MULTISCALE MODEL. SIMUL.
Vol. 7, No. 3, pp. 1192–1219

TRANSITION PATH THEORY FOR MARKOV JUMP PROCESSES*

PHILIPP METZNER†, CHRISTOF SCHÜTTE†, AND ERIC VANDEN-EIJNDEN‡

Abstract. The framework of transition path theory (TPT) is developed in the context of continuous-time Markov chains on discrete state-spaces. Under assumption of ergodicity, TPT singles out any two subsets in the state-space and analyzes the statistical properties of the associated reactive trajectories, i.e., those trajectories by which the random walker transits from one subset to another. TPT gives properties such as the probability distribution of the reactive trajectories, their probability current and flux, and their rate of occurrence and the dominant reaction pathways. In this paper the framework of TPT for Markov chains is developed in detail, and the relation of the theory to electric resistor network theory and data analysis tools such as Laplacian eigenmaps and diffusion maps is discussed as well. Various algorithms for the numerical calculation of the various objects in TPT are also introduced. Finally, the theory and the algorithms are illustrated in several examples.

Key words. transition path theory, Markov jump process, committor function, network, graph theory, reactive trajectories, probability current, transition rate

AMS subject classifications. 60J27, 65C40, 65C50

DOI. 10.1137/070699500

1. Introduction.

Continuous-time Markov chains on discrete state-spaces have an enormous range of applications. In recent years, especially, with the explosion of new applications in network science, Markov chains have become the tool of choice not only to model the dynamics on these networks but also to study their topological properties [2, 26]. In this context, there is a need for new methods to analyze Markov chains on large state-spaces with no specific symmetries, as is relevant for large complex networks. This paper proposes one such method.

A natural starting point to analyze a Markov chain is to use spectral analysis. This is especially relevant when the chain displays metastability, as was shown in [8, 12] in the context of time-reversible chains. By definition, the generator of a metastable chain possesses one or more clusters of eigenvalues near zero, and the associated eigenvectors provide a natural way to partition the chain (and hence the underlying network) into cluster of nodes on which the random walker remains for a very long time before finding its way to another such cluster. This approach has been used not only in the context of Markov chains arising from statistical physics (such as glassy systems [4, 7] or biomolecules [30]) but also in the context of data segmentation and embedding [32, 23, 29, 3, 14, 11, 21]. The problem with the spectral approach, however, is that not all Markov chains of interest are time-reversible and metastable, and, when they are not, the meaning of the first few eigenvectors of the generator is less clear.

In this paper, we take another approach which does not require metastability

*Received by the editors August 7, 2007; accepted for publication (in revised form) September 17, 2008; published electronically January 9, 2009.
http://www.siam.org/journals/mms/7-3/69950.html

†Department of Mathematics and Computer Science, Free University Berlin, Arnimallee 6, D-14195 Berlin, Germany (metzner@math.fu-berlin.de, schuette@math.fu-berlin.de). These authors' research was supported by the DFG Research Center MATHEON "Mathematics for Key Technologies" (FZT86) in Berlin.

‡Courant Institute of Mathematical Sciences, New York University, New York, NY 10012 (eve2@cims.nyu.edu). This author's research was partially supported by NSF grants DMS02-09959 and DMS02-39625, and by ONR grant N00014-04-1-0565.

and applies for non-time-reversible chains as well. The basic idea is to single out two subsets of nodes of interest in the state-space of the chain and ask what the typical mechanism is by which the walker transits from one of these subsets to the other. We can also ask what the rate is at which these transitions occur, etc. The first object which comes to mind to characterize these transitions is the path of maximum likelihood by which they occur. However, this path can again be not very informative if the two states one has singled out are not metastable states. The main objective of this paper, however, is to show that we can give a precise meaning to the question of finding typical mechanisms and rates of transition even in chains which are neither metastable nor time-reversible. In so doing, we shall exploit the framework of transition path theory (TPT) which has been developed in [16, 34, 25] in the context of diffusions. In a nutshell, given two subsets in state-space, TPT analyzes the statistical properties of the associated reactive trajectories, i.e., the trajectories by which transition occur between these sets. TPT provides information such as the probability distribution of these trajectories, their probability current and flux, and their rate of occurrence. In this paper, we shall adapt TPT to continuous-time Markov chains and illustrate the output of the theory via several examples. For the sake of brevity, we will focus only on continuous-time Markov chains, but we note that our results can be straightforwardly extended to the case of discrete-time Markov chains. We choose illustrative examples motivated by molecular dynamics and chemical physics, but the tools of TPT presented here can also be used for data segmentation and data embedding. In this context, TPT may also provide an alternative to Laplacian eigenmaps [29, 3] and diffusion maps [11], which have become very popular recently in data analysis.

The remainder of this paper is organized as follows. In section 2 we present the framework of TPT for Markov jump processes. In section 3 we discuss the algorithmic aspects related to the numerical calculation of the various objects in TPT. We focus especially on the reaction pathways whose calculation involve techniques from graph theory which are nonstandard in the context of Markov chains. In section 4 we illustrate the theory and the algorithms in several examples arising in molecular dynamics and chemical kinetics. Finally, in section 5 we make a few concluding remarks.

2. Theoretical aspects.

2.1. Preliminaries: Notation and assumptions.

We will consider a Markov jump process on the countable state-space S with infinitesimal generator (or rate matrix) $L = (l_{ij})_{i,j\in S}$:

$$\begin{cases} l_{ij} \geq 0 & \forall i, j \in S,\ i \neq j, \\ \sum_{j\in S} l_{ij} = 0 & \forall i \in S. \end{cases} \tag{2.1}$$

Recall that if the process is in state i at time t, then $l_{ij}\Delta t + o(\Delta t)$ for $j \neq i$ gives the probability that the process jumps from state i to state j during the infinitesimal time interval $[t, t + \Delta t]$, and this probability is independent of what happened to the process before time t. We assume that the Markov jump process is irreducible and ergodic with respect to the unique, strictly positive invariant distribution $\pi = (\pi_i)_{i\in S}$, the solution of

$$0 = \pi^T L. \tag{2.2}$$

We will denote by $\{X(t)\}_{t\in\mathbb{R}}$ an equilibrium sample path (or trajectory) of the Markov jump process, i.e., any path obtained from $\{X(t)\}_{t\in[T,\infty)}$ by pushing back the initial

condition, $X(T) = x$, to $T = -\infty$. Following standard conventions, we assume that $\{X(t)\}_{t\in\mathbb{R}}$ is right-continuous with left limits (*càdlàg*) (i.e., at the times of the jumps the process is assigned to the state it jumps into rather than to the one it jumped from).

We will be interested in studying certain statistical properties of the ensemble of equilibrium paths. In principle, this requires us to construct a suitable probability space whose sample space is the ensemble of these equilibrium paths. Such a construction is standard (see, e.g., [10]), and we will not dwell on it here since, by assumption of ergodicity, the statistical properties of the ensemble of equilibrium paths that we are interested in can also be extracted from almost any path in this ensemble via suitable time averaging. This is the viewpoint that we will adopt in this paper since it gives an operational way to compute expectations from a trajectory generated, e.g., by numerical simulations.

Below, we will also need the process obtained from $\{X(t)\}_{t\in\mathbb{R}}$ by time reversal. We will denote this time-reversed process by $\{\tilde{X}(t)\}_{t\in\mathbb{R}}$ and define it as

$$\tilde{X}(t) = X^*(-t), \quad \text{where} \quad X^*(t) = \lim_{s\to t-} X(s). \tag{2.3}$$

By our assumptions of irreducibility and ergodicity, the process $\{\tilde{X}(t)\}_{t\in\mathbb{R}}$ is again a càdlàg Markov jump process with the same invariant distribution as $\{X(t)\}_{t\in\mathbb{R}}$, π, and infinitesimal generator $\tilde{L} = (\tilde{l}_{ij})_{i,j\in S}$ given by

$$\tilde{l}_{ij} = \frac{\pi_j}{\pi_i} l_{ji}. \tag{2.4}$$

Finally, recall that if the infinitesimal generator satisfies the detailed balance equations

$$\forall i, j \in S: \quad \pi_i l_{ij} = \pi_j l_{ji}, \tag{2.5}$$

then $\tilde{L} \equiv L$ and, hence, the direct and the time-reversed process are statistically indistinguishable. Such a process is called *reversible*. We do not assume reversibility in this paper.

For the algorithmic part of this paper, it will be convenient to use the notation and concepts of graph theory. We will mainly consider directed graphs $G = G(S, E)$, where the vertex set S is the set of all states of the Markov jump process and two vertices i and j are connected by a *directed edge* if $(i, j) \in E \subseteq (S \times S)$.

We also recall the following definition.

DEFINITION 2.1. *A* directed pathway $w = (i_0, i_1, i_2, \ldots, i_n)$, $i_j \in S$, $j = 0, \ldots, n$, *in a graph G is a finite sequence of vertices such that $(i_j, i_{j+1}) \in E$, $j = 0, \ldots, n-1$. A directed pathway w is called* simple *if w does not contain any self-intersections (loops), i.e., $i_j \neq i_k$ for $j, k \in \{0, \ldots, n\}$, $j \neq k$.*

We will later consider several forms of induced directed graphs.

DEFINITION 2.2. *Let $E' \subset E$ be a subset of edges of a graph $G = G(S, E)$; then we denote by $G[E'] = G(S', E')$ the* induced subgraph*, i.e., the graph which consists of all edges in E' and the vertex set*

$$S' = \{i \in S : \exists j \in S \text{ such that } (i, j) \in E' \text{ or } (j, i) \in E'\}.$$

DEFINITION 2.3. *Whenever a $|S| \times |S|$-matrix $C = (C_{ij})$ with nonnegative entries is given, the* weight-induced directed graph *is denoted by $G\{C\} = G(S, E)$. In this graph the vertex set S is the set of all states of the Markov jump process, and two vertices i and j are connected by a* directed edge $(i, j) \in E \subseteq (S \times S)$ *if the corresponding* weight C_{ij} *is positive.*

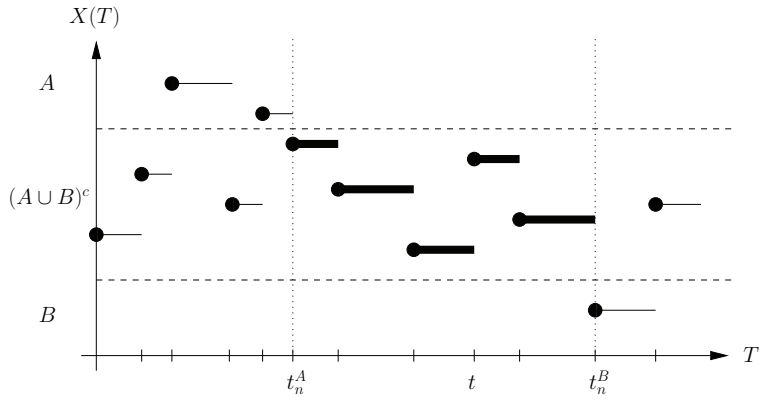

FIG. 1. *Schematic representation of a piece of an ergodic trajectory. The subpiece connecting A to B (shown in thick black) is a reactive trajectory, and the collection of reactive trajectories is the ensemble of reactive trajectories.*

2.2. Reactive trajectories. Let A and B be two nonempty, disjoint subsets of the state-space S. By ergodicity, any equilibrium path $\{X(t)\}_{t\in\mathbb{R}}$ oscillates infinitely many times between set A and set B. We are interested in understanding how these oscillations happen (mechanism, rate, etc.). If we view A as a reactant state and B as a product state, then each oscillation from A to B is a reaction event, and so we are asking about the mechanism, rate, etc., of these reaction events. To properly define and characterize the reaction events, we proceed by pruning out of each equilibrium trajectory $\{X(t)\}_{t\in\mathbb{R}}$ the pieces during which it makes a transition from A to B (i.e., the reactive pieces), and we ask about various statistical properties of these reactive pieces. The pruning is done as follows (see also Figure 1 for a schematic illustration).

