visited along reactive paths and corresponds to a limiting step of the overall slow transition, the global dynamics, including the MFPT, are highly sensitive even to small perturbations of the transition probabilities corresponding to this node [123].

The principles that we have used in our analysis of the LJ₃₈ system can be applied to yield insight into the dynamics of an arbitrary discrete- or continuous-time Markov chain. It is particularly useful to identify the nodes that comprise the TSE, for which the local dynamics have a critical effect in determining the global dynamics [118], and to identify the favoured nodes that mediate the dominant pathways for the productive transition between two endpoint states. For Markov chains where the transition probabilities or rates depend on an external parameter [134], such as the temperature in physical systems, perturbations may significantly alter the dynamical behavior. For instance, a switching effect may be observed in systems with alternative competing mechanisms

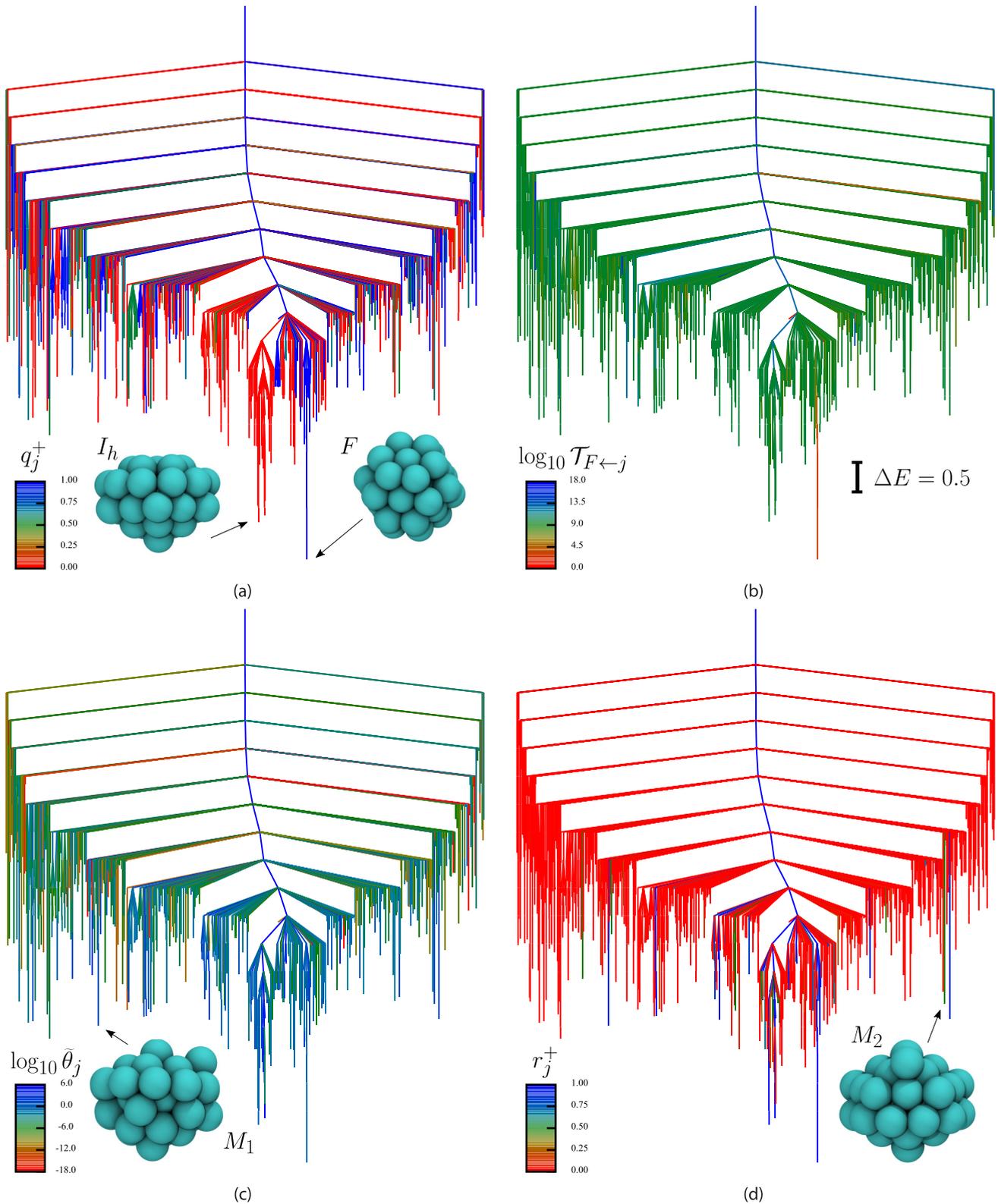


FIG. 5. Disconnectivity graphs [132] showing the dynamical properties of nodes in the Markov chain for the transition of the LJ_{38} cluster from an incomplete icosahedron (I_h) to a face-centered cubic (F) structure, computed using state-reduction algorithms as described in Sec. IV. The vertical axis represents the potential energy, and the threshold increment is $\Delta E = 0.5$ (in reduced units). (a) Committor probabilities q_j^+ for nodes j . (b) MFPTs $\mathcal{T}_{F \leftarrow j}$ for transitions to the F state. (c) Expected numbers of node visits along reactive paths that leave I_h and reach F without returning to I_h , $\tilde{\theta}_j$ [Eq. (41)]. (d) Reactive visitation probabilities r_j^+ [Eq. (44)]. Nodes M_1 and M_2 both have high visitation probabilities, but only the latter has a committor probability close to 0.5. The M_2 structure is therefore a dynamical bottleneck that has a critical role in modulating the overall transition.

for a given $\mathcal{A} \leftarrow \mathcal{B}$ transition, with different reactive pathways and dynamical bottlenecks being favoured in separate parameter regimes. When the system exhibits rare event dynamics, the origins of switching behavior can likely be traced to a small number of influential states. The quantities discussed in the present work, and especially the reactive visitation probability [118] derived herein, provide a convenient means to rigorously assess which regions of the state space are kinetically relevant with respect to a particular $\mathcal{A} \leftarrow \mathcal{B}$ process of interest. Our proposed methodology, which allows for the treatment of models with metastable states, is therefore essential to analyzing the features of a general Markov chain, and to understanding differences in the dynamical behavior of related models.

V. CONCLUSIONS

We have described state-reduction algorithms for the numerically stable analysis of first-passage processes in finite discrete- and continuous-time Markov chains exhibiting metastability, for which the systems of linear equations to be solved are severely ill-conditioned [41–44]. Since a separation of characteristic timescales is a ubiquitous feature of Markov chains representing realistic dynamical processes [38,46–59], our methodology provides a valuable approach to analyze complex systems in practical applications. The limiting factor affecting the viability of the state-reduction procedures presented here is the available computer memory. Nonetheless, the methodology remains feasible for sparse networks comprising several thousand nodes [108]. For larger networks, metastability can be exploited to lump [18,135–137] the nodes of the Markov chain without introducing significant error in the representation of the slow dynamics [38]. We have illustrated our approach with numerical results for a CTMC representing a structural transition in a model atomic cluster at a low temperature, which is not amenable to analysis by standard linear algebra methods [38].

We have presented an iterative formulation of the GT algorithm [73–78] (Sec. II B) that incorporates a backward pass phase, which enables the MFPTs for transitions from all nonabsorbing nodes to the absorbing state to be determined simultaneously (Sec. II D). The procedure requires storing a subset of elements of the transition probability matrix, and (optionally) waiting times for nodes, during the forward pass phase. If the MFPTs for transitions from all nonabsorbing nodes of the Markov chain are of interest, then our proposed variation of the GT algorithm (Algorithm 1) is preferable to previous formulations [76,107] that compute the MFPT for a single transition from a particular node. For example, the new version is particularly advantageous when computing the

optimal coarse-grained transition probabilities or rates for a given partitioning of the Markov chain [8,138,139], which requires the matrix of MFPTs for all pairwise transitions between nodes [116]. Efficient and robust computation of the stationary, committor [79–82], and absorption probabilities can also be incorporated into this GT algorithm (Sec. II C).

We then derived a state-reduction algorithm to compute the fundamental matrix of an absorbing Markov chain (Sec. III B), the elements of which are the expected number of node visits along first-passage paths. This procedure provides a numerically stable route to compute the variance of the FPT distribution [Eq. (25)], a key global dynamical property that is otherwise challenging to obtain in a robust manner [69]. Together with the committor probabilities, the expected number of node visits allows for the straightforward evaluation of key dynamical properties that characterize the direct transition process to the absorbing state at a nodewise level of detail (Sec. III C). In particular, we have derived expressions for the reactive visitation probabilities of nodes (Sec. III D), that is, the probability that a node is visited along a trajectory that hits the absorbing state without first re-entering the initial state (cf. Fig. 4). We considered reactive visitation probabilities for both the nonequilibrium [17] and equilibrium [112] (i.e., steady state) path ensembles [Eqs. (44) and (46), respectively]. The expected number of times that nodes are visited along reactive paths [Eqs. (41) and (47)] can be obtained similarly.

The methodology presented herein can be used to gain fundamental insight into dynamical processes on finite Markov chains, including for models with metastable states. The separation of the first-passage path ensemble into nonreactive and reactive components [17] allows for the individual nodes and edges that are critical in facilitating the productive transition process to be readily identified. Patterns in quantities such as the reactive visitation probabilities of nodes may be evident for different classes of network structure [140]. The effects of important features of the network, such as an edge that is associated with a large net reactive flux, or a node that constitutes a dynamical bottleneck, can be investigated by applying perturbations [141]. Thus, our methodology provides a framework to assess the robustness of a Markovian network [142–144] and to probe the relationship between network topology and dynamics at both a local and global scale [145,146].