First, given a trajectory $\{X(t)\}_{t\in\mathbb{R}}$ we define a set of last-exit-before-entrance and first-entrance-after-exit times $\sigma = \{t_n^A, t_n^B\}_{n\in\mathbb{Z}}$ as follows.

DEFINITION 2.4 (exit and entrance times). *Given a trajectory $\{X(t)\}_{t\in\mathbb{R}}$, the last-exit-before-entrance time t_n^A and the first-entrance-after-exit time t_n^B belong to σ if and only if*

$$\begin{aligned} &\lim_{t\to t_n^A-} X(t) = x_n^A \in A, \quad X(t_n^B) = x_n^B \in B \\ &\forall t \in [t_n^A, t_n^B) : X(t) \notin A \cup B. \end{aligned} \tag{2.6}$$

By ergodicity, we know that the cardinality of σ is almost surely (a.s.) infinite. It is also clear that the times t_n^A and t_n^B form an increasing sequence, $t_n^A \le t_n^B \le t_{n+1}^A$, for all $n \in \mathbb{Z}$. Notice, however, that we may have $t_n^A = t_n^B$ for some $n \in \mathbb{Z}$ corresponding to events when the trajectory jumps directly from A to B. If, on the other hand, $t_n^A < t_n^B$, then the trajectory visits states outside of A and B when it makes a transition from the former to the latter.

Next, given the set σ, we define the following.

DEFINITION 2.5 (reactive times). *The set R of reactive times is defined as*

$$R = \bigcup_{n\in\mathbb{Z}} (t_n^A, t_n^B) \subset \mathbb{R}. \tag{2.7}$$

Finally, we denote by $t_n^1 \equiv t_n^A \le t_n^2 \le \cdots \le t_n^{k_n} \le t_n^B$ the set of all of the successive jumping times of $X(t)$ in $[t_n^A, t_n^B]$, i.e., all of the times in $[t_n^A, t_n^B]$ such that

$$\lim_{t\to t_n^k-} X(t) \neq X(t_n^k) =: x_n^k, \qquad k = 1, \ldots, k_n \in \mathbb{N}, \tag{2.8}$$

and we define the following.

DEFINITION 2.6 (reactive trajectories). *The ordered sequence*

$$P_n = [x_n^A, x_n^1, x_n^2 \ldots, x_n^{k_n} \equiv x_n^B] \tag{2.9}$$

consisting of the successive states visited during the nth transition from A to B (including the last state in A, x_n^A, and the first one in B, $x_n^B \equiv x_n^{k_n}$) is called the nth reactive trajectory. The set of all such sequences,

$$P = \bigcup_{n\in\mathbb{Z}} \{P_n\}, \tag{2.10}$$

is called the set of reactive trajectories.

(Note that we have $k_n = 1$ when the trajectory hops directly from A to B at time $t_n^A = t_n^B$, in which case $P_n = [x_n^A, x_n^B]$.)

Since the equilibrium trajectory $\{X(t)\}_{t\in\mathbb{R}}$ used in the construction above is part of a statistical ensemble, the sets R, P_n, and P are also random sets whose statistical properties are induced by those of the ensemble of equilibrium trajectories. In the next sections we obtain explicit expression for various expectations involving these random sets. Using ergodicity, these expectations can be computed a.s. from a single trajectory via time averaging, even though in this case σ, R, P_n, and P are fixed sets. As already explained above, the second viewpoint is the one we will take in this paper since it gives operational definitions to all of the statistical quantities we are interested in.

2.3. Probability distribution of reactive trajectories. A first object relevant to quantify the statistical properties of the reactive trajectories is the following definition.

DEFINITION 2.7. *The distribution of reactive trajectories $m^R = (m_i^R)_{i\in S}$ is defined so that for any $i \in S$ we have*

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathbf{1}_{\{i\}}(X(t))\mathbf{1}_R(t)\mathrm{d}t = m_i^R, \tag{2.11}$$

where $\mathbf{1}_C(\cdot)$ denotes the characteristic function of the set C.

The distribution m^R gives the equilibrium probability that the system is in state i at time t and that it is reactive at that time; i.e., m_i^R can also be expressed as

$$m_i^R = \mathbb{P}(X(t) = i \;\&\; t \in R), \tag{2.12}$$

where $\mathbb{P}$ denotes probability with respect to the ensemble of equilibrium trajectories. To avoid confusion, note that the random objects in (2.12) are $X(t)$ and R: the time t in this expression is fixed, and m_i^R does not depend on t since we look at equilibrium reactive trajectories.

How can we find an expression for m^R? Suppose we encounter the process $X(t)$ in a state $i \in S$. What is the probability that $X(t)$ is reactive? Intuitively, this is the probability that the process came from A rather than from B times the probability that the process will reach B rather than A in the future. This indicates that the following objects will play an important role.

DEFINITION 2.8. *The discrete forward committor $q^+ = (q_i^+)_{i\in S}$ is defined as the probability that the process starting in $i \in S$ will first reach B rather than A. Analogously, we define the discrete backward committor $q^- = (q_i^-)_{i\in S}$ as the probability that the process arriving in state i last came from A rather than B.*

The forward and backward committors both satisfy a discrete Dirichlet problem:

$$\begin{cases} \sum_{j\in S} l_{ij} q_j^+ = 0 & \forall i \in (A\cup B)^c, \\ q_i^+ = 0 & \forall i \in A, \\ q_i^+ = 1 & \forall i \in B \end{cases} \tag{2.13}$$

and

$$\begin{cases} \sum_{j\in S} \tilde{l}_{ij} q_j^- = 0 & \forall i \in (A\cup B)^c, \\ q_i^- = 1 & \forall i \in A, \\ q_i^- = 0 & \forall i \in B. \end{cases} \tag{2.14}$$

Here $L = (l_{ij})_{i,j\in S}$ and $\tilde{L} = (\tilde{l}_{ij})_{i,j\in S}$ denote the infinitesimal generator forward and backward in time, respectively. For the reader's convenience, we recall the derivation of these equations in the appendix. The committor q_i^+ is related to hitting times with respect to the sets A and B by

$$q_i^+ = \mathbb{P}_i(\tau_B^+ < \tau_A^+). \tag{2.15}$$

Here $\mathbb{P}_i$ denotes probability conditional on $X(0) = i$, $\tau_A^+ = \min\{t > 0 : X(t) \in A\}$ denotes the *first entrance time* of the set A, and $\tau_B^+ = \min\{t > 0 : X(t) \in B\}$ denotes the *first entrance time* of the set B; q_i^- can be defined similarly using the time-reversed process as

$$q_i^- = \tilde{\mathbb{P}}_i(\tau_B^- > \tau_A^-), \tag{2.16}$$

where $\tilde{\mathbb{P}}_i$ denotes probability with respect to the time-reversed process conditional on $\tilde{X}(0) = i$, $\tau_A^- = \inf\{t > 0 : \tilde{X}(t) \in A\}$ denotes the *last exit time* of the subset A, and $\tau_B^- = \inf\{t > 0 : \tilde{X}(t) \in B\}$ denotes the *last exit time* of the subset B.

We have the following theorem.

THEOREM 2.9. *The probability distribution of reactive trajectories defined in* (2.11) *is given by*

$$m_i^R = \pi_i q_i^+ q_i^-, \qquad i \in S. \tag{2.17}$$

Proof. Denote by $x_i^{AB,+}(t)$ the first state in $A \cup B$ reached by $X(s)$, $s \geq t$, conditional on $X(t) = i$. Similarly, denote by $x_i^{AB,-}(t)$ the last state in $A\cup B$ left by $X(s)$, $s \leq t$, conditional on $X(t) = i$, or, equivalently, the first state in $A\cup B$ reached by $\tilde{X}(s)$, $s \geq -t$. In terms of these quantities, (2.11) can be written as

$$m_i^R = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathbf{1}_{\{i\}}(X(t)) \mathbf{1}_A(x_i^{AB,-}(t)) \mathbf{1}_B(x_i^{AB,+}(t)) \mathrm{d}t.$$

Taking the limit as $T \to \infty$ and using ergodicity together with the strong Markov property, we deduce that

$$m_i^R = \pi_i \, \mathbb{P}_i(\tau_B^+ < \tau_A^+) \tilde{\mathbb{P}}_i(\tau_B^- > \tau_A^-),$$

which is (2.17) by definition of q^+ and q^-. □

Notice that $m_i^R = 0$ if $i \in A \cup B$. Notice also that m^R is not a normalized distribution. In fact, from (2.12)

$$Z_{AB} = \sum_{j \in S} m_j^R = \sum_{j \in S} \pi_j q_j^+ q_j^- < 1 \tag{2.18}$$

is the probability that the trajectory is reactive at some given instance t in time, i.e.,

$$Z_{AB} = \mathbb{P}(t \in R). \tag{2.19}$$

The distribution

$$m_i^{AB} = Z_{AB}^{-1} m_i^R = Z_{AB}^{-1} \pi_i q_i^+ q_i^- \tag{2.20}$$

is then the normalized distribution of reactive trajectories which gives the probability of observing the system in a reactive trajectory and in state i at time t conditional on the trajectory being reactive at time t.

REMARK 2.10. *If the Markov process is reversible (i.e., $\pi_i l_{ij} = \pi_j l_{ji}$), then $q_i^+ = 1 - q_i^-$ and the probability distribution of reactive trajectories reduces to*

$$m_i^R = \pi_i q_i^+ (1 - q_i^+) \qquad \textit{(reversible process).} \tag{2.21}$$

2.4. Probability current of reactive trajectories. In this section we are interested in the probability current of reactive trajectories, i.e., the average rate at which they flow from state i to state j. A precise definition amounts to counting how many reactive trajectories jump from state i to state j on average in a time interval of length $s > 0$ and then computing the limit as $s \to 0+$ of the ratio between this average number and s. In formula, this reads as follows.

DEFINITION 2.11. *The probability current of reactive trajectories $f^{AB} = (f_{ij}^{AB})_{i,j \in S}$ is defined so that for all pairs of states (i, j), $i, j \in S$, $i \neq j$, we have*

$$\begin{aligned} \lim_{s \to 0+} \frac{1}{s} \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} & \mathbf{1}_{\{i\}}(X(t)) \mathbf{1}_{\{j\}}(X(t+s)) \\ & \times \sum_{n \in \mathbb{Z}} \mathbf{1}_{(-\infty, t_n^B]}(t) \mathbf{1}_{[t_n^A, \infty)}(t+s) \mathrm{d}t = f_{ij}^{AB}. \end{aligned} \tag{2.22}$$

In addition, we set $f_{ii}^{AB} = 0$ for all $i \in S$.

In (2.22), the factor $\sum_{n \in \mathbb{Z}} \mathbf{1}_{(-\infty, t_n^B]}(t) \mathbf{1}_{[t_n^A, \infty)}(t+s)$ is used to prune out of the time average all of the times during which $X(t)$ and $X(t+s)$ are both not reactive. It has this complicated looking form because we want the flux f_{ij}^{AB} to be nonzero even if $i \in A$: for any $i \notin A$ the pruning factor in (2.22) can be replaced by $\mathbf{1}_R(t) \mathbf{1}_R(t+s)$, but this is not adequate if $i \in A$ because $X(t_A^n) \notin A$ by construction. For $i \notin A$, f_{ij}^{AB} can be also be defined as

$$f_{ij}^{AB} = \lim_{s \to 0+} \frac{1}{s} \mathbb{P}(X(t) = i \ \&\ X(t+s) = j \ \&\ t \in R \ \&\ t + s \in R). \tag{2.23}$$

We have the following theorem.