ACKNOWLEDGMENTS

D.J.S. gratefully acknowledges the Cambridge Commonwealth, European and International Trust for a PhD scholarship. D.J.W. gratefully acknowledges support from the EPSRC.

APPENDIX

All algorithms described herein are implemented in DISCOTRESS [86], a highly flexible C++ program for simulation and analysis of arbitrary discrete- and continuous-time Markov chains, including models featuring metastable states, which are numerically challenging. DISCOTRESS is freely available software under the GNU General Public License. For code, documentation, and tutorials, see Ref. [86].

Algorithm 1: State-reduction algorithm to simultaneously compute the MFPTs, committor probabilities, and absorption probabilities for all transient (i.e., nonabsorbing) nodes in a DTMC or CTMC. This algorithm is illustrated in Fig. 2. The steps of the Grassmann-Taksar-Heyman (GTH) algorithm [66,67] are also incorporated into this procedure, so that the stationary distribution is computed if the Markov chain is irreducible.

input: discrete-or continuous-time transition probability matrix \mathbf{T} with state space $\mathcal{S} \equiv \mathcal{Q} \cup \mathcal{A}$
 set of absorbing nodes \mathcal{A} and set of transient nodes $\mathcal{Q} \equiv \mathcal{A}^c$
 set of initial nodes $\mathcal{B} \subseteq \mathcal{A}$ (note that nodes of the set $\mathcal{I} \equiv (\mathcal{A} \cup \mathcal{B})^c$ are prioritized for elimination)
 $|\mathcal{Q}|$ -dimensional vector of mean waiting (or lag) times $\boldsymbol{\tau}$ for nodes $j \in \mathcal{Q}$

output: $|\mathcal{Q}|$ -dimensional vector of $\mathcal{A} \leftarrow j$ MFPTs $\mathcal{T}_{\mathcal{A}}$ for nodes $j \in \mathcal{Q}$
 $|\mathcal{Q}|$ -dimensional vector of $\mathcal{A} \leftarrow \mathcal{B}$ committor probabilities \mathbf{q}^+ for nodes $j \in \mathcal{Q}$
 $|\mathcal{A}| \times |\mathcal{Q}|$ -dimensional matrix \mathbf{B} of $i \in \mathcal{A} \leftarrow j \notin \mathcal{A}$ absorption probabilities B_{ij}
 $|\mathcal{S}|$ -dimensional stationary distribution vector $\boldsymbol{\pi}$ for all nodes (exists if the chain is irreducible)

initialize $\mathcal{T}_{\mathcal{A}}, \mathbf{q}^+, \mathbf{B}, \boldsymbol{\pi}, \mathbf{L}, \mathbf{U};$
 $\mathbf{T}^{(0)} \leftarrow \mathbf{T}, n \leftarrow 1;$
 /* forward pass phase to eliminate all transient nodes by renormalization */
while $n \leq |\mathcal{Q}|$ (i.e. $n \notin \mathcal{A}$) **do**
for $i \in \mathcal{S}, j \notin \mathcal{A}$ **do**
 $L_{nj} \leftarrow T_{nj}^{(n-1)} / (1 - T_{nn}^{(n-1)}), U_{in} \leftarrow T_{in}^{(n-1)} - \delta_{in};$ LU decomposition of transition matrix
 $T_{ij}^{(n)} \leftarrow T_{ij}^{(n-1)} + U_{in} L_{nj};$ eliminate node by graph transformation
 $\tau_j^{(n)} \leftarrow \tau_j^{(n-1)} + \tau_n^{(n-1)} L_{nj};$ renormalize waiting times
 $n \leftarrow n + 1;$
if all nodes of the set $\mathcal{I} \equiv (\mathcal{A} \cup \mathcal{B})^c$ have been eliminated with this iteration **then** $q_b^+ \leftarrow 0 \forall b \in \mathcal{B};$
 /* compute committor probabilities for all intermediate nodes */
for $j \leftarrow n, n + 1, \dots, |\mathcal{Q}|$ (i.e. $j \in \mathcal{I}$) **do**
 $q_j^+ \leftarrow \sum_{a \in \mathcal{A}} T_{aj}^{(n)};$
 $B_{ij} \leftarrow T_{ij}^{(n)} \forall i \in \mathcal{A}, j \notin \mathcal{A};$ compute absorption probabilities
 /* If the Markov chain is irreducible, eliminate and then restore all
 but one of the absorbing nodes, needed to compute the stationary distribution (GTH algorithm) */
while $|\mathcal{Q}| < n < |\mathcal{S}|$ (i.e. $n \in \mathcal{A} \setminus |\mathcal{S}|$) **do**
 $L_{nj} \leftarrow T_{nj}^{(n-1)} / (1 - T_{nn}^{(n-1)}), U_{in} \leftarrow T_{in}^{(n-1)} - \delta_{in};$
 $T_{ij}^{(n)} \leftarrow T_{ij}^{(n-1)} + U_{in} L_{nj};$ eliminate node by graph transformation
 $n \leftarrow n + 1;$
 $\pi_n \leftarrow 1, \mu \leftarrow 1;$ at this point, only the $|\mathcal{S}|$ th node remains
while $|\mathcal{Q}| < n < |\mathcal{S}|$ (i.e. $n \in \mathcal{A} \setminus |\mathcal{S}|$) **do**
 $T_{ij}^{(n-1)} \leftarrow T_{ij}^{(n)} - U_{in} L_{nj};$ restore node (i.e., undo graph transformation)
 $n \leftarrow n - 1;$
 $\pi_n \leftarrow L_{n,|\mathcal{S}|} + \sum_{k=n+1}^{|\mathcal{S}|-1} \pi_k L_{nk}, \mu \leftarrow \mu + \pi_n;$ GTH step
 /* compute MFPT and stationary probability for the $|\mathcal{Q}|$ th node, which was the last transient node to be eliminated */
 $\mathcal{T}_{\mathcal{A}n} \leftarrow \tau_n^{(n)};$
 /* backward pass phase to compute MFPTs and stationary probabilities for all other transient nodes */
while $n \geq 1$ (i.e. $n \notin \mathcal{A}$) **do**
 $T_{ij}^{(n-1)} \leftarrow T_{ij}^{(n)} - U_{in} L_{nj};$ restore node (i.e., undo graph transformation)
 $\tau_j^{(n-1)} \leftarrow \tau_j^{(n)} - \tau_n^{(n-1)} L_{nj} \quad n \leftarrow n - 1;$
 $\mathcal{T}_{\mathcal{A}n} \leftarrow \tau_n^{(n)} + \sum_{\gamma \notin \mathcal{A}} \mathcal{T}_{\mathcal{A}\gamma} T_{\gamma n}^{(n)};$ compute MFPT for restored node
 $\pi_n \leftarrow L_{n,|\mathcal{S}|} + \sum_{k=n+1}^{|\mathcal{S}|-1} \pi_k L_{nk}, \mu \leftarrow \mu + \pi_n;$ GTH step
 $\pi_j \leftarrow \pi_j / \mu \forall j;$ renormalization of the stationary distribution (GTH)
deallocate \mathbf{L}, \mathbf{U}
return $\mathcal{T}_{\mathcal{A}}, \mathbf{q}^+, \mathbf{B}, \boldsymbol{\pi}$

Algorithm 2: State-reduction algorithm to robustly compute elements N_{ij} , the expected number of times that the i th node is visited along a first-passage path initialized at node j prior to absorption, for all transient nodes $i, j \in \mathcal{Q}$. This procedure is illustrated in Fig. 3.

input: discrete-or continuous-time transition probability matrix \mathbf{T} with state space $\mathcal{S} \equiv \mathcal{Q} \cup \mathcal{A}$
 set of transient nodes $\mathcal{Q} \subset \mathcal{S}$
 set of absorbing nodes $\mathcal{A} \subset \mathcal{S}$
 set of dummy nodes \mathcal{Q}^* , where $|\mathcal{Q}^*| = |\mathcal{Q}|$

output: $|\mathcal{Q}| \times |\mathcal{Q}|$ -dimensional fundamental matrix \mathbf{N} associated with the absorbing Markov chain

/ define a network with augmented state space \mathcal{S}^* that includes transient, absorbing, and dummy nodes */*
 $\mathcal{S}^* \leftarrow \mathcal{S} \cup \mathcal{Q}^*$;
/ set the initial $i \leftarrow j$ edge weights, N_{ij}^* , of the augmented network with state space \mathcal{S}^* */*
 $N_{ij}^* \leftarrow T_{ij} \quad \forall i \in \mathcal{S}, j \in \mathcal{Q}$;
 $N_{ij}^* \leftarrow 1 \quad \forall i \in \mathcal{Q}^*, j \in \mathcal{Q}$;
 $N_{ij}^* \leftarrow 1 \quad \forall i \in \mathcal{Q}, j \in \mathcal{Q}^*$;
 $N_{ij}^* \leftarrow 0 \quad \forall i \in \mathcal{S}, j \in \mathcal{A}$;
 $N_{ij}^* \leftarrow 0 \quad \forall i, j \in \mathcal{Q}^*$;