THEOREM 2.12. *The discrete probability current of reactive trajectories is given by*

$$f_{ij}^{AB} = \begin{cases} \pi_i q_i^- l_{ij} q_j^+ & \textit{if } i \neq j, \\ 0 & \textit{otherwise.} \end{cases} \tag{2.24}$$

Proof. Using the same notation as in the proof of Theorem 2.9, (2.22) can also be written as

$$f_{ij}^{AB} = \lim_{s\to 0^+} \frac{1}{s} \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathbf{1}_{\{i\}}(X(t))\mathbf{1}_{\{j\}}(X(t+s)) \times \mathbf{1}_A(x_i^{AB,-}(t))\mathbf{1}_B(x_j^{AB,+}(t+s))\mathrm{d}t. \tag{2.25}$$

Taking the limit $T \to \infty$ and using ergodicity, we deduce that

$$f_{ij}^{AB} = \lim_{s\to 0^+} \frac{1}{s} \pi_i q_i^- \mathbb{E}_i[q^+_{X(s)}, \mathbf{1}_{\{j\}}(X(s))],$$

where $\mathbb{E}_i$ denotes the expectation conditional on $X(0) = i$. To take the limit $s \to 0^+$ we use

$$\forall \Phi : S \mapsto \mathbb{R} : \lim_{s\to 0^+} \frac{1}{s}(\mathbb{E}_i[\Phi(X(s))] - \Phi(i)) = \sum_{j\in S} l_{ij}\Phi(j),$$

and we are done since $i \neq j$. ∎

This result implies an expected property, namely the conservation of the discrete probability current or flux in each node.

THEOREM 2.13. *For all $i \in (A \cup B)^c$ the probability current is conserved, i.e.,*

$$\sum_{j\in S}(f_{ij}^{AB} - f_{ji}^{AB}) = 0 \quad \forall i \in (A \cup B)^c. \tag{2.26}$$

Proof. By the definition of f^{AB} for $i \in (A \cup B)^c$,

$$\begin{aligned}\sum_{j\in S}(f_{ij}^{AB} - f_{ji}^{AB}) &= \pi_i q_i^- \sum_{j\neq i} l_{ij} q_j^+ - \pi_i q_i^+ \sum_{j\neq i} \frac{\pi_j}{\pi_i} l_{ji} q_j^- \\ &= -q_i^- q_i^+ \pi_i l_{ii} + q_i^- q_i^+ \pi_i \tilde{l}_{ii} = 0,\end{aligned}$$

where we used $\sum_{j\in S} l_{ij} q_j^+ = 0$ if $i \in (A \cup B)^c$ from (2.13) and $\sum_{j\in S} \tilde{l}_{ij} q_j^- = 0$ if $i \in (A \cup B)^c$ from (2.14). ∎

For later use we should also mention that conservation of the current in every state $i \in (A \cup B)^c$ immediately implies the following total conservation of the current:

$$\sum_{i\in A,\, j\in S} f_{ij}^{AB} = \sum_{j\in S,\, i\in B} f_{ji}^{AB}, \tag{2.27}$$

where we used that $f_{ij}^{AB} = 0$ if $i \in S$ and $j \in A$, and $f_{ij}^{AB} = 0$ if $i \in B$ and $j \in S$.

2.5. Transition rate and effective current. In this section we derive the average number of transitions from A to B per time unit or, equivalently, the average number of reactive trajectories observed per time unit. More precisely, let $N_T^-, N_T^+ \in \mathbb{Z}$ be such that

$$R \cap [-T, T] = \bigcup_{N_T^- \leq n \leq N_T^+} (t_n^A, t_n^B); \tag{2.28}$$

that is, $N_T^+ - N_T^-$ is the number of reactive trajectories in the interval $[-T, T]$ in time.

Then we have the following definition.

DEFINITION 2.14. *The transition rate k_{AB} is defined as*

$$k_{AB} = \lim_{T\to\infty} \frac{N_T^+ - N_T^-}{2T}. \tag{2.29}$$

We have the following theorem.

THEOREM 2.15. *The transition rate is given by*

$$k_{AB} = \sum_{i\in A,\, j\in S} f_{ij}^{AB} = \sum_{j\in S,\, k\in B} f_{jk}^{AB}. \tag{2.30}$$

Proof. From (2.25) we get

$$\sum_{i\in A,\, j\in S} f_{ij}^{AB} = \lim_{s\to 0^+} \frac{1}{s} \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathbf{1}_A(X(t)) \sum_{j\in S} \mathbf{1}_B(x_j^{AB,+}(t+s))\mathrm{d}t. \tag{2.31}$$

Let us consider the integral; we can always restrict our attention to generic values of T such that there is no $n \in \mathbb{Z}$ for which $T = t_n^A$ or $T = t_n^B$. The integrand in this expression is nonzero *if and only if* $X(t) \in A$, $X(t+s) \in A^c$ and $t+s \in R$, i.e., if $t_n^A \in (t, t+s)$ for some $n \in \mathbb{Z}$. But this means that the integral of $\mathbf{1}_A(X(t))\mathbf{1}_B(x_j^{AB,+}(t+s))$ on every interval $t \in (t_n^A - s, t_n^A)$ is equal to s and the only contributions to the integral in (2.31) come from the intervals in $[-T, T] \cap \cup_{n\in\mathbb{Z}} (t_n^A - s, t_n^A)$. But these are exactly $N_T^+ - N_T^-$ intervals such that the whole integral amounts to $(N_T^+ - N_T^-)s$. From (2.31) and (2.28), this implies the first identity for the rate k_{AB}. The second identity follows from (2.27). ∎

Notice that the rate can also be expressed as

$$k_{AB} = \sum_{i\in A,\, j\in S} f_{ij}^+, \tag{2.32}$$

where we have the following definition.

DEFINITION 2.16. *The effective current is defined as*

$$f_{ij}^+ = \max(f_{ij}^{AB} - f_{ji}^{AB}, 0). \tag{2.33}$$

Identity (2.32) follows from (2.30) and the fact that for all $i \in A$: $f_{ij}^+ = f_{ij}^{AB}$ since $f_{ji}^{AB} = 0$ and $f_{ij}^{AB} > 0$ if $i \in A$. The effective current gives the net average number of reactive trajectories per time unit making a transition from i to j on their way from A to B. The effective current will be useful to define transition pathways in section 2.7.

REMARK 2.17. *If the Markov process is reversible, then the effective current reduces to*

$$f_{ij}^+ = \begin{cases} \pi_i l_{ij}(q_j^+ - q_i^+) & \text{if } q_j^+ > q_i^+, \\ 0 & \text{otherwise} \end{cases} \qquad \text{(reversible process)}, \tag{2.34}$$

and the reaction rate can be expressed as

$$k_{AB} = \tfrac{1}{2} \sum_{i,j\in S} \pi_i l_{ij}(q_j^+ - q_i^+)^2, \qquad \text{(reversible process)}. \tag{2.35}$$

The last identity can also be written as $k_{AB} = -\sum_{i\in S,\, j\in B} \pi_i l_{ij} q_i^+$ *(for reversible processes!), which in turn is identical to the expression that we know from Theorem* 2.15*:*

$$k_{AB} = \sum_{\substack{i\in S,\, j\in B \\ i\neq j}} \pi_i l_{ij}(1 - q_i^+), \qquad \text{(reversible process)}.$$

2.6. Relations with electrical resistor networks. Before proceeding further, it is interesting to revisit our result in the context of electrical resistor networks [15]. Recall that an *electrical resistor network* is a directed weighted graph $G(S,E) = G\{C\}$, where $C = (c_{ij})$ is an entrywise nonnegative symmetric matrix (see Definition 2.3), called the *conductance matrix* of G. The reciprocal r_{ij} of the conductance c_{ij} is called the *resistance* of the edge (i,j). Establishing a voltage $v_a = 0$ and $v_b = 1$ between two vertices a and b induces a voltage $v = (v_i)_{i\in S\setminus\{a,b\}}$ and an electrical current F_{ij} which are related by Ohm's law:

$$F_{ij} = \frac{v_i - v_j}{r_{ij}} = (v_i - v_j)c_{ij}, \qquad i,j \in S,\ i \neq j. \tag{2.36}$$

Furthermore, Kirchhoff's current law, that is,

$$\sum_{j\in S} F_{ij} = 0 \quad \forall i \in S \setminus \{a,b\}, \tag{2.37}$$

requires that the voltages have the property

$$v_i = \sum_{j\neq i} \frac{c_{ij}}{c_i} v_j \quad \forall i \in S \setminus \{a,b\}, \tag{2.38}$$

where $c_i = \sum_{j\neq i} c_{ij}$. A *reversible* Markov jump process, given by its infinitesimal generator L, can be seen as an electrical resistor network by setting up the conductance matrix C via

$$c_{ij} = \pi_i l_{ij} \quad (j \neq i), \tag{2.39}$$

where $\pi = (\pi_i)_{i\in S}$ is the unique stationary distribution. Now observe that (2.38) reduces to

$$0 = \sum_{j\in S} l_{ij} v_j \quad \forall i \in S \setminus \{a,b\}. \tag{2.40}$$

But this means that the forward committor q^+ with respect to the sets $A = \{a\}$ and $B = \{b\}$ can be interpreted as a voltage (see (2.13)). Moreover, a short calculation shows that the effective flux, defined in (2.33), pertains to the electrical current.

2.7. Dynamical bottlenecks and reaction pathways. The transition rate k_{AB} is a quantity which is important to describe the global transition behavior. In this section we characterize the local *bottlenecks* of the ensemble of reactive trajectories which determine the transition rate. In order to get a detailed insight into the local transition behavior we characterize reaction pathways by looking at the amount of reactive trajectories which is conducted from A to B by a sequence of states.

We use the notation of graph theory introduced at the end of section 2.1. Let $G(S,E) = G\{f^+\}$ be the weight induced directed graph associated with the effective current $f^+ = (f_{ij}^+)$, $i_j \in S$. A simple pathway in the graph G, starting in $A \subset S$ and ending in $B \subset S$, is the natural choice for representing a specific reaction from A to B because any loop during a transition would be redundant with respect to the progress of the reaction.

DEFINITION 2.18. *A* reaction pathway $w = (i_0, i_1, \ldots, i_n)$, $i_j \in S$, $j = 0, \ldots, n$, *from A to B is a simple pathway such that*

$$i_0 \in A,\ i_n \in B,\ i_j \in (A \cup B)^c, \qquad j = 1, \ldots, n-1.$$

The crucial observation which leads to a characterization of bottlenecks of reaction pathways is that the amount of reactive trajectories which can be conducted by a reaction pathway per time unit is confined by the minimal effective current of a transition involved along the reaction pathway.

DEFINITION 2.19. *Let $w = (i_0, i_1, \ldots, i_n)$ be a reaction pathway in $G\{f^+\}$. We define the* min-current *of w by*

$$c(w) = \min_{e=(i,j)\in w} \{f_{ij}^+\}. \tag{2.41}$$

The dynamical bottleneck *of a reaction pathway is the edge with the minimal effective current*

$$(b_1, b_2) = \operatorname*{argmin}_{e=(i,j)\in w} \{f_{ij}^+\}. \tag{2.42}$$

We call such an edge (b_1, b_2) a bottleneck.

Here and in the following we somewhat misuse our notation by writing $e = (i, j) \in w$ whenever the edge e is involved in the pathway $w = (i_0, i_1, \ldots, i_n)$, i.e., if there is an $m \in \{0, \ldots, n-1\}$ such that $(i, j) = (i_m, i_{m+1})$.

Now it is straightforward to characterize the "best" reaction pathway, that is, the one with the *maximal min-current*.

REMARK 2.20. *Notice that the problem of finding a pathway which maximizes the minimal current is known as the* maximum capacity augmenting path problem [1] *in the context of solving the maximal flow problem in a network.*

In general, one cannot expect to find a unique "best" reaction pathway because the bottleneck corresponding to the maximal min-current could be the bottleneck of other reaction pathways too.