/ eliminate all transient nodes of the augmented network by renormalization */*
 $\mathcal{E} \leftarrow \emptyset$; set of nodes that have been eliminated (initially empty)

for $n \in \mathcal{Q}$ **do**
 $\mathcal{E} \leftarrow \mathcal{E} \cup \{n\}$;
 $N_n^* \leftarrow \sum_{\gamma \in \mathcal{S} \setminus \mathcal{E}} N_{\gamma n}^*$ ($\equiv 1 - N_{nn}^*$); confers numerical stability
for $i, j \in \mathcal{S}^* \setminus \mathcal{E}$ **do**
 $N_{ij}^* \leftarrow N_{ij}^* + (N_{in}^* N_{nj}^* / N_n^*)$; renormalization preserves $\sum_{\gamma \in \mathcal{S} \setminus \mathcal{E}} N_{\gamma n}^* = 1 \quad \forall n \in \mathcal{Q} \setminus \mathcal{E}$

/ once all transient nodes have been eliminated, the edge weights for transitions between dummy nodes in the remaining network are the elements of the fundamental matrix \mathbf{N} */*

initialize \mathbf{N} ;
 $N_{ij} \leftarrow N_{ij}^* \quad \forall i, j \in \mathcal{Q}^*$;
return \mathbf{N}

<p>[1] S. Redner, <i>A Guide to First-Passage Processes</i> (Cambridge University Press, Cambridge, UK, 2012).</p> <p>[2] R. Metzler, G. Oshanin, and S. Redner, <i>First-Passage Phenomena and Their Applications</i> (World Scientific, Singapore, 2014).</p> <p>[3] S. Iyer-Biswas and A. Zilman, <i>Adv. Chem. Phys.</i> 160, 261 (2016).</p> <p>[4] M. Castro, M. López-García, G. Lythe, and C. Molina-París, <i>Sci. Rep.</i> 8, 15054 (2018).</p> <p>[5] Z. Zhang, A. Julaiti, B. Hou, H. Zhang, and G. Chen, <i>Eur. Phys. J. B</i> 84, 691 (2011).</p> <p>[6] Z. Zhang, Y. Sheng, Z. Hu, and G. Chen, <i>Chaos</i> 22, 043129 (2012).</p> <p>[7] Z. Zhang, T. Shan, and G. Chen, <i>Phys. Rev. E</i> 87, 012112 (2013).</p> <p>[8] A. Kells, V. Koskin, E. Rosta, and A. Annibale, <i>J. Chem. Phys.</i> 152, 104108 (2020).</p> <p>[9] S. Park, M. K. Sener, D. Lu, and K. Schulten, <i>J. Chem. Phys.</i> 119, 1313 (2003).</p> <p>[10] S. X. Sun, <i>Phys. Rev. Lett.</i> 96, 210602 (2006).</p>	<p>[11] B. Harland and S. X. Sun, <i>J. Chem. Phys.</i> 127, 104103 (2007).</p> <p>[12] T. Oettel, V. V. Bulatov, A. Donev, M. H. Kalos, G. H. Gilmer, and B. Sadigh, <i>Phys. Rev. E</i> 80, 066701 (2009).</p> <p>[13] S. Hwang, D.-S. Lee, and B. Kahng, <i>Phys. Rev. Lett.</i> 109, 088701 (2012).</p> <p>[14] M. Manhart and A. V. Morozov, <i>Phys. Rev. Lett.</i> 111, 088102 (2013).</p> <p>[15] M. Manhart, W. Kion-Crosby, and A. V. Morozov, <i>J. Chem. Phys.</i> 143, 214106 (2015).</p> <p>[16] M. Manhart and A. V. Morozov, <i>Proc. Natl. Acad. Sci. USA</i> 112, 1797 (2015).</p> <p>[17] M. von Kleist, C. Schütte, and W. Zhang, <i>J. Stat. Phys.</i> 170, 809 (2018).</p> <p>[18] J. G. Kemeny and J. L. Snell, <i>Finite Markov Chains</i> (Van Nostrand, New Jersey, 1960).</p> <p>[19] N. G. van Kampen, <i>Stochastic Processes in Physics and Chemistry</i> (Elsevier, Amsterdam, Netherlands, 1992).</p> <p>[20] J. R. Norris, <i>Markov Chains</i> (Cambridge University Press, New York, 1997).</p>
---	---