DEFINITION 2.21. *Let $\mathbf{W}$ be the set of all reaction pathways and denote the maximal min-current by c_{max}. Then we define the set of the* dominant reaction pathways *$\mathcal{W}_\mathcal{D} \subseteq \mathbf{W}$ by*

$$\mathcal{W}_\mathcal{D} = \{w \in \mathbf{W} : c(w) = c_{max}\}.$$

REMARK 2.22. *To guarantee uniqueness of the bottleneck, we henceforth assume that the nonvanishing effective currents are pairwise different, i.e., $f_e^+ \neq f_{e'}^+$ for all pairs of edges $e = (i, j)$, $e' = (i', j')$ with $f_e^+, f_{e'}^+ > 0$. Nevertheless, we are aware that in applications the situation could show up where more than one bottleneck exists because the corresponding currents are more or less equal. This ambiguity is taken into account in an ordered decomposition of the set of all reaction pathways described at the end of this section.*

Let $G[\mathcal{W}_\mathcal{D}] = G(S_\mathcal{D}, E_\mathcal{D})$ be the directed graph induced by the set $\mathcal{W}_\mathcal{D}$, i.e., the graph whose vertex/edge set is composed of all vertices/edges that appear in at least one of the pathways in $\mathcal{W}_\mathcal{D}$. The next lemma shows that the graph $G[\mathcal{W}_\mathcal{D}] = G(S_\mathcal{D}, E_\mathcal{D})$ possesses a special structure which is crucial for the definition of a *representative dominant reaction pathway.*

LEMMA 2.23. *Let $b = (b_1, b_2)$ denote the unique bottleneck in $G[\mathcal{W}_\mathcal{D}]$. Then the graph $G(S_\mathcal{D}, E_\mathcal{D} \setminus \{b\})$ decomposes into two disconnected parts $G[\mathfrak{L}]$ and $G[\mathfrak{R}]$ such that every reaction pathway $w \in \mathcal{W}_\mathcal{D}$ can be decomposed into two pathways $w_\mathfrak{L}$ and $w_\mathfrak{R}$,*

$$w = (\underbrace{i_{l_1}, \ldots, i_{l_n} = b_1}_{=w_\mathfrak{L}}, \underbrace{b_2 = i_{r_1}, \ldots, i_{r_m}}_{=w_\mathfrak{R}}), \tag{2.43}$$

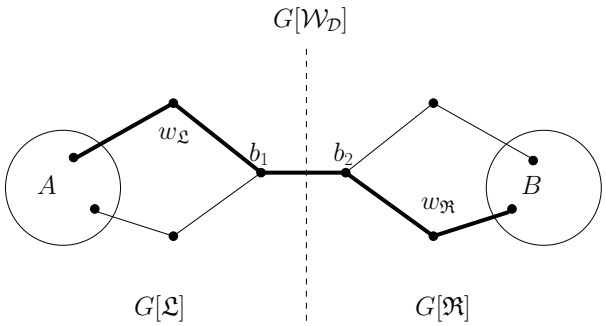

FIG. 2. *Schematic representation of the decomposition of $\mathcal{W}_\mathcal{D}$. A reaction pathway w (shown in thick black) can be decomposed into two simple pathways $w_\mathfrak{L}$ and $w_\mathfrak{R}$.*

where $w_\mathfrak{L} \in \mathfrak{L}$ is a simple pathway in $G[\mathfrak{L}]$ starting in $i_{l_1} \in A$ and ending in $\{b_1\}$ and $w_\mathfrak{R} \in \mathfrak{R}$ is a simple pathway in $G[\mathfrak{R}]$ starting in $\{b_2\}$ and ending up in $i_{r_m} \in B$. Whenever we have $\mathfrak{L} = \emptyset$, i.e., $(b_1 \in A)$, then $G[\mathfrak{L}] = (\{i_{l_1}\}, \emptyset)$; if $\mathfrak{R} = \emptyset$, then $G[\mathfrak{R}]$ is defined likewise.

Here and in the following we write $w_\mathfrak{L} \in \mathfrak{L}$ (and $w_\mathfrak{R} \in \mathfrak{R}$, respectively) if we want to express that for every edge $e \in w_\mathfrak{L}$ we have $e \in \mathfrak{L}$. The lemma expresses the natural property that the graph $G[\mathcal{W}_\mathcal{D}] = G(S_\mathcal{D}, E_\mathcal{D})$ can be decomposed into two disconnected graphs by removing the bottleneck; see Figure 2 for a schematic illustration.

Proof. It immediately follows from the definition of $\mathcal{W}_\mathcal{D}$ that the bottleneck b is involved in every dominant reaction pathway because otherwise there would exist a pathway $w \in \mathcal{W}_\mathcal{D}$ such that $c(w) > c_{max}$, which leads to a contradiction. By definition, a reaction pathway does not possess any loops. Consequently, the bottleneck b separates $\mathcal{W}_\mathcal{D}$, which proves the assertion. ∎

According to the lemma, the set of dominant reaction pathways $\mathcal{W}_\mathcal{D}$ can be represented as

$$\mathcal{W}_\mathcal{D} = \mathfrak{L} \times \mathfrak{R} := \{(w_\mathfrak{L}, w_\mathfrak{R}) : w_\mathfrak{L} \in \mathfrak{L},\, w_\mathfrak{R} \in \mathfrak{R}\}\,. \tag{2.44}$$

In Figure 2 we give a schematic representation of the decomposition of $\mathcal{W}_\mathcal{D}$.

Next, we address the most likely case in applications where more than one dominant reaction pathway exists. By definition, each dominant reaction pathway conducts the same amount of current from A to B, but they could differ, e.g., with respect to the maximal amount of current which they conduct from the set A to the *bottleneck*, respectively. Now observe that the simple pathways in the set $\mathfrak{L}$ could be seen as reaction pathways with respect to the set A and the B-set $\{b_1\}$. Hence, $\mathfrak{L}$ itself again possesses a set of dominant reaction pathways $\mathcal{W}_\mathcal{D}(\mathfrak{L})$, and so on. This motivates the following recursive definition of the *representative* dominant reaction pathway.

DEFINITION 2.24. *Let $\mathcal{W}_\mathcal{D} = \mathfrak{L} \times \mathfrak{R}$ and suppose $b = (b_1, b_2)$ is its (unique) bottleneck. Then we define the* representative *dominant reaction pathway w^* of $\mathcal{W}_\mathcal{D}$ by*

$$w^* = (w^*_\mathfrak{L}, w^*_\mathfrak{R}), \tag{2.45}$$

where $w^*_\mathfrak{L}$ is the representative dominant pathway of the set $\mathcal{W}_\mathcal{D}(\mathfrak{L})$ with respect to the set A and the B-set $\{b_1\}$ and $w^*_\mathfrak{R}$ is the representative of $\mathcal{W}_\mathcal{D}(\mathfrak{R})$ with respect to

the A-set $\{b_2\}$ *and the set* B. *If* $\mathfrak{L} = \emptyset$ *and* $G[\mathfrak{L}] = (\{i\}, \emptyset)$, *then* $w^*_{\mathfrak{L}} = \{i\}$; *if* $\mathfrak{R} = \emptyset$, *then* $w^*_{\mathfrak{R}}$ *is defined likewise.*

Notice that the representative w^* is unique under the assumption made in Remark 2.22. Furthermore, it follows immediately from the recursive definition of w^* that

$$\begin{aligned} w^* &= \operatorname*{argmax}_{w \in \mathcal{W}_{\mathcal{D}}} \min_{\substack{e=(i,j)\in w,\\ (i,j)\neq(b_1,b_2)}} \{f^+_{ij}\} \\ &= \operatorname*{argmax}_{w \in \mathcal{W}_{\mathcal{D}}} \min_{\substack{e=(i,j)\in w,\\ (i,j)\neq(b_1,b_2)}} \{f^+_{ij} - c_{max}\}. \end{aligned} \tag{2.46}$$

Finally, we turn our attention to the *residuum current* which results from updating the effective current of each edge along the representative pathway $w^*_1 = w^*$ by subtracting the min-current $c^{(1)}_{max} = c_{max}$. That is, the residuum current is defined as

$$f^{r,1}_{ij} = \begin{cases} f^+_{ij} - c^{(1)}_{max} & \text{if } (i,j) \in w^*_1, \\ f^+_{ij} & \text{otherwise.} \end{cases} \tag{2.47}$$

The graph $G_1 = G\{f^{r,1}_{ij}\}$ induced by the residuum current satisfies the current conservation property in analogy to (2.26). It again possesses a bottleneck, say $\tilde{b}$, a set of dominant pathways, and a representative pathway, say w^*_2. If we denote the min-current of w^*_2 with respect to the residuum current by $c^{(2)}_{max}$, then it should be clear that $c_{max} = c^{(1)}_{max} > c^{(2)}_{max}$ holds. The property (2.46) of w^*_1 guarantees that $c^{(2)}_{max}$ is maximal with respect to all possible residuum currents. We can obviously repeat this procedure by introducing the residuum current $f^{r,2}_{ij}$ by subtracting $c^{(2)}_{max}$ from $f^{r,1}_{ij}$ along the edges belonging to w^*_2, and so on. The resulting iteration terminates when the resulting induced graph $G_{M+1} = G\{f^{r,M+1}_{ij}\}$ no longer contains reaction pathways and leads to an ordered enumeration $(w^*_1, w^*_2, \ldots, w^*_M)$ of the set $\mathbf{W}$ of all reaction pathways such that

$$\begin{aligned} &c^{(i)}_{max} > c^{(j)}_{max}, \qquad 0 \le i < j \le M, \\ &\sum_{i=1}^{M} c^{(i)}_{max} = k_{AB}, \end{aligned} \tag{2.48}$$

where the last identity simply follows from the following equation for the rates $k_{AB}(G_i)$ associated with the graphs $G_1, \ldots, G_M$:

$$k_{AB}(G_i) = k_{AB}(G_{i-1}) - c^{(i)}_{max},$$

where G_0 denotes the original graph $G\{f^+_{ij}\}$, and $k_{AB}(G_{M+1}) = 0$.

REMARK 2.25. *The composition of the total rate into fraction coming from currents along reactive pathways is quite a general concept in graph theory. We herein just presented a specification of it. We refer the interested reader to, e.g.,* [1, *section* 3.5].

2.8. Relation with Laplacian eigenmaps and diffusion maps. Let us briefly comment about the relevance of our results in the context of data analysis (in particular, data segmentation and embedding, i.e., low-dimensional representation). Recently, two classes of methods have been introduced to this aim: Laplacian eigenmaps [32, 23, 29, 3, 14] and diffusion maps [11, 21]. The idea behind these approaches

is quite simple. Given a set of data points, say $S = \{x_1, x_2, \ldots, x_n\}$, one associates a weight induced graph with weight function $w(x, y)$. This graph is constructed locally, e.g., by connecting all points with equal weights that are below a cut-off distance from each other. These weights are then renormalized by the degree of each node, which means that $w(x, y)$ can be reinterpreted as the stochastic matrix of a discrete-time Markov chain. Alternatively, it is also possible to interpret the weights as rates and thereby build the generator of a continuous-time Markov chain. In both cases, the properties of the chain are then investigated via spectral analysis of the stochastic matrix or the generator. In particular, the first N eigenvectors with leading eigenvalues, say, $\phi_j(x)$, $j = 1, \ldots, N$, can be used to embed the chain into $\mathbb{R}^N$ via $x \mapsto (\phi_1(x), \ldots, \phi_N(x))$. The eigenvectors can also be used to segment the original data set into important components (segmentation).

As explained in the introduction, the spectral approach is particularly relevant if the Markov chain displays metastability, i.e., if there exists one or more clusters of eigenvalues which are either very close to 1 (in the case of discrete-time Markov chains) or 0 (in the case of continuous-time Markov chains). When the chain is not metastable, however, the meaning of the first few eigenvectors is less clear, which makes the spectral approach less appealing. In these situations, TPT may provide an interesting alternative. For instance, if several points (or groups of points) with some specific properties can be singled out in the data set, then, by analyzing the reaction between pairs of such groups, one will disclose global information about the data set (for instance, the committor functions between these pairs may be used for embedding instead of the eigenvectors). The current of reactive trajectories and dominant reaction pathways will also provide additional information about the global structure of the data set which is not considered in the spectral approach.

In this paper, we will not, however, develop these ideas any further.

3. Algorithmic aspects. In this section we explain the algorithmic details for the computation of the various quantities in TPT. Given the generator L and the two sets A and B, the stationary distribution $\pi = (\pi_i)_{i \in S}$ is computed by solving (2.2), whereas the discrete forward and discrete backward committors, $q^+ = (q_i^+)_{i \in S}$ and $q^- = (q_i^-)_{i \in S}$, are computed by solving (2.13) and (2.14). Solving these equations numerically can be done using any standard linear algebra package. These objects allow one to compute the probability distribution of reactive trajectories $m^R = (m_i^R)_{i \in S}$ in (2.17), its normalized version $m^{AB} = (m_i^{AB})_{i \in S}$ in (2.20), the probability current of reactive trajectories $f^{AB} = (f_{ij}^{AB})_{i,j \in S}$ in (2.24), and the effective current $f^+ = (f_{ij}^+)_{i,j \in S}$ in (2.33). This also gives the reaction rate k_{AB} via (2.30) or (2.32). Next, we focus on the computation of the bottlenecks and representative dominant reaction pathways which is less standard.

3.1. Computation of dynamical bottlenecks and representative dominant reaction pathways. From the definition in (2.42) of the bottleneck $b = (b_1, b_2)$ associated with the set of dominant reaction pathways $\mathcal{W}_{\mathcal{D}}$, it follows that

$$f_e^+ > f_b^+ \quad \forall e \in E_{\mathcal{D}},\, e \neq b,$$

where $f^+ = (f_{ij}^+)_{i,j \in S}$ is the effective current and $E_{\mathcal{D}}$ is the edge set of the induced graph $G = G[\mathcal{W}_{\mathcal{D}}]$. This observation leads to a characterization of the bottleneck which is algorithmically more convenient. Let $E_{sort} = (e_1, e_2, \ldots, e_{|E|})$ be an enumeration of the set of edges of $G = G\{f^+\}$ sorted in ascending order according to their effective current. Then the edge $b = e_m$ in E_{sort} is the bottleneck if and

only if the graph $G(S, \{e_m, \dots, e_{|E|}\})$ contains a reaction pathway but the graph $G(S, \{e_{m+1}, \dots, e_{|E|}\})$ does not. The bisection algorithm stated in Algorithm 1 is a direct consequence of this alternative characterization of the bottleneck and is related to the *capacity scaling algorithm* [1, section 7.3] for solving the maximum flow algorithm. For an alternative algorithm in the context of distributed computing which is based on a modified Dijkstra algorithm; see [18].

ALGORITHM 1. Computation of the bottleneck

Input: Graph $G = G\{f^+\}$.
Output: Bottleneck $b = (b_1, b_2)$.
(1) Sort edges of G according to their weights in ascending order
$\Longrightarrow E_{sort} = (e_1, e_2, \dots, e_{|E|})$.
(2) **IF** the edge $e_{|E|}$ connects A and B **THEN RETURN** bottleneck $b := e_{|E|}$.
(3) Initialize $l := 1$, $r := |E|$.
(4) **WHILE** $r - l > 1$
(5) Set $m := \lfloor \frac{r+l}{2} \rfloor$, $E'(m) := \{e_m, \dots, e_{|E|}\}$.
(6) **IF** there exists a reaction pathway in $G(S, E'(m))$
(7) **THEN** $l := m$ **ELSE** $r := m$.
(8) **END WHILE**
(9) **RETURN** bottleneck $b := e_l$.

We also have the following lemma.

LEMMA 3.1. *The computational cost of Algorithm* 1 *in the worst case is* $\mathcal{O}(n \log n)$, *where* $n = |E|$ *denotes the number of edges of the graph* $G = G\{f^+\}$.

Proof. Assume that $n = 2^k$, $k > 1$. First, notice that the sorting of the edges of $G = G\{f^+\}$ can be performed in $\mathcal{O}(n \log n)$. In the worst case scenario, the edge $e_1 \in E_{Sort}$ is the bottleneck.[1] When this is the case, the number of edges in the jth repetition of the while-loop would be

$$\frac{n}{2^j},$$

and we would have $k - 1$ repetitions. The cheapest way to determine whether there exists a reactive trajectory is to perform a breadth-first search starting in A; the computational cost of that step depends only linearly on the number of edges to be considered, such that we deduce for the worst case effort $T(n)$ of the entire procedure

$$\begin{aligned} T(n) &= \mathcal{O}(kn) + \mathcal{O}\left(\frac{n}{2}\right) + \mathcal{O}\left(\frac{n}{4}\right) + \cdots + \mathcal{O}\left(\frac{n}{2^{k-1}}\right) \\ &= \mathcal{O}\left(kn + n\left(\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^{k-1}}\right)\right) \\ &= \mathcal{O}(kn), \end{aligned}$$

which by noting that $k = \log(n)$ ends the proof. ∎

The algorithm for computing the unique representative pathway w^* of the set of dominant reaction pathways is a direct implementation of the recursive definition of

[1] We are aware that the edge e_1 could never be the bottleneck unless all effective currents are equal, which by Remark 2.22 is excluded. Nevertheless, the following reasoning with respect to e_1 leads only to a slight overestimation of the computational cost.

w^* given in (2.45). Recalling that $\mathcal{W}_{\mathcal{D}}$ can be decomposed as stated in (2.44) and assuming that f^+ takes different values for every edge (i,j), we end up with Algorithm 2. A rough estimation of the computational cost of this algorithm is $\mathcal{O}(mn\log n)$, where m is the number of edges of the resulting representative pathway w^* and $n=|E|$.

ALGORITHM 2. Representative pathways

Input: Graph $G=G\{f^+\}$, set A, set B.
Output: Representative $w^*=(w^*_{\mathfrak{L}},w^*_{\mathfrak{R}})$ of $\mathcal{W}_{\mathcal{D}}(G)$.
(1) Determine bottleneck $b=(b_1,b_2)$ in G via Algorithm 1.
(2) Determine decomposition $\mathcal{W}_{\mathcal{D}}(G)=\mathfrak{L}\times\mathfrak{R}$.
(3) Set $w^*_{\mathfrak{L}}:=\begin{cases} b_1 & \text{if } b_1\in A,\\ \text{result of the recursion with } (G[\mathfrak{L}],A,\{b_1\}) & \text{if } b_1\notin A.\end{cases}$
(4) Set $w^*_{\mathfrak{R}}:=\begin{cases} b_2 & \text{if } b_2\in B,\\ \text{result of the recursion with } (G[\mathfrak{R}],\{b_2\},B) & \text{if } b_2\notin B.\end{cases}$
(5) **RETURN** $(w^*_{\mathfrak{L}},w^*_{\mathfrak{R}})$.

4. Illustrative examples. In this section we illustrate the discrete TPT in three examples. The first is the discrete equivalent of a diffusion, which we chose because the results of TPT are transparent in this case. This example also establishes a link to the case of continuous state-space. The second example deals with a problem in molecular dynamics, the trialanine molecule, and shows that TPT allows us to characterize reaction pathways between molecular conformations. There is an additional difficulty in this example, namely that the process is given by an incomplete observation of the system in a certain time interval, meaning that we have to deal with the issue of reconstructing the generator of the process given the time series. The third example we consider is a nonreversible Markov process arising from the modeling of a genetic toggle switch in chemical kinetics.

4.1. Discrete analogue of a diffusion in a potential landscape. In [25], TPT for diffusion processes was illustrated in the example of a particle whose dynamics is governed by the stochastic differential equation

$$\begin{cases} \mathrm{d}x(t) = -\dfrac{\partial V(x(t),y(t))}{\partial x}\mathrm{d}t + \sqrt{2\beta^{-1}}\,\mathrm{d}W_x(t),\\ \mathrm{d}y(t) = -\dfrac{\partial V(x(t),y(t))}{\partial y}\mathrm{d}t + \sqrt{2\beta^{-1}}\,\mathrm{d}W_y(t),\end{cases} \tag{4.1}$$

where $(x(t),y(t))\in\mathbb{R}^2$ denotes the position of the particles, $V(x,y)$ is the potential, $\beta>0$ is a parameter referred to as the inverse temperature, and $W_x(t)$ and $W_y(t)$ are two independent Wiener processes, i.e., Gaussian processes with mean zero and covariance $\mathbb{E}W_x(t)W_x(s)=\mathbb{E}W_y(t)W_y(s)=\min(t,s)$. For $V(x,y)$ in [25] we chose the three-hole potential

$$\begin{aligned} V(x,y) = {} & 3e^{-x^2-(y-\frac{1}{3})^2} - 3e^{-x^2-(y-\frac{5}{3})^2}\\ & - 5e^{-(x-1)^2-y^2} - 5e^{-(x+1)^2-y^2}\\ & + \frac{2}{10}x^4 + \frac{2}{10}(y-\tfrac{1}{3})^4 \end{aligned} \tag{4.2}$$

which has been already considered in [20, 27, 13]. As one can see in Figure 3 the

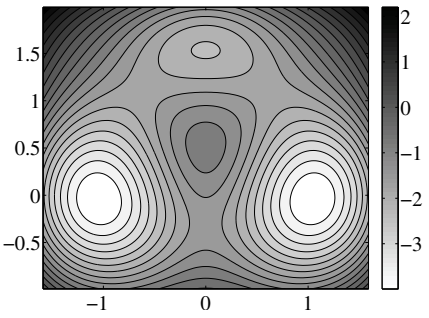

FIG. 3. *Level sets of the three-hole potential.*

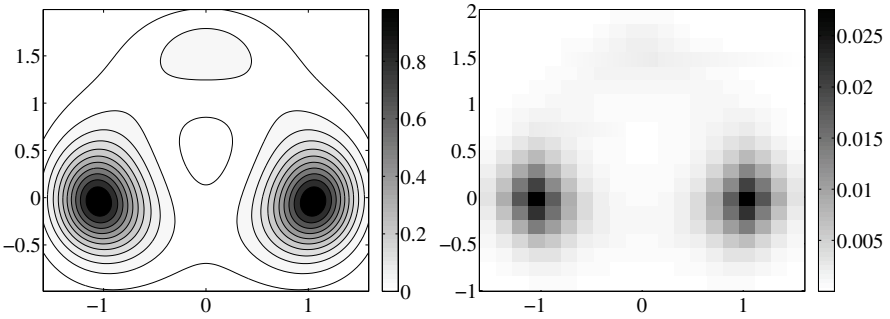

FIG. 4. *Left: Contour plot of the equilibrium density function* $\exp(-\beta V(x))$. *Right: Box plot of the stationary distribution* $(\pi_{(x,y)})_{(x,y)\in S}$. *Results for* $\beta = 1.67$ *and a* 20×20 *mesh discretization.*

potential (4.2) has two deep minima approximately at $(\pm 1, 0)$, a shallow minimum approximately at $(0, 1.5)$, three saddle points approximately at $(\pm 0.6, 1.1)$, $(-1.4, 0)$, and a maximum at $(0, 0.5)$. The process defined by (4.1) is ergodic with respect to the Gibbs measure

$$\mathrm{d}\mu(x, y) = Z^{-1} \exp(-\beta V(x, y))\mathrm{d}x\mathrm{d}y, \tag{4.3}$$

where $Z = \int_{\mathbb{R}^2} \exp(-\beta V(x, y))\mathrm{d}x\mathrm{d}y$ is a normalization constant. If β is small enough, then the measure is strongly peaked on the deep minima of the potential (see the left panel of Figure 4), and the system displays metastability; i.e., the particle makes transitions between the vicinity of these minima only very rarely. In [25] it was shown that TPT can be used to describe the mechanism of the transition and compute their rates. In particular, it was shown that transitions preferably occur by the upper channel visible in Figure 3 when β is very small but that they proceed by the lower channel when β is somewhat increased. The reasons for this entropic switch were elucidated in [25], and we refer the reader to this paper for details. Our purpose here is to apply TPT on a discrete analogue of (4.1).