- [21] C. M. Grinstead and J. L. Snell, *Introduction to Probability* (American Mathematical Society, Providence, RI, 1997).
- [22] H. M. Taylor and S. Karlin, *An Introduction to Stochastic Modeling*, 3rd ed. (Academic Press, London, UK, 1998).
- [23] L. J. S. Allen, *An Introduction to Stochastic Processes with Applications to Biology* (Prentice Hall, Upper Saddle River, NJ, 2003).
- [24] J. Goutsias and G. Jenkinson, *Phys. Rep.* **529**, 199 (2013).
- [25] N. Masuda, M. A. Porter, and R. Lambiotte, *Phys. Rep.* **716-717**, 1 (2017).
- [26] G. R. Bowman, V. S. Pande, and F. Noé, *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*, 1st ed. (Springer, Netherlands, 2014).
- [27] F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, and T. R. Weikl, *Proc. Natl. Acad. Sci. USA* **106**, 19011 (2009).
- [28] J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schütte, and F. Noé, *J. Chem. Phys.* **134**, 174105 (2011).
- [29] V. S. Pande, K. Beauchamp, and G. R. Bowman, *Methods* **52**, 99 (2010).
- [30] B. E. Husic and V. S. Pande, *J. Am. Chem. Soc.* **140**, 2386 (2018).
- [31] A. Mardt, L. Pasquali, H. Wu, and F. Noé, *Nat. Commun.* **9**, 5 (2018).
- [32] J. A. Joseph, K. Röder, D. Chakraborty, R. G. Mantell, and D. J. Wales, *Chem. Commun.* **53**, 6974 (2017).
- [33] K. Röder, J. A. Joseph, B. E. Husic, and D. J. Wales, *Adv. Theory Simul.* **2**, 1800175 (2019).
- [34] E. M. Hanks, M. B. Hooten, and M. W. Allredge, *Ann. Appl. Stat.* **9**, 145 (2015).
- [35] H. A. Simon and A. Ando, *Econometrica* **29**, 111 (1961).
- [36] C. D. Meyer, Jr., *SIAM Rev.* **17**, 443 (1975).
- [37] D. P. Heyman and D. P. O’Leary, What is fundamental for Markov chains: First-passage times, fundamental matrices and group generalized inverses, in *Computations with Markov Chains*, edited by W. J. Stewart (Springer, New York, 1995), pp. 151–161.
- [38] T. D. Swinburne, D. Kannan, D. J. Sharpe, and D. J. Wales, *J. Chem. Phys.* **153**, 134115 (2020).
- [39] D. J. Hartfiel and C. D. Meyer Jr., *Linear Algebra Appl.* **272**, 193 (1998).
- [40] C. R. MacCluer, *SIAM Rev.* **42**, 487 (2000).
- [41] D. P. Heyman and A. Reeves, *ORSA J. Comp.* **1**, 52 (1989).
- [42] B. Philippe, Y. Saad, and W. J. Stewart, *Oper. Res.* **40**, 1156 (1992).
- [43] C. D. Meyer Jr., *SIAM J. Matrix Anal. Appl.* **15**, 715 (1994).
- [44] J. L. Barlow, *SIAM J. Matrix Anal. Appl.* **22**, 230 (2000).
- [45] Y. Saad, *Numerical methods for large eigenvalue problems* (SIAM, Philadelphia, PA, 2011).
- [46] D. J. Aldous and M. Brown, Inequalities for rare events in time-reversible Markov chains I, in *IMS Lecture Notes in Statistics, Vol. 22: Stochastic Inequalities*, edited by M. Shaked and Y. L. Tong (Institute of Mathematical Statistics, Ohio, 1992), pp. 1–16.
- [47] P. Heidelberger, *ACM Trans. Model. Comput. Simul.* **5**, 43 (1995).
- [48] P. Glasserman, P. Heidelberger, P. Shahabuddin, and T. Zajic, *Oper. Res.* **47**, 495 (1999).
- [49] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein, *J. Phys. A: Math. Gen.* **33**, L447 (2000).
- [50] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein, *Commun. Math. Phys.* **228**, 219 (2002).
- [51] S. Juneja and P. Shahabuddin, *Manage. Sci.* **47**, 547 (2001).
- [52] J. Beltrán and C. Landim, *J. Stat. Phys.* **140**, 1065 (2010).
- [53] E. Vanden-Eijnden and J. Weare, *Commun. Pure Appl. Math.* **65**, 1770 (2012).
- [54] O. Benois and M. Mourragui, *J. Stat. Phys.* **153**, 967 (2013).
- [55] C. Hartmann, R. Banisch, M. Sarich, T. Badowski, and C. Schütte, *Entropy* **16**, 350 (2014).
- [56] M. Sarich, R. Banishc, C. Hartmann, and C. Schütte, *Entropy* **16**, 258 (2014).
- [57] M. K. Cameron, *J. Chem. Phys.* **141**, 184113 (2014).
- [58] T. Gan and M. Cameron, *J. Nonlinear Sci.* **27**, 927 (2017).
- [59] C. Pérez-Espigares and P. I. Hurtado, *Chaos* **29**, 083106 (2019).
- [60] D. P. Heyman, *SIAM J. Alg. Discr. Meth.* **8**, 226 (1987).
- [61] W. Grassmann and D. A. Stanford, Matrix analytic methods, in *Computational Probability*, edited by W. Grassmann (Springer, New York, 2000), pp. 153–203.
- [62] C. A. O’Cinneide, *Numer. Math.* **65**, 109 (1993).
- [63] C. A. O’Cinneide, *Numer. Math.* **73**, 507 (1996).
- [64] D. P. O’Leary and Y.-J. J. Wu, *SIAM J. Matrix Anal. Appl.* **17**, 470 (1996).
- [65] I. Sonin, *Adv. Math.* **145**, 159 (1999).
- [66] W. K. Grassmann, M. I. Taksar, and D. P. Heyman, *Oper. Res.* **33**, 1107 (1985).
- [67] T. J. Sheskin, *Oper. Res.* **33**, 228 (1985).
- [68] J. Kohlas, *Zeit. Oper. Res.* **30**, A197 (1986).
- [69] T. Dayar and N. Akar, *SIAM J. Matrix Anal. Appl.* **27**, 396 (2005).
- [70] D. P. Heyman, *SIAM J. Matrix Anal. Appl.* **16**, 954 (1995).
- [71] D. P. Heyman and D. P. O’Leary, *SIAM J. Matrix Anal. Appl.* **19**, 534 (1998).
- [72] I. Sonin and J. Thornton, *SIAM J. Matrix Anal. Appl.* **23**, 209 (2001).
- [73] S. A. Trygubenko and D. J. Wales, *Mol. Phys.* **104**, 1497 (2006).
- [74] S. A. Trygubenko and D. J. Wales, *J. Chem. Phys.* **124**, 234110 (2006).
- [75] D. J. Wales, *Int. Rev. Phys. Chem.* **25**, 237 (2006).
- [76] D. J. Wales, *J. Chem. Phys.* **130**, 204111 (2009).
- [77] J. D. Stevenson and D. J. Wales, *J. Chem. Phys.* **141**, 041104 (2014).
- [78] R. S. MacKay and J. D. Robinson, *Philos. Trans. R. Soc., A* **376**, 20170232 (2018).
- [79] C. Dellago, P. G. Bolhuis, and P. L. Geissler, *Adv. Chem. Phys.* **123**, 1 (2002).
- [80] W. E, W. Ren, and E. Vanden-Eijnden, *Chem. Phys. Lett.* **413**, 242 (2005).
- [81] A. M. Berezhkovskii and A. Szabo, *J. Chem. Phys.* **150**, 054106 (2019).
- [82] Q. Li, B. Lin, and W. Ren, *J. Chem. Phys.* **151**, 054112 (2019).
- [83] K. L. Chung, *J. Res. Natl. Bur. Stand.* **50**, 302 (1953).
- [84] K. L. Chung, *Illinois J. Math.* **5**, 431 (1961).
- [85] G. L. Sriwastav and S. N. N. Pandit, *Nav. Res. Logist. Q.* **25**, 653 (1978).
- [86] D. J. Sharpe, DISCOTRESS, <https://github.com/danieljsharpe/DISCOTRESS>.
- [87] P. Coolen-Schrijner and E. A. van Doorn, *Probab. Eng. Inf. Sci.* **16**, 351 (2002).