In order to construct this analogue, we exploit the well-known fact that a diffusion process can be approximated by a Markov jump process after discretization of state-space (see, e.g., [17]). Here we approximate the dynamics (4.1) on a two-dimensional, rectangular domain $\Omega = [a, b] \times [c, d] \subset \mathbb{R}^2$ via a birth-death process on the discrete state-space (mesh) $S = ((a + h\mathbb{Z}) \times (c + h\mathbb{Z})) \cap ([a, b] \times [c, d])$, where $h > 0$ is the uniform mesh width. For clarity, in the present example we will denote by (x, y) the

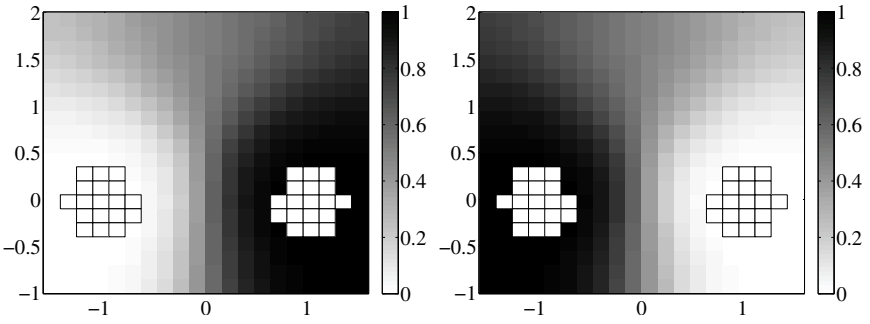

FIG. 5. *Box plot of the discrete committors. Left: Forward committor q^+. Right: Backward committor q^-. Results for $\beta = 1.67$ and a 20×20 mesh discretization.*

state that is denoted by i. Then the generator is given in terms of its action on a test function f as

$$\begin{aligned}(Lf)(x,y) = {} & k_x^+(x+h,y)(f(x+h,y)-f(x,y)) \\ & + k_x^-(x-h,y)(f(x-h,y)-f(x,y)) \\ & + k_y^+(x,y+h)(f(x,y+h)-f(x,y)) \\ & + k_y^-(x,y-h)(f(x,y-h)-f(x,y)),\end{aligned} \tag{4.4}$$

where

$$k_x^+(x+h,y) = \begin{cases} \dfrac{\beta^{-1}}{h^2} - \dfrac{1}{2h}\dfrac{\partial V(x,y)}{\partial x} & \text{if } x \in (a,b), \\ 0 & \text{if } x = b, \\ \dfrac{1}{h} & \text{if } x = a, \end{cases}$$

$$k_x^-(x-h,y) = \begin{cases} \dfrac{\beta^{-1}}{h^2} + \dfrac{1}{2h}\dfrac{\partial V(x,y)}{\partial x} & \text{if } x \in (a,b), \\ 0 & \text{if } x = a, \\ \dfrac{1}{h} & \text{if } x = b \end{cases}$$

and the coefficients k_y^+ and k_y^- are defined analogously with respect to $\partial V(x,y)/\partial y$. In the left panel of Figure 4 we show the level sets of the density function $\exp(-\beta V(x,y))$ associated with the Gibbs measure (4.3). In the right panel of Figure 4 we illustrate the stationary distribution $\pi = (\pi_{(x,y)})_{(x,y)\in S}$ of the birth-death process as a *box plot*.

We now present the results of TPT in this example. The panels in Figure 5 show the box plots of the forward committor q^+ (left panel) and the backward committor q^- (right panel). The set $A \subset S$ is chosen such that it sufficiently covers the region around the left minimum. The set B is defined analogously for the right minimum. The symmetry of the potential together with the symmetry of the sets A and B implies that the particular $\frac{1}{2}$-committor surface, defined as the set $\{(x,y) \in S : q^+_{(x,y)} = 0.5\}$, should correspond to the symmetry axis in y-direction, which is confirmed in Figure 5. Notice how the presence of the shallow minima in the upper part of the potential spreads the "level sets" of q^+ in this region. This follows from the fact that the reactive trajectories going through the upper channel get trapped in the shallow well

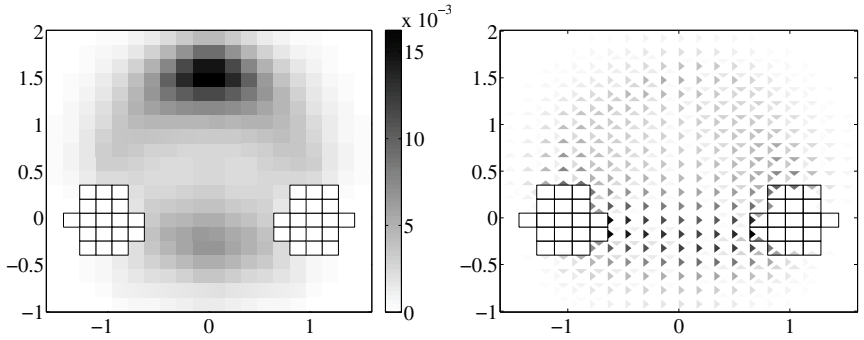

FIG. 6. *Left: Box plot of the discrete probability distribution of reactive trajectories* m^{AB}. *Right: Visualization of the effective current* f^+ *between mesh points (boxes). An edge* $((x,y),(x',y'))$ *with positive effective current* $f^+_{((x,y),(x',y'))}$ *is depicted by a triangle pointing from the box which corresponds to the state* (x,y) *towards the box identified with* $(x',y') \in S$. *The darker the color of a triangle, the higher is the effective current.*

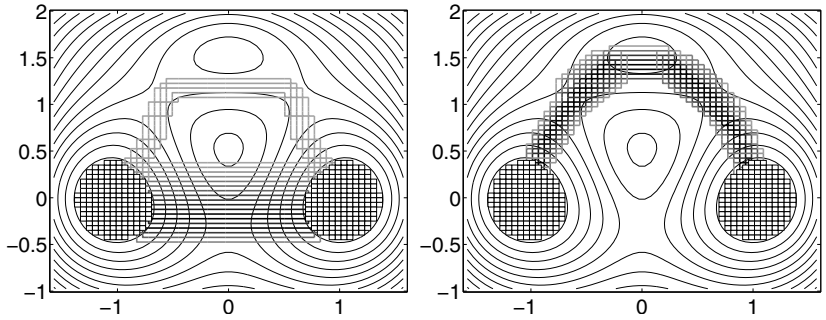

FIG. 7. *Reaction pathway families for two different temperatures. Both families cover about* 50% *of the probability flux of reactive trajectories. The pathways are colored according to the values of their min-currents. The darker the color, the more current that is conducted by the corresponding reaction pathway. Left: Reaction pathway family at a high temperature* $\beta = 1.67$. *Right: Reaction pathway family at a low temperature* $\beta = 6.67$. *Results for a* 60×60 *mesh discretization; for the sake of illustration the mesh is chosen finer than before.*

for a long period of time before exiting towards the set B. Next, we turn our attention to the probability distribution of the reactive trajectories, shown in the left panel of Figure 6. One can see that the distribution has a peak in the upper shallow minima, whereas the effective current, visualized in the right panel of Figure 6, suggests that most of the reactive trajectories prefer the lower channel. This again can be explained by the fact that the reactive trajectories going through the upper channel get trapped in the shallow well, whereas the reactive trajectories in the lower channel just need to overcome the barrier. We end this example by discussing the family of dominant reaction pathways resulting from the procedure described at the end of section 3.1. In the left panel of Figure 7 we plot the family of reaction pathways which covers about 50% of the probability flux of reactive trajectories at the temperature $\beta = 1.67$. The pathways are colored according to the values of their min-currents. The darker the color, the larger is the current conducted by the corresponding reaction pathway. At the high temperature ($\beta = 1.67$, left panel), the reaction occurs mostly via the lower channel, whereas at the low temperature ($\beta = 6.67$, right panel) it occurs mostly via the upper channel. This is consistent with the results presented in [27, 25].

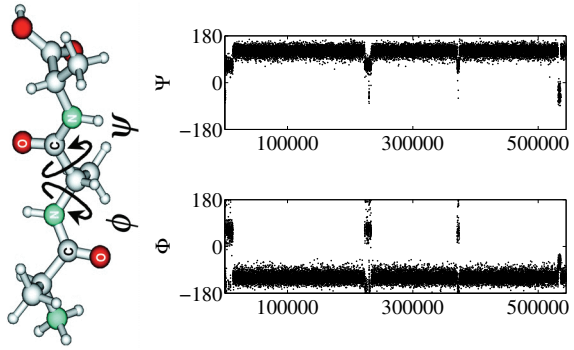

FIG. 8. *Left: The trialanine molecule shown in ball-and-stick representation and the two torsion angles* Φ *and* Ψ*. Right: Projection of the original time series (all atomic positions) onto the torsion angle space spanned by* Φ *and* Ψ*, which reveals the metastable behavior.*

4.2. Molecular dynamics: Trialanine. In this example we use discrete TPT to study conformation changes of the trialanine molecule which is shown in ball-and-stick representation in the left panel of Figure 8. Unlike in the first example, here the process is implicitly given by a time series of two torsion angles. The time series used herein was generated in vacuum using the hybrid Monte Carlo method [9] with 544.500 steps with GROMACS force field [5, 22] at a temperature of 750K. The integration of the subtrajectories of the proposal step was realized with $\tau = 1$ fs time steps of the Verlet integration scheme. Hybrid Monte Carlo is based on a discrete-time Markov chain in continuous state-space. Before explaining how we constructed a Markov jump process with discrete state-space (and especially its generator) out of this time series, let us give some background about this example.

4.2.1. Metastability and conformation states. A conformation of a molecule is understood as a mean geometric structure of the molecule which is conserved on a large time scale compared to the fastest molecular motions. From the dynamical point of view, a conformation typically persists for a long time (again compared to the fastest molecular motions) such that the associated subset of configurations is *metastable* [31]. In the right panel of Figure 8 we show the projection of the time series of the torsion angles Φ and Ψ which clearly reveals the metastable behavior. The Ramachandran plot of the time series in the left panel of Figure 9 illustrates the dependency of the conformation states on the two torsion angles. At first glance, the molecule attains three conformations in the torsion angle space.

4.2.2. Generator estimation. The first step towards the application of discrete TPT is to determine a coarse grained model of the dynamics in the torsion angle space based on the given time series. We discretized the two-dimensional torsion angle space with an equidistant box discretization and identified each element of the time series with the box by which it is covered. Assuming that the resulting discrete time series is Markovian, we estimated a reversible Markov jump process on the discrete state-space of boxes which most likely explains the discrete time series. This is done using an efficient generalization of the method recently presented in [6]; for details see [24]. The idea behind this method is to determine a generator such that it maximizes the discrete likelihood of the given incomplete observation which is accomplished by an *expectation-maximization* algorithm.

In the following, we denote by $\tilde{L} = (\tilde{l}_{(\Phi,\Psi),(\Phi',\Psi')})_{(\Phi,\Psi),(\Phi',\Psi')\in S}$ the infinitesimal

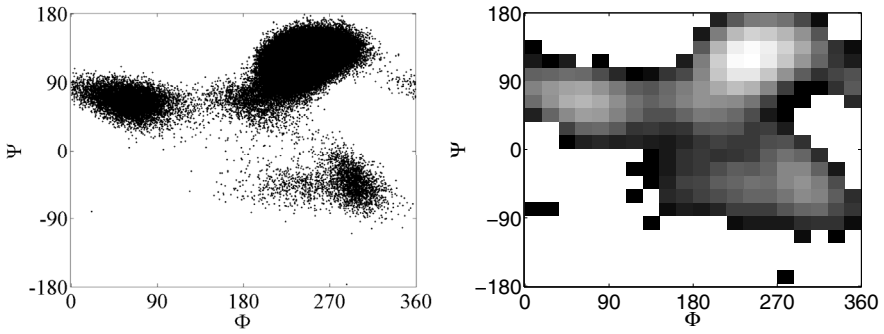

FIG. 9. *Left: Ramachandran plot of the torsion angles. Right: Box plot of the Gibbs energy,* $-\log(\pi_{(\Phi,\Psi)})$*, where* $(\pi_{(\Phi,\Psi)})_{(\Phi,\Psi)\in S}$ *is the stationary distribution computed from the estimated generator* $\tilde{L}$*. Results for an equidistant discretization of the torsion angle space into* 20×20 *boxes.*

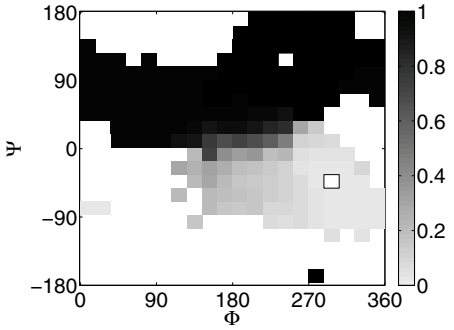

FIG. 10. *The forward committor* q^+ *computed via* (A.2)*. As the set A we chose the box (shown as a white box with black boundary) which covers the peak of the restricted stationary distribution on the lower conformation. The set B for the upper conformation (shown as a white box) was chosen analogously.*

generator of the estimated Markov jump process. For the sake of illustration, we show in the right panel of Figure 9 the *Gibbs energy*, $-\log \pi_{(\Phi,\Psi)}$, where $(\pi_{(\Phi,\Psi)})_{(\Phi,\Psi)\in S}$ is the stationary distribution computed from the estimated generator $\tilde{L}$ with respect to a 20×20 box discretization. The lighter the color of the boxes, the more probable it is to encounter the equilibrated process in the corresponding state. As one can see, the estimated process spends most of its time in three nonoverlapping regions which correspond to the three conformations, respectively.

4.2.3. Analysis within TPT. We were interested in the reaction pathways between the main conformations—the upper right one and the lower right one. As the set B we chose the box in which the Gibbs energy restricted on the upper right conformation attains its minimum. The set A was selected analogously with respect to the lower conformation. The discrete forward committor q^+ is given in Figure 10. Comparison of the distribution of reactive trajectories m^{AB} (illustrated in the left panel of Figure 11) with the family of dominant reaction pathways (right panel of Figure 11) again reveals that m^{AB} is insufficient to describe the effective dynamics from A to B. Again, this is explained by noting that whenever a reactive trajectory makes a transition from A to B via the upper left conformation it gets trapped in that upper left conformation, and thus it is more likely to encounter a reactive trajectory

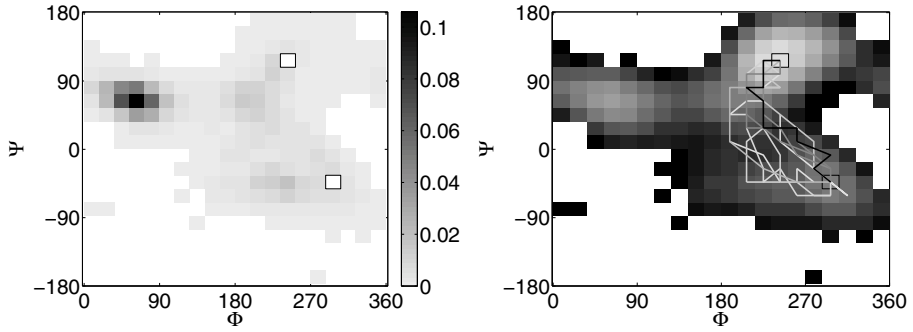

FIG. 11. *Left: Box plot of the discrete probability distribution of reactive trajectories* m^{AB}. *Right: Family of dominant reaction pathways which cover* 40% *of the transition rate. The darker the color of a pathway, the more current it conducts from A to B. For the sake of illustration, the dominant reaction pathways are embedded in the box plot of the Gibbs energy.*

there than in the direct channel.

The results shown here do not change significantly when the mesh of discretization boxes is refined.

4.3. Chemical kinetics. In the last example, we consider a Markov jump process which arises as a stochastic model of a genetic toggle switch consisting of two genes that repress each others' expression [28].

The expression of each of the two respective genes results in the production of a specific type of protein; gene G_A produces protein P_A and gene G_B protein P_B. Denote the number of available proteins of type P_A by x and of type P_B by y; the model for the toggle switch proposed in [28] is a birth-death process on the discrete state-space $S = (\mathbb{Z} \times \mathbb{Z}) \cap ([0, d_1] \times [0, d_2]), d_1, d_2 0$, whose generator is given in terms of its action on a test function f as

$$\begin{aligned}(Lf)(x,y) &= c_1(x+1,y)(f(x+1,y) - f(x,y)) \\ &\quad + \frac{x}{\tau_1}(f(x-1,y) - f(x,y)) \\ &\quad + c_2(x,y+1)(f(x,y+1) - f(x,y)) \\ &\quad + \frac{y}{\tau_2}(f(x,y-1) - f(x,y)),\end{aligned} \tag{4.5}$$

where

$$c_1(x+1,y) = \begin{cases} \dfrac{a_1}{1+(y/K_2)^n} & \text{if } x \in [0, d_1), \\ 0 & \text{if } x = d_1, \end{cases}$$

$$c_2(x,y+1) = \begin{cases} \dfrac{a_2}{1+(x/K_1)^m} & \text{if } y \in [0, d_2), \\ 0 & \text{if } y = d_2. \end{cases}$$

We refer the reader to [28] for the biological interpretation of the parameters in (4.5). For our numerical experiments, we used the parameters $a_1 = 156$, $a_2 = 30$, $n = 3$, $m = 1$, $K_1 = K_2 = 1$, and $\tau_1 = \tau_2 = 1$, consistent with [28]. With these parameters the system's dynamical behavior is as follows: There are two "metastable" states; in the first of these, only gene G_A is expressed and protein P_A is produced until a

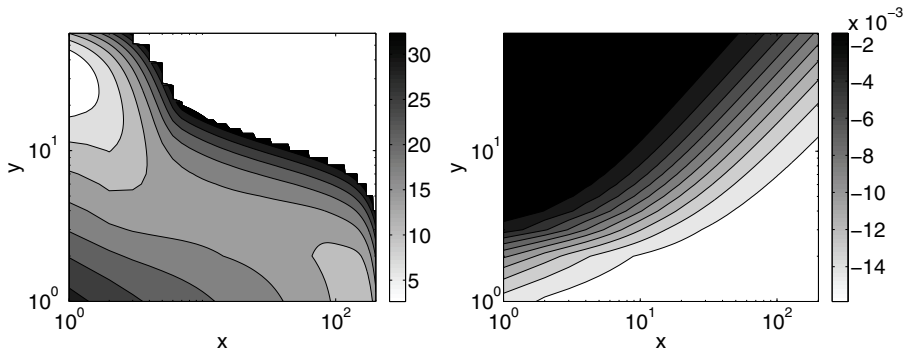

FIG. 12. *Left: Contour plot of the Gibbs energy,* $-\log \pi_{(x,y)}$*, of the birth-death process* (4.5) *on the state-space* $S = \mathbb{Z} \times \mathbb{Z} \cap ([0, 200] \times [0, 60])$*. The white region in the right upper part of the panel indicates the subset of states with almost vanishing stationary distribution (all boxes with distribution less than machine precision have been colored white). Right: Contour plot of the eigenvector of the first nontrivial right eigenvalue of* L*. Results for* $a_1 = 156$*,* $a_2 = 30$*,* $n = 3$*,* $m = 1$*,* $K_1 = K_2 = 1$*, and* $\tau_1 = \tau_2 = 1$*.*

certain number (around $x = 155$ for the parameters chosen) is reached which then is rather stable, while gene G_B is repressed and almost no protein P_B is produced (so that typically $y = 0$ or $y = 1$). After some rather long period of fluctuation in this metastable state the system is able to exit from it which leads to expression of gene G_B and repression of G_A. Then the system gets into a metastable state where the number of protein P_B fluctuates around a certain nonvanishing number ($y = 30$ for our parameters) and P_A is rather not produced (typically $x = 0$ or $x = 1$).

For the sake of illustration, we illustrate in the left panel of Figure 12 the Gibbs energy, $-\log \pi$, of the birth-death process instead of its stationary distribution π itself. Moreover, we neglected all states with almost vanishing stationary distribution (depicted by the white region), and, in order to emphasize the states of interest, we chose a log-log representation. The color scheme is chosen such that the darker the color of a region, the more probable it is to find the process there. One can clearly see that the process spends most of its time near the two metastable core sets $(x, y) \in \{(155, 0), (155, 1)\}$ and $(x, y) \in \{(0, 30), (1, 30)\}$.

We were interested in the reaction from the set $A = \{(155, 0), (155, 1)\}$ towards the set $B = \{(0, 30), (1, 30)\}$. The different shapes of the level sets of the discrete forward and discrete backward committors, as shown in the left and right panels of Figure 13, indicate the high nonreversibility of the birth-death process. Notice that the geometry of the level sets of the forward committor q^+ looks very similar to the geometry of the eigenvector associated with the first nontrivial right eigenvalue of L, as plotted in the right panel of Figure 12. Finally, the edges of the three most dominant reaction pathways are plotted in the right panel of Figure 14. Again, the reaction pathways deviate from the channel which is suggested by the distribution m^{AB} of reactive trajectories, shown in the left panel of Figure 14.

5. Conclusion. We developed the framework of transition path theory (TPT) in the context of continuous-time Markov chains on discrete state-space. Under assumption of ergodicity, TPT analyzes the statistical properties of the ensemble of reactive trajectories between some start and target sets, and it yields properties such as the probability distribution of the reactive trajectories, their effective probability current, and their rate of occurrence and the dominant reaction paths.

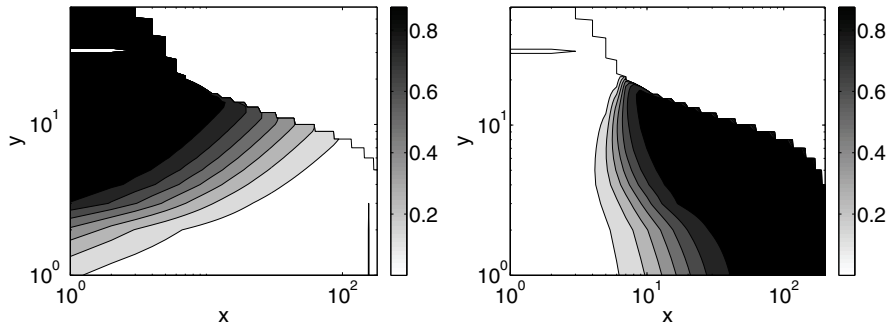

FIG. 13. *Contour plots of the discrete forward and discrete backward committors. Due to the logarithmic scaling, the set* $A = \{(155, 0), (155, 1)\}$ *is depicted as a vertical black line and the set* $B = \{(0, 30), (1, 30)\}$ *as an ellipsoid. Left: Discrete forward committor* q^+*. Right: Discrete backward committor* q^-*.*

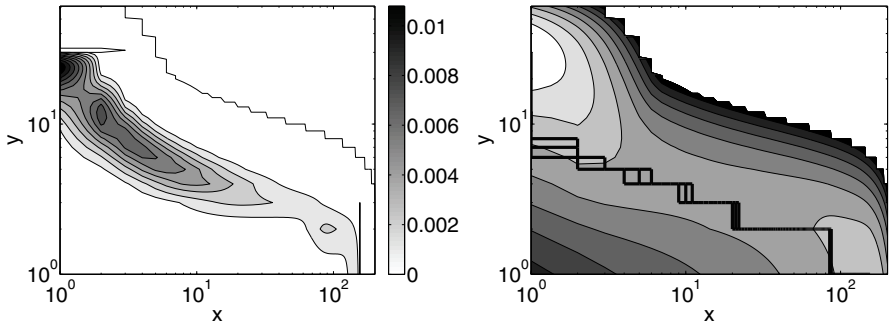

FIG. 14. *Left: Contour plot of the distribution of reactive trajectories* m^{AB}*. Right: Edge plot of the three dominant reaction pathways which cover approximately* 6% *of the current.*

Whenever the generator of the Markov chain is given, the computational tasks related to TPT are those of solving linear equations of the dimension of the state-space and some ordered max-min flux problems on directed graphs. The efficient solution of the latter task has been discussed in detail including links to related literature. As emphasized the assumption of pairwise different effective currents can be relaxed; however, one should be aware that the pathological situation of very many bottlenecks carrying the same current can cause inefficiency. Thus, at least for nonpathological cases there are efficient algorithms from numerical and discrete mathematics for the two computational tasks, such that TPT can also be applied to rather large state-spaces.