- [88] S. A. Serebrinsky, *Phys. Rev. E* **83**, 037701 (2011).
- [89] J. Medhi, Stochastic processes, in *Stochastic Models in Queueing Theory*, edited by J. Medhi (Academic Press, San Diego, 2003), pp. 1–46.
- [90] D. P. Heyman, *J. Appl. Probab.* **32**, 893 (1995).
- [91] T. Dayar and W. J. Stewart, *SIAM J. Sci. Comput.* **17**, 287 (1996).
- [92] W. J. Stewart, Numerical methods for computing stationary distributions of finite irreducible Markov chains, in *Computational Probability*, edited by W. Grassmann (Springer, New York, 2000), pp. 81–111.
- [93] J.-H. Prinz, M. Held, J. C. Smith, and F. Noé, *Multiscale Model. Simul.* **9**, 545 (2011).
- [94] J. J. Hunter, *Linear Algebra Appl.* **549**, 100 (2018).
- [95] Z. Zheng, G. Xiao, G. Wang, G. Zhang, and K. Jiang, *Math. Probl. Eng.* **2017**, 8217361 (2017).
- [96] N. Singhal, C. D. Snow, and V. S. Pande, *J. Chem. Phys.* **121**, 415 (2004).
- [97] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed. (SIAM, Philadelphia, PA, 2003).
- [98] Y. Saad, *Linear Algebra Appl.* **34**, 269 (1980).
- [99] H. D. Simon, *Linear Algebra Appl.* **61**, 101 (1984).
- [100] C. D. Meyer Jr., *SIAM Rev.* **31**, 240 (1989).
- [101] E. Seneta, *SIAM J. Matrix Anal. Appl.* **19**, 556 (1998).
- [102] E. Meerbach, C. Schütte, and A. Fischer, *Linear Algebra Appl.* **398**, 141 (2005).
- [103] M. Haviv, *SIAM J. Numer. Anal.* **22**, 952 (1987).
- [104] Y. Q. Zhao and D. Liu, *J. Appl. Probab.* **33**, 623 (1996).
- [105] T. Dayar, H. Hermanns, D. Spieler, and V. Wolf, *Numer. Linear Algebra Appl.* **18**, 931 (2011).
- [106] A. Miliias-Argeitis and J. Lygeros, *J. Chem. Phys.* **138**, 184109 (2013).
- [107] T. D. Swinburne and D. J. Wales, *J. Chem. Theory Comput.* **16**, 2661 (2020).
- [108] M. Athènes and V. V. Bulatov, *Phys. Rev. Lett.* **113**, 230601 (2014).
- [109] E. Vanden-Eijnden, Transition path theory, in *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1*, edited by M. Ferrario, G. Ciccotti, and K. Binder (Springer, Berlin, 2006), pp. 453–493.
- [110] P. Metzner, C. Schütte, and E. Vanden-Eijnden, *J. Chem. Phys.* **125**, 084110 (2006).
- [111] W. E and E. Vanden-Eijnden, *J. Stat. Phys.* **123**, 503 (2006).
- [112] P. Metzner, C. Schütte, and E. Vanden-Eijnden, *Multiscale Model. Simul.* **7**, 1192 (2009).
- [113] W. E and E. Vanden-Eijnden, *Annu. Rev. Phys. Chem.* **61**, 391 (2010).
- [114] J. J. Hunter, *Spec. Matrices* **4**, 151 (2016).
- [115] J. G. Kemeny and J. L. Snell, *Theory Probab. Appl.* **6**, 101 (1961).
- [116] D. Kannan, D. J. Sharpe, T. D. Swinburne, and D. J. Wales, *J. Chem. Phys.* **153**, 244108 (2020).
- [117] M. K. Cameron and E. Vanden-Eijnden, *J. Stat. Phys.* **156**, 427 (2014).
- [118] D. J. Sharpe and D. J. Wales, *J. Chem. Phys.* **153**, 024121 (2020).
- [119] G. Hummer, *J. Chem. Phys.* **120**, 516 (2004).
- [120] R. B. Best and G. Hummer, *Proc. Natl. Acad. Sci. USA* **102**, 6732 (2005).
- [121] P. G. Bolhuis, *J. Stat. Phys.* **145**, 841 (2011).
- [122] A. Berezhkovski, G. Hummer, and A. Szabo, *J. Chem. Phys.* **130**, 205102 (2009).
- [123] J. J. Hunter, *Linear Algebra Appl.* **410**, 217 (2005).
- [124] R. E. Funderlic and C. D. Meyer, *Linear Algebra Appl.* **76**, 1 (1986).
- [125] G. D. Zhang, *SIAM J. Matrix Anal. Appl.* **14**, 1112 (1993).
- [126] D. J. Sharpe and D. J. Wales, *J. Chem. Phys.* **151**, 124101 (2019).
- [127] D. Shukla, C. X. Hernández, J. K. Weber, and V. S. Pande, *Acc. Chem. Res.* **48**, 414 (2015).
- [128] E. Suárez, J. L. Adelman, and D. M. Zuckerman, *J. Chem. Theory Comput.* **12**, 3473 (2016).
- [129] J. P. K. Doye, M. A. Miller, and D. J. Wales, *J. Chem. Phys.* **110**, 6896 (1999).
- [130] J. P. K. Doye, M. A. Miller, and D. J. Wales, *J. Chem. Phys.* **111**, 8417 (1999).
- [131] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, *J. Comput. Phys.* **17**, 10 (1975).
- [132] O. M. Becker and M. Karplus, *J. Chem. Phys.* **106**, 1495 (1997).
- [133] M. Athènes, S. Kaur, G. Adjanor, T. Vanacker, and T. Jourdan, *Phys. Rev. Materials* **3**, 103802 (2019).
- [134] V. Betz and S. Le Roux, *Stochastic Processes Appl.* **126**, 3499 (2016).
- [135] P. Buchholz, *J. Appl. Probab.* **31**, 59 (1994).
- [136] W. E, T. Li, and E. Vanden-Eijnden, *Proc. Natl. Acad. Sci. USA* **105**, 7907 (2008).
- [137] J. A. Ward and M. López-García, *Appl. Netw. Sci.* **4**, 108 (2019).
- [138] G. Hummer and A. Szabo, *J. Phys. Chem. B* **119**, 9029 (2015).
- [139] A. Kells, Z. É. Mihálka, A. Annibale, and E. Rosta, *J. Chem. Phys.* **150**, 134107 (2019).
- [140] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
- [141] B. Barzel and O. Biham, *Phys. Rev. E* **80**, 046104 (2009).
- [142] W.-X. Wang and G. Chen, *Phys. Rev. E* **77**, 026101 (2008).
- [143] B. Mirzasoleiman, M. Babaei, M. Jalili, and M. A. Safari, *Phys. Rev. E* **84**, 046114 (2011).
- [144] M. Bellingeri and D. Cassi, *Physica A* **489**, 47 (2018).
- [145] B. Barzel and A.-L. Barabási, *Nat. Phys.* **9**, 673 (2013).
- [146] U. Harush and B. Barzel, *Nat. Commun.* **8**, 2181 (2017).