As demonstrated, the TPT framework has many interesting relations to other topics in the Markov chain and network literature; we discussed the relation to electric resistor network theory and data segmentation tools such as Laplacian eigenmaps and diffusion maps. Future investigations should work out these and other relations in more detail.

Appendix. Discrete committor equations. The discrete forward and discrete backward committors play a central role in TPT. Recall that for a state $i \in S$ the discrete forward committor q_i^+ is defined as the probability that the Markov jump process starting in state i will reach B rather than A. In other words, q_i^+ is the first

entrance probability of the process $\{X(t),\, t \geq 0,\, X(0) = i\}$) with respect to the set B *avoiding* the set A. The usual step in dealing with entrance or hitting probabilities with respect to a certain subset of states is the modification of the process such that these states become *absorbing* states. Let $L = (l_{ij})_{i,j \in S}$ be the infinitesimal generator of a Markov jump process and $A \subset S$ be a nonempty subset. Suppose we are interested in the process resulting from the declaration of the states in A to be absorbing states. Then the infinitesimal generator $\hat{L} = (\hat{l}_{ij})_{i,j \in S}$ of the modified process is given by [33]

$$\hat{l}_{ij} = \begin{cases} l_{ij}, & i \in A^c,\, j \in S, \\ 0, & i \in A,\, j \in S. \end{cases} \tag{A.1}$$

From this viewpoint, now it is simple to prove the following theorem.

THEOREM A.1. *Let q_i^+ be the probability of reaching B before A, provided that the process has started in state $i \in S$. Then the discrete forward committor $q^+ = (q_i^+)_{i \in S}$ satisfies the equations*

$$\begin{cases} \sum_{k \in S} l_{ik} q_k^+ = 0 & \forall i \in (A \cup B)^c, \\ q_i^+ = 0 & \forall i \in A, \\ q_i^+ = 1 & \forall i \in B. \end{cases} \tag{A.2}$$

Proof. If we make the states in the set A absorbing states, then the discrete forward committor q^+ is the first entrance probability with respect to the set B under the modified process. Thus q^+ satisfies the discrete Dirichlet problem [33]

$$\begin{cases} \sum_{k \in S} \hat{l}_{ik} q_k^+ = 0 & \forall i \in B^c, \\ q_i^+ = 1 & \forall i \in B \end{cases} \tag{A.3}$$

or, equivalently,

$$\begin{cases} \sum_{k \in S} l_{ik} q_k^+ = 0 & \forall i \in (A \cup B)^c, \\ q_i^+ = 0 & \forall i \in A, \\ q_i^+ = 1 & \forall i \in B, \end{cases} \tag{A.4}$$

which ends the proof. □

Observe that if we substitute the "boundary conditions" into the equations in (A.2), then we end up with a linear system

$$U q^+ = v, \tag{A.5}$$

where the matrix $U = (u_{ij})_{i,j \in (A \cup B)^c}$ is given by

$$u_{ij} = l_{ij}, \quad i, j \in (A \cup B)^c,$$

and an entry of the vector $v = (v_i)_{i \in (A \cup B)^c}$ on the right-hand side of (A.5) is defined by $v_i = -\sum_{k \in B} l_{ik}$ for all $i \in (A \cup B)^c$. Now we can prove the following lemma.

LEMMA A.2. *If the matrix U is irreducible, then the solution of* (A.2) *is unique.*

Proof. By the definition of the matrix U there exists at least an index $k \in (A \cup B)^c$ such that

$$|u_{kk}| > \sum_{j \neq k} u_{kj}.$$

But this implies that U is weakly diagonally dominant. Together with its assumed irreducibility, this implies that it is invertible [19]. ∎

Next, we turn our attention to the discrete backward committor q_i^-, $i \in S$, which is defined as the probability that the process arriving at state i came from A rather than from B. The crucial observation is now that $q^- = (q_i^-)_{i \in S}$ is the discrete forward committor with respect to the *reversed time process*.

THEOREM A.3. *The discrete backward committor $q^- = (q_i^-)_{i \in S}$ satisfies the linear system of equations*

$$\begin{cases} \sum_{k \in S} \tilde{l}_{ik} q_k^- = 0 & \forall i \in (A \cup B)^c, \\ q_i^- = 1 & \forall i \in A, \\ q_i^- = 0 & \forall i \in B, \end{cases} \tag{A.6}$$

where $\pi = (\pi_i)_{i \in S}$ is a stationary distribution and $\tilde{l}_{ik} = \pi_k l_{ki}/\pi_i$ is the generator of the reversed time process (see (2.4)). Moreover, if the Markov jump process is reversible, then the backward committor is simply related to the forward committor by

$$q^- = 1 - q^+. \tag{A.7}$$

Proof. The derivation of (A.6) is a straightforward generalization of the one of (A.2). To derive (A.7), note that if the Markov jump process is reversible, then the detailed balance condition

$$\pi_i l_{ij} = \pi_j l_{ji} \quad \forall i, j \in S$$

is satisfied and the discrete backward committor solves

$$\begin{cases} \sum_{k \in S} l_{ik} q_k^- = 0 & \forall i \in (A \cup B)^c, \\ q_i^- = 1 & \forall i \in A, \\ q_i^- = 0 & \forall i \in B. \end{cases} \tag{A.8}$$

On one hand, the solution of the discrete Dirichlet problem (2.14) is unique (see (A.2)). On the other hand, a short calculation shows that $1 - q^+$ also satisfies (2.14). Consequently, we have $q^- = 1 - q^+$, which ends the proof. ∎

REFERENCES

[1] R. K. AHUJA, T. L. MAGNANTI, AND J. B. ORLIN, *Network Flows*, Prentice–Hall, Englewood Cliffs, NJ, 1993.

[2] R. ALBERT AND A.-L. BARABÁSI, *Statistical mechanics of complex networks*, Rev. Modern Phys., 74 (2002), pp. 48–97.

[3] M. BELKIN AND P. NIYOGI, *Laplacian eigenmaps for dimensionality reduction and data representation*, Neural Comput., 6 (2003), pp. 1373–1396.

[4] G. Ben Arous, A. Bovier, and V. Gayrard, *Aging in the random energy model under Glauber dynamics*, Phys. Rev. Lett., 88 (2002), 087201.
[5] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen, *GROMACS: A message-passing parallel molecular dynamics implementation*, Comput. Phys. Comm., 91 (1995), pp. 43–56.
[6] M. Bladt and M. Sörensen, *Statistical inference for discretely observed Markov jump processes*, J. R. Stat. Soc. Ser. B Stat. Methodol., 67 (2005), pp. 395–410.
[7] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein, *Metastability in stochastic dynamics of disordered mean-field models*, Probab. Theory Related Fields, 119 (2001), pp. 99–161.
[8] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein, *Metastability and low lying spectra in reversible Markov chains*, Comm. Math. Phys., 228 (2002), pp. 219–255.
[9] A. Brass, B. J. Pendleton, Y. Chen, and B. Robson, *Hybrid Monte Carlo simulations theory and initial comparison with molecular dynamics*, Biopolymers, 33 (1993), pp. 1207–1315.
[10] L. Breiman, *Probability*, Classics Appl. Math. 7, SIAM, Philadelphia, 1992.
[11] R. C. Coifman and S. Lafon, *Diffusion maps*, Appl. Comput. Harmon. Anal., 21 (2006), pp. 5–30.
[12] P. Deuflhard, W. Huisinga, A. Fischer, and Ch. Schütte, *Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains*, Linear Algebra Appl., 315 (2000), pp. 39–59.
[13] P. Deuflhard and Ch. Schütte, *Molecular conformational dynamics and computational drug design*, in Applied Mathematics Entering the 21st Century, J. M. Hill and R. Moore, eds., SIAM, Philadelphia, 2004, pp. 91–119.
[14] D. L. Donoho and C. Grimes, *Hessian eigenmaps: New locally linear embedding techniques for high-dimensional data*, Proc. Natl. Acad. Sci. USA, 100 (2003), pp. 5591–5596.
[15] P. G. Doyle and J. L. Snell, *Random Walks and Electric Networks*, Mathematical Association of America, Washington, D.C., 2000.
[16] W. E and E. Vanden-Eijnden, *Towards a theory of transition paths*, J. Stat. Phys., 123 (2006), pp. 503–523.
[17] C. W. Gardiner, *Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences*, Springer-Verlag, Berlin, 2004.
[18] A. Gupta, M. Zangril, A. Sundararaj, P. A. Dinda, and B. B. Lowekamp, *Free network measurements for adaptive virtualized distributed computing*, in Proceedings of the 20th IEEE International Parallel and Distributed Processing Symposium, 2006.
[19] W. Hackbusch, *Elliptic Differential Equations: Theory and Numerical Treatment*, Springer-Verlag, Berlin, 1992.
[20] S. Huo and J. E. Straub, *The MaxFlux algorithm for calculating variationally optimized reaction paths for conformational transitions in many body systems at finite temperature*, J. Chem. Phys., 107 (1997), pp. 5000–5006.
[21] S. Lafon and A. B. Lee, *Diffusion maps and coarse-graining: A unified framework for dimensionality reduction, graph partitioning, and data set parameterization*, IEEE Trans. Pattern Anal. Mach. Intell., 28 (2006), pp. 1393–1403.
[22] E. Lindahl, B. Hess, and D. van der Spoel, *GROMACS* 3.0*: A package for molecular simulation and trajectory analysis*, J. Mol. Model., 7 (2001), pp. 306–317.
[23] M. Meila and J. Shi, *A random walk's view of spectral segmentation*, in Proceedings of the 8th International Workshop on Artificial Intelligence and Statistics, 2001.
[24] P. Metzner, E. Dittmer, T. Jahnke, and Ch. Schütte, *Generator estimation for Markov jump processes*, J. Comput. Phys., 227 (2007), pp. 353–375.
[25] P. Metzner, Ch. Schütte, and E. Vanden-Eijnden, *Illustration of transition path theory on a collection of simple examples*, J. Chem. Phys., 125 (2006), 084110.
[26] M. E. J. Newman, *The structure and function of complex networks*, SIAM Rev., 45 (2003), pp. 167–256.
[27] S. Park, M. K. Sener, D. Lu, and K. Schulten, *Reaction paths based on mean first-passage times*, J. Chem. Phys., 119 (2003), pp. 1313–1319.
[28] D. M. Roma, R. O'Flanagan, A. Ruckenstein, A. M. Sengupta, and R. Mukhopadhyay, *Optimal path in epigenetic switching*, Phys. Rev. E (3), 71 (2005), 011902.
[29] S. T. Roweis and L. K. Saul, *Nonlinear dimensionality reduction by locally linear embedding*, Science, 290 (2000), pp. 2323–2326.
[30] Ch. Schütte, A. Fischer, W. Huisinga, and P. Deuflhard, *A direct approach to conformational dynamics based on hybrid Monte Carlo*, J. Comput. Phys., 151 (1999), pp. 146–168.
[31] Ch. Schütte, W. Huisinga, and P. Deuflhard, *Transfer operator approach to conformational dynamics in biomolecular systems*, in Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems, B. Fiedler, ed., Springer-Verlag, Berlin, 2001, pp. 191–223.

[32] J. Shi and J. Malik, *Normalized cuts and image segmentation*, IEEE Trans. Pattern Anal. Mach. Intell., 22 (2000), pp. 888–905.
[33] R. Syski, *Passage Times for Markov Chains*, IOS Press, Amsterdam, 1992.
[34] E. Vanden-Eijnden, *Transition path theory*, in Computer Simulations in Condensed Matter: From Materials to Chemical Biology, Vol. 2, Lecture Notes in Phys. 703, M. Ferrario, G. Ciccotti, and K. Binder, eds., Springer-Verlag, Berlin, 2006, pp. 439–478